

# Problem Set Two

Devin Williams

February 3, 2025

## Main Tools of a Data Scientist

- Measurement
  - ◊ How insights and policies are constructed.
    - ◊ You can't construct a policy without measuring what the policy is meant to accomplish in the end.
- Statistical programming languages
  - ◊ R
  - ◊ Python
  - ◊ Julia
  - ◊ Stata
  - ◊ SAS
  - ◊ SPSS
  - ◊ Matlab
  - ◊ JavaScript
- Visualization tools
  - ◊ Very important tools for data scientists in different coding languages.
    - ◊ ggplot2 (R)
    - ◊ matplotlib (Python)
    - ◊ Plots.jl (Julia)
    - ◊ Tableau
    - ◊ Modeling
- Big Data management software
  - ◊ There is often data sets that are too being for things like R or Python to handle. There are multiple other programs that can help analyze data if the data set is too large.
    - ◊ RDDs
    - ◊ SQL
- Data collection tools
  - ◊ There is many different ways to collect data as a data scientist. Using these tools makes data analysis easier in many different ways.
    - ◊ Web scraping
      - ◊ Can be done in almost every coding language available for use.