IMPROVING DATA SCIENCE LABS FOR PROMOTING STUDENTS' COMPUTATIONAL ACTION AND SOCIAL JUSTICE AWARENESS: A DESIGN-BASED RESEARCH STUDY

BY

KARLE FLANAGAN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Curriculum and Instruction
in the Graduate College of the
University of Illinois Urbana-Champaign, 2022

Urbana, Illinois

Doctoral Committee:

Professor Gloriana González, Chair
Professor Michelle Perry
Assistant Professor Christina Krist
Assistant Professor Michael Tissenbaum

ABSTRACT

Data science education is important for all students because they need data science literacy to succeed in their future careers and to be able to make informed decisions as citizens. In this curricular study, I redesigned two data science labs so that they were centered around social justice issues and included scaffolds and artifacts that encourage communication in many forms. The redesigned labs contained coding questions, individual written reflection questions, and group discussion questions. Using the theory of distributed cognition, I designed the scaffolds and artifacts in the labs to help students engage in multimodal communication. In other words, students discussed the data science work that they did and the implications of this work in writing and through talking.

This study was done using design-based research (DBR) to allow for multiple iterations to improve the scaffolds and artifacts in the labs. Through the DBR process, I was able to document the principles of distributed cognition that I used to design the labs, as well as the changes that I made to the labs and the reasons behind making those changes. Through using the Toulmin Argumentation Pattern to analyze the group discussion questions and Thematic Analysis to analyze the individual reflection questions, I found evidence that the students engaged in computational action (CA) and social justice awareness. In their group discussions, most groups used the data analysis that they did as the grounds or warrants for their arguments. In other words, they used the data science that they did to justify their claims and the scaffolds provided data for these claims. The students also engaged in data science practices by acting as real data scientists while working on authentic problems related to social justice issues.

In their individual reflections, the students reflected on their analysis, the implications of this analysis, and were able to connect the work they did inside the classroom to the world

outside of the classroom. I found themes related to social justice, data science concepts, and the connection to the outside world which showed evidence that the students engaged in computational action and social justice awareness. This work provides examples of two labs that are curricular innovations with social justice components in a data science course. I also identified five design principles for creating labs that focus on multimodal communication and social justice. This study also illustrates methods that can aid in the understanding of how to improve data science labs.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Dr. Gloriana González. Without her, this work would not have been possible. I am forever grateful for her mentorship and guidance throughout my time in the College of Education, as well as the feedback and encouragement she has provided me throughout every milestone. Words cannot express how much I appreciate her support from Day 1.

I would also like to thank my committee members: Dr. Michelle Perry, Dr. Stina Krist, and Dr. Mike Tissenbaum. Their feedback, expertise, and perspectives have been invaluable throughout the journey of this dissertation. I am grateful to each of them for the time they have spent with me.

I would also like to thank the Statistics Department for being patient with me as I navigated life as a Ph.D. student and a faculty member for the past four and a half years. While this has not been easy, they have supported me in so many ways and I am grateful to be a part of a department that values education. I would also like to thank the College of Education for welcoming me with open arms and allowing me to explore research areas that I am passionate about.

This work would also not be possible without the colleagues that I have worked with over the past few years who have become my friends. Thanks to Julia Nagel who helped make this research possible by assisting with data collection and making my life easier in every way that she could. Thanks to Wade Fagen-Ulmschneider who has been an amazing co-teacher and someone who challenges me to be the best version of myself. Thanks to Ellen Fireman who taught me everything I know about teaching statistics. Thanks to Kelly Findley for answering all of my questions and sharing his knowledge throughout this process. Thanks to the other Ph.D.

students who have encouraged me during the difficult moments and made this journey fun, especially Laura Placzek.

Outside of the university, I would also like to thank my friends and family who have supported me throughout the years. There are too many to list individually. A special shoutout to my best friend, Erin Cornelius. I am so blessed to have a friend like her.

I would also like to thank my parents, Tony and Natalie Laska, who have always believed in me and encouraged me to never stop learning. A special shoutout to my dad who when I asked if I should consider doing a Ph.D., he said "Yeah, that sounds like a great idea!" A special shoutout to my mom who never complains when I have to work while visiting her at the lake and has been my biggest cheerleader since the day I was born. I would also like to thank my in-laws, Karen and Terry Flanagan, two amazing educators that have provided so much support.

Thank you to my students who challenge me to be a better teacher and a better person each semester. I am incredibly grateful to be a part of their journey in learning statistics and data science. They make my job the best job in the world.

And last, but most certainly not least, thanks to my husband, Steve Flanagan. The love and support he has provided throughout this process is unmatched. Thank you, Steve. For everything.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

This is a curricular study using design-based research methods to improve the labs in a college-level data science course. Using the theory of distributed cognition, I designed the scaffolds and artifacts in the labs for students to develop communication skills and increase awareness of using data science for social justice. The goals of the labs are for students to engage in computational action and social justice awareness by engaging with coding exercises, individual reflections, and group discussion questions in the labs.

Statement of Problem

One of the key needs of society today is for citizens to be literate in data science. Data science is a new field that looks at how statistics, computer science, and communication intersect. Data science courses go a step further than statistics courses by integrating computation and programming throughout the course, rather than using it as a secondary tool. According to Moses (2019), one accepted goal of higher education is to prepare students for the real world and its needs. Nowadays, students in all disciplines need basic data science skills to be successful at their jobs. Also, the demand and need for data scientists has expanded well beyond the tech industry to many other disciplines, for example business and the humanities (Irizarry, 2020). There is widespread acknowledgment that the job market for people who have data science skills is strong, and there is evidence that demand for this type of labor far exceeds supply (Baumer, 2015). According to Van Der Aalst (2016), data science is the profession of the future and if data science is not a part of organizations, the organizations will not survive. Data science courses can be the first step in preparing students for doing data science in their jobs.

However, data science skills and conceptual understanding of data science topics are not only necessary to prepare students for being employees, but also for being citizens (Finzer,

2013). In other words, it is important for college students to become critical consumers and producers of data who know how to answer real world questions, think about the implications of data science, and make decisions under uncertainty both in the workforce and in their everyday life. The skills that students learn in their data science courses should empower them to understand the world, question the status quo, and think about important issues of social justice and equity. In their 2020 book, *Data Feminism*, D'ignazio and Klein discussed how data science needs feminism and that we need to rethink the way that data science is taught. D'ignazio and Klein (2020) stressed that students should think about data, their analysis, and their display in a way that is informed by feminist activism and critical thought so that they can use data science to make the world a more just place. Additionally, it is important for students to think about the absence of data for significant problems and why certain data are not being collected. Also, it is important for students to think about how data is being generated and the implications of who this harms and benefits.

According to Engel (2017), massive amounts of data on important societal topics are accessible to the general public, covering a wide range of important topics including "migration, employment, social (in)equality, demographic changes, crime, poverty, access to services, energy usage, living conditions, health and nutrition, education, human rights, and many others" (Engel, 2017, p. 45). Understanding these topics and being able to think critically about them is essential for civic engagement in modern society, however, analyzing this type of data is not usually part of regular statistics instruction at the K-12 or college level (Engel, 2017). Data science courses can help students explore these issues and think about the implications of doing data science, as well as who benefits from it and who does not. Data science courses can also help students think about how these datasets were collected and generated. This also includes

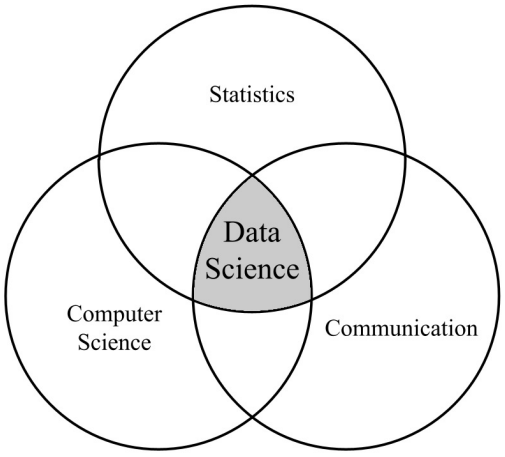
thinking about what groups of people benefit from collecting and not collecting certain types of data and what groups of people this data collection harms. Overall, data science education and data science courses can play a major role in helping students prepare for the workforce and more importantly, prepare for being citizens who think critically about issues of equity, social justice, and making the world a better place.

What is Data Science?

Data science is a new field that requires students to do statistics, use computation, and be able to communicate their results. Because it is such a new field, there is not a general consensus on how to define data science yet. Many scholars and researchers have provided their definitions of data science (Song & Zhu, 2016; Irizarry, 2020). In general, data science can be thought of as the discipline that educates people who can address challenges in the big data era that we are in today (Song & Zhu, 2016). Data science is often represented as the intersection of doing statistics, using computation, and the ability to communicate your findings, which is how I think about it throughout this dissertation (Figure 1).

Figure 1

What is Data Science?



Various stakeholders and researchers have different views on whether or not data science should be its own academic field or whether it should be a subdomain of another field, most commonly statistics or computer science (Engel, 2017). However, stakeholders and researchers seem to agree that data science education is a very important part of K-12 and post-secondary education. At many universities, data science is often housed in Statistics Departments, but more and more introductory data science courses are being created as data science grows into its own discipline. In data science classrooms, technology tools, such as programming languages, can help students learn statistics and better understand statistical concepts. By learning statistics, computation, and information science in parallel, students are not only able to understand basic statistics better, but they are also able to use the programming, statistical practices, and communication skills they learn in class and transfer them to job situations, as well as situations in the real world.

In situations both in the workforce and in society, it is important that students can communicate their findings by explaining what they did and why. Jobs require their data scientists to communicate their results by presenting reports that include visualizations and explaining what they did in a way that business partners or people without a background in data science can understand. Also, if students are using data science to question the status quo and think about social justice issues, it is important that they can communicate their results to others in a way that makes sense to people without specialized knowledge and the general public.

Because communication is multimodal, it can involve basic visualizations and explaining concepts using both verbal and written communication. Doing data science without explaining how and why it was done leaves certain communities not engaged and discussions not being had (D'ignazio & Klein, 2020). It is important that students communicate their results in a way that

evokes emotion to get people interested and involved in the data science they are doing and issues that they are investigating. Despite the importance of communication in data science, this often gets pushed aside in data science courses, especially when courses are large because it can be difficult to assess. However, the communication piece of data science is just as important as the statistics and computer science, which is why it is emphasized in this study.

Rationale for Studying Labs in a Data Science Class

Throughout this study, I designed and implemented labs in an introductory data science class that emphasize real-world skills, solving problems, and communication. According to Rao et al. (2018), most popular online data science courses require prior programming experience in Python or other tools, which can be intimidating to non-computer science majors. The course in this study does not have any prerequisites and is designed for non-computer science majors who have never had any programming or statistics prior to taking the course. In order to successfully help students learn Python, it is important that they get time to practice programming and using Python to solve real world problems. Most of the practice comes from the labs that they complete in their lab sections each week.

Introductory data science courses typically have three hours of lecture and a lab section each week. In the course in this study, the lab sections are 80 minutes long and occur once per week outside of lecture. The lab sections have 30 students at most, whereas the lectures have about 300 students. These lab sections provide an opportunity for students to get practice with programming, verbal communication, and written communication through working in small groups on longer data science programming problems. This can be difficult for students to do in lecture, especially with 300 students. The lab sections allow the students to collaborate with their peers and get practice talking about data science. The labs also allow the students to get practice

actively programming themselves rather than passively listening to the instructor or watching the instructor program in lecture.

Data science labs are a unique space. Labs in science courses and statistics courses are very common and there are many researchers looking at how to implement these labs in ways that maximize students learning. Coker (2017) described how science labs should be designed so that students can carry out their own experiments. Science labs should not follow a set of instructions that lead to a predetermined finding (Coker, 2017, p. 14). Guardiola et al. (2010) described how they added a lab component to a statistics course that previously had no technological component. The lab component had students use software to solve problems that were similar to the problems the students saw in lecture (Guardiola, 2010). In this study, I consider data science labs to be a combination of science labs where students do their own experiments like Coker (2017) described and statistics labs where students use software to solve practice problems related to what they learn in lecture. The goal of the labs I redesigned is for students to solve real-world problems using real-world tools. The real-world problems involve important social justice issues, and the real-world tools are the artifacts included in the labs, such as Python, Python libraries, and the datasets. This study is an example of how to create data science labs with specific goals in mind and this study will add to the limited literature on data science labs.

Communicating in the Labs

The Guidelines for Assessment and Instruction in Mathematical Modeling Education (GAIMME) Report emphasized the importance of communication in mathematics, specifically mathematical modeling. GAIMME also mentioned the importance of giving students the opportunity to communicate and to develop their mathematical communication (Garfunkel et al.,

2016). In this curricular study, I redesigned two data science labs so that they are centered around social justice issues and include scaffolds and artifacts that encourage communication in many forms. Prior to this study, these labs did not have any scaffolds that encouraged communication and the students often completed them on their own. The scaffolds and artifacts that were included in the redesigned labs allowed the students to collaborate and work in small groups. Without the lab sections, communication would be a difficult skill to develop in the lectures alone. GAIMME also emphasized the importance of assessing individual written communication, along with giving students the opportunity to communicate with each other (Garfunkel et al., 2016). The scaffolds in the redesigned labs contained reflection questions in which students individually communicated their thoughts in writing and discussion questions that required the students to talk to each other. Many of these questions focused on thinking about the implications of the data analysis the students had done. In the redesigned labs, I expected the students to discuss many of the ideas in D'ignazio and Klein's *Data Feminism*. For example, I expect the students to discuss who benefits from the analysis they do and who does not, along with who is doing data science and who is not. This type of communication assessment is not present in the typical homework and exams in the course, again showing the importance of the new labs.

In addition to the open-ended questions on the redesigned labs, the redesigned labs also contain questions and activities that had students use visualization as a tool to help communicate their findings. GAIMME mentions oral and visual communication as ways that students can effectively communicate their findings. The development of mathematical communication skills directly transfers to other courses and the job market (Garfunkel et al., 2016). These skills are applicable to any job, and they are key skills that employers look for when hiring. Also,

GAIMME emphasized the idea that the skills not only help students get jobs, but also help them "navigate the global community in which we live" (Garfunkel et al., 2016, p. 90). In other words, the lab environment is necessary to allow students to get practice communicating mathematically which will help them on the job market and in their lives.

Data Science Practices in the Labs

The redesigned labs aimed at engaging students in data science practices. In other words, the labs intended to give the students a space to act as real data scientists while doing the exercises and working with their peers. Some of these data science practices included communicating about the analysis that they have done. My intention was for students to engage in data science practices through oral discussion, in writing, and through creating representations of data such as visualizations. Another important data science practice was using the work they did to justify claims about social justice issues. My intention was for the students to get practice doing analysis to answer real world questions that they had about the social justice issues that were present in the labs. By engaging in data science practices, the intention was for the students to engage in both computational action and social justice awareness.

Explanation of Theories and Constructs

The goal of this study is to understand how to design labs in a data science course, using principles from distributed cognition that help students engage in computational action and social justice awareness. This was done using design-based research or DBR. DBR is a relatively new approach to research that started in the early 1990s by Alan Collins (1992) and Ann Brown (1992). Brown (1992) and Collins (1992) started doing DBR to study teaching and learning in a natural environment, rather than a lab. DBR is a method used primarily for education research that is situated in a real educational context, focuses on the design and testing of an intervention,

is iterative, collaborative, flexible, and has a practical impact on the practice of teaching (Anderson & Shattuck, 2012; Barab & Squire, 2004). The goal of DBR is for researchers to study teaching and learning in a natural environment and for researchers to make improvements to interventions based on student feedback (Collins, 1992; Brown, 1992). I used the DBR methodology because this study took place in a classroom, contained two iterations, and student feedback played a key role. One outcome of this study was the identification of design principles for improving labs in data science courses.

I designed the labs that the students worked on using principles of distributed cognition. Distributed cognition (DC) is a theory that looks at how knowledge is distributed or allocated among individuals and their surroundings (Hutchins, 1995; Rogers, 1997). Much of the work done on DC can be attributed to Edwin Hutchins and his 1995 book, *Cognition in the Wild*. Hutchins is considered a pioneer in the study of DC because he developed this framework to study ship navigation. He did this by spending a month aboard the ship, the U.S.S. *Palau*, and studying the culture of the navigation team. The work of Hutchins is in the tradition of others such as Lave (1988) who have investigated learning as situated by conducting ethnographies. Anderson et al. (1996) wrote about the importance of situated learning in mathematics. They described the disconnect between the mathematics that is done in the classroom and the mathematics that students do in the workplace. Anderson et al. (1996) said that situated learning allows students to see the connection between mathematics and the real world. Similarly, the scaffolds and artifacts in the labs that I designed using DC principles also allowed students to see this connection because they were doing authentic work with others.

With distributed cognition, the unit of analysis is not only the individual, but the entire system in which the individual is working, such as the other individuals involved, the setting,

artifacts, and culture (Hutchins, 1995). Throughout the study, the labs were viewed as a system consisting of the students, their peers, their tools, and the instructors. This was similar to how Hutchins viewed the navigation team as a system consisting of the team members, the artifacts they used for navigation, and their culture. The labs were designed to promote students working together with other people and artifacts, rather than working alone as individuals. Because the labs are a complex setting that involve so much more than just the individual students, DC is a helpful theory to use to design them.

One goal of the labs is for students to engage in computational action. Computational action (CA) is a relatively new construct for computing education that emphasizes the idea of ensuring that the skills that students learn inside the classroom can help them outside of the classroom (Tissenbaum et al., 2019). CA allows students to recognize themselves as members of a community of scientists who can design solutions to problems that are important to them (Tissenbaum et al., 2019). CA also heavily relies on student choice, allowing students to use data science to explore questions that interest them and have discussions with others about these questions. The work of Tissenbaum et al. (2019) is in the tradition of other initiatives that look at broadening participation in computer science and ensuring that the computing students do in the classroom is relevant in their lives outside of the classroom. Peckham et al. (2007) described how it is not only important to get more people involved with computer science, but it is also important that we teach computer science in a way that connects to students from many different backgrounds. I attempted to do this in the labs by creating scaffolds that allow students to communicate about the implications of their analysis and become critical consumers of data who ask important questions.

However, it is important to acknowledge that a large part of CA involves having students analyze real data that they care about. In large classes, it can be difficult to find data that everyone cares about and is interested in. I attempted to do this by using datasets and examples that have traditionally been known to be interesting to students. I also attempted to insert agency into the questions they were answering in the labs to allow the students to have some choice over what analysis they did and the discussions that they had. Despite these limitations, CA allows me to tie together communication and social justice, which is why it is being emphasized throughout this study. DBR, DC, CA, and social justice are the methods, theories, and constructs that frame this study and are included in the research questions.

Research Questions

1) How can design-based research be used to create labs that use principles of distributed cognition in the context of a data science course?
    a) What scaffolds and artifacts are included in the labs?
    b) What principles of distributed cognition are used in the design of these scaffolds and artifacts and why?
    c) What adaptations to the labs (scaffolds and artifacts) result from the DBR process and why?
2) What evidence do students show of engaging in computational action during and after these labs?
    a) What evidence is there that the scaffolds and artifacts help students engage in computational action?
    b) How do students apply data science practices to question the status quo and consider social justice issues?

c) What do students perceive that they are learning through these labs that will be useful in the real world?

Explanation of the Research Questions

The first set of research questions involves looking at the design process of creating the labs. I designed the labs using principles from distributed cognition that emphasize communication, including verbal and multimodal communication. Overall, the first set of research questions address the curriculum of the labs, leading to a set of design principles that could be applied to the design of data science labs. Specifically, this shows what types of scaffolds and artifacts are used to encourage communication, as well as what DC principles motivate these scaffolds and artifacts. This set of questions also looks at the iterative process of DBR and how the original scaffolds and artifacts change after the first iteration and getting student feedback. The second set of research questions involves looking at the outcome of the innovation. The goal of the labs is for students to engage in CA and social justice awareness and this research question addresses how that is accomplished. Specifically, this provides evidence about whether and how the scaffolds and artifacts help the students engage in CA and social justice awareness. This also helps me understand what the students perceive they are learning that helps them both in the workplace and in society.

Positionality Statement

My experience as a doctoral student is unique. I have worked as a full-time faculty member at a large university for almost nine years. As a former undergraduate who wanted to become a high school mathematics teacher, when I got hired to teach introductory statistics at a university, I felt as if I had gotten my dream job. My goal in teaching statistics has been to help students see the discipline of statistics as something that can be useful to them, rather than

something that they are afraid of. For most of my career, I have taught introductory statistics to non-majors and have been successful, for the most part, at helping them appreciate statistics as a tool to help them make decisions under uncertainty. My teaching evaluations and campus-level awards have recognized my dedication and teaching skills. Recently, I had the opportunity to develop and teach an introductory data science course that combines the power of statistics with computation using Python. This is the course that is being studied in this dissertation, which puts me in the position of the course developer, the instructor, and the researcher all at the same time. I acknowledge that this can lead to bias because I am so close to the course, however, I think that with a DBR study, this can also be a strength.

Using DBR allows me to do meaningful research in my own classroom that not only helps my students in their quest to learn data science, but it also helps me as an instructor understand my students better and use their feedback to make changes to the labs. I understand the course, the content, and the students more than most researchers and I have the power to make any changes that I think will benefit them. I believe that data science is even more powerful than statistics on its own because it includes computer science and communication. Both of these enhance learning statistics and I want to use all three components of data science to make the course meaningful for students in multiple ways.

For the past few years, I have wanted to add a social justice component to my courses. After the Summer of 2020, I started reflecting on the ongoing Black Lives Matter movement and thought about how I can incorporate social justice issues into my classroom. I realized that while I sometimes do use example problems related to social justice issues in my classes, I never have the students discuss these issues at all. I originally thought that mathematics, statistics, and data science were not political and that there was always a right answer in these subjects. However,

after taking classes in my doctoral program such as *Sociopolitical Perspectives on Mathematics and Science Education* with Dr. Rochelle Gutiérrez, I have learned that all subjects are political and that it is important to prepare my students to help make the world a more just place. In the data science classroom, having these discussions can be difficult because of the size of the class. However, the lab sections and labs are the perfect place to revamp the assessments to include more social issues and discussion.

When I first started teaching this data science course, I had no idea how important the lab sections would be. I originally thought we should not have lab sections because I thought the students would not sign up for the class if it had three hours of lecture and 80 minutes of lab each week. However, other faculty members from the Computer Science Department insisted that labs are the best way for students to learn data science. The first semester I taught this course, my co-teacher and I led the labs and I quickly learned how important they were. They helped foster a community and were a great space for students to get practice using what they learned in lecture in a safe environment where they could ask questions and work with others. This was where so much of the learning happened. Ever since that first semester, I have wanted to include more types of questions and communication throughout these labs, however with graduate school and teaching a full load, I never got around to revamping the labs until this study.

This dissertation has motivated me to try new curricular innovations to help the students communicate in the labs, and it has also motivated me to think about what datasets I am using, what types of questions I am asking, and how to incorporate social justice issues into the data science classroom. As a white cisgender female who has grown up quite privileged, I understand that many of my students may come from backgrounds that are different from mine. I think it is important that all students not only do data science but think about the implications of doing data

science, think about who is doing data science, and who benefits from it. The different perspectives present in the discussion sections and lab groups will help the students engage in rich conversations about a variety of issues.

As a woman in STEM, I understand some of the challenges that the women in the class face, however, I lack understanding of what it is like to be an under-represented race in a STEM course. I want to understand more about how all students think about the social implications of data science and engage in communication throughout the labs. Overall, there are limited opportunities to have conversations about data science and social justice, so this is my attempt to insert this into college courses. In my training, I was never given this opportunity as a student, but I am trying to open spaces for my students to have this experience.

Summary

Data science courses can be incredibly valuable for college students today and help transform our society to become critical consumers and producers of data. They can be used to help them better understand statistical topics by allowing them to use programming as a tool to explore these topics thoroughly. Data science courses also help students be prepared for a workforce in which basic data science skills are required. Lastly, data science courses can help students become citizens who use data science to think about social justice issues, question the status quo, and make society more just. Being able to communicate results and explain findings from data science analysis is especially important for students, but this is often not emphasized in data science courses. This study focuses on how to design labs in an introductory data science course using principles from distributed cognition. The scaffolds and artifacts based on DC principles used in these labs will encourage communication through talking with others, writing, and visualization. The goal of the labs is for students to engage in computational action and

social justice awareness. In other words, the goal is for students to engage in data science practices that can help them in the workplace and in society. Because this work took place in an actual classroom and the goals are computational action and social justice awareness, this study was done using design-based research in order to get feedback from the students and optimize their learning opportunities.

CHAPTER 2: LITERATURE REVIEW

Data science is a new field that has been growing rapidly for the past decade. Because it is so new, there is not extensive literature about data science education in college. Throughout this literature review, I include some studies and literature from K-12 education, however the focus of this study is on introductory data science in college. I specifically focus on two themes throughout this study: communication and social justice. I emphasize that communication is multimodal, meaning that it can take on many forms such as written communication, verbal communication, and data visualization as communication. I also emphasize that social justice is an important component of data science education.

Data Science Education History

Data science education is becoming more and more common in colleges and universities today because of the importance of statistical literacy and data analytics skills in everyday life. Students need data science education to handle our society's increase in the reliance on data (Finzer, 2013). Data science is an interdisciplinary field that has been growing rapidly since around 2007 when the world observed a new wave of technology solutions built around data (Ramamurthy, 2016). Between 2007 and 2009, multiple large companies such as Google, Facebook, and Yahoo!, met to discuss the new data science that they had been using internally at their companies to handle the uptick in data (Ramamurthy, 2016). This included data modeling, data aggregation and data mining. During this time, leaders in industry realized that traditional science and engineering majors were behind in these data science skills used at technology companies (Ramamurthy, 2016). In 2009, a proposal was written for a National Science Foundation (NSF) grant to help tackle this problem by creating a data forward course and a certificate program for students. The project, Timely Introduction of Data-intensive Computing

(TIDE) began in 2010 (Ramamurthy, 2016). This can often be seen as the start of data science education.

The Guidelines for Assessment and Instruction in Statistics Education (GAISE) created three reports for recommendations for teaching introductory statistics. The first report, created in 2007, was for grades PreK-12 and the second report, published in 2016, was for the college level. More recently, the Guidelines for Assessment and Instruction in Statistics Education (GAISE) introduced the GAISE II in 2020. All reports are endorsed by the ASA (American Statistical Association) and have been used as the foundation for many introductory statistics courses here in the United States. GAISE II emphasizes the importance of including data science in statistics education. The introduction described how it is essential that students "leave high school prepared to live and work in a data driven world" (Franklin & Bargagliotti, 2020, p. 2). The document highlights the COVID-19 pandemic and how communicating with data should be at the forefront of data science education. Although the GAISE II is primarily focused on PreK-12 education, many of the ideas introduced in this document can be seen in college-level introductory data science courses as well. These include communication, visualization, storytelling, skepticism, and statistical reasoning. The importance of data science education can be seen at both the K-12 levels and post-secondary levels; however, this study specifically focuses on introductory data science at the college level.

After the launch of TIDE, schools with strong computer science and statistics programs started introducing data science courses and concepts into their curricula (Ramamurthy, 2016). This involved creating introductory data science courses as well as adding data science concepts or lessons to existing courses. Many of these schools also established data science certificate programs. Data science education also extended beyond just the universities. MOOCs (massive

open online courses) hosted on platforms such as Coursera, EdX, and Udacity have several data science certificates, programs, and courses which are targeted for beginners and the general public (Ramamurthy, 2016). As the need for data science education expanded, so did the need for data science skills in the workforce. Nowadays, there is a gap between the amount of data science job positions and the amount of people with data science training who can fill them.

As universities worked to address this gap, they realized that all students, regardless of their field, can benefit from data science education. Originally, data science was seen as the intersection of statistics and computer science. The data science education developers at the University of California Berkeley (UC Berkeley) said that they whole heartedly believe that computation and inference (statistics) are "natural allies" and should be taught together, as a blend, in a modern undergraduate curriculum (Adhikari et al., 2021). Data science is more than just the combination of statistics and computer science, however there is not a clear definition of what data science actually is. Dichev and Dicheva (2017) defined data science as "an interdisciplinary field about scientific processes and systems to extract knowledge or insights from data in various forms" (p. 2151). More specifically, many educators and researchers see data science as an interdisciplinary field that combines statistics and inferential thinking, computer science involving coding, visualizing, and computing with data, and domain expertise with a focus on information science, communication, and ethics (Engel, 2017). Because of this consensus, educators often design their courses around this definition when thinking about their learning outcomes and goals.

Introductory Data Science Curriculum

In order to have a sense of the typical content in introductory data science courses, I reviewed Dichev and Dicheva's detailed description of their introductory data science course

along with other courses with syllabi available online. Dichev and Dicheva (2017) described their introductory data science course in detail and this course provides a general overview of how most introductory data science courses have been designed. They covered introduction to computation using Python, starting with simple basics and arithmetic, then transitioning to more complicated computer science topics such as control flow, conditionals, and data structures[1]. Next, they covered data collection and data processing using data frames, working with CSV (comma separated values) files, cleaning data, and basic EDA (exploratory data analysis). They also covered standard introductory statistics topics such as descriptive statistics, inference, and linear regression. Lastly, there was a large focus on communication and visualization using simple visual displays of data such as boxplots, histograms, and scatterplots.

Another platform course that provides a model for introductory data science courses is UC Berkeley's "Data 8." Data 8 is one of the first introductory data science courses designed for students of all majors with no prerequisites. In other words, any student, regardless of their statistics and computer science background can take Data 8. This allows more and more students, especially those in non-technical fields, to be exposed to data science and many schools are following suit and creating courses like Data 8 that require no prior knowledge in programming and statistics. This course plays a critical role in providing broadly accessible and relevant data science instruction to the UC Berkeley undergraduate population and beyond (Adhikari et al., 2021). A key focus in Data 8 is on ethics, specifically the ethics of data collection and analysis. UC Berkeley has a designed an entire data science program around their introductory course and

[1] Control flow is a concept in computer science that represents the order in which the computer executes code. Usually, code is read by the computer from top to bottom, but we can write more complicated code that controls the flow depending on how we want the computer to execute it. Conditionals are a way for computers to make decisions off of conditions. Conditional statements are handled by IF statements in Python. Data structures are different formats for organizing, processing, and storing data. All three of these are common topics in introductory computer science.

integrated data science throughout the campus through "connector courses." Connector courses are data science courses that immerse students in a particular domain. Faculty from any department can design a connector course using the same infrastructure as Data 8. These connector courses are focused on a specific field and very project oriented (Adhikari et al., 2021). UC Berkeley is a leader in undergraduate data science education, and they have worked to share their ideas and materials with others by hosting and organizing a conference centered on data science education, The National Workshop on Data Science Education, that has been in existence since 2018. This conference brings people from all over the United States together to collaborate and discuss undergraduate data science education.

Along with being a leader in promoting data science education, UC Berkeley has also discussed what they call the provocative question of whether or not the new data science curriculum should "entirely displace classical curricula in computer science and statistics, or should it simply live side by side with those curricula?" (Adhikari et al., 2021, p. 22). At many universities, it has been unclear in which department data science should exist. Although it is interdisciplinary, data science as a field is different from statistics and computer science on their own. A statistics major often lacks computational power and a computer science major often lacks statistical training and inferential thinking (Finzer, 2013). Data science is also different from information science (IS) on its own. Furner (2015) argued that information science is not only a science, but that there is much more to it than just information and science. IS focuses heavily on ethics, policy, interpretation, communication, and storytelling (Furner, 2015). Data science is unique because it contains elements of statistics, computer science, and information science, but does not go as in depth on each of them as the individual majors do.

Data Science Labs

Despite there being no universal definition of data science, there are some common design elements that show up in many introductory data science courses, including the course in this study. The first is that data science courses are generally taught with a lecture and what most universities are referring to as labs. The labs allow the students to practice what they learn in lecture. These labs can be done in a discussion section, like in UC Berkeley's Data 8, or completed out of class like in the class that Dichev and Dicheva (2017) designed. Throughout the labs, it is common for the students to get practice using industry-standard tools. In other words, while working on the labs, the students are able to practice using tools that data scientists who work in industry use regularly. Both Data 8 and Dichev and Dicheva's course use Python for the computation part of data science. Over the last couple of decades, Python has emerged as one of the most popular tools for data science (VanderPlas, 2016). Specifically, this is because Python has large, active third-party libraries that make the analysis and visualization of large datasets relatively simple (VanderPlas, 2016).

To my knowledge, there is limited literature on data science labs and what goes on in them. However, there are some studies that look at statistics labs, which have similarities to data science labs. Gould et al. (2010) created a set of labs for their introductory statistics course that helped the students work with data in a realistic context. Although these labs were done using Fathom[2], a data analysis software, instead of Python, the researchers still created the labs so that the students can work with real data and use it to discover analysis procedures and better understand statistical concepts. The labs helped the students learn to see statistical analysis as

[2] Fathom is a dynamic data software that is mainly used by high school students to explore modeling with mathematics. It is a point and click program that allows you to quickly build and manipulate data visualizations.

inquiry and discovery, rather than a recipe (Gould et al., 2010). These labs, like data science labs, were a supplement to lecture, involved looking at basic data visualizations such as bar charts and two-way tables, and had the students work together in smaller groups. Despite these similarities, there was no mention of communication or social justice issues in the labs designed by Gould et al. There were no scaffolds that encouraged written or verbal communication and the focus was on using real data, rather than data that involved social justice issues.

Other researchers have looked at how to design labs in statistics courses as well. Nolan and Speed (2001) authored a book called *Stat labs: mathematical statistics through applications* that illustrated how to teach mathematical statistics through in-depth case studies that students completed as labs. These labs involved working with real datasets and understanding a problem by using statistical techniques to investigate it (Nolan & Speed, 2001). These labs gave a lot of background on the topic and contained many open-ended questions, as well as reflection questions about the analysis. A few of the case studies contained topics related to social justice such as voting behavior, however, this did not seem to be the focus. Again, the focus seemed to be on using real datasets, rather than communication and social justice. These labs also used R, another statistical software, and not Python. Overall, it seems that focusing on designing data science labs that encourage communication about issues related to social justice is a gap in the literature that should be further explored.

Benefits to Data Science Education

Data science education has many benefits to all students and having basic data science knowledge helps students both in their personal and professional lives. Dichev and Dicheva (2017) described how data science literacy will soon become an important asset for any type of profession, such as technology, science, finance, journalism, politics, and marketing. Even

though most undergraduates will not become data scientists, they still need to know basic data science (Dichev & Dicheva, 2017). In other words, data science skills will help students succeed in their jobs and set them apart from their peers. Data science also helps all students better understand society and their environment. Engel (2017) said that enlightened citizens who have the power to study evidence-based facts as well as manage, analyze, and think critically about data are the best remedy for a world that is guided by fake news or false claims.

Nowadays, massive amounts of data are available to the general public on important societal topics (Engel, 2017). Many of these include social justice issues such as social inequality, crime, poverty, access to services, healthcare, climate change, and human rights. In order to fully participate in society, it is necessary for students to be aware of and understand these topics and data science can help them do that. According to Engel, this is "essential for civic engagement" as social media and access to data have shaped our political discourse (Engel, 2017, p. 45). In order to fully understand these topics, students need data science so that they can understand statistical arguments that have been presented or do statistical analysis themselves. They also need to know how to manipulate and analyze large datasets, as well as communicate their results to others. Franklin and Bargagliotti (2020) also commented on how data are now readily available to the general public and can be used to help gain insights and make recommendations on world issues in their GAISE II document. They argued that now is a critical time for data literacy and that this type of education should start as early as possible (Franklin & Bargagliotti, 2020). Overall, data science education is becoming more and more common because researchers, educators, and students are realizing how important it is. It helps students be prepared for their jobs and become data literate citizens who can help make society a better place.

Critiques to Data Science Education

Despite the many benefits to data science education, there are a few critiques to this new field as well. One of them stems from the idea that in general, data science does not seem to have a natural home in terms of school subjects (Finzer, 2013). At universities, many departments seem hesitant about including data science into their curricula. For example, mathematics departments already often house statistics education and math professors may think that statistics and data science take away from students' ability to focus on abstract math (Finzer, 2013). Biological sciences are already overwhelmed with how much content they have to teach students, so it is difficult for them to imagine adding data science (Finzer, 2013). The social sciences are often worried about the quantitative nature of data science and how their students and faculty will handle it (Finzer, 2013). Figuring out how to make changes at universities so that all students leave with a strong understanding of basic data science presents a challenge since it can be difficult to add to currently existing departments and majors.

Another critique to data science education is that because it is so new, teachers and professors are often not equipped with the skills to teach students data science or incorporate it into their lessons. Oftentimes, teachers are not comfortable using data-driven lessens and policy makers' attempts to strengthen data science education are uninformed (Finzer, 2013). This can be seen in both the K-12 level and the university level, particularly in non-mathematical departments. This makes data science seem daunting to teachers and college professors who already have so much on their plates. Another complication is the limited access that students and teachers have to collecting data (Finzer, 2013). In general, finding good datasets for students to use can be challenging and time consuming for professors. The data should be relevant, real, and relatable to the students; however, such data is often messy or not easily available, making

teachers revert back to standard datasets that may not be interesting to students. Engel (2017) described how important it is for citizens to be critical consumers of data in the media and how students should be exposed to the misuses of statistics so that they can learn about effective ways to overcome them. Finding these specific types of examples can also be incredibly difficult for teachers and professors.

With regard to the technology involved in data science, there may be issues of equity with regard to access to that technology. Oftentimes, students need their own laptop or computer to complete their homework for data science classes and there may be students who do not have them. Kross et al. (2020) mentioned that traditional data science courses target students in technical majors and that many of these students have access to technology and resources to help them learn data science. There is a push to move beyond students in technical majors and expand data science to students in a wide variety of disciplines (Kross et al., 2020). This could accentuate issues of equity in technology. Lastly, Adhikari et al. (2021) described how the infrastructure of Data 8 is complicated and students often need a lot of help outside of classes. Some universities may not have the bandwidth of course staff to help the students outside of class. Because students may not get help outside of class, building labs that rely on collaboration is important. The labs in this study provide students with opportunities to learn from their peers and ask questions. The focus on peer-to-peer communication will hopefully decrease the amount of help that students will need from course staff outside of class. Despite these challenges and critiques, data science education is important and necessary. These challenges call for thinking about improving the design of data science courses and labs to help create a data-literate society and this study is one step towards achieving this goal.

Communication in Data Science

One of the ways to create a data literate society is to make sure that the data scientists and the students coming out of data sciences classes can communicate effectively with non-technical people. Communication in data science is multimodal. In other words, there are multiple ways that students can communicate. For example, students can communicate through writing, visual displays of data, and through group discussion. The GAISE II document highlighted that data scientists need to be so much more than data crunchers (Franklin & Bargagliotti, 2020). While technical skills are important, data scientists need to question the status quo, make decisions under uncertainty, and understand that the art of communication with data is essential (Franklin & Bargagliotti, 2020). Carmichael and Marron (2018) discussed how communication is not generally emphasized in STEM education, despite this being in demand from industry and academic employers. Because it is in demand from employers, communication is a key concept that should be emphasized in data science classrooms.

Because communication is multimodal, students should have the opportunity to practice both oral and written communication. Lemke (1990) described the idea of "talking science" and advocated that learning science means being able to talk science. The same idea holds true for data science. The work of Lemke (1990) is situated in the context of other researchers who commented on the importance of discourse and classroom talk (e.g., Sinclair & Coulthard, 1975; Mehan, 1979; Cazden, 1986; Edwards & Westgate, 1986). The idea of talking science is more than just talking about science, instead it also involves things such as observing, describing, questioning, challenging, designing experiments, evaluating, and teaching (Lemke, 1990). Lemke (1990) mentioned the importance of students learning to communicate the language of science so that they can become members of the scientific community. In general, teachers and

professors are already members of that community who speak the language of science, or in this case, data science. Lemke (1990) argued that "communication is always the creation of community" (p. 12) and the classroom is a great place to start creating a community of learners who know how to talk science. However, in order to do this, students must be able to observe those who already know how to talk science do this and they must also be given the opportunity to practice. The students in this study can see the professors talking data science in the lecture and the labs offer a safe place for them to practice talking data science with their peers and through writing.

Because communication is multimodal, students should also have the opportunity to create visualizations to communicate data science. Data visualization is an entire area of data science that can be explored. While both labs include standard visualizations to help students understand their results and communicate them, how to create good data visualizations is not the main focus of this study. Instead, I focus on communication being multimodal and data visualization being one important way to communicate. For example, students can use histograms to show the distribution of data, or they can use a boxplot to show if the data has outliers. Xyntarakis and Antoniou (2019) mentioned that data visualization allows us to communicate about big data and often, visualizing big data gets mixed with data analytics. Throughout the labs, the students analyze large datasets or create simulations that can be difficult to describe in words. Standard visualizations such as boxplots, histograms, and scatterplots can help the students communicate their results to the general public, including non-technical people.

Dodge (2021) argued that visualization is one of the biggest parts of communicating with data. Visualization can help summarize patterns, interpret analytical results, and promote visual thinking and reasoning to understand these patterns, extract meaningful information, and develop

knowledge (Dodge, 2021). D'ignazio and Klein (2020) highlighted the idea that data visualizations should not only look professional, have a clean design, and only show the facts, instead, data visualizations should engage the emotions. Data science courses should reject the binary that reason and emotion cannot be in the same visualization (D'ignazio & Klein, 2020). Regular data alone often does not evoke any emotion until you see it in a visual form. For this reason, it is important that data visualization makes an appearance throughout the labs. I designed the labs in this study so that the students will have multiple opportunities to practice communicating their thoughts through individual reflection, visualization, and group discussion. The labs allow them to get experience engaging in statistical thinking and talking about data science in multiple ways.

Justice in Data Science

Along with communication, issues of social justice should play a large role in data science education. In their 2020 book, *Data Feminism*, D'ignazio and Klein discussed the importance of including social justice issues and principles from feminism in data science. They emphasize that data science needs intersectional feminism quite badly because in today's world, data is power. Throughout the book, D'ignazio and Klein described seven principles of data feminism: examining power, challenging power, rethinking binaries and hierarchies, elevating emotion and embodiment, embracing pluralism, considering context, and making labor visible (D'ignazio & Klein, 2020). All of these principles work together to operationalize feminism for data science and insert feminism ideas into data science. Data Feminism looks at all parts of feminism: political, social, and economic equality of the sexes and activist work. It also implies that feminism should be intersectional (D'ignazio & Klein, 2020). This book demonstrates a way of thinking about data, data analysis, and visual displays of data that is informed by feminist

activism and critical thought (D'ignazio & Klein, 2020). Many of these ideas can be seen throughout the labs in this study.

The notion of counter data is important in Data Feminism. Counter data is data that could have been collected, but was not (D'ignazio & Klein, 2020). D'ignazio and Klein used the analogy that data is the new oil. In other words, it is an untapped natural resource that can make you a huge profit if you figure out how to capture it, refine it, and make it accessible to the general public. Oftentimes, data is collected with the intention of making a profit for companies, and issues that are important to the general public or marginalized communities are often not attended to by large, powerful corporations. By understanding what counter data is, why it is not collected, and who it can benefit, students can think about the implications of data science and use it to make the world a more equitable place.

Lastly, data justice involves questioning the status quo. Regarding data science, this means that students should be questioning who is doing data science, as well as who is not. When doing data science is seen as the mastery of technology and skills, entire communities are often not engaged, which means important discussions are not being had and certain perspectives are being ignored (D'ignazio & Klein, 2020). Data Feminist design prioritizes the participation in data science of people who have been most marginalized by the system and looks at the people who could be the most harmed. Another question that should always be asked when doing data science is whose goals are prioritized and whose are not. D'ignazio and Klein (2020) mentioned that large companies or those who have the most money are often prioritized. This leads to important data not getting collected or analyzed. Finally, it is important to think about who benefits from data science and who does not. In order to be more inclusive, data feminism

requires an expanded definition of data science (D'ignazio & Klein, 2020), which involves changing the way that we think about and teach data science.

D'ignazio and Klein (2020) are building on the work of others who have critiqued science in general and traditional feminism. One example is Donna Haraway who argued for a change in how science and technology are viewed. Haraway (2013) discussed the idea of traditional feminism and how it operates under the assumption that all men are one way, and all women are another. She recommended that it is best for feminists move away from this idea and to blend identities. She also mentioned how new technologies can have benefits, but oftentimes women and people of color do not see these benefits (Haraway, 2013). More recently, Ruha Benjamin (2019) discussed the role of technology in perpetuating racism. Benjamin (2019) gave examples of how search engines like Google are created and used to discriminate against certain groups of people. Some of these included labeling Black neighborhoods as areas of high crime and reinforcing racist stereotypes. She also described this happening in other ways like automatic soap detectors not recognizing darker skin types. Like Haraway, Benjamin (2019) argued that both old and new technologies were created to sustain capitalism and that it is important that people understand why this is a problem.

Similarly, Safiya Noble addressed the issue of racism and bias in Google search algorithms in her 2018 book *Algorithms of Oppression*. Noble described how people claim that search engine algorithms are neutral, but this has proven to be untrue over and over again. These biases specifically have a negative impact among women of color (specifically Black girls) and other marginalized populations (Noble, 2018). Noble referred to Google as an "information monopoly" (p. 24) that has the power to prioritize web search results based on promoting their own business interests and those interests of multinational corporations. She described how this

is problematic and that we should not blindly be trusting the information that we get from these search engines. Noble (2018) also talked about the importance of having a data literate society. She also mentioned that the public is often unaware of these biases, which lead to problems such as racial and gender profiling, misrepresentation, and even economic relining (Noble, 2018, p. 28). One of the key ways that these issues can be addressed is through proper data science education that includes looking at these issues of data justice.

D'ignazio and Klein (2020) and Safiya Noble (2018) were some of the first people to discuss the idea of data justice, but after their books were published, others started talking more about data justice and about the importance of implementing it within data science curricula. For example, Green (2021) discussed how data scientists must orient the work that they do around addressing social justice issues and getting involved in the political process. Green (2021) mentioned the idea that many data scientists think of themselves as politically neutral and that this is problematic. It is important that data science is thought of as a form of political action and a way to make the world a better place. Green (2021) also mentioned that including politics and social justice in data science will require data scientists to restructure their values and practices.

These ideas about integrating social justice and data science have been making their way into both K-12 education and higher education. For example, the University of Virginia, in collaboration with The Equity Center, launched a summer research program for undergraduate students from groups that have been historically underrepresented in data science (University of Virginia, 2022). This nine-week program provides students with mentored research, technical skills training, data ethics and justice seminars, career exploration, relationship building, and both personal and professional development. Students can work on a variety of data justice projects that cover a variety of disciplines. For example, some projects involve topics like health

data and privacy, race disparities in criminal records in local cities, housing justice, and how to make machine learning algorithms less biased. Similarly, there are other initiatives outside of universities that aim to help integrate data science education and social justice. For example, the Data Science for Social Good (https://www.datascienceforsocialgood.org/) organization also trains data scientists to do projects that have a positive impact on society. Like the University of Virginia, Data Science for Social Good has a summer fellowship program through Carnegie Mellon University that lasts 12 weeks. The students in this program partner with nonprofit organizations and government agencies to work on data science problems that have high impact. Some of the topics of the projects include education, public health, criminal justice, and public safety.

Boenig-Liptsin et al. (2022) describe another example of how to incorporate data justice into data science education through their Data Science Ethos Lifecycle. They described the connection between ethical thinking and practicing data science, as well as how it is our responsibility to teach data science in a way that centers around social justice issues and real-world data science projects. The Ethos Lifecycle involves four dimensions that connect technology and the human: positionality, power, sociotechnical systems, and narratives (Boenig-Liptsin et al., 2022, p. 3). Many of these align with the principles of Data Feminism described by D'ignazio & Klein (2020). Another example of including data justice in education looks at taking a humanistic stance towards data science education at the K-12 level. Lee et. al (2021) described a framework that looks at how students understand the connection between data science and social justice issues. The framework has three layers: personal, cultural, and sociopolitical (Lee et al., 2021). Each layer has a description of the goals related to the layer and data science practices and each layer has example questions for instructors and researchers. The example

questions include items that discuss topics such as how data science affects marginalized communities, accessibility in data science, how cultural values impact how students interpret data. The goal of the framework presented by Lee et al. (2021) is for data science education to be human-centered starting at the earliest levels.

Justice is a journey and the uncomfortableness that comes from this journey is par for the course (D'ignazio & Klein, 2020). It is important that students are exposed to issues of justice and have the opportunity to analyze data related to social justice and have discussions about these issues. The labs in this study attempt to allow students to explore social justice issues and have individual reflection and group discussion with their peers about these issues. D'ignazio and Klein (2020) said that data are part of the problem, but they are also the solution to the problem. Data science instructors are tasked with the challenge of educating students so that they know how to use data in a productive way to help challenge institutional systems of power and make a positive impact on their communities.

Summary of the Literature

Not long ago, data was incredibly expensive, but now it is an untapped resource that is backing up (Finzer, 2013). Statistics education has a big role to play in the data science era as it is expanding to more and more fields (Finzer, 2013). Data science is an interdisciplinary field that involves statistics and computer science, as well as information science, communication, and data justice. There are many benefits to helping students become data literate, such as preparing students for their jobs and helping them to make a difference in the world by becoming critical consumers and producers of data who question the status quo. Because data science education is so new, there are very few studies done looking at how to improve communication in data science classrooms and labs. The labs in this study also contain social justice components and

allow students to have discussions about these issues. This study will help add to the literature on data science education, communication in data science, and justice in data science. Next, I discuss the theories and constructs that frame the work done in this study.

CHAPTER 3: THEORETICAL FRAMEWORK

In this chapter, I describe the theoretical framework that guides this study. First, I introduce the theory of distributed cognition. Many researchers agree that cognition is situated, embodied, and distributed. In this study, I specifically focus on cognition being distributed, by looking at the research done surrounding distributed cognition, with an emphasis on the work done by Edwin Hutchins studying cognition *in the wild*. Additionally, I explore how other researchers used distributed cognition as a framework in different contexts, including education. Lastly, I describe how and why designing data science labs from a perspective of distributed cognition can be beneficial. Secondly, I draw on the construct of computational action, which pertains to computational identity and digital empowerment. I describe the computational action construct in detail, the criteria for engaging in computational action, as well as what computational action brings to data science education. Overall, I use the theory of distributed cognition as to design the labs with the goal being for students to engage in computational action.

What is Distributed Cognition?

Distributed cognition (DC) is a theory that looks at how knowledge is allocated among individuals[3] and their surroundings. With distributed cognition, the unit of analysis is not the individual, but the entire system in which the individual is working, such as the other individuals involved, the setting, artifacts, and culture. There was a cognitive revolution in the late 1950s where cognitive anthropology drifted away from society and practice and started investigating what knowledge individuals have and how they obtain it (Hutchins, 1995). Specifically, cognitive psychologists were focused on the mind of individuals, what they knew, and what they

[3] When I refer to an individual or individuals, I am referring to a person or people.

needed to know to function in society. In more recent years, researchers started looking beyond the individual and realized that systems of socially distributed cognition may have interesting cognitive properties of their own that differ from the cognitive properties of the individuals within the systems (Hutchins, 1995). Hutchins (1995) was one of the first to study this in depth and when doing this, he emphasized the importance of studying systems in their natural habitat. He promoted the idea of *cognition in the wild* which referred to human cognition outside of the laboratory. To study this, it is important to not only look at the cognition of individuals in research labs, but also the cognition of entire systems that exist in the real world. Hutchins referred to this as naturally occurring, culturally constituted human activity, stressing the important role that culture plays when understanding cognition in the wild. Throughout his work, Hutchins made sure to highlight the distinction between cognition in the lab and cognition in the real world, which is dependent on surroundings that are unpredictable and can change frequently.

Cognition in the Wild

Edwin Hutchins' (1995) book, *Cognition in the Wild,* helped inform the definition of DC. Hutchins helped develop the theory of DC and his studies of ship navigation and aviation illustrate the components of the theory. Throughout his book, he describes how he spent a month aboard the ship, the U.S.S. *Palau*, studying the culture of the navigation team. The ship was the setting in which the studying took place and Hutchins explored what the navigation team knew and how they knew it. Because he was immersed in the community of navigators and working alongside the crew aboard the ship for a long period of time, his work was considered ethnographic. This allowed him to study the culture and social coordination of the navigation team in depth, in a way that is different from only a few observations. His ethnographic journey studying this team contrasted with studying the naturally situated cognition of individuals. After

his first journey at sea, Hutchins came to the conclusion that cognition is not only naturally situated, but also socially distributed (Hutchins, 1995). Hutchins (1995) was one of the first researchers to study the cognition of a team or group of individuals, rather than individuals on their own.

Throughout his book, Hutchins gave many examples of how DC could be used as a framework to analyze different tasks that were done by individuals while navigating the ship, *Palau*. These tasks required specific knowledge and Hutchins described how each person's individual job was integrated into the entire task of navigating the ship. Every crew member's job was a small part of the whole mission, but also essential to the success of the whole mission. These different jobs or tasks involved knowledge that was distributed among other crewmates as well as the tools that were available at the time. Additionally, in the very first chapter of *Cognition in the Wild*, Hutchins described a terrifying incident when the *Palau* was approaching a harbor and suddenly lost steam. The crew had to work together in a very short amount of time to navigate the ship into the harbor without going too fast or hitting sailboats and buoys. While describing this experience, Hutchins highlighted that the success of the crew was due to the distributed cognition amongst the crew working together as a system. The interdependence of the crewmates highlighted the importance of looking at other individuals when studying DC, making it one part of the system and framework for researching how certain tasks are accomplished.

