

Problem Set 1-Econ 5253

Jhinuk Banerjee

January 29, 2026

1 Introduction: Part (5)

My interest in this data science class is mainly in how I can apply these tools and techniques to economics research, particularly in my field of interest. Being a third-year PhD student, I am working on my dissertation. One of my papers is about studying economic insecurity and gender in politics. Over the years, I have realized how crucial advanced data science skills are for handling large-scale datasets.

I was motivated to take this class because my research requires data manipulation and analysis techniques, which I will be able to learn here. This class will provide me with hands-on experience in data science, which I will be able to use in my future research work. Working with big datasets like IPUMS USA, State Legislative Election Returns, and various other political and economic databases has made me realize the importance of these computational tools. I want to expand my abilities beyond Stata to include more advanced programming capabilities. I had taken a class on Data Science last semester, and I am hoping that both these classes are going to be very helpful for my understanding of R and data analysis.

For my class project, I am not very sure what I am going to work on. I am still thinking of an interesting idea where I can apply the data science techniques that I will learn in this class.

My goals for this class include becoming proficient in R and more, improving my coding efficiency, learning best practices for coding and analysis, and learn how to handle big and complex datasets. After graduation, I plan to pursue a career in public or private sector or think tanks, where I can use my research for policy-oriented ground-level application.

2 Equation: Part (6)

$$a^2 + b^2 = c^2 \quad (1)$$