Incorporating Student Projects into the Introductory Statistics Class

Katherine Taylor Halvorsen
Smith College, Northampton, MA
khalvors@smith.edu

January 13, 2018

Overview of this talk

- Setting for these projects
- Define what I mean by a student project
- Outline some types of student projects
- Overview of how I conduct student projects in my courses
- Some examples of projects
- How projects fulfill the 2016 College GAISE Report recommendations
- How projects can support learning statistics

Setting for my course projects

- Smith is a women's college with about 2600 students
- The academic year consists of two 13-week semesters
- We use projects in:
  - A basic statistics course for social science majors
  - A course for science, engineering, math, and data science students that has a calculus prerequisite
- Class size 40 meets MWF for 70 min; lab 80 min
- Text is: *Introductory Statistics with Randomization and Simulation,* from OpenIntro

What is a student project?

- Students produce a statistical study
- Usually conducted over entire semester
- Working semi-independently
- Typically working in small groups of 3
- Find an area of research that interests them
- Formulate a question, collect and analyze data
- Interpret and present their results to their peers and to faculty.

Some variations on student projects

- Allow individual projects
- Coordinate projects with a plant BIO course
  - Stat students act as stat consultants to BIO groups
- Use a single class-wide project in a given semester
- Service learning projects

Overview of how I teach projects:

- Announce in syllabus; worth 20% of final grade
- “Class Projects” handout describes scope of project, gives general rules, outlines proposal
- Moodle serves as a resource site for
  - instructions
  - forms
  - examples
- Moodle provides the location for collecting deliverables
- Some class and lab times are reserved for projects
- Final report presented in class as talk with slides or poster

Class Projects Description

The class project gives you the opportunity to apply the statistical methods you learn in class to a research question that interests you. Please work with classmates in teams of three to formulate a research question and write a proposal explaining how you will conduct your research. You will review proposals written by other classmates, and you will receive reviews of your proposal from your peers, and from the instructor. After your research proposal has been approved, you will carry out your data collection and analysis. Finally, you will write a report describing your methods and results and give a short presentation of your work to the class.

Project Proposal

- Title Page (who)
- Introduction (what and why)
- Methods (when, where, how)
  - General issues (primary vs secondary data)
  - Sampling
  - Data description: units, variables, predictor, response
  - Statistical analyses
  - Anticipate outcomes
- Appendices (if needed)
  - Data collection forms
  - Third-party Information
  - CITI training certificate

Methods: General Issues

- Type of primary study
  - Original experiment
  - Survey
  - Other observational study
- Primary vs secondary data
  - Collected by students for our class
  - Collected by students in another setting (summer research, another class, etc.)
  - Obtained data from another faculty member's research
  - Internet data

Methods: Sampling

- What are the **units** in your study (objects, people, animals, countries, studies)?
- What is the **population** (or populations) from which the units are drawn?
- **How many units** will you collect?
- **Explain how you will collect the data**. Provide enough detail so that someone could follow your instructions and collect the data for you.
- Will you employ **randomization** in your study or was it used when the data was collected by someone else?
- Will **sampling bias** in the selection of your units be an issue for you?

Methods: Data description

- What variables will you measure on each unit (you should measure a minimum of two variables).
- How will you collect measurements from the units?
- What are your response (outcome) and explanatory (predictor) variables?
- Are your variables quantitative (continuous or discrete) or categorical (nominative or ordinal)?

Methods: Statistical Analysis

- Do you plan to look for an association between your variables or a cause and effect relationship?
  - Give your best guess about the types of relationships you expect to see between the response and each explanatory variable.
- Note that your statistical analysis **must include** at least one of the following tools:
  - an estimate for the difference between 2 means
  - an estimate for the difference between 2 proportions
  - a test for a difference between 2 (or more) means
  - a test for a difference between 2 (or more) proportions
  - a test for the independence of two categorical variables
  - a test for the significance of the slope(s) in a regression model

Methods: Anticipated Outcomes

- Provide a sketch of one table and one or two graphs you could use to display your data after it has been collected.

- What do you expect the graphs to look like?

Some Moodle Resources

- Project proposal instructions
- Research question form
- Examples of research questions used in past
- Sample executive summary
- Sample final research report
- Sample slide presentation (or poster)
- *A Guide to Writing Technical Reports*
- List of project deliverables and due dates
- List of presentations, dates, participants

Moodle Deliverables (One per project group)

- Initial research question and justification
- Proposal
- Project data
- Summary statistics
- Inferential analyses
- Executive Summary
- Final paper, slides, and/or poster
- “Reflections on the project” questionnaire (individual responses)

MTH 220: Project Research Question

Name(s) of project investigators and email addresses:

1. _____ _____
2. _____ _____
3. _____ _____

What is your research question?

Topic or title of proposal:

Type of proposal: Experimental Observational

Type of data: **Primary Data (you collect it)** **Secondary Data (someone else collected it)**

Type of units: **Human Subjects** **Non-human Vertebrate Subjects**

Other (explain):_____

What is your statistical question?

Response/Dependent variable:

What variable will you use?

What are the units of measurement for this variable?

Explanatory/Predictor/Independent variables:

What variables will you use?

