Economic Data Analysis

Part 1.0 | Data Types

Taylor Weidman

Types of Data

Data comes in all shapes, sizes, and types.

- Value Type: numerical data, categorical data
- Unit of Observation: cross-section data, time series data, panel data
- Number of Variables: univariate data, bivariate data, multivariate data

Types of Data: Numerical Data Examples

Numerical data is data that’s best recorded in numerical form, continuous or discrete.

- Annual income (continuous)
- Hours worked (discrete, maybe)
- Annual GDP (continuous)
- Number of moves (discrete)

Data Types

Categorical Variables

Data Types

Categorical Variables

- Nominal: categories cannot be ordered / ranked

Data Types

Categorical Variables

- Nominal: categories cannot be ordered / ranked
  - Blood types (A, B, AB, O)
  - Colors (Red, Blue, Green)
  - Eye color

Data Types

Categorical Variables

- Nominal: categories cannot be ordered / ranked
  - Blood types (A, B, AB, O)
  - Colors (Red, Blue, Green)
  - Eye color
- Ordinal: categories have order / rank but not a meaningful scale

Data Types

Categorical Variables

- Nominal: categories cannot be ordered / ranked
  - Blood types (A, B, AB, O)
  - Colors (Red, Blue, Green)
  - Eye color
- Ordinal: categories have order / rank but not a meaningful scale
  - Education levels (High School, Bachelor's, Master's, PhD)
  - Survey responses (Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)
  - Size categories (Small, Medium, Large)
  - Pain scales (No Pain, Mild, Moderate, Severe)

Data Types

Categorical Variables

- Nominal: categories cannot be ordered / ranked
  - Blood types (A, B, AB, O)
  - Colors (Red, Blue, Green)
  - Eye color
- Ordinal: categories have order / rank but not a meaningful scale
  - Education levels (High School, Bachelor's, Master's, PhD)
  - Survey responses (Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)
  - Size categories (Small, Medium, Large)
  - Pain scales (No Pain, Mild, Moderate, Severe)
- Binary: only two categories

Data Types

Categorical Variables

- Nominal: categories cannot be ordered / ranked
  - Blood types (A, B, AB, O)
  - Colors (Red, Blue, Green)
  - Eye color
- Ordinal: categories have order / rank but not a meaningful scale
  - Education levels (High School, Bachelor's, Master's, PhD)
  - Survey responses (Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)
  - Size categories (Small, Medium, Large)
  - Pain scales (No Pain, Mild, Moderate, Severe)
- Binary: only two categories
  - Economics major (YES/NO)
  - Human (YES/NO)

Data Types

Numerical/Quantitative Variables

- Discrete: countable numbers with meaningful intervals
  - Number of Children in a Household (1, 2, …)
  - Number of Siblings
- Continuous: quantities measurable on the reals
  - Household income in USD

