EXPERIENCE WITH USING ELECTRONIC SURVEYS TO MEET THE SIX RECOMMENDATIONS OF THE GAISE COLLEGE REPORT

Bill Rybolt and Joe Aieta
Babson College, United States
rybolt@babson.edu

For several years we have used web-based surveys to collect data from the students in our introduction to probability and statistics course at Babson College. We believe that the use of such surveys and the data produced from them is an ideal way to meet the recommendations of the GAISE College Report. Our surveys are designed to collect data representing all four measurement scales: categorical, ordinal, interval, and ratio. In this presentation we give examples of the questions we ask and the surveys from which they come. As we go through each of the six recommendations of the GAISE report, we demonstrate how our use of our survey data contributes to achieving that objective and indicate areas needing improvement.

INTRODUCTION

The six recommendations of the GAISE report (2005) are:

1. Emphasize statistical literacy and develop statistical thinking
2. Use real data
3. Stress conceptual understanding rather than knowledge of procedures
4. Foster active learning in the classroom
5. Use technology for developing conceptual understanding and analyzing data
6. Use assessments to improve and evaluate student learning

Babson College is a business-oriented college located near Boston, Massachusetts. The introduction to probability and statistics course is typically taken in the spring semester of the freshman year. During the spring semester, we typically have a dozen sections of this course with an average of just over thirty students per section. All freshman students are issued laptop computers for use during their time at Babson. A complete Microsoft office suite of programs is included as well as programs such as Minitab which we use for statistical computations. Our classrooms are wired for Internet access and a few have wireless Internet access.

Several years ago we began using web based surveys to collect data from the students for use in the course. These surveys were originally designed to serve three purposes: 1) to give students experience with generating, cleansing, analyzing, and interpreting real data for which they have an intuitive feel, 2) to give the instructors information to allow us to better understand our students, and 3) to give us a mechanism to better evaluate and improve our course. Initially we had four sections participating in this process generating a data set with roughly one-hundred usable responses. We have been expanding the use of these surveys and hope to have all dozen sections participating during spring of 2006.

GENERATING, CLEANSING, ANALYZING, AND INTERPRETING REAL DATA

The major survey is given during the first week of class and consists of roughly twenty-five questions. These data are used during class to illustrate statistical computations and concepts. They are also used to generate questions for use on take-home and in-class exams.

Let us consider an example of a question and describe how we use it to meet the six recommendations of the GAISE report.

- About how many hours of sleep (to the nearest tenth of an hour) did you get last night?

If we use a textbox without constraints on the survey form, then we get responses such as: 3 to 4, less than 7, three and a half, and unsure. When students are presented with such original data for analysis, then they must cleanse the data. Clearly “three and a half” becomes 3.5 and “unsure” becomes “missing data.” What about 3 to 4? Does it become 3.5? What do we do with “less than 7”?

As students transition from cleansing to analyzing and interpreting the data, they engage in the first five of the GAISE recommendations. This simple exercise requires statistical thinking,

real data, conceptual understanding, active learning, and employs technology (laptops and Minitab) to calculate statistics, construct histograms, draw box plots, and think about the process.

In our course we have two midterm exams and a final exam. Each exam typically has two parts: a take-home part and an in-class part. For the take-home part, they are given a subset of the survey data. The students are then asked to use Minitab to perform a series of calculations on the data set and generate graphs and tables which they are to paste into a Word document. The take-home part is generally assigned one week before the in-class part of the exam. Thirty to forty percent of the in-class portion is based on the take-home part and requires that the students to use the printout of the Minitab session window, the Minitab project file, and the Word document to answer questions.

For example, consider the hours of sleep data, students might be asked: 1) what percent of the students get between 5 and 8 hours of sleep?, 2) what is the mean and standard deviation of the data?, 3) what is the 95% confidence interval?, and 3) is there evidence at a 2% level of significance that the average student gets less than 7 hours of sleep?

To foster active learning and reinforce the use of technology we use a different subset of the data and have the students practice with similar questions before the exam. The above tasks are strong on recommendations #2, #4, and #5. However, one might argue that we are weak on recommendation #1 and #3. To address this concern we hope to add tasks which place more on emphasis on conceptual understanding.

Thus, students might be presented with the following assignment. You are to produce a short paragraph description of student sleep habits at Babson. Your material is to be placed in a pamphlet to be given to parents when they visit Babson. Your material may include a table or graph if you deem it appropriate, but your presentation must be clear to the typical parent. First a student must identify the concept they are trying to communicate and then decide how they should present the results. It is really helpful to give parents means and standard deviations, confidence intervals, or even the results of a hypothesis test? Hopefully, students might even ask themselves, what do parents really want to know about the sleep habits at Babson?

