Crim 250:
Statistics for the social sciences

Introduction

Review

- Meet the toolkit: from the video you followed.

- Review Chapter 1 of DVB

- What is R?

- Exploratory data analysis using an example: Prisoner count in PA

Outline

- Exploratory data analysis

  - Review code on website.

Breakout room activity

Exploratory data analysis

- What variables are in the dataset?

- What type of variables are they?

  - Quantitative (Continuous or Discrete)

  - Categorical (Nominal, Ordinal, or Binary)

- What does the distribution of each variable look like?

- How are the variables related to each other?

- Are there any "red flags" in the data?

- What are some interesting questions the dataset can answer?

- What questions can the dataset not answer?

Statistical thinking

- Teaches you to think like a statistician rather than someone who uses statistics.

- Not an automated, mechanical procedure.

- Not an art, but it does require creativity.

One-variable EDA: Questions we want to answer

- Variable type

- Important quantities:

  - shape

  - center

  - spread

  - unusual observations

- The shape and unusual observations are best identified through a visualization, while the center and spread can be found with quantitative summaries.

< Go to website to 3. EDA.R >