Economic Data Analysis

Part 1.1 | Categorical Variables

Taylor Weidman

Example 1.1 | Coffee Shops by State

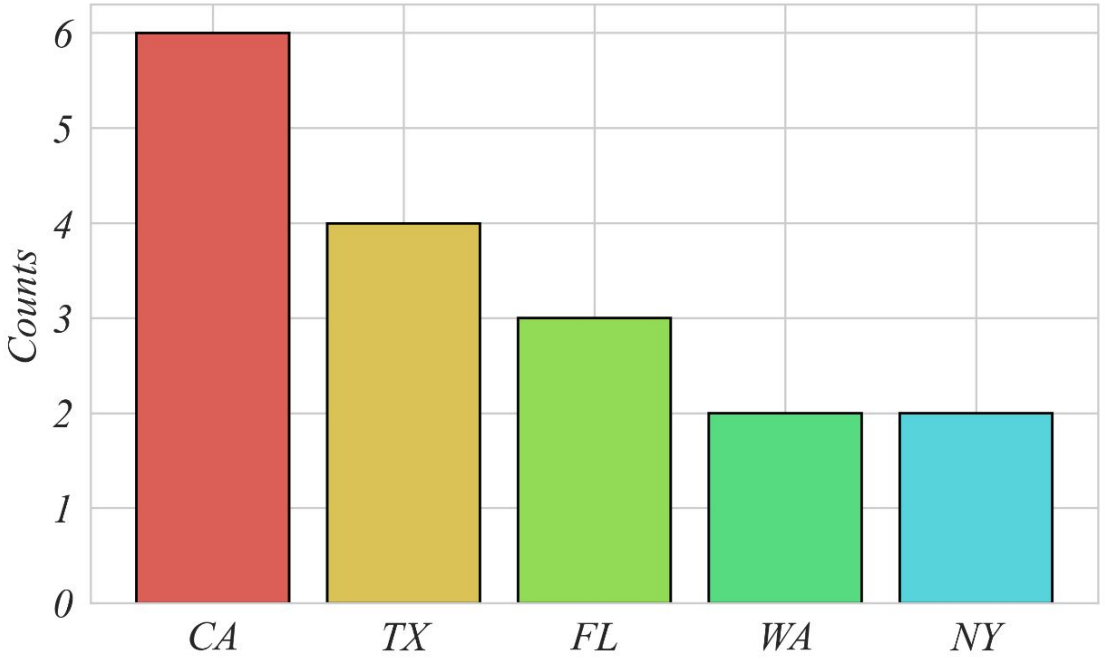
Summarize the locations of coffee shops in Part_1_1_Coffee_Shops.csv

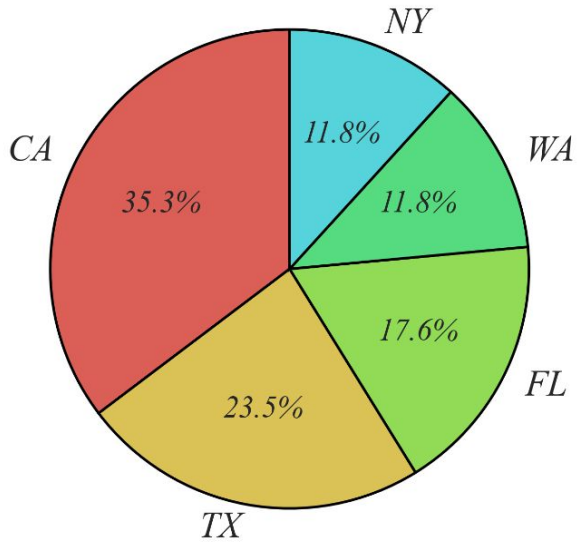
Categorical Variables: Coffee Shop Locations by State

Q. Which state has the most locations?

Categorical Variables: Coffee Shop Locations by State

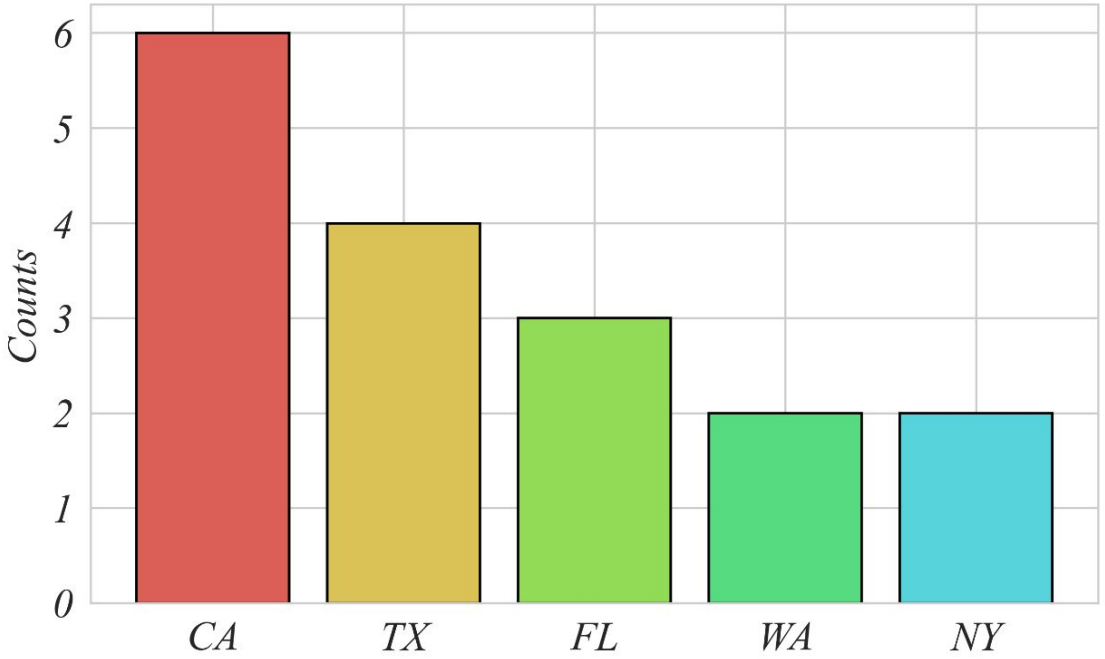
Q. Which state has the most locations?

