ECON 0150 | Economic Data Analysis

The economist's data analysis pipeline.

Part 1.0 | Variable Types

Dimensions of Data

Data comes in all shapes, sizes, and types.

Variable Type

- *Categorical data*
- *Numerical data*

Data Structure

- *Cross-sectional data*
- *Time series data*
- *Panel data*

Number of Variables

- *Univariate (n=1)*
- *Bivariate (n=2)*
- *Multivariate (n>2)*

Variable Types: Categorical

... data that's best recorded in categories

Binary: only two categories

- *Economics major (YES/NO)*
- *Human (YES/NO)*

Nominal: categories cannot be ordered / ranked

- *Blood types (A, B, AB, 0)*
- *Colors (Red, Blue, Green)*

Ordinal: categories have order / rank but not a meaningful scale

- *Education levels (High School, Bachelor's, Master's, PhD)*
- *Survey responses (Strongly Disagree, Neutral, Agree, Strongly Agree)*
- *Size categories (Small, Medium, Large)*

Variable Types: Numerical

... data that's best recorded in numerical form

Discrete: countable numbers with meaningful intervals

- *Number of children in a household (1, 2, ...)*
- *Number of siblings (1, 2, ...)*

Continuous: quantities measurable on the reals

- *Household income*
- *US GDP*

Categorical Variables: Examples

... hands on examples for all three

Binary: two categories (yes/no; true/false)

- *Employment Status (Employed, Unemployed)*

Nominal: no inherent order

- *Employment Sector (Agriculture, Services, Unemployed).*

Ordinal: meaningful order without meaningful intervals.

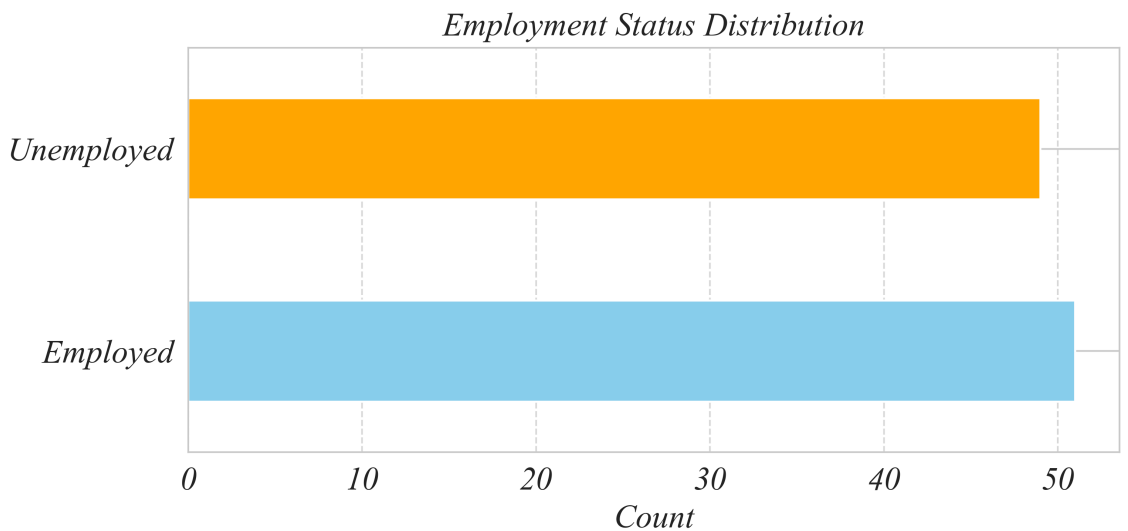
- *Economic Optimism (Very Pessimistic to Very Optimistic).*