On more fundamental level we hope to have the students address the wording of the question and the choice of the survey sample. Students would be asked if the data generated by the question are appropriate for answering the task above. If not, then how should the question be worded? Another concern for the students to address is, even if the question is worded correctly, can the data be used to describe the sleep habits of Babson students? Of what population does the data really represent the sleep habits?

In order to use the data from these surveys for class work and for homework assignments, take-home exam questions, and in-class exam questions, we find it necessary to relax slightly the "Use real data" recommendation of the GAISE report. Although we start with real data, we sometimes modify it slightly to achieve specific learning objectives. For example, to demonstrate the need to cleanse data, we occasionally change a few of the values before presenting the data to the students. We then ask the students to cleanse the data.

It would be nice, if we always got data which required the types of cleansing that we wish to illustrate. A purist would say that we are not using real data. Our counter argument is that we have made the data more valuable as a learning tool. We introduce the type of errors and need for cleansing that we have found in past surveys, and the students are likely to find in future surveys. We feel that this is a much better way to achieve a learning objective than to spend hours trying to find "real" data to make the point.

A second reason that we modify the data is to preserve the anonymity of the students. We ask questions such as: 1) what is your height?, 2) In what country were you born?, and 3) what was your highest score on the SAT I math exam? We either modify or organize the data in such a manner that individual students can not be identified. Again we feel that this modified data is just as valuable as "real data."

UNDERSTANDING THE STUDENTS BETTER

The GAISE report emphasizes the benefits from statistical thinking, and conceptual understanding using real data. We believe that it is important to include items on our surveys which help the instructors to better understand the students and the students to better understand

each other. For example we ask: how many minutes per day do you spend 1) reading and sending e-mail?, 2) surfing the Web?, 3) talking on a cell phone?, 4) text messaging on a cell phone?, and 5) instant messaging? These questions help us to understand how the students spend their time. The last question has helped us to appreciate the important role that instant messaging has come to play in student lives.

Of course whether we use box plots, scatter diagrams, two sample t-tests, or regression to investigate the relationships between these activities, we are trying to structure our investigations so as the foster recommendations #1 to #5. We have data involving both categorical and ratio data, and we give students pairs of variables. When they investigate the relationship, they need to decide whether to use chi-square, two sample *t*-test, or regression. Next the students are asked to do the calculations using Minitab and to explain the results first to a statistical expert and then to an average person. We are still trying to decide how to modify our exercises to better meet recommendations #1 and #3.

Other examples of questions which help us to understand students include: 1) how frequently do you read a newspaper? 2) how much time do you spend playing video games? and 3) how much time per week do you spend on academic activities outside of the classroom? Some of theses questions found on surveys given at the end of the semester and used in future classes.

IMPROVING AND EVALUATING STUDENT LEARNING

Recommendation #6 of the GAISE report is to use assessment to improve and evaluate student learning.

To help us improve and evaluate student learning, we give a short five- to ten-minute parameterized web based quiz at the end of each class. This quiz is based on the main ideas in that class and helps the students to identify the key concept and requires that they use the technology to analyze data and demonstrate simple understanding. If they do not get a ten on the quiz, they are encouraged to retake the quiz until they get a ten. After they take the quiz the first time, they are encouraged to help each other master the material. Since each quiz asks a slightly different question, casual cheating is not possible. If a significant number of students do not master the quiz, this tells the instructor that more effort is needed to help the students master that material.

We use an "EDU Review and Quiz Improvement Survey" to determine what we can do to improve the quizzes. We also employ a "Course Improvement Survey" to help us better understand the students as well as to improve the course.

One of the problems associated with the use of technology and the Internet during class is the temptation to engage in other activities and to multitask. Examples of questions we have used to help us better understand and deal with this issue are : 1) distractions such as e-mail, Web surfing, computer games, instant messaging, audio and video material, is engaged in by about __ percent of the students during class, 2) these distractions are best handled by ___, and 3) during the exam, the percentage of students engaging in inappropriate behavior is___ .

CONCLUSION

Our surveys generate significant amounts of data which we have used in a variety of ways to meet the requirements of the GAISE report. We currently require that all of our students participate in generating and using real data. In each class, students use technology to analyze and interpret real data. We are modifying the ways we view and use this data to place more emphasis on developing conceptual understanding and promoting statistical thinking.

ACKNOWLEDGEMENTS

This work was made possible by a grant from the Davis Educational Foundation.

REFERENCES

GAISE College Report. (2005). Guidelines for assessment and instruction in statistics education. The American Statistical Association, http://www.amstat.org/education/gaise/.
Members of the GAISE Group: Aliaga, M., Cobb, G., Cuff, C., Garfield, J. (Chair), Gould, R., Lock, R., Moore, T., Rossman, A., Stephenson, B., Utts, J., Velleman, P. and Witmer, J.