In the situation where the ship lost steam, Hutchins emphasized that no single individual, even the navigator, could have kept control of the ship and brought it safely to anchor by themselves alone. As he described how they successfully brought the ship to anchor, he talked about how the knowledge to do this was distributed among the crew members. There were many types of knowledge and on-the-spot thinking required to perform this task. Some of that thinking

involved the crewmates working together, some involved them talking out loud, some thinking occurred inside their heads, and some occurred in small groups. The leaders of the navigation team, such as the navigator and the quartermaster, had to use their prior knowledge and experience to decide very quickly when to drop the anchor. This was essentially a life-or-death situation because if they did not drop the anchor at just the right time, the ship may have crashed.

While describing this situation, Hutchins used the entire navigation team, a system, as the unit of analysis, rather than the individual members of the team. The navigation team members did not have time for arguing and they had to react quickly and together. Navigating a ship in this type of situation is very difficult to do because stopping the engine may not stop the ship, inertia makes it slow to respond to changes in the propeller speed or rudders, and the ship is a massive object involving multiple people acting to control it (Hutchins, 1995). Furthermore, Hutchins (1995) stated that the situation was anything but routine, the ship was not fully under the control of the crew, and that lives were possibly in danger. This instance not only showed DC in action but highlighted that cognition in the wild is very different from cognition in the laboratory. Even though people may state what they would do in this situation, Hutchins' account unpacks real time decisions and how the navigation team made these decisions. This added another level of understanding of how the team accomplished tasks. This type of ethnographic research studied from the perspective of DC allowed Hutchins to analyze this team more completely than if he were to use previous cognition frameworks that did not take the setting into account. Through this example, it is clear that the setting is a crucial part of the system and the framework of DC.

Overall, Hutchins shared throughout the book how complex navigation is. It requires a lot of precision, tools, and knowledge as well as people acting alone, interacting in small groups,

and the entire team working together as a whole. Because navigation is both a cognitive and computational system, computation is a theme throughout the book. Hutchins mentioned the importance of studying computation within navigation as a key part of the system because computations are performed frequently to ensure that the navigation is on track and safe. If something happens out of the ordinary, new computations may be necessary. This computational knowledge of what computations need to be done, when to do them, and how to do them is distributed among different crew members and artifacts. Throughout the early part of *Cognition in the Wild,* Hutchins detailed how some of the artifacts available on the *Palau* made computations easier, such as the compass rose, fathometer, and charts constructed by previous navigators. All of these are considered internal tools located within the ship that are not absolutely necessary for success of the task, but help with different tasks, many of which are computational. When studying situations of DC, it is important to look at how internal tools, external tools, and the culture of the system you are studying play a role in completing tasks. In addition to internal artifacts, some artifacts are external and can be found in nature. Hutchins gave the example of landmarks which can help determine a ship's position, location, and bearings. These internal and external artifacts, both individually and together, as a part of the system, make navigation and the computations required for navigation easier for the team. Exploring how these artifacts play a role in the system helped Hutchins better understand navigation culture, how the system worked, and how the team used them to navigate the ship. To look at how the team accomplished this without the artifacts would not have given Hutchins a clear picture of their cognition, demonstrating the importance of including artifacts when using the framework of DC.

As mentioned previously, Hutchins also emphasized the idea of culture multiple times throughout his book. Part of his ethnographic journey was to study the culture on the ship, how cognition was distributed among different parts of the system, and how this contributed to the success of the team. He described the different roles each person played on the ship and how each role was essential in the navigation journey. Specifically, Hutchins said that from the point of view of the average citizen, a sailor is a sailor, but in the Navy and aboard *Palau*, the different titles and distinctions signified important subcultural identities (Hutchins, 1995). When discussing navigation culture, Hutchins said that it is not incredibly different depending on the ship or group of people navigating the ship. Hutchins described a situation where his colleague was doing research similar to his on another ship and the cultural differences that they observed were minor (Hutchins, 1995). Hutchins himself also went aboard other ships in addition to the *Palau* and the framework of individual and group cognition being distributed still held.

Although ship navigation has a specific culture, there are differences in crews on different ships. For example, one crew may not work as efficiently or may not know each other as well as another crew. One crew may also use different tools and artifacts to do the same tasks as another crew. Hutchins described the difference between European navigation and Micronesian navigation to further explain this idea. Micronesian cultures do not read or write in the language of European navigators so they have to memorize information that European cultures can write down or with which they can create charts. Micronesian navigators also used tools found in nature for a longer time than European navigators, rather than manufacturers' tools. In other words, advances in navigation were made differently and at different times depending on the culture or group of people doing the navigation. However, the knowledge was still distributed within the system. Not only should settings and artifacts be taken into

consideration when using the framework of DC, but also it is important to consider the culture of the people working together to accomplish the task.

Although many references to Hutchins studying DC refer back to his time aboard the *Palau*, he also used DC as a framework in other scenarios, such as the cockpit of an airplane. In this study, Hutchins and Klausen (1996) talked about the features of a system of distributed cognition from the point of view of a commercial airline cockpit. They reiterated the idea that the individual pilot is part of a much bigger system that performs the task of flying the airplane. They reminded the reader that when people are flying on a commercial airplane as a passenger, the question of interest should not be whether a particular pilot is performing well, but whether or not the system that is composed of the pilot, co-pilots, air traffic control, and technology in the cockpit are performing well (Hutchins & Klausen, 1996). Not only should the passengers consider this, but researchers analyzing how a plane is flown should also consider this. Hutchins and Klausen again showed that entire systems have cognitive properties and that they are best studied through DC using ethnographic methods where the researcher is up close and personal with the members of the system in their natural environment. This allows them to observe all individuals involved, the artifacts used, the culture of the team, and the natural setting.

The data in the airplane study is only from a small part of the flight but is very rich. Hutchins and Klausen (1996) collected audio data, video data, created transcripts, created documentation of what happened in the cockpit, created a translator to summarize what was said in a way that someone who does not know aviation could understand, and kept track of what tools the pilots used and with whom they interacted, when, and how. This thorough data collection led to cultural understanding of how the team flew the plane, what knowledge was needed to perform this task, and how this knowledge was distributed among the people and both

internal and external artifacts in the cockpit. Like navigating a large ship, Hutchins and Klausen realized that flying a commercial plane cannot be done by any individual alone and that the framework of DC worked well to study it. Throughout the paper, Hutchins and Klausen described how different properties of cognitive systems such as cognitive labor, access to information, and information storage, are distributed among different people and artifacts. Examples of the people in the cockpit include the pilots, co-pilots, and air traffic control, while examples of the artifacts in the cockpit include the instrument panel and flight controls. Similar to navigation, changes in the medium of representation of task relevant information can have important consequences for the cognitive functioning of the cockpit system (Hutchins & Klausen, 1996). Understanding these properties and how the system works to perform the task of flying the plane is an important step in understanding how these types of systems operate, how the knowledge is distributed, and how changes to the system could impact the team's ability to accomplish the task.

Distributed Cognition in Other Contexts

While Hutchins was a pioneer of DC, other researchers have also explored the theory of DC as well. Specifically, they elaborated on Hutchins' work and looked at what DC was and how it should be used to help better understand cognitive systems. These perspectives also helped add to understanding the DC theory. It is also important to note that many of these researchers' ideas were similar to those of Hutchins or based on his work. Rogers (1997) described how the theoretical and methodological base of the DC framework is derived from the cognitive sciences, cognitive anthropology, and the social sciences. Rogers (1997) also referred to distributed cognition as 'dcog' and specified that it is not a methodology that can be used, but instead a framework used to help explain the interactions between people and artifacts. This is an

important distinction. With this distinction, Rogers (1997) claimed that DC is not a methodology and therefore does not determine how one studies the task but rather it can be used as a framework that helps researchers see the world through this lens. This allows them to better understand how knowledge is distributed among different parts of a system.

Multiple people further explored Hutchins' idea of internal and external representations within the distributed cognition system. Internal and external representations are artifacts that are included in the system. Zhang and Patel (2006) described DC as a cognitive system whose structures and processes are distributed between internal and external representations (or artifacts), across a group of individuals, and across space and time. Here he specified that the more we study the interaction between the elements in the system, the more we can use DC to learn how these things interact. Giere (2007) elaborated on the importance of external representations. He described them as representations of aspects of the world that are not localized in a person's brain or in a computer, but rather, somewhere external to these locations (Giere, 2007). Including these types of artifacts in the framework made DC new and, in some ways, more powerful than other types of cognition. In addition to external representations, it is important to note that the internal representations are an important part of the distributed cognition system as well. Sutton (2006) emphasized that the DC framework does not assume both internal and external representations do the same thing, have the same characteristics, or are identical. Instead, all the parts of a system, including these artifacts studied using DC complement each other and work together (Sutton, 2006). Learning how the parts of each system, such as artifacts and individuals, work together to perform a task helps us understand the system and the individuals in the system more thoroughly.

Like Hutchins, Lave (1988) described the need for studying cognition outside of the lab, thereby emphasizing Hutchins' desire for studying cognition in the wild as well as the impact of setting on a system's ability to perform a task. Lave referred to studying distributed cognition in confined spaces such as laboratories as a "claustrophobic view of cognition" and said we should move away from this because cognition in everyday practice is distributed differently depending on the situation (Lave, 1988). Achiam et al. (2014) used DC, along with the notion of affordance to explain how visitors interacted with exhibits in a natural history museum to learn and make sense of the exhibits. They used DC to explore visitors' encounters with exhibits, interactions, explanations, and connections made while at the museum (Achiam et al., 2014). The visitor and the exhibit formed a distributed cognition system, where knowledge, practice, and meaning making occurred with both internal and external representations such as the entire exhibit and its components and the individual's mind, culture, and thoughts. By studying the participants in the context of the museum exhibit rather than a laboratory it was possible to see even more artifacts and how they impacted the individual's meaning-making and learning through the framework of DC. This is similar to studying students' communication in data science labs.

Giere and Moffat (2003) gave a very simple example of a system where the cognition was also distributed among people and artifacts. This involved people solving the following multiplication problem: 456*789. Giere and Moffat (2003) mentioned that people like Hutchins who advocated for distributed cognition liked to talk about cognitive tasks such as this. The cognitive system performing the multiplication task would be the person, their knowledge, and any external representations they used to complete the task such as a paper and pencil, all of which were necessary to complete the task (Giere & Moffat, 2003). Although this is an example

of a very simple system of DC, the properties of this distributed system are similar to many of the properties that Hutchins and others have described.

Using the DC theory is not just limited to analyzing complex tasks such as navigating a ship or flying a plane. Another example of applying the DC theory is the work of Rogers and Ellis (1994), which occurred around the same time frame as Hutchins, in which they explored collaborative work environments in the engineering and information technology (IT) setting. They argued that existing theoretical frameworks were not sufficient for studying collaborative work environments because it made accounting for both social and cognitive interactions difficult while using one framework. The focus of the analysis done by Rogers and Ellis (1994) was on the relationships between the individuals and artifacts and how they coordinated and worked together in the IT setting. During the same time period, Boland Jr. et al. (1994) also studied how to use DC to design ways to help support IT in an organizational setting. Both Rogers and Ellis and Boland Jr. et al. saw that knowledge in an IT setting was distributed among the people, artifacts, culture, and setting and that looking at the whole system rather than the individuals within the system allowed researchers to understand how the system worked to accomplish tasks.

Distributed Cognition and Education

As has been seen with navigating a ship and flying a plane, using the framework of DC more fully allows researchers to analyze how the teams involved completed their tasks. Similarly, classrooms should be analyzed through the framework of DC in order to more completely see how teachers teach and students learn. In the early 2000s, DC started making its way into education research. Before then, it was a common belief in education that cognition resided in the individual head (Karasavvidis, 2002). Karasavvidis (2002) sought to investigate

the implications for teaching and learning, given the material and social dimensions of DC. He did this through examples that looked at both the material (for instance, the artifacts used) and social (for instance, the cultural) aspects of DC.

Regarding the material aspect, he used simple examples, such as students solving a correlation question. Karasavvidis (2002) compared students doing this by hand using paper, a pencil, and a graph to students solving the same question using a computer spreadsheet. In both cases, the knowledge was distributed among the students and their different materials (or artifacts) and the type of knowledge necessary to solve the problem differed depending on the artifacts being used. The computer spreadsheet as an aid to get to the solution changed what it meant to solve correlation problems (Karasavvidis, 2002). It made the problem faster, helped reinforce the knowledge, and lowered the mental processing and cognitive labor. The learning process changed so much that Karasavvidis (2002) recommended the learning goals of this problem were changed. According to Karasavvidis (2002), we cannot introduce tools such as computers into classrooms and keep the same learning objectives. Since the artifacts changed, the way the cognition was distributed also changed. The use of the computer spreadsheet led to a lighter cognitive load for the individual as the spreadsheet did most of the computational work. This is just one example showing the need for DC as a framework to study education. If Karasavvidis did not look at the system as a whole (including the changing artifacts), but rather only considered cognition as what goes on inside the individual's mind, the change of cognitive load may have been noticed but not understood. By understanding these changes, researchers, such as Karasavvidis, can share these observations with teachers, who can then use them to inform their teaching.

Karasavvidis (2002) also described the social aspect of DC and said that students should be allowed to collaborate with each other more since this was how they would be learning outside of the classroom. Karasavvidis (2002) emphasized that the process of learning was inherently social in nature and that students' proficiency was distributed. He advised that teachers should focus on more practical and situated tasks that do not have only one solution or even tasks that do not have solutions at all to help encourage collaboration among students and better mirror learning in real life (Karasavvidis, 2002). Hutchins also believed that the crewmen on the ship, *Palau,* did not need to know how to do each individual's task, but they did need to be able to work together to navigate the ship. Similarly, students should not need to be experts in each piece of information studied in class, but rather how to work together as a part of a system to accomplish real world tasks set forth by the teacher. In recent years, collaboration and problems with multiple solutions seem to be emphasized more often in education, which leads to the need of the framework of DC by researchers so that they can study the system of the classroom as a whole.

Distributed cognition has also been used as a framework for studying teaching and learning in more recent years. Cognitive processes involved in the mastery of tasks and the process of learning are not an individual matter, but instead they are distributed among the teachers, students, and cultural artifacts used in the activities (Cole & Engestrom, 1993; Angeli, 2008; Salomon, 1992). Examples of educational tasks that can be framed using DC are tasks such as learning to read (Snow, 2011), using computers (e.g., Angeli, 2008), other technology (e.g., Evans et al., 2011), and even foreign language learning (Narciss & Koerndle, 2008). Research done using the distributed cognition framework often involved teaching and learning using technology. To understand learning in technology enhanced classrooms, it is important to

view the classroom as a system that is cognitively distributed and study the entire classroom as a system, including but not limited to the individuals within it (Angeli, 2008; Salomon, 1992). Angeli (2008) gave examples of multiple studies that showed that the framework of DC worked well to help researchers understand the human aspects of cognition and computers, suggesting that DC could be used as a guide to integrating technology into education successfully. Learning and classrooms are very complex and have uncertainty, so it is important to try to understand how the different components (people, technology, and artifacts) interact with each other.

Technology is often found in mathematics classrooms specifically. Evans et al. (2011) examined and coded how elementary school children communicated when solving a geometric puzzle in a computer supported collaborative learning (CSCL) context and group setting. They used DC as a way to understand CSCL, children's mathematical learning, and the use of technology in a classroom. They specifically looked for instances of distributed cognition in what the students said, their gestures, and the artifacts they used. Some of their findings included that learners were more likely to discuss and articulate their ideas in the CSCL setting, but there was less gestural communication in the CSCL settings. They saw that students who were more hands-on may struggle more in the CSCL space because of this. They also found a natural emergence of group leaders (Evans et al., 2011). In other words, certain students naturally took control of the group as leaders. If the researchers were only looking at the individual's cognition, these observations would be missed.

The distributed cognition approach is a viable framework to understand the relationships and interactions between the students and their graphing calculators. Similar to what Karasavvidis (2002) said about computers, graphing calculators can ease the cognitive burden and enable performance. Therefore, DC perspectives can explain the reasons why the use of a

graphing calculator will not hinder learning mathematics. If anything, it allows the learner to focus more on problem solving by lessening the cognitive load (Tajuddin et al., 2009). Sivasubramaniam (2004) also looked at graphing calculators and DC. He believed that graphing calculators were more powerful tools than using paper and pencil to produce graphs. He made three very important points in his study: 1) Technology including graphing calculators cannot replace the human mind, 2) Graphing calculators should be viewed as tools to help solve mathematical problems, not a tool that solves mathematical problems, 3) The skills that math students attain may be attained without using a graphing calculator, but the graphing calculator accelerates the process of attaining these skills for many students (Sivasubramaniam, 2004). In other words, technology plays an important role in distributed cognition systems, such as classrooms, and it is clear that more research needs to be done looking at teaching and learning through the perspective of DC, so such aspects are not missed.

Limitations of Distributed Cognition in Education

There are many benefits of using distributed cognition as a framework to study education. It allows us to understand learning more thoroughly because we are branching out beyond solely looking at what goes on inside the human mind. Also, it allows us to see how culture, environment, and artifacts play a role in how people learn and accomplish tasks. Studying teaching and learning through a DC perspective solves many problems that researchers have had in the past, such as understanding how artifacts play a role in teaching and learning. However, there are some challenges that come with using this framework in education. Narciss and Koerndle (2008) pointed out that the classroom is very different from a cockpit of a plane or a naval ship. In many classrooms, if students are not used to working collaboratively, the teacher needs time to help get them comfortable doing this. However, this is not an option for the crew

on a ship or plane. Hence, more research needs to be done on how to best prepare teachers for leading distributed cognition classrooms (Angeli, 2008). Martin et al. (2019) started looking at this by investigating distributed scaffolding, which included features of distributed cognition such as how artifacts and other individuals play a role in students learning in the classroom. Teachers have to attend to many students all at once in an environment where student learning and cognition is distributed among the teacher, the other students, the scaffolds, and other materials. Distributing support across multiple tools, resources, and agents can address the challenge of supporting multiple students with very diverse needs in a classroom (Martin et al., 2019). It is clear that the DC framework could be used in other studies to analyze the system of the classroom when utilizing distributed scaffolding. However, such distributed scaffolding is still an incredibly difficult task. It would be useful to explore how teaching with distributed scaffolding in addition to teaching using DC as a framework could be a part of professional development for teachers.

Karasavvidis (2002) drew attention to traditional methods of teaching and learning that make it difficult to design assessments for classrooms where cognition is distributed. Because most of the current educational practice was founded on the assumption that cognition resides in the individual head, the idea of the individual student as the sole bearer of all cognition is manifested in the teaching and learning methods, including assessment or exams (Karasavvidis, 2002). Karasavvidis (2002) mentioned how unrealistic it would be to assume that students would learn in the real world the way they do in most classrooms by the teacher lecturing, the students storing that information in their minds, and reproducing it on an exam. Assessment in these new types of classrooms would have to be rethought to include group work and resources that

previously were not allowed on exams that were meant to be done individually. Designing such assessments effectively would require more teacher education and professional development.

Lastly, it is definitely important to mention issues of equity that may arise in a classroom where cognition is distributed among students, teachers, and artifacts. Cobb (2006) was one of the only researchers to mention the idea that distributed cognition theorists have only given a small amount of focus to issues of equity, specifically in students' mathematical learning. Cobb (2006) mentioned that some students may not be engaged in the activities the students do in class due to cultural differences or lack of access to certain artifacts to help them learn. He also mentioned that this limitation can be addressed by focusing on both a distributed perspective and emphasizing the sociocultural perspective of student learning. Teachers should not only focus on the current learning environment they are operating in their classrooms, but also students' activities and home lives outside of school when designing their activities. With any framework, it is important to discuss and try to remedy any limitations or weaknesses, however, despite the few that have been mentioned, DC is a framework that has positively changed the way we think about education and has helped us better understand student learning. As we research DC in other contexts, such as statistics and data science classrooms, it is important to keep these limitations in mind.

The Practice of Data Science and Distributed Cognition

The practice of data science is a context for learning from a distributed cognition perspective. As seen through the many examples given from previous researchers, such as Hutchins, distributed cognition is a framework that looks at how knowledge is distributed or allocated among individuals and their surroundings. Data science is a discipline that is relevant to many aspects of life. Data scientists work at companies, at hospitals, with sports teams, at

universities, and in many other industries. Students who major in data science often combine it with another field because it is a useful tool to help us understand the world in which we live. As data scientists work in their selected field, their knowledge is distributed among their workplace environment, the colleagues with whom they work, the tools that they use, and the culture of the workplace. Understanding how these aspects of their jobs work together helps us understand how data scientists make decisions and do their jobs.

A data scientist working at a company can be compared to the navigators aboard a ship as Hutchins described. For example, data scientists who work at an insurance company are a small group of individuals who help the company complete their goals of selling insurance and helping customers. They have a very specific role that they play, but they also have to understand how their role fits into the system and how some of the other units work. This is similar to how each small team on the *Palau* played a very important role in navigating the ship and how they had to have knowledge about the other teams and how their work helped with accomplishing the main task of navigation. What the data scientists do influences how the company makes decisions and the company would have a hard time being successful without data scientists. Also, the data scientists who are not in job situations also are a part of an even bigger system. People who are applying data science to understand important issues are a small part of a much bigger system that consists of their culture, artifacts, other people, and their knowledge.

Data scientists in all fields engage in data science practices. Many data science practices involve ideas from distributed cognition. For example, creating representations and using them to explain a concept is a data science practice. Data scientists at companies use artifacts such as programming languages to create representations of data that they can use to explain something to others. Communicating using data science representations is also a data science practice. The

goal of the labs in this study is for students to engage in data science practices and act as real data scientists. They do this through using real tools and real datasets related to social justice issues. They explore these issues in their labs and have opportunities to communicate with their group members.

Distributed Cognition in Data Science Labs

Similarly, the idea of data science education can also be studied from a distributed perspective. We have seen through the examples of studies with graphing calculators and technology how DC plays a role in mathematics education and data science education has many similarities to mathematics education. Specifically in this study, I look at how the labs in a data science course can be designed using principles of DC. I am looking at how this design perspective can improve how students communicate in the labs. In the labs, knowledge is distributed between the individual students, their group members, the artifacts that they have access to, as well as the scaffolds in the labs themselves. Martin et al. (2019) mentioned that distributing support across multiple artifacts and scaffolds can help support many students in a classroom. I hypothesize that by using DC to design the labs, the students will have multiple opportunities to communicate and reflect with their peers.

Scaffolds and Artifacts

I designed the labs so that students work together with the scaffolds and artifacts to get practice doing data science and understanding how what they learn in the labs can help them in their jobs and their lives. One specific goal is to improve communication in the labs and include group discussion. The scaffolds designed from a DC perspective can help the students achieve the goal of improving communication. Since there are multiple small groups in each lab and only one TA, the scaffolds were carefully designed using DC principles to facilitate multiple different

types of communication. There are three main artifacts in the labs: 1) the datasets that the students are analyzing, 2) Python, and 3) Jupyter notebooks that allow students to type code and text into the same document. These artifacts also contributed to the facilitation of communication. Like the artifacts, there are also three types of scaffolds in each of the labs. The first type are coding questions, which allow the students to use programming to visualize their findings. Because communication is multimodal, visualization is an important form of communication. These coding questions involve using two artifacts, Python and Jupyter notebooks. The second type of scaffolds are individual reflection questions. These questions allow the students to use their knowledge from the lab, from the lecture, their specific field, and their lives to think about the implications of the data analysis they are doing. They also help them get practice with written communication. The students type their answer to these questions directly into the Jupyter notebooks, making them a key artifact in this part of the lab.

The last type of scaffolds are group discussion questions. By designing the discussion questions drawing from a DC perspective, I intended for the students to practice communicating with each other, working together, and having discussions about data justice. These questions are the most open-ended, and I planned to give the students freedom think about what questions they want to answer from the data provided, thereby acting as real data scientists. This would then achieve what Karasavvidis (2002) discussed when he explored the social dimension of DC. They also encourage students to think about the analysis they did and the social implications of it, along with how data science can be used as a tool to help understand social justice issues. Through all of these types of questions, the students are learning from their group members (their ideas and the discussion), the scaffolds (different types of questions), and the artifacts involved in the labs (the datasets, Python, and Jupyter notebooks). This type of learning environment was

designed using the DC framework since the unit of analysis is a system, rather than individual students. I intended to have the students focus more on communication in many forms throughout the labs.

Thinking about any of the parts of the system individually does not give us a clear picture to design the labs because learning data science in this context is enhanced by including other people, artifacts, and scaffolds. Designing the labs using DC principles aims to achieve what Karasavvidis (2002) discussed when he described how traditional exams are not realistic contexts for demonstrating statistical knowledge. Labs allow students to demonstrate their knowledge in multiple ways and the labs provide a more realistic context for doing data science because they involve artifacts, scaffolds, and collaboration with other people. Data science is a very collaborative discipline that can be embedded into other subjects, is useful in many aspects of the real world, and can be used to help make the world a better place. Designing and implementing these labs is an example of doing education research "in the wild" since we are making changes in an actual classroom, rather than studying innovations in a laboratory. There are very few, if any, studies on data science curriculum design using DC, making this a gap in the literature that should be explored. Next, I will describe the main construct in this study, *computational action*.

What is Computational Action?

Computational action (CA) is a relatively new construct for computing education that focuses on encouraging students to use the programming or computing that they learn in the classroom in their lives outside of school. Tissenbaum et al. (2019) described this idea by saying that "while learning about computing, young people should also have opportunities to create with computing that have direct impact on their lives and their communities" (p. 34). Tissenbaum et

al. (2019) described two components to computational action: computational identity and digital empowerment. Computational identity allows students to recognize themselves as scientists who can design solutions to problems that are important to them. Computational identity also allows students to see themselves as a member of a community of others who can do the same. Digital empowerment allows students to have the belief that they can put their computational identity in action to think about issues that are important to them (Tissenbaum et al., 2019). Both computational identity and digital empowerment make up CA.

In the past, subjects like mathematics and computer science have been seen as topics that students only explore and utilize in the classroom and have no relation to students' lives outside of school. Tissenbaum et al. (2019) argued that the focus in computer science education is on the basics which includes the nuanced elements of computation, such as variables, loops, conditionals, parallelism, operators, and data handling. Students often fail to see how they can use their computer science knowledge and skills to do anything outside of class. Researchers have expressed similar concerns about statistics education. For a while, scholars have recommended that basic introductory statistics courses are reformed to include more data analysis and exploration and less theory and recipes (Gal, 1997). Data science education with the goal of computational action can help address this problem. By having students learn statistics through doing computer science and using real world datasets in the labs, it will be easier for students to see how these skills can be used in their lives and in their jobs.

Computational Action in Data Science Labs

Computational action helps students answer the age-old question of "When am I going to use this in real life?" Data science already uses computer programming skills, but I am proposing a focus on CA because it can help students go beyond performing operations by allowing them to

choose problems that are relevant to them and will help build their identity. CA foregrounds student choice and allows students to choose to do work on topics they are interested in. The data science labs will contain many elements of student choice and multiple scaffolds that have the students think about issues that matter to them and problems that they would like to solve. In designing the labs, I intend for students to engage in computational identity by positioning them as members of a group whose ideas are important. They also allow the students to put their computational identity into action by thinking critically about the implications of collecting data and analyzing it to answer questions that are important to them. The individual reflection and group discussion questions let the students think about being critical consumers of data who use data science to make the world more equitable.

Criteria for Engaging in Computational Action

Tissenbaum et al. (2019) developed a set of criteria for engaging in CA (computational identity and digital empowerment) and I will describe how each of these are seen in the labs. The first criterion for computational identity says that students should feel that they are designing solutions, rather than working towards predetermined correct solutions (Tissenbaum et al., 2019). The scaffolds in the labs foster creativity in how the students answer the coding questions. They also have been designed for the students to think about and discuss questions that are open-ended and do not have right answers. The second criterion for computational identity says that students need to feel that the work they are doing is authentic and applicable to scientific communities. The artifacts used in the labs are industry standard tools as both Python and Jupyter are commonly used by data scientists. The labs in this study are centered around social justice issues. By working on problems involving social justice issues, the students are able to

discuss authentic problems that do not have one right answer. The labs allow the students to share their unique perspective on these issues with their group members.

The first criterion for digital empowerment says that the activities students do should be authentic and relevant (Tissenbaum et al., 2019). Both labs involve the students working with or simulating real data to answer real questions. The communication parts of the labs are also authentic because data scientists need to be able to communicate their results and have discussions about the work they do. The second criterion for digital empowerment says that the students need to see how their work could impact their lives and communities (Tissenbaum et al., 2019). This is difficult to do in an introductory course, however, designing the labs around social justice issues allows students to see the connection between what they are doing in the classroom and the world outside of the classroom. The coding questions in the labs are simple, yet powerful and I intend for them to have a lot of impact. Also, during the individual reflection and group discussion parts of the lab, the students can discuss how data science can impact their lives and communities through the social justice examples. The third and final criterion for digital empowerment says that the students should feel confident that this computational work puts them in a position to do new computational work (Tissenbaum et al., 2019). Both the individual reflection and group discussion parts of the labs involve thinking about and discussing new data science problems that the students are interested in solving. The students discuss what data they would need to collect and how they could go about doing the analysis. They also discuss the implications of collecting this data, doing the analysis, and who benefits from it. Overall, designing the labs with the goal of the students engaging CA helps make the labs relevant and meaningful for the students, as well as prepares the students for the opportunity to use data science to make the world a more just place.

Summary of Theoretical Framework

The field of data science education is unique because it is so new. This study looks at how to design curriculum for data science labs. The DC theory allows me to think about how students learn data science and communicate with others with the help of scaffolds and artifacts. The CA construct allows me to have a goal for the curriculum that I am designing. The artifacts and scaffolds designed using the DC principles help the students engage in CA and social justice awareness. Using this theory and this construct together allows me to design curriculum with meaningful goals in mind.

CHAPTER 4: METHOD

Course Description

This study took place in a data science course at a large midwestern University. The course is an introductory data science course that focuses on combining statistics and inferential thinking with computation using Python. This course is a 15 week-long four-credit-hour undergraduate course. The course has no pre-requisites and is designed for students who have no prior knowledge in statistics or programming. It also fulfills the quantitative reasoning general education requirement (this is generally called a "math gen-ed") at the university. In addition, the course is a first-semester requirement for students who are majoring in Statistics and students who are majoring in Information Science. The course covers introductory statistics topics such as experimental design, descriptive statistics, basic data visualization, probability, inference, hypothesis testing, and a small amount of machine learning towards the end. The students learn these concepts through programming in Python.

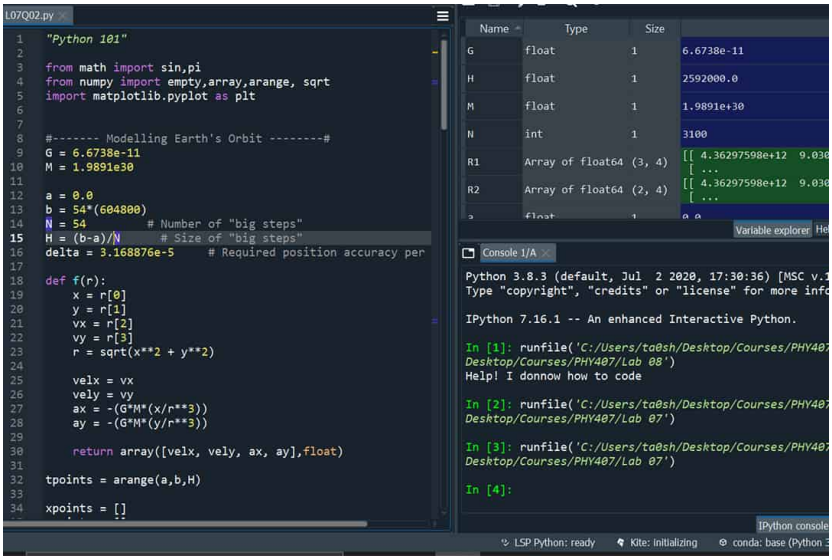
Python

There is a healthy debate over what the best programming language for learning data science is, however, many universities are choosing to use Python for their data science courses (Grus, 2019). Python is an object-oriented high level programming language that is easy to use. Python is a general programming language that can perform most, if not all, of the data analysis operations that a data scientist might need, especially when using the Pandas library (Brunner & Kim, 2016). Python is considered open source, meaning that anyone can use, study, change, and distribute the software and its source code. Python has many advantages and was chosen for the

course in this study[4] because of the reasons that Grus (2019) described: it's free, relatively simple to understand and program in, and has many useful data science libraries, such as Pandas which allows students to easily analyze and manipulate data. It also is an industry standard tool that is widely used by different companies. Figure 2 shows what the Python interface looks like.

Figure 2

Python Interface



Source: https://thevarsity.ca/2020/11/15/how-to-learn-to-code-latex-python-and-r/

Discussion Sections

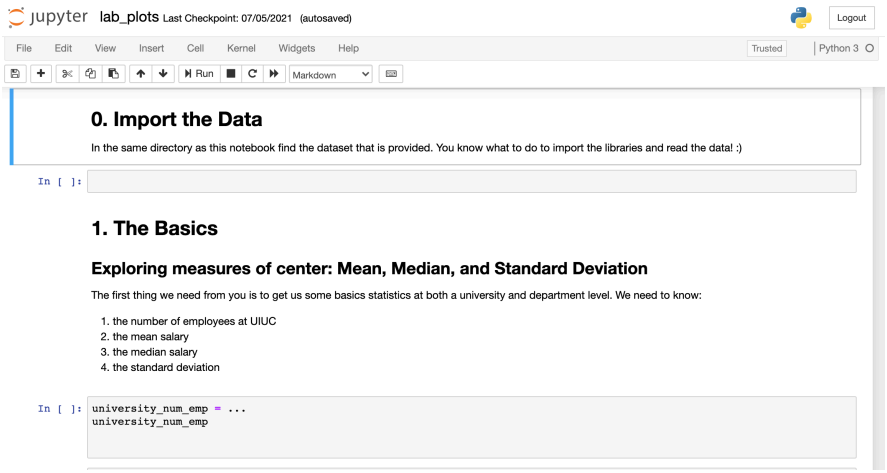
The data science course in this study is a large lecture course where the students attend 50 minutes of lecture three times each week and they spend 80 minutes in a discussion section once each week. The lecture usually has over 300 students enrolled, but the discussion sections are capped at 30 students maximum. These discussion sections are run and led by graduate teaching assistants (TAs). During the discussion sections, students work in groups of two or three to

[4] The choice of using Python was part of the original course design, which occurred before this study. Many introductory data science courses at other universities are also choosing to use Python.

complete an interactive lab that was created using Jupyter notebooks. Jupyter notebooks allow the students to program and enter text responses within the same document. Figure 3 shows what the Jupyter notebooks look like to the students. This data science course is a course that teaches students introductory statistics through programming in Python. During the lab sections, the students get practice programming in Python to answer statistical questions, explore data, and be creative. They also work with students in small groups and get help from TAs. The labs give the students an opportunity to try new problems on their own and practice communicating their findings to others.

Figure 3

Jupyter Notebook Interface



Content of the Labs

The two labs that I revised in this study were completed during Week 5 and Week 10 of the semester. The first lab covered visual displays of data and descriptive statistics. I will refer to the first lab as Lab A. See Appendix A for the version of Lab A that the students got, Appendix B for Lab A coded with scaffolds and artifacts, and Appendix C for a table describing each scaffold and artifact, including the data I got from them and why they were included. The second

lab covered sampling, simulation, and the Central Limit Theorem and I will refer to it as Lab B (see Appendix F for Lab B). When selecting the labs for this study, I chose labs that could be revised to include social justice issues and real datasets. I also chose labs that occurred after the first few weeks of the semester so that the students would be familiar with Python and have learned about basic descriptive statistics such as mean, median, and standard deviation. This allowed me to focus on social justice and communication when designing the lab, rather than teaching them Python or teaching them basic statistics. In both labs, the students used statistics and Python to think about important issues and how they can use data science to answer questions regarding equity and social justice. The scaffolds and artifacts in both labs encourage students to communicate their findings through talking and writing. Both labs also have scaffolds that encourage students to have discussions with their peers and think about the implications of their analysis.

During Lab A, the students used Python to analyze a real dataset involving salary data by looking at descriptive statistics and standard visual displays such as boxplots and histograms. They used these statistical tools to answer specific questions that involve using programming, short written answers, and group discussion. Throughout the lab, they also thought about the implications of the data collection and analysis as well as explored questions that they personally found interesting and thought about how they might go about answering them. During Lab B, the students used Python for simulating different scenarios involving interview data and jury data. For the interview data, the students looked at how resumes that are identical except for the name can get different amounts of attention. The jury data allowed the students to explore random sampling and whether juries truly come from random samples when considering ethnicities of

the jurors. Both scenarios involved creating simulations that showed the students what they were expected to get and how that was similar or different from what was actually reported.

Participants

The targeted participants for this study were students in the data science course that I previously described. All students enrolled in the in-person lab sections were emailed about participating in the study. In the email, the students were told that they could attend "office hours" with a member of the research team to ask questions if they had any. A member of the research team hosted these office hours online and was available to answer any questions. Since the goal was to recruit as many participants as possible, no specific age, race, sex, or socioeconomic status was targeted. There were 17 students who gave consent in Fall 2021 and 26 students who gave consent in Spring 2022. I gave the students gender neutral pseudonyms that I used throughout the results section.

Recruitment and Compensation

The email that the students received also contained a link for them to click on to fill out consent for either one and/or both parts of the study. Part 1 of the study did not involve extra work outside of the required course assignments as the activities were embedded into the Stat 107 course. The students were told that Part 1 involved completing two labs and completing four surveys (one before and after each lab). See Appendix D for an example of the pre-lab and post-lab surveys. The students were also told that when completing the labs, their conversations would be audio-recorded, but they would not be video-recorded. Part 2 of this study did involve extra work outside of the required course assignments. Part 2 involved giving consent to be interviewed outside of class after all course grades had been finalized. See Appendix E for the interview protocol. The students were told that the interviews would be audio-recorded and 30

minutes long. Everyone who consented to being interviewed was not interviewed. Interviewees were selected randomly. The students could participate in Part 1 only, Part 1 and Part 2, or choose not to participate in Part 1 or 2 and not release their data for research purposes.

If the students consented to release their data for research purposes and completed all of Part 1 (four surveys and two labs), they were automatically entered in a lottery drawing to win one of five $25 Amazon e-gift cards. Selection of winners was completely random and conducted at the end of each semester. If they consented to be interviewed for Part 2, they were automatically entered in a second lottery drawing to win one of two $50 Amazon e-gift cards. In other words, if they consent to Part 1 and Part 2, they were entered into two lottery drawings.

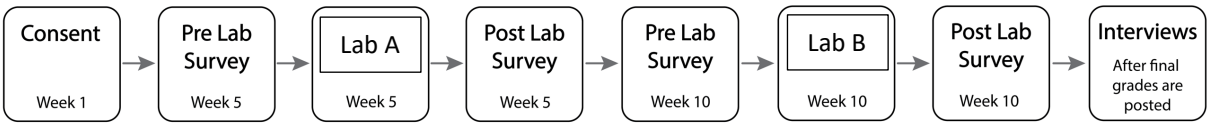
Design-Based Research (DBR)

This study was done using design-based research over two semesters (Fall 2021 and Spring 2022) where the first semester research informed the second semester research. I specifically studied two labs, Lab A and Lab B, during each semester. The students who consented completed the following tasks in the following order. During Fall 2021, the students completed a survey before each lab. Next, they completed the two labs in their discussion sections and the conversations they had while completing those labs were audio-recorded. Afterwards, everyone took a survey once the lab was completed. At the end of the semester, ten students who gave consent were randomly selected to be interviewed about their experience completing the labs. Of those ten students, seven were able to be interviewed. All of the data collected in Fall 2021 informed the changes made to the labs for Spring 2022. During Spring 2022, the students completed a survey before doing each lab, then they completed the labs in their discussion sections which were also audio-recorded, and then took a survey when they were done with the labs. At the end of the semester, ten students who gave consent were also

randomly selected to be interviewed about their experience completing the labs. Of those ten students, seven were able to be interviewed. Figure 4 shows the sequence that the students followed during both semesters.

Figure 4

Sequence of Student Data Collection



Because this study is situated in a real educational context, focused on the design and testing of an intervention, iterative, collaborative, flexible, and has a practical impact on the practice of teaching, I chose to use design-based research (DBR) as the method. Figure 5 shows a conceptual way to think about the implementation of the labs.

Figure 5

Conceptual Implementation of the Labs

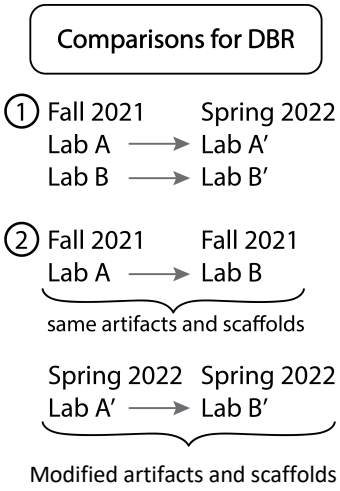


Figure 5 shows that I can make multiple comparisons. Throughout the study, I was able to compare Lab A with Lab B, Lab A' with Lab B', as well as Lab A with Lab A', and Lab B

with B'. Although Lab A and Lab B covered different topics, they had similarly designed scaffolds and artifacts. In other words, their content is different, but the design components are the same. I used student feedback through student interviews to improve Lab A' and Lab B' in Spring 2022. Next, I define DBR and discuss it by looking at how education researchers used DBR for their work. While doing this, I include underlying assumptions, benefits, and critiques to this approach. I review the literature of DBR in this section because it helped me to frame the discussion about the methods applied for designing a new curriculum for the labs.

What is Design-Based Research?

Throughout the years, education researchers have worked tirelessly to improve the teaching and learning that goes on in classrooms. However, until recently, many of these studies have been unilateral and disappear with the researcher once the experiment has been concluded (Anderson & Shattuck, 2012). As stated earlier, *design-based research* or DBR is one methodology that focuses on moving from research in the laboratory to the practical impact of improving the teaching and learning that occurs in the classroom. In other words, it is a method that helps improve both educational theory and practice. DBR is unique because it contains aspects of multiple disciplines, including developmental psychology, cognitive science, learning sciences, anthropology, and sociology (Sandoval & Bell, 2004). DBR has dramatically increased since 2000 and is especially prominent in the learning sciences research. Design-based research (DBR) is a methodology used primarily for education research that is situated in a real educational context, focuses on the design and testing of an intervention, is iterative, collaborative, flexible, and has a practical impact on the practice of teaching.

Characteristics of DBR

DBR is a relatively new approach to research. In the early 1990s, Alan Collins (1992) and Ann Brown (1992) started conducting a new type of research that they referred to as "design experiments." The idea behind these design experiments was for researchers to be able to study teaching and learning in a natural environment. Since then, design experiments have filled a gap in the category of experimental methodologies that is needed to improve educational practices (Collins et al., 2004). Design experiments are now more commonly known as design-based research (DBR) in the education research community, and DBR has grown in popularity and significance. Despite this, researchers are still working towards agreement on what constitutes design-based research, why it is important, and what methods should be used for carrying it out (Barab & Squire, 2004). Anderson and Shattuck (2012) used Google Scholar to identify articles that have been published and are related to DBR and education. They found 1,940 articles from the search but reviewed 47 of them from 2002 to 2011. This collection of 47 articles contained the articles that were cited most often each year. Because DBR is an emerging research framework, 88% (14 out of 16) of the philosophical or expository articles about DBR were written from 2002 to 2006, and 74% (23 out of 31) of the empirical studies were written later, from 2007 to 2011 (Anderson & Shattuck, 2012). These results showed that the design is moving from theoretical discussion, or theory, to practice (Anderson & Shattuck, 2012). Anderson and Shattuck (2012) also found that DBR is mostly happening in the US, as 73% of the articles in the study were published there. There are six main characteristics of DBR studies: 1) situated in real educational contexts, 2) focuses on design and testing of an intervention, 3) iterative, 4) collaborative, 5) flexible, and 6) have a practice impact on the practice of teaching. Together, these characteristics make DBR both unique and beneficial to students, teachers, and researchers.

Situated in Real Educational Contexts. Design-based research is situated in real educational contexts. Anderson and Shattuck (2012) found that DBR is mainly occurring in classrooms, as all of the studies took place in educational contexts and all of the papers were published in education related journals. It also seems to be especially popular for use in K–12 schools, in the learning sciences discipline, and with technological interventions such as graphing calculators and computer programs. DBR has two goals that occur simultaneously: developing successful learning environments and using these environments as "natural laboratories" to study teaching and learning, rather than studying them in a research laboratory (Sandoval & Bell, 2004). Collins et al. (2004) said that laboratory studies are "effective for identifying the effects of particular variables, but they often neglect variables critical to the success of any intervention" (p. 20), such as the variables that are unique to each classroom like the norms, resources, and types of students. It is important that the research team recognizes the characteristics and culture of the classroom setting. Cobb et al. (2003) mentioned that DBR is one of the few research methodologies that addresses the complexity of educational settings. This step outside of the laboratory and into the classroom helps researchers to do research that directly helps teachers and students.

Focuses on Designing and Testing an Intervention. Design-based research focuses on the design and testing of an intervention. The type of interventions can vary, but in the investigation by Anderson and Shattuck (2012), 68% of the interventions involved the use of mobile and online technologies. In addition to technological interventions, other common types of educational interventions may include learning activities, types of assessments, and changes in policies or administrative activity. A key part of DBR focuses on the design of these interventions. Kelly et al. (2014) stressed the importance of careful documentation when

designing the interventions to be used as evidence to help inform redesign of the next phase, if necessary. These interventions should be designed for examining learning in naturalistic contexts and DBR should help researchers develop a methodological toolkit for observing evidence-based claims from examining these natural contexts (Barab & Squire, 2004). Sandoval (2004) argued that the design and study of a particular intervention can help researchers understand whether an intervention works and more importantly, how that intervention works, and why that intervention works. Collins et al. (2004) also mentioned that it is important for the research team to focus on why interventions work and how the setting plays a role in the study. By both designing and studying the intervention and how it affects students, this also helps give a clear picture of the research issues or problems that need to be addressed in the study. The design and testing of these interventions is a key feature of the quality, effectiveness, and results of DBR.

Iterative. Design-based research is an iterative methodology. In other words, the interventions go through different phases of design in an attempt to make them the best they can possibly be. The knowledge generated during each phase of the DBR process is used to refine the design and implementation of the intervention in the next phase, which is why DBR is not only considered iterative, but adaptive (Anderson & Shattuck, 2012; McKenney & Reeves, 2014). This makes DBR different from other types of educational research which typically involve a single cycle of data collection and analysis focused on producing knowledge (Shah et al., 2015). The iterative component of DBR is important since design-based researchers recognize that hardly anything in life is done perfectly the first time. The same can be said about research and Anderson and Shattuck (2012) often joked about how the iterative adjustments and improvements to interventions in DBR can be referred to as "research through mistakes." Anderson and Shattuck (2012) described how the design practice of anything, even situations

outside of the research work (making cars, pharmaceutical drugs, etc.) rarely involves creating a perfect product on the first try. Instead, creating an optimal product involves testing it and making improvements at least a few times. By updating the interventions used in DBR, this creates a more authentic comparison to how things work in the real world and the natural environment. Iteration and making improvements to interventions based on feedback from the people using them encourages researchers to recognize the importance of improving their mistakes rather than pretending they do not exist. Hence, iteration is a key component to making DBR successful.

Collaborative. Design-based research is a collaborative methodology. When engaging in DBR, it is important that researchers and practitioners work together to have the most impact. Many times, teachers are too busy to do research and are not trained as researchers (Anderson & Shattuck, 2012). Similarly, some researchers are disconnected from the classroom and not necessarily trained as teachers. This collaboration highlights the strengths of both partners and looks at the best of both worlds, allowing them to work together to meet the same goal of helping students learn and teachers teach effectively. There is a growing emphasis on the idea that teachers should be important members of the design team in DBR and as professional contributors rather than subjects in a study who have to follow a script (Linn et al. 1999; Penuel et al. 2007). This is because teachers' experience in the classroom can provide valuable insight that the researchers may not otherwise have without them. In addition to the collaboration between teachers and researchers, oftentimes in DBR, participants are not "subjects" assigned to treatments, but are instead treated as co-participants in both the design and sometimes even the analysis (Barab & Squire, 2004). For example, many times the data that the researchers get from the students influences the design of the next iteration in DBR studies. Like the researchers' and

the teachers' opinions, the students' opinions are valued and taken into consideration throughout the process, making collaboration a key feature of DBR.

Flexible. Design-based research is flexible and adaptable. Because DBR is an iterative methodology that is based on making changes to improve the intervention, the researchers and teachers should be flexible in terms of design. In other words, the design plan should be flexible to accommodate the inevitable refinements necessary in the design processes (Wang & Hannafin, 2005). This flexibility allows the research team to adapt or even completely change the plan if something does not work or does not seem to be going well. It is important for DBR studies to be flexible because in classrooms, teachers and students are expected to be flexible and adapt their plans according to their needs. The collaborative and flexible nature of DBR makes it a unique methodology that can have an actual impact on teaching and learning sooner than other methodologies. If something in the study needs to be changed, the research team can do that at any time without having to wait until the original plan of the study is finished. Flexibility is a key component that helps make DBR so impactful in the classroom since classrooms are generally not stable and are subject to a lot of uncertainty.

Practical Impact on the Practice of Teaching. Because of all of the previously described characteristics, DBR is heavily focused on making real changes in classrooms to help students learn. In other words, design-based research has a practical impact on the practice of teaching. Anderson and Shattuck (2012) described how in their own keynote talks, many times they would try to challenge participants to think of one research outcome that has made a difference in a real education setting. They shared how it is both surprising and depressing that many educators cannot think of a single research outcome (Anderson & Shattuck, 2012). This is because, in general, research done in a laboratory or done as a highly controlled experiment is

difficult to replicate in an actual classroom where conditions are constantly changing. Educational settings, such as classrooms, are incredibly complex and it is difficult to predict what will happen within them. Designers of DBR studies need to try their best to account for the influence of social factors and dynamics that affect both participants, such as students and teachers, and the processes, such as school culture, physical characteristics of classrooms, etc. (Wang & Hannafin, 2005). DBR helps researchers and teachers better understand existing situations in classrooms and if necessary, change them into better situations, hence having a practical impact on the practice of teaching.

Other Common Characteristics. There are a few other points that were brought up multiple times in the DBR literature. The first is that many scholars argue that DBR studies should be classified as mixed methods studies, containing both quantitative and qualitative components. However, although this is common, it is not a requirement, and instead the focus lies on whether or not the methods are justified and whether or not the research team deems them best for answering the research questions. The second point is that many times, DBR is compared to other types of research, specifically Action Research. Anderson and Shattuck (2012) mentioned that both practitioners and researchers often have trouble differentiating between Action Research and DBR because they share many epistemological, ontological, and methodological assumptions.

Both Action Research and DBR are iterative and focus on identifying a problem, assessment, and analysis in an applied educational setting, along with the implementation and evaluation of some type of change or intervention to address a problem (Ford et al., 2017). One key difference between the two is that in Action Research, the researcher and teacher are one in the same, whereas in DBR, the teacher and the researchers work together and complement each

other's strengths. According to Shah et al. (2015), another key difference is that Action Research emphasizes knowledge generation about what works or how to improve what is working immediately with minimal or no concern for why it works, whereas DBR does look at why interventions help students. However, despite these differences, Action Research and DBR share many positive qualities and are great steps towards improving the practice of teaching and improving the quality of learning for students. I chose DBR for this study because I specifically wanted to see a cycle of improvement by looking at the labs in two back-to-back semesters. I also wanted to use student feedback when making changes to the labs and DBR allowed me to do this.

Assumptions of DBR

Along with the characteristics of DBR that have been described, there are also some assumptions to be made when using DBR as the methodology in a research study. One example of an assumption of the approach is that the outcome of the study is related to the intervention. In other words, researchers should take great care to acknowledge and control for confounding variables that could potentially mix up their results during their analysis. Another example of an assumption of DBR is that students are honest in their feedback and are metacognitively aware enough to accurately give feedback to researchers as to how they learn. Similarly, it is also assumed that the teachers and researchers are honest in their feedback as well and do not allow their biases to influence their decisions. Lastly, a third assumption is that the design team will know what to do with the feedback they get from the participants and that they will make the best changes for supporting student learning. Different DBR studies will have different assumptions and it is important for the research team to be clear about what they are assuming.

There are many benefits to using DBR that can be seen by looking at some exemplary studies that have been done using DBR.