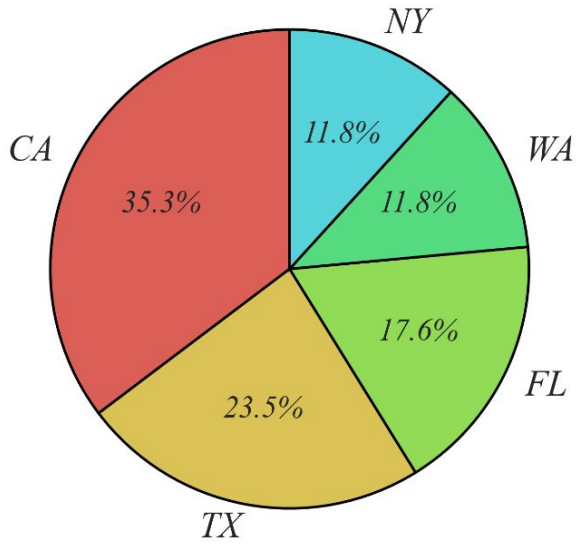




Categorical Variables: Coffee Shop Locations by State

Q. Which state has the most locations?

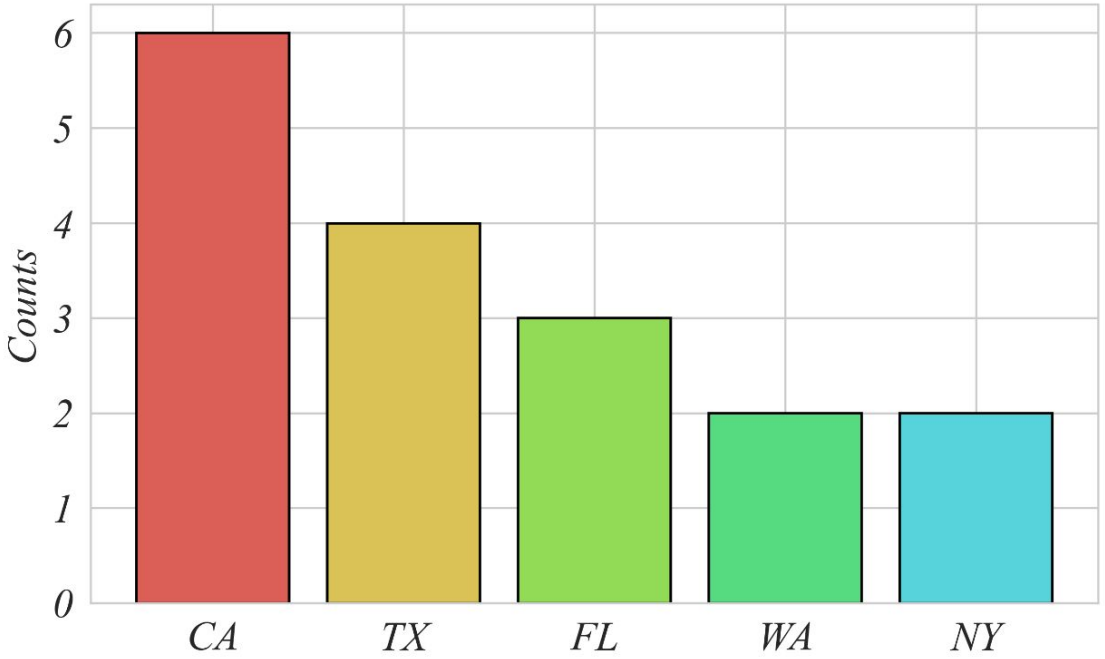


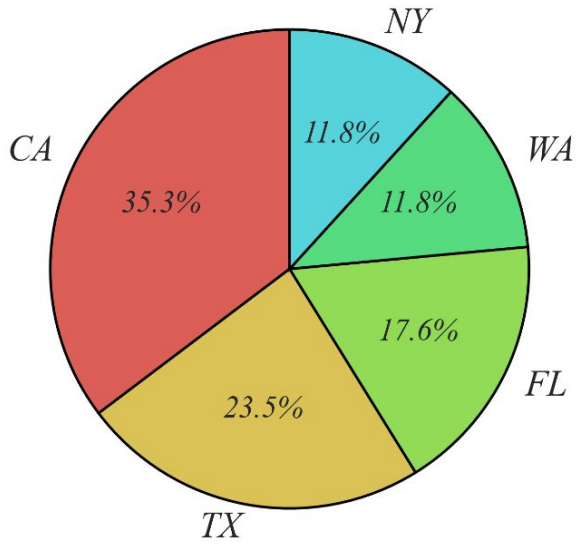


