MATH 141: Introduction to Probability and Statistics - Fall 2021

Syllabus

About the Class

Instructor: Alex John Quijano (He/Him/His) **Email:** aquijano@reed.edu
Office: Library 394
Office Hours: *MWF 3PM - 4PM and TuTh 4PM - 5PM (or by appointment) - Library 394*
Day & Time: Lectures: *F03: MWF 9:00 AM – 9:50 AM, F04: MWF 10:05 AM – 10:55 AM*
Labs: *F33: Tu 8:50 AM – 10:10 AM, F31: Tu 10:25 AM – 11:45 AM, F32: Tu 12:00 PM – 1:20 PM*
Rooms: Lectures: PHYSIC 240A
Labs: ETC 205
Office Hours - LAs: Maxwell - MW 4:30 PM - 6:00 PM (ETC 105b)
Gillian - MTu 7:30 PM - 9:00 PM (reed-edu.zoom.us/j/94604917095)
Robin - Thu 4:30 PM - 6:00 PM (ETC 105b)

Description: An introduction to exploratory data analysis, regression modeling, hypothesis testing, point and interval estimates, and resampling methods. First, we discuss the basic idea of study design and data exploration. Next, we introduce linear and logistic regression modeling. Then, we introduce the concept of inference. After an introduction of hypothesis testing, we proceed in discussing statistical inference for proportions, two-way tables, and means. Finally, bringing together the ideas of regression modeling and inference, we discuss inference for linear and logistic regression. This course is designed for a student to learn statistical concepts and their applications to the natural sciences, sociology, and the humanities. This is a data-driven approach to introductory statistics and focused on training a student to give clear and organized written and verbal explanations of statistical ideas to a wide audience. In the interest of scientific transparency, replicability, and reproducibility, part of this course introduces the R programming language for data exploration, visualization, and statistical computing.

Distribution Requirements: This course can be used towards your Group III, “Natural, Mathematical, and Psychological Science” requirement. It accomplishes the following learning goals for the group:

1. Use and evaluate quantitative data or modeling, or use logical/mathematical reasoning to evaluate, test or prove statements.
2. Given a problem or question, formulate a hypothesis or conjecture, and design an experiment, collect data, or use mathematical reasoning to test or validate it.
3. Collect, interpret, and analyze data.

This course does not satisfy the “primary data collection and analysis” requirement.

Learning Objectives: Upon completing MATH 141, students should understand the following:

- The basic structure of data.
- The evaluation and interpretation of descriptive statistics.

- The concept of statistical modeling and its application to data.
- The formulation of a hypothesis for a particular study design.
- Basic problems of inference for proportions, two-way tables, and means.
- Basic concepts of random variables and probability theory.

Learning Outcomes: Upon completing MATH 141, students should be able to do the following:

- Understand and explain the basic structure of data, how it is collected, evaluate its limitations, and how they affect the results of inference, and represent data graphically and compute statistical properties of data.
- Able to use linear or logistic regression models appropriately, understand and explain the concept of statistical inference, and understand and explain the basic concepts of random variables and probability theory.
- Able to apply methods to analyze single variables or the relationship between two variables, and interpret results correctly in the context of a given data.
- Able to critique claims and evaluate decisions based on data, and write and run simple R codes for data analysis, statistical modeling, and basic visualizations.
- Complete a written project report demonstrating a comprehensive understanding of data exploratory analysis, statistical modeling, and statistical inference.

Lecture Structure: Lectures will occur synchronously during their scheduled time. The first 30 minutes of class will be a lecture/presentation based on the reading assignment assigned for that day. The last 20 minutes of class are dedicated to questions and discussions.

Lab Structure: Labs are dedicated to R programming tutorials for the purpose of helping students to complete their lab reports. The first 20 minutes of the lab will be a demonstration of R coding in relation to the given topic of the current week. The remaining 60 minutes of lab time will be dedicated for students to start their lab assignments, ask questions to the instructor and their fellow peers.

Class Materials and Resources

Class Website: The syllabus, tentative schedule, lecture slides, homework, labs, project information, and all other class materials are posted on the course website, reed-statistics.github.io/math141-fall2021**.**

Textbook: *OpenIntro: Introduction to Modern Statistics* (2021) by Mine Çetinkaya-Rundel and Johanna Hardin, First Edition. The textbook is free and open-source. We will use this textbook for all assignments (see tentative topic schedule).

Computing: This class will use the R programming language and the R Studio Integrated Development Environment (IDE). There are several ways to use R and R studio. Below are two ways to access R and R Studio.

- Reed College has an R Studio server where you can log in using your Kerberos credentials. Go to this link, rstudio.reed.edu.

- Downloading and Installing on your own computer. First, you need to install R; r-project.org. Next, you can install R Studio; rstudio download.