Benefits of DBR in Education Research

This first example describes how collaboration and flexibility are both benefits of using DBR. Kali (2016), a learning scientist, described a framework called Design Researchers' Transformative Learning (DLTR) that can be used in conjunction with DBR. Kali (2016) claimed that boundary crossing within teacher-researcher partnerships was an important feature of the DRTL framework. In other words, the collaboration between teachers and researchers was a key feature of this framework. Kali (2016) described a DBR study that was done in a large undergraduate biology course that used the DRTL framework and allowed the research team to learn through this process. The motivation for this study came from the instructors of the course who felt like they had more to give to their students than just a traditional class that involved a lot of lecturing and not a lot of student engagement. Because DBR is collaborative, the researchers and instructors were able to work together and use their expertise to design a study in which they hoped to improve student learning to be more meaningful.

To begin, a design team was formed that included the instructors, two science education researchers, and two design researchers (Kali, 2016). This team created a design that involved gradually changing the course throughout three years in three separate stages. The first stage of the study did not involve any change in how the course was taught as a large lecture, however the students were given access to a website that included "interactive visualizations, video recordings of the course lectures, self-feedback questions, and links to relevant sections of an online version of the course textbook" (Kali, 2016, p. 8). The second stage was the same as the first except the use of the website was mandatory and the third stage actually eliminated the

lectures and replaced them with the course website and mini conferences that occurred each week to discuss the topics in more detail. The goal of the study was to determine how students and instructors describe learning in the new teaching model of the biology course (Tal & Tsaushu, 2018). They collected most of their data through interviews and found evidence of deep and meaningful learning through the website.

Because the research team used DBR, they were able to understand how the students learned as well as what motivated them and then used this information to inform the changes they made to the course. Kali (2016) described how DBR helped them confirm or refute their initial assumptions. I chose to highlight Kali's study because it is one of the few DBR studies that specifically focused on undergraduate college students. In this study, the research team had a hypothesis that they would be able to find relationships between students using the course website, their attitudes towards biology, and their understanding of the content (Kali, 2016). However, within each stage, they did not find any interesting relationships between students' use of the website and their learning outcomes, although they did find this relationship between stages (Kali, 2016). So, it was clear that the third stage did promote deeper learning, however it was difficult to determine why this was the case.

The research team decided to investigate this further. Since the students chose whether they wanted to take the class with the new design or not, the researchers were able to determine why. Originally, the design team anticipated that students took the class because they thought they would get a higher grade in the new design. However, they determined through the interviews that most of the students signed up because they were very interested in the subject, excited to try something new, and wanted to make sure they had a deep understanding of the content. This was another variable that allowed them to further understand why and how the

intervention worked in promoting meaningful learning. However, this result was unexpected for the researchers because it went against their initial assumption. Despite this result, it did help them better understand what students valued in their learning process, and it allowed the research team to have a better grasp on how motivation affects student learning outcomes. They used this information to help the instructors understand the students, potentially make other changes to the course, and to analyze some of the data through a new lens which took this new variable into account. The research team had to modify their original plan in this study and because DBR is flexible and collaborative, this was possible.

A study done by Zydney et al. (2020) also shows how the flexibility of DBR can be very beneficial. This exploratory study was done to design, implement, and assess a blended synchronous learning environment that involved protocols in a graduate education course. According to the study, protocol pedagogy is a "student-centered method of teaching that uses structured discussions with the intention of fostering meaningful interactions" (Zydney et al., 2020, p. 2). This study had three phases and the goal was to examine the influence of protocols on the students' and instructors' experiences in this environment through qualitative data analysis (surveys, observations, and interviews). These data sources were designed to assess the learner experience and understand how to adapt the classroom environment to better meet student and instructor needs. Through each iterative phase, the research team would come together and discuss the decisions made to design the learning environment. Through each iteration, the design decisions were tested to improve the underlying design propositions. The research team came together multiple times throughout the analysis and after each phase, data was collected from students and the design was improved based on that feedback (Zydney et al., 2020). It is important to understand how and why some of these changes were implemented.

Ultimately, the research team changed two main aspects of the protocol: timing and structure of the activities. These changes were based on student feedback. The students felt rushed by time constraints they had for doing activities, so they extended their time and gave time ranges instead of time limits for the activities to offer more flexibility. The structure of the activities changed from focusing on different topics to focusing on different students' ideas. They reported that this allowed the conversation to flow more naturally. Student roles and the use of technology also changed throughout the phases of the study due to the feedback from the participants (Zydney et al., 2020). After the first phase, students expressed that they thought it would be better to give the student discussion facilitators more control over the discussion. The web conferencing tool also changed and allowed the students to use breakout rooms to work in groups or work independently if needed. Through iterative qualitative data analysis, this study provided recommendations on how to design and change classroom protocols and what technology could be helpful to do this. All of the changes made throughout the study were student and instructor driven and could be implemented throughout the process due to DBR being flexible, as well as iterative and collaborative.

These same benefits are exemplified in Bakker and Van Eerde's (2015) research regarding statistics education. Their work started as a part of Bakker's doctoral dissertation on DBR in statistics education and was motivated by stakeholders being dissatisfied with what students learned about statistics and how technology played a role in the teaching and learning of statistics (Bakker & Van Eerde, 2015). In an attempt to address this dissatisfaction, Bakker and Van Eerde (2015) realized that one common problem that students have when they are first learning statistics is that they tend to see individual data points in a dataset, rather than thinking of the dataset as a whole and the distribution created from samples. They developed a plan to

design educational materials involving technology to help students understand distributions. The design team consisted of researchers designing the materials, teachers giving the instruction, and pre-service teachers who acted as assistants and helped with data collection. Due to the collaborative nature of DBR, the group met weekly to assess their progress, talk about challenges, and make adaptations. In addition to DBR being collaborative, another key point that Bakker and Van Eerde (2015) stressed was that DBR is flexible. They described how their original research question actually changed because they realized that they needed to be more specific and that the concept that they set out to study was too difficult for seventh graders. The flexibility of DBR allowed them to adapt their research question to better suit their needs and they emphasized that this happens frequently with DBR.

Bakker and Van Eerde's final research question was "how can we promote coherent reasoning about distribution in relation to data, variability, and sampling in a way that is meaningful for students with little statistical background?" (Bakker & Van Eerde, 2015, p. 28) and their subjects were eighth graders instead of seventh graders. The research team designed ten lessons that all contained learning goals, activities, and assumptions about students' potential learning processes. They referred to these as hypothetical learning trajectories or HLTs. After each lesson, the HLTs were evaluated, discussed, and the research team used these discussions to inform the next lesson design. For each lesson, the researchers collected data that included student work, field notes, and the audio and video recordings of class activities, including mini-interviews (Bakker & Van Eerde, 2015). To analyze this data, they divided the lesson into episodes and for each episode the researchers watched the videos, read the transcripts, and came up with "conjectures" about student learning to use as codes for the analysis.

These conjectures that they generated were tested against other episodes and the rest of the collected data in the next round of analysis. If the conjectures were confirmed in the next round, they stayed on the list of codes and if they were refuted, they were removed. This process of generating and testing was repeated for each lesson (Bakker & Van Eerde, 2015). These codes allowed them to better understand student learning and how it changed as more lessons were completed. Through this iterative design and analysis, the researchers were able to see what activities promoted logical reasoning about distributions for these students. Bakker and Van Eerde have shown that because DBR is flexible, iterative, and collaborative, it can be very useful in studying statistics education.

As discussed, many of the key features of DBR can be potential benefits to researchers, teachers, and students. The iterative design takes into account that researchers and teachers are not perfect and allows the design team to learn from their mistakes and improve without having to conclude one study and launch another one. The collaborative nature of DBR allows for multiple perspectives from a diverse group of people with different specialties who all have the same goal of improving education. The flexibility and adaptability of DBR allows the research team to make changes to the research design if they feel it is beneficial. Because this type of research is situated mainly in classrooms and in real educational contexts, a huge strength is that it can have a practical impact on the practice of teaching. Despite the benefits of DBR, there are also some challenges and critiques to acknowledge.

Critiques to DBR

One of the challenges to DBR pertains to it being an iterative research methodology. We know that designs are rarely implemented perfectly and that there is always room for improvement. The challenge becomes how to know when to stop. In other words, it is difficult to

know when, if ever, the research is truly completed (Anderson & Shattuck, 2012). Additionally, the iterative aspect makes these studies generally longer than other types of research. Anderson and Shattuck (2012) mentioned that DBR projects are susceptible to the research team running out of resources, funding, or time. However, there are ways to mitigate this. For example, Herrington et al. (2007) addressed this concern openly by describing a four-year plan for DBR to be used in a doctoral dissertation. Another option for a doctoral student would be to complete some phases during their graduate program and plan to complete others after they graduate. Because of the length of time it takes to complete multiple iterations, another challenge is the complexity of the DBR process. Collins et al. (2004) noted that DBR studies produce a lot of data, which can be difficult to manage, work with, and determine what is useful. However, this rich data is often very meaningful and can help researchers and teachers make the intervention the best it can possibly be.

Also, many of the examples of DBR talk about the idea of researcher bias. Because the researchers and teachers are both heavily involved in the design and implementation of DBR studies, there is some concern about bias, reliability, and validity. It is difficult to guarantee that the researchers are reliable and can make credible and trustworthy claims (Anderson & Shattuck, 2012; Barab & Squire, 2014). Hence, a certain kind of wisdom is needed to walk this narrow line between objectivity and bias (Anderson & Shattuck, 2012). Unlike most researchers, design-based researchers are not solely observing interactions in the classroom but instead many times they are actually "causing" these interactions that they are making claims about, which could seem like a conflict of interest (Barab & Squire, 2004). It seems that one way to help with reliability is to make sure the team is aware of these issues and understands the importance of minimizing all types of bias. Barab and Squire (2004) specifically mentioned the importance of

validity and discussed how DBR embraces Messick's (1992) idea of consequential validity. Messick (1992) argued that the validity of a claim is based on the changes it produces in a system and that these changes can be considered evidence of validity. Barab and Squire (2004) specifically mentioned that design-based researchers should be clearer about the kinds of claims they make from DBR and also be very open about the limitations of their findings.

One important critique that Bergold and Thomas (2012) mentioned was the notion that marginalized communities are less likely to participate in research, specifically DBR where the researchers and teachers collaborate so closely. They discussed how DBR is a type of participatory research that involves both inquiry and action and is mainly led by the participants, such as teachers and students. The participants are usually community members, and they have control over what is done and how it is done in the study. Therefore, if these participants do not include members of marginalized communities, the study is designed and implemented with these biases. Bergold and Thomas (2012) described the dilemma that marginalized communities are "in a very poor position to participate in participatory research projects, or to initiate such a project themselves" (p. 197). So oftentimes their voices are not heard. They also often have limited resources available, making it even more difficult to participate in research. It is important that DBR studies seek out a variety of people with whom to collaborate. Bergold and Thomas (2012) also mentioned how despite this criticism, the goal of many participatory research studies is to make the voices of these marginalized communities heard, support them, and foster empowerment by involving them in these research studies.

Another important point is that it can be challenging to clearly compare the multiple iterations of DBR. Unlike purely quantitative studies, most DBR studies do not produce "measurable effect sizes that demonstrate what works" (Anderson & Shattuck, 2012, p. 8).

Although oftentimes there are not hard numbers to make an explicit comparison, DBR studies do provide very rich qualitative data such as thorough descriptions of the contexts of the studies, a detailed description of how the intervention was created, the implementation journey of the intervention including the challenges, and a descriptive notation on the design principles that emerged from the study. Many times, they also provide important quantitative data as well and through both of these types of data, the research team is able to understand how interventions are created and why interventions may or may not be successful in that particular context.

Lastly, another challenge is concerned with the issue of replicability in DBR. Replicability in research is oftentimes a central idea and because DBR takes place in natural contexts, usually classrooms which are incredibly dynamic and unique, these natural contexts are often very difficult to replicate or reproduce. Barab and Squire (2004) emphasized that it is impossible and oftentimes undesirable to manipulate cultural contexts, however this does make it difficult for design-based researchers to replicate others' findings. Although these studies are difficult to replicate, they can still provide insight to other researchers on how to design and test an intervention and what impact it has on teaching and learning. While the environment may not be the same, it is more realistic and applicable to teachers and students in the real world (where the environment is never the same). However, the basis of the theories and interventions can be replicated and then adapted to each particular context. To be clear, DBR involves so much more than simply describing the design and the conditions of the research. Again, this distinction reiterates the idea that DBR can help researchers develop theories about why interventions work rather than only figure out that an intervention works.

Despite these challenges and critiques, the benefits outweigh them, particularly when the goal of the study is to focus on both the outcome and the why of the outcome. In this study, the

goal is to make changes in an actual classroom and understand the reasons for those changes. It has been shown that DBR may be a challenging way to approach research, but according to Bakker and Van Eerde (2015), DBR is also very rewarding considering the products such as interventions or theories that come out of it and the insights that can be gained by the research team. In this study, the curricular innovation of the labs is one of such potential products. Some may argue that the interventions developed in DBR studies could be described as very small changes to very specific contexts, however, Anderson and Shattuck (2012) reminded their readers that even small changes can make a big difference in students' learning and their experiences. Hence, DBR does make a difference, especially at the level of individual teachers, students, and schools.

DBR and Data Science Classrooms

In Anderson and Shattuck's 2012 DBR literature exploration, one feature that they looked at was in which subjects or programs the DBR studies were taking place. Science was the most common discipline with 51% of the studies reviewed falling into this category (Anderson & Shattuck, 2012). Mathematics and computers were identified, but with only 9% and 7% of the studies, respectively, falling into these categories (Anderson & Shattuck, 2012). The discipline of statistics was not included for the studies they reviewed. After reviewing the literature on DBR, I have found that there are few, if any, studies besides the studies from Bakker's doctoral dissertation, involving statistics education or data science education and DBR, making this a gap in the literature that should be further explored. Using DBR to study data science labs may yield an understanding of classroom complexities that we lack, especially in this novel field. There is not enough research done on what actually happens in a lab, especially data science labs. This

study intends to contribute to both curriculum development and seeing the implementation in the context of the two labs.

DBR and Mixed Methods Analysis

Because I am specifically focusing on how to design the labs to help students engaged in CA and social justice awareness, as well as how students engaged in CA and social justice awareness, I needed both DBR and a mixed methods approach for designing the study and analyzing the data. I used DBR as the methodology because it is flexible and iterative. I wanted to design an intervention that I could test and use student feedback to make improvements. The first research question looks at how this was done by describing the design of the labs, how the students interact with the labs, and the adaptations. In order to understand the data that I collected in this study I used a mixed methods approach to help me analyze the data using multiple techniques to get a full picture of how the students engaged in CA and social justice awareness.

I answered the second research question by using mixed methods. Tashakkori and Creswell (2007) defined mixed methods research as "research in which the investigator collects and analyzes data, integrates the findings, and draws inferences using both qualitative and quantitative approaches or methods in a single study or a program of inquiry" (p. 4). Through the combination of the interviews, surveys, audio-recordings, and the scaffolds in the labs, I got both qualitative and quantitative data. A key factor in mixed methods research is that the methods are integrated to help answer the research questions, rather than using two types of data without integrating the findings (Tashakkori & Creswell, 2007). Both the quantitative and qualitative data in this study help me better understand how the students engaged in CA and social justice

awareness. All of the data sources are integrated to answer that overall question. In the next section, I elaborate on the data sources and data analysis techniques that were used in this study.

Data Sources and Analysis Methods

Throughout the duration of this study, four main data sources were collected to answer the research questions: surveys, interviews, data from the labs that included written answers to individual reflection questions, and audio recordings of group discussion questions. Because this is a DBR study with two iterations, the data were collected once in Fall 2021 and then again in the Spring 2022. The data from Fall 2021 informed the changes made in the labs for Spring 2022. The first research question is about the DBR process, specifically, the study design and the adaptations of the labs. The second research question involves looking at how the students engaged in CA as well as social justice awareness through analyzing the data collected throughout this study. Table 1 shows each of the research questions, the sub-questions, and how those map to the data sources and methods. Next, I unpack this table and describe the data sources and how the data from those sources were analyzed.

Table 1

Research Questions, Sub-Questions, Data and Methods

Research Question	**Research Sub Questions**	**Data Sources**	**Methods**
1. How can design-based research be used to create labs that use principles of distributed cognition in the context of a data science course?	1a. What scaffolds and artifacts are included in the labs? 1b. What principles of distributed cognition are used in the design of these scaffolds and artifacts and why? 1c. What adaptations to the labs (scaffolds and artifacts) result from the DBR process and why?	• All versions of the labs • All versions of the labs • Lab Submissions • Student Interviews	• During the design process, I examined what the students did in the first iterations of the labs and looked at their interview data for feedback.
2. What evidence do students show of engaging in computational action during and after these labs?	2a. What evidence is there that the scaffolds and artifacts help the students engage in CA? 2b. How do students apply data science practices to question the status quo and consider social justice issues? 2c. What do students perceive that they are learning through these labs that will be useful in the real world?	• Audio Recordings of Group Discussions • Lab Submissions for Individual Reflection questions • Student Interviews • Student Interviews	• **Data Reduction Rubric** for evaluating the interviews to decide which groups to analyze • **Toulmin Argumentation Pattern** for Group Discussion Questions • **Thematic Analysis** for Individual Reflection Questions • **Thematic Analysis** for Interviews

Data Sources

Surveys. The students completed surveys before and after completing each lab (see Appendix D). The pre-lab surveys contained Likert-scale questions to get a baseline for how much prior knowledge the students have about the topics covered in the labs. The same Likert-scale questions were also asked on the post-lab surveys. In the pre-lab survey, there were also questions that let me know how the students have been communicating in prior labs since one of the goals of this study is for students to have more conversations. The post-lab survey asked multiple short answer questions about communication to determine how they communicated and what they liked and disliked about the communication portions of the lab. Lastly, in the pre-lab survey there was an open-ended question asking the students what they hoped to learn from completing this lab and in the post-lab survey there were short answer questions asking about their takeaways from the labs and if they had any remaining questions after completing them. Table 2 shows the question topics and question types for both surveys.

Table 2

Types of Questions in the Surveys

Pre-Lab Survey		Post-Lab Survey	
Question Topics	Types	Question Topics	Types
Background	Multiple Choice	Knowledge Gained	Likert-scale
Prior Knowledge	Likert-scale	Communication	Likert-scale, short answer
Prior Communication	Multiple Choice	Lab Take-Away	Short answer
What do they hope to learn?	Short Answer	Likes and Dislikes	Short answer

After analyzing all of the data, I realized that I got a clear picture of what went on in the labs and how students felt about them through the interviews and audio-recordings. Because the surveys gave me the same information that I got from the other data sources, I focused on the other data sources.

Labs. Throughout each semester, the students completed two labs (Lab A and Lab B) that were used as a part of the data collection in this study (see Appendix A and Appendix F). The labs contain three different types of questions for students to answer. The first type are coding questions in which the students are instructed to use Python to complete a specific task. The students coded using Jupyter notebooks as their code editor. Jupyter notebooks are a tool that started in 2013 and have been widely adopted by many different communities and are especially popular with data science. Jupyter is a free, open-source, interactive web tool known as a computational notebook, which researchers can use to combine software code, computational output, explanatory text, and multimedia resources in a single document (Perkel, 2018, p. 146). They help users communicate while coding because they combine code, text, and execution results with visualizations (Pimentel et al., 2019). Jupyter notebooks were ultimately designed to help make data analysis easier to share, document and reproduce (Pimentel et al., 2019). In this study, they allow me to see students' code, the output it produces, and allow the students to type written text all in the same document.

Throughout the labs there are also individual reflection (IR) and group discussion (GD) questions. The students' answers to the IR questions can be embedded into the same lab using Jupyter notebooks. The individual reflection questions were questions that encouraged the students think about on their own and write about individually. To study what type of communication occurred during the labs and to analyze the GD questions, I audiotaped the

groups of students who worked together and gave consent to participate in Part 1 of the study. I transcribed the audio tapes and then they were used for analysis. During the transcription process, I put explanatory text in brackets. I primarily focused on analyzing the responses to the GD and IR questions when answering the second research question. While the coding was important, almost all of the students were able to get all of the coding questions correct.

Interviews. The last type of data that I collected in this study was from individual interviews with students at the end of each semester (see Appendix E). Throughout these interviews, my goal was to discover how the communication between students went during the labs, what they liked and disliked about the labs, what they learned from the labs, and how they thought these labs would be useful in their lives and in their jobs. The goal of the interview data was to help see if the students engaged in CA and social justice awareness. Like the audio recordings from the labs, the interviews were also audio recorded and transcribed for analysis. The interviews allowed the students to elaborate on their thoughts from the pre-lab and post-lab surveys. In the next section, I discuss the analysis methods that I used to analyze the data from the labs and interviews.

Analysis Methods

Toulmin for GD Questions

Background and Data Reduction. Data science is a collaborative discipline that involves multi-modal communication and working with others to solve problems. During the labs, the students answered coding questions, written individual reflection questions, and group discussion questions. The labs were designed to have the students engage in computational action (CA) by engaging with exercises in lab that mimic working as a

data scientist in the real world. The goal was for students to see the connection between what they were doing in the labs and the world outside of the classroom. To help facilitate this, both labs were centered around social justice issues to help students engage in data science practices to engage in social justice awareness. One way that I saw the students engage in CA and social justice awareness was through the group discussions that they had during the labs. One of the most important types of data that I gathered was the audio recordings from the labs. This allowed me to hear what went on in the students' group discussions. These audio recordings were collected and then transcribed for analysis.

Audio recordings were chosen as a data source because they allowed me to know exactly what happened during the labs. Surveys and interviews gave student accounts of what happened in the labs, rather than evidence of what actually happened. These are helpful, but it is also important to know what students said during the group discussions. The labs asked students group discussion questions that were designed to help them engage in CA and social justice awareness. The students discussed their answers to these questions in their groups and the audio data allowed me to determine whether the goals were met. Before starting this analysis, I needed a data reduction technique to help me focus the analysis. To decide what groups to analyze, I looked at the individual students in each group. For each student, I assessed them based on their interviews. To do this, I created a rubric that looked at three topics: (1) Quality of answers to interview questions, (2) computational action and (3) social justice. This rubric is included in Table 3.

Table 3

Rubric for Data Reduction

Topic	Scores
Quality of Answers to Interview Questions	1. Did not answer most of the questions. 2. Gave one-word answers for most of the questions 3. Answered the questions but did not think too much about most of them. 4. Gave thorough answers to most questions. 5. Went above and beyond- gave great answers and elaborated on most of them.
Computational Action (CA)	1. Did not see the connection between the labs and the real world/their lives 2. Mentioned the connection briefly 3. Saw the connection and mentioned it multiple times 4. Saw the connection, mentioned it frequently, and discussed it in detail 5. Indirectly discussed evidence of engaging in both parts of CA: computational identity and digital empowerment.
Social Justice Awareness	1. No mention of anything related to social justice 2. Mentioned briefly 3. Mentioned frequently 4. Mentioned frequently and described how they want to learn more 5. Mentioned frequently and talked about taking action (thinking about mobilizing for change)

Each student was given a score for each of the three categories ranging from 1 (not doing well with the topic) to 5 (going above and beyond for the topic). Since there were three topics, each with a possibility of five points, the students were given a total score out of 15 points. I used the following labels to categorize their scores, shown in Table 4.

Table 4

Labels for Rubric Scores

Score	Label
12-15	Excellent
9-11	Good
7-8	Satisfactory
6 or Below	Needs Improvement

I gave each student a score and had a second coder who has experience doing statistics education research do the same for four random students. I met with the second coder and explained the rubric and answered any questions that they had. We agreed on three out of four scores for the random students, meaning our reliability was 75%. After obtaining reliability, I chose three groups from Fall 2021 (F21) and three groups from Spring 2022 (S22) to analyze. For the F21 groups, I looked at their discussions from Lab B. Lab A was too long in F21 so no one got to the group discussion questions, meaning there was no audio data regarding GD questions to analyze. For the S22 groups, I looked at the discussions from both Lab A and Lab B.

The groups that I chose to analyze had a mixture of students who scored "Excellent," "Good," and "Satisfactory" on the rubric. All of the groups that I analyzed had at least one student who scored "Excellent" or "Good" on the rubric and a few groups had one student who scored "Satisfactory" in addition to the students who scored "Excellent" or "Good." I chose a mixture of students because I wanted to analyze groups that had students who seemed to engage in CA and social justice awareness, but also did not want to pick the outliers (groups with students who demonstrated excellent performance). By including students who got a variety of scores on the rubrics, I was able to identify groups to analyze intentionally. The groups that I picked were balanced yet had substantial discussions. In other words, I did not want to pick only groups with only students who scored "Excellent" or groups that did not have much discussion. It is important to note that not everyone who participated in the discussions was interviewed.

Toulmin Argumentation Pattern. I analyzed the transcripts of the group discussions using the Toulmin Argumentation Pattern (TAP). This is an analysis method that identifies the parts of an argument that people make. The TAP shows the structure of an argument in terms of an interconnected set of a claim, grounds (or data), warrant, qualifier, rebuttal, and backing

(Toulmin, 1958). I used the TAP because I wanted to highlight the arguments that the students made collectively as a group to show that they were engaging in CA and social justice awareness. Because the students were working together in small groups, the unit of analysis was each group. The students produced collective arguments during their group discussions. In collective argumentation, a group of students provide the components of an argument (Forman et al., 1998; Moore-Russo et al., 2011; Yackel, 2002). During the group discussions, either one or multiple students would make a claim and the other students would add the other parts of the argument.

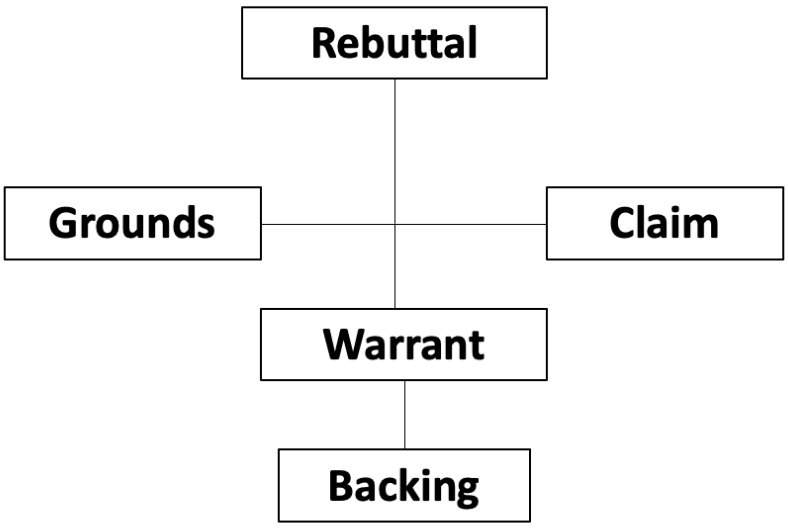
Collective argumentation works well with the framework of DC because in *Cognition in the Wild*, Hutchins (1995) pointed out that the group is more important than individuals. All of the individuals in the system depend on each other and the social organization makes the point of using the other group members and artifacts to have distributed cognition (Hutchins, 1995, p. 178). This was seen in the group discussions because the knowledge in the labs was distributed among the students, their group members, the artifacts that they were given, and the artifacts that they created. The students in each group also created collective arguments together by answering the questions provided in the scaffolds.

I did the analysis of the group discussion questions in two parts. The first part involved identifying the main arguments that the groups made by looking at the different parts of the arguments mentioned above. First, I identified the claims that the groups made. Many of these were claims about data science and social justice issues. I also looked at what grounds (sometimes referred to as data) they used to support their claims. I identified the warrants that linked the grounds to the claim. These warrants explained why the grounds are relevant. I also looked for backing or additional support for the warrant. There were some cases where there was

no evidence of backing and some cases where the backing was implied and not explicitly stated by the students. When the backing was not explicitly stated, I put it in a cloud in the Toulmin Diagram to show that it was implied. This is consistent with what others have done to show this analysis in the Toulmin Diagram (e.g., Hollebrands et al., 2010; Moore-Russo et al., 2011). I took note of the qualifiers that students used, which showed the strength of the claims. There were a few qualifiers, but these did not help me answer the research questions, so they were not included in the results, despite that they were identified in the analysis. Lastly, I identified any rebuttals, or acknowledgement of other viewpoints where the claim may not be true, that came up in their arguments. There were very few rebuttals, but I did include the ones that came up. A diagram of the parts of the Toulmin Pattern for Argumentation that I used in my analysis is shown in Figure 6.

Figure 6

Toulmin's Argument Pattern



The second part of the analysis involved evaluating each group based on how they did with each of the goals. The purpose of this analysis was to answer the second research question which looks at how students engaged in CA and social justice awareness. To evaluate the groups, I created a short rubric that allowed me to identify whether they met the two goals based on their group discussions. The rubric looked at whether they needed improvement, were sufficient, great, or went above and beyond with CA and social justice awareness. This rubric is shown in Table 5.

Table 5

Rubric for Goals

Computational Action (CA)		Social Justice (SJ)	
Category	Description	Category	Description
Needs Improvement	Did not have discussions about issues, just shared their answers	Needs Improvement	No mention of anything related to social justice
Sufficient	Had brief discussions about the questions and coding they did	Sufficient	Mentioned social justice issues briefly
Great	Had deep back and forth discussions, used their analysis in their discussion	Great	Had deep back and forth discussions about social justice
Above and Beyond	Same as Great, but discussions go beyond the issues in the labs	Above and Beyond	Same as Great, but discussions about social justice go beyond issues in the labs (outside examples, thinking about mobilizing for change)

I had the same second coder from that I discussed previously look at the group discussions of four groups, two from the fall and two from the spring, and use the rubric to decide how they did

with CA and social justice awareness. We agreed on four out of the four groups that the second coder analyzed (100% agreement). After that, I analyzed the remaining groups.

I chose the Toulmin Argumentation Pattern as my analysis method for the group discussion questions because it allowed me to determine whether the students were using the artifacts and scaffolds, as well as whether they engaged in CA and social justice awareness. By identifying their arguments, I was able to see connections between what they talked about and the scaffolds and artifacts, showing that they used them. I expected to see the scaffolds and artifacts in the grounds and the warrants of the arguments. I was able to see what type of arguments they made about social justice (engaging in social justice awareness), as well as whether or not they discussed social justice issues. I also was able to see what types of connections they made between the data science they were doing in the labs and real-world issues (engaging in CA). More specifically, I was able to see how they made arguments collectively as a group, which helped them engage in computational identity and how they used real world data and things they created in the labs as justifications for their claims. This helped them engage in CA, specifically digital empowerment. Using the TAP as an analysis method allowed me to use the data that I collected to answer the second research question.

Thematic Analysis for IR Questions

Background. Along with the audio recordings from the group discussion questions, I also collected data from the students' lab submissions. This included both their code as well as their answers to written individual reflection questions. These data allowed me to see how students individually answered questions and their thoughts on the analysis they did. The individual reflection questions showed me what the students took away from the analysis and their immediate thoughts on the implications of their analysis. The audio recordings showed how

the students talked about data science with their group members and the interviews showed how they felt about the lab after they completed it. These individual reflection questions added another layer of understanding of how the students understood and reflected on the scenarios and analysis from the labs. Like the group discussion questions, the individual reflection questions in the labs were designed to help the students engage in CA and social justice awareness. The students typed their individual answers, and these data allowed me to determine whether the goals were met. For individual reflection questions and the interviews, I used Braun and Clarke's Thematic Analysis to analyze the data.

Thematic Analysis. Thematic Analysis is a widely used method for identifying, analyzing, and reporting patterns, or themes within qualitative data (Braun & Clarke, 2006). Although it has been around since the 1970s and is widely used, until recently, there was no clear agreement about what Thematic Analysis is and how it should be used (Braun & Clarke, 2006). Nowadays, it is considered a good method for research where you are trying to find out something about people's views, experiences, opinions, or knowledge from a set of qualitative data. In other words, if you want to find patterns in your data, Thematic Analysis is a great method for qualitative data analysis. Oftentimes this is good for data like interview transcripts, survey responses, or written replies. Thematic Analysis gives the researcher a lot of flexibility in interpreting the data and is a method that works well for analyzing large data sets because it sorts the data into broad themes. In 2006, Braun and Clarke gave a step-by-step guide to doing Thematic Analysis using six phases, or steps. The steps included 1) familiarizing yourself with your data, 2) generating initial codes, 3) searching for themes, 4) reviewing themes, 5) defining and naming the themes, and 6) producing the report.

The first step (familiarizing yourself with your data) helps researchers get to know their data. It usually involves transcribing if the data is audio data, taking initial notes, and reading through the data or transcriptions. For the IR questions, I completed this step by reading through each student's individual reflection responses and taking notes on them. The second step (generating initial codes) involves coding the data by coming up with descriptive phrases or codes that describe what the data is saying. For the IR questions, I went through each student's responses and highlighted ideas that I thought were relevant or could be potentially interesting. I categorized these into codes (which I will describe in the Results section). The third step (searching for themes) involved looking over the codes that I created and identifying patterns to start coming up with themes. Themes are broader codes and a lot of times, the researcher will have several codes that fit into a single theme (Braun & Clarke, 2006). The fourth step (reviewing themes) is where the researcher makes sure that all themes are useful and accurately represent their data. This is where they look at their data again and make sure their themes make sense. At this point, if the researcher realizes there are problems with the themes, they should edit them. For this analysis, this was my final check before I started officially defining and naming the themes. The fifth step (defining and naming the themes) involves saying what is meant by each theme and how it helps answer the research questions. If there are words in the theme names that require a definition, this step is where they are added. The last step (producing the report) is where the researcher writes the analysis of the data. I talked about each of the themes in detail in the Results section. I talked about how often the themes up, what they mean, and gave examples of my data as evidence of the themes.

The purpose of these six phases is to identify patterns in a dataset that will help researchers answer their research questions. Braun and Clarke (2006) also described how

Thematic Analysis is "theoretically-flexible." In other words, it can be used within a wide variety of disciplines and frameworks to answer many different types of research questions. Throughout this process, I found themes in the students' written reflections that show that they are thinking about social justice issues and engaging in CA. I chose Thematic Analysis because I wanted a qualitative method that worked well for large amounts of data and would allow me to find evidence for whether or not the students were engaging in CA and social justice awareness from using the artifacts and individual reflection question scaffolds. By identifying codes and themes, I was able to see how the students made the connection between the data science that they were doing in the labs and the real world, which helped them engage in CA. I saw how the students explained the data science that they were doing in writing, which also helped them engage in CA. I was also able to see how students thought they could use data science to address issues of discrimination and racism, which helped them grapple with social justice awareness.

Reliability Coding. To establish reliability, I created a codebook with the themes I came up with. The codebook contained the codes that made up each theme, descriptions of each theme, and one example. I coded two of the nine groups and had a second coder code the same two groups. We met and compared the coding that we both did. At our first meeting, we realized that there was a missing theme. There were some parts of the transcript that we were both unsure how we should code them. After our discussion, we determined that we should add another theme, so we added that to the codebook and then recoded the two groups with the new codebook. We also coded a third group with the new codebook to be sure we were not missing anything else. At the next meeting we discussed the codes and looked at agreements and disagreements. There were 60 agreements and seven disagreements for the themes (89.5% agreement). All disagreements

were discussed until 100% agreement was reached on all coding. After obtaining reliability, I finished coding the remaining groups.

Summary of Methods

DBR is a new type of research methodology that has the potential to work well for studying problems in data science education, just as it has in the learning sciences. Willingham and Daniel (2021) stated that researchers are often "frustrated and saddened that teachers do not make greater use of research findings in their practices" (p. 33). This study gives me the power to make meaningful changes in a classroom while contributing to the literature of teaching and learning data science at the same time. Using DBR and analyzing my data using a mixed methods approach allows me to develop the labs based on the needs of the students, implement them, get feedback, and then adapt and re-test them to get more data. Overall, this study allows me to use DBR to develop labs that help students engage in CA and social justice awareness while using mixed methods to help understand how the students engaged in CA and social justice awareness. I designed the labs using scaffolds that encouraged communication and the labs were centered around topics related to social justice. The techniques used to analyze the data helped me to understand how the students felt about the labs and whether or not they were effective in helping them engage in CA and social justice awareness.

CHAPTER 5: RESULTS

In this section, I describe the results of the study according to the two research questions. Sandoval (2013) said that DBR has a dual commitment to improving teaching and understanding how students learn. The two research questions in this study explore both of these ideas. Through the first research question (RQ1), I describe how I designed the labs through the DBR process. The improvements that were made to the labs helped improve teaching data science. Through the second research (RQ2) question, I look at what happened in the labs and how the students engaged in CA and social justice awareness. In other words, I was able to understand how students learned in connection with the goals of the labs. Throughout this section, I describe how I used the data I collected throughout the study to answer each of the three sub questions for both RQ1 and RQ2.

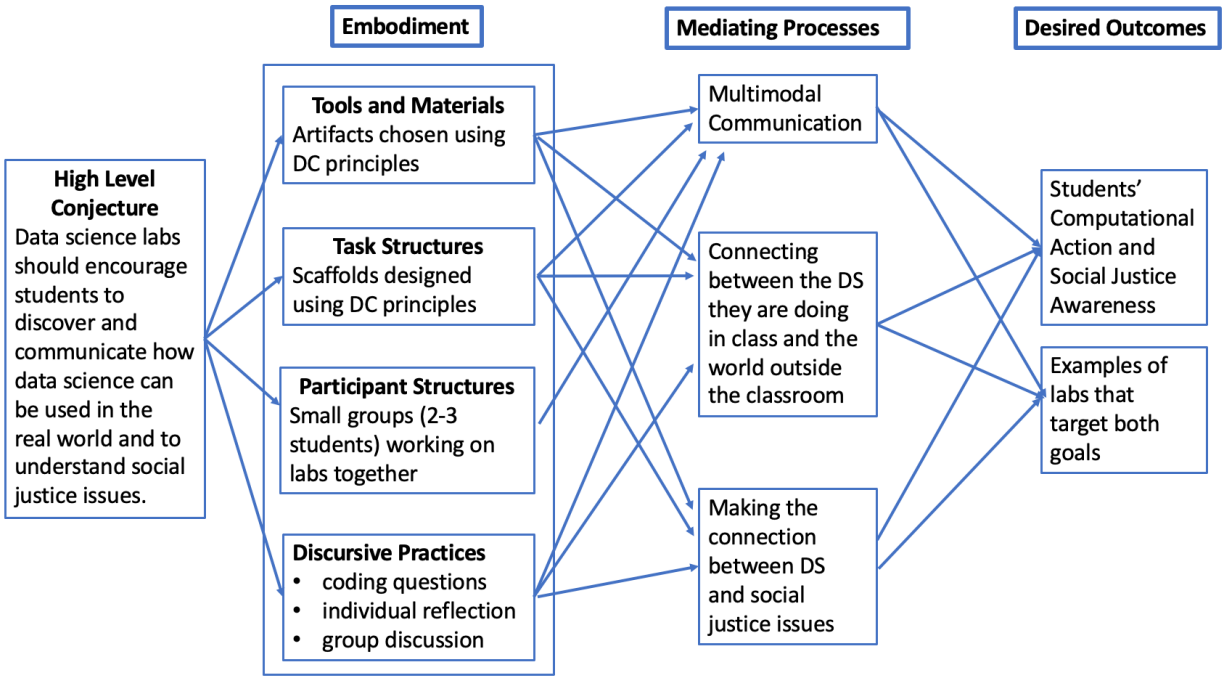
RQ1 Results

RQ1 asked: *How can design-based research be used to create labs that use principles of distributed cognition in the context of a data science course?* I answered this question by describing at the scaffolds and artifacts that are included in both of the labs, outlining which principles of distributed cognition were used to design the labs, and documenting what changes were made to the labs throughout the DBR process. The three sub questions (1a, 1b, and 1c) from RQ1 allowed me to examine each of these ideas in detail. When thinking about how to design the labs, it is important to think about the goals of the labs before designing them. I created a *conjecture map* to show how I intended the labs to lead the students to accomplishing the two goals, CA and social justice awareness. Sandoval (2013) described conjecture mapping as a technique for conceptualizing design-based research by mapping out how the ideas in a study work together to produce the desired outcomes. Conjecture maps consist of high-level

conjectures about learning in a specific context and how that learning can be supported. Those conjectures connect to the embodiment that describes the features of the learning environment. The features of the learning environment help shape the mediating processes, which explain how the students produce the desired outcomes (Sandoval, 2013). Figure 7 shows a conjecture map for this study showing the high-level conjecture, embodiment, mediating processes, and desired outcomes.

Figure 7

Conjecture Map for this Study



The high-level conjecture guided my work and summarized the intended goals of the labs. The embodiment described the features of the learning environment. In this case, I included artifacts and scaffolds that I chose and designed using DC principles. The students worked in small groups as they completed these labs by answering coding questions, individual reflection questions, and group discussion questions. The mediating processes are the hypothesized actions,

such as multimodal communication and making connections between data science, the students' lives, and social justice issues, that lead to the desired outcomes. The desired outcomes are the students' development of CA and social justice awareness, as well as examples of two labs that help students develop both CA and social justice awareness. The scaffolds and artifacts in the labs were designed with the intention of helping the students engage in CA and social justice awareness. The next sections describe the scaffolds and artifacts and their purpose, how I created them, and how I adapted throughout the DBR process.

1a. What scaffolds and artifacts are included in the labs?

I designed both labs to include scaffolds and artifacts for applying data science practices to solve problems. Scaffolds are the specific questions that I included in the labs that consisted of coding questions, individual reflection questions, and group discussion questions. Artifacts are the tools that the students used to complete the labs. These were technological tools, such as Python, Jupyter Notebooks, and GitHub, as well as tools that helped students understand the context of the problems such as datasets and studies. I designed the scaffolds in the labs to help guide the students through the process of completing the labs without significant instruction needed from the TAs. The artifacts included in the labs allowed the students to use real world tools to explore the datasets and analyze them to help them tell a story or solve a problem. Table 6 shows what artifacts are included in each lab and each artifact's definition and purpose.

Table 6

Types of Artifacts in the Labs

Artifact	Definition and Purpose	Lab A	Lab A'	Lab B	Lab B'
Python	Python is a computer programming language that is accessible, efficient and reliable, and supported by a large community. The Python programming language is what the students use to do the computational part of the lab. In other words, the students code in Python and use it to do data science.	x	x	x	x
Git and GitHub	Git is a free and open-source version control system designed to handle projects that have multiple people working on them. Students use git to retrieve the blank lab from the instructor and they submit their completed lab to GitHub.	x	x	x	x
Jupyter Notebooks	Jupyter Notebooks are coding notebooks that allow the students to type code and text in the same document. Using Jupyter Notebooks allows us to ask both coding questions with answers that involve code and individual reflection questions with answers that involve text.	x	x	x	x
Python Libraries	Python libraries are a set of built in modules that contain bundles of code that can be used over and over again in different	x	x	x	x

Table 6 (cont.)

	programs. The Python libraries that the students in this study use (pandas, random, matplotlib, etc.) allow them to have access to built-in functions to help them successfully analyze data. This saves time so that students do not have to write these functions themselves.				
Salary Dataset	Lab A revolves around the salary dataset. This dataset gives the salaries for all employees at a large university. This is a real dataset that students can use to answer questions they have about salaries of employees at a large university.	x	x		
Gender Salary Dataset	The Gender Salary Dataset is a subset of the salary dataset that also includes each employee's gender for two departments at a large university. This allows students to stratify by gender when analyzing the data.	x	x		
Name Discrimination Study Paper	The second part of Lab B involves simulating a name discrimination study. There is a paper written about this study. This paper shows students that this is a real-world example and allows them to get more details about it if they wanted to.			x	x
Resume Dataset	The Resume Dataset consists of each name, whether or not it was considered Black sounding or			x	x

Table 6 (cont.)

	white sounding, and whether or not it received a callback. For the name discrimination study, students needed a dataset to do the simulation. This allowed them to compare what they expected to get based on what the employers claimed versus what they actually got for the number of callbacks for the resumes with white sounding names and resumes with Black sounding names.

These artifacts were crucial for the students' completion of the labs. They intended to help students engage in CA and social justice awareness by giving them the opportunity to work with real world tools to help them answer important questions about social justice issues. In addition to these artifacts, I also included carefully constructed scaffolds in the labs to encourage students to engage in CA and social justice awareness through coding, individual written reflection, and group discussion. Each lab had multiple types of scaffolds that I adapted through the DBR process. Table 7 categorizes the scaffolds based on type. I listed the type of scaffold in the first column, a description of that type of scaffold in the second column, and a few examples of that type of scaffold in the third column. For a list of the individual scaffolds and artifacts in both labs, see Appendix C.

Table 7

Types of Scaffolds in the Labs

Scaffold	Description	Examples
Coding Questions		
Visualization Questions	These questions have students use Python to create simple visualizations.	-Create a frequency histogram of the salaries in the dataset. (Lab A) -Load the second dataset and create two boxplots of the male salaries and female salaries in this dataset. (Lab A) -Create a histogram of the results of your simulation. (Lab B)
Descriptive Statistics Questions	These questions have students use Python to calculate basic descriptive statistics.	-Next, calculate the overall mean, median, and SD of the salaries. (Lab A) -Next, calculate the mean of the dataframe that you created from the simulation. (Lab B)
Multi-step Coding Questions	These are coding questions that involve more than one line of code. They are more complex than the previous types of coding questions. The goal of these questions is for students to think about an algorithm for solving the problem before starting to code.	-There was a certain function you learned that will help you group all of the people in each department so you can find departmental aggregates. In other words, you want to group by 'Primary Department'. (Lab A) Next, instead of just grouping by a single aggregate, use Python list syntax to aggregate by count, mean, and median to find the mean and median salary, as well as the count for how many people are in each department. (Lab A) -Let's start by doing this simulation 100 times. In other words, we are simulating picking 100 juries. Store the results in a dataframe called df. (Lab B) Simulate this 100 times and store the

Table 7 (cont.)

		results in a dataframe. The variable or column in the dataframe should be named panel. (Lab B)
Unstructured Coding Questions	These types of questions allow students to use Python to answer questions that they personally find interesting or important, like a data scientist would. Unlike the other types of coding questions, these do not give any directions and each student should have a unique answer.	-Think about two questions that you have that have not been answered. Record them below. Then, answer at least one of these questions using Python and either dataset. (Lab A)
Individual Reflection Questions		
Statistical Questions	These questions examine students' knowledge of statistical concepts.	-Write a few sentences answering the following questions: Is the mean or median larger? Why do you think this might be the case? What does the standard deviation tell you in this context? (Lab A) -Explain why it is valuable to look at visual displays of salary data in addition to descriptive statistics like the mean, median, and SD. Write at least 3 sentences. (Lab A)
Impact Questions	These questions have students examine how the data science they are doing and the questions they are answering impact them personally or impact others.	-As a student, do you think these are fair salaries or are they too low or too high? Why do you think this? How does this dataset impact you? Write down your thoughts below. (Lab A) -Give an example of a way that we can use data science to help address issues of racism. This can be something you discussed in your group or an example you are interested in. (Lab B)
Discussion Reflection Questions	These questions have the students write about the discussions they had with	-Write down something you learned from your group discussion! Write at least 3 sentences. (Lab A and Lab B)

Table 7 (cont.)

	their peers.	-Write down the most interesting part of your group discussion. (Lab A and Lab B) -Write down something that surprised you from your group discussion. (Lab A and Lab B)
Expert Position Questions	These questions position the students as experts and are longer than the other individual reflection questions. The students are instructed to write a paragraph instead of a few sentences.	-Pretend that you are a defense attorney and a data scientist. Write a memo to the Supreme Court positioning yourself as a data scientist arguing whether or not you think the jury with 3 Black men was randomly selected. Justify your decision and include guidelines for the future. (Lab B)
Group Discussion Questions		
Coding Discussion Questions	These questions ask the students to discuss the results of the coding they did.	-Share the question you chose to answer with Python and the results with your group. (Lab A) -Discuss the results with your group. Some questions to think about: Do the boxplots look similar or different? Are there any outliers? (Lab A)
Statistics Discussion Questions	These questions have the students discuss a statistical concept.	-Discuss with your group whether you think a histogram or a boxplot or both best visualize the salary data. There is no right answer to this question. Explain why histograms, boxplots, or both are important and what they can tell us about the data. (Lab A) -Interpret the results of your histogram. How does your histogram provide evidence for or against the claim that the jury was not fair? What does this tell us about this case? (Lab B)
Impact Questions	These questions have the	-Who benefits from collecting this

Table 7 (cont.)

	students discuss the implications of their analysis. They have them reflect on the importance of using data science to explore social justice issues and think critically about the work they are doing.	salary data? Who does this data harm? Why do you think this salary data is public? Can you think of any reasons that this could be problematic? (Lab A) -This is an example of how we can use statistics to help us solve real world problems. Discuss with your group how simulations and data science can be used to help address issues of racism specifically. (Lab B) -The Smith trial happened in 1993 and the name study occurred in the early 2000s. Discuss with your group whether or not you think similar events still occur today and why. Reflect on how data science can be used to educate people about this. (Lab B)

The labs included three types of scaffolds: coding questions, individual reflection questions, and group discussion questions. Table 7 further divides these questions into descriptive categories. The scaffolds were intended to guide the students through the lab and give them context for the problems they are solving and encourage them to talk about social justice issues. The labs were intended to allow students to act as real data scientists by practicing multimodal communication through coding, writing, and discussion with their peers. All of the scaffolds and artifacts included in the labs were created and chosen using principles of distributed cognition and I describe these principles in the next section.

1b. What principles of distributed cognition are used in the design of these scaffolds and artifacts and why?

I used ideas from distributed cognition (DC) to design the scaffolds and choose the artifacts for both of the labs in this study. Using the DC theory, I summarized some of the ideas

into three principles of distributed cognition (which I will refer to as DC principles) that I used to guide the creation of the labs. The three principles are the Shared Knowledge Principle, the Internal Artifacts Principle, and the Collaboration Principle. For each of these principles, I named them based on my interpretation of the theory and described what they are. This included a description and the assumptions I made about the principles. I discussed how they connect to the DC theory and how each principle was specifically applied to the labs. Lastly, I reflected on what students said about the scaffolds and artifacts that were designed using the principles.

The Shared Knowledge Principle. The first DC principle that I will discuss is the Shared Knowledge Principle. The Shared Knowledge Principle says that the labs should include artifacts and scaffolds that elicit knowledge sharing among students to help them engage in computational action (CA) and learn more about social justice issues. The assumption is that every member of the group possesses knowledge that can be shared with the other group members. This knowledge sharing helps the students accomplish two of the goals of the labs: computational action and social justice awareness. Both CA and social justice awareness are the two primary goals that this principle helps achieve. The third goal of the labs, communication, is the means for the students to share their knowledge, but communication is not the main focus for this particular principle. Although implicitly, by sharing their knowledge, they are developing communication skills. These three goals are the mediating processes in the Conjecture Map shown in Figure 8.

By designing the labs to contain scaffolds and artifacts that encourage knowledge sharing, I considered that each student would bring a different set of knowledge to their group. Each student has knowledge that comes from their own domain (or major) and from their unique life experiences. They bring this knowledge with them as they work together in groups to discuss

the coding questions in the labs and engage actively in the group discussion questions. By sharing their knowledge with others and learning from their group members as they share the knowledge that they have, the students build on their previous knowledge. The scaffolds and artifacts in the labs directly encourage students to share their knowledge and learn from each other. This knowledge can be about data science, programming, social justice issues, particular fields of study, or other topics. Learning about and discussing these topics helps the students engage in CA and social justice awareness.

Connection to Goals. The first goal of using the Shared Knowledge Principle to create scaffolds and artifacts for the labs is for the students to engage in computational action. There are two parts to CA: computational identity and digital empowerment. The Shared Knowledge Principle specifically was created to help students engage in computational identity. Computational identity says that the students feel that they are a part of a group where each member's ideas and thoughts are important when designing creative solutions to a problem. Through working together and answering questions that elicit knowledge sharing, the students are contributing to their group, learning how to work together to solve problems and think about real social justice issues, and learning from each other. Because they are discussing and thinking about social justice issues, engaging CA should also help the students grapple with social justice awareness.

Both labs are centered around social justice issues and the scaffolds and artifacts in the labs encourage students to share knowledge that they have about the issues in the labs and other social justice issues with their group members. One way to learn about social justice issues is to have discussions about them. In my experience, this is uncommon in data science classes and many math related classes. When designing these labs, I chose two social justice topics that

allowed students to learn about the topics through coding and then learn more about them through discussion. The first topic was salary discrepancies among genders and other factors. I chose a dataset that allowed students to discover the discrepancies and have discussions about the implications of their findings. The second topic was jury selection and racial discrimination. The students were able to simulate the scenario as a real-world event in Python and discuss the implications of doing this. These labs gave the students a space to have these discussions and learn from each other. Students shared their perspectives on these issues and learned from the other students in their group who may have had different perspectives and ideas on social justice issues. Sharing knowledge amongst group members is an idea from Distributed Cognition (DC) and something that was considered when designing all versions of the labs. The next section describes how the Shared Knowledge Principle connects to the DC theory.

Connection to Theory. Distributed Cognition is a theory that looks at how knowledge is distributed among individuals and their surroundings (Hutchins, 1995). The Shared Knowledge Principle says that one of the ways that knowledge is distributed is among the group members and the scaffolds and artifacts in the labs elicit knowledge sharing among the students. One example of this idea comes from Hutchins' *Cognition in the Wild*. In *Cognition in the Wild*, Hutchins talks about a situation where the ship that the crew was navigating lost steam. The crew members all had to use their knowledge to work together to stop the ship so that it would not crash into the harbor. Hutchins described how no single crew member could have done this alone and how the knowledge was distributed among the crew members. The crew members do not know how to do every task on the ship, instead they each have a unique role that helps contribute to successful navigation. During this situation, the crew members had to share the knowledge they had about their job so that everyone could help stop the ship. The crew members learned

from each other and relied on each other to complete the task, similar to how I expect the students learn from each other and rely on each other to complete the labs.