Both graphs work, but bar graphs are easier to read.

Categorical Variables: Coffee Shop Locations by State

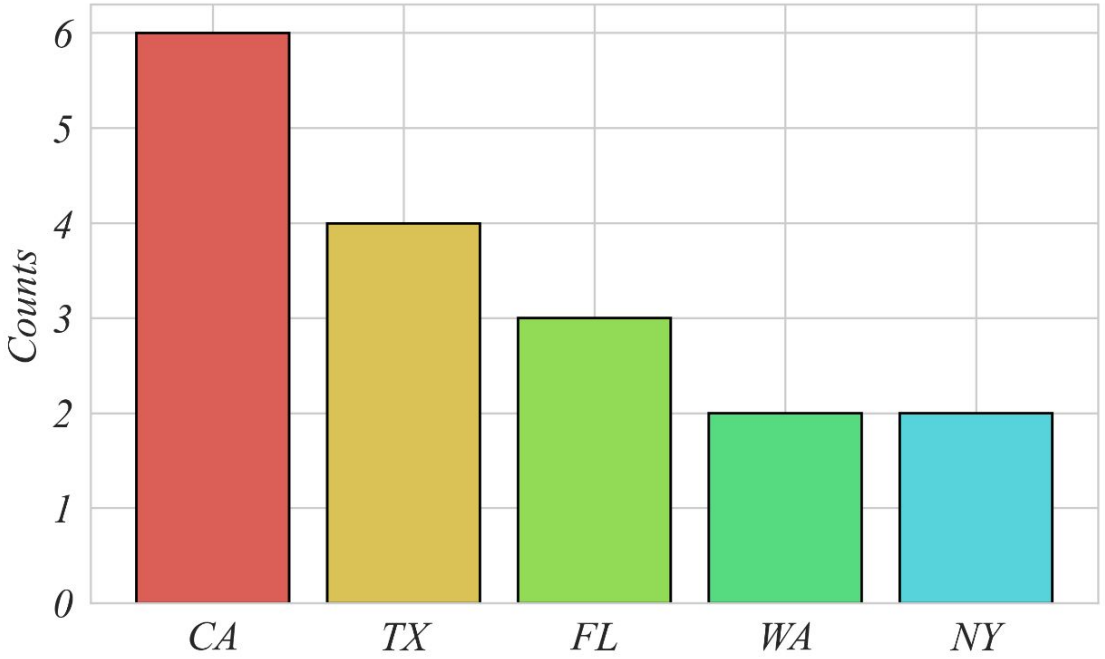
Q. Does FL or WA have more locations?