For more information and resources, visit Reed R resources.
Gradescope: Homework, lab, project assignments, and exams are submitted through Gradescope. Please sign-up as a student with your Reed email using this Gradecope entry code: 2R8X6V.

Slack Workspace: Please signup for the Class Slack Workspace using your Reed email during the first week of class. We will be using Slack Workspace: math141-quijano.slack.com as the main real-time communication tool; from general announcements and question-answering to direct messages. Please check the Slack Workspace regularly. Concise and specific messages are helpful. If you prefer communicating through email, note that the instructor has set up an email filter for this course and you must put the "MATH 141" keyword in your subject line. It is easy for the instructor to get notice of your email if you put the keyword in the subject line.

Class Assessment

Homework: Homework assignments will be posted on the class website every week and due in one week. Problems selected for homework will be similar to those exercises written in the textbook and readings.

Lab Reports: Lab assignments will be posted on the class website every other Tuesday and due in two weeks. Problems selected for the creation of the lab report will be similar to those labs presented in the textbook. *Amendment: Starting October 25, 2021, the lab exercises will be integrated into the homework assignments, meaning there will be one assignment submission every week.*

Midterm and Final Exam: There will be a midterm exam in the middle of October with topics from the first day of class. A comprehensive final exam will be conducted during finals week in December.

Project: The final report is a group research project that demonstrates a comprehensive understanding of basic data analysis, statistical modeling, and inference. Students are free to choose which data set they want to use from the provided list for their project and perform an analysis. The timeline of the final report is done in five phases:

- Phase 1: Group formation
- Phase 2: Data selection, study design, data exploration, and framing research questions
- Phase 3: Hypothesis formulation, more data exploration, and testing out methods as appropriate.
- Phase 4: Testing out more methods, methods fine-tuning, and writing the interpretation, discussion, conclusions of the results in the context of the data
- Phase 5: Putting the entire report together as one scientific narrative

Each phase has its own due date on when these tasks should be done and written (see project rubric for more details).

Attendance: It is strongly recommended that you attend classes and labs promptly. Participation is an important part of learning statistics. Be prepared to participate in the discussion by doing the assigned readings and homework every week.

Grading: Each assignment will contain questions where each is graded according to the grading guide. The project submissions are also graded this way. Each project phase is graded according to the grading guide.

Grading guide for conceptual or mathematical questions:

- *5 – Outstanding*; showed full understanding of the material. Congratulations!
- *4 – Excellent*; showed almost full understanding but with minor errors. Well done!
- *3 – Acceptable*; showed some understanding but okay despite a few errors. Good!
- *2 – Needs Improvement*; showed some potential but it needs more work. Okay!
- *1 – Needs Major Improvement*; at least you tried, E for effort!
- *0 – Incorrect or no submission*; meh.

Grading guide for multiple choice questions:

- 1 – Correct.
- 0 – Incorrect.

Extra Credit: Throughout the course, there will be opportunities for extra credit. You can submit at most two extra credit assignments. The extra credit assignment grade is added to your lab report grades. Extra credit can be from any of these two categories:

- *A critique on the statistical method, visualization, and/or analysis from a chosen article or news source*, which involves writing a critical essay (2-3 pages and single-spaced) regarding the statistical analysis of a chosen scientific article or news source. The essay must include a summary of the article, a description of the data used, and statistical method used, and a description of a better statistical method, a better way to visualize the results, or a comment on statistical errors/pitfalls if it exists.
- *Create an informative and visually appealing visualization of complex data*, which involves creating a visualization of a chosen data set. You can use tools (R, Python, etc.) to create the visualization. The resulting visualization must include a half-page description of how to read/interpret it and what are its weaknesses.

Class Expectations[1]

Inspirational Talk: Learning statistics and the R programming language are like learning two new languages simultaneously – like Spanish, French, Mandarin, or Tagalog. Learning

[1] Some of the statements in this section are borrowed from the Reed Center for Teaching and Learning Syllabus Policy Blurbs.

probability and statistics can help you on your research/thesis and it can help you statistically assess and critique someone's research work, arguments or claims. It can be difficult and frustrating when learning R if you have little or no experience with computer programming. Many experienced statisticians, mathematicians, and computer scientists still get frustrated occasionally when writing R codes or any programming language. Part of the R learning experience is frustration and self-denial. These are valid emotions and will slowly fade over time. Once it clicks you will feel joy and excitement but probably with a hint of skepticism making sure the code does what it is supposed to do. If you find yourself stuck or taking a lot of time to successfully execute a code snippet, talk to your fellow peers, ask questions, and most importantly send the instructor a message. Take a break and do something fun - like eating - and try R coding again. You can do this!