Another example of how the Shared Knowledge Principle connects to the DC theory comes from Achiam et al. (2014). They used DC to study how visitors in a museum interacted with exhibits to make sense of them. They studied the participants in the museum and how they interreacted with each other and the parts of the exhibits. Each of the visitors have knowledge that they bring to the exhibit and the exhibits were designed so that the visitors could use their knowledge and learn from the artifacts provided at the exhibits. The exhibits in the Achiam et al. (2014) study were created to elicit knowledge sharing among the participants similar to how the scaffolds and artifacts in the labs elicit knowledge sharing among the members of the groups in the labs. The artifacts in the exhibits were meant to encourage knowledge sharing, discussion, and learning, however the participants needed to use them and engage with them to learn. This is similar to the scaffolds and artifacts in the labs. The students need to use them and engage with them to learn.

A third example of how the Shared Knowledge Principle connects to the DC theory involves education research. Karasavvidis (2002) talked about the social aspect of DC. He advised teachers to focus on more practical and situated tasks that do not have a single solution. He also recommended that the scaffolds in these tasks should encourage students to work together and share their knowledge. Because a big part of DC is looking at how people interact with others, Karasavvidis (2002) wanted teachers to encourage as much knowledge sharing as possible among students. When I designed both labs, I included real data and practical problems. Many of the scaffolds, especially the discussion questions, do not have a single solution or a correct answer. The artifacts related to social justice and the discussion questions encourage

students to discuss the social justice topics and share their unique knowledge. Next, I will discuss the specific scaffolds and artifacts that were designed using the Shared Knowledge Principle. In other words, I will give examples of how this principle was applied to the labs.

Examples of Scaffolds. The artifacts and scaffolds in the labs were included to encourage students to share their knowledge with their group members and contribute something unique to their conversations. In other words, I wanted the students to understand the benefit of working in small groups and make sure that each student felt that they could contribute to the group as well as learn from their group members. In both labs, I included scaffolds that encouraged students to be creative, explore their own interests, and engage in multimodal communication to share their knowledge. This allowed them to develop CA. For example, in Lab A, one scaffold asked students to use the salary dataset to create visualizations and calculate descriptive statistics for their own department (for example, if their major was History, they would do this for the History department). Another scaffold asked them to discuss their findings with their group members. Through these discussion questions about the analysis they did on their own department, each individual student is contributing something unique to the group discussion. Even if there were two students in one group who are in the same department, their interpretations or thoughts are still unique. I chose to highlight this example because it was an example where each student was able to create something unique and the scaffolds encouraged them to share it.

Another example of how scaffolds that promoted knowledge sharing helped students engage in CA comes from discussion reflection questions. These were individual reflection (IR) questions that had the students write about what they learned from their discussion with their peers. These scaffolds were placed throughout the labs after the group discussion questions. The idea was that in order to answer them, students had to participate in the group discussions by

either sharing the knowledge they have or by learning from their peers. I chose to discuss this example because these scaffolds required students to write down something from their discussion. If they did not share their knowledge and learn from other students during the discussion, they would not be able to answer this question. By reflecting on the discussions that they had, the students are engaging in CA because they are thinking critically about real world issues and discussing them with their peers.

Examples of Artifacts. In addition to knowledge sharing scaffolds that helped students engage in CA, I also included artifacts that encouraged students to share their knowledge about the topics. These specific artifacts were the social justice examples and datasets that were included in the labs. One of the main artifacts that was included because of the Shared Knowledge Principle was the Salary Dataset from Lab A. The Salary Dataset was one artifact that was chosen to help students analyze salary discrepancies based on gender, department, and other factors. By using this artifact, I was able to create scaffolds that allowed students to share the knowledge that they found from analyzing this data with their group members. Because they were looking at discrepancies and discussing them, they were also grappling with social justice awareness. A second example of an artifact that helped students engage in social justice awareness by sharing knowledge is the Jury Selection Scenario from Lab B.

Lab B was themed around a trial in which the jury selection was being questioned for being truly random. In the actual trial, there were only three Black people in the jury pool when the expected value was eight (the population was 8% Black). To decide if the jury was selected randomly, students simulated picking a jury over and over again from the population and seeing how many Black people ended up on the jury. This example was another artifact that was chosen to help students learn about random sampling, expected values, and errors. This social justice

example allowed me to create scaffolds that asked students to share the knowledge that they learned from analyzing this particular example, as well as share their knowledge about other social justice issues involving racial discrimination. As described, social justice examples, along with scaffolds that elicited knowledge sharing about these topics were included in both labs. Students have different perspectives and knowledge about the social justice issues presented in both labs. By designing scaffolds using the Shared Knowledge Principle and choosing artifacts that help create these scaffolds, the students can engage in CA by learning more about social issues and getting experience discussing them with their peers.

Student Perceptions. After seeing how the Shared Knowledge Principle was applied to the labs, it is also beneficial to look at what students said about the artifacts and scaffolds designed using this principle. To investigate this, I looked at the end of semester interviews. Throughout the interviews, a common theme was that students mentioned that they liked the social justice parts of the lab. They liked them because they felt like they had the opportunity to share knowledge with their peers. They enjoyed these discussions and they specifically liked when the labs included scaffolds where each group member could contribute something unique to the discussion. For example, one student said: "I thought it was a really good way to get people thinking about social justice issues, especially if you're in STEM, you don't really think about that as often." This was regarding the artifacts related to social justice. In both iterations (Fall 2021 and Spring 2022) of implementing the labs, all of the students we interviewed responded positively when asked how they thought about the labs including the artifacts about social justice.

Students also responded positively to the scaffolds in the labs that encourage knowledge sharing. Earlier, I described a scaffold that had the students each work on answering questions

specifically related to their own department. After they finished the analysis, they were encouraged to share their results and their knowledge with their group. In Fall 2021, six out of seven students mentioned that they liked this question (one student did not remember it). And in Spring 2022, five out of seven students mentioned they liked this question (two did not remember this specific question). One student mentioned that they had knowledge to share about one of the topics in the lab. This student said: "I personally liked the lab, because I'm familiar with like, you know, wage inequality in some fields. So that was something I liked talking about." In other words, this student came in with knowledge about the topic and was able to share it with their group members. Overall, most students reacted positively to the scaffolds and artifacts designed using the Shared Knowledge Principle.

The Internal Artifacts Principle. The next DC Principle is the Internal Artifacts Principle. The Internal Artifacts Principle says that the labs should give students the opportunity to use internal artifacts, such as datasets and computational tools, to engage in CA and learn more about social justice issues. This principle assumes that the internal artifacts in the labs are not absolutely necessary for accomplishing the tasks of the labs, but they do help make this easier for students. The internal artifacts included in the labs are Python Libraries, Github, Jupyter Notebooks, and the two social justice examples that I described previously that include specific contexts (salaries and jury selection) and datasets. Choosing artifacts for the lab with the Internal Artifacts Principle in mind is intended to help the students engage in CA and social justice awareness. They engage in CA through using real world tools to analyze data in a social justice context.

Connection to Goals. The two goals associated with this principle, CA and social justice awareness, are highly interconnected. There are two parts to CA: computational identity and

digital empowerment. The Internal Artifacts Principle specifically helps students engage in digital empowerment. Digital empowerment says that students should critically engage with authentic and relevant work that impacts their lives and interests them. During the labs, the students are supposed to engage in digital empowerment through working on problems that are situated in the context of social justice. Both of the labs are centered around social justice issues and contain examples that are authentic and relevant. The first lab is about salary discrepancies and other factors such as gender. The second lab is about jury selection and racial discrimination. These contexts and datasets are artifacts that allow students to engage in meaningful work that is interesting and impacts peoples' lives. The social justice contexts of the labs also helped make it easier to include scaffolds that encourage students to share knowledge, communicate, and work together to think about the impact of the analysis that they are doing in the labs.

Digital empowerment also teaches students how to bring the work they have done inside the classroom to other situations outside of the classroom. The social justice examples and datasets included in the lab cover topics that are relevant to life outside of the classroom. Students can apply what they learn from exploring these topics and examples in class to similar situations in the real world. For example, although the trial and jury selection example in Lab B occurred in 1993, racial bias in jury selection is still a problem that we are facing today. The context of Lab B allowed students to develop the programming skills to simulate jury selection for a specific population to determine if there is evidence that the jury was randomly selected. In addition to the contexts, the other artifacts included in the lab (Python libraries, GitHub, and Jupyter Notebooks) are all tools that real data scientists that work in industry use in their day-to-day jobs. These artifacts can easily be used in other contexts beyond just the classroom. Throughout these labs, students are working with real tools and analyzing real data, making the

context authentic. These tools help make the programming more efficient for students, meaning that the work can be done faster and with less lines of code, but the tools are not absolutely necessary. They could do these calculations by hand or in another tool like Excel, although by using the specific artifacts that were chosen for the labs, the students can do the calculations quickly and efficiently. Overall, using internal artifacts to help accomplish a task with a specific learning outcome is an idea from Distributed Cognition (DC). This idea was considered when designing all versions of the labs. The next section will describe how the Internal Artifacts Principle connects to the DC theory.

Connection to Theory. Distributed Cognition is a theory that looks at how knowledge is distributed among individuals and their surroundings (Hutchins, 1995). The Internal Artifacts Principle says that students should use helpful internal artifacts to help accomplish the specific goals described above. This idea comes from Hutchins' *Cognition in the Wild*. In the book, Hutchins described how some of the artifacts available on the *Palau* made computations easier, such as the compass rose, fathometer, and charts constructed by previous navigators. All of these were considered internal tools located within the ship that the navigation team could use. They were not necessary for the task of navigation, but they did help with many different parts of the tasks, especially ones that were computational. The artifacts that I intentionally included in the labs to help with computation (Python libraries, GitHub, and Jupyter Notebooks) are not absolutely necessary for engaging in CA, however they are very helpful because they allow the students to use tools that align with life outside of the classroom. The other artifacts involving social justice (the contexts and datasets for the two labs) also were not absolutely needed to engage in social justice awareness (there are other examples that could have been chosen). However, I felt that these examples worked best for designing both coding questions and group

discussion (GD) questions around them.

Another example of how the Internal Artifacts Principle connects to the DC theory comes from Karasavvidis (2002). Karasavvidis (2002) studied DC in the context of education and looked at how students solved a correlation question by hand using a paper and pencil, as well as solving the same questions using a computer spreadsheet. In both cases, the knowledge was distributed among the students and the different helpful internal artifacts. The computer spreadsheet served as an aid to help make the problem faster by lowering the mental processing and cognitive labor. Karasavvidis (2002) argued that this also helped reinforce the knowledge because the students could focus on more than just the calculations. This plays a very similar role to the artifacts that I used in the lab, such as Python Libraries. Students could solve these problems without them, just how the students solved the correlation problem with pencil and paper. But by using them, this lessens the cognitive load, which in turn allows them to think more critically about the context and the implications of this problem.

Examples of Scaffolds and Artifacts. Next, I describe how the Internal Artifacts Principle was applied to the labs and give examples of this. I included many internal artifacts to help students complete the labs and get practice coding, as well as think critically about the social justice issues that were included in the labs and engage in CA (specifically digital empowerment). These artifacts helped the students engage in CA by allowing them to work with real tools that data scientists use and think critically about social justice issues. The Salary Dataset was one artifact that was chosen for Lab A to help students analyze salary discrepancies based on gender, department, and other factors. By using this dataset, students were able to explore what discrepancies exist. Through the scaffolds that were motivated by this context, the students also discussed why these discrepancies exist and their implications. By using a relevant

example with real world data, students were able to analyze the data and use their computational skills to make sense of the social justice issues present in the lab. This was intended to allow students to engage in digital empowerment as they think about real world issues that can impact people's lives.

Another example of how the Internal Artifacts Principle was applied to the labs is the decision to include the trial and jury selection example as the main context for Lab B. This specific example was another artifact that was chosen to help students discover whether or not it was likely that the jury was selected randomly. The students did this through coding exercises that involved simulation. Here, students were able to explore racial discrimination in jury selection for a high stakes trial. This is an incredibly relevant social justice topic that can be applied to the world outside of the classroom. Similar to the previous example, by using this context, students are able to use data science to make sense of and learn more about the social justice issues present in this lab. This also helps the students engage in digital empowerment since they are again thinking about real-world issues that impact people's lives.

A third example of how I applied the Internal Artifacts Principle to the labs is the decision to use GitHub, Python Libraries, and Jupyter Notebooks in both labs. The students use these tools to engage in digital empowerment because they can make a connection between what they are doing in class and the real world. These are all tools that data scientists use every day to help make analyzing data more efficient. GitHub allows the students to easily retrieve the lab from the instructor, get it on their personal computer, and then save their work each time they work on it. This makes it easier for students to work on their analysis, stop working, and store their work until they want to come back to it. Python libraries have built in functions that allow the students to perform operations on the data in one line of code. Including these in the lab

makes it so that students do not have to spend large amounts of time writing their own functions. This time that is saved allows them to spend more time thinking about the implications of their analysis and the social justice aspects of the labs. Juypter notebooks allow the students to type both code and written content into the same document. This helps the students be able to connect the coding that they are doing to the reflection on the social justice issues. Overall, these three artifacts help the students engage in CA and learn about social justice, which could be done in other ways, such as simulation or giving students descriptive statistics. However, I chose these particular artifacts so that the students could have an authentic and interesting context and be able to get practice working with real data, as well as thinking critically about their analysis and the decisions that they made. I also wanted them to think about the implications of their data analysis on the real world.

Student Perceptions. After seeing how the Internal Artifacts Principle was applied to the labs, it is also beneficial to look at what students said about the artifacts chosen with this principle in mind. To investigate this, I looked at the end of semester interviews. Overall, the students appreciated the opportunity to grapple with the social justice issues in the labs. Many of them specifically talked about how they felt having real world social justice examples in a data science lab. In Fall 2021, all but one student recognized that a major theme of both of the labs was social justice. Six out of seven students that were interviewed enjoyed having labs that included artifacts relating to social justice. They enjoyed being able to apply the data science skills they learned in lecture to real world contexts. Seven out of seven students thought that including these social justice topics in the labs is important. However, there was one student from this set of interviews who did not realize that Lab A, which was the lab about salaries, was related to social justice. There was also one student who said that it was difficult to talk about

social justice issues, but they recognized that it was important. They said: “It was a little hard to just talk about the intersection of data science with social justice issues. It was kind of uncomfortable. But I felt like it was much needed.” This is an interesting point that although this may be uncomfortable at first, it is beneficial for students and this particular student recognized this. In Spring 2022, all of the students recognized that social justice was a theme in both labs. These students also said that they liked having labs that included artifacts relating to social justice, they liked being able to apply the skills they learned to real world problems and believe that including the social justice artifacts is important.

In addition to the social justice contexts, I also looked into whether students mentioned the other internal artifacts included in the labs (Python libraries, GitHub, and Jupyter notebooks). Most students did not specifically mention these internal artifacts, however many of them mentioned that they liked using Python and thought it was a useful tool to help make analyzing data and visualizing data easier. For example, one student from Fall 2021 said: “I think in general I've gotten a lot better at Python through this class. And I feel like, there are kind of a lot of places where you can use Python outside of just this class.” This student recognized that the labs have helped them become more proficient with Python and that knowing Python can be useful outside of the classroom. Another student from Fall 2021 did mention Python libraries specifically when talking about how the content of the labs will be useful in their future job. They said: “A lot of C+ and Java applications have these requirements, such as knowing Python, and specifically like learning pandas as one of their libraries. So I feel like as a one-semester course, there was a lot of experience with pandas[5] library, and also a little bit with, data visualizations. So it's definitely very helpful for future jobs.” They are saying that Python

[5] Pandas is the main Python library used in this course.

libraries are useful in other programming applications. Overall, the students recognized the connection between many of the artifacts used in the labs and their importance in life outside of the classroom.

The Collaboration Principle. Next, I discuss the Collaboration Principle. The Collaboration Principle says that the labs should contain scaffolds that promote collaboration among students to give them the opportunity to engage in CA and to practice communicating about data science through talking. The assumption is that the individual students are a part of a bigger system that includes their group members. The students are a key part of this system, and they must work together to complete tasks. In other words, collaboration among group members is necessary and they need to communicate with each other in order to collaborate. The scaffolds included in the labs encourage the students to collaborate because if I am assuming the knowledge is distributed, the students will learn more from collaborating with others, rather than working alone by themselves. Some of the scaffolds in the labs actually contain questions that are impossible to answer without collaboration, emphasizing how important I thought collaboration was when designing the labs. The scaffolds created using the Collaboration Principle have two main goals for the students: improving communication and enaging in CA. The third goal of the lab, social justice, is also connected to this principle since many of the scaffolds that require collaboration are related to social justice. The three goals of the labs are highly interconnected for the scaffolds developed using the Collaboration Principle.

Connection to Goals. Collaboration is intended to help students improve their communication skills. There were many scaffolds in the labs that encouraged students to work together and practice communication through talking and writing about the data science that they are doing. These questions helped foster a culture in the labs where communication is necessary

and important, working together is more beneficial than working individually, and that learning data science is a journey that they are all on together, rather than an isolating journey that they take alone. All of this helps the students engage in CA. There are two parts to CA: computational identity and digital empowerment. The Collaboration Principle specifically aims at helping the students engage in computational identity. During the labs, they engage in computational identity through collaborating with each other and answering questions that involve communicating with their group members. Computational Identity says that the students are a part of a group where each member's ideas and thoughts are important when designing creative solutions to a problem. By including scaffolds that foster collaboration, the students are able to engage in computational identity as they form relationships with their group members, learn from them, and communicate with them during the labs. Many of these scaffolds that help students engage in computational identity often include discussions about social justice issues.

Students' collaboration within groups models a data science practice that they could apply in their future jobs. Data scientists are almost always a part of a bigger team that works together and communicates frequently. This is also the case for many other jobs. One part of computational action is the idea that the students are learning skills in the classroom that can be applicable to their lives and jobs outside of the classroom. Collaboration with others is something that all students, regardless of their major or careers goals, will have to do in their jobs and their lives. Collaboration and working together is an idea from Distributed Cognition (DC) and something that was considered when designing all versions of the labs. The next section will describe how this principle connects to the DC theory.

Connection to Theory. The Collaboration Principle says that the labs should include scaffolds that encourage and promote collaboration to accomplish the two goals described in the

previous section. In his book, *Cognition in the Wild*, Hutchins (1995) described how each person's individual job was integrated into the entire task of navigating the ship and how the crew members had to work together to be successful. In other words, the crew members had to collaborate in order to complete the tasks. I also emphasized the importance of working together through the scaffolds in the labs by creating questions that could only be answered by collaboration. Hutchins also talked about how it is important to study cognition outside of the lab so that you can see how team members work together in their natural environment. He emphasized that observing the crew members on the ship partaking in navigation and collaborating was better than bringing the crew members to a lab. During this study, I also studied students in their natural environment (their lab sections) instead of bringing them in to an unfamiliar lab. This allowed me to see how they collaborate naturally in their system, rather than a controlled setting.

Lave (1988) also described how many studies on student learning do not look at students in their natural environment, but instead have students come to a lab to be evaluated. Having students collaborate in a research lab is not as authentic as having students collaborate in the classroom. The aim of this study was to make the research as authentic as possible by having students work together naturally in their system. Lave (1988) also talked about how cognition should be studied in the wild because the setting can impact the systems' ability to perform tasks and work together. Hence, this can affect collaboration. I also studied students in their lab sections and used audio recordings to understand when and how collaboration was happening among the group members.

Karasavvidis (2002) studied DC in the context of education and explored the social and collaborative aspect of DC. He emphasized that learning is social, and that students' proficiency

is distributed. He advised that teachers should focus on situated tasks that do not have one correct answer to mirror real life and encourage collaboration. I included many scaffolds in the labs that did not have one right answer and encouraged students to collaborate and discuss different solutions. Another example of an education researcher who studied DC and focused on collaboration was Evans et al. (2011). Evans et al. (2011) looked at how children communicated when solving a geometric puzzle in a CSCL context. The students were working in groups and the task was designed so that students were supposed to collaborate with each other. They found that learners were more likely to discuss and articulate their ideas in the CSCL setting. Our labs are very similar to a CSCL setting because the students are using computers to complete the labs, working in groups, and collaborating to come up with a solution. The authentic setting and scaffolds that encourage collaboration seem better for accomplishing the goals than an inauthentic setting and questions that can be answered individually. Next, I will describe how the Collaboration Principle was applied to the labs and show how it helped create scaffolds that allowed students to engage in CA and communicate.

Examples of Scaffolds. I included scaffolds that encourage collaboration among group members to help the students get practice communicating data science, communicating about social justice issues, and engaging in CA (specifically computational identity). I also included scaffolds that specifically required students to collaborate through talking. These were all of the group discussion questions and the discussion reflection questions. The group discussion questions had the students discuss something verbally with their group. The topics were usually something related to their code, something related to a statistical concept, their opinion on something, or a reflection on an analysis that they did. For example, in Lab B, one of the discussion questions asked: "This is an example of how we can use statistics to help us solve real

world problems. Discuss with your group how simulations and data science can be used to help address issues of racism specifically." Here, the students are being encouraged to discuss a social justice topic. By explicitly asking them to talk about this, they are getting practice discussing how data science can be used to address real world issues.

After the group discussion question, there was also an individual reflection question asking students to describe one way that we can use data science to address issues of racism. This question had students use what the learned in the lab to think about another application from the real world. Their collaboration through the previous group discussion question should help them be able to think critically about this question, come up with ideas to answer it, and communicate their thoughts in writing. I chose to highlight this scaffold as an example of a scaffold that was created using the Culture of Collaboration Principle in order to foster communication between students because it had two parts (group discussion and an individual reflection question that was based on the group discussion). Also, it is open ended, yet gives students a direction on what to talk about, which hopefully would yield more collaboration, leading to more practice communicating. It was also directly related to social justice.

Another example of this from Lab B is a scaffold that asks students to "Write down something that surprised you from your group discussion." The group discussion was in regard to the resumes with white sounding names and Black sounding names. In this section of the lab, the students did a simulation to see if there was a significant difference between callbacks for resumes with white sounding names and Black sound names. They saw through the simulation, that resumes with Black sounding names were significantly less likely to get callbacks. In order to answer this question, the students must have not only collaborated, but also communicated during the previous group discussion question since it is directly asking about that. In other

words, this type of scaffold requiring communication encourages students to draw on their group discussions to think about their responses for the scaffolds that depend on them. I chose to highlight this scaffold because it specifically involves social justice and requires the students to listen to what others said and describe it in writing.

In addition to helping students improve their communication skills, the scaffolds designed using the Culture of Collaboration Principle also help students engage in CA. The scaffolds in the lab also help the students form groups that they feel comfortable having discussion with. Being an active part of a small group that communicates, collaborates, and works together to solve a problem helps students develop computational identity. For example, in Lab A, we asked students to pick a question that they have about the Salary Dataset and answer it. Once they do this, there is a group discussion question that asks: "Share your question and results with the group." This not only helps students communicate, but it also helps them get comfortable sharing their own analysis and their own ideas with the group. By sharing something unique that they did, this encourages other students to ask follow-up questions. This gives the students experience working in a group together to solve problems. I chose to highlight this scaffold because it helps the students engage in CA through communication, which are both of the goals that his principle helps students meet.

Student Perceptions. After seeing how I applied the Collaboration Principle to the labs, it is also beneficial to look at what students said about the scaffolds chosen with this principle in mind. To investigate this, I looked at the end of semester interviews. Overall, the students said that they liked the questions that required discussion and collaboration, but not all students they felt like their groups engaged in this. A few students said they wished their groups collaborated and worked together more. Here is what one student said about collaboration: "I think when it

comes to groups, group collaborations are really important, and I notice that when I work in a group, I learn things better." They specifically pointed out that collaboration is important because when they collaborate, they learn, and that working with others is more beneficial than working alone. Another student said: "I think the collaboration was pretty good. I think how my group does it is that we all kind of work on one part of the lab together. And then after we're all done go over what we did, what answers we got, if we had any kind of problems with it. And I think because at least everyone in my group kind of works at the same pace, I feel like it works pretty well, in a collaborative sense." This student said that their group worked on the same part of the lab at the same time so that they could help each other and collaborate if anyone got stuck. A third student specifically said that they enjoy collaborating with their group. They said: "I like that most of the labs are very focused on collaborating with your group. I know most labs are like that. But these two especially had a lot of group discussion, individual reflection questions. And that kind of gave me an opportunity to look at what my partner thinks because most of my partners are in stats. So it's cool to see their differing perspectives on social justice issues." This student also pointed out that their group members are statistics majors (they were a non-STEM major) and that they had different perspectives on the social justice issues. The scaffolds designed using the Collaboration Principle allowed them to learn from each other.

As stated previously, there were some students who expressed issues with the collaborative nature of the labs. Specifically, these students mentioned that it was difficult to collaborate while working on the coding questions and that not all groups had good collaboration. For example, one student liked the collaboration in the group discussion questions but felt like collaborating with coding was difficult because the students were all at different levels. They did mention that for the discussion questions, this was not an issue. They said: "But

I think the reflection questions aren't like that because most people are on the same page, it's just that there's such a big disparity with coding skills, that a lot of times there's a problem." In other words, they thought that it was difficult to collaborate on the coding questions because they did not believe they were at the same coding level as their group members. Another student saw another group that they were not a part of collaborating well and wished all groups were like that one. They said during the interview: "I feel like the group work is kind of hit or miss. There's like one group in my section that's, super talkative, and they all like to help each other. And I'm not always with that group. I wish all of the groups are like this." In other words, they said that some groups collaborated well, but not all of the groups did. It is important to look at how the students felt about the scaffolds and artifacts created from the principles because if there are issues that come up, I can understand why they came up, which will help me fix them as a part of the DBR process.

Summary of DC Principles. I used three DC Principles to design the scaffolds and artifacts in the labs in order to help students improve their communication, engage in social justice awareness, and engage in CA. Table 8 summarizes each DC Principle, the description, and what scaffolds and artifacts were created using them.

Table 8

DC Principles

DC Principle	Description	Scaffolds and Artifacts
Shared Knowledge Principle	The labs should include artifacts and scaffolds that elicit knowledge sharing among students to help them engage in computational action (CA) and learn more about social justice issues.	Individual reflection questions about group discussion, group discussion questions about coding questions
Internal Artifacts Principle	The labs should give students the opportunity to use internal	Python Libraries, Jupyter Notebooks, GitHub, the two

Table 8 (cont.)

	artifacts, such as datasets and computational tools, to engage in CA and learn more about social justice issues.	social justice examples, the salary dataset
Collaboration Principle	The labs should contain scaffolds that promote collaboration among students to give them the opportunity to engage in CA and to practice communicating about data science through talking.	Group discussion questions, the two social justice examples, the salary dataset

The scaffolds and artifacts included that were designed using the Shared Knowledge Principle elicit knowledge sharing among students which in turn allowed them to engage in CA and social justice awareness. The artifacts chosen using the Internal Artifacts Principle were specifically chosen to help the students engage in CA and social justice awareness. These artifacts gave students practice working with real world data science tools and allowed them to explore relevant social justice issues. Lastly, the scaffolds designed using the Collaboration Principle helped students get practice communicating through talking about the implications of their data analysis. This also helped students engage in CA and social justice awareness. In the next section, I talked about how the scaffolds and artifacts evolved throughout the DBR process.

1c. What adaptations to the labs (scaffolds and artifacts) result from the DBR process and why?

Design-based research (DBR) is an iterative process that allows researchers to make meaningful changes in classrooms. In this study, the DBR process was a journey in which I created data science labs and then revised them based on feedback from the students and observing what happened during the labs while the students completed them. I designed these labs with three goals in mind: computational action (CA), social justice awareness, and

communication. However, researchers never design perfect interventions. The beauty of using the DBR methodology is that I was able to make adaptations to the labs to improve them after the first iteration. The first lab (Lab A) covered visual displays of data, descriptive statistics, and exploratory data analysis through analyzing a dataset of salaries. The second lab (Lab B) covered simulation through looking at jury selection and resumes. Lab A and Lab A' were very different from each other, whereas Lab B and Lab B' had less changes. Since Lab A was the first lab that I redesigned, it needed the most adaptations as I learned the most from this iteration.

Below I describe the problems I noticed in the first iteration of the labs (Fall 2021). After describing the problems, I describe what type of data showed me there was a problem, an example of a scaffold that had this problem, how that scaffold was intended to connect to either the DC principles or learning outcomes, and the strategies I used to attempt to solve these problems in the second iteration (Spring 2022). These strategies were adaptations to the labs that I categorized into three categories: additions, deletions, and modifications. Additions are new scaffolds that I added to the labs that did not exist in the previous version. Deletions are existing scaffolds that I deleted completely. Modifications are changes that modify an existing rather than deleting it altogether. Lastly, I summarized each problem with the implications for solving them.

Problem 1: Lab A was too long. The data or evidence that the lab was too long came from the audio recordings of the labs. None of the groups who did Lab A were able to finish the lab during their section. Most of the groups (5 out of 6 groups) did not even get to the halfway point in the lab, implying that it was way too long to complete during the 80-minute lab section. Only one group was able to get slightly more than halfway through the lab. Because most of the discussion questions were at the end of the lab, none of the students actually got to this important part of the lab. Hence, they were unable to engage in the group discussions. The GD questions

were all examples of scaffolds that had this problem. In all of the audio recordings from Lab A, none of the students discussed any of the group discussion questions. The only discussion in these audio recordings was about how to do the coding required for the lab. Also, some of the students mentioned in the recordings that they were unable to finish the questions. For example, one student said: "There is no way we can finish all of these questions." A second student from a different group said "Well, I guess we'll have to do the rest of these [questions] on our own."

This problem violated some of the DC Principles that I used to design the scaffolds. The length of the lab made it so that the students could not share knowledge or collaborate about anything other than code. This violated both the Shared Knowledge Principle and the Collaboration Principle. Although the students were communicating about code, they were missing a huge part of both CA and social justice. Because they were not having group discussions, they were not engaging in computational identity or digital empowerment (the two components of CA). Also, most of the discussion on social justice comes from the GD and IR questions. Since the students did not get to these questions, they missed out on learning from their peers about social justice and discussing social justice issues. In other words, they did not have time to have a meaningful discussion on the social justice topics in the labs.

Overall, I used modifications and deletions to adjust the length of Lab A. Since the lab was too long, I had to cut parts out. I discuss the exact questions I cut out and why in the next sections. I also modified some questions for various reasons which I describe below. Some of the reasons for cutting and modifying questions included redundancy, unclarity, and the question being superficial and not producing a meaningful discussion. By realizing that Lab A was too long, I was able to keep this in mind when designing Lab B. I made sure that Lab B was not as long as Lab A, and it also showed me how to modify Lab A'. It is important to note that although

lab A did not give great data from the audio recordings since the students only discussed code, it was useful because it got me to Lab A'. Lab A being too long prompted me to make changes to Lab A' (described below) and design Lab B using what I learned from Lab A.

Problem 2: Some questions in the lab were repetitive. When analyzing the data from the audio recordings of the labs, I noticed that some of the questions in Lab A were repetitive, as I was asking the same thing multiple times. In other words, there were questions that asked the students to do the same thing, just with a different dataset. There were also some questions that covered topics that had been covered by previous questions. Since I did not want Lab A' to be too long, I often cut out repetitive questions. The data or evidence that some questions were repetitive came from the labs and the students' responses. The students were writing similar code and written responses to the repetitive questions.

One example of a repetitive scaffold from Lab A was a section where I had the students explore three different departments: English, Psychology, and Electrical & Computer Engineering (ECE). They were asked to analyze these three departments and compare different descriptive statistics about their salaries. I chose these three departments to highlight how different they were. The English Department is one of the departments in the College of Liberal Arts and Sciences with the lowest average salary, Psychology is more of an average department in terms of salary in the College of Liberal Arts and Sciences (it is also a common major for students in the course), and then ECE is a department with high salaries in the College of Engineering. This coding question asked the students to find the mean and median salary of these departments and there was also an individual reflection question about this exercise. However, there were other questions in Lab A that had the students do the same coding exercises and compare departments. Hence, this question was repetitive.

When designing this question, I connected it the learning outcome of engaging in CA. Knowing how to calculate the descriptive statistics for subsets of data is a practice of data science because it is something that data scientists do regularly. The goal of this question was for students to engage in CA by doing something in lab that data scientists in the workforce do. However, because this was repetitive, this goal was met in another part of the lab. So, although I believe this was a good question, when looking at what to cut to save time, this made the most sense since the coding and learning outcomes were repetitive. This question was deleted because the students already got practice finding descriptive statistics earlier in the lab. I also deleted the IR question associated with these coding questions since I deleted the coding questions. I also noticed that in Lab A', multiple groups looked at comparing salaries in different departments when answering one of the scaffolds that asked students to think about a question they had about the dataset and then use Python to answer it. There was also a question in Lab A that asked students to look at their home department and share their results with their groups. Overall, by deleting the question mentioned above, this shortened the lab, and the students were not missing out on these coding exercises, ideas, and concepts. Also, Lab A' no longer had repetitive questions.

Problem 3: The order of the IR and GD questions affected how well the students answered them. I noticed two major placement issues that affected the responses of the GD and IR questions when looking at sequencing. First, I saw from the audio recordings that when all of the GD and IR questions were at the end of the lab, students left the lab before completing them. This could be for multiple reasons such as that they ran out of time or they wanted to be done with the lab and skip the discussion. Next, I saw from the audio recordings that IR questions directly before GD questions did not elicit as good of a group discussion as the GD questions

directly before IR questions. My original hypothesis was that having the IR questions directly before the GD questions would help the students would gather their thoughts and write them out first so that then they would be ready and more prepared for a group discussion. However, students were more likely to skip the group discussion or not have a thorough discussion when this was the case.

Next, I give an example of scaffolds with each placement issue. For the first placement issue, I originally put a group of two GD and two IR questions at the end of Lab A. Due to the lab being too long, I quickly saw that no students got to these questions. In general, for Lab A, I put most of the coding questions at the start of the lab and most of the GD and IR questions towards the end of the lab. My idea was that the latter part of the lab would be for reflecting and discussing the coding that they did at the beginning of the lab. For the second placement issue, in Lab A, I put many of the IR questions right before the GD questions. My original hypothesis was that if the students wrote about the IR questions, this would help jump start their ideas for group discussion and that they would have a more productive discussion. In Lab B, I put many of the GD questions right before the IR questions. Contrary to my original hypothesis, this seemed to elicit both better group discussions and more thorough written responses to the IR questions. They seemed to have more ideas for the IR questions that many times were connected to the discussions they had with their groups.

A large part of engaging in CA, communication, and social justice is for the students to collaborate and share knowledge through the group discussion questions. If students are not having thorough discussions, this violates both the Shared Knowledge Principle and the Collaboration Principle. If the students are not doing the discussion questions, they are not collaborating or sharing knowledge about the social justice issues in the labs. By using an

intentional ordering of questions to elicit the most thorough responses, solving this problem helped the students engage in each of the goals for the lab. They were engaging in CA because they are communicating with each other about social justice issues, thinking about the implications of the data science that they are doing, and reflecting on their perspectives and the perspectives of their group members.

I implemented two strategies to address the placement problems that were both modifications to the labs. First, I wanted to modify Lab A' so that it did not have all of the coding questions at the beginning and all of the GD and IR questions at the end. I did this by attempting to put each type of question (coding, IR, and GD) throughout the whole lab. In other words, I had all three types of questions at the beginning, middle, and end of the lab. I did this for Lab B and Lab B' as well. Most of the other labs in the class consisted mainly of coding questions, so when I originally only had the coding questions at the beginning of the labs, the students may not have realized that the IR and GD questions were an important part of the lab. By starting each lab with a GD or IR question, this lets the students know that these were an important part of the lab and that they should expect them throughout the entire lab.

To address the issue of ordering GD and IR questions, I first wanted to confirm that the GD questions before the IR questions elicited better responses for both. I used this ordering in Lab A', Lab B, and Lab B.' In other words, I put some GD questions directly before IR questions in Lab A' (and some IR questions directly before GD questions for comparison) and did this throughout all of Lab B and B'. I saw the same pattern consistently: students had better discussions when the GD questions were first. By better discussions, I mean that the students were a lot less likely to skip the GD questions completely and they had longer back and forth conversations about social justice issues. All group members were also more likely to contribute

to the conversation. I hypothesize that the reason for better discussions when the GD questions are first is because the group discussion allows them to brainstorm together by collaborating and sharing knowledge. This gives them more ideas for their individual written reflections. Overall, the placement of questions can affect student responses and using the multiple iterations of DBR allowed me to confirm the idea I saw in the first iteration.

The last three problems are all related to each other. They are all centered around the idea that some questions did not get thoughtful or meaningful discussions or reflections. Although the problems are similar, the reasons why the problems occurred are different, so I decided to highlight them as three separate sections.

Problem 4: Questions that were categorized as being both group discussion (GD) questions and individual reflection (IR) questions did not get discussion from most groups. Originally, I thought it would be a good idea to categorize some of the questions as both group discussion and individual reflection so that they could discuss the prompt with their group members and be able to write about it. However, when I had questions as both IR and GD, students only wrote about them and did not discuss them. The data that I used to determine this problem came from the audio recordings of the group discussions in the labs. All of the questions were categorized as either coding questions, group discussion, or individual reflection questions. The students could see what each question was categorized as so that they knew how to answer it. There was one question in Lab A that I categorized as both GD and IR, meaning that they were supposed to write a reflection about the question and discuss it with their groups. The question was at the beginning of Lab A and it said: "Thinking about the variables in the dataset, what are two questions that you'd like to use data science to answer (ex. What is the mean salary of my home department?). Type them below and then share at least one of your questions with

your group and why you want to find the answer to it." Only 2 out of 6 of the groups spent time talking about the question out loud. The other 4 groups answered the question in writing but did not discuss it with their group members.

Because this question was supposed to be discussed within the groups and most students did not do this, this question violated the Collaboration Principle and the Shared Knowledge Principle. Most groups did not talk to each other or share any knowledge when answering this question. In order to address this issue, I ended up deleting this scaffold altogether because it did not elicit good discussion and was a bit confusing as to what it was asking. It was also repetitive since there was a scaffold towards the end of the lab that had the students think about a question they had, then use Python to answer it, and then discuss it with their group members. I also realized that it was probably confusing because this question had two labels. Students may have thought that they could choose to either write about it or discuss it. There was also no credit attached to the group discussion, and the lab was already too long. Any of these could be reasons why the students chose not to discuss this question. After Lab A, I only labeled each question as either coding, GD, or IR to make it clearer. In other words, each question had one specific category.

Problem 5: Some group discussion questions did not elicit any meaningful discussion among students. Similar to the last problem, the GD questions that had this problem were questions where the students either skipped the questions completely or just shared their answers and did not have a discussion. I discovered this from listening to the audio recordings of the group discussions in the labs. Also, two students from Fall 2021 mentioned that the wording of the discussion questions was confusing. This was in response to us asking what they did not like about the labs. This could be another explanation of why students skipped certain GD

questions. One example of a GD question with this problem asked students to:

> Discuss with your group whether you think a histogram or a boxplot or both best visualize the salary data. There is no right answer to this question. Explain why histograms, boxplots, or both are important and what they can tell us about the data. Why is it valuable to look at visual displays of salary data in general (as opposed to just looking at descriptive statistics like the mean and SD)?

This question did not get good discussion from the groups. Of the six groups that were recorded in F21, none of them discussed this question in detail or discussed all parts to the question. Two out of the six groups briefly discussed the first part of the question. The remaining four groups did not have a discussion at all. Instead they just had one student say that they thought either histograms or boxplots are better, but nothing else.

Table 9 shows the conversations that the two groups had who briefly discussed the first part of the question. Both of these groups did discuss the first part of the question and explained why they thought either a histogram or a boxplot was better, but afterwards, they moved on to the next question.

Table 9

Group 1 and Group 2 Conversations

Group 1 Conversation	Group 2 Conversation
Student 1: It looks like that, but I don't know if this is right. Student 2: The histogram would probably best explain the data because of how it's skewed. Student 1: I thought the same thing. The histogram visualizes better because you can see the shape and the skewness.	Student 1: What do you say for the group discussion? Student 2: I think the boxplot is better because you can see the range better. Student 3: Yes, there are also a lot of outliers, and you can see them in the boxplot.

These were very quick discussions where their points were not thoroughly explained, elaborated on, and they did not answer all of the parts of the question. The GD questions were supposed to help students with communication and CA. Some of the specific GD questions were also intended to help students engage in social justice awareness through knowledge sharing. However, since most groups did not discuss these questions or did not have meaningful discussion or communication, this scaffold violated the Collaboration Principle and Shared Knowledge Principle. This scaffold also did not help most students engage in CA.

I used two strategies to fix this problem. First, I modified the question shown in the example. I consolidated it into a single question that highlighted the most important part of the question that no one discussed previously. The most important part was the last part that asked why it is important to look at both visual displays of data and descriptive statistics in an analysis. The new modified question asked: "Discuss why it is valuable to look at visual displays of salary data in addition to descriptive statistics like the mean, median, and SD." This is now a single question with clear instructions, rather than a confusing multi-part question. It seemed like some of the questions had too many parts, so it was confusing to the students on whether they had to answer all of the parts or just some of them. In the following labs, I tried to modify the multi-part questions to become single questions with straightforward instructions. Also, for this particular question, the importance of visual displays of data was discussed in a later question so I did not include this scaffold in Lab A'.

Next, I changed some of the GD questions that did not elicit good group discussion to include an incentive for them to talk. I did this by adding IR questions that have them reflect on their group discussions. I think one of the problems with the original labs in Fall 2021 was that the group discussion questions had no points attached to them. If I was not audio recording the

labs, there would be no way to know if students had the discussions because there were no questions that asked them to write about what they said in their group discussions. Since multimodal communication was a goal in the labs, in Lab A', Lab B, and Lab B', I added individual reflection questions that asked students to write about something from their group discussion. This encouraged students to engage in the group discussion since they could not answer these IR questions if they did not partake in the discussion.

One example of a new IR question that was added comes from Lab A'. In Lab A' I added an IR question that said: "Write down something you learned from your group discussion! Write at least 3 sentences." This connected the GD and IR questions and helped the students engage in CA by reflecting on the coding they did and the discussions they had with their peers. A second example comes from Lab B. In Lab B, I added an IR question that asks students to: "Write down the most interesting part of your group discussion or the part you were most surprised by. Write at least 3 sentences." This question also requires them to reflect on their group discussions and allows them to engage in CA for the same reason as the previous question.

Another example of an IR question that was added to Lab A' was a question that asked the students to write about what they did for a prior coding question. The prior coding question asked them to use Python to answer a real question that they specifically had about the salary dataset. This gave students practice coding in Python to answer questions, then communicating their results in both writing in the IR questions and through talking in the GD questions. This again is a practice of data science that helps the students engage in CA. Overall, these IR questions cannot be answered without doing the group discussions. By reflecting on their group discussions, students must collaborate and share knowledge during those discussions in order to be able to reflect. I learned from this problem that GD questions should be clear and there should

be IR questions directly following them that incentivize students to participate in the group discussions.

Problem 6: Some individual reflection questions did not elicit thoughtful responses from students. Similar to some of the GD questions that I discussed in the previous section, there were some IR questions that did not elicit thoughtful responses from students. Some of these questions covered important topics that I wanted students to think critically about. The data that showed this was a problem was the labs and the students' answers to the individual reflection questions. For some of the IR questions, students wrote one-word answers, very short answers, or answers that were not complete sentences. One example of this comes from an IR question in Lab A that asked the students to: "Write a few sentences answering the following questions: Is the mean or median larger? Why do you think this might be the case? What does the standard deviation tell you in this context?" This question did not elicit thoughtful written responses. There were 18 students who gave consent for us to analyze their data in Fall 2021. Of those 18 students, 11 of them only answered the first part of the question which asked: "Is the mean or median larger?" For those that answered this part of the question, they all gave answers such as just "mean" or "mean is larger." Three students skipped this question completely.

This question was intended to review a prior concept (descriptive statistics) and give the students practice writing about the data science that they did, which is a part of computational action. I knew that I wanted to end the labs with longer IR questions that would help the students engage in CA by communicating their thoughts in writing and thinking about the big picture takeaways regarding the social justice issues in the labs. So to address this problem, I used deletions, modifications, and additions. For the deletions, I ended up deleting the question from the example. I wanted the students to think deeply about the IR questions so if the questions did

not elicit thorough responses, I did not prioritize keeping them in the lab. Also, the concept that this question was getting at (descriptive statistics) was not one of the main concepts from this lab. It was a review and since the lab was too long, I ended up choosing to delete it.

Next, I made some modifications. In order to help the students understand what kind of written responses I expected from the individual reflection questions, I modified some of the questions to include the directions: "Write at least 3 complete sentences." I included this in Lab A' and Lab B' so that the students had guidance and clearer directions with how to respond to the IR questions. Based on the students' responses to the IR questions, they did not understand that I expected them to write in complete sentences and to write more than one sentence. Also, one student mentioned in their interview that they weren't used to writing in math classes. They said: "I've never had to write anything in a math class, so I wasn't sure what was expected for the reflection questions. I'd suggest being specific or giving an example." This was more evidence that prompted these modifications. As a result, in Labs A' and B', there were only three students who did not follow the directions of writing at least three sentences for the IR questions. This was three students out of the 21 students who gave consent in Spring 2022.

Lastly, I made some important additions to the labs to address this problem. As previously stated, I wanted the labs to contain at least one scaffold that had the students write a longer individual reflection so that they could reflect on the big picture takeaways from the labs and the social justice issues. In Lab A', I had the students write a paragraph style response to the last IR question. Here is how I asked the question: "Write a paragraph style response (at least 5 sentence) summarizing what you learned from working with the salary data. We have listed a few questions below to give you some ideas on what to write about if you need them." Because this was the first paragraph style response, I listed four prompts that they could use to start

thinking about what to write. By having them write more than prior IR questions, this allowed them to think more critically about their answers and get practice communicating their thoughts in writing about their takeaways from the labs, hence engaging in CA.

In Lab B, I did the same thing except this time, I put the students in the position of an expert answering the paragraph style question. I gave them two options and they could pick which one they wanted to write about. Thinking as an expert allowed the students to engage in CA and social justice awareness because they were acting as though they were a real data scientist examining an important, real-world issue. This paragraph style question allowed them to use the knowledge they gained inside the classroom and apply it to a real-world situation outside of the classroom. By acting as an expert, they are getting practice engaging in data science practices. Because this question elicited great responses, I included it in Lab B' as well. Overall, addressing this problem improved the quality of the IR questions and the quality of responses received from students. By answering questions in writing and thinking critically about the analysis that they are doing, students are engaging in CA since this is an important part of data science.

Additional Problems. When talking about problems and the methods I used to solve them, it is also important to point out that there could be some potential underlying problems that either cannot be solved or need further study to determine if they are problems. Problems that cannot be solved have to do with the typical organizational issues in higher education courses. The labs are run by graduate TAs and most labs have 30 students enrolled. In other words, there are no TAs or instructors in the current organization of the labs who can mediate conflict since the instructors are not present in the labs and the TAs have 30 students to attend to. Also, when students are not used to talking about social justice issues, oftentimes they feel uncomfortable

engaging in these discussions with their peers, especially if they are people they do not know. It would be great if the TAs or course staff could help guide these discussions and encourage knowledge sharing and collaboration, but due to the nature of this labs, this is not possible. The students also have limited time to finish the coding portion of the labs as well as have meaningful discussions. It is unknown whether or not students skipped discussion questions because they felt rushed or felt like they did not have enough time. There could be other reasons for this as well, but to determine the reasons, further study is needed.

Summary of Problems and Adaptations to the Labs. The six problems that I described and the adaptations I made to the labs to attempt to solve these problems allowed me to document important ideas to remember when designing data science labs. First, it is important to consider the length of time students have to complete the lab. If the labs are too long, the students will run out of time and be forced to skip certain questions. The labs should give students plenty of time to finish the coding sections and have meaningful discussions with their groups. When thinking about length, it is also important to create questions that specifically help you accomplish the goals of the labs, keeping in mind that questions should not be too repetitive. Each question should be mapped to a specific goal.

It is also important to place each type of question throughout the entire lab. For example, I had coding questions, GD, and IR questions. These should all be seen at the start of the lab, the middle of the lab, and the end of the lab to let students know that they are all equally important. I also noticed that putting GD questions directly before IR questions motivates the students to engage in their group discussions so that they are able to answer the IR questions. Questions should also be clearly labeled as one category (coding, GD, or IR), not both. Overall, it is important to be clear about directions for each question. For IR questions, include directions that

tell students to write in complete sentences and the minimum number of sentences you would like them to write. For GD questions, try to make them as concise as possible since multi-part questions can be confusing and oftentimes students will only pick one part to answer. Lastly, it is important to make sure that throughout the lab, students are engaging in the practice of data science. This includes giving them examples to practice acting as a real data scientist in all types of questions. Table 10 shows each of the problems and the adaptations used to help solve them.

Table 10

Summary of Problems and Strategies to Solve Them

Problem	Additions	Modifications	Deletions
1. Lab A was too long.		x	x
2. Some questions were repetitive.			x
3. The order of the IR and GD questions affected how well students answered them.		x	
4. Questions that were identified as being both GD and IR did not get discussion from most groups.			x
5. Some group discussion questions did not elicit any meaningful discussion among students.	x	x	
6. Some individual reflection questions did not elicit thoughtful responses from students.	x	x	x

Overall, the DBR process allowed me to support the students as they engaged in CA and social justice awareness by implementing these labs, getting feedback from the students to make these improvements and adaptations, and implementing them again in the second iteration. The additions to the labs helped students thoroughly dive into each question and added more elements of multimodal communication, such as written individual reflection from the position of an expert. The deletions from the labs helped make them an appropriate length. The modifications to the labs improved the quality of the scaffolds and helped make them clearer for the students. All three of these adaptations combined helped improve the labs for the students. Looking at how the labs evolved gives an example of how DBR can be used to design labs where

the goal is for students to engage in CA, explore social justice issues, and practice multimodal communication. In the next section, I answer the second research question which looked at what happened in the labs and how the students met the goals.

RQ2 Results

RQ2a and RQ2b asked: "What evidence is there that the scaffolds and artifacts help the students engage in computational action (CA)?" and "How do students apply data science practices to question the status quo and consider social justice issues?" In order to answer both of these questions, I looked at the group discussions and the individual reflection questions from the labs. Because these two questions were both related to the two goals of the labs and I used the same analysis methods for both of them, I answered them together instead of separately.

Overview of Group Discussion Question Analysis

For the group discussion questions, the unit of analysis is each group, and I used the Toulmin Argumentation Pattern for analysis. I found evidence that the students engaged in data science practices during the labs by using the representations that they created from the scaffolds and data from the artifacts as warrants to their claims. By participating in group discussions about data science, the students engaged CA because part of engaging CA is doing authentic work and acting as real data scientists. Discussing the analysis that they did is something that data scientists typically do in their jobs. Also, the scaffolds encouraged them to think about how their work in class can be useful outside of the classroom. The students also engaged in CA by working with real data and using real-world examples. Some of the artifacts such as the datasets and real-world examples in the lab were centered around social justice issues. The students discussed these issues while thinking about their implications during their group discussions and

individual reflections. In other words, the students engaged in social justice awareness through working with these artifacts and scaffolds in the labs.

Table 11 shows the three group discussion questions the students answered in Lab A. The first question asked students why it is important to look at visual displays of data in addition to descriptive statistics. The benefits of looking at both visual displays of data and descriptive statistics in an analysis is an important data science concept and understanding it helped the students engage in CA. The second question had the students share the results of an individual analysis that they did on their own departments with their group members. Sharing analysis with others is an important data science practice that aimed to help the students engage in computational identity, which is a part of CA. The third question expected the students act as real data scientists by picking a question that they want to answer and using what they know to answer it. Then they were instructed to share what they did and what they found with their group. This question also was intended to help them engage in CA by using data science to answer questions that they have and then share them with their group. Many of the students talked about the social justice issues in their discussions so through answering these questions, the groups were expected to engage in CA and social justice awareness.

Table 11

Group Discussion Questions for Lab A

Group Discussion Questions	
GD Question 1	Explain why it is valuable to look at visual displays of salary data in addition to descriptive statistics. Share what you said with your group members.
GD Question 2	Share the results of the departmental analysis with the people at your table. Do you notice any similarities or differences between departments?
GD Question 3	Pick a question you want to answer. Share the question and results with your group.

Three groups answered these discussion questions from Lab A. Table 12 shows the details about these three groups, including the number of students, the pseudonyms of the students, what semester they completed Lab A, and how they did according to the rubric that evaluated how they engaged in CA and social justice awareness.

