ECON 4803 Introduction to Data Science for Economics

Georgia Institute of Technology

Fall 2026

Course Information

Instructor: Dr. Stefan Faridani

Drop-in Hours: see Canvas

Class time: see Canvas or OSCAR

Classroom location: see Canvas or OSCAR

Course Description

Governments, companies, and charities are increasingly using “big" data to guide decision making. This course builds up foundational concepts and skills to help data scientists answer questions about the economic world. In the first half of the course we learn the fundamentals of data analysis in R, causal inference, and prediction. In the second half we learn what makes a dataset “big” and how to make accurate predictions about economic variables when the data contains many predictors. We will use real datasets to study topics which may include: the markups charged by online retailers, rents charged by airbnb units, salaries paid to athletes, and the prices of used cars. No prior experience with R or coding is necessary or expected.

Recommended Preparation

ECON 2250: Statistics for Economics (or equivalent or permission of instructor)

This course assumes that students have already seen some basic probability and statistics concepts. But no prior experience with coding or R is expected. I will review basic probability and statistics, so it’s okay if you are rusty at first.

Required Textbook

Data Analysis for Social Science: A Friendly and Practical Introduction, by Elena Llaudet and Kosuke Imai. (Princeton University Press). First Edition.

Optional Textbooks

An Introduction to Statistical Learning: with Applications in R, by Gareth

James, Daniela Witten, Trevor Hastie, Rob Tibshirani. (Springer Texts in Statistics). First Edition.

Data Analysis for Business, Economics, and Policy, by Gábor Békés and Gábor Kézdi. (Cambridge University Press). First Edition.

The optional text books are helpful, but I do not expect students to have a copy in hand. However, the Llaudet and Imai textbook is required and you will be expected to have access to a copy.

Course Materials

All lecture notes, homework, important dates, and class announcements will be posted on canvas. Please visit canvas for the latest updates. Students are expected to have access to: the required textbook, a laptop computer, RStudio, and Canvas. **Please bring your laptop to every lecture.**

Exams

There will be two midterm exams (see canvas for dates) during normal lecture hours and one final exam. The date of the final exam is on the GT final exam matrix. The final exam is cumulative. The exam dates are subject to change. If there is any change regarding the exam dates, students will be notified at least 2 weeks in advance.

Homework

About 5-7 take-home problem sets will be assigned throughout the semester. R and Rstudio software is required.

Homework assignments and due dates will be posted on canvas. Late homework will only receive half credit unless permission for a specific extension is obtained in advance. Homework submitted after solutions are posted will receive no credit.

Homework must be submitted through Canvas. Paper and emailed homeworks will not be accepted. Scans or photographs of handwritten homework are only acceptable if they are easily readable. Any homework not legible or incomplete will receive partial or no credit.

Your **lowest** homework grade will automatically be dropped from the Homework part of your final grade. For each assignment, a random subset of problems will be chosen to be graded for correctness and the rest will be graded based on completion. Always show all of your work for every problem. If your response to a problem does not show its steps, it may not receive full credit even if the final answer is correct.

You may consult with other students in this class about the homework. But you must submit **your own original work**. No copying is allowed. You must also **include the names** of any other students that you collaborated with. You may not consult with anyone not enrolled in this class.

Grading

Participation (5%) + Homework (25%) + Midterm Exam I (20%) + Midterm Exam II (20%) + Final Exam (30%)

Your final grade will be assigned as a letter grade according to the following scale: A: 90-100%, B: 75-89%, C: 60-74%, D: 50-59%, F: 0-49%

Class Participation

You are expected to attend the class unless you have a compelling reason not to do so. Topics in this course are cumulative. Missing one topic will make subsequent topics more challenging. You are expected to come to class on time having done the readings and problem sets due that day. Please bring your laptop computers to lectures for in-class activities. A few lectures will be given virtually due to Dr. Faridani's pre-arranged research travel. Students will be informed about any virtual lectures in advance.

In-Class Exercises

Lectures sometimes include in-class coding exercises in R. The purpose of these exercises is to help you quickly pick up basic R techniques and to introduce you to the main datasets of the class. The exercises are meant to be fast and straightforward. Homeworks are sometimes based on in-class exercises. Occasionally you will need to finish in-class exercises in order to start the homework. In-class exercises are not graded. If you find them difficult, please email me and I will help you!

Accommodations for Students with Disabilities

If you are a student with learning needs that require special accommodation, contact the Office of Disability Services at (404) 894-2563 or https://disabilityservices.gatech.edu/ as soon as possible to make an appointment to discuss your special needs and to obtain an accommodations letter. Please also e-mail me as soon as possible in order to set up a time to discuss your learning needs.

Academic Integrity

Georgia Tech aims to cultivate a community based on trust, academic integrity, and honor. Students are expected to act according to the highest ethical standards. For information on Georgia Tech's Academic Honor Code and Student-Faculty Expectations, please visit https://policylibrary.gatech.edu/student-life/academic-honor-code and https://catalog.gatech.edu/rules/21/ .

A student suspected of cheating or plagiarizing on a quiz, exam, or assignment may be reported to the Office of Student Integrity, who will investigate the incident and identify the appropriate penalty for violations.

You are encouraged to discuss the homework problems with other students enrolled in this class and to compare answers. However, each student must complete and submit their own individual work. Copying other students' work, discussing solutions with individuals not enrolled in the course, or using the internet or AI to solve homework problems outright are not permitted.

Attendance Policy

Please read the Georgia Tech Attendance Policy. While formal attendance will not be taken, you are expected to attend class. In particular:

> "All students are responsible for obtaining an understanding of each instructor's policy regarding absences; all students are expected to attend announced quizzes, laboratory periods, and final examinations. Although it is recognized that occasionally it may be necessary for students to be absent from scheduled classes or laboratories for personal reasons, students are responsible for all material covered in their absences, and they are responsible for the academic consequences of their absences. Students should discuss planned absences with their instructors as soon as possible after the beginning of an academic term. Work missed may be made up at the discretion of the instructors."

Student-Faculty Expectations

Please read the Georgia Tech Student-Faculty Expectations Agreement. I will expect you to follow its description of student conduct just as you can expect me to follow it as well.