CHALLENGES FOR INDEPENDENT STUDENT PROJECTS:
A PERSPECTIVE OF THE TEACHER

Layla Guyot, Amanda Walker and Alexander White
Texas State University, San Marcos, TX
l_g244@txstate.edu

Including independent student projects in an introductory statistics course allows students to engage in the process of data analysis by applying and connecting concepts learned throughout the semester (Chance, 2002). Course projects provide opportunities for the students to investigate data in a context that is meaningful to them and to practice statistical analysis. From our experience, students may face different obstacles during the data analysis process such as posing ambitious statistical questions, collecting data from reliable sources, or going beyond the scope of the course. These challenges require the teacher to be an expert in many different fields. Case studies of student independent projects are discussed, enlightening challenges encountered from the perspective of the teacher and what worked to support students' opportunities to learn.

BACKGROUND

Including independent student projects in an introductory statistics course allows students to engage in the process of data analysis by applying and connecting concepts learned throughout the semester (Chance, 2002; da Silva & Pinto, 2014). Course projects provide opportunities for students to gain an authentic experience of statistical analysis. For example, projects promote statistical reasoning by not only applying statistical concepts but choosing what method to apply and how. The transition between simply applying statistical methods and recognizing which method is appropriate to use requires the development of statistical thinking and practices that grow out of experience (Pfannkuch & Wild, 2000). Thus, offering an authentic experience of statistical practice is recommended through projects (Anderson & Loynes, 1987; Chance, 2002; da Silva & Pinto, 2014).

In our semester-long projects, students are given the opportunity to experience the different steps of statistical process as recommended by the Guidelines for Assessment and Instruction in Statistics Education (GAISE) (Aliaga et al., 2005): formulating a question, collecting data, conducting statistical analysis, and interpreting the data. Students can explore a context that is of particular interest and collect data that is meaningful to them as suggested by Cobb and Moore (1997). They are required to gather data themselves, from surveys, experiments or from sources on the internet. Students are performing statistical analysis on data collected by themselves and are producing a final report that summarizes their methods and findings.

According to da Silva and Pinto (2014) projects help connect statistical concepts to practical applications. In particular, Spence, Bailey and Sharp (2017) found that students improve their knowledge in identifying appropriate analysis by conducting projects. Students also develop general professional skills such as communication, collaboration, or meeting deadlines (Chance, 2002; da Silva & Pinto, 2014). During statistical process, students may be challenged at different stages. da Silva and Pinto (2014) report that students are mainly struggling with the choice of an appropriate sampling method, the phase of data collection and the presentation of summary statistics in tables.

We share our experience of implementing independent student projects in an introductory statistics course by presenting two case studies of student projects, illustrating common challenges encountered from the perspective of the teacher, and how we, as teachers, found ways to support students' opportunities to learn.

CASE STUDIES

We introduce two case studies of projects that occurred during two of our introductory statistics courses. In the next section, we define challenges encountered from the perspective of the teacher and illustrate each challenge using instances from the case studies. We also discuss what decisions were taken by the teacher to address each challenge in order to support students' opportunities to learn.

Angela is a Wildlife Biology major and is enrolled in an Honors Elementary Statistics course. She is examining a possible correlation between human population density in Texas and the white winged-dove population.

Rose is a Sociology major and is enrolled in a College Algebra and Statistics course. She is investigating if temperatures are related to homicide rates in the United States.

COMMON CHALLENGES

The role of the teacher during the projects is to guide students through an authentic experience of statistical practice that allows them to work in a creative and independent way. From our experience of implementing projects in introductory statistics courses, we have faced challenges throughout the different steps of the statistical process by ensuring the authenticity of the project with the choice of genuine research questions and questioning the collection, organization, analysis, and the interpretation of the data. The teacher can also be challenged by the management of the projects over the semester.

Authenticity

Students are encouraged to work on a research question that is relevant to them, related to their hobby, major or a career they are interested in pursuing. However, students can pose ambitious or unreasonable research questions that are beyond or below the expectations of the project. The role of the teacher is to evaluate the appropriateness or difficulty of the research question and to balance the tension between class requirements and the research project.

