STAT 140: Introduction to the Ideas and Applications of Statistics

Course Information

- Meeting Time: TTH 8:30AM-9:45AM; W 8:30AM-9:20AM
- Course Website: marievozanne.github.io/STAT_140_Homepage.html
- Email: mozanne@mtholyoke.edu
- Office: 403 Clapp Laboratory
- Office Hours: To be decided by the class. This document will be updated when office hours have been set.

Course Description

We live in an increasingly data-driven world, where it is important to leverage the available information to make decisions and also to question the data/conclusions that are presented regularly in the news, for example. The main objective of this course is to equip you with the skills to enable you to think critically about what data can and cannot tell us, and to be an informed consumer and producer of quantitative/statistical analyses.

STAT 140: Introduction to the Ideas and Applications of Statistics lays the foundation for descriptive and inferential statistics, including constructing statistical models from data. You will learn to produce meaningful graphical and numerical data summaries, apply basic probability models, and employ statistical inference methods using computational tools. Topics of study include: basic descriptive and inferential statistics, visualization, study design, and multiple regression.

Key Learning Goals:

1. Students should become critical consumers of statistically-based results reported in popular media, recognizing whether reported results reasonably follow from the study and analysis conducted.
2. Students should be able to recognize questions for which the investigative process in statistics would be useful and should be able to answer questions using the investigative process.
3. Students should be able to produce graphical displays and numerical summaries and interpret what graphs do and do not reveal. Students should also understand how graphical summaries can be manipulated to mislead readers.
4. Students should recognize and be able to explain the central role of variability in the field of statistics.
5. Students should recognize and be able to explain the central role of randomness in designing studies and drawing conclusions.
6. Students should gain experience with how statistical models, including multivariable models, are used.
7. Students should demonstrate an understanding of, and ability to use, basic ideas of statistical inference, both hypothesis tests and interval estimation, in a variety of settings.

8. Students should be able to interpret and draw conclusions from standard output from statistical software packages.
9. Students should demonstrate an awareness of ethical issues associated with sound statistical practice.

Time Commitment: This is a 4 credit course. I recommend that you budget 12 hours a week outside of class for reading, homework assignments, and study. This time commitment may vary over the semester, but if you are regularly spending much more time than 12 hours a week on this course outside of lecture, please let me know.

Textbook

The textbook for this course is *OpenIntro Statistics* (4th Edition) by Diez, Çetinkaya-Rundel, and Barr (ISBN-13: 978-1943450077). This book is available as a free pdf through `https://www.openintro.org/stat/os4.php`. If you prefer, a paper copy can be purchased (Amazon, $20). Reading will be assigned before class, and you will be expected to be able to discuss what you read at the beginning of class.

Computing

We will be using the open source `R` statistical programming language in this class. `R` is an important programming language, both in academic and industry statistics and data science positions. In this class, you will learn enough about `R` to conduct basic data analysis tasks. Note, however, that the main focus of this course will be learning statistical methods. No prior programming experience is needed; `R` code will be presented in a way that is accessible to all students.

It will be important to bring your laptop to class, as we will be using `R` on most days. If you do not have a working laptop or you forget yours, please let me know - the department has laptops available for you to use in class.

Assignments

Homework: Homework is the most effective way to reinforce concepts learned in class. There will be regular homework assignments. Often, questions will relate to material in the reading that will be covered in class. Homework is due at the start of class, and unless indicated otherwise, will be accepted with a 25% penalty if turned in within the shorter of 48 hours or the next class period (and no credit otherwise). Extensions may be possible, but need to be requested well before the deadline. About one third of the credit for homework assignments is awarded for effort, and substantial partial credit can be awarded even if there are mistakes in your work. This means that you should turn in whatever work you have done for partial credit!

Exams: Two midterm exams will be given during the semester. There will also be quizzes throughout the semester; some will be announced, while others may be unannounced (i.e. "pop" quizzes). The two midterm exams and quizzes will comprise 50% of your grade. A final exam will be given during the final examination period (worth 20% of your grade). All exams will be closed book, although some may include an open-notes take home portion. You may bring a calculator to the in-class portions of the exams. During exams and quizzes, you may only communicate with the instructor/proctor.

Project: There will be one project (worth 15% of your grade) that you will work on over the course of the semester, culminating in a short final paper due at the end of the term. In the first two weeks of class, you will identify a question and data set of interest that you will use throughout the class. You can use a problem/data set that you find in a journal article, newspaper article, or that you have collected as part of research for another class/independent study/thesis project. If you have trouble identifying a problem/data set, I will help you find something. Throughout the semester, you will apply methods that we learn in class to these data and reflect on whether any of these methods help you address the question you had in mind. The point of these exercises is to practice what we are learning in class while reflecting on the strengths and limitations of the material. In your final paper, you will summarize the work that you did throughout the semester AND you will include a discussion of an advanced method that we have not learned that you think would better address your problem of interest. Note, you will NOT implement this advanced method, but I expect you to consider its strengths and limitations (e.g., what assumptions does it rely on?) and to argue for its suitability for your question and data. This final paper is intended to broaden your knowledge of statistical methods so that you will (a) want to study more statistics (yay!) or (b) know what questions to ask as you work with data or collaborate with statisticians/data scientists (also yay!).

Extra Credit: Extra credit can be earned by: (a) attending an out-of class lecture and writing a critique of it (will be announced), (b) pointing out a mistake in the textbook or class notes, or (c) sharing an interesting data set or news article (that uses statistics) with me. Extra credit will be applied when a student is near a letter grade boundary.

Grading: When grading your written work, I am looking for technically correct solutions and clearly explained reasoning. Numerically correct answers, alone, are not sufficient on homework, tests, or quizzes. You should turn in neat, organized work, with brief, clear answers that explain your thinking. If I cannot read or follow your work, I cannot give you full credit for it. The grade weights are as follows.

	Percentage
Homework	15%
Midterm Exams and Quizzes	50%
Project	15%
Final Exam	20%

Extra help: In addition to my scheduled office hours, office hours you schedule, or emails to me, there are other resources available to you. There will be drop-in student help hours during several evenings each week. The times and locations for these will be posted on the course website. Your fellow students are also an excellent resource - remember that you can work on assignments and study with your classmates!

Policies

Attendance: Attendance in lecture is critical for success in this class. If your absence is unavoidable, please notify me in advance and I will do my best to accommodate you. Note, it is unusual to have more than one necessary absence during the semester.

Collaboration: You are encouraged to *work with other students in the class* to review material, complete homework, and study for exams and quizzes. You must, however,

write up your assignments on your own. I suggest that after working a problem with a partner or a group, you go back and do the problem on your own, without looking at the solution. This is an excellent way to check your understanding. Copying verbatim paragraphs or large amounts of `R` code from another student is neither acceptable nor useful to you as you learn the material; any such copying will receive no credit and may receive a penalty. You may not communicate with anyone but the instructor/proctor during exams or quizzes. All students, staff, and faculty are bound by the Mount Holyoke College Honor Code. Any instances of dishonesty, plagiarism, etc., will be reported.

Note, portions of this syllabus were adapted from one developed by Professor Evan Ray.