Econ 106
Data Analysis in Economics
UC Riverside
Fall 2024

Instructor: Veronica Sovero
Email: veronica.sovero@ucr.edu
Office: Sproul 3121

Lecture Meeting Times: Mondays and Wednesdays, 2:00-3:20pm

Lecture Location: Spieth 2200

Office Hours: Wednesdays 10:30-11:30am (Sproul 3121) , Fridays 10:30-11:30am (zoom)

I respond to emails/Slack/Canvas messages during regular work hours (Monday to Friday, 9am to 5pm). If you have not heard from me within one working day, please send a follow up message (chances are I missed it in my inbox).

Course Description

Data literacy has become an increasingly in demand skill in the workforce. This course introduces exploratory data analysis using R, the most common statistical programming language in data science. Although the course will be taught using R, an emphasis will be placed on general rules and best practices for programming that apply to many programming languages, ranging from STATA to PYTHON. The first part of the course teaches students how to collect, clean, and prepare data for analysis. The second part of the course teaches students how to generate meaningful displays of quantitative and qualitative data from a variety of sources (spatial, text, etc.). An emphasis will be placed on data visualization, which is the graphical representation of information and data. The quarter culminates with the completion of a comprehensive final project addressing an economic topic of the student's choosing.

Course Learning Outcomes:

- Learn how to collect and prepare economic data for analysis
- Generate and correctly interpret tables, charts, graphs, and maps
- Develop competency on the use of software tools that can be used for data analysis (R)
- Apply data analytic skills to an economics-related research project

Prerequisites

Statistics (Econ 101)

No prior programming experience is required (beginners are welcome!)

Textbooks
"R for Data Science: Import, Tidy, Transform, Visualize, and Model Data"- Hadley Wickham, Garrett Grolemund [RDS] (available at https://r4ds.hadley.nz/)

I will also occasionally post links to other online resources.

Software

The emphasis throughout the class will be placed on providing hands-on experience with data analysis using R Studio. You will have access to R during your lab section. I highly encourage you to download R and Rstudio on your own computer as well. It is free.

Grading

Assignment	Points
Lab Assignments	40
Poll Everywhere	10
Project Milestones	30
Final Project	20
TOTAL	**100**

Grade Distribution

A+: 98-100%	B+: 88-89.99%	C+: 78-79.99%	D+: 68-69.99%
A: 94-97.99%	B: 84-87.99%	C: 74-77.99%	D: 64-67.99%
A-:90-93.99%	B-: 80-83.99%	C-: 70-73.99%	D-: 60-63.99%
			F: <59.99%

Lab Section
During lab sessions, students will either work on a lab assignment or a research project milestone.

Lab Assignments
Lab assignments will give students the opportunity to practice how to analyze data and communicate findings. The writeup and r script are two necessary components to every submission. Your script will be graded on whether your code uses the functions covered during lecture, is well commented, and is well organized.

Poll Everywhere
I will use Poll Everywhere during every lecture to make class more interactive. Coding is learned through repeated practice, so it is also one of your best tools for learning the material. You are welcome to discuss the poll questions with your classmates- it is a great way to solidify your understanding of the material.

Research Project
Students will have the opportunity to complete an applied research project to address an economic topic of their own choosing. I will provide a list of suitable datasets. You are also free to find your own data (with my approval). There will be graded project milestones that will be completed over the course of the quarter. These assignments are essential for being able to complete the final project at the end of the quarter. I will post more detailed guidelines on Canvas.

Participation/Attendance
Students are strongly encouraged to attend lectures with a computer that has Rstudio installed on it. I will post R scripts for every lecture, which will allow for you follow along on your own computer. Sometimes Poll Everywhere questions will be based on coding exercises.

Late Work Policy
To make it easier to keep track of deadlines, the due date for all assignments will be Sundays at 11:59pm. Late submissions are allowed with a 5% late penalty for 24 hours after the due date. Any work submitted more than 24 hours after the due date will receive a zero.

Of course, unexpected events will inevitably happen over the course of the quarter (you get sick, you had a family issue, etc.). To allow for some flexibility, I will drop the one lowest lab score. For Poll Everywhere, you will need to answer at least half of the questions correctly to receive full credit. Please note that I will not drop any project milestone scores- this is because every milestone is essential to complete. Please reach out to me ASAP if you have a serious documented reason for missing more than one week's worth of assignments. This is the only circumstance where I consider granting extensions or removing late penalties.

Academic Integrity
I expect that you will complete all exams and assignments in the manner in which they are intended and inform me of suspected acts of academic misconduct by your peers. By doing so, you will affirm the integrity and intellectual work of the University and the degree it represents. Should you choose to commit academic misconduct in this course, you may receive a failing grade. In all cases, you will be reported the Student Conduct and Academic Integrity Programs (SCAIP) and you will be held accountable according to the policies set forth by the University. For more information about the University's Academic Integrity policies, visit

http://conduct.ucr.edu/policies/academicintegrity.html. Here are some examples of violations of academic integrity in this class:

- Using ChatGPT or other generative AI to generate code without acknowledging its usage on assignments
- Using ChatGPT or other generative AI for written portions of assignments (not code)
- Turning in assignments that represents the work of another person or student (for example copying the project of a previous student in 106)

Please ask for clarification if you have any concerns regarding the definition of academic integrity regarding this course.

Tentative Schedule

Week	Dates	Topic	Lab	Project Milestone
1	9/30, 10/2	Course Introduction, R Basics	R setup	
2	10/7, 10/9	Data Wrangling (dplyr)	Project ideas	
3	10/14, 10/16	Visualization (ggplot2)	1	
4	10/21, 10/23	Data Cleaning (Numeric, Factor Variables)	2	
5	10/28, 10/30	Exploratory Analysis		MS 1
6	11/4, 11/6	Tidy Data	3	
7	11/11, 11/13	**No Class Monday,** Text as Data		MS 2
8	11/18, 11/20	Spatial Data	4	
9	11/25, 11/27	Visualization for communication		MS 3
10	12/2, 12/4	work on projects	5	
Finals week		**Final projects due Sunday, December 8th at 11:59pm**		