Rose has reviewed some of the literature exploring the association between temperatures and homicide rates. She decides to mimic a study that was conducted in another country, claiming that warm temperatures do affect homicide rates. The teacher is concerned with the validity of the arguments coming from the literature and that the methods of statistical analysis employed by the study may be beyond the scope of the course. These two aspects challenge the authenticity of the project.

To ensure the authenticity of the project, the teacher could recommend a research question that is more appropriate and within the scope of the course or the teacher could work with students individually to teach more advanced procedures or discuss precautions in interpretations and limitations of the projects. Concerning Rose's research question, the teacher decided that pursuing the question will provide the opportunity for a productive discussion about different elements involved in the statistical process and Rose continued to explore her research question.

Data Collection

Once students have determined their research questions, they are required to gather real data. They can collect data from online resources, surveys, experiments, or any other sources. Different aspects of data collection might be challenging for the teacher to guide students. For example, the teacher may have to help students towards the accessibility of raw data, the identification of a sample, or the search of specific methodologies for some contexts.

To collect data, Rose decides to use a stratified sample by dividing the United States into five different regions: West, Midwest, Northeast, Southeast and Southwest. She chose one state randomly from each region to constitute her sample. The idea of stratified sample was introduced by the teacher as a sampling method. However, the analysis of such sample with the correction of standard errors was not covered in this course. Rose collects data about the number of homicides and the annual temperature for each state included in her sample by researching government-funded websites. Even though Rose collects data that comes from reliable sources, the teacher wants her to pay close attention to the definition of the two variables she collected.

The teacher has to make sure that students question and understand to what extent the sample they choose is appropriate and the data they collect is reliable. The teacher and Rose discussed the relative small size of her sample and decided she should collect data from all the states. In addition, the teacher interrogated Rose about the meaning of the variables involved by noting the difference between reporting the number of homicides and the homicide rate. Also, only considering the average annual temperature loses the information about the variation in temperatures in a state, and how the range of temperatures may vary from location to location.

Data Organization

In an introductory statistics course, students are usually given real data that are presented nicely and ready to analyze. In real life, statisticians have to deal with a data management phase which includes treating missing data, coding and recoding variables, or organizing data using statistical software.

Angela is using the census data from 1960 to 2010, taken every 10 years and she has two different sources for white-winged dove data, coo-counts and nest estimates per hectare. The coo-count data are estimates taken by volunteers walking in a 24 mile radius counting the number of white-winged doves seen or heard. Wildlife biologists agree that the coo-counts do not match the more reliable nest estimates. However, the nest data only goes back to 1990 and have several missing years, while the coo-count data stems from 1967. The teacher and Angela are concerned about what data to use for the analysis and how to match the white-winged dove data to census data since they are not given on the same years.

Treating missing data is not an easy task and the teacher can raise the issue of interpretation in such cases. For Angela's projects, the teacher recommended using the more reliable nest estimate data rather than the coo-count data to account for less bias, and to try and locate Texas human population data for the years in between the census. If population data cannot be located, the teacher suggested that Angela could interpolate the human population using population growth models.

Data Analysis

To apply statistical methods for analysis, students are often given examples and problems in class for which the method has already been identified. Through their project, students gain experience in choosing an appropriate method to analyze their own data. However, the latter may be outside the scope of the course or specific to the context and domain of application.

After organizing her data, Angela investigates a possible correlation between the Texas human population and white-winged dove population. Examining the scatterplot, it appears the relationship seems to be polynomial. However, only simple linear regression and quadratic regression were covered by the teacher in the course.

Introducing additional statistical methods can be an alternative for the teacher to guide students during statistical analysis. Angela had explored various models that do not fit her data and the teacher decided to extend the topics already learned, by creating a mini-lesson to demonstrate how to explore and interpret other models with the coefficient of determination and residual plots, using an additional software package.

Data Interpretation

The interpretation of the data and the results of the statistical analysis requires caution. Careful attention to detail should be given about the different steps of the statistical process.