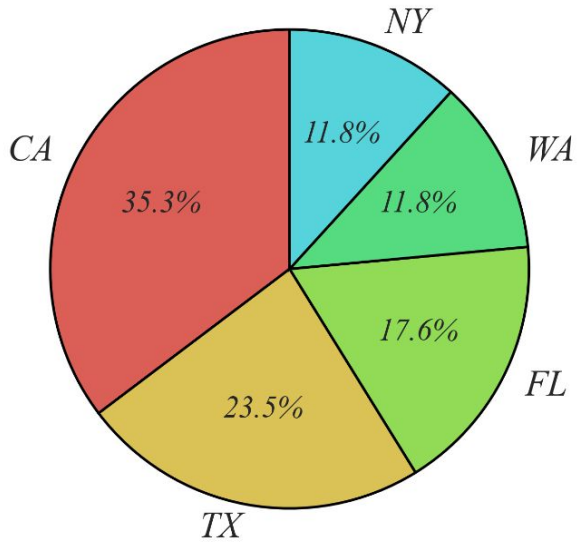




Categorical Variables: Coffee Shop Locations by State

Q. Does FL or WA have more locations?

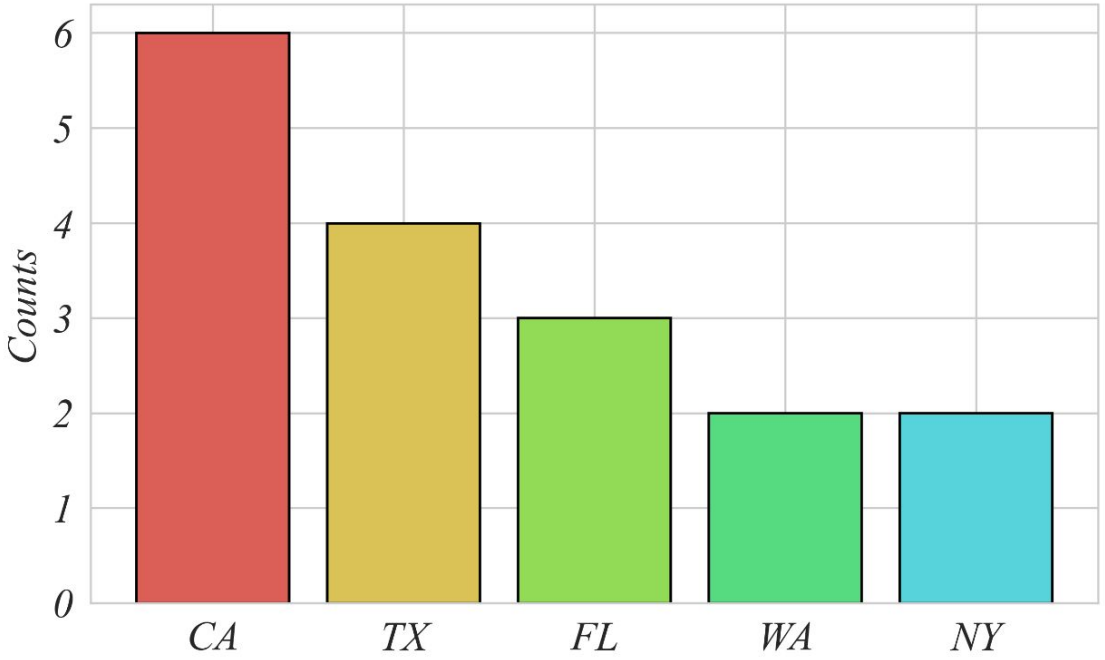


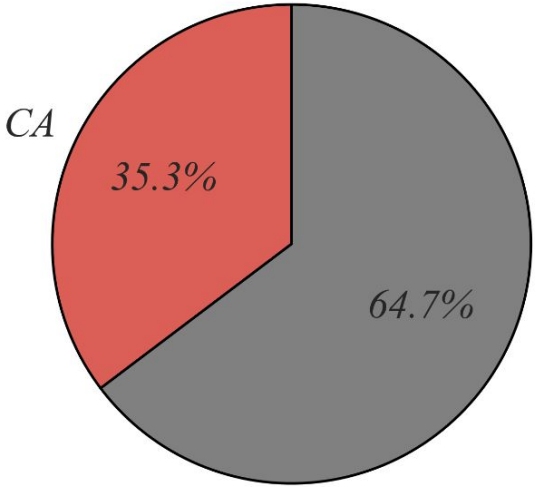


Bar graphs make it especially easy to compare between categories.