What units of measurement for each of these variables?

What comparisons will you make in your data?

Reflections on the Project
MTH 220: Introduction to Probability and Statistics

We would like to know whether you learned anything new from the process of producing your project, and we would like to know how well your group worked together. Would you please complete the following set of questions about your project. Please sign and date this form before you hand it in.

1. Have you produced a project in a course or on a research team before you completed the project for MTH 220? If yes, please explain.

2. What have you learned from the **process of conducting your research** for the MTH 220 project?

3. What have you learned from the **process of producing the poster** for the MTH 220 project?

4. In your opinion, how well did your group work together in producing your project?

5. Please estimate how much of the total work for the project was completed by each group member.

Group member 1: _____ Per cent of project work_____

Group member 2: _____ Per cent of project work_____

Group member 3: _____ Per cent of project work_____

The Data Science Part

- Our courses have an 80-minute weekly lab that is used primarily to teach R in the first half of the semester and to explore and analyze their project data in the second half.
- Students use R statistical software for project data management and analysis.
  - Students use OpenIntro labs written by Andrew Bray that are freely available on the internet for learning R.
  - Students learn RMarkdown for report writing and homework.
- I intend to add web scraping skills to the lab to help students who collect their data from the internet.

The Challenges

- Using human subjects
  - Requires CITI training and IRB approval
- Missing data
  - Use complete cases or
  - Fill in if only a few missing values
  - Find a substitute variable
- Students missing deadlines is rarely a problem
- Lack of cooperation within a group
  - Meet with group to discuss professional behavior
- Enormous amount of grading
  - Only one paper per deliverable per group
  - Use student TA and grading rubrics for interim deliverables
  - Check grading on about 10 %
  - Grade final paper myself

Examples of Small Group Projects

- Who becomes a Smith math major?
- Do Daykin Animal Shelter callers get the information they seek?
- Is there a difference in flight distances between Smith campus squirrels and Child's Park squirrels?
- A Study of Spatial Differences in Claw Thickness and Strength in *Carcinus maenas*
- Parasite-induced galls in the goldenrod plant *Solidago altissima*: Effects on plant health and the relationship between gall size and secondary parasitism
- Do basketball teams that use more "assists" score higher points?

Example of a Class-Wide Project

- Committee on Educational Technology survey, Fall 2002 MTH 107, 24 students
  - CET provided survey
  - Students formed teams and did in-person interviews
  - Students entered data into Excel and did analyses in MINITAB
  - Provided written report to CET

Examples of Service-Learning Projects

- Northeastern University College of Business Administration, Introductory Statistics Course, Analysis of Professional Accounting Alumni survey.
  - Students entered data, analyzed, and provided summary report.
- Students analyzed results of a survey conducted by the local realtors association
- Survey of apartment rents in Northampton for the city housing office
  - Students designed survey with the city administrator
  - Did phone interviews in a local relator's office after hours

Course projects fulfill the 2016 College GAISE Report recommendations

- Projects teach statistics as an investigative process.
- Students must formulate a question, collect and analyze data, and interpret and present their results.
- The process requires students to think statistically and to understand and apply the concepts they learn in the course.
- The project requires active engagement and hence constitutes active learning.
- Students use technology to collect, store, share, clean, explore, and draw inferences from data.
- Projects are inherently multivariate.
- Students receive regular assessments throughout the process. Some are verbal and some written, both from the instructors and from their peers through peer reviews.

Projects support learning statistics

- "The sorts of skills that teachers want for their students – such as the ability to analyze and think critically – require extensive factual knowledge."
  - Projects build a factual knowledge base
- Projects build and reinforce neural connections between concepts
  - Projects require students to apply concepts in real time and make connections between what they see in class and what they have to do for their own projects.
- Projects provide intrinsic motivation to learn statistical concepts
  - Students choose their research question, capturing their attention, providing a sense of purpose and the camaraderie of a shared social experience.

Many Thanks to:

- Douglas Zahn, Florida State University, Emeritus 2005, who provided the first full description of a class project.
- Steve Wang, Swarthmore College, who in 2002 encouraged me to hold Poster Sessions at end of the course and provided his project instructions describing his poster sessions.
- Nick Horton, Smith and Amherst Colleges, who pushed me into switching from teaching with MINITAB to teaching with R.

Recommended Reading

- Ambrose, S.,Bridges, M.,DiPietro, M.,Lovett, M., & Norman, M. *How learning works: 7 research-based principles for smart teaching*, Jossey-Bass, 2010.
- Bailey, B., Spence, D., and Sinn, R., *Implementation of Discovery Projects in Statistics*. Journal of Statistics Education, Vol. 21, No. 3, 2013.
- Halvorsen, K.T. and Moore, T.L. Motivating, Monitoring, and Evaluating Student Projects, *Teaching Statistics: Resources for Undergraduate Instructors,* T.L. Moore, editor, MAA, 2000.
- Lang, James M., *Small Teaching*, James M. Lang, Jossey-Bass, 2016.
- Willingham, D. *Why Don't students like school? A cognitive scientist answers questions about how the mind works and what it means for the classroom,* Jossey-Bass, 2014.