Rose performs linear regression to investigate the association of temperatures and homicides rates in the United States. She finds that there is a positive moderate correlation between these two variables. The teacher reminds Rose to be cautious about her conclusions because of the definition of her variables. In addition, since correlation does not imply causation, nothing can be concluded about the effect of temperatures on homicide rates but only that there seems to be an association.

While interpreting the results of data analysis, the teacher can raise several issues that may have occurred during the statistical process. Indeed, students should be encouraged to discuss their interpretations in light of the method used to collect the data, question the reliability of the data, the appropriateness of the statistical analysis or if the assumptions of the models used are met. They need to understand and summarize the limitations of their interpretations.

Management

As the project is intended to be conducted throughout the semester, it is important to start working on it as early as possible. However, it can be very challenging for students to formulate a research question before they learn and experience what type of questions can be answered through statistical analysis.

Angela comes to the teacher with issues about her data only two weeks before the final report of her project is due. It gives little time for the teacher to help her develop a strategy to treat the missing data or investigate new models.

Throughout the semester, the teacher has to set different deadlines and check with students to report their advancement on their project. Valuable feedback from the teacher can ensure that the students can learn and benefit from the authentic experience given by the project.

PROMOTING STUDENTS' OPPORTUNITIES TO LEARN

The role of the teacher is crucial in monitoring independent student projects. Indeed, the teacher provides structure and support for conducting projects (da Silva & Pinto, 2014) and needs to develop and "refine the art of implementing projects" (Spence et al., 2017, p. 258). However, the teacher needs to balance between creating structure and relinquishing responsibility to the students. We combine previous literature (Chance, 1997) and our experience to identify key elements for implementing projects while promoting students' opportunities to learn.

First, expectations must be clearly set. A detailed syllabus could describe the different steps of the project with deadlines, rubric and grading systems emphasizing the effort that needs to be put at each phase of the project. Examples of successful projects could be given. Second, multiple checkpoints between the teacher and students should provide students with constructive and constant feedback for possible adjustments to the methods involved in the project. Projects can be integrated into the course by sharing challenges and opening discussion to the whole class. The teacher needs to decide if mini-lessons of statistical methods that are outside the scope of the course should be implemented during class or individually during office hours. Third, students should be encouraged to collaborate with others by working in groups if they have common interests, getting reviews and feedback on their work, or contacting experts from different fields. Indeed, the teacher may have a limited knowledge of the context and an expert from a specific domain (advisor, faculty member, or graduate students) could help make connections between statistical concepts and the context of application. Finally, students should be given the opportunity to reflect on how to improve their project and learn from what was done.

CONCLUSION

Implementing projects in an introductory statistics course can be challenging from both the perspective of the student and of the teacher. The teacher is required to be an expert in many different fields to guide students through the different steps of the statistical process. However, we found the teacher can overcome these challenges by raising issues and awareness about the limitations of the project. We highlighted with two case studies some significant and productive opportunities for the students to learn while conducting projects.

REFERENCES

Aliaga, M., Cobb, G., Cuff, C., Garfield, J., Gould, R., Lock, R., . . . Utts, J. (2005). *Guidelines for assessment and instruction in statistics education (GAISE): College report*.

Anderson, C. W., & Loynes, R. M. (1987). *The teaching of practical statistics*. John Wiley & Sons, Inc..

Chance, B. L. (1997). Experiences with authentic assessment techniques in an introductory statistics course. *Journal of Statistics Education*, *5*(3), 1-15.

Chance, B. L. (2002). Components of statistical thinking and implications for instruction and assessment. *Journal of Statistics Education*, *10*(3).

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly, 104*(9), 801-823.

da Silva, M., Porciúncula, M., & Pinto, S. S. (2014). Teaching Statistics through Learning Projects. *Statistics Education Research Journal*, *13*(2).

Pfannkuch, M., & Wild, C. J. (2000). Statistical thinking and statistical practice: themes gleaned from professional statisticians. *Statistical science*, 132-152.

Spence, D. J., Bailey, B., & Sharp, J. L. (2017). The impact of student-directed projects in introductory statistics. *Statistics Education Research Journal*, *17*(1).