Table 12

Groups Who Completed Lab A

Group	Number of Students	Names	Semester	Computational Action	Social Justice Awareness
Group 1	3	Jordan, Logan, Jayden	Spring 2022	Sufficient	Sufficient
Group 2	3	Bailey, Alex, Casey	Spring 2022	Above and Beyond	Sufficient
Group 3	2	Marley, Rowen	Spring 2022	Above and Beyond	Needs Improvement

Next, I describe what I found from each of the three groups who completed Lab A and each of the six groups who completed Lab B. I separated this analysis by lab because it is possible that how well the groups did could change depending on the lab since the labs covered different topics. For each group, I summarized their discussions and talked about their scores according to rubric that looked at engaging in CA and social justice awareness. I also gave examples of how the Toulmin analysis showed that they engaged in CA and social justice awareness. Lastly, I described the characteristics of a typical group that completed each lab by looking at the similarities that I noticed between the groups.

Toulmin Analysis for Lab A Group Discussions

Group 1 Summary. Group 1 scored in the "sufficient" category for both CA and social justice awareness according to the rubric that evaluated how well each group did with these two goals. In other words, there was evidence that the students in this group engaged in CA and

social justice awareness but did not go above and beyond in their discussions. Even though Group 1 answered the questions and demonstrated evidence of using what they created to justify claims, their discussions of social justice issues could have been further developed. Group 1 mentioned the salary discrepancies between men and women, however they did not elaborate on the implications of this. One of the most interesting parts of their group discussion was when one group member, Logan, talked about how what they did for analysis prompted them to want more data and to want to do more analysis. Logan said: "I also want to look at how long they've been teaching, like a full professor teaches for longer than an assistant professor so how much does your salary go up for each year you teach." Here, Logan was saying that they would like to have more data that says what each person's title is and look at how the salaries compare between the titles. Although Logan said they would like to do this, they did not actually do this. The scaffold encouraged Logan to think about collecting more data and doing more analysis as well as share these ideas with the group, which is one example of how the scaffold helped this group engage in CA.

Computational Action. There was evidence that Group 1 engaged in CA from answering GD Question 2, which had the students share their departmental analysis with the group. Figure 8 shows the Toulmin Diagram showing Group 1's collective argument. Table 13 also shows the conversation that they had when answering GD Question 2. In each table, I labeled the turns for each excerpt from "1" to "n."

Table 13

Group 1's Conversation for GD Question 2

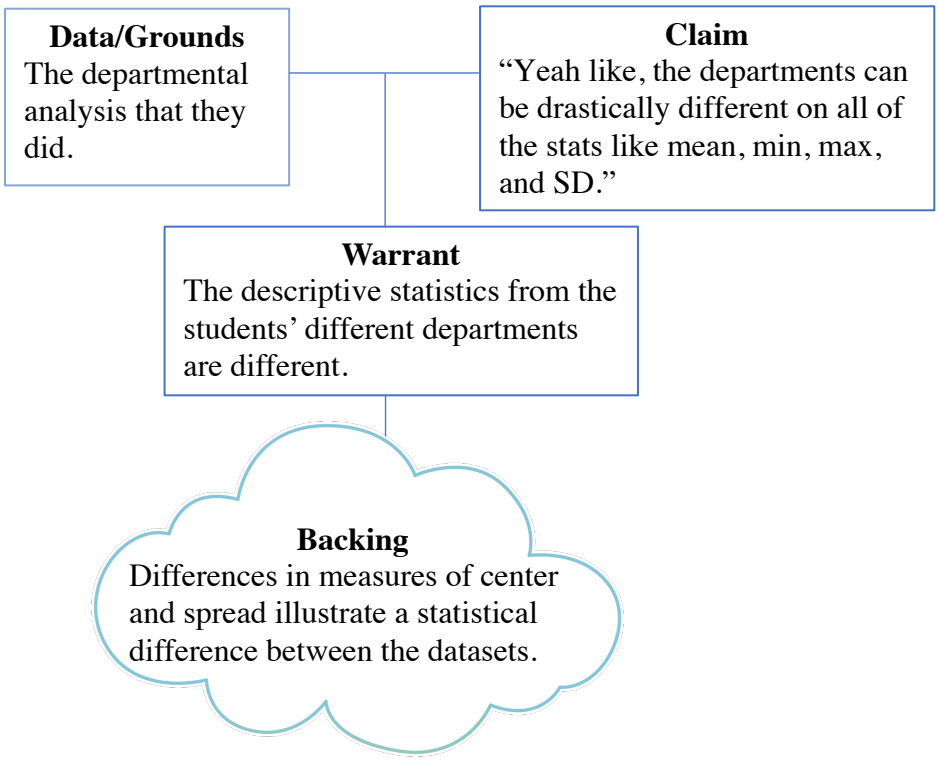
Turn Number	Speaker	Turn
1	Logan	Then we have to compare them, so 210,000 was my max
2	Jayden	Thirty-three thousand was my lowest and 180,000 is my highest
3	Logan	That's unexpected, it's such a big difference.

Table 13 (cont.)

4	Jayden	Yeah, because that's stats and you'd think a mathematician would make more.
5	Jordan	Okay so my department was also different from yours both. I think there are a lot of differences.
6	Logan	Yeah like, the departments can be drastically different on all of the stats like mean, min, max, and SD.

Figure 8

Toulmin Diagram for Group 1, GD Question 2



Group 1 claimed that there were a lot of differences between departments. They specifically mentioned differences between the descriptive statistics (mean, minimum, maximum, and standard deviation) in Turn 6 from Table 13. The grounds were the departmental analysis they did from the scaffold. The scaffold had the students find the minimum and maximum salaries in their departments and then share their results with their group members. The warrant was that the descriptive statistics were different for each group member. For

example, when Logan and Jayden both gave the maximum salaries for their departments in Turns 1 and 2 from Table 13. The backing was the statistical principle that says that differences in measures of center and spread illustrate a difference among the different datasets for each students' department.

I found that GD Question 2 helped Group 1 engage in CA because each student created a subset of the original dataset containing only the data for their department. They conducted their own analysis on their specific dataset and shared their perspective and results with the group. Each student in this group contributed to the claim that the departments had different descriptive statistics. Logan and Jayden shared their minimum and maximum salaries and Logan mentioned that the difference was large and unexpected (Turn 3, Table 13). Jordan also mentioned that there were a lot of differences (Turn 4, Table 13) and Logan said that the departments all had different descriptive statistics (Turn 6, Table 13). Students' interactions within Group 1 illustrates a case of distributed cognition because the knowledge was shared among the group members. Because each student contributed something unique to the group to help justify the claim, there is evidence that the students engaged in computational identity. Computational identity is a specific part of CA that focuses on students working together and being an important part of a group. The students also engaged in digital empowerment (the other part of CA) because they used real data to do their own analysis and then shared it with their group members. Data scientists often do their own analysis and then have to share the results with others, just like the students did here.

Social Justice Awareness. In addition to engaging in CA, students in Group 1 also engaged in social justice awareness by answering GD Question 3. GD Question 3 asked them to choose a question they want to answer and share it with their group. They discussed the main social justice issue in this lab which was the salary discrepancies between men and women.

Figure 9 shows the Toulmin Diagram of Group 1's collective argument to this question. Table 14 shows the conversation they had when answering GD Question 3.

Figure 9

Toulmin Diagram for Group 1, GD Question 3

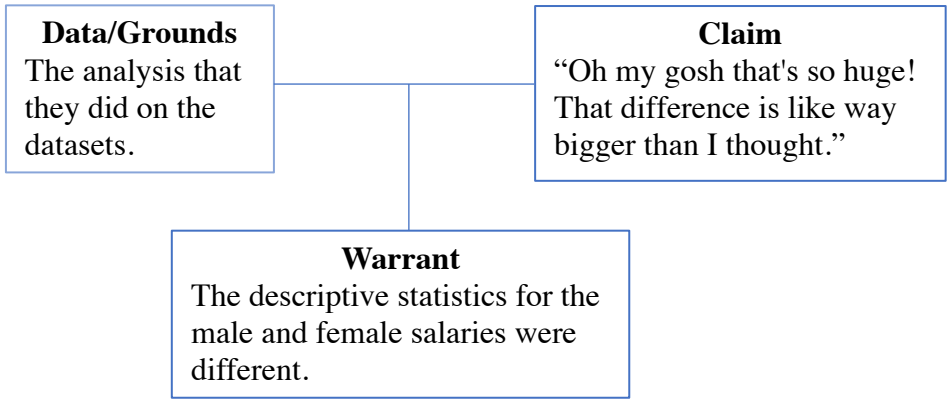


Table 14

Group 1's Conversation for GD Question 3

Turn Number	Speaker	Turn
1	Jordan	I wanted to see if the highest salary in the gender dataset was male or female and it was male.
2	Jayden	What were they?
3	Jordan	Males were just under $130,000 and females were almost $100,000
4	Jayden	Oh my gosh that's so huge! That difference is like way bigger than I thought.
5	Jordan	I wasn't really surprised the males were higher, but I didn't think it would be like $30,000 more. How about you?
6	Logan	I also want to look at how long they've been teaching, like a full professor teaches for longer than an assistant professor so how much does your salary go up for each year you teach.

In Table 14, Jordan wanted to see if male employees had a higher salary than female employees in general. In Turn 3, Jordan shared what they found by saying that the male average was just under $130,000 and the female average was just under $100,000. Logan and Jayden

reacted to these findings in Turns 4 and 6 from Table 14. Jayden thought that the difference was huge, and Logan mentioned that they were not surprised that males earned more. Although Logan said that they were not surprised that the males earned more, they did mention that difference between men and women salaries was way larger than they expected. This group claimed that males made a lot more than females and used the data as grounds. The warrant was that the descriptive statistics for the men and women were different. There was no evidence of a backing in this case. Group 1 did analysis in Python and then used that analysis to justify their claim. Their claim was about the salary discrepancy between men and women. GD Question 3 helped Group 1 use data science to show their claims and this work helped them become aware of the social justice issues in the lab, hence engaging in social justice awareness.

Group 2 Summary. Group 2 scored in the "above and beyond" category for CA and the "sufficient" category for social justice awareness according to the rubric that evaluated how well each group did with these two goals. In other words, there was evidence that this group engaged in CA and went above and beyond the expectations that I had for them. They did this by doing more analysis than the scaffold asked them to do and discussing the additional analysis. The work that they did from the scaffold made two students, Alex and Casey, curious about other questions in addition to the questions asked by the scaffold. They brought this up during the group discussion and the other student, Bailey, reacted to it. This group also engaged in social justice awareness; however, they did not elaborate on the implications of the social justice issue in the lab during their group discussions, which is why they were not in the "above and beyond" category for this goal.

Computational Action. There was evidence that Group 2 engaged in CA by answering GD Question 2, which asks the students to share their departmental analysis with the group.

Figure 10 shows the Toulmin Diagram of Group 2's collective argument. Table 15 also shows the conversation that they had when answering GD Question 2.

Figure 10

Toulmin Diagram for Group 2, GD Question 2

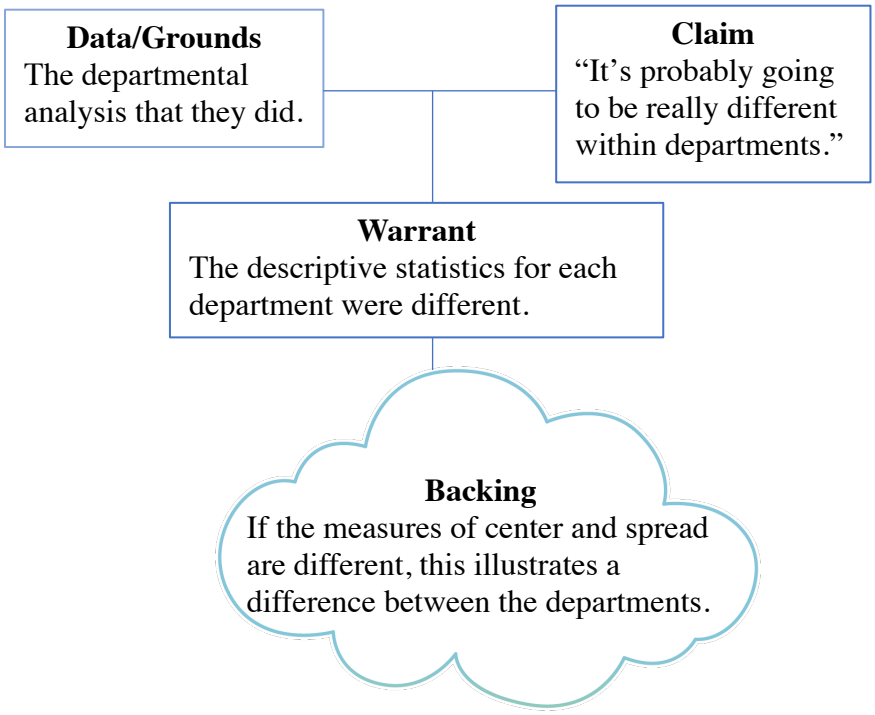


Table 15

Group 2's Conversation for GD Question 2

Turn Number	Speaker	Turn
1	Alex	I got very surprising results for my min and max.
2	Bailey	I'm not as surprised about my min and max results.
3	Casey	What does your histogram look like?
4	Alex	It's probably going to be really different within departments. Mine has gaps but I think that's just the difference between lecturers and professors.
5	Bailey	What was your min and max?
6	Alex	My max was 215,000 but my min was only 9,000 which is weird because mine is the math department.
7	Bailey	That's weird are they counting TAs?
8	Alex	They might be.
9	Casey	My min is 50,000 and my max is 366,000

Table 15 (cont.)

10	Bailey	My max is 260,000 and my min is 40,000
11	Casey	I also looked at some other departments too because I was curious. Accountancy seems similar to business, but most are different.
12	Bailey	They are both Business so that makes sense.
13	Alex	I looked at computer science and the salaries are so high. The average is over 120. Like compared to the rest of the university, CS (computer science) pays so much more.
14	Bailey	That's crazy! CS is so good.

Like Group 1, Group 2 claimed that there were a lot of differences between departments. The grounds were the departmental analysis they did where each student in the group found the minimum and maximum salary in their own department. The warrant was that the descriptive statistics were different among departments. For example, Bailey asked Alex and Casey what they had found for the minimum and maximum salaries in their departments in Turn 5 from Table 15. Casey and Alex responded with their numbers in Turns 6 and 8 from Table 15. Bailey also gave their minimum and maximum salaries in Turn 10 from Table 15. The backing was that differences in descriptive statistics imply that there is a difference between the departments. Similar to Group 1, Group 2 engaged in computational identity because they collectively answered this question by each doing analysis, sharing it with the group, and reacting to their group members' analysis. Each group member played an important role in the discussions and shared something unique. Group 2 also engaged in digital empowerment because they used real data to do their own analysis and then share it with their group. The reason this group scored "Above and Beyond" rather than "Sufficient" for CA is because while answering GD Question 2, they ended up doing more analysis than what was required.

GD Question 2 motivated two of the members of Group 2 to continue exploring this question with departments they were interested in outside of their majors. Unlike Group 1, Group 2 was able to do some of this extra analysis and discuss it. They looked at different majors in the

College of Business and the Computer Science major to give even more justifications to their claim. In Turn 11 from Table 15, Casey mentioned being curious about other departments and how they looked at the Accountancy Department because it was similar to their department. In Turn 13 from Table 15, Alex described how they did some analysis on the Computer Science Department and noticed how high the salaries were. This extra work that Alex and Casey did showed that this group went above and beyond what the scaffold asked them to do. This showed more evidence that the scaffold helped them engage in CA by having the agency to pursue this additional investigation.

Social Justice Awareness. In addition to engaging in CA, there was also evidence that the scaffolds and artifacts helped Group 2 engage in social justice awareness. One example of this was the claim that they made when answering GD Question 3, which had them pick a question they want to answer and share it with their group. Group 2 discussed the main social justice issue in this lab which was the salary discrepancies between men and women. Figure 11 shows the Toulmin Diagram of Group 2's collective argument to this question. Table 16 also shows the conversation that they had when answering GD Question 3.

Figure 11

Toulmin Diagram for Group 2, GD Question 3

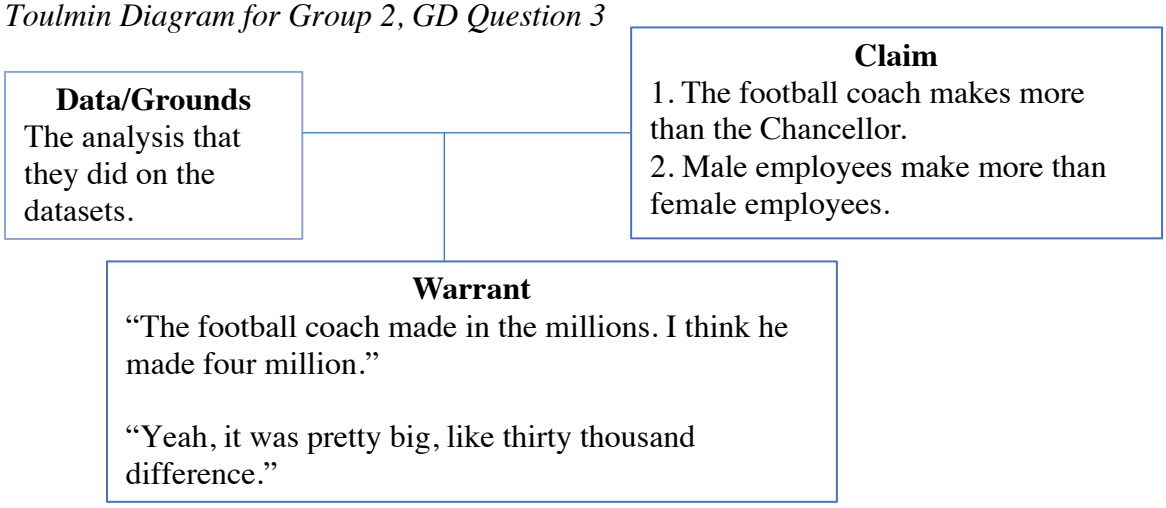


Table 16

Group 2's Conversation for GD Question 3

Turn Number	Speaker	Turn
1	Bailey	My question was does anyone make more the Chancellor and some people in sports did. The football coach made in the millions. I think he made four million.
2	Alex	My two questions were what if we only looked at data where the faculty was professor and my other was what primary college has the highest variance and then I answered the first one and calculated the mean salary of professors.
3	Casey	Mine was really simple. Is there a difference between the mean of the male and female employees?
4	Bailey	I'm guessing there was.
5	Casey	Yeah, it was pretty big, like thirty thousand difference.
6	Alex	Wow yeah that's not okay.

For GD Question 3, each student picked a unique question that they wanted to answer. Bailey wanted to know whether the Chancellor made more than employees involved in athletics. Alex wanted to know what department had the highest variance in professor salaries. Casey wanted to know if there was difference in salaries of male and female employees. Group 2 made two claims: 1) The football coach made more than the Chancellor, and 2) Male employees made more than female employees. The grounds for both claims were the analysis that they did. Bailey described the results of their analysis (Table 16, Turn 1) and Casey described the results of their analysis (Table 16, Turn 5). The warrants were the calculations and numbers they shared with their groups that came from their individual analyses. There was no evidence of a backing. Bailey and Casey both gave the results of their analysis, while Alex said what question they answered on their own but did not share their results of that question.

There was evidence that Group 2 engaged in social justice awareness through sharing the results of their analyses. They targeted their analysis to study whether men earned more than women. Then, they used the results of this analysis to justify their claim regarding the social

justice issue in this lab. They did not react to the claims except when Alex said "that's not okay" in Turn 6 from Table 16. This was in reference to men earning $30,000 more than women. By using the scaffold which allowed them to explore something they were interested in using data science and the data they were given, Group 2 engaged in social justice awareness because they did discover that there was difference between male and female salaries. However, they did not gain more than awareness and they did not discuss this issue any further than what Alex said in Turn 6 from Table 13.

Group 3 Summary. Group 3 scored in the "above and beyond" category for CA and the "needs improvement" category for social justice awareness according to the rubric that evaluated how well each group did with these two goals. In other words, there was evidence that this group engaged in CA and went above and beyond the expectations that I had for them. They did more than the scaffolds asked them to do and had back-and-forth discussions about the analysis that they did. They also showed that the group discussions helped them understand important data science concepts. Despite their discussions, they did not mention any social justice issues in their discussions. In other words, they did not mention the salary discrepancies between men and women like the other groups did, which is why they scored in the "needs improvement" category for social justice.

Computational Action. There was evidence that Group 3 engaged in CA by answering GD Question 1, which asked the students to explain why looking at visual displays of data in addition to looking at descriptive statistics is important. Figure 12 shows the Toulmin Diagram showing Group 3's collective argument. Table 17 also shows the conversation that they had when answering GD Question 1.

Table 17

Group 3's Conversation for GD Question 1

Turn Number	Speaker	Turn
1	Rowen	It's definitely good to look at both because otherwise you aren't getting the whole picture.
2	Marley	Yeah, I think with the histogram, you can see the outliers better.
3	Rowen	Yeah. I also said visual displays are easier to comprehend because you can see the data in relation to each other. And like what you said with outliers, boxplots show them easily.
4	Marley	Yeah, with boxplots, that's the first thing I see.

Figure 12

Toulmin Diagram for Group 3, GD Question 1

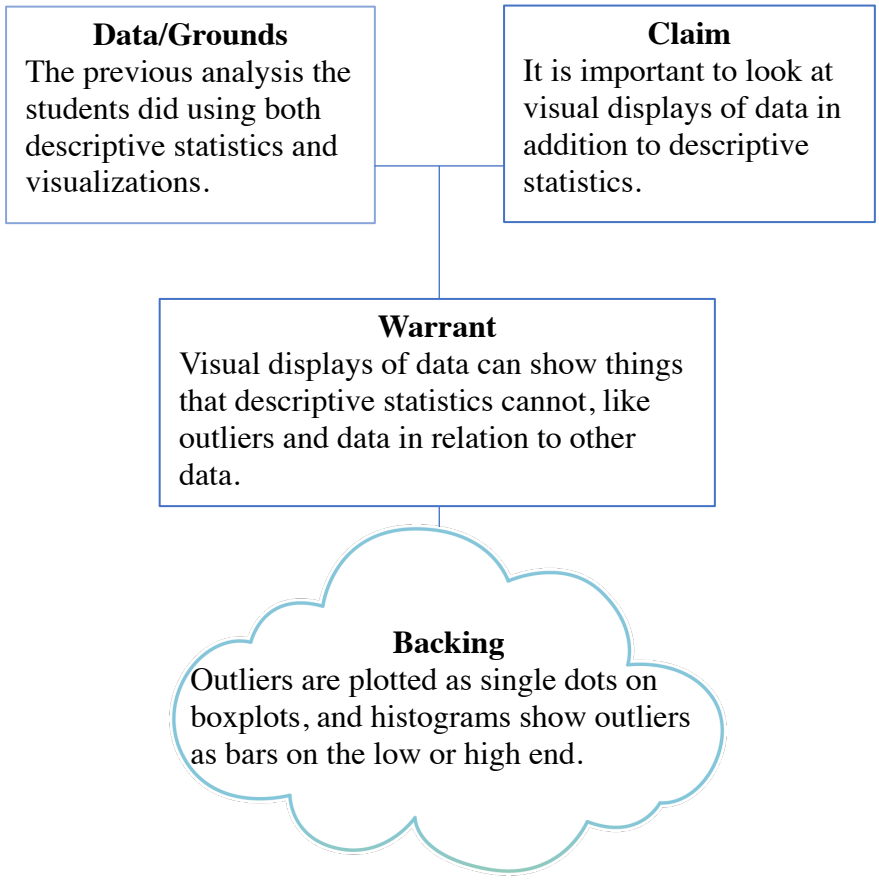
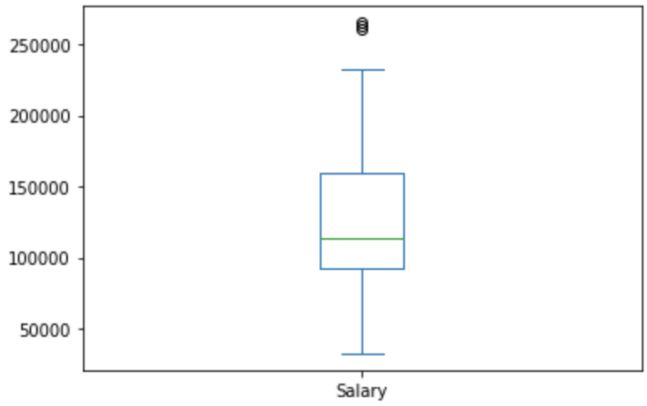


Figure 13

Example of a Boxplot with Outliers Produced by a Student



Group 3 made the claim that it is important to look at visual displays of data in addition to descriptive statistics, which is a key idea in data science. Rowen made this claim in Turn 1 from Table 17 and Marley continued this discussion in Turn 2 from Table 17. The input of these two group members contributed to the elaboration of a collective argument. The grounds were the previous analysis that they did that involved using both descriptive statistics and visual displays of data. This analysis helped them see the importance of looking at visual displays of data and descriptive statistics. The warrant was the discussion they had about how visual displays of data can show things that descriptive statistics cannot. In Turns 2 and 3 from Table 17, they stressed the importance of looking at visual displays of data because there are statistics and data points (outliers) that can be seen more easily through boxplots and histograms. Figure 13 shows an example of a boxplot with outliers (the dots). This is an important idea in data science that they understood due to the coding questions in the lab that asked them to look at both descriptive statistics and visualizations. Looking at both visual displays of data and descriptive statistics is a data science practice. By understanding this data science practice, the students in Group 3 engaged in CA. This idea is something that will be very important in any

data science analysis that they do in the future, both inside and outside the classroom.

There was also evidence that Group 3 engaged in CA from GD Question 2, which asked them to share the results of the analysis of their major. Marley and Rowen both had the same major, and instead of skipping this question completely or briefly discussing it, they used the scaffold to explore another department. They looked at the Statistics Department and the Advertising Department in their analysis that they did together to answer this question. They mentioned how they were surprised that Advertising professors made more on average than Statistics professors. They also hypothesized why this might be the case in their discussion shown in Table 18.

Table 18

Group 3's Conversation for GD Question 2

Turn Number	Speaker	Turn
1	Marley	This is crazy. I would have thought Advertising would be much less but I'm getting Statistics actually is.
2	Rowen	Me too. With stats being math, I thought the stereotype was that math gets paid more.
3	Marley	Yeah, and the stat professors seem to have to do more work than the advertising professors since most classes are basically just projects.

They thought that Statistics professors would earn higher salaries than Advertising professors since advertising classes have a lot of projects and the students work together in groups, as opposed to statistics courses where the instructor lectures for the majority of the time. Here, Marley and Rowen used data science and the artifacts to investigate an issue, discussed it, and think about the implications of the salary disparity. The discussion was one example of how students engaged in CA and this exchange was also one of the reasons they scored "Above and Beyond" on the rubric for CA.

Lastly, there was evidence that Group 3 engaged in CA from GD Question 3. GD Question 3 had students pick a question they wanted to answer and share it with their groups. Marley made the claim that they should have used standard deviation as a measure of spread instead of variance. Figure 14 shows the Toulmin Diagram for this question and Table 19 shows the conversation that Marley and Rowen had about this scaffold.

Figure 14

Toulmin Diagram for Group 3, GD Question 3

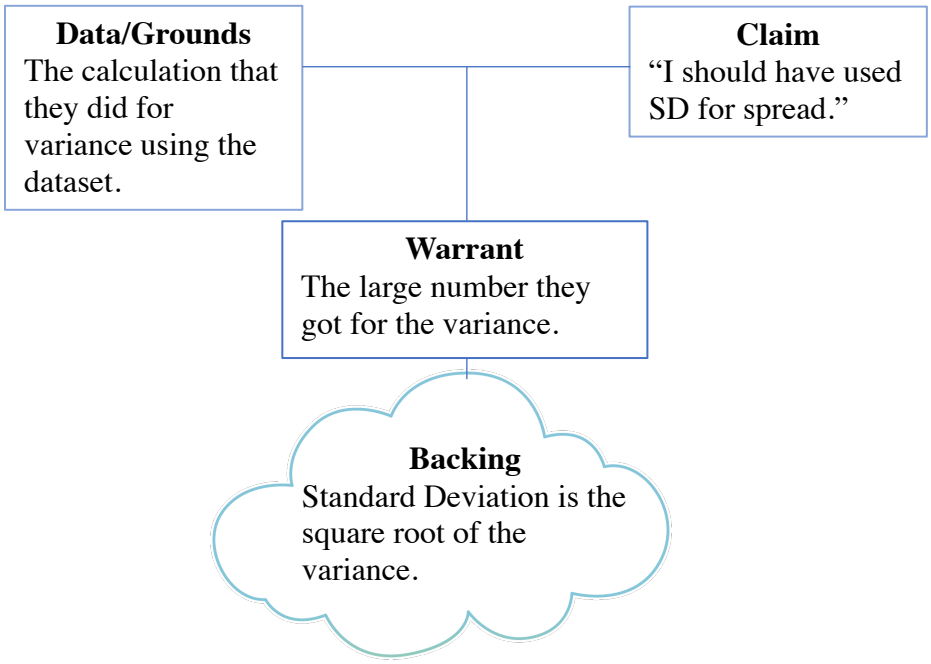


Table 19

Group 3's Conversation for GD Question 3

Turn Number	Speaker	Turn
1	Rowen	Yeah. What about yours?
2	Marley	So, I just got that the largest salary is this and then the variance is this.
3	Rowen	Why is the variance so big?
4	Marley	I think I should have used SD for spread because variance is squared. Then it would make more sense.
5	Rowen	Yeah, since the variance is huge, the SD would be easier to understand since you're not squaring it.

The grounds were the calculations that they did using the dataset, and the warrant was the number they got for the variance. The backing was that the standard deviation is the square root of the variance. The backing is significant because this is a data science principle that this group understood. Group 3 engaged in CA because they did an analysis, shared it by making a claim, and then through the group discussion, realized what they did wrong. I was able to see this from the argument that they made because they claimed that Marley should have used standard deviation instead of variance. They discussed why this was the case and said that the number that Marley got was too big. This prompted them to realize that they calculated the wrong statistic. Marley calculated the variance instead of the standard deviation. The variance is in units squared and is more difficult to interpret because of this. Marley acknowledged this in Turn 4 from Table 19 and Rowen confirmed that the standard deviation (SD) is easier to understand because it is dollars, rather than dollars squared. This exchange illustrated how the knowledge of variance and standard deviation was distributed among the group members. The understanding of variance and standard deviation is an important data science concept.

Social Justice Awareness. Group 3 did not engage in social justice awareness because they did not talk about any social justice issues in their group discussions. Marley and Rowen did not mention the salary difference between men and women at all during their discussions. However, this seems to be due to the nature of the lab, rather than the fault of the group. No group discussion question directly asked them to talk about the salary difference between men and women. Despite this limitation of the lab, the discrepancies came up in the other two groups' discussions. Although Group 3 did not mention the difference between men salary and women salary, they did mention the difference between the departments. They focused their analysis on that instead of the differences between men and women.

Similarities Between Groups Who Completed Lab A

The three groups who completed Lab A had a lot of similarities. The first thing I noticed about the group discussions was that everyone in the groups talked and participated in the conversations about data science. There were not any students who were excluded from the discussions or did not participate. Because the groups discussed data science concepts, social justice issues, and the implications of the analysis that they did, there was evidence that the students engaged in computational identity. Each student was an important part of the group who contributed unique ideas to theses discussion. The students added their unique perspectives to the conversations, shared their knowledge with their group members, and made collective arguments. GD Questions 2 and 3 especially helped the students engage in computational identity and these questions encouraged them to share knowledge. The scaffolds had the students do something unique and then share their results with their groups and explain what they did. These types of group discussion questions also helped the students engage in digital empowerment because they were able to use the artifacts in the labs such as the datasets and Python to answer questions that they had about the data.

All groups used the data analysis that they did as the warrants and grounds for their claims in questions that asked them to do data analysis. The scaffolds provided evidence or data for these claims and the backings are data science ideas and practices involving the topics covered in class. In other words, they used what they created using data science and the artifacts that they had to justify their claims to answer the questions provided by the scaffolds. Using what they created to justify their claims also showed that they engaged in digital empowerment. Because the students used representations that they created to help justify a claim, they engaged in data science practices and acted as real data scientists. Because some of the claims were about

social justice issues and some of their discussions involved looking at how data science can be used in connection with social justice issues, the students engage in social justice awareness.

Although the students did engage in social justice awareness, the social justice discussions were limited ("sufficient" for two groups, "needs improvement" for one group). The two groups that acknowledged the salary discrepancies did mention it briefly but did not reflect its implications or discuss it in detail. The third group did not discuss the issues of salary discrepancies at all. However, I hypothesize that these limited discussions are due to the nature of the lab and the topic of the lab, rather than the students in the groups. The group discussion questions did not require students to discuss this issue. The discussion questions were open ended, and the students could choose what they wanted to discuss. If I do a third iteration of this lab in the future, it would be beneficial to incorporate more social justice issues or ask an explicit question about it for the group discussion questions. Overall, the Toulmin analysis of the group discussion questions showed that there was evidence that the groups used the artifacts (Python, Jupyter, the dataset) and scaffolds to engage in CA and some social justice awareness. Next, I describe how the scaffolds and artifacts in Lab B did the same.

Toulmin Analysis for Lab B Group Discussions

Table 20 shows the five group discussion questions that the students answered in Lab B. The goal of these questions was for the students to engage in CA and social justice awareness. GD Questions 1, 2, and 4 were intended to help the students engage in CA. GD Questions 1 and 4 asked the students to think about the analysis that they were going to perform and the implications of their analysis. GD Question 2 asks the students to interpret the results of a histogram that they created. Using visual displays of data to help explain analysis is an important data science practice. GD Questions 3 and 5 were intended to help the students engage in social

justice awareness. GD Questions 3 and 5 ask students to think about how social justice issues and data science are related. In other words, these scaffolds encouraged the students to discuss how data science can be used to educate people about social justice issues.

Table 20

Group Discussion Questions for Lab B

Group Discussion Questions	
GD Question 1	Because 8% of the eligible population was Black, 3 black people on a panel of 100 might seem low. Does this difference (8% vs. 3%) seem big to you? Do you think this could be due to chance?
GD Question 2	Interpret the results of your histogram. How does your histogram provide evidence for or against the claim that the jury was not fair? What does this tell us about the case? Do you think this could have happened by chance? If so, why? If not, why did some people claim that it did?
GD Question 3	This is an example of how we can use statistics to help us solve real world problems. Discuss with your group how simulations and data science can be used to help address issues of racism specifically.
GD Question 4	When this study was done in real life, the white sounding names had 10.33% callbacks and the black sounding names had 6.87% callbacks. Some of the companies claimed this difference was due to chance. Do you think this is a significant difference? Why do you think this happened? Why or why not is this problematic?
GD Question 5	Discuss with your group whether or not you think similar events still occur today and reflect on how data science can be used to educate people about this.

Six groups answered these discussion questions from Lab B. Table 21 shows the details about these six groups, including the number of students, the names of the students, and what semester they completed Lab B.

Table 21

Groups Who Completed Lab B

Group	Number of Students	Names	Semester	CA	Social Justice Awareness
Group 1	2	Dakota, Harper	Fall 2021	Great	Great
Group 2	2	Sawyer, Kendall	Fall 2021	Needs Improvement	Needs Improvement
Group 3	3	Cameron, Avery, Jude	Fall 2021	Great	Great
Group 4	3	Jordan, Logan, Jayden	Spring 202	Great	Sufficient
Group 5	3	Bailey, Alex, Casey	Spring 2022	Great	Sufficient
Group 6	2	Marley, Rowen	Spring 2022	Great	Sufficient

For each group in Lab B, I summarized their discussions and talked about how they scored according to rubric that looked at engaging in CA and social justice awareness. I also give examples of how the Toulmin analysis showed that they engaged in CA and social justice awareness. Lastly, I describe the characteristics of a typical group that completed each lab by looking at the similarities that I noticed between the groups. Since there were more groups that we analyzed for Lab B than Lab A, I sorted them by how they did with CA and social justice awareness according to the rubric. For each group, I discuss how they did on the rubric that looked at how they engaged in CA and social justice awareness. I also give examples of how the Toulmin analysis showed that they engaged in CA and social justice awareness, and I give a summary of each group. Lastly, I describe the characteristics of a typical group that completed this lab by looking at the similarities between groups.

Groups that were "Great" with both CA and Social Justice Awareness

Group 1 Summary. Group 1 did not complete the first and last group discussion question for Lab B. They ran out of time for the last one and skipped the first one. Despite this,

Group 1 had thorough discussions about the questions that they did discuss. For the trial example, they discussed how the jury selection probably was not due to chance. There was evidence that this group engaged in social justice awareness because they acknowledged that choosing a jury with a non-random method was problematic. They also discussed how they have heard of situations outside of class that are similar to the examples that were brought up in the labs. This is one example of evidence of how they engaged in CA by seeing the connection between the data science they did in the labs and real social justice issues.

Computational Action. One way Group 1 engaged in CA was by answering GD Question 2, which had them interpret the results of their histogram. Figure 15 shows the Toulmin Diagram for this question and Table 22 shows the conversation that Group 1 had about this scaffold.

Figure 15

Toulmin Diagram for Group 1, GD Question 2

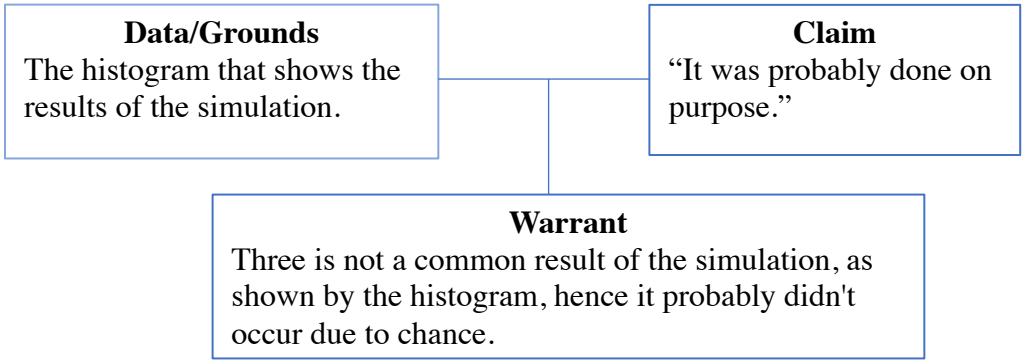


Figure 16

Dakota's Histogram

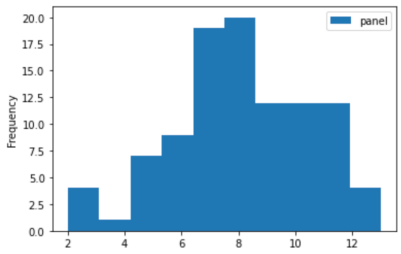


Table 22

Group 1's Conversation for GD Question 2

Turn Number	Speaker	Turn
1	Dakota	So, I guess the histogram proves that the 3% Black jurors is really rare, and that's why it was probably done on purpose. Like there is hardly any time where you would get 3 Black jurors.
2	Harper	Yep, I agree. With the histogram, it seems like it couldn't have been just because. It's showing we should have gotten like 8 or 9.
3	Dakota	That's kind of messed up that this can happen and that like, they cannot do random selection when that's the rule.

Group 1 claimed that the jury was not randomly selected. The grounds were the histogram that they created from the simulation. Dakota's histogram is shown in Figure 16. The warrant was that three was not a common result, meaning it probably did not happen by chance. There was no evidence of a backing in this conversation. Dakota said that the histogram shown in Figure 16 proves that 3% Black jurors is rare (Table 22, Turn 1). They also referenced that the histogram showed that three Black jurors does not occur very often. Harper also acknowledged that the histogram helped make the claim (Table 22, Turn 2) and noted that the expected value was eight or nine, not three. In other words, this group used a representation that they created using data science (the histogram) to justify their claim. There was evidence that the students engaged in CA because of this scaffold, which asked them to create a histogram, interpret the results of it, and discuss these results with their group members. Group 1 made the collective argument that the jury was not randomly chosen and used the simulation and histogram that they created to justify the claim. These actions mimicked what real data scientists do in their jobs. They are often creating visualizations, interpreting them, and explaining them to others. Because

the students acted as real data scientists and made a claim that they justified using data science, they engaged in CA.

Social Justice Awareness. In addition to engaging in CA, this group also engaged in social justice awareness. GD Question 2 also helped this group engage in social justice awareness because they acknowledged that non-random jury selection is problematic when Dakota said it is "kind of messed up" that they can "not do random selection when that's the rule" in Turn 3. They were able to use the analysis that they did to identify that the jury selection was not random, which is a problem. Another example of how Group 1 engaged in social justice awareness was through answering GD Question 4, which asked them what they thought about the resume study. Figure 17 shows the Toulmin Diagram of Group 1's collective argument. Table 23 shows the conversation that the students had when they answered this question.

Figure 17

Toulmin Diagram for Group 1, GD Question 4

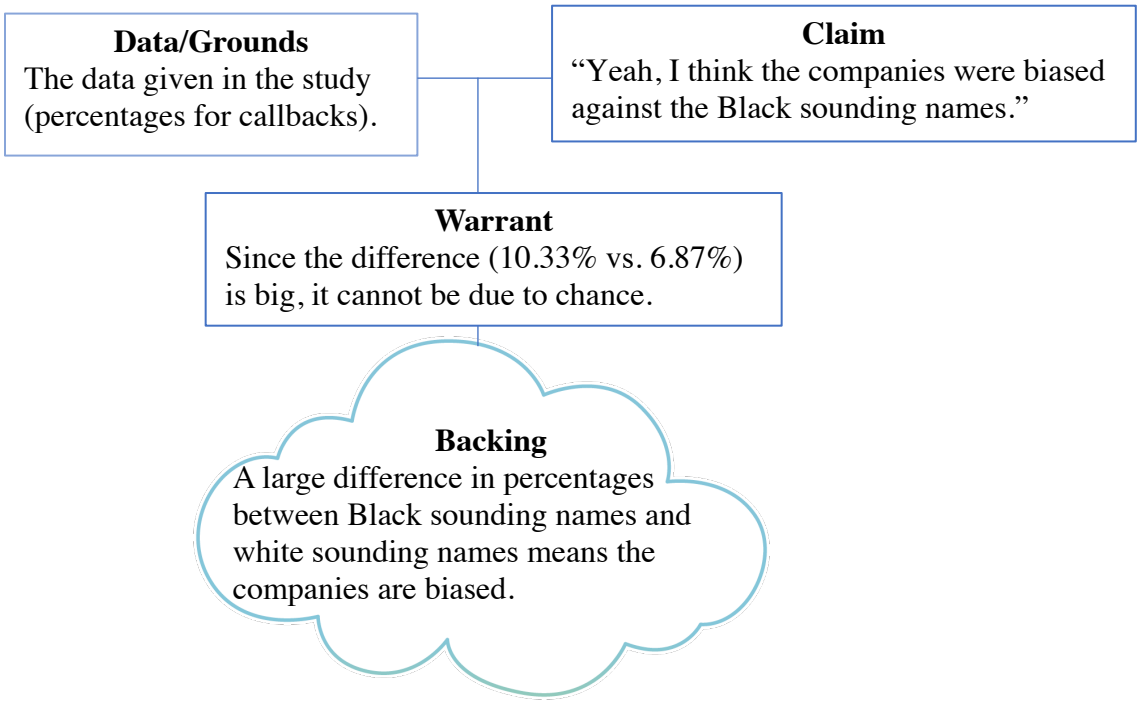


Table 23

Group 1's Conversation for GD Question 4

Turn Number	Speaker	Turn
1	Dakota	Okay so I think that difference seems really big. Do you?
2	Harper	I agree, that's like a 4% difference and I feel like based on the previous question that it wasn't due to chance.
3	Dakota	Yeah, I think the companies were biased against the black sounding names. I've heard of this happening like with companies being biased.
4	Harper	Yeah and this is just showing it with data.

Group 1 claimed that the companies were biased against the Black sounding names. The grounds come from the data in the study and the warrant is the difference in percentages. The backing is that a large difference in percentages between Black sounding names and white sounding names implies the companies were biased. By using the artifacts provided in this scaffold (Python and the resume study), the students collectively made the claim that the difference between Black and white sounding name callbacks was too big to be due to chance. Harper said this in Turn 2 from Table 23 and Dakota confirmed it in Turn 3 from Table 23. Their claim showed that they understood this issue. They acknowledged that this difference in percentages was a problem and that the companies are at fault for being biased, hence, the scaffolds and artifacts helped the students engage in social justice awareness about this particular issue.

Group 3 Summary. Group 3 mentioned that they were not sure if the difference between getting three Black people on the jury compared to the expected eight Black people on the jury was big or not during the early part of their discussion. They said they would need to do more analysis to determine this and after creating a histogram and doing a simulation, they learned that the difference was most likely not due to chance. This was evidence that they engaged in CA

because they understood that it would be best to analyze the data before making a claim. They also mentioned ideas from class in their discussion about their histograms, such as that the more times you run the simulation, the more accurate it is. By articulating that they understood these data science practices, they also engaged in CA. This group also engaged in social justice awareness by giving a solution to the resume problem and discussing why the bias with names is problematic. They mentioned that events that are similar to the resume study still occurred, but that there have been improvements in recent years.

Computational Action. There was evidence that Group 3 engaged in CA by answering GD Question 2, which had them interpret the results of their histogram. Figure 18 is the Toulmin Diagram showing Group 3's collective argument. Table 24 shows the conversation Group 3 had when answering GD Question 2.

Table 24

Group 3's Conversation for GD Question 2

Turn Number	Speaker	Turn
1	Cameron	I think 100 is not that big of a sample. Let's maybe try 1000 since my histogram isn't looking very normal
2	Avery	There's a chance that it's 3. It's a small chance, but there's a chance.
3	Cameron	If you run it a bunch of times it could be better. So if you run it 10,000 times it is much more clear. Look at these histograms. 100 is not great.
4	Jude	It does look like getting 3 is unusual with the histogram and it's definitely not the mean.
5	Avery	Yep, now that we have the histogram, it's easy to see how unusual 3 is.

Figure 18

Toulmin Diagram for Group 3, GD Question 2

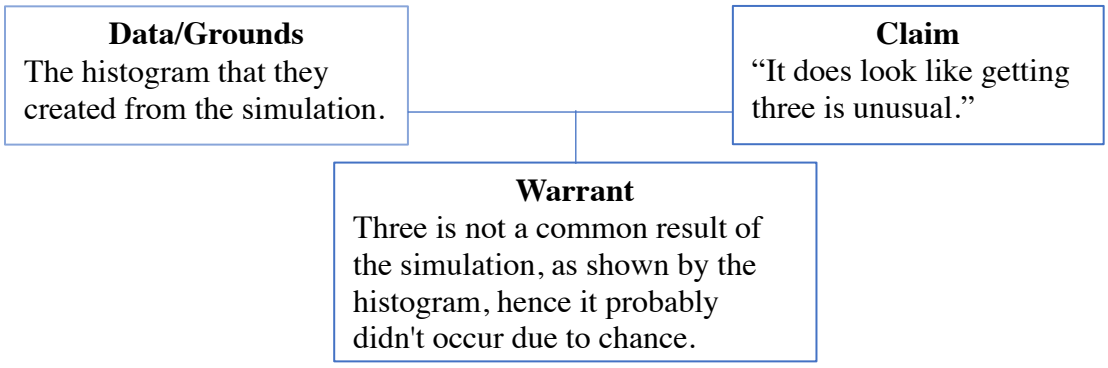
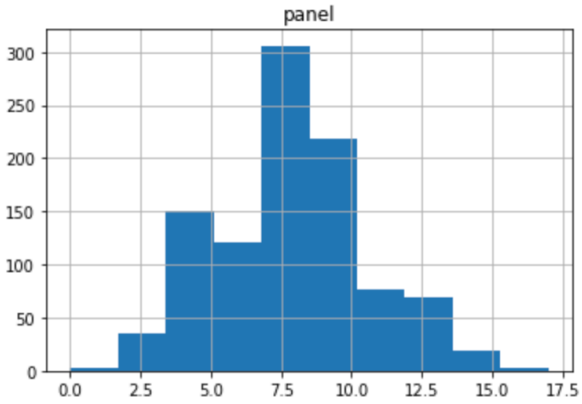


Figure 19

Cameron's Histogram



Just like Group 1, Group 3 claimed that having three Black men on the jury was not due to chance and the grounds was the histogram created from the simulation. The warrant was the idea that getting three Black people on the jury was not a common result, hence the jury was probably not chosen randomly. Group 3 had no evidence of a backing. Cameron looked at sampling 1,000 juries and produced this histogram shown in Figure 19. Cameron's histogram showed that getting three Black people on the jury was rare. Jude said this in Turn 4 from Table 24 and Avery specifically mentioned in Turn 5 from Table 24 that the histogram is what showed

them that getting three Black people is rare. This group engaged in CA in multiple ways through this question.

In their discussion, they used the histogram as a justification for their claim. In other words, they used the histogram as evidence to show the claim was true. They also understood the idea that the more simulations they ran, the more accurate their results will be, so they ran more simulations, even though the question did not ask them to. Cameron's histogram was created using 1,000 simulations even though the question asked them to do 100 simulations. This is an important data science concept that they understood and used when they answered GD Question 2. Lastly, they mentioned the idea that getting three Black people on the jury is possible, but not probable. In Turn 2 from Table 24, Avery acknowledged that getting three Black people is possible and that there is a chance that it could happen, but that it is a small chance. The difference between something being possible and probable is an important data science idea. This group used the histogram to explain this idea. Using something a representation that they created to justify a claim showed evidence that this group engaged in CA. There was also evidence that the scaffolds and artifacts helped them engage in social justice awareness.

Social Justice Awareness. Group 3 also engaged in social justice awareness by answering GD Question 5 which asked them to reflect on how data science can be used to educate people about social justice issues and whether or not similar issues still occur today. Figure 20 shows the Toulmin Diagram for Group 3's collective argument. Table 25 shows the conversation that Group 3 had when answering GD Question 5.

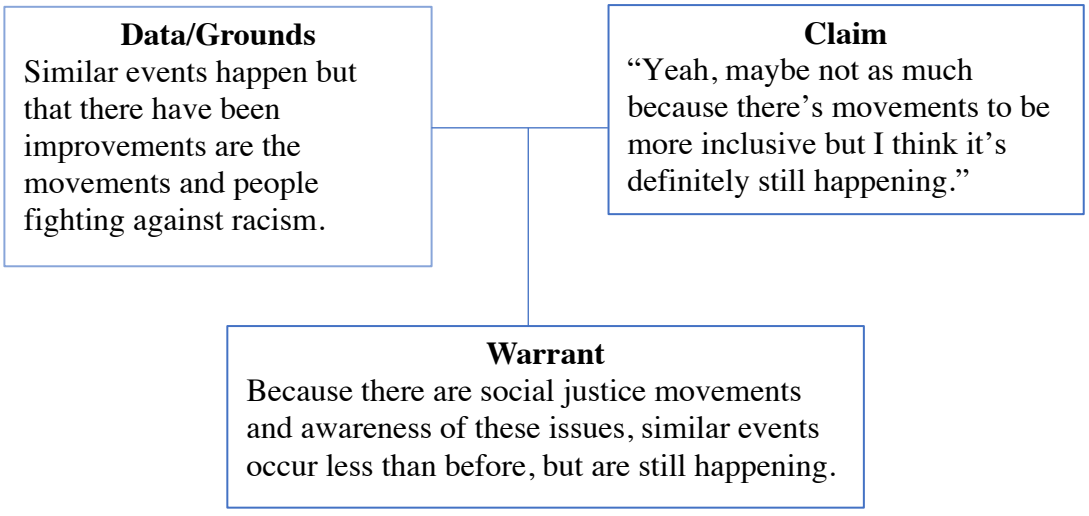
Table 25

Group 3's Conversation for GD Question 5

Turn Number	Speaker	Turn
1	Cameron	OK part three. Looks like this is just discussion. Are there similar events that occur today? Yes, for sure.
2	Jude	I think yes and no.
3	Cameron	Yeah, maybe not as much because there's movements to be more inclusive but I think it's definitely still happening.
4	Avery	Yeah, because the thing about the resumes is that they were identical except for the name. So, any difference means the names were a factor.
5	Jude	I also don't think that this would happen a lot today because there are a lot of people fighting against this stuff but like in some places, I feel like it could.
6	Cameron	I don't think it would be that different today, but I think there would be some difference with it being a little better.
7	Jude	Would you hire a resume that doesn't have a name on it?
8	Avery	I don't think we could.
9	Jude	I was saying in order to fix this issue have resumes not include names.
10	Avery	Yeah, you could do that. I think the bigger problem is that the bias exists. There are people at the company who are doing something wrong

Figure 20

Toulmin Diagram for Group 3, GD Question 5



Group 3 claimed that similar events do occur today but stated that there have been improvements with racism. In other words, their claim and argument showed that they were aware of this social justice issue. Cameron and Jude said that they could see similar issues happening today. Cameron said this in Turn 3 from Table 25, Jude said this in Turn 5 from Table 25. This group engaged in social justice awareness because they acknowledged that discrimination based on your name is still a problem. This was shown through their collective argument in Figure 20. They also mentioned that society has taken some steps in the right direction in their argument. Cameron mentioned that there are movements to be more inclusive (Table 25, Turn 3) and Jude mentioned that there are people fighting against this (Table 25, Turn 5). This group actually gave a suggestion for a solution at the end of the conversation and talked about why the bias with the names was problematic. In Turn 9 from Table 25, Jude suggested that the resume reviewers should not be able to see the names when they are reviewing them. This showed that the scaffolds and artifacts helped the members of this group engage in social justice awareness and even helped them start thinking about solutions to the problems, which also showed they were engaging in CA.