Academic Honor Principle: We are committed to adhering to the standards regarding academic honesty contained in the honor principle and the values of mutual trust, concern, and respect for oneself and for others upon which the Reed community depends. In class, give your undivided attention to others. If you don't agree with what someone else has to say, you are encouraged to express your point of view, but do so respectfully, and support your claims with textual evidence.

In your written work, follow the conventions of an appropriate citation for your respective discipline or major. Please consult with the instructor if you have questions about proper citation.

Academic Support: We expect you to participate in the class through lectures, discussion, labs, and other engagements. We also expect you to make use of opportunities to get help outside of class (office hours, Slack, email, tutoring) if you need help. Concise and specific messages are the most helpful.

The Writing Center offers free appointments and experienced peer tutors who are there to help you at any stage of the writing process. I strongly encourage even experienced writers to take advantage of these services. For more information, start here: Reed Writing Resource.

Library: Reed's subject librarians Librarians can help you locate and access subject-specific resources for projects, classes, and thesis. Do not hesitate to turn to them!

Late Assignments and Incompletes: You are expected to turn in all completed assignments on time. Circumstances that may disallow you to turn in your work on time – such as a medical reason – are understandable. Please let the instructor know if you are unable to submit your work and have missed the deadline way beyond its original posted date. Because every assignment is an important aspect of your learning in this class, we will discuss when you will turn in the assignment as well as decide upon an acceptable consequence for your turning it in late. We are committed to successfully helping you learn statistical concepts from this course.

Collaboration Policy: It is encouraged that students participate in discussions regarding homework and lab assignments. However, each student must take responsibility and ownership of their work and submit their work individually.

Accommodations: We will make every effort to accommodate students whose personal obligations lead to scheduling conflicts. Please speak with the instructor during the first two weeks of class regarding any potential accommodations that may arise.

If you have a disability or think you may have a disability, you may also want to meet with Disability & Accessibility Resources (DAR) to request an official accommodation. You can find more information about DAR, including contact information, here: Reed Disability Resources.

If you have already been approved for accommodations through Disability & Accessibility Resources, please meet with the instructor outside of class so we can develop an implementation plan together.

Inclusion and Diversity: The natural and mathematical sciences are often viewed as objective disciplines. Science is a method for us to understand how the world works. However, it is historically built from a small set of privileged populations that often ignores the biases. We acknowledge that there may be some parts in this course that have overt and covert biases. Science is a human endeavor, and the pursuit of knowledge and skill must incorporate a diverse set of experiences.

We value all students regardless of their background, country of origin, race, religion, ethnicity, sexual orientation, disability status, etc. We are committed to providing a climate of excellence and inclusiveness within all aspects of this course. If you have any concerns, issues, or challenges, you are encouraged to discuss with the instructor (set up a meeting by email or a direct message in the Slack Workspace) with the assurance of full confidentiality except for academic integrity code violations or sexual harassment.

Pandemic Related Statements: You probably know people who have lost their jobs, have tested positive for COVID-19, have been hospitalized, or perhaps have even died. You all have increased (or possibly decreased) work responsibilities and increased family care responsibilities - you might be caring for extra people (young and/or old) right now, and you are likely facing uncertain job prospects (or have been laid off). We are fully committed to making sure that you learn everything you were hoping to learn from this class. Please let the instructor know or your academic advisor if you are facing difficulty or issues because of the pandemic.

Note that this is an in-person class and masks are required indoors. Therefore, when your health allows, you are expected to be present and engaged in class. At the same time, each community member has an individual responsibility to help prevent the spread of the coronavirus. Following public health guidance is part of living in an honorable community. The following recommendations should guide your decision about coming to class:

- Self-isolation is the recommended course of action for anyone experiencing flu-like symptoms, whether due to possible coronavirus or to other illnesses. Please stay at home if you feel sick and contact the Health and Counseling Center (HCC) or your healthcare provider to discuss. This is especially important if you think you may have an infectious disease.
- You should not attend class if you have tested positive for COVID-19 in the last 10 days, or if you have received notification or advice from the college or a health professional (including HCC staff) to quarantine or self-isolate.
- The CDC suggests that people with the following symptoms may have COVID: fever or chills, cough, shortness of breath or difficulty breathing, fatigue, muscle or body aches, headache, new loss of taste or smell, sore throat, congestion or runny nose, nausea or vomiting, diarrhea. As always, please consult a medical professional (members of the HCC or otherwise) if you have any questions about your health or health safety.
- If you suspect or know you have been exposed to a case of COVID-19, contact the HCC right away to discuss your next steps. For more information, visit the CDC's webpage on isolation and quarantine.

If you need to miss a class, or series of classes, due to illness, self-isolation, and/or quarantine, you are responsible for emailing the instructor as soon as possible. You are also responsible for coordinating with the instructor to complete work that you might miss due to absences.