Categorical Variables: Coffee Shop Locations by State

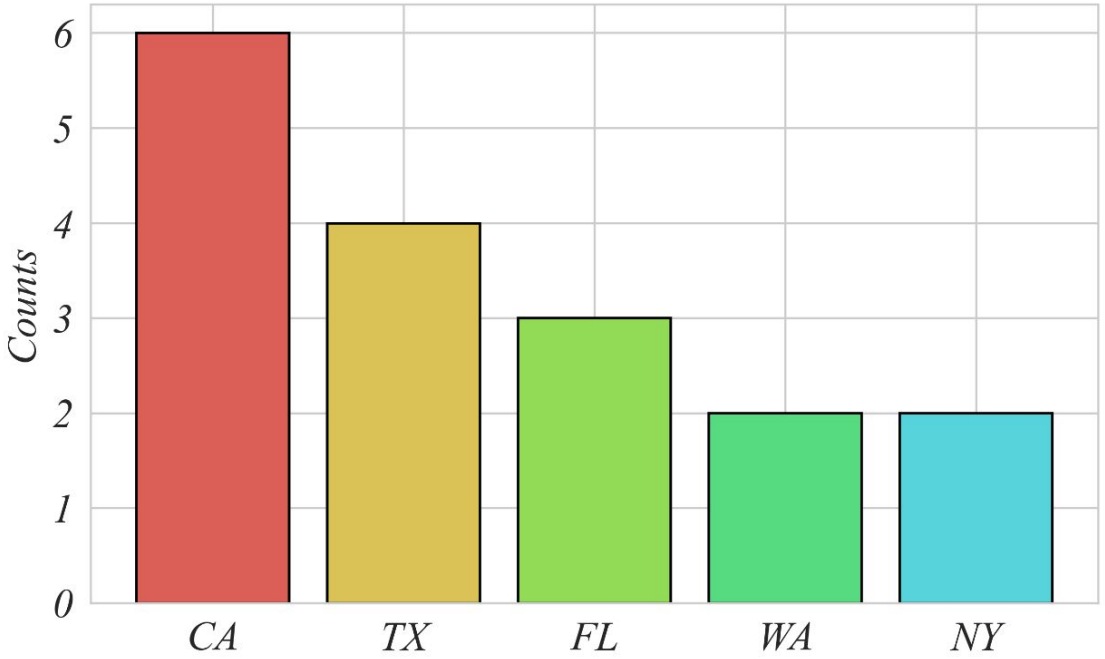
Q. What share of coffee shops are located in CA?

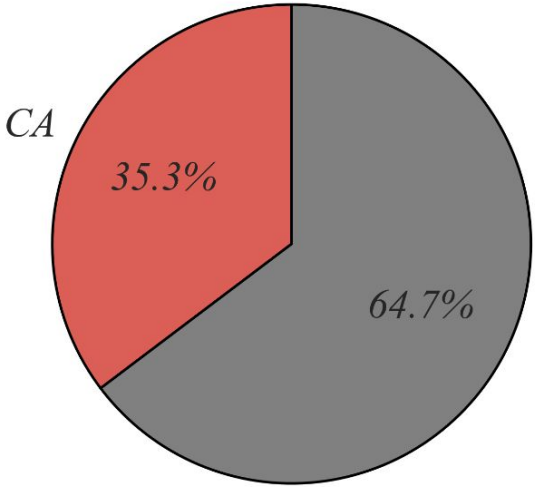




Categorical Variables: Coffee Shop Locations by State

Q. What share of coffee shops are located in CA?





Pie graphs can be used (with caution) for binary categorical variables.

Categorical Variables

- Use ***bar graphs*** to compare values with each other.
- Use ***pie graphs*** to compare one value against the total.