Groups that were "Great" with CA and "Sufficient" with Social Justice Awareness

Group 5 Summary. Group 5 completed all of the group discussion questions for Lab B except for the last one because they ran out of time. When discussing the jury example, this group agreed that the difference between eight and three was big, however they did not say why or what the implications of this were when they answered GD Question 1. However, when they discussed the histograms that they created for GD Question 2, they went into a lot more detail than most groups and talked about many important points. They discussed exactly what the histogram showed in relation to the scaffold, and why it showed that the jury was not randomly

selected. Group 5 was also the only group that talked about all parts of the histogram and what they meant. In other words, they talked about both the *x*-axis and the *y*-axis. They took their knowledge from class about the histogram and applied it to this real-life scenario. They did acknowledge social justice and mentioned that they have heard about issues that were similar to the resume study, however they did not elaborate on these ideas, instead they just mentioned that they were familiar with them.

Computational Action. There was evidence that Group 5 engaged in CA in their discussion of GD Question 2, which had them interpret the results of their histogram. Figure 21 shows the Toulmin Diagram of Group 5's collective argument. Table 26 shows the conversation that Group 5 had when answering GD Question 2.

Figure 21

Toulmin Diagram for Group 5, GD Question 2

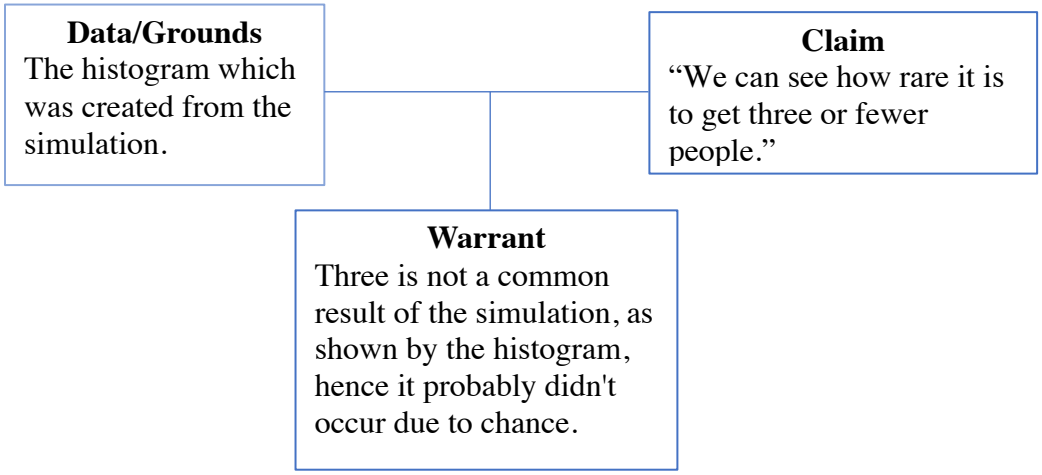


Table 26

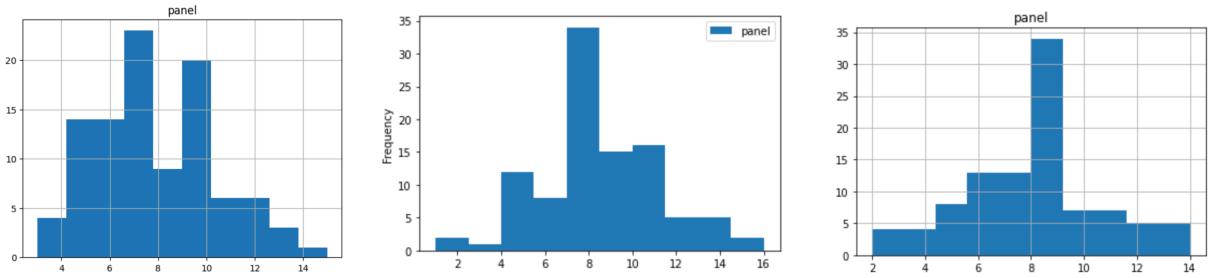
Group 5's Conversation for GD Question 2

Turn Number	Speaker	Turn
1	Casey	We all have different histograms.
2	Alex	Yeah because it's random every time.
3	Bailey	They're all definitely higher than three.
4	Casey	What does this tell us?
5	Bailey	They lied.
	Casey	Eight percent of the population is black and 3% were chosen I don't understand what the histogram is telling us.
6	Bailey	We can see how rare it is to get three or fewer people
7	Alex	Which one is the people one?
8	Bailey	I think it's x.
9	Alex	Then what's y?
10	Bailey	I think it's frequency of getting less than three.
11	Alex	I think it's rare because my histogram is centered around six to eight.
12	Bailey	Yeah mine is too it's pretty rare.
13	Alex	There's a slight possibility it would happen by chance but it's very slight.

Figure 22

Casey, Alex, and Bailey's Histograms Side by Side

Casey's Histogram Alex's Histogram Bailey's Histogram



Like many other groups, Group 5 claimed that the jury was not randomly selected. The grounds were the histograms they created from the simulations shown in Figure 22. The warrant was that the histogram showed that three was not a common result of the simulation which meant

that the jury probably was not randomly selected. There was no evidence of a backing, and you can see that three was not a common result in any of the histograms in Figure 22. This scaffold helped the group engage in CA because it asked them to create a histogram and interpret the results of it. Creating a representation using programming and then explaining it is a common data science practice. Their collective argument shown in Figure 21 is evidence that Group 5 engaged in CA by using data science representations to justify a claim about a social justice issue.

Group 5 also was the only group who explained the meanings of the *x*-axis and *y*-axis of the histogram using the context of the problem. Alex asked what the *x*-axis and *y*-axis represented in Turns 7 and 9 from Table 26, and Bailey explained that the *x*-axis represented the number of people, and the *y*-axis represented the frequency. Group 5 described what the *x* and *y* axes showed and used the histogram to justify why getting three Black people is rare. From this conversation, we can see that there was one student who was confused, Alex, and the other students in the group explained why the histogram helped justify the claim. Group 5's discussion about the two axes on the histogram exemplified how knowledge was distributed in the group discussions. Here, the students shared knowledge about the histograms and what they were showing. They also acknowledged the idea that there was a slight possibility that the jury could have been chosen randomly, but it was very small. Alex talked about this in Turn 13 from Table 26 and described the chance of the jury being chosen randomly as "slight." By recognizing this distinction between an event being possible, but not probable and using the histogram to explain it, Group 5 engaged in CA through this discussion.

Social Justice Awareness. In addition to engaging in CA, Group 5 also engaged in social justice awareness and was sufficient with this goal. They did this by answering GD Question 4

which asks them about the resume simulation. Figure 23 shows the Toulmin Diagram of Group 5's collective argument and Table 27 shows the conversation that Group 5 had when answering GD Question 4.

Figure 23

Toulmin Diagram for Group 5, GD Question 4

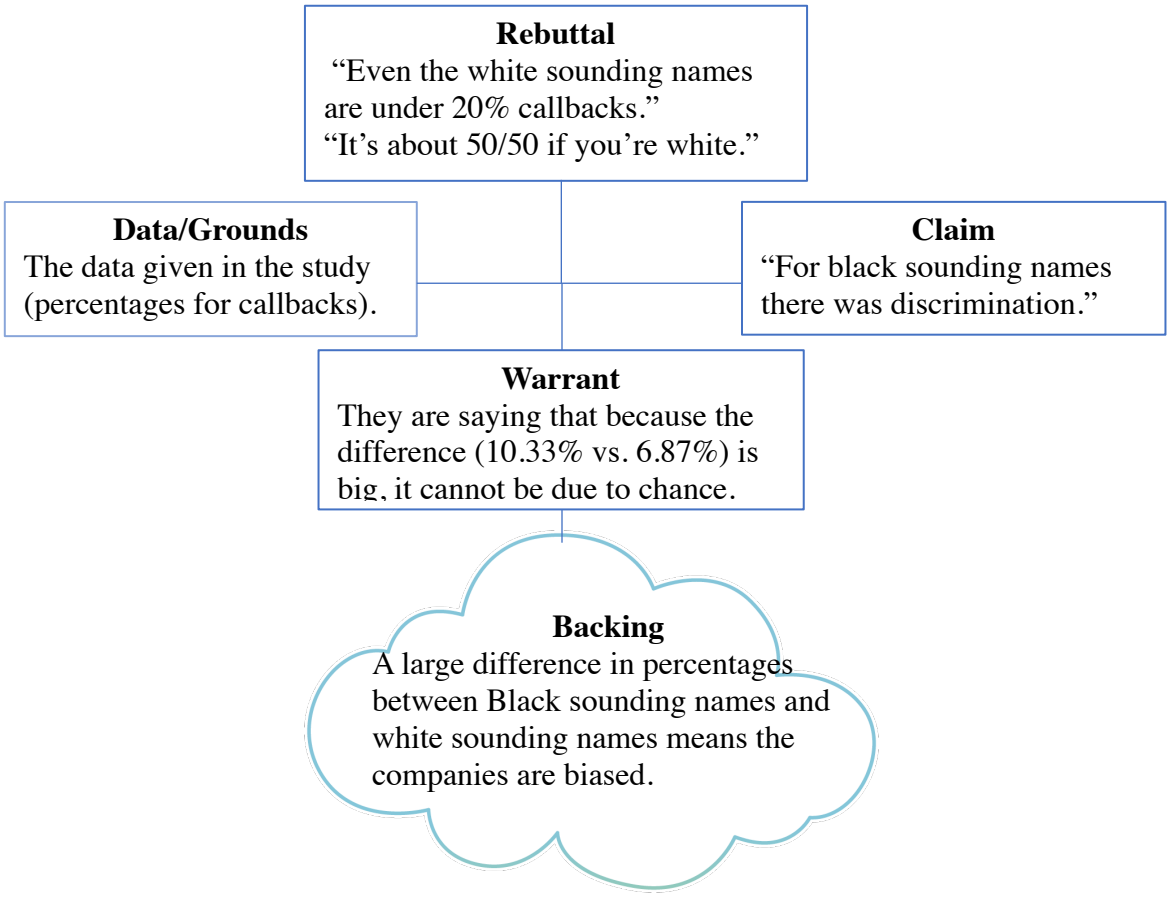


Table 27

Group 5's Conversation for GD Question 4

Turn Number	Speaker	Turn
1	Alex	Yep, so the probability is quite small to get this difference.
2	Bailey	I've heard of this.
3	Casey	I would say it's significant, right?

Table 27 (cont.)

4	Bailey	Yeah. The difference is huge between Black and white.
5	Casey	Even the white sounding names are under 20% callbacks.
6	Bailey	It's a bit surprising for both.
7	Alex	It's about 50/50 if you're white.
8	Casey	For Black sounding names there was discrimination.

Group 5 claimed that there was a significant difference between the white and Black sounding names. Casey said that the difference is significant (Table 27, Turn 3) and Bailey confirmed this (Table 27, Turn 4). The grounds were the data given in the study and the warrant was that the difference between percentages is big. The backing was that this large difference implies that the companies were biased. Group 5 did engage in social justice awareness because their claim was related to social justice since they acknowledged the discrimination; however, they did not elaborate on anything besides acknowledging that there was a difference between Black and white sounding names. The rebuttal in Group 5 came from Casey in Turn 5 from Table 27. Casey mentioned that it was also surprising how low the percentage of callbacks was for the white sounding names was even though it was a lot higher than for the Black sounding names. Bailey mentioned that the percentages are surprising for both groups (Table 27, Turn 6). Although they did discuss their thoughts on this, noticing that the white sounding names had a low percentage of callbacks was the not the point of the lab, which is why this group did not score "Above and Beyond" in the social justice awareness category.

Group 6 Summary. Group 6 completed all of the group discussion questions for Lab B except for the last one because they ran out of time. Group 6 was skeptical about whether or not the difference between the observed value (8%) and the expected value (3%) for Black people on the jury was big or not. After looking at their histograms, they came to the collective agreement

that the difference is big and that the jury was not randomly chosen. They justified this using their histograms, which was evidence that they engaged in CA. This group also showed that they engaged in some social justice awareness in their discussion about how data science can be used to address racism. They understood that data science can be used to address racism, however they did not elaborate on exactly how this was done.

Computational Action. One way this group engaged in CA was by answering GD Question 2, which had them interpret the results of their histogram. Figure 24 shows the Toulmin Diagram of Group 6's collective argument and Table 28 shows the conversation that they had about GD Question 2.

Table 28

Group 6's Conversation for GD Question 2

Turn Number	Speaker	Turn
1	Rowen	Since I got three out of 100 only a few times, I think three is pretty low, but I guess it's possible.
2	Marley	The histogram helped differentiate human action from random chance. That's essentially what we're doing here.
3	Rowen	It seems some of them fall within the range, but it's only a few.
4	Marley	My 1.9 says the probability is 3% which is low. I think this means the jury wasn't fair.
5	Rowen	We've already found that a difference of five percent in our expected value is unlikely but not so unlikely that it isn't possible. A difference of four percent seems believable, but since this is five, I'd say the jury wasn't fair.

Figure 24

Toulmin Diagram for Group 6, GD Question 2

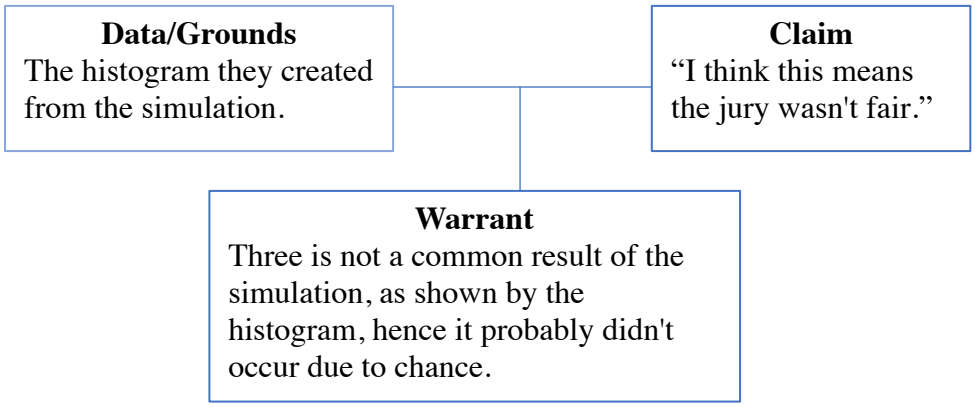
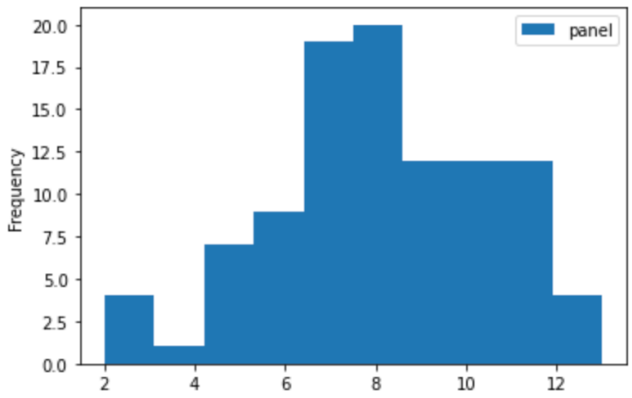


Figure 25

Rowen's Histogram



Like many other groups, Group 6 claimed that the jury was not randomly selected. The grounds were the histogram they created from the simulation and the warrant was the idea that three was not a common result, meaning that the jury was probably not chosen randomly. Rowen's histogram in Figure 25 shows that three was not a common result. Group 6 used the histogram as the data to justify the claim, but they did this in multiple ways. They looked at how many times three showed up and saw that it did happen, but it was very uncommon. Rowen

mentioned this in Turn 1 from Table 28 by saying that they only got three a few times (see Figure 25). Group 6 also looked at the difference between 3% (their observed value) and 8% (their expected value) on the histogram and saw that the difference was very big. Rowen specifically mentioned that the difference is 5% in Turn 5 from Table 28, which is too big to be due to chance. The students engaged in CA because they created a histogram (from the scaffold) and then discussed how it helped them justify their claim in their collective argument. They were able to do this in more than one way and with various tools. Using a visualization like a histogram to justify a claim is something that data scientists do frequently. Their claim and justifications shown in Figure 24 showed that Group 6 engaged in CA by acting as data scientists.

Social Justice Awareness. In addition to engaging in CA, this group also engaged in social justice awareness. One example of how they did this was through answering GD Question 3, which looks at how students can use data science to address issues of racism specifically. Figure 26 shows the Toulmin Diagram of Group 6's collective argument and Table 29 shows their conversation when answering GD Question 3.

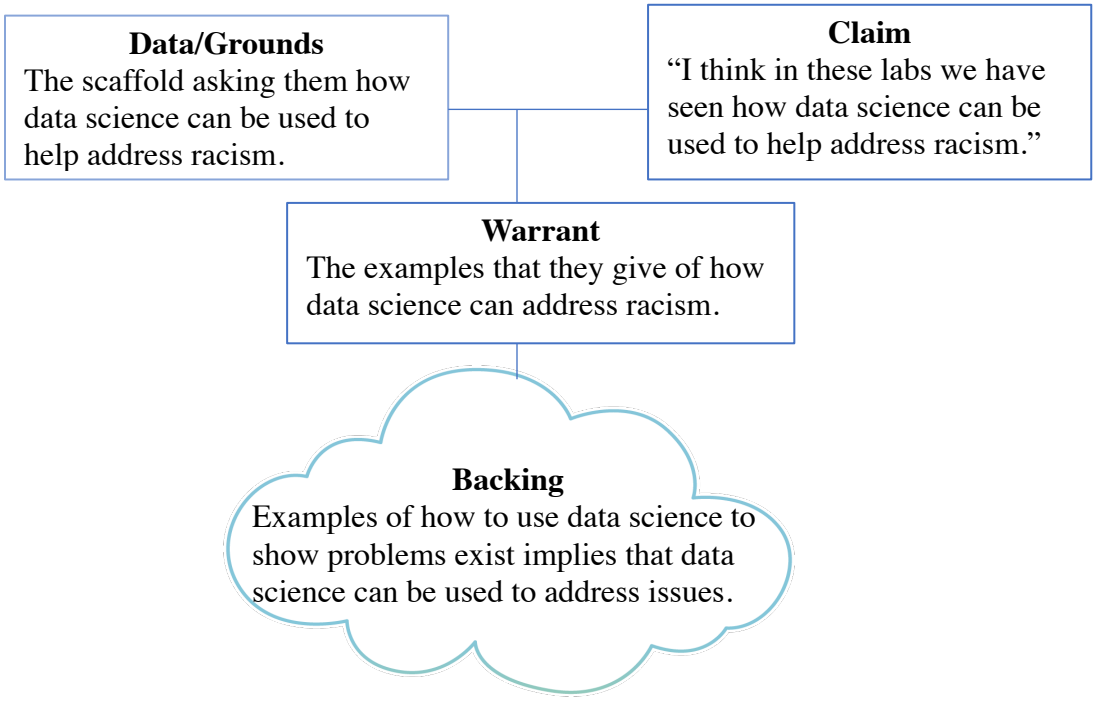
Table 29

Group 6's Conversation for GD Question 3

Turn Number	Speaker	Turn
1	Rowen	I think in these labs we have seen how data science can be used to help address racism because we can do simulations and stuff to prove it.
2	Marley	Yeah, I've never really thought about this before until now.
3	Rowen	And with data science you can see the exact numbers for discrimination and use them.

Figure 26

Toulmin Diagram for Group 6, GD Question 3



Group 6 claimed that data science can be used to address issues of racism. Rowen claimed this in Turn 1 from Table 29. The grounds were the scaffold, which asked them how data science can be used to address racism. Group 6 gave examples of how you can use data science to address racism as the warrant, however these were vague examples. Rowen mentioned that we can do simulations (Table 29, Turn 1) and see the exact numbers in Turn 3 from Table 29. Marley mentioned that they had never thought of the intersection of racism and data science until doing this lab (Table 29, Turn 2). This showed that the scaffold prompted them to think about this key idea. This conversation and collective argument did show that Group 6 did engage in social justice awareness; however, they only gave vague examples as the warrant. They said in general that the examples in the lab show how data science can be used to address issues of racism in Turn 1 from Table 29 but did not give any unique other examples outside of the lab. So

overall, they did engage in social justice awareness, but did not elaborate on anything they said in this specific discussion.

Group 4 Summary. Group 4 completed all of the group discussion questions for Lab B except for the last one because they ran out of time. Their discussion about the histogram that they created was very similar to the other groups. They used a representation that they created (their histograms) to justify their claim about whether or not the jury was randomly selected. They also described what their histograms looked like by taking note of the range of the *x*-axis. For the questions regarding social justice issues, Group 4 group had shorter discussions but did talk about all the important points, which is why they scored sufficient for social justice awareness on the rubric.

Computational Action. One way this group engaged in CA was by answering GD Question 2, which had them interpret the results of their histogram. Figure 27 shows the Toulmin Diagram of Group 4's collective argument and Table 30 shows their conversation when answering GD Question 2.

Figure 27

Toulmin Diagram for Group 4, GD Question 2

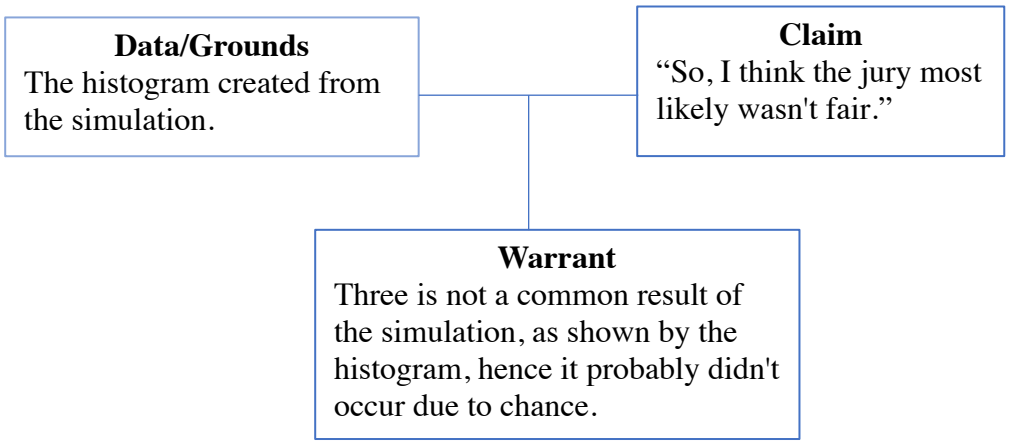


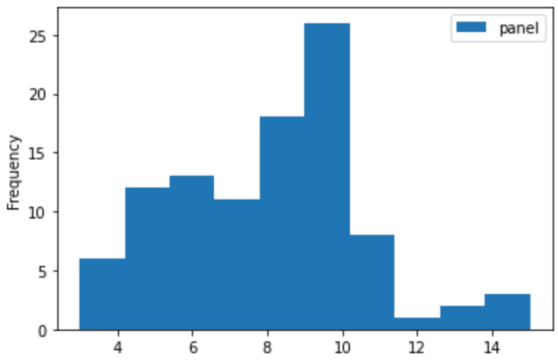
Table 30

Group 4's Conversation for GD Question 2

Turn Number	Speaker	Turn
1	Jordan	If you look at mine, three doesn't even show up once for me so it's technically possible but not very likely.
2	Logan	Mine shows from two all the way to 16.
3	Jordan	Mine is from four to 16.
4	Jayden	It still could happen right?
5	Jordan	Yeah, it could but the probability of that happening is very small. So, I think the jury most likely wasn't fair.

Figure 28

Jordan's Histogram



Group 4 claimed that the jury selection was not fair. The grounds were the histograms they created from the simulations. The warrant was that three was not a common result of the simulation, meaning that the jury was most likely not chosen randomly. There was no evidence of a backing. This group talked about the range of the *x*-axis of the histogram. Logan said their range was from 2 to 16 (Table 30, Turn 2) and Jordan said their range was from 4 to 16 (Table 30, Turn 3). While talking about the range, they also discussed how getting three Black people to be on the jury was very unlikely. Jordan described how you can see that three is unlikely because of the histogram. Jordan's histogram is shown in Figure 28. Jordan said that three did not show

up on his histogram (Table 30, Turn 1) and this is evident from Figure 28. Jordan also made the distinction that getting three Black people on the jury is possible, but not probable (Table 30, Turn 5). This exchange showed that Group 4 was using the histogram to justify their claim, and their discussions were similar to the previous groups. The students did engage in CA because of the scaffold, which asked them to create a histogram and interpret the results of it. The students used the histogram as data to justify their claim. They created something using data science, interpreted what they created, and used it to justify a claim about a real-world issue.

Social Justice Awareness. One way Group 4 engaged in social justice awareness was by answering GD Question 4 which asked them to think about the resume scenario and decide if they thought the difference in percentages of callbacks for Black sounding names and white sounding names was significantly different. Figure 29 shows the Toulmin Diagram of Group 4's collective argument and Table 31 shows their discussion.

Figure 29

Toulmin Diagram for Group 4, GD Question 4

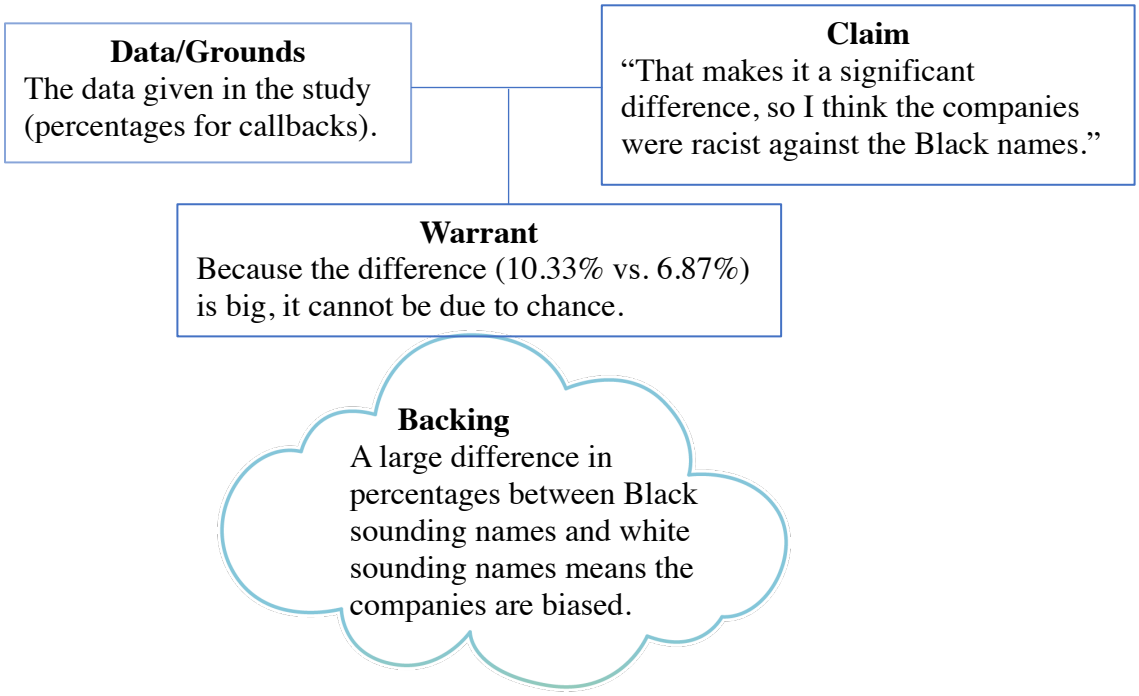


Table 31

Group 4's Conversation for GD Question 4

Turn Number	Speaker	Turn
1	Logan	What did you say for this?
2	Jordan	I just said 4% is a big difference because the applications are identical other than the name.
3	Jayden	That makes it a significant difference, so I think the companies were racist against the Black names.
4	Jordan	Yeah, I could understand if it was like 1%, but 4% is really sketchy.

Group 4 claimed that there was a significant difference between the white and Black sounding names. The grounds were the data from the study and the warrant is that the difference between percentages is big. The backing was that a large difference in percentages implied that the companies were biased. This group did engage in social justice awareness because they understood the problem and made a claim about the social justice issue. Jordan explained in Turn 2 from Table 31 that the difference in percentages was big since the applications are identical except for the name. Despite this, the group did not elaborate on anything besides acknowledging that there was a difference. Jordan did say that the large difference was "sketchy," but the discussion was short (Table 31, Turn 4). They seemed to really understand the resume study and mentioned the key idea that they resumes were identical except for the names, however they did not mention the implications of the social justice issue.

Groups that were "Needs Improvement" with both CA and Social Justice Awareness

Group 2 Summary. Group 2 was the only group that scored "Needs Improvement" in both categories on the rubric. There was some evidence that this group engaged in CA and social justice awareness, but there were also some large problems in their reasoning. When Group 2 answered GD Question 4, there was evidence that they engaged in CA and social justice

awareness because they understood that the simulation showed that the percentage for the white sounding names was close to the expected value (10%) and the percentage for the Black sounding names was not close to the expected value. Kendall articulated that both should be close to 10% (Table 32, Turn 3) and Sawyer said that this meant there was discrimination based on your name (Table 32, Turn 4). Understanding this key idea shows that they became aware of this social justice issue and were able to use the simulation that they created to make sense of the results. Group 2's conversation for GD Question 4 is shown in Table 32.

Table 32

Group 2's Conversation for GD Question 4

Turn Number	Speaker	Turn
1	Kendall	One of them is close to ten percent and the other is also close to ten percent
2	Sawyer	You're trying to think about the actual percentage though, right? So, are you writing your answer based on 10.3 percent?
3	Kendall	Yes. They both should be close to 10% and 10.3 is but the other one is not.
4	Sawyer	Got it so there is discrimination based on your name.
5	Kendall	Yeah, that's what the simulation showed.

Although there was evidence that they did engage in CA and social justice awareness when answering GD Question 4, some other questions presented problems with the goals. For example, when Group 2 answered GD Question 2 which asked them to interpret their results of their histogram, they did not interpret the results correctly. Table 33 shows their conversation for GD Question 2.

Table 33

Group 2's Conversation for GD Question 2

Turn Number	Speaker	Turn
1	Sawyer	I mean mine looks fair.
2	Kendall	What does it look like? Yeah, I mean it looks random, so I think it's fair.
3	Sawyer	I feel like there's a big portion here, and I feel like that's fair, and it might be an unpopular opinion. Is that how you're supposed to read it?
4	Kendall	I think so.
5	Sawyer	I just said that mine shows it to be fair. A lot of the frequencies and densities below nine provides evidence the jury is more likely to be even. This may have claimed to be unfair because of one test case.
6	Kendall	It does seem like it is less than 8, but it could happen due to chance, because it's not too far off but then I also said that if you think about it not being by chance then it couldn't be supported because it's only 100 simulations.

Table 33 shows that Group 2 misunderstood the idea of something being possible, but not probable. They argued that the jury selection was fair because their histogram showed that getting three Black people on the jury was possible. Sawyer said it looked like the jury was randomly selected because they see values below nine on the histogram (Table 33, Turn 5). Kendall also confirmed what Sawyer said in Turn 6 from Table 33 and claimed that 100 simulations was not enough to prove that the jury was not selected randomly. Group 2 did not understand that the histogram showed that getting three Black people was actually very rare. Here they used the histogram as the warrant a way to justify their incorrect claim that the jury was randomly selected. Because they did not understand the point of the question and did not interpret the histogram correctly, I thought they needed improvement for engaging in CA.

Group 2 also needed improvement for social justice awareness because while they did claim that data science can be used to educate people about racism, they spent most of their discussion focusing on the idea that there could be a "legitimate reason" why Black people were not getting callbacks. Table 34 shows their conversation about GD Question 5 which asked whether similar events occur today.

Table 34

Group 2's Conversation for GD Question 5

Turn Number	Speaker	Turn
1	Sawyer	I think it's asking what plays the role in how the actual percentage happens and how computers do it because computers don't have bias, but humans do have bias.
2	Sawyer	I feel like data science could help educate but then at the same time it's not about numbers it about bias. It's about people's belief and how they interact with people it's nothing about how a computer can just give the average amount of black people in a country.
3	Kendall	I think that's the point. You can see what the expected value is and then you can get actual data from real life. You can use data science to get percentages from that data frame
4	Kendall	Now I think about the other side of the reasoning. Like why there could be ten percent and six percent for the resumes. Is there a legitimate reason?
5	Sawyer	I feel terrible saying this. But I understand why there could be difference in selecting people like instead of the ten percent split like what if they don't meet the job criteria, so they don't receive a call back. Like what if it has nothing to do with their race, but they just aren't qualified.

Group 2 mentioned that maybe the resume study actually was fine because the people with Black sounding names may not have been qualified for the jobs that the resumes were sent to. Sawyer mentioned this in Turn 5 from Table 34. However, they missed the point here again

because the resumes were identical except for the names. Focusing on finding what Kendall called a "legitimate reason" for why Black people would get less callbacks (Table 34, Turn 4) was not the goal of this question. Talking about the people being qualified also does not make sense since the resumes sent out were identical except for the name. Overall, this group had some good discussion, but the problematic discussions made it clear that this group needed improvement.

Similarities Between Groups Who Completed Lab B

The six groups who completed Lab B had a lot of similarities, and I described them in this section. Just like in Lab A, I noticed that everyone in the groups who completed Lab B talked and participated in the conversations about data science. There were not any students who were excluded from the discussions or did not participate. Because the groups discussed data science concepts, multiple social justice issues, and the implications of analyzing data related to these issues, they engaged in computational identity (a part of CA) because each student was an important part of the group who contributed unique ideas to the discussion. The students added their unique perspectives to the conversations, shared their knowledge with their group members, and made collective arguments. Lab B contained GD Questions that specifically ask about how data science and social justice are related, such as GD Question 5. These questions helped the students with computational identity and sharing knowledge because these scaffolds had the students do something unique and then share their results with their groups and explain what they did. These types of group discussion questions also helped the students engage in digital empowerment (another part of CA) because they were able to use the artifacts in the labs such as the datasets and Python to answer questions that they had about the data.

Most groups used the data analysis that they did as the grounds or warrants for their arguments. In other words, the Toulmin analysis showed that the groups used the data science that they did to justify their claims and that the scaffolds provide data for these claims. Also, the backings are data science principles involving topics they have covered in class. A common theme among all of the groups was that for GD Question 2, the students used their histogram as grounds to justify why the jury was not randomly selected. The Toulmin analysis showed that this scaffold gave them the opportunity to engage in CA because they used something that they created to justify a claim about a social justice issue.

The last group was the only one that seemed to use their justifications incorrectly and did not make claims that made sense. The other groups were able to make valid claims through using the artifacts in the labs and the representations that they created from the scaffolds. Lab B was centered around two social justice issues and most groups engaged in awareness. Lab B was more focused on social justice issues than Lab A. I included scaffolds and artifacts directly related to social justice which helped make the discussions focused on social justice. A lot of the groups did not answer every part of the question, mainly due to running out of time at the end. However, they still seemed to engage in CA and social justice awareness. So overall, the Toulmin analysis showed that the GD Question scaffolds and artifacts in both labs helped the students engage in CA and social justice awareness. Next, I look at how the individual reflection questions also helped the students engage in CA and social justice awareness.

Overview of Individual Reflection Question Analysis

In addition to analyzing the group discussions, I also look at the individual reflection (IR) questions from the labs to answer how the students used the scaffolds and artifacts to engage in CA and social justice awareness. For the individual reflection questions, the unit of analysis is

each student and I use Thematic Analysis for analysis. For the Thematic Analysis, I identify four themes that were common among all of the students who participated in the study and completed Lab A. I describe each theme in detail and explained which codes went into each theme. Next, I describe how each theme connects to the two goals of the lab (CA and social justice awareness). And lastly, I look at each individual reflection question in the lab and discuss which themes were most common in the responses to each question and give examples of what students said.

Thematic Analysis for Lab A

Theme 1: Importance of Visualizing Data. The students mentioned that visualizing data is important for many reasons. Oftentimes, descriptive statistics alone are not enough to understand trends and important aspects of data. Basic data visualization is a key part of exploratory data analysis. This is an idea from data science that is important. A data science practice is to visualize data in ways that make it easier to understand. The students engaged in visualization through creating simple visualizations such as boxplots, histograms, and scatterplots in this lab. The codes that make up this theme are codes that acknowledge that visual displays of data make it easier to see outliers, easier to see the spread or distribution of the data, and easier to see the center of the data and how close the numbers are to the center. This theme also includes students' acknowledgement of the power of visual displays of data beyond the use of descriptive statistics. As students recognized the importance of visual displays of data and the role they play in data analysis, they engaged in CA. This is a key concept in data science and the IR questions allow us to see how well students understood this concept. Understanding data science concepts and being able to articulate them in writing is one way students can engage in CA.

Theme 2: Social Justice. I categorized any code that showed that students acknowledged any aspect of social justice issues or mentioned social justice as this theme. Social justice awareness is also a goal of both labs, and it is very related to CA. The two goals are interconnected throughout each lab. The main social justice issue in Lab A was the salary differences among men and women in the dataset. The codes that created this theme include the students explicitly saying that salary data are important to collect, students discussing that there are differences between salaries of men and women, students acknowledging that the differences are problematic, and students mentioning that this analysis can harm some people and benefit some people. This theme is clearly connected to the goal of justice awareness. Also, since students thought about issues outside of the classroom involving social justice, such as the discrepancies in pay among men and women, as well as the benefit and harm their analysis can create, this helped them engage in CA.

Theme 3: Thinking About Mobilizing for Change. This theme takes CA and social justice awareness a step further. Here, the students talked about ways they can advocate for change, collect more data, or use counter data in a future analysis. The codes that created this theme were codes where the students talked about counter data, collecting more data to get a more thorough picture of what was going on, and advocating for making changes based on the data analysis. This theme is also clearly connected to the goal of social justice awareness. Because the students thought about counter data and other data they could collect to understand the problem better, they engaged in CA when answering questions that involved this theme.

Theme 4: Data Science Discoveries. A common idea that I saw in the IR questions was students using the analysis they did to discover something. For example, many students talked about how the analysis that they did showed them something that they did not know before. In

other words, the analysis that they did on the data helped them learn new information. Students also talked about how their analysis produced surprising results, or results that they did not expect. The codes that created this theme were codes that specifically looked at students talking about what they learned from their analysis. This theme is connected to CA because the students used the analysis that they did in the lab to discover something new. Using data science as a tool to learn is a part of CA.

Individual Reflection Questions in Lab A. Next, I looked at each individual reflection question from Lab A and talked about which themes came from the responses to each of these questions. I also gave examples of specific responses related to each theme. Table 35 shows each individual reflection question and a count of which themes surfaced in the written responses to those questions from the students.

Table 35

Lab A Individual Reflection (IR) Questions and Themes

IR Question	Theme 1: Importance of Visualizing Data	Theme 2: Social Justice	Theme 3: Thinking About Mobilizing for Change	Theme 4: Data Science Discoveries
IR Question 1: Explain why it is valuable to look at visual displays of salary data in addition to descriptive statistics like the mean, median, and SD. Write at least 3 sentences.	8			
IR Question 2: Write down something you learned from your group discussion! Write at least 3 sentences.				8
IR Question 3: Describe which question you answered and how you answered it.		1		8
IR Question 4: Write a paragraph style response (at least 5 sentences) summarizing what		8	5	2

Table 35 (cont.)

you learned from working with the salary data. We have listed a few questions below to give you some ideas on what to write about if you need them.

In the descriptions in the next section, I highlight the most common themes for each question. I defined most common themes as themes that showed up in more than five students' IR Question responses. For IR Question 1, Theme 1 was the most common, for IR Questions 2 and 3, Theme 4 was the most common, for IR Question 4, Themes 2 and 3 were the most common.

IR Question 1. IR Question 1 asked students to describe the importance of visual displays of data. The goal was for them to understand that it is important to include visual displays of data in their analysis, in addition to descriptive statistics. The students were supposed to explain why this is true. This scaffold was included in the lab to help the students engage in CA. By being able to articulate this data science concept in writing, this question was intended to help the students engage in CA since they were getting practice explaining concepts and justifying why data science is important. Theme 1, the Importance of Visualizing Data, was the most common theme in the responses to this question. The students described why visualizing data is important. They mentioned that you could see outliers better and that looking at visual displays of data in addition to descriptive statistics gives a better picture of the data as a whole. Table 36 shows what Logan from Group 1 and Alex from Group 2 said for IR Question 1.

Table 36

Examples of Responses to IR Question 1

Student	IR Question 1 Response
Logan	It is valuable to create visual displays as it allows information to be viewed at easier. Descriptive statistics represented in a visual format are represented in a way that is more appealing in terms of data visualization. This type of format allows you to see the outliers and any other aspect that may not be apparent through viewing numbers.
Alex	Individual numbers such as mean, median, and standard deviation are somewhat abstract. Visual displays show a more complete picture of the shape of data. It is a lot easier to grasp the number of outliers for data with a boxplot for example than with a standard deviation.

Logan referred to visual displays of data in their response and they highlighted that visualizing data allows us to see things that may not be apparent from just looking at descriptive statistics, such as outliers. This is a key data science concept and shows evidence that Logan engaged in CA from IR Question 1. Alex also discussed outliers in their response to IR Question 1. Alex said that visualizations help us see the shape of the data and whether or not there are outliers. Also, Alex mentioned that it is easier for people to interpret visual displays of data compared to descriptive statistics, which are just numbers, not pictures. By explaining why visual displays of data are important and understanding that visual displays of data help us see a clearer picture of the data, this is evidence that Alex engaged in CA. It was interesting that many of the students specifically mentioned that visual displays of data allow you to see outliers easier. Most of the students said this was true for boxplots, but some also said that histograms allow you to see outliers easier. Understanding the importance of looking descriptive statistics in addition to visual displays of data is an important concept in the practice of data science.

IR Question 2. IR Question 2 instructed students to write down something that they learned from the group discussion they had just before answering this question. During their

group discussions, the students discussed the analysis that they did to answer a question that they were interested in about their own department. They were supposed to write down something that they learned from their group discussion that they found interesting. This could have been something that their peers taught them or something that they learned about their own analysis that was brought out by the group discussion. The idea here was for the students to be able to explain what they learned from the group discussion in writing and summarize it in a few sentences. A lot of students talked about something that they learned about the dataset specifically. This scaffold was included in the lab to help the students engage in CA. Discussing data analysis with group members and then being able to write about it is something that data scientists do regularly. In other words, this is a data science practice. By summarizing what was said in a group discussion in writing, the students got practice communicating about data science concepts, which in turn helped them engage in CA.

Theme 4, Data Science Discoveries, was the most common theme in the responses to this IR question. The students used the analysis that they did or that their group members did to discover something that they did not know before. Most of the students wrote about something that they learned about their own department or something that they learned from their group members about their department. By having these discussions and describing what they found interesting, the students shared knowledge with each other. Table 37 shows what Rowen from Group 3 and Logan from Group 1 said for IR Question 2.

Table 37

Examples of Responses to IR Question 2

Student	IR Question 2 Response
Rowen	Overall, statistics professors in general makes a lot less money than advertising professors. This is unexpected to me because I would expect a mathematician/ math professor to make more money than an advertising professor because of how skill-based stats is compared to advertising. Also, the range is strangely similar, as the lowest salary is around $140,000 away from the highest.
Logan	Econ has the highest paid salary with ~275K. It's interesting to see the differences across multiple concentrations.

Rowen found that Advertising professors make more than Statistics professors from their discussion. They mentioned that this was surprising and gave a reason that they were surprised. Rowen's reasoning was that statistics is a type of math and math is more "skill-based" compared to advertising. Because Rowen and their group used Python to answer a question they had and then wrote about the discussion that they had about the results, IR Question 2 helped them engage in CA since doing analysis and writing about it is a data science practice. Similarly, Logan learned that the Economics Department has the highest salary and mentioned that they found the data to be interesting. Logan did the calculation about the highest salary using Python, shared these results with their group, and then noted that there were differences among the departments. Each of these actions helped Logan engage in CA.

IR Question 3. IR Question 3 asked students to describe a question that they had about the Salary Dataset and write about how they answered their question. For this question, the students were supposed to write about the data analysis that they did for a unique question that they had. The previous scaffold asked the students to use their data science skills to answer a question that they had that was not answered in the lab. IR Question 3 question wanted them to summarize the question and what they did to answer it. This scaffold was included in Lab A to

help the students engage in CA. By answering a real question that they had using data science and this real-world dataset and then writing about the results, the students acted as real data scientists and engaged in CA. They acted as though they were data scientists engaging in data science practices, rather than students who were answering pre-made questions on an assignment. Many of the students investigated the discrepancies between male and female salaries, which also helped them engaged in social justice awareness.

Theme 4, Data Science Discoveries, was the most common theme in the responses to this IR question. The students used their data science skills and the artifacts given in the lab (Python, the dataset, etc.) to answer a real question that they had. In their responses to IR Question 3, they described what question they had and what the answer was. This scaffold was designed to help them discover something that they did not know before. Table 38 shows what Logan from Group 1 and Marley from Group 3 said for IR Question 3.

Table 38

Examples of Responses to IR Question 3

Student	IR Question 3 Response
Logan	What can be observed is that males have a higher, mean, higher median, and a lower standard deviation. What this means is that males are more likely than females to have a higher salary, with a deeper concentration and less of a spread.
Marley	I answered both (highest and lowest salary), but they were asking for the opposite of the same thing. To start, I asked which professor made the highest salary. I did this by making a new data frame (df6) and then using the formula on the cheat sheet (df.nlargest). Since I was only looking for who made the most money (only one person) I had n= 1 and then I was looking at salary (my value). I then printed this answer (print(df6)) to find the information regarding the professor. I then did the opposite instead of largest I used smallest."

Logan wanted to investigate and compare the descriptive statistics between male and female salaries. In their response, Logan noted that males generally have higher salaries and have

a lower spread. Logan did not talk about the implications of their analysis; however, IR Question 3 did not ask them to. There was evidence that Logan did engage in social justice awareness by noticing this difference in salaries. There was also evidence that Logan engaged in CA from using Python to answer a real question they had. There was also evidence that Marley engaged in CA from answering IR Question 3. Marley described their exact steps to answering their question, including the Python code that they used to answer the question. Marley justified why they used the coding that they did and how that code changed when they were looking at largest salary and the smallest salary. Summarizing their question and being able to articulate exactly how they solved it in writing helped Marley engage in CA.

IR Question 4. IR Question 4 asked students to write a paragraph style response summarizing what they learned from working with the salary data in Lab A. This was the final reflection question of the lab. The students were supposed to reflect on what they learned through the lab and the analysis they did. This scaffold also encouraged them to reflect on the implications of their analysis. We gave them the suggestion of talking about counter data and who this analysis could benefit and harm, which many of them did. This scaffold was included in the lab to help the students engage in CA and social justice awareness. Writing about the data science that they did and thinking about the big picture analysis helped them engage in CA because they also engaged in data science practices. Real data scientists do this often in reports. Also, by thinking about the implications of their analysis, they were able to engage in social justice awareness.

IR Question 4 had two themes that were the most common in the responses to this question. The first theme that was the most common was Theme 2, Social Justice. Many of the students mentioned social justice ideas in their responses to this question. Specifically, they

talked about the difference in salaries between men and women. They also talked about who this data could benefit, who it could harm, and that this data is important to collect. Table 39 shows what Jordan from Group 1 and Alex from Group 2 said for part of IR Question 4.

Table 39

Examples of Responses to IR Question 4

Student	IR Question 4 Response
Jordan	Collecting this salary data helps workers and may harm the university. The reasoning for this is that employers historically benefit from workers not knowing each other's salaries because it gives them extra bargaining power.
Alex	It allows people to compare their salaries and determine whether there is some sort of injustice. If not, it assures people that injustice between salaries does not exist; if so, it allows people to advocate for themselves or others. This is ever present in the data regarding gender.

Jordan had an interesting perspective about the salary data being public. They mentioned that this analysis could benefit workers and harm the employer, which in this case was a university. Jordan said this could harm the employer because the employees could discuss their salaries and have more bargaining power for higher salaries. Since Jordan thought about the implications of their analysis, they engaged in CA and social justice awareness. Alex also talked about how this data analysis can benefit people, specifically those who may be being discriminated against. Alex argued that making this data public and available for analysis allows people to see how they compare to their peers. Alex also mentioned the difference between men salaries and women salaries, hence there was evidence that Alex engaged in social justice awareness. By thinking about the implications of their analysis in the real world, Alex also engaged in CA.

The second theme that was the most common was Theme 3, Thinking About Mobilizing for Change. Some of the students took the social justice aspect of the labs beyond the classroom

and talked about collecting more data, advocating for change, and the implications of collecting counter data. Table 40 shows what Rowen from Group 3 and Alex from Group 2 said for part of IR Question 4.

Table 40

Examples of Responses to IR Question 4

Student	IR Question 4 Response
Rowen	Some counter data that I would like to collect is the amount of time a professor is spending at the university, the research that they conduct, and the amount of time that they spend doing research, and the time they actually spend teaching classes/what classes that they teach. It is important to collect this or attempt to collect this data because it gives a reason as to why certain professors are making more money than others.
Alex	If it is proven the gap in salaries is based on gender, people can advocate for this to change. This could push the administration to make changes because the analysis shows it's not fair.

Rowen mentioned that they would like to collect more data, specifically about the employees and the amount of time they are spending on work, what research they are doing, and what classes they teach. Rowen said they wanted to collect this data because it would add to our understanding of why certain employees have high salaries than others. Thinking about this as a real issue, rather than just an assignment helped the students think about future work they could do as data scientists, and therefore, they engaged in CA. Alex's response was an example of a student talking about the implications of the analysis and how people who are being discriminated against can use this data to advocate for themselves and get the administration to make changes. Thinking about how people can use data science to advocate for change helped the students engage in both CA and social justice awareness. Some students did point out that although the men had higher salaries than the women, there are other factors that should be considered when making this argument. None of them denied that the men had higher salaries

and that this was problematic, however they did mention that looking at things like tenure and job titles could help give a clearer picture. Overall, they had a variety of reasons that they thought the employees could benefit from the analysis or be harmed by it, and by acknowledging these reasons, they engaged in CA and social justice awareness.

Summary of Lab A's Thematic Analysis. The IR Questions from Lab A allowed the students to reflect on the analysis that they did, the implications of this analysis, and how it connected to the world outside the classroom. The students' responses to the IR Questions showed evidence that the lab helped them engage in both CA and social justice awareness. There was also evidence through their responses that the students engaged in data science practices throughout the lab. This included creating visualizations, communicating about data science through writing and talking, and answering real questions that they had. Engaging in data science practices helped them engage in both CA and social justice awareness and the IR Questions allowed them to communicate about this in writing.

Thematic Analysis for Lab B

For the Thematic Analysis, I identified five themes that were common among all of the students who participated in the study and complete Lab B. I described each theme in detail and explained which codes went into each theme. Next, I described how each theme connects to the two goals of the lab (CA and social justice awareness). And lastly, I looked at each individual reflection question in the lab and discussed which themes were most common in the responses to each question and I gave examples of what students said.

Theme 1: Possible vs. Probable. Many students talked about the idea of something being possible and how that was different from it being probable in their individual reflection question responses. Technically, even with random selection, any results are possible, even

unlikely ones. However, the further you get from the expected value, the lower the probability of obtaining those results. Although the probability of obtaining these results is low, it is never impossible. This is an idea from data science that is important for data scientists to understand. A data science practice is to use data science tools to determine whether or not something is likely to occur, while understanding that technically, with any element of randomness, anything is possible. Student responses varied in whether or not they acknowledged this and still came to the correct conclusion or whether or not they used this as a reasoning for coming up with the wrong conclusion. Because this was such a specific idea that came up in almost every group, the code was the theme. Sometimes in Thematic Analysis, a single code can become its own theme. It is also important to note that none of the scaffolds specifically asked about this concept (see the Appendix for a list of all of the scaffolds included in both labs), but it did come up in many of the responses. By acknowledging this idea and understanding the difference between something being probable and possible, the students exhibited CA. This is a key concept in data science and these IR questions allowed us to see how well students understood this concept. Understanding data science concepts and being able to articulate them in writing is one way that students can engage in CA.

Theme 2: Social Justice. Social justice awareness was also a goal of both labs, and it is very related to CA. I categorized any code that showed that students acknowledged that there was a social justice issue as this theme. The codes that created this theme were codes where the students realized that these problems exist, realized that they are problematic, and realized that certain people are ignoring them. Also, thinking about the issues in any way that did not involve data science was put into this theme, since most of these issues involved something related to racism or discrimination. This theme is clearly connected to the students engaging in social

justice awareness. Also, having students think about issues outside of the classroom involving social justice helped them to engage in CA.

Theme 3: Thinking About Mobilizing for Change. This theme takes CA and social justice awareness a step further. It came up in Lab A and was also relevant for Lab B. For this theme, the students talked about ways they could fix social problems and how they can make changes to address these problems. Specifically in this lab, they talked about how people in positions of power can make the changes. The codes that fell into this theme were codes that acknowledged that something needed to change and codes that gave suggestions for fixing the problems. This theme is clearly connected to the students wrestling with social justice. Also, by having students think about solutions to problems, rather than just acknowledging that there is a problem, there was evidence that they demonstrated CA. It is important to note that although there was evidence from this theme that some students thought about ways to mobilize for change, there was no evidence that they actually did any of the things they thought about.