Binary Categorical Variables

Exercise: summarize binary_categorical_dataset.csv

Summary

- *Use a **Bar Plot** or a **Pie Plot***
- *Horizontal bar works well*

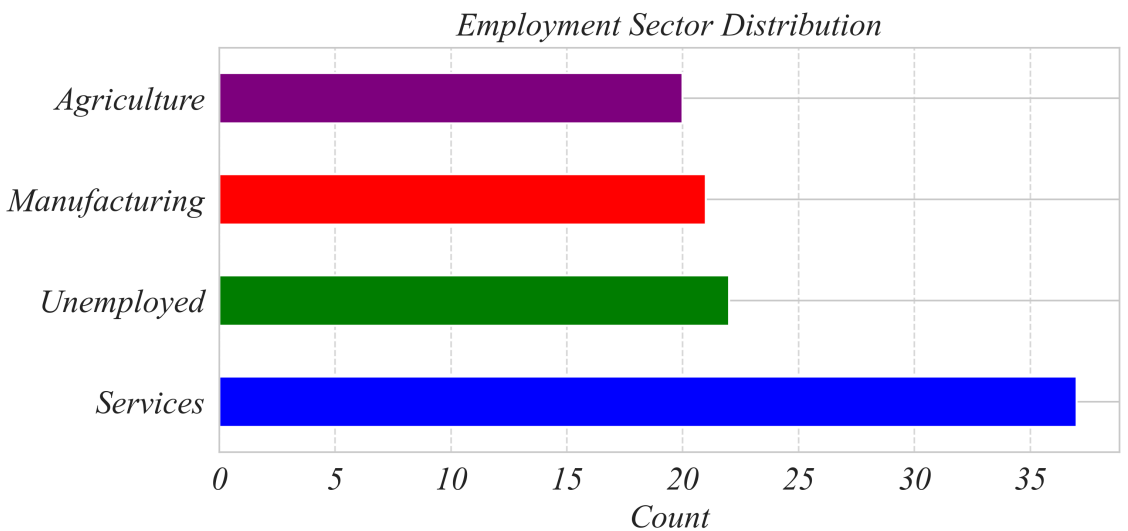


Nominal Categorical Variables

Exercise: Summarize nominal_categorical_dataset.csv

Summary

- *Use a **Bar Plot***
- *Similar to Binary*
- *More categories*

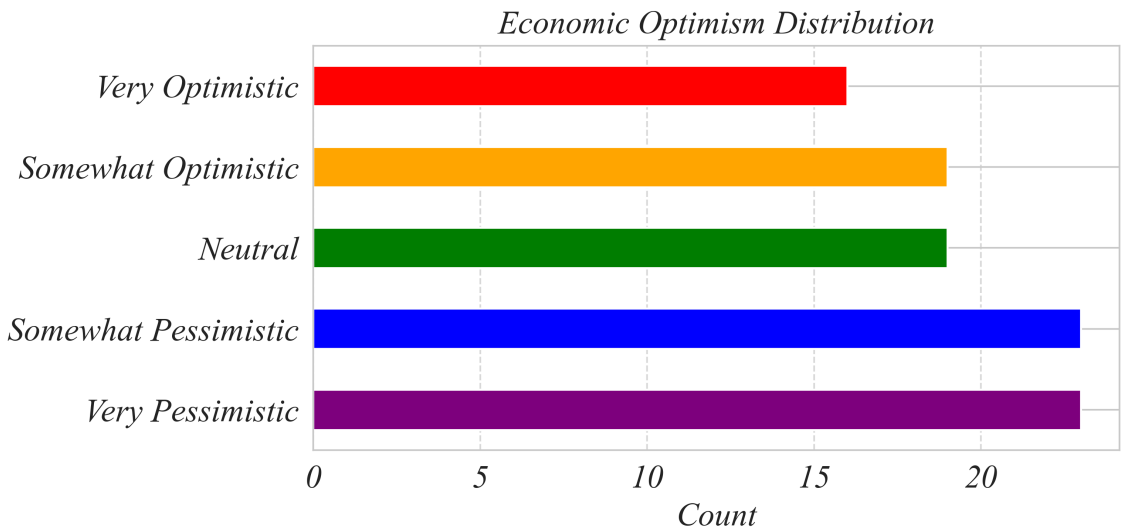


Ordinal Categorical Variables

