Problem Set 2

Samuel Hunt

February 2025

1 Main Tools of a Data Scientist

- Measurement
- Statistical Programming Languages
  - R
  - Python
  - Julia
  - Stata
  - SAS
  - SPSS
  - Matlab
  - JavaScript
- Visualization Tools
  - ggplot2 (R)
  - matplotlib (Python)
  - Plots.jl (Julia)
  - Tableau
- Big Data Management Software
  - Resilient Distributed Datasets
  - SQL
- Data Collection Tools
  - Web Scraping