Theme 4: Data Science Claims. A common idea that appeared in the responses to the individual reflection questions was students using representations or visualizations that they created to make a claim. For example, many students referenced simulations that they did or visual displays of data that they created (histograms and boxplots) when making a claim about an issue. These representations that they created helped them write about both data science concepts and the social justice issues that the labs discussed. This theme involved using artifacts that were internal to the labs to help make a claim about something. The codes that fell into this theme were codes that specifically looked at using something the students created or data from the scaffold to make a claim about the issue in the labs. This theme is connected to CA because the students are using the tools that they create in the lab to help make a claim about a real-world

issue. Because the labs were related to social justice, this theme is also connected to social justice awareness since the claims students are making are also related to social justice.

Theme 5: Data Science and the Outside World. Another common theme was students making the connection between the work they were doing in the labs and life outside of the classroom. Specifically in this lab, students talked about how they could use data science to address racism. In Lab B, the students looked at two social justice examples and got practice using data science to help identify problems and reflect on them. For this theme, the students used what they did in the classroom to think about how they could apply it beyond the examples presented in the labs. This theme involved thinking externally about issues beyond what they did in the lab. The codes that created this theme were codes that specifically looked at using data science to address issues of racism. Students often did this by giving examples of racism and describing how data science could be used to showcase these examples. This theme is connected to CA because the students made the connection between the work they did inside the classroom and the world outside of the classroom. This is a large part of CA. Because the labs are related to social justice, this theme is also connected to social justice awareness since the students thought about how data science is related to other social justice issues.

Theme 6: Data Science Concepts. Theme 6 was discovered through the reliability coding during the analysis. This theme illustrates the idea of students explaining data science concepts. As part of their answers, the students sometimes explained statistical or programming concepts as a part of their written reflection questions. The codes that were included in this theme were codes where students explained different concepts in their own words. For example, they would use the data given in the scaffold to explain the reasoning behind data science concepts. This theme is connected to CA because the students got practice explaining data

science concepts and ideas in writing. This is something that data scientists do all of the time. Oftentimes they explain concepts to business partners or people at their place of employment who are not data scientists. This helps emphasize students' understanding of concepts by being able to explain them in writing.

Individual Reflection Questions in Lab B. Next, I looked at each IR question from Lab B and talked about which themes came from the responses to each of these questions. I also gave examples of specific responses related to each theme. In the descriptions in the following sections, I highlighted the most common themes for each question. For most common themes, I picked themes that showed up in more than five students' IR responses. Table 41 shows each individual reflection question and a count of which themes showed up in the written responses to those questions from the students.

Table 41

Lab B Individual Reflection (IR) Questions and Themes

IR Question	Theme 1- Possible vs. Probable	Theme 2- Social Justice	Theme 3- Mobilizing for Change	Theme 4- Data Science Claims	Theme 5- Connection to Real World	Theme 6- Explaining DS Concepts
IR Question 1: Write a few sentences summarizing what your group members said during your discussion. Did people think this difference could have been due to chance or not?	7	9	2	7		1
IR Question 2: Write down the most interesting part of your group discussion.	5	2		13		
IR Question 3: How did your results change? What does this				4		12

Table 41 (cont.)

tell us about the more simulations you run?						
IR Question 4: Give an example of a way that we can use data science to help address issues of racism. This can be something you discussed in your group or an example you are interested in.			1		15	2
IR Question 5: Write down something that surprised you from your group discussion.	4	10		4		2
IR Question 6: Think about how these percentages (that they just calculated) compare to the actual percentages (10.33% vs. 6.87%). What does this say about discrimination based on your name?	2	6		9		
IR Question 7: Option 1: Pretend you are a data scientist arguing whether or not there was discrimination based on how applicants' names sound. Write a memo to the HR Department of one of the companies positioning yourself as a data scientists arguing whether employers are biased against certain names. Justify your decision using what you've already done and include guidelines for the future.	2	8	9	10		4

Table 41 (cont.)

Option 2: Pretend that you are a defense attorney and a data scientist. Write a memo to the Supreme Court positioning yourself as a data scientists arguing whether or not you think the jury with 3 Black men was randomly selected. Justify your decision and include guidelines for the future.

In the descriptions below, I highlighted the most common themes for each question. For IR Question 1, Themes 1, 2, and 4 were the most common, for IR Question 2, Theme 4 was the most common, for IR Question 3, Theme 6 was the most common, for IR Question 4, Theme 5 was the most common, for IR Question 5, Theme 2 was the most common, for IR Question 6, Themes 2 and 4 were the most common, and for IR Question 7, Themes 2, 3, and 4 were the most common.

IR Question 1. IR Question 1 asked the students to summarize what they said in a group discussion. This question referred to a discussion about the trial data and whether or not getting three Black people on the jury could have been due to chance or not. The students were asked to summarize what they learned from their group discussions. This scaffold was included in the lab to help the students develop CA and social justice awareness. If they understood the jury simulation correctly, they should have realized from their analysis that the jury was probably not chosen randomly. This is a social justice example that allows the students to use the simulation and the histogram that they created to explain why the jury likely was not randomly chosen. Using representations that they created using data science is a data science practice and engaging in this data science practice helped them develop CA. Also, by summarizing what was said in a

group discussion in writing, the students also engaged in CA since they got practice communicating about data science concepts.

There were three themes that appeared most often in the responses for IR Question 1. The first theme that appeared a lot for IR Question 1 was Theme 1: Possible vs. Probable. The idea behind IR Question 1 was that the jury selection was not fair. Random selection should result in more than three Black people on the jury. A lot of groups understood this idea and also pointed out that while getting three Black people on the jury is not probable, it is possible. The distinction between probable and possible is a key data science concept that this scaffold highlighted. By being able to articulate this in writing, the students engaged in CA because they got practice communicating about data science which is something data scientists do often. Table 42 shows what Jayden from Group 4 and Bailey from Group 5 said for part of IR Question 1.

Table 42

Examples of Responses to IR Question 1

Student	IR Question 1 Response
Jayden	The most interesting part of our group discussion was that we all had pretty low chances that only 3 black jury members would be part of the jury, but it is possible.
Bailey	We also discussed that, while there is the possibility that this could have been due to chance, we thought that it was unlikely, especially since no Black people made it to the final 37.

Jayden's response showed that they understood that the probability of getting three Black people on the jury was low, yet they acknowledged that it technically was possible. Bailey also mentioned the same idea. They understood that there was a possibility that this could happen, but it was very unlikely. It is also important to note that there were some students who did not understand the difference between possible and probable. They thought that because getting

three Black people on the jury was possible, that mean the jury was fair. They misunderstood the idea that just because something is possible does not necessarily mean it was probable.

The second theme that appeared the most in the responses for IR Question 1 was Theme 2: Social Justice. The students realized that the jury selection most likely was not random and that the difference between what they observed and expected was not due to chance. Many of them recognized that this is problematic, hence they engaged in social justice awareness. Table 43 shows what Dakota from Group 1 and Cameron from Group 3 said for part of IR Question 1.

Table 43

Examples of Responses to IR Question 1

Student	IR Question 1 Response
Dakota	The panel has roughly 8 black people, so this provides evidence that it isn't fair. I don't think this happened by chance and people are choosing to not give this too much attention.
Cameron	Given that 0 of the 100 randomly selected people made it into the final 37 considered for Smith's trial, and that Smith himself was a black man, it seems more likely that there were other factors at play.

Dakota's response showed that they recognized that there was a problem that the jury was not fair and acknowledged that people are ignoring it. Cameron also said that because the defendant was Black, there were likely to be other reasons why the jury had so few Black people, such as discrimination. Both of these students' responses are evidence that they engaged in social justice awareness from IR Question 1.

The last theme that appeared the most in the responses for IR Question 1 was Theme 4: Data Science Claims. For this question, the students used the representations that they created using data science to make claims about whether or not the jury selection was fair. Many of them used the histogram they created or the simulation that they did as justifications to claim that the jury selection was not random. Table 44 shows what Jordan from Group 4 and Rowen from

Group 6 said for part of IR Question 1.

Table 44

Examples of Responses to IR Question 1

Student	IR Question 1 Response
Jordan	Since the expected value of black people in the jury was supposed to be 8, but ended up being 3, we discussed that the expected value was almost 3x the actual value, which seemed a little off to us.
Rowen	This case absolutely needs to be revisited. By analyzing the data (real numbers!) involved in this case, it is clear that Smith should have at least 8 people to truly represent and align with the population of Black people in the United States. It is crucial that this case is revisited on a national (Supreme Court) level with a fair representation in the jury.

Jordan had calculated the expected value from the simulation to be eight and in their response, they compared that to the value they observed which was three. They noted how different these numbers were as evidence that the jury was not chosen fairly. Here Jordan used the data from the lab and the simulation to make a claim. Rowen also mentioned analyzing the data to make the claim that this case should be revisited since there is strong evidence that the jury selection was not fair.

IR Question 2. IR Question 2 asked the students to write down the most interesting part of their group discussion. This question referred to a discussion about the trial data and whether or not getting three Black men on the jury could have been due to chance or not. The students were asked to recall the part of their discussion that they found most interesting. This scaffold was included in the lab to help the students engage in CA. It encouraged them to recall their group discussions and summarize what they discussed about this social justice issue. By summarizing what was said in a group discussion in writing, the students also engaged in CA since they got practice communicating about data science concepts.

The theme that appeared the most often in the responses for IR Question 2 was Theme 4: Data Science Claims. For this question, the students used the representations that they created using data science to make claims about whether or not the jury selection was fair. Many of them used the histogram they created or the simulation that they did as justifications to claim that the jury selection was not random. Table 45 shows what Harper from Group 1 and Sawyer from Group 2 said for part of IR Question 2. I also included an interesting response from Marley from Group 1.

Table 45

Examples of Responses to IR Question 2

Student	IR Question 2 Response
Harper	The histogram most definitely seems to go in favor of the claim that there was some foul play or discrimination at work, as the histogram shows that majority of the time it is three or more being drawn rather than three or less.
Sawyer	In the histogram, the occurrence of the jury panel having one to three people is rare, the vast majority of the time the number would five to 10 range.
Marley	It was interesting how we all had slightly different graphs, but still the same overall takeaway. It is also interesting how data is able to allow us to draw conclusions as to how the jury selection was most likely a deliberate choice. However, the debate of race and justice has always been a touchy subject.

Harper used the histogram to show to that getting three or less Black people was very rare. In other words, Harper used a representation they created to make a claim. Sawyer also did this by mentioning that getting one to three Black people was rare and that they could see this from the histogram. Sawyer also pointed out that the histogram showed that the expected value should be between five and ten Black people. Using these representations that they created to make a claim helped them engage in CA. Marley also had some interesting comments for IR Question 2. They were the only student that acknowledge that each of their histograms were different because simulations were random. This is an advanced data science concept that I was

impressed that this student understood. Marley also acknowledged that it can be uncomfortable to talk about social justice issues like the one in this lab.

IR Question 3. IR Question 3 asked the students how the results of their simulation changed the more times they ran the simulation. This question was referring to a coding question that asked the students to simulate picking juries 100 times and 10,000 times. The students were supposed to observe what happens when you increase how many simulations you do. This scaffold was included in the lab to help the students understand an important data science concept which says that the more simulations you run, the closer your average is to the expected value. In other words, you can get more accurate results and less error when running more simulations. By recognizing and understanding this important concept, the students engaged in CA. Knowing how to get less error also helps the students make stronger arguments.

The theme that appeared the most often in the responses for IR Question 3 was Theme 6: Explaining Data Science Concepts. For this question, the students explained common data science concepts in their answer. Most of the students explained that the more simulations you run, the closer your mean is to your expected value. Table 46 shows what Sawyer from Group 2, Cameron from Group 3, and Jayden from Group 4 said for part of IR Question 3.

Table 46

Examples of Responses to IR Question 3

Student	IR Question 3 Response
Sawyer	The number for the mean got closer to 8, the expected number. This tells me that the more simulations you run, the closer to the expected value you get.
Cameron	This tells us that with the more simulations you run, the closer to reality your results become. This is because your results average out, and outliers have less of an impact on the end result.
Jayden	The first mean was 8.27 and the second mean was 8.0266. This tells us that as the more simulations we run, the closer we get to our expected value.

Sawyer, Cameron, and Jayden were able to understand this data science concept that says that the more simulations you run, the mean gets closer and closer to the expected value. Each of the three students in Table 46 were able to express this concept in writing. Cameron included the idea of outliers in their explanation. They said that outliers have less of an impact the more simulations you run. Jayden gave the actual numbers that they got from their simulation to explain the idea that the mean gets closer and closer to the true expected value the more simulations you run. Understanding this key idea and building simulations to show it helped the students engage in CA.

IR Question 4. IR Question 4 asked the students to give an example of how data science can be used to address racism. The goal of this question was to allow students to take what they learned in the lab and apply it to a situation outside of the lab. The lab involved using simulation to address a specific social justice issue. This scaffold had them think of another example where data science in general can be used to address racism specifically. This scaffold was included in the lab to help the students engage in CA and social justice awareness. By using what they did in lab and applying it to a situation outside of the lab, they engaged in CA. Because the question specifically asked about racism and data science, they also engaged in social justice awareness. Being able to articulate their ideas in writing is also another way students can engage in CA because this is something that data scientists do regularly.

The theme that appeared the most often in the responses for IR Question 4 was Theme 5: Connection to the Real World. This theme emphasized the idea of making a connection between what the students did in the classroom and the outside world. Many of the students gave specific examples of how data science can be used to address racism which were motivated by the

examples in the labs. Table 47 shows what Sawyer from Group 2, Jude from Group 3, Bailey from Group 4, and Jayden from Group 4 said for part of IR Question 4.

Table 47

Examples of Responses to IR Question 4

Student	IR Question 4 Response
Sawyer	An area where data science can be used is to prevent gerrymandering in states where it often put minority groups at a disadvantage. Data science can be used to draw more equal districts and promote greater voting access.
Jude	You can prove or fact check claims that people make when they say they want to make things more inclusive. For example, if a school says they admit a certain number of black students, you can fact check that claim using data science.
Bailey	We can use data science and statistics to look into college or private school admissions. We can look into seeing whether minorities have less of a chance of getting in or why or whether or not any of that is true. We can look into whether or not some places are letting in certain people based on race and not letting others in even though they have better credentials due to trying to artificially create diversity.
Jayden	I never thought about ways that we can use data science for this use, but I think that it's great that we are able to do so.

Sawyer gave an example of how we can use data science to address racism (gerrymandering), but Sawyer did not say exactly how data science can be used to draw more equal districts and promote greater voting access. Jude gave a little bit more detail about how you can check claims using data science, similarly to how they did in the lab. Bailey also gave a very detailed answer for what data we could use and a problem we could investigate. They talked about looking at college and private school admissions and checking if certain races have a disadvantage or a lower probability of getting in. Lastly, I thought Jayden's comment was very interesting because they acknowledged that the lab gave them a new perspective on how data science is applicable to social justice. All of these students saw the connection between data science and different social justice issues, hence there was evidence that they engaged in social

justice awareness after answering this question.

IR Question 5. IR Question 5 asked the students to write down something that surprised them from their group discussions. Through this question, students discussed whether or not they thought that the difference in percentage of callbacks for white sounding names vs. Black sounding names was significantly different. The goal of this question was to allow students to reflect on another social justice issue and use data science to back up their claims. This scaffold was included in the lab to help the students engaged in CA and social justice awareness. Because the students analyzed and reflected on another social justice issue, there is evidence that they engaged in CA because they saw another example of how data science can be applied to situations outside of the classroom. Because the scenario they reflected on was another social justice issue, they also engaged in social justice awareness. Being able to articulate their ideas in writing is also another way students can engage in CA because this is something that data scientists do regularly.

The theme that appeared the most often in the responses for IR Question 5 was Theme 2: Social Justice. Here, the most common theme was the students engaging in social justice awareness. As the students answered this question, they included explanations that were related to data science and unrelated to data science about why this happened and many of them wrote about the implications of the problem. Table 48 shows what Cameron from Group 3 and Rowen from Group 6 said for part of IR Question 5.

Table 48

Examples of Responses to IR Question 5

Student	IR Question 5 Response
Cameron	In my personal opinion, this may have happened for similar reasons as those that likely had a hand in the previous scenario, which was racial biases. This is problematic because there should be very miniscule to no differences in

Table 48 (cont.)

	identical resumes between White and Black people; since there is a noticeable difference, it is likely that in the real world, Black people with resumes on par with their White competitors are less likely to be hired, which should not be the case.
Rowen	I don't think this was due to chance because the resumes are identical, and the percentage change was extreme. My guess is to what happened was some sort of internal bias. This is extremely problematic because the people "black" sounding names are put at immediate disadvantage.

Multiple students mentioned that they had heard of the resume study before or heard of similar events occurring, however they had not analyzed data regarding this study. Cameron realized that there was a significant difference in callback rates and discussed how this is problematic. They said that people who have the same qualifications could get less callbacks just because they are Black, which is problematic. Cameron acknowledged that there are racial biases that are the root of this issue. Rowen also realized that this problem was occurring from the data and the simulation. Rowen also acknowledged that the people with Black sounding names are at a disadvantage and that this is problematic. By thinking about the analysis that they did using it to make claims about social justice issues, the students engaged in social justice awareness.

IR Question 6. IR Question 6 asked the students to think about the difference in percentages of callbacks for Black sounding names and white sounding names. They were instructed to write about what this means for discrimination based on your name. With IR Question 6, the students reflected on a coding simulation that they did. Assuming the resumes were all equally qualified, the students simulated randomly picking resumes for callbacks. Their simulation showed that the percentage of Black and white sounding names that received callbacks should have been very close. They were then supposed to compare those expected values from the simulation to the actual values and see that it is a large difference. Lastly, they were supposed to reflect on what this means. This scaffold was included in the lab to help the

students engage in CA and social justice awareness. By building a simulation to see what should have happened and then comparing that to what actually happened, the students are learning how to use data science to explore cases of discrimination, which helps them engage in both CA and social justice awareness. By explaining what they did and reflecting on the implications of this analysis in writing, the students also engaged in CA since data scientists do this regularly.

There were two themes that appeared the most often in the responses for IR Question 6. The first was Theme 2: Social Justice. As the students answered this question, they included explanations that were related to data science and unrelated to data science about why the difference in callback percentages was so big. Many of them wrote about the implications of the problem. Table 49 shows what Marley and Rowen from Group 6 and Alex from Group 2 said for part of IR Question 6.

Table 49

Examples of Responses to IR Question 6

Student	IR Question 6 Response
Marley	This shows that there was an intentional bias against black individuals over a very serious matter.
Alex	The percentage of white sounding names who got call backs is close to the population value, so they are accurately represented. However, the group with the black sounding names were extremely underrepresented, which indicates a high likelihood of discrimination.
Rowen	I think that these percentages are pointing towards the fact that name discrimination is a real thing. My name is _____ which would most likely not cause discrimination against me (but the possibility is still there because it is a feminine sounding name) but it might put me at an unfair advantage to those who don't have traditional "white sounding" names.

Marley used their simulation to show that there was a difference between Black and white sounding names. They believed this bias was intentional and pointed out that this was a serious problem. Alex realized that the white sounding names percentage was very close to what

you would expect, but the percentage of Black sounding names who got callback was a lot lower. Alex said this shows there was discrimination. These responses were evidence that the students engaged in both CA and social justice awareness. One student, Rowen, mentioned that their name is a typical white sounding female name. They said they may be at a disadvantage for having a female name but may have more of an advantage than someone who did not have a typical white sounding name. I thought this was interesting that the student thought about their own name and how this situation would affect them.

The second theme that appeared the most often in the responses for IR Question 6 was Theme 4: Data Science Claims. For this question, the students used the data science that they did to make claims about whether or not the difference in callbacks could have been due to chance. Many of them used the simulation they created analysis they did to justify the claim that the difference was significant. Table 50 shows what Dakota from Group 1 and Marley from Group 6 said for part of IR Question 6.

Table 50

Examples of Responses to IR Question 6

Student	IR Question 6 Response
Dakota	These percentages were a lot closer than the actual percentages. According to the simulation data they are about equal to get a call back.
Marley	The results from what I ran today was 10% for black sounding names and 9.9% for white sounding names which is a much closer percentage to one another. However, the actual percent of 10.33% and 6.87% has a very large gap.

Dakota recognized that the simulation percentages for callbacks for Black and white sounding names were very close to each other, and the actual percentages were quite different. They said that the simulation was evidence that the difference was significant. Marley also explained how their simulation showed something drastically different than what happened in

real life. Marley gave the exact percentages from the simulation and the problem. This scaffold and the responses showed evidence that the students engaged in both social justice awareness and CA.

IR Question 7. IR Question 7 was a longer IR question that put students in the position of an expert. They were supposed to write at least five sentences pretending they were a data scientist using their work to argue about one of the two social justice scenarios in the lab. The students were able to pick whether they wanted to talk about the jury scenario or the resumes and callbacks scenario. Regardless of which situation they chose to write about, they were instructed to justify their decision using the data science that they did and include guidelines for the future. This scaffold was included in the lab to help the students engage in CA and social justice awareness. Here, by acting as a real data scientist, they engaged in CA. They also wrote about social justice issues, so there was evidence that they engaged in social justice awareness. This question was included to allow students to get practice formally writing about the data science they did and using it as justifications to a claim. They were instructed to write a paragraph style response, rather than a few sentences like the previous IR Questions.

There were three themes that appeared most often in the responses for IR Question 7. The first theme that appeared a lot for IR Question 7 was Theme 2: Social Justice. The responses from this question were centered around social justice issues so this was a major theme. Table 51 shows what Jude from Group 3 and Sawyer from Group 2 said for part of IR Question 7.

Table 51

Examples of Responses to IR Question 7

Student	IR Question 7 Response
Jude	I have found that black sounding names appear to get a callback at a much lower rate than what they should be getting a call back. They should be receiving a call back about 9% of the time, but they do at about 6% of the time. The white sounding names get admitted at about the same rate as they should. This shows blatant racism on the part of your company.
Sawyer	Hello HR Department. It has recently been discovered that there is discrimination in the hiring process. There is discrimination based on the name of applicants. Only 6.87% of black sounding names got callbacks, when 10.33% of white sounding names got callbacks. Based on data science, if this was a fair process, it should be an even 10% for any name to get a call back. This is a problem that needs to be addressed as soon as possible.

Jude acknowledged that the Black sounding names getting callbacks at lower rates was a problem. They also said that this was blatant racism. Sawyer also acknowledged that there was discrimination from the resume example. They also said that this is a problem and that it needed to be addressed as soon as possible.

The second theme that appeared a lot for IR Question 7 was Theme 3: Thinking About Mobilizing for Change. In IR Question 7, a lot of students took social justice a step further and mentioned how to fix the problems they described or gave recommendations on how to fix these issues. Table 52 shows what Marley and Rowen from Group 6 said for part of IR Question 7.

Table 52

Examples of Responses to IR Question 7

Student	IR Question 7 Response
Marley	We urge you to look at this data and keep a 3rd party data scientist on staff to review decisions by. This scientist can help highlight patterns and large bias gaps that you might not be aware of to ultimately hold you accountable.
Rowen	This case absolutely needs to be revisited. By analyzing the data (real numbers!) involved in this case, it is clear that Smith should have at least 8 people to truly represent and align with the population of Black people in the United States. It is crucial that this case is revisited on a national (Supreme Court) level with a fair representation in the Jury.

Marley suggested that the company hire a data scientist to help hold the company accountable and Rowen recommend that the case is revisited, and that Smith gets a fair trial with a representative jury. Both of these students are recommending that something is done about this problem, hence thinking about mobilizing for change. IR Question 7 was also the only question that had a common theme of thinking about mobilizing for change.

The third theme that appeared a lot for IR Question 7 was Theme 4: Data Science Claims. The students used the representations that they created in the labs to make claims about the social justice issues. Table 53 shows what Avery from Group 3 and Bailey from Group 5 said for part of IR Question 7.

Table 53

Examples of Responses to IR Question 7

Student	IR Question 7 Response
Avery	Based off the simulations in multiple trials, there is a common theme of applicant with white sounding names having a higher chance of getting a call back.
Bailey	Through the research conducted, it is very difficult to conclude the disproportionately low percentage of black men represented in this jury was randomly selected. We performed two simulations. In the first, we repeated the simulation 100 times; in the second, we repeated the simulation 10000. Regardless of the size, it is very reasonable to conclude that there is a high inprobability [*sic*] of generating a random simulation with only 3 black men represented.

Avery said that their claim was based on the simulating the jury for multiple trials, hence referencing their simulation. Bailey used the simulation as the justification for why getting three Black people on the jury was not fair. Both of these students used the simulation as a justification for their claim that the jury was not randomly chosen.

Summary of Lab B's Thematic Analysis. Overall, the IR Questions from Lab B allowed the students to reflect on the analysis that they did, the implications of this analysis, and

how it connected to the world outside the classroom. IR Question 7 allowed the students to summarize what they learned from doing this analysis and helped them think of ways to use data science to advocate for change or make suggestions. The students' responses to the IR Questions showed evidence that the lab helped them engage in both CA and social justice awareness. There was also evidence through their responses that the students engaged in data science practices throughout the lab. This included creating representations of data such as simulations and histograms, communicating about data science through writing and summarizing those discussions, and answering real questions that they had. Engaging in data science practices helped them engage in both CA and social justice awareness and the IR Questions in Lab B helped them to communicate this through writing.

RQ2c Results

RQ2c asked: "*What do students perceive that they are learning through these labs that will be useful in the real world?"* To answer this question, I look at interview data from two of the questions that specifically asked students how they thought the content of the labs would help them in their future jobs and in their lives. This is a key idea from computational action. One important idea from CA is that students should be doing work inside of the classroom that is applicable to their lives outside of the classroom. To understand what students perceived that they are learning in the labs that will be useful in the real world, I used Thematic Analysis to identify four themes that were common among all of the students that we interviewed who each completed both labs. I describe each theme in detail and explain which codes went into each theme. After describing each theme, I unpack how each theme connects to the two main goals of the lab (CA and social justice awareness). And lastly, I give examples of what students said during the interviews to show what they said about each theme.

Theme 1: Data Science in the Future. One theme that came up during the student interviews was the idea that students felt that data science will help them in their jobs. Some of the students specifically mentioned that their career goal was to be a data scientist. Others mentioned that although their career is not going to be a data scientist, they did feel like learning data science would be useful for their specific career. Overall, the students thought that in general, data science would be useful for them in the future, both in their jobs and in their lives. The remaining three themes talk more specifically about what parts of data science will be useful. The codes that made up this theme were codes where students acknowledged that they want to do data science in the future, codes where students mentioned data science being directly related to their jobs, and codes where students mentioned that data science will be useful in their careers and their lives. Because students are recognizing the importance of data science in their lives and in their jobs, they are engaging in CA. They understand that being able to do data science is a valuable skill to have, not only for this class, but for them as citizens outside of the classroom.

Table 54 shows the two interview questions that I used to analyze these data. Both questions asked students whether the content of the labs will help them in their jobs and their lives. Both questions also asked them to explain their answers.

Table 54

Interview Questions for RQ2c

Question Number	Interview Questions
Question 1	Overall, do you think the content of the labs will help you in your future job? If so, how? If not, why not?
Question 2	Overall, do you think the content of the labs will help you in your life? If so, how? If not, why not?

The students often started their response to these questions by talking about the importance of data science and how they feel it is something that will be useful to them in the future. Table 55 shows two example responses from Rowen and Jude's answers to Question 1.

Table 55

Example Responses to Rowen and Jude's Question 1

Student	Response
Rowen	These labs would help me a lot because eventually, this is what I'm going to do. Building on the data set and then just analyzing it and learning different ways to clean it and you know, just make visualizations out of it. And I believe it's good that I took this course now because I feel more involved in the subject and more clear with what I want to do eventually.
Jude	Oh yeah, these labs will for sure be useful for me. Again, I'll be working at the department of justice and basically with my job, I'll literally just be touching on things like that all the time.

Rowen is saying that they want to do data science in the future, so this class is directly related to their career. The labs have helped them feel confident about their decision to become a data scientist and more comfortable working with data, specifically creating simple visualizations. Jude wants to work at the Department of Justice and sees a direction connection between what they are doing in the labs and in their future job.

Theme 2: Technical Skills. One way that students mentioned the content of the labs being useful outside of the classroom was that the labs help the students gain various technical skills that are important. They specifically mentioned the importance of being able to visualize data and create simple data visualizations. They also mentioned using real world tools like Python, Python libraries, and Jupyter. Some students talked about how different statistical concepts are important as well. The codes that created this theme included codes where the students talked about the importance of data visualization, the importance of learning statistics

and different statistical concepts, as well as the importance of learning how to use real world tools like Python, Python libraries, and Jupyter. In general, students also talked about how being able to analyze data and present that analysis is something that will be important in their lives outside of this class. This theme is directly connected to the students engaging in CA. By having students think about the importance of the technical skills they have learned in the labs and how that connects to their lives, they are engaging in CA. They are also using real world tools that data scientists use in their day to day lives, which is an important part of CA.

Many of the students talked about the importance of learning different technical skills and acknowledged that these skills will help them in their lives. Table 56 below shows two example responses from Harper, Cameron, and Jayden's answers to Question 1.

Table 56

Three Example Reponses to Question 1

Student	Response
Harper	I think it will help me in my job. I think in general I've gotten a lot better at Python through this class. And I feel like, it's kind of a lot of places where you can use Python. And then I also feel like a lot of the stat stuff is pretty useful, because it's not like super difficult, but it's like, stuff that's you can apply it to like a lot of different areas.
Cameron	They definitely will. A lot of C+ and Java applications have these requirements, such as knowing Python, and specifically like learning pandas as one of their libraries. So I feel like as a one-semester course, there was a lot of experience with pandas library, and also a little bit with, data visualizations. So it's definitely very helpful for future jobs.
Jayden	I think it will because in my opinion, whatever I end up doing, it programming will be involved because that's where we are heading in the future with technology.

Harper said they have gotten better at Python, which is something that they will use outside of the classroom. They also mentioned that the statistics concepts are useful as well. By realizing that Python and statistics are applicable to different fields and useful outside of the

classroom, the students are engaging in CA. Cameron mentioned that knowing Python, pandas, and data visualization is important for future jobs. They specifically mentioned more complex computer science applications and how knowing the things they learned through this class will help them with that. By seeing the connection between what we are doing in class and their jobs, they are engaging in CA. Lastly, Jayden believes that regardless of what their career ends up being, programming will be helpful because that is where technology is headed in the future. By recognizing that programming is an important skill that could be useful to any career, they are engaging in CA.

Theme 3: Social Justice. I categorized any code that showed that students acknowledged any aspect of social justice issues or mentioned social justice as this theme. Social justice awareness is also a goal of both labs, and it is very related to CA. The codes that created this theme included codes where the students mentioned social justice or any social justice issue. I also included codes that referenced students questioning the status quo as codes that fell into the social justice theme. This theme is clearly connected to the students engaging in social justice awareness. Also, by having students think about issues outside of the classroom involving social justice and data science, they are engaging in CA.

Many did mention the idea of social justice in their answers, and they did this in various ways and with various examples. Table 57 below shows two example responses from Cameron, Avery, and Bailey's answers to Question 2.

Table 57

Three Students' Response to Question 2

Student	Response
Cameron	I think if we're just talking about like, the two social justice labs, I think, yes. Because they make me more aware of the inequality that exists in CS and stats and things that I'm not really familiar with. And also like the political justice system and how it's really skewed and biased, and just like kind of making me more aware of the bias that are in certain aspects of society.
Avery	I can't say that I will definitely have to prove that having less than three people on a jury that are black will ever come up in my actual life. But learning how to think about data science and injustice, and learning how to think about how I should examine a data set and what I should be looking to find in my examination of a data set will definitely be something I will use.
Bailey	If I ever want to investigate something like, you know, like a claim a company makes, for example, about like, we hire 50% women or something. You could check that if you want.

Cameron is saying that the two labs in particular helped them become more aware of bias and inequality. By acknowledging that the labs helped them see this, they are engaging in social justice awareness. Avery's comment was interesting because they acknowledged that they probably will not have to do exactly what they did in the lab in real life. However, the lab did help them think about the connection between data science and injustice and that they seem themselves using the skills they learned in the lab outside of class even if it's not in an identical scenario. Being able to acknowledge this and verbalize it shows that they are engaging in social justice awareness and CA. Bailey that they now know how to check claims that companies make because of the labs. This is a clear extension of what they did in the labs. Seeing this connection between examples they did in lab and similar situations in the real world shows they are engaging in social justice awareness and CA.

Theme 4: Problem Solving. This theme looks at how the labs teach students how to approach problems differently and using problem solving skills to answer questions. Both of these ideas are useful both inside the classroom and outside of the classroom. This theme was the

least common theme, but it did come up a few times and had important ideas. The codes that made up this theme were codes where students mentioned that the labs helped them to approach problems different or anything about improved problem solving. This theme is directly connected to engaging in CA because problem solving skills are important in all careers and in students' lives as they navigate issues.

Some students mentioned that the labs helped them see problems in a different way and think about issues in a way that they had not before. Table 58 below shows two example responses from Avery's answer to Question 2 and Jude's answer to Question 1

Table 58

Example Responses from Avery and Jude for Question 1

Student	Response
Avery	I don't think my bosses are going to be like "okay create a simulation about a trial or social justice issue." But I think these labs in a different way, do teach you how to approach problems differently and actually use data science to solve them. I think my brain got used to a different way of thinking. Instead of being in an advertising class where they are like "memorize this and that and spit it back out to me", I was actually able to practice problem solving so I think I'm going to use that in the future no matter what my job is.
Jude	Actually, yes, I can think of like a few instances where people have been like, oh, I don't think that's true. And like now, if somebody says something that I think is wrong, I'm comfortable saying that is not true because I now know how to think about claims in a way that I didn't before. And even if it is true, like you don't have a source for it.

Avery said that the problems in the labs are different from problems in traditional math classes where you have to memorize something and then reproduce it on an assessment. The labs allowed them to practice problem solving skills using data science, which is a skill that is important in many careers. Recognizing this shows that Avery is engaging in CA. Avery previously mentioned a similar idea that the exact problems that they did may not come up outside of class, but the skills they learned while solving those problems in lab can help them

when solving similar problems outside of class. Jude said that now they approach claims by other people differently. They want to make sure that people have data to back up their claims. This is one way that the labs helped change their way of thinking and one way that they have engaged in CA.

Summary of Thematic Analysis. Overall, the students enjoyed the labs and thought that they were useful. They described how doing data science in school can be useful in the future because it provides them with technical skills, such as visualization, and helps them get practice problem solving. The labs also help the students engage in social justice awareness because they are exploring multiple social justice issues in the labs. The students understand that data science can be a useful tool for examining social justice issues. Despite the students thinking that data science is useful, there were not many specific examples of how the students can directly use what they learned from the labs in their jobs. It may be too early for them to fully be able to answer this question since they have only done two labs. Perhaps if I revised all of the labs, they would be better equipped to answer this question.

There is also some bias with this interview question since we are asking them explicitly what will be useful so they may be saying what they think we want to hear. Because I was both the instructor of the course and the researcher, I had a research assistant do the interviews so that the students did not feel pressured to tell me what they thought I wanted to hear. In my experience, students are very honest, and I tried to give them multiple avenues to state their opinions through the anonymous surveys and the interviews with the research assistant. In the future, it would be interesting to do a similar study in another context where I am not the instructor of the course to lessen some of the power issues.

CHAPTER 6: DISCUSSION

In this curricular study, I examined how I could use DBR to create data science labs using principles of DC that helped students engage in CA and social justice awareness. I found that the students engaged in CA and social justice awareness by 1) using the scaffolds and artifacts in the labs designed using DC principles, 2) discussing and writing about important social justice issues, and 3) engaging in data science practices through coding, writing, and group discussion. In this section, I discuss how I make sense of the findings. I discuss how the DC framework played a role in the design of the labs. Next, I discuss the importance of the intersection between CA and social justice. Lastly, I describe five Design Principles that I created as a guide to creating labs using DC principles that help students engage in CA and social justice awareness.

DC Framework

Designing the labs using DC principles allowed for knowledge to be shared through the group discussions. In Lab A, some of the GD questions had the students do analysis that was specific to their home department or analysis that was based on a question of their choice that they were interested in. After doing these analyses, the students were instructed to then share the results with the rest of the group. In Lab B, the students answered GD questions that allowed them to share their own unique perspectives about social justice issues and how data science can be used to address racism. There was evidence that the students learned from each other, and that the knowledge was shared among the entire group in the GD questions. There was also reference to knowledge sharing in the individual reflection questions. Students wrote about the most interesting parts of their group discussions and wrote about what they learned from the group discussion questions.

The scaffolds (GD and IR questions) and artifacts (social justice datasets) that were included in the labs helped create an atmosphere in the labs where the knowledge was shared among the group members. Although the classroom and a ship like *Palau* are very different environments, the idea of knowledge being shared among different people and artifacts remains the same. Hutchins (1995) described how the navigation team worked together with different artifacts to accomplish the simple goal of moving the ship. Although the goal was simple, the process was complex, and it involved many people working together and being willing to help each other. Similarly, in the labs, the students worked together with the scaffolds, artifacts, and each other to accomplish the somewhat simple goal of completing the labs. However, the process of engaging in CA and social justice awareness is also complex, just like the process of navigation.

The goals require the students to work together and be willing to learn from each other. For example, when Group 5 discussed what the histograms that they created showed about the data in Lab B, Alex did not know what the *x* and *y* axes of the histogram represented. The other group members explained this to them. Also, when Group 3 discussed the analysis that they did in Lab B, Marley used the variance instead of the standard deviation to calculate the spread. Through Rowen's discussion with Marley, they realized that standard deviation would be the more appropriate statistic to measure spread in this case. These are two examples where the knowledge is distributed among the group members and students helped each other understand certain concepts. Hutchins (1995) described how crew members corrected each other and helped others understand the different perspectives of the different roles involved in navigation. This is similar to how the students in the group helped each other and were able to learn about different perspectives of social justice and data science. Students used the scaffolds, artifacts, and made

collective arguments with other people throughout the labs. The knowledge was distributed among the artifacts and the group members to help them engage in both CA and social justice awareness. Hence, using the DC framework was a valuable way to design the labs.

CA and Social Justice

The intersection of both goals of the labs (CA and social justice awareness) was very important in both the design of the labs and the student experience. My original hypothesis was that if I put these elements together, they would provoke student learning. A large part of CA involves allowing students to do authentic and meaningful work that is connected to their lives outside of the classroom. One way to do this is through social justice examples. My goal was to give students the space to see the connection between data science and social justice through working with real examples and discussing the implications of the analysis with their peers. I said earlier that I did not have this opportunity as a student, but that I am very passionate about giving my students this opportunity as an instructor. I did not want either of the goals to be an afterthought, instead I wanted them both to be important and through designing and implementing the labs, I discovered that they can be interconnected.

I saw that students were able to combine CA and social justice awareness when answering multiple types of scaffolds in the labs. Counter data is an idea from *Data Feminism* (D'ignazio & Klein, 2020) that involves looking at what data were not collected and how people could benefit from these data being collected. Many students mentioned the idea of collecting counter data in Lab A. By thinking about other data they could collect and the implications of collecting this data, the students engaged in both CA and social justice awareness. Safiya Noble (2018) mentioned the idea that it is important to have a data literate society. Engaging in CA and social justice awareness is one way that students can become data literate. By using data science

to examine social justice issues in Lab B, the students were able to engage in both CA and social justice awareness. For example, when the student reflected on whether the difference between callbacks of white sounding names and Black sounding names was significantly different, they used representations to show how the expected values differed from the observed values. By creating representations using a real-world example, the students engaged in CA. However, by discussing the implications of this analysis, they also engaged in social justice awareness. Many of the scaffolds allowed the students to meet both learning goals since engaging in social justice awareness through doing data science directly relates to engaging in CA.

The labs were designed so that the learning goals were for the students to engage in CA and social justice awareness. However, my design framework did not consider critical issues. In other words, it would be beneficial to use critical frameworks (e.g., critical race theory, feminist perspectives) for designing the prompts of the labs. Despite social justice awareness being a key learning goal in the labs, some students still came into the labs with prior biases and ideas regarding the social justice topics. I saw that in Lab B, some members from Group 2 disregarded the idea that the resumes were identical and discussed the possibility of the people with Black sounding names being less qualified for the jobs than those with white sounding names. Using a critical framework to design the scaffolds in the labs could help redirect those discussions and help students understand that this type of thinking can be problematic.

The frameworks that I have are insufficient to mediate the issues of social justice that come up in student discussions. Using other critical frameworks as lenses to create scaffolds and artifacts adds another layer that could be fundamental to this work. For example, Safiya Noble (2018) used a Black intersectional feminist approach in *Algorithms of Oppression* to examine how Google search algorithms perpetuate racism, specifically harm women of color. Also, Ibram

X. Kendi (2019) explored similar issues in his book *How to be an Antiracist*. Kendi (2019) talked about how racism is everywhere and that it is important for people not only to not be racist, but that they should also be antiracist and actively fighting against racism. He described how being antiracist means that people should be self-aware, critical, and constantly examining their actions and how they can improve. Using frameworks such as these to design the labs would help mediate the discussions that students have and encourage students to think beyond social justice awareness and explore how critical issues are connected to data science.

Design Principles

After designing these labs through the DBR process, I created a set of *design principles* for designing labs in a data science class that promote multimodal communication, contain principles from DC, and help students engage in CA and social justice awareness. Many of the ideas in these design principles are similar to ideas and practices involved in different educational fields, such as the learning sciences. For example, one of the principles involves giving students the agency to make their own decisions during the labs. This builds on the idea of *epistemic agency*, which says that students should be able to shape their learning and knowledge building in the classroom (Miller et al., 2018). Another one of the principles focuses on giving students a variety of different types of questions throughout the labs. This builds on the idea of multimodality in education, which says that different modes should be emphasized during learning to diversify instruction, such as images, written words, and spoken words (Leeuwen, 2015). By building from important ideas in education and analyzing the data in this study, I created these five design principles for data science labs. These design principles can be used by other instructors to create data science labs similar to the ones that I designed in this study that encourage students to engage in CA and social justice awareness.

Chieu et al. (2011) created a set of five design principles using DBR. These design principles can be used to "optimally exploit interactive rich-media technologies in the implementation of virtual settings, such as online learning environments and online communities of practice, for supporting teachers' learning to notice and interpret important aspects of instructional practice" (Chieu et al., 2011, p. 11). Moreno and Mayer (2000) also created and tested six instructional design principles for helping students understand scientific systems using multimedia. Similarly, I created five design principles from this DBR study that I describe in detail. Chieu et al. (2011) described their design principles as guidelines that facilitate making decisions while creating learning conditions. The design principles that I created are also guidelines that can help other instructors while creating their own labs. Table 59 describes the five design principles that result from this study.

Table 59

Design Principles

Principle	Description
Multimodal Communication	Allow students to engage in multimodal communication through coding questions, individual reflection questions, and group discussion questions throughout the labs
Discussion Reflection	Place group discussion questions directly before individual reflection questions in the labs
Exploration	Include scaffolds that have the students explore their own unique interests
Student Agency	Allow students the opportunity to make their own decisions in the labs
Expert Positioning	Include paragraph style individual reflection questions that put students in the position of an expert

Multimodal Communication Principle

The Multimodal Communication Principle involves having a variety of question types throughout the labs. Before this study, the labs in this class were longer coding exercises. They were more in depth than homework assignments, but essentially were just a series of coding

questions for students to complete during their lab section. The coding questions were the only types of questions that occurred throughout the lab. Although the students did help each other with these coding questions, the labs lacked reflection and group discussion about important issues outside of how to code. My intention was to redesign the labs to allow students to engage in multimodal communication. Although coding questions are important, I did not want them to be the only type of questions in the lab. Since the main goals of the labs are for students to engage in CA and social justice awareness, I wanted to add questions that allowed students to communicate. It is important for data scientists and citizens to be able to explain the work they have done to others. Because communication is multimodal, I thought about what other ways students can do this besides through coding questions.

The first method of communication that I thought of was talking. I specifically wanted these labs to include discussion among group members. Because I thought discussion was important, I created group discussion questions and inserted them throughout the labs. The group discussion questions asked the students to spend some time discussing topics with their group members. These topics included statistical topics, reasoning for making the decisions that they made, and some allowed the students to share something unique that they had done with their group. By talking with their group members, the students were acting as real data scientists, hence engaging in CA.

Oftentimes discussion can be left out of classes that involve computation. The students communicated by helping each other solve the coding questions, but the group discussion questions allowed them to discuss the implications of the scenarios and reflect. The students were able to share their ideas and learn from their classmates. This was an important part of the multimodal communication. In addition to group discussion questions, I also decided to include

written individual reflection questions where the students respond to prompts in writing. The individual reflection questions allowed students to get practice writing about the work they were doing, as well as writing about the group discussions they were having. They also allowed students the opportunity to think about the social justice issues they were exploring during the lab and reflect on the implications of the data science they were doing, and the scenarios that were presented.

Discussion Reflection Principle

The Discussion Reflection Principle involves the placement of the types of questions in the lab. I originally hypothesized that having individual reflection questions before group discussion questions would allow students to think about something individually before sharing their thoughts with other people. I imagined that this would prompt them to bring their unique thoughts to the group and that the groups would have more thorough discussions. To test this theory, I tried both orders. In Lab A, I put individual reflection questions before group discussion questions and in Lab B, I put the group discussion questions before the individual reflection questions. After analyzing the audio recordings of the group discussion and the individual reflection questions, I actually found the opposite of my original hypothesis to be true.

I saw that having group discussion questions first seemed to elicit more thorough discussions. The students were able to use their perspective of the group discussions to write their individual reflections. There seemed to be diversity within people's individual reflection questions even if they were in the same group. Overall, the students were more likely to write about their group discussions if I put the group discussion questions before the individual reflection questions. Doing this also helps the students prioritize the group discussion questions.

Because the individual reflection questions depend on the group discussions, this makes the students more likely to engage with and participate in the group discussion questions.

Exploration Principle

The Exploration Principle involves including scaffolds that have the students explore their own interests. For example, it is important to include scaffolds that allow the students to explore something unique that is related to their own interests and allows them to share their knowledge with the group. Allowing the students to make connections to their major or own experiences in the lab will help students understand how data science is useful in the real world. I included questions in Lab A and Lab B that allow them to explore their own majors and think about unique social justice issues that are related to data science. According to the interviews, this was one of their favorite parts of the labs. The students were very interested in their major and how data science can benefit them. Thinking about their own unique majors also helped with the group discussions because it gave them something unique to talk about. This allowed the students to share their knowledge with their group members and engage in computational identity by seeing themselves as an important part of the group with knowledge to share.

Student Agency Principle

The Student Agency Principle gives students the opportunity to make their own decision in the labs. In other words, it is important to include questions where the students can pick a question that they want to answer and use data science to solve it. Many assessments in math-related fields tell students exactly what to do and test on how well they can carry out those requests. In the labs, I added questions that had the students think of something that they wanted to know the answer to and use data science to answer it. This helped the students engage in CA because they were acting as real data scientists by coming up with a question that they were

interested in finding the answer to and using the knowledge that they learned from the course to answer it. These types of questions that allow students the opportunity to make their own decisions about what to look into promote student agency.

Allowing students to make choices about coding questions in the lab gives them practice making their own decisions, similarly to how they will have to do this in their jobs. To test this principle, I allowed them to do this in the labs and had interview questions regarding this. The students said they felt like what they were learning was important and relevant outside of the class when they could use what they learned in class to answer a real question that they had about the data. They mentioned that this was a unique part of the labs that they often do not get to do in other classes.

Expert Positioning Principle

The Expert Positioning Principle allows students to get practice communicating via writing through a paragraph style response reflecting on their labs in the position of an expert. I discovered that having these paragraph style responses really encouraged them to write something thoughtful about social justice issues and think about the implications of the problem they are solving. To test this, I initially only did this in Lab B but added it to Lab A' and will keep it in Lab B'. I saw that students in Lab A often wrote very short answers to the individual reflection questions. These specific paragraph style prompts allowed students to explore social justice issues from the perspective of an expert and produced much deeper and thought-out responses.

Summary of Principles

These design principles can help other instructors create programming labs that encourage multimodal communication. They were designed with the idea that they can help the

groups that need improvement for social justice awareness and CA. These were motivated by the elements of the labs that worked best. From Lab A, this included scaffolds where students had to do analysis on their own department and scaffolds where students could pick what question they wanted to answer. A few examples from Lab B were scaffolds where students were put in a position as an expert and scaffolds where students had to reflect on how data science can be used to address racism. Figure 30 shows the principles and how they relate to the learning goals in the lab.

Figure 30

Design Principles and Learning Goals

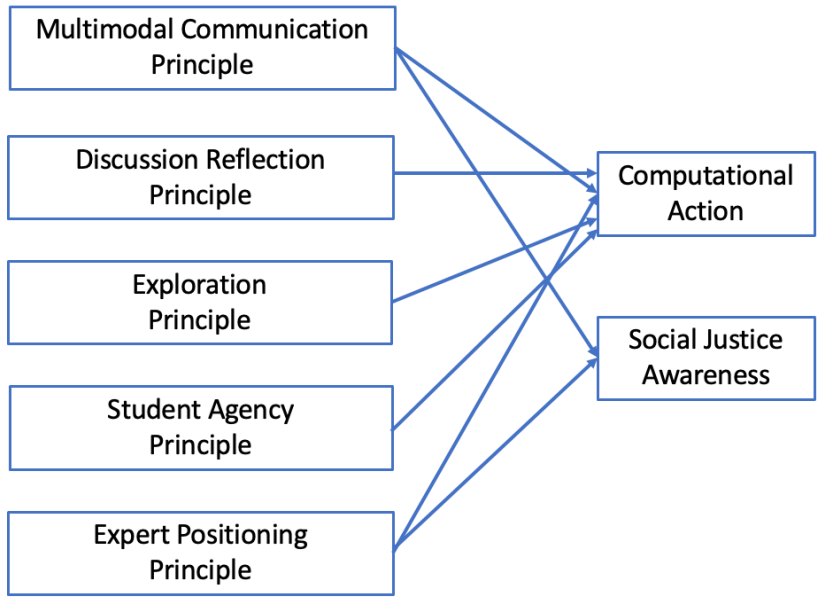


Figure 30 shows that the Multimodal Communication Principle helps students engage in CA and social justice awareness. By engaging in multimodal communication about social justice issues, the students act as real data scientists by talking and writing about the analysis they do. The Discussion Reflection Principle helps the students engage in CA because they can have discussions about their analysis and then reflect on it in writing. The Exploration Principle helps

the students engage in CA because they are able to explore their own interests and answer real questions that they have. The Student Agency Principle allows the students to make their own decisions in the labs, which also allows them to act as real data scientists and engage in CA. Lastly, the Expert Positioning Principle helps the students engage in CA because they are acting as an expert and this principle helps them engage in social justice awareness because the questions are centered around social justice issues. Together, these five design principles helped the students meet the learning goals in both of the labs.

Implications

This work has implications for research, implications for teaching, and implications for labs in general. For research, this study is a case of how we can use DC as a framework for studying labs. The DC framework works well for studying data science labs because the students work together with the artifacts and each other to solve problems in a natural environment. Knowledge is distributed among different parts of the system and designing the scaffolds to encourage this helps the students share knowledge and collaborate with each other. This study also adds to the literature on how we can use DC in education related research. Lastly, this study is evidence that DBR is a useful method to study how DC can be used in education related research. DBR allows for multiple cycles of implementation and allows researchers to work with educators to make improvements in how they design labs. Because DBR is iterative, we can change different design components based on student feedback and observations to help the students meet the goals.

This study also has implications for teaching. This work helped me as an instructor make meaningful changes in my classroom. This is one of the goals of DBR. The DC framework also allowed me to see how the students work together in the system to work with other students and

artifacts. This allowed me to design the scaffolds to help encourage communication and collaboration. As the instructor of this course, I was skeptical of including labs in this data science course at first. I was worried that students would rather just doing the programming exercises and labs on their own and that they would not sign up for the class if it was required to sign up for a lab section. I quickly realized that the labs are the most important part of the class, and that programming is very difficult and isolating to do on your own. The labs give the students a space to share knowledge with each other, work together, and learn from their peers. The labs also allow students to get practice communicating, which is a key part of data science.

Lastly, this work has implications for labs in general. There is not a lot of research looking at how to design data science labs or what happens in data science labs. This is because data science labs are new. The labs that I designed in this study are college labs that are a hybrid of science labs where students do experiments and mathematics/statistics labs where students get practice solving problems using software. These data science labs are a new case of statistics labs. The students are not only programming and solving problems, but they are getting practice communicating and reflecting on the work that they are doing. The labs allow them to communicate in writing through the IR questions and through talking in the GD questions. The design principles can help other instructors make labs that foster communication and encourage students to work together to solve problems.

Limitations

There were some limitations in this study that need to be addressed. With the current model, the TAs cannot fully facilitate good group discussions. TAs are not trained to encourage students to work together. The current role of the TAs is to answer student questions. However, because the TAs cannot facilitate group discussions, there may be some groups who do not have

good discussions. The scaffolds alone would have to encourage students to work together and have good discussions, which is not always the case. Also, self-guided labs are good up until a point, but students have prejudices and biases that sometimes make them disregard data science findings. For example, when Group 2 completed Lab B, they talked about how they thought that some people with Black sounding names may not be qualified for the jobs. They missed the point that the resumes were identical and got caught up in trying to think of a justification for why the difference in percentages was okay. This is just one lab and social justice needs to be a theme in other courses for there to be a larger impact on students.

Another limitation that came up in this study was that although students did think the labs were valuable, they did not give too much thought on how they could use the labs in the future. They did not give specific examples of how these labs would be useful in their jobs and in their lives. Instead, they gave somewhat vague suggestions of how they could connect the data science that they did to social justice issues. This may be a difficult issue to address in class, but I could give suggestions for programmatic changes like internships to help students make this connection. If students got practice doing similar data science work outside of class and in a real-world context, they may see the connections better and be able to give examples of how the work they did in class relates to their lives and their jobs.

Lastly, I mentioned that CA involves giving students the opportunity to work on real world problems that they care about and are interested in. In a class with hundreds of students, it can be challenging to find data and examples that all students find interesting. When designing the labs in this study, I chose topics and examples that students historically had found interesting and included scaffolds where students could decide what questions they wanted to answer. In future iterations of these labs, I hope to give students multiple examples involving different

social justice issues and allow them to choose which examples they would like to analyze and form groups based on their preferences. Another idea that I had was to reserve some lab time for students to give presentations about the analysis that they did and the questions that they answered. If groups are analyzing different data and answering different questions, the students will learn from each other's presentations and engage in knowledge sharing, which is an important part of DC.

Future Research

As both a researcher and an educator, I would like to continue this work by doing more iterations of DBR and designing other labs. Doing DBR allows me to make meaningful changes in the classroom and continue to make improvements to the labs. Currently in this study, I was able to redesign two of the 14 labs that the students do throughout the semester. I would love to do what I did for these two labs to the entire course. It would be great if the students could get practice communicating in multiple ways and working with real data involving social justice issues in every lab. It is important that students get a large variety of examples showing how data science is connected to social justice. Both of these labs occurred in the middle of the semester, but it would be ideal if we started having labs like these during the first week of the class so that students understand that communicating and working together is something that is expected throughout the entire semester.

In the future, I would also like to add more scaffolds to the labs that help the students start thinking about solutions to problems. Right now, the students analyze the data and through this analysis, they engaged in social justice awareness. However, it is important that as educators, we also allow our students to take the idea of social justice awareness a step further. It is important for the students to get exposed to these issues and gain awareness, but it is also

important for them to think about how the work they are doing can be used to facilitate change and make a greater impact beyond the scope of the course. Right now, the theme of thinking about mobilizing for change came up frequently in IR Question 7 in Lab B. When students responded to this question, they acted as experts and provided recommendations for solutions to the problems they addressed. Some students also mentioned the idea of solutions to problems in the GD Questions. I would like to add more scaffolds that encourage this to both Lab A and Lab B, as well as future labs that I redesign.

Also, in this study, I used gender neutral pseudonyms for each of the participants. Some may argue that gender and race can have an impact on how students identify issues of social justice. Another way that future work can be done with this data is by looking at it from a different perspective involving race or gender frameworks. This could allow us to see more nuances in the data and see the connection between students' gender and race and their thoughts on social justice issues.

Lastly, I would like to extend these ideas to other data science courses and possibly other statistics and computer science courses. The design principles can be used to create labs in other courses that are similar to Lab A and Lab B from this study. This can help extend the goal of engaging in CA and social justice awareness to similar courses and follow-on courses. If students are only getting practice doing these types of labs in one course, it may not make a large impact. However, if multiple courses expand on these ideas and give students the space to have important discussions, the labs can have a much greater impact. I am excited to continue working on improving data science education by adding data justice to the data science curriculum in as many ways as possible.

REFERENCES

Achiam, M., May, M., & Marandino, M. (2014). Affordances and distributed cognition in museum exhibitions. *Museum Management and Curatorship*, *29*(5), 461-481.

Adhikari, A., DeNero, J., & Jordan, M. I. (2021). Interleaving Computational and Inferential Thinking: Data Science for Undergraduates at Berkeley. *arXiv preprint arXiv:2102.09391*.

Almurashi, W. A. (2016). An introduction to Halliday's systemic functional linguistics. *Journal for the Study of English Linguistics, 4*(1), 70-80.

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher, 41*(1), 16–25.

Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. *Educational researcher*, *25*(4), 5-11.

Angeli, C. (2008). Distributed cognition: A framework for understanding the role of computers in classroom teaching and learning. *Journal of Research on Technology in Education*, *40*(3), 271-279.

Bakker, A. (2004). *Design research in statistics education: On symbolizing and computer tools* (Doctoral dissertation).

Bakker, A., & Van Eerde, D. (2015). An introduction to design-based research with an example from statistics education. In *Approaches to qualitative research in mathematics education* (pp. 429-466). Springer, Dordrecht.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13*(1), 1-14.

Baumer, B. (2015). A data science course for undergraduates: Thinking with data. *The American Statistician*, *69*(4), 334-342.

Bergold, J., & Thomas, S. (2012). Participatory research methods: A methodological approach in motion. *Historical Social Research/Historische Sozialforschung*, 191-222.

Bieda, K. N., Sela, H., & Chazan, D. (2015). "You are learning well my dear" shifts in novice teachers' talk about teaching during their internship. *Journal of Teacher Education, 66*(2), 150-169.

Boenig-Liptsin, M., Tanweer, A., & Edmundson, A. (2022). Data Science Ethos Lifecycle: Interplay of ethical thinking and data science practice. *Journal of Statistics and Data Science Education*, 1-13.

Boland Jr, R. J., Tenkasi, R. V., & Te'Eni, D. (1994). Designing information technology to support distributed cognition. *Organization Science*, *5*(3), 456-475.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, *2*(2), 141-178.

Brunner, R. J., & Kim, E. J. (2016). Teaching data science. *Procedia Computer Science*, *80*, 1947-1956.

Carmichael, I., & Marron, J. S. (2018). Data science vs. statistics: two cultures? *Japanese Journal of Statistics and Data Science*, *1*(1), 117-138.

Cazden, C. (1986). Classroom discourse. In M. Wittrock (Ed.), Handbook of research on teaching (3rd ed., pp. 432-463). New York: Macmillan.

Chieu, V. M., Herbst, P., & Weiss, M. (2011). Effect of an animated classroom story embedded in online discussion on helping mathematics teachers learn to notice. *Journal of the Learning Sciences*, *20*(4), 589-624.

Cobb, P. (2006). Mathematics learning as a social process. In *New mathematics education research and practice* (pp. 147-152). Brill Sense.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.

Coker, J. S. (2017). Student-Designed Experiments: A Pedagogical Design for Introductory Science Labs. *Journal of College Science Teaching*, *46*(5).

Cole, M., & Engeström, Y. (1993). A cultural-historical approach to distributed cognition. *Distributed cognitions: Psychological and educational considerations*, 1-46.

Collins, A. (1992). Toward a design science of education. In *New directions in educational technology* (pp. 15-22). Springer, Berlin, Heidelberg.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, *13*(1), 15-42.

Dichev, C., & Dicheva, D. (2017). Towards data science literacy. *Procedia Computer Science*, *108*, 2151-2160.

D'ignazio, C., & Klein, L. F. (2020). *Data Feminism*. MIT press.

Dodge, S. (2021). A data science framework for movement. *Geographical Analysis*, *53*(1), 92-112.

Edwards, A., & Westgate, D. P. (2005). *Investigating classroom talk*. Routledge.

Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal*, *16*(1), 44-49.

Evans, M. A., Feenstra, E., Ryon, E., & McNeill, D. (2011). A multimodal approach to coding discourse: Collaboration, distributed cognition, and geometric reasoning. *International Journal of Computer-Supported Collaborative Learning*, *6*(2), 253.

Finzer, W. (2013). The data science education dilemma. *Technology Innovations in Statistics Education*, *7*(2).

Ford, C., McNally, D., & Ford, K. (2017). Using Design-Based Research in Higher Education Innovation. *Online Learning*, *21*(3), 50-67.

Franklin, C., & Bargagliotti, A. (2020). Introducing GAISE II: A guideline for precollege statistics and data science education. *Harvard Data Science Review*, *2*(4).

Furner, J. (2015). Information science is neither. *Library Trends*, *63*(3), 362-377.

Gal, Y. (1997). *The assessment challenge in statistics education*. IOS press.

Garfunkel, S. A., Montgomery, M., Bliss, K., Fowler, K., Galluzzo, B., Giordano, F., ... & Zbiek, R. (2016). *GAIMME: Guidelines for assessment & instruction in mathematical modeling education*. Consortium for Mathematics and its Applications.

Giere, R. N. (2007). Distributed cognition without distributed knowing. *Social Epistemology*, *21*(3), 313-320.

Giere, R. N., & Moffatt, B. (2003). Distributed cognition: Where the cognitive and the social merge. *Social Studies of Science*, *33*(2), 301-310.

González, G., & Herbst, P. (2013). An oral proof in a geometry class: How linguistic tools can help map the content of a proof. *Cognition and Instruction, 31*(3), 271-313.

Gould, R., Davis, G., Patel, R., & Esfandiari, M. (2010, July). Enhancing conceptual understanding with data driven labs. In *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on*

Teaching Statistics, Ljubljana, Slovenia. *Voorburg, The Netherlands: International Statistical Institute*.

Green, B. (2021). Data science as political action: Grounding data science in a politics of justice. *Journal of Social Computing*, *2*(3), 249-265.

Grus, J. (2019). *Data science from scratch: first principles with python*. O'Reilly Media.

Guardiola, J. H., Duran-Hutchings, N., & Elsalloukh, H. (2010). Are Statistics Labs Worth the Effort? Comparison of Introductory Statistics Courses Using Different Teaching Methods. *Numeracy*, *3*(1), 5.

Haraway, D. (2013). A manifesto for cyborgs: Science, technology, and socialist feminism in the 1980s. In *Feminism/postmodernism* (pp. 190-233). Routledge.

Herrington, J., McKenney, S., Reeves, T., & Oliver, R. (2007, June). Design-based research and doctoral students: Guidelines for preparing a dissertation proposal. In *EdMedia+ Innovate Learning* (pp. 4089-4097). Association for the Advancement of Computing in Education (AACE).

Hollebrands, K. F., Conner, A., & Smith, R. C. (2010). The nature of arguments provided by college geometry students with access to technology while solving problems. *Journal for Research in Mathematics Education*, *41*(4), 324-350.

Hutchins, E. (1995). *Cognition in the Wild* (No. 1995). MIT press.

Hutchins, E., & Klausen, T. (1996). Distributed cognition in an airline cockpit. *Cognition and Communication at Work*, 15-34.

Irizarry, R. A. (2020). The role of academia in data science education. *Harvard Data Science Review*, *2*(1).

Kali, Y. (2016). Transformative learning in design research: The story behind the scenes. *Transforming Learning, Empowering Learners*, 4-5.

Karaali, G., & Khadjavi, L. S. (Eds.). (2019). *Mathematics for social justice: Resources for the college classroom* (Vol. 60). American Mathematical Soc.

Karasavvidis, I. (2002). Distributed cognition and educational practice. *Journal of Interactive Learning Research*, *13*(1), 11-29.

Kelly, A. E., Lesh, R. A., & Baek, J. Y. (Eds.). (2014). *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching*. Routledge.

Kendi, I. X. (2019). *How to be an antiracist*. One world.

Kross, S., Peng, R. D., Caffo, B. S., Gooding, I., & Leek, J. T. (2020). The democratization of data science education. *The American Statistician*, *74*(1), 1-7.

Lave, J. (1988). *Cognition in practice*. Cambridge University Press.

Lee, V. R., Wilkerson, M. H., & Lanouette, K. (2021). A call for a humanistic stance toward K–12 data science education. *Educational Researcher*, *50*(9), 664-672.

Leeuwen, T. V. (2015). Multimodality in education: Some directions and some questions. *Tesol Quarterly*, *49*(3), 582-589.

Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Ablex Publishing Corporation.

Linn, M. C., Shear, L., Bell, P., & Slotta, J. D. (1999). Organizing principles for science education partnerships: Case studies of students' learning about 'rats in space' and 'deformed frogs'. *Educational Technology Research and Development, 47*(2), 61–84.

Martin, J. R., & Rose, D. (2007). *Working with discourse: Meaning beyond the clause (2nd ed.)*. London: Continuum.

Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. New York: Palgrave Macmillan.

Martin, N. D., Tissenbaum, C. D., Gnesdilow, D., & Puntambekar, S. (2019). Fading distributed scaffolds: the importance of complementarity between teacher and material scaffolds. *Instructional Science*, *47*(1), 69-98.

McKenney, S., & Reeves, T. C. (2012). *Conducting Educational Design Research*. New York: Routledge.

Mehan, H. (1979). Learning lessons: Social organization in the classroom. Cambridge, MA: Harvard University Press.

Messick, S. (1992). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(2), 13–23.

Miller, E., Manz, E., Russ, R., Stroupe, D., & Berland, L. (2018). Addressing the epistemic elephant in the room: Epistemic agency and the next generation science standards. *Journal of Research in Science Teaching*, *55*(7), 1053-1075.

Moore-Russo, D., Conner, A., & Rugg, K. I. (2011). Can slope be negative in 3-space? Studying concept image of slope through collective definition construction. *Educational studies in Mathematics*, *76*(1), 3-21.

Moreno, R., & Mayer, R. E. (2000). A learner-centered approach to multimedia explanations: Deriving instructional design principles from cognitive theory. *Interactive multimedia electronic journal of computer-enhanced learning*, *2*(2), 12-20.

Moses, L. E. (2019). Statistical concepts fundamental to investigations. In *Medical Uses of Statistics* (pp. 5-26). CRC Press.

Narciss, S., & Koerndle, H. (2008). Benefits and constraints of distributed cognition in foreign language learning: Creating a web-based tourist guide for London. *Journal of Research on Technology in Education*, *40*(3), 281-307.

Noble, S. U. (2018). Algorithms of oppression. In *Algorithms of Oppression*. New York University Press.

Nolan, D., & Speed, T. P. (2001). *Stat labs: mathematical statistics through applications*. Springer Science & Business Media.

O'Donnell, M. (2011). Introduction to systemic functional linguistics for discourse analysis. *Language, Function and Cognition, 12*, 1-8.

Pimentel, J. F., Murta, L., Braganholo, V., & Freire, J. (2019, May). A large-scale study about quality and reproducibility of jupyter notebooks. In *2019 IEEE/ACM 16th International Conference on Mining Software Repositories (MSR)* (pp. 507-517). IEEE.

Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). Teachers: An analysis of the co-design process. *Learning, 2*(1), 51–74.

Perkel, J. M. (2018). Why Jupyter is data scientists' computational notebook of choice. *Nature*, *563*(7732), 145-147.

Ramamurthy, B. (2016, February). A practical and sustainable model for learning and teaching data science. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education* (pp. 169-174).

Rao, A., Bihani, A., & Nair, M. (2018, October). Milo: A visual programming environment for Data Science Education. In *2018 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)* (pp. 211-215). IEEE.

Rogers, Y. (1997). *A brief introduction to distributed cognition*.

Rogers, Y., & Ellis, J. (1994). Distributed cognition: an alternative framework for analysing and explaining collaborative working. *Journal of Information Technology*, *9*(2), 119-128.

Salomon, G. (1992). New challenges for educational research: Studying the individual within learning environments. *Scandinavian Journal of Educational Research*, *36*(3), 167-182.

Sandoval, W. A., & Bell, P. (2004). Design-based research methods for studying learning in context: Introduction. *Educational Psychologist*, *39*(4), 199-201.

Shah, J. K., Ensminger, D. C., & Thier, K. (2015). The Time for Design-Based Research is Right and Right Now. *Mid-Western Educational Researcher*, *27*(2).

Sinclair, J., & Coulthard, M. (1975). Towards an analysis of discourse. London: Oxford University Press.

Sivasubramaniam, P. (2004). Distributed cognition and the use of graphing calculators in the learning of mathematics. In *Proceedings of the 2nd National Conference on Graphing Calculators*, Penang, October 4 (Vol. 6, pp. 93-103).

Song, I. Y., & Zhu, Y. (2016). Big data and data science: what should we teach? *Expert Systems*, *33*(4), 364-373.

Snow, C. E. (2011). The unavoidable need for distributed cognition in teaching literacy. *South African Journal of Childhood Education*, *1*(2), 10.

Sutton, J. (2006). Distributed cognition: Domains and dimensions. *Pragmatics & Cognition*, *14*(2), 235-247.

Tajuddin, N. A. M., Tarmizi, R. A., Konting, M. M., & Ali, W. Z. W. (2009). Instructional Efficiency of the Integration of Graphing Calculators in Teaching and Learning Mathematics. *Online Submission*, *2*(2), 11-30.

Tal, T., & Tsaushu, M. (2018). Student-centered introductory biology course: evidence for deep learning. *Journal of Biological Education*, *52*(4), 376-390.

Tashakkori, A., & Creswell, J. W. (2007). *The new era of mixed methods*.

Tishkovskaya, S., & Lancaster, G. A. (2012). Statistical education in the 21st century: A review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education*, *20*(2).

Tissenbaum, M., Sheldon, J., & Abelson, H. (2019). From computational thinking to computational action. *Communications of the ACM*, *62*(3), 34-36.

Toulmin, S. (1958). *The uses of argument*. Cambridge: Cambridge University Press.

University of Virginia (2022). *A summer undergraduate research program at the University of Virgini*a. The Data Justice Academy. https://datascience.virginia.edu/data-justice-academy

Van Der Aalst, W. (2016). Data science in action. In *Process Mining* (pp. 3-23). Springer, Berlin, Heidelberg.

VanderPlas, J. (2016). *Python data science handbook: Essential tools for working with data*. O'Reilly Media, Inc.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, *53*(4), 5-23.

Willingham, D.T., & Daniel, D.B. (2021). Making education research relevant. *Education Next, 21*(2).

Xyntarakis, M., & Antoniou, C. (2019). Data science and data visualization. In *Mobility Patterns, Big Data and Transport Analytics* (pp. 107-144). Elsevier.

Zhang, J., & Patel, V. L. (2006). Distributed cognition, representation, and affordance. *Pragmatics & Cognition*, *14*(2), 333-341.

Zydney, J. M., Warner, Z., & Angelone, L. (2020). Learning through experience: Using design based research to redesign protocols for blended synchronous learning environments. *Computers & Education*, *143*, 103678.

APPENDIX A: LAB A

Topics 1: Visual Displays of Data (Week 5)

Welcome to Lab_Plots!
You may not know this, but the salaries of the employees at the University of Illinois are publicly available! We have curated this data in a nice, structured dataset called salaries.csv for you to explore. The goal of this lab is to work with real UIUC salary data to explore its properties, answer important questions, and to think about the implications of collecting and analyzing this data. You will also create visual displays of data to help communicate your findings. Throughout the lab, it is important to think about being a *critical consumer of data* who can not only use statistics and programming to analyze data but can also think about the why part of data science both in the classroom and in the world. Let's get started!

Exploratory Data Analysis (EDA)
As we discussed in lecture, the first step of any data analysis is to get familiar with your dataset. Think about what this data can tell you and what variables are included. Data scientists always start with this step.

Coding: Import the dataset and take a quick look at it. Look at the variable names, how many rows and columns there, and anything else you may notice!
Group Discussion: Thinking about the variables in the dataset, what are three questions that you'd like to use data science to answer (ex. What is the mean salary of my home department?). Type them below and then share them with your group. Share at least one of your questions and why you want to find the answer to it.

Descriptive Statistics
Coding: Next, let's do some overall exploratory data analysis to get some baseline data to compare to. Before you do the calculations, guess what you think the average salary is at UIUC. Enter your answer below. Then find the overall mean salary, the overall median salary, and the overall SD of the salaries.

Individual Reflection: Write a short paragraph answering the following questions: Is the mean or median larger? Why do you think this might be the case? How is the standard deviation is related to your answer?

Visual Displays of Data
Now, we are a bit more familiar with the dataset and some summary statistics. Looking at overall descriptive statistics helps us summarize all of the observations in a column, rather than having to scroll through all of the observations! However, descriptive statistics alone often don't tell the whole story.

In lecture, we discussed Anscombe's Quartet (a set of 4 datasets with the exact same descriptive statistics that look completely different when graphed). This showed us that visual displays of data are also a necessary and important part of exploratory data analysis. (insert a visual of Anscombe's Quartet).

Coding: Let's look at one of the simplest, yet powerful visual displays of data, a histogram. Histograms are used to show the overall shape of the data and they allow us to see frequencies. Create a frequency histogram of the salaries at UIUC.

Group Discussion: Look at the histogram below and compare it to your histogram you just created. This is a histogram of US salaries, whereas the histogram you created is for UIUC salaries. How do they compare? Why do you think they are similar or different? Discuss your thoughts with your group. (insert histogram of US salaries)

Coding: Next, let's look at another simple, yet power visualization: a boxplot! Create a boxplot of the overall salary data at UIUC.

Group Discussion: Discuss with your group whether you think a histogram or a boxplot or both best visualize the salary data. Explain why histograms, boxplots, or both are important and what they can tell us about the data. Why is it valuable to look at visual displays of salary data in general?

Departmental Data

One of the interesting properties of the salary dataset is that you can see what department each professor is from. Let's explore whether or not there are salary discrepancies among departments.

Let's look at 3 different departments: The Department of English, The Department of Psychology, and The Department of Electrical and Computer Engineering.

Coding: Find the mean, median, and SD for these departments.

Individual Reflection: Are the descriptive statistics drastically different? Think about why or why not. Are there any confounding variables that could be driving these differences? Write down your thoughts below.

Coding: Create a histogram of the salaries in your home department. Also, record the maximum and minimum salary in your home department.

Group Discussion: Share the results with the people at your table. Do you notice any similarities or differences between departments?

Coding: After hearing everyone at your table talk about their department, pick the department from your table that you think is the most different from yours. Create side by side boxplots of your department and the department of someone else at your table.

Individual Reflection: What do you notice about the two boxplots? What are the similarities and differences? What might this mean for the faculty members in this department? What might this mean for the students? For the university? Record some thoughts below.

Gender and Titles
Data can often reveal systemic problems or discrimination. For example, in many companies, men and women are promoted at different rates. Let's look at a subset of the salary dataset to investigate whether or not there is a difference in salaries between faculty who identify as men and women. We have this data for the STAT and CS departments (Karle and Wade's home departments) and have compiled it in a dataset called statcssalaries.csv.

Coding: Load the dataset and create two side by side boxplots of the male salaries and female salaries in the STAT and CS Department. Have two people in your group do this for the STAT Department and have the other two people in your group do this for the CS Department.

Group Discussion: Share the results with the rest of the group. Some questions to think about: Do the boxplots look similar or different? Are there any outliers? Do you think these results present a problem? What are the implications of this analysis for faculty, students, and the university?

Coding: You'll also notice that the titles are available in this new dataset! Calculate the average, median, and SD of your title (see below). Also, create a histogram and boxplot of salaries for your assigned title.

Group Member #1: Instructor: Instructors normally do not have a terminal degree (PhD) and are mainly hired to teach.
Group Member #2: Teaching Assistant Professor: Teaching assistant professors are non-tenure track teaching faculty with PhDs.
Group Member#3: Assistant Professor: Assistant professors
Group Member #4: Professor: This is the highest title for faculty and means the person has a PhD, is on the tenure track, and has been promoted multiple times.

Group Discussion: Share your results with the group.

Individual Reflection: As a student, do you think these are fair salaries or are they too low or too high? Why do you think this? How does this dataset impact you? Write down your thoughts below.

Exploring Your Own Interests
At this point of the lab, we have investigated a lot of questions, however, these have been questions that we have told you to answer. As a data scientist, it is important to be able to use the data science skills that you learn in the classroom to answer real questions that you have.

Individual Reflection: Think about two questions that you have that have not been answered. Record them below.

Coding: Now, answer at least one of these questions using Python and the salary dataset. Type your code below.

Group Discussion: Share your question and results with the group. If you finish before everyone is done, think of more questions and try to answer them.

Counter Data

As a data scientist, it is important to think about the implications of collecting and not collecting certain data. Data that is not collected in an analysis is called *counter data*.

Many times, data is collected to make a profit. Data has even been referred to as "the new oil" because it is oftentimes seen as an uncapped natural resource and if you can figure out how to capture and refine it, it can lead to a massive profit.

Group Discussion: Chat in your groups about the following questions: What is some counter data that you'd like to collect in regard to salaries? Why do you think it's important to collect this data? Share your thoughts with your group. Some examples of questions that would need counter data are listed below:

Do Small Liberal Arts Colleges pay a lot less than R1 institutions?

How does UIUC compare to other Big Ten institutions?

Are there salary disparities if we stratify based on ethnicity?

How has this data changed over time?

Individual Reflection: Now think about how you would actually go about collecting this data to answer a question that you have. Respond below with the question you'd want to answer, counter data you'd want to collect, what resources you'd need to collect it, and if this is realistic to do.

Beyond Stat 107

As a Data Scientist, it's important to think about who data analysis can benefit and who it can harm. It's also important to think about why the data is collected and why counter data is not collected.

Group Discussion: Who benefits from collecting this salary data? Who does this data harm? Why do you think this salary data is public? Can you think of any reasons that this could be problematic?

Individual Reflection: Let's think beyond the salary data! Thinking about the data science skills that you have mastered: experimental design, descriptive statistics, and basic visualization, what is one real world question that you'd like to answer using data? It can be about anything, a hobby or a specific passion or issue. You won't be doing this, but instead thinking about how you'd do this. Write down your thoughts below.

Individual Reflection: Communication is a key part of data science and it's important that all of us are critical consumers of data. What have you learned through doing this lab? Was there anything you wish you could have discussed or could have discussed more?

APPENDIX B: ANNOTATED LAB A

Student Worksheet	Artifact	Scaffold
Exploratory Data Analysis (EDA)		
Coding: Import the dataset and take a quick look at it. Look at the variable names, how many rows and columns there, and anything else you may notice!	A1. The Salary Dataset	
Group Discussion: Thinking about the variables in the dataset, what are three questions that you'd like to use data science to answer (ex. What is the mean salary of my home department?). Type them below and then share them with your group. Share at least one of your questions and why you want to find the answer to it.		S1. Beginning Questions
Descriptive Statistics Coding: Next, let's do some overall exploratory data analysis to get some baseline data to compare to. Before you do the calculations, guess what you think the average salary is at UIUC. Enter your answer below. Then find the overall mean salary, the overall median salary, and the overall SD of the salaries.	A2. Python A3. Jupyter Notebooks	
Individual Reflection: Write a short paragraph answering the following questions: Is the mean or median larger? Why do you think this might be the case? How is the standard deviation is related to your answer?		S2. Mean vs. Median
Visual Displays of Data Coding: Let's look at one of the simplest, yet powerful visual displays of data, a histogram. Histograms are used to show the overall shape of the data and they allow us to see frequencies. Create a frequency histogram of the salaries at UIUC.		S3. Histogram
Group Discussion: Look at the histogram below and compare it to your histogram you just created. This is a histogram of US salaries, whereas the histogram you created is for UIUC salaries. How do they compare? Why do you think they are similar or different? Discuss your thoughts with your group. (insert histogram of US salaries)		S4. Histogram Comparison
Coding: Next, let's look at another simple, yet power visualization: a boxplot! Create a boxplot of the overall salary data at UIUC.		S5. Boxplot

Group Discussion: Discuss with your group whether you think a histogram or a boxplot or both best visualize the salary data. Explain why histograms, boxplots, or both are important and what they can tell us about the data. Why is it valuable to look at visual displays of salary data in general?		S6. Boxplot Comparison
Departmental Data		
Let's look at 3 different departments: The Department of English, The Department of Psychology, and The Department of Electrical and Computer Engineering. Coding: Find the mean, median, and SD for these departments.		S7. Descriptive Statistics for Departments
Individual Reflection: Are the descriptive statistics drastically different? Think about why or why not. Are there any confounding variables that could be driving these differences? Write down your thoughts below.		S8. Departmental Differences
Coding: Create a histogram of the salaries in your home department. Also, record the maximum and minimum salary in your home department.		S9. Home Department Histogram
Group Discussion: Share the results with the people at your table. Do you notice any similarities or differences between departments?		S10. Home Department Comparison
Coding: After hearing everyone at your table talk about their department, pick the department from your table that you think is the most different from yours. Create side by side boxplots of your department and the department of someone else at your table.		S11. Side by Side Boxplots
Individual Reflection: What do you notice about the two boxplots? What are the similarities and differences? What might this mean for the faculty members in this department? What might this mean for the students? For the university? Record some thoughts below.		S12. Implications for People
Gender and Titles		
Coding: Load the dataset and create two side by side boxplots of the male salaries and female salaries in the STAT and CS Department. Have two people in your group do this for the STAT Department and have the other two people in your group do this for the CS Department.		S13. Salaries Stratified by Gender

Group Discussion: Share the results with the rest of the group. Some questions to think about: Do the boxplots look similar or different? Are there any outliers? Do you think these results present a problem? What are the implications of this analysis for faculty, students, and the university?		S14. Gender and Salary Discussion
Coding: You'll also notice that the titles are available in this new dataset! Calculate the average, median, and SD of your title (see below). Also, create a histogram and boxplot of salaries for your assigned title.		S15. Titles and Salaries
Group Discussion: Share your results with the group.		S16. Title and Salary Discussion
Individual Reflection: As a student, do you think these are fair salaries or are they too low or too high? Why do you think this? How does this dataset impact you? Write down your thoughts below.		S17. Students' Thoughts on Salary Analysis
Exploring Your Own Interests		
At this point of the lab, we have investigated a lot of questions, however, these have been questions that we have told you to answer. As a data scientist, it is important to be able to use the data science skills that you learn in the classroom to answer real questions that you have. Individual Reflection: Think about two questions that you have that have not been answered. Record them below.		S18. Unanswered Questions
Coding: Now, answer at least one of these questions using Python and the salary dataset. Type your code below.		S19. Answering Unanswered Questions
Group Discussion: Share your question and results with the group. If you finish before everyone is done, think of more questions and try to answer them.		S20. Sharing Your Questions
Counter Data		
Group Discussion: Chat in your groups about the following questions: What is some counter data that you'd like to collect in regard to salaries? Why do you think it's important to collect this data? Share your thoughts with your group. Some examples of questions that would need counter data are listed below:		S21. Counter Data

Do Small Liberal Arts Colleges pay a lot less than R1 institutions? How does UIUC compare to other Big Ten institutions? Are there salary disparities if we stratify based on ethnicity? How has this data changed over time?		
Individual Reflection: Now think about how you would actually go about collecting this data to answer a question that you have. Respond below with the question you'd want to answer, counter data you'd want to collect, what resources you'd need to collect it, and if this is realistic		S22. Data Collection
Beyond Stat 107		
As a Data Scientist, it's important to think about who data analysis can benefit and who it can harm. It's also important to think about why the data is collected and why counter data is not collected. Group Discussion: Who benefits from collecting this salary data? Who does this data harm? Why do you think this salary data is public? Can you think of any reasons that this could be problematic?		S23. Benefits and Harms of Analysis
Individual Reflection: Let's think beyond the salary data! Thinking about the data science skills that you have mastered: experimental design, descriptive statistics, and basic visualization, what is one real world question that you'd like to answer using data? It can be about anything, a hobby or a specific passion or issue. You won't be doing this, but instead thinking about how you'd do this. Write down your thoughts below.		S24. Beyond Salary Data
Individual Reflection: Communication is a key part of data science and it's important that all of us are critical consumers of data. What have you learned through doing this lab? Was there anything you wish you could have discussed or could have discussed more?		S25. Lab Reflection

KEY: S1 is read as "Scaffold 1" and A1 is read as "Artifact 1"

APPENDIX C: DESCRIPTION OF SCAFFOLDS AND ARTIFACTS FOR LAB A

Scaffold or Artifact	What is?	Why is it included?	What data do I get from it?
A1	The salary dataset- the main dataset used throughout the lab	It is important that the students get practice working with real datasets. This is an Illinois-themed data set and historically has been interesting to them. It allows students to explore important issues (salary discrepancies)	Having a dataset for the students to analyze allows me to evaluate their code throughout the lab. Part of CA is to make sure the students know how to use coding to answer real questions.
A2	Python- programming language	It is a major component of the class and is a tool that is both used frequently and industry and that helps students analyze large datasets easily.	All of the coding done in the lab will be data that helps me identify whether or not the students understand the questions and can use Python to answer them.
A3	Jupyter Notebooks- an interface where students can type code and text in the same document.	These allow students to type code that runs and type text into the same document. This allows me to intertwine the analysis and reflection and discussion so that it is all connected.	All of the typed data (Python code and typed reflections) will be collected through the Jupyter notebooks. They allow me to capture what is not heard on the audio recordings.
S1: Group Discussion	Beginning Questions- the students think about questions they would like to use data science to answer	These get students to start thinking about the dataset and what questions they have right away. This is also the first discussion question to help them get used to having group discussions.	Here, I get audio data. This allows me as the researcher to understand my students' interests. I can use this information to help refine the labs later.
S2: Individual Reflection	Mean vs. Median- the students think about the difference between the mean and median and	One of the goals of the lab is for students to understand the statistical techniques they are using. It is important that students not only know how to calculate the mean, median, and SD, but also that they understand the difference and how they are related to each	Here, I get written data. This allows me to know if they understand these statistical concepts and whether or not I need to go more in depth on them in lecture. Understanding these statistical concepts in the lab will allow the students to use them outside of class.

	how it's related to SD.	other. This can help them in the future when analyzing new datasets and communicating their findings.	
S3: Coding Questions	Histogram- the students create a histogram of the overall salaries	This is the first time that students are creating a visual display of data. Visual displays of data are a great way to communicate data science findings so it's important that the students know how to create these.	Here, I get code as the data. This allows me to see if the students know how to use Python to create a histogram. Knowing how to do this in the lab, empowers them to do this with other data outside of the lab.
S4: Group Discussion	Histogram comparison- the students look at a histogram of US salaries and compare it to the histogram of UIUC salaries	The students may not know much about salary data in general since most of them have never had a salary. This allow them to see how the data at UIUC compares to national data and think about why they may be similar or different.	Here, I get audio data. This allows the students to start thinking about the data beyond the context of the course.
S5: Coding Questions	Boxplot- the students create a boxplot of the overall salaries	Boxplots are another visual display of data that can be used as a way for students to communicate their findings.	Here, I get code as the data. This allows me to see if the students know how to use Python to create a boxplot. Knowing how to do this in the lab, empowers them to do this with other data outside of the lab.
S6: Group Discussion	Boxplot comparison- the students think about the value of different visual displays of data and how they compare	It is important for students to think about which visual displays of data are best for certain datasets and why. This is a part of CA since the students are using the salary data in the lab to think about how to visualize data in general.	Here, I get audio data. This helps me understand whether or not students understand the different visual displays of data and see the importance of creating them to communicate data science results.
S7: Coding Questions	Descriptive Statistics for Departments- the students calculate the mean, median, and SD for 3 very	This is to promote social justice awareness through the idea that salaries can differ drastically based on the department, regardless of other variables (years of experience, etc.).	Here, I get code as the data and this allows me to see if the students can calculate the statistics for subsetted data which is a useful skill for data analysis. It allows the students to dive deeper into the data.

	different departments		
S8: Individual Reflection	Departmental Differences- the students can express the thoughts they have about departmental differences in writing	This is also to promote social justice awareness by having the students think about why they are seeing these differences in the departments. It also allows the students to gain CA by thinking about confounding variables- a common problem that is often present in data analysis but not understood.	Here, I get written data. I can see if the students understand confounding variables and how they feel about the differences between the departments.
S9: Coding Question	Home Department Histogram- the students create another histogram from subsetted data and focus on their own department	This allows the students to explore something personal to them and see how the department that they are in compares to the whole university. Oftentimes, data scientists are focusing on social justice issues in their communities and can compare their data to state, national, or world data. This helps students gain CA since they are thinking about issues that directly affect them.	Here, I get code as data. This allows the students to get more practice with creating histograms from subsetted data.
S10: Group Discussion	Home Department Comparisons- here the students discuss their results from Scaffold 9 with the rest of their groups	This promotes gaining CA since the students are doing data analysis and communicating their results to others. The students are most likely in different departments so their own individual analysis will be different from their group members. It's important for the students to practice analyzing data and communicating their results.	Here, I get audio data. I can see how the students communicate their individual results and whether or not this promotes rich discussion.
S11: Coding Questions	Side by Side Boxplots- the students create side by side boxplots on the same plane to compare their	This promotes comparison and the understanding of statistical concepts such as IQR and outliers through boxplots. This also promotes CA because it is foregrounding student choice since they get to decide which other department to use.	Here, I get code as data. This allows the students to create two boxplots to show comparison as opposed to one. Comparison is something data scientists do often.

	department to a very different department		
S12: Individual Reflection	Implications for People- here the students are thinking about what their findings mean, who they effect, and what the implications are for different people	This promotes thinking about the implications of the data analysis they do, which is an idea from Data Feminism. The students are looking at these results from multiple perspectives and thinking about who benefits from this analysis and who does not.	Here I get written data. This allows me to see if the students are thinking critically about the implications of their data analysis rather than just doing it.
S13: Coding Question	Salaries Stratified by Gender- the students are using boxplots to compare the salaries by gender in the two departments	This allows the students to practice subsetting data and create boxplots of the subsetted data. It allows the students to start thinking about potential issues of discrimination and get practice using data to discover if there are any of these issues. It also promotes CA because the students are using real data to answer important questions about systemic problems or discrimination.	Here, I get code as data. This scaffold has the students create boxplots from subsetted data with a new dataset.
S14: Group Discussion	Gender and Salary Discussion- the students have the opportunity to discuss their findings with the group, thinking about the implications of this analysis for faculty, students, and the university	This promotes communication and CA since the students are discussing an important topic together and thinking about the implications of their analysis and seeing how it can be related to issues outside of the classroom.	Here, I get audio data that helps me determine if the students are gaining CA and communicating.

S15: Coding Question	Titles and Salaries- the students are using histograms and descriptive statistics to compare the salaries by title in the two departments	This allows the students to practice subsetting data and creating histograms of the subsetted data. It allows them to think about other issues of discrimination (first gender, now title) and again they get practice using data to discover if there are any of these issues. It also promotes CA because the students are using real data to answer important questions about systemic problems or issues of discrimination.	Here, I get code as data. This scaffold has the students create histograms and calculate descriptive statistics from subsetted data with a new dataset.
S16: Group Discussion	Title and Salary Discussion- the students have the opportunity to discuss their findings with the group, thinking about the implications of this analysis for faculty, students, and the university	This promotes communication and CA since the students are discussing an important topic together and thinking about the implications of their analysis and seeing how it can be related to issues outside of the classroom. They are also working as a part of a team since they are each responsible for one part of the analysis and depending on each other to understand the entire question.	Here, I get audio data that helps me determine if the students are gaining CA and communicating.
S17: Individual Reflection	Students' Thoughts on Salary Analysis- here the students are reflecting on their analysis and discussion and sharing their thoughts	This allows the students to take all of the information they have, form an opinion, and write about it. This helps them gain CA because they are thinking about how this impacts them as well as others. It also helps them get practice communicating their overall thoughts in writing.	Here, I get written data. This helps me see if the students are gaining CA and thinking critically about the issues presented in this lab and how they affect their lives.
S18: Individual Reflection	Unanswered Questions- the students think about	For the first time in the lab, the students can think about questions other than the ones provided. This promotes CA because it is	Here, I get written data. This helps me see if the students are gaining CA by asking questions that they find interesting.

	two remaining questions they have	allowing the students to choose questions that they think would be interesting to answer.	
S19: Coding Question	Answering Unanswered Questions- the students use data science to answer at least one of the questions they had	Here the students are gaining CA because they are answering a question that is important to them using data science and the techniques they learned in this lab.	Here, I get code as data. This allows me to see if the students are gaining CA by using the tools they learned in the lab to answer a question.
S20: Group Discussion	Sharing Your Questions- here the students discuss the work they did answering a new question	The students are gaining CA by explaining what question they wanted to answer and how they answered it. The students are all learning through each other in this discussion. This also foregrounds student choice, hence helping them gain CA.	Here, I get audio data that helps me determine if the students are gaining CA and communicating.
S21: Group Discussion	Counter Data- the students discuss counter data that is interesting to them	This promotes the idea of counter data from Data Feminism and helps the students start thinking critically about it.	Here, I get audio data that helps me see how the students take this concept from the lab and apply it to their lives.
S22: Individual Reflection	Data Collection- the students describe how they would collect data to answer a question that they have	This promotes CA by having the students think about data they would like to collect and how they would collect it and analyze it. These are all things they could have to do in their jobs or in their lives to use data science to answer questions.	Here, I get written data. This helps me see how students can prepare for analysis when thinking about a question they would like to answer.
S23: Group Discussion	Benefits and Harms of Analysis- the students think about the lab they just completed and who benefits from this	This promotes CA as the students are thinking about the benefits of the work they have done and who could potentially be harmed by it. It's important to think about these ideas when doing any type of data science.	Here, I get audio data that helps me see how the students are thinking critically about how what they are doing affects others.

	analysis and who it could harm		
S24: Individual Reflection	Beyond Salary Data- the students write about a question outside of the lab that they would be interested in using data science to answer	This is helping them gain CA by thinking about how they can use what they learned from the lab and apply it to their lives or their jobs.	Here, I get written data. This helps me see how the students are gaining CA.
S25: Individual Reflection	Lab Reflection- the students reflect on what they learned through the lab and whether or not they wished they could have done anything more	This helps me as the researcher and instructor to see what the students took away from the lab and get suggestions on what to add for the next adaptation.	Here, I get written data. This helps me see what the students learned and whether or not I should make any adjustments to the lab for the next semester.

APPENDIX D: PRE-LAB AND POST-LAB SURVEYS

Pre-Lab Survey

1. Before taking this course, did you have any experience with…
    a. Statistics? Yes/No
    b. Computer Science? Yes/No
    c. Data Science? Yes/No

2. What year are you in school? Freshman, Sophomore, Junior, Senior, Other

3. Are you a STEM (science, technology, engineering, and math) major? Yes/No

4. How much do you agree with this statement? I feel confident that I could explain what a histogram is and how it can be useful.
    a. Answer on a scale of 1-5, 1 being strongly disagree and 5 being strongly agree
5. How much do you agree with this statement? I feel confident that I could explain what a boxplot is and how it can be useful.
    a. Answer on a scale of 1-5, 1 being strongly disagree and 5 being strongly agree
6. How much do you agree with this statement? I feel confident that I can use Python to create visual displays of data.
    a. Answer on a scale of 1-5, 1 being strongly disagree and 5 being strongly agree
7. How much do you agree with this statement? I know what counter data is and understand why it is important.
    a. Answer on a scale of 1-5, 1 being strongly disagree and 5 being strongly agree
8. What do you hope to learn from doing this lab? It can be anything! (short answer)
9. In previous labs, have you worked with other people?
    a. I mainly have worked by myself in previous labs.
    b. I mainly work with one other person in previous labs.
    c. I mainly work with a small group in previous labs.

Post-Lab Survey

1. How much do you agree with this statement? I feel confident that I could explain what a histogram is and how it can be useful.
    a. Answer on a scale of 1-5, 1 being strongly disagree and 5 being strongly agree
2. How much do you agree with this statement? I feel confident that I could explain what a boxplot is and how it can be useful.
    a. Answer on a scale of 1-5, 1 being strongly disagree and 5 being strongly agree
3. How much do you agree with this statement? I feel confident that I can use Python to create visual displays of data.
    a. Answer on a scale of 1-5, 1 being strongly disagree and 5 being strongly agree
4. How much do you agree with this statement? I know what counter data is and understand why it is important.

    a. Answer on a scale of 1-5, 1 being strongly disagree and 5 being strongly agree
5. How much do you agree with this statement? My group communicated well during the questions that asked for discussion during the lab.
    a. Answer on a scale of 1-5, 1 being strongly disagree and 5 being strongly agree
6. How (if at all) do you think what you learned in this lab will help you in your future job? (short answer)
7. How (if at all) do you think what you learned in this lab will help you in your life? (short answer)
8. What was your biggest takeaway from doing this lab? (short answer)
9. Did you like the discussion parts of the lab? Why or why not? (short answer)
10. Did you like the individual reflection parts of the lab? Why or why not? (short answer)
11. Do you feel like this lab helped you become a better Python programmer? Why or why not? (short answer)
12. Do you have any remaining questions about the lab, data science, or anything? (short answer)

APPENDIX E: INTERVIEW PROTOCOL

Participant Information

Institution: _____

Interviewee (Title and Name): _____

Interviewer: _____

Introductory Protocol

To facilitate our notetaking, we would like to audio tape our conversations today. Only researchers on the project will have access to the audio files, which will be eventually destroyed after they are transcribed. In addition, you have signed a consent form devised to meet our human subject requirements. Essentially, this document states that: (1) all information will be held confidential, (2) your participation is voluntary, and you may stop at any time if you feel uncomfortable, and (3) we do not intend to inflict any harm. Thank you for your agreeing to participate.

We have planned this interview to last no longer than 30 minutes. During this time, we have several questions that we would like to cover. If time begins to run short, it may be necessary to interrupt you in order to push ahead and complete this line of questioning.

Before we begin the interview, do you have any questions? [Discuss questions]
If any questions (or other questions) arise at any point in this study, you can feel free to ask them at any time. I would be more than happy to answer your questions.

Introduction Example

Hi, I'm Karle Flanagan and I'm a PhD student at UIUC who will be doing the interview. You have been selected to speak with us today because you have been identified as someone who has completed the two labs (lab_plots and lab_clt) in the study Data science students' development of computational action: A design-based research study. This interview does not aim to evaluate your data science skills or experiences. Rather, we are trying to learn more about your perceptions of the lab and how to improve it for future semesters. Let's get started!

Part 1: Background Questions

10. What is your major?
11. What type of job do you plan to have after you graduate?
12. Before taking this course, did you have any experience in statistics, computer science, or data science?

Part 2: Student Perceptions

1. How would you describe the communication between your group members throughout the labs (lab_plots and lab_clt)?
2. In your opinion, what were the major strengths of the labs?

3. In your opinion, what were any weaknesses of the labs?
4. How (if at all) do you think the content of the labs will help you in your future job?
5. How (if at all) do you think the content of the labs will help you in your life?

APPENDIX F: LAB B

Welcome to this week's lab! Lab_Justice

We plan on continuing to explore simulation in Python this week, however, this time, we are going to simulate some real-world events that have actually happened in the past. The goal is for you to see how we can use data science to think about issues of equity and social justice. As usual, in addition to coding, we want you to get practice having discussions about the data science you are doing and think about how data science can be a useful tool that can help you in your future job and as a citizen.

Let's get started!

First, it's important that you are able to work with others during this lab! Form a group of anywhere between 2 and 4 total students and enter their information below:

Group Member #1 Name:
Group Member #1 NetID:
Group Member #1 Major:

Group Member #2 Name:
Group Member #2 NetID:
Group Member #2 Major:

Group Member #3 Name:
Group Member #3 NetID:
Group Member #3 Major:

Puzzle #1: Jury Selection + Simulation

The Sixth Amendment to the U.S. Constitution provides the right to an "impartial jury" in criminal prosecutions, but what exactly does this mean? The Supreme Court has said that juries must be drawn from a representative cross-section of the community. In other words, juries should be randomly selected from the eligible population. However, there are many instances in history where it doesn't seem like this was the case.

Take for example, the case of *Berghuis v. Smith. In 1993,* Smith (a black man) was convicted of second-degree murder by an all-white jury and sentenced to life in prison. The jury was selected from a panel of about 100 randomly selected people. Only 3 of them were Black and none of those three made it into the final 37 considered for Smith's trial. The county population was approximately 8 percent Black at the time of the trial. Remember, jury panels are supposed to be selected at random from the eligible population.

Group Discussion: Because 8% of the eligible population was Black, 3 black people on a panel of 100 might seem low. Does this difference (8% vs. 3%) seem big to you? Do you think this could be due to chance?

Individual Reflection: Write a few sentences summarizing what your group members said during your discussion. Did people think this difference could have been due to chance or not?

This case was appealed and sent to the Supreme Court. Some people claimed that the overall percentage disparity (between 8% and 3%) was small and reflected no attempt to include or exclude a specified number of black people from the jury." They claimed this "small" difference could simply be due to chance. Critics and Smith claimed that the system of jury selection caused the lack of fair and reasonable representation of black people on the jury.

Coding: Let's do a simulation in Python to understand what's going on. Write a simulation to show the expected distribution. Remember, from the case: we are trying to see if the difference between 8% and 3% is small and reflects no attempt to include or exclude a specified number of African Americans." In other words, we want to write a function to randomly sample 100 people from a population that has 8% Black people. The function should return the number of Black people in each jury. We can do this multiple times and see how rare it is to get only 3 Black people in the jury.

Write the Function.

Coding: Store the results in a dataframe so we can analyze this. Let's start by doing this simulation 1000 times. In other words, we are simulating picking 1000 juries. Store the results in a dataframe called df.

Simulate this 1000 times and store the results in a dataframe.

Coding: Create a histogram of the results.

Create a histogram.

Group Discussion: Interpret the results of your histogram. What does this tell us about this case? Do you think this could have happened by chance? If so, why? If not, why did some people claim it did?

Individual Reflection: Write down the most interesting part of your group discussion.

Coding: What was the expected value for the number of Black people in the jury? Enter your answer here. Next, calculate the mean of the dataframe that you created from the simulation.

Coding: Run the simulation again, but this time for 10,000 people. Store this data in a dataframe called df2. Create a histogram and find the mean (you can copy and paste your previous code, no need to do everything again from scratch!).
Simulation for 10,000 people.

Store the results in df2.

Create a histogram of the results.

Find the mean number of black people on the randomly selected juries.

Individual Reflection: How did your results change? What does this tell us about the more simulations you run?

Group Discussion: This is an example of how we can use statistics to help us solve real world problems. Discuss with your group how simulations and data science can be used to help address issues of racism specifically.

Individual Reflection: Give an example of a way that we can use data science to help address issues of racism. This can be something you discussed in your group or an example you are interested in.

Coding: Lastly, find the probability that we will get 3 or less Black men on the jury using df2. You may want to think back to conditionals- that will help here.

Puzzle #2: Names and Resumes + Simulation!

This next simulation comes from a study looking at female names. You can read the fully study here if you'd like: https://www.nber.org/system/files/working_papers/w9873/w9873.pdf

Here's a summary of what happened: Researchers created a standard resume and sent it out to about 500 different "Help Wanted" ads in Chicago and Boston. The resumes that they sent were identical, except they changed the names and email addresses to have either a "white sounding" name like Emily or a "black sounding" name like Lakisha, according to the study. Assuming that this resume has about a 20% chance of getting a call back and half of the resumes had white sounding names and half had black sounding names, let's do this simulation. We will simulate the results and then look at what actually happened.

Group Discussion: Before we start coding, let's think about this. When this study was done in real life, the white sounding names had 10.33% callbacks and the black sounding names had 6.87% callbacks. Some of the companies claimed this difference was due to chance. Do you think this is a significant difference? Why do you think this happened? Why or why not is this problematic?

Individual Reflection: Write down something that surprised you from your group discussion.

Coding: Let's do a simulation in Python to understand what's going on. Since the content of the resumes are identical, we should expect to get an equal number of call backs for both types of

names. In other words, picking the callbacks should be like random sampling. Write a simulation to randomly sample 20% of 500 resumes 10,000 times. Half of the 500 resumes should be represented as 0s (representing white sounding names) and half of the resumes should be represented as 1s (representing black sounding names). Write a function that returns the number of black sounding names and white sounding names that got selected. Simulate the random selection 10,000 times, and then store your results in a dataframe called df3.

Write a function. Do the simulation 10,000 times.

Store the results of the simulation in a dataframe (df3).

Coding: Calculate the average percent of resumes that got callbacks for each group of names in the simulation (these should be close to 10% for each group of names).

Calculate the average percent of resumes with black sounding names that got callbacks.

Calculate the average percent of resumes with white sounding names that got callbacks.

Individual Reflection: Think about how these percentages compare to the actual percentages (10.33% vs. 6.87%). What does this say about discrimination based on your name.

Beyond Stat 107

Group Discussion: The Smith trial happened in 1993 and the name study occurred in the early 2000s. Discuss with you group whether or not you think similar events still occur today and why. Reflect on how data science can be used to educate people about this.

Individual Reflection: Think about your takeaways from doing these simulations in this lab. Choose one of the following prompts to write a paragraph style response explaining what the simulation showed in either Puzzle 1 or Puzzle 2.

Option 1: Puzzle 1- Pretend that you are a defense attorney and a data scientist. Write a memo to the Supreme Court positioning yourself as a data scientist arguing whether or not your think the jury with 3 black men was randomly selected. Justify your decision and include guidelines for the future.

Option 2: Puzzle 2- Pretend that you are a data scientist arguing whether or not there was discrimination based on how applicants' names sound. Write a memo to the HR Department of one of the companies positioning yourself as a data scientist arguing whether or employers are biased against certain names. Justify your decision using what you've already done and include guidelines for the future.

Remember, you only need to do either Option 1 or Option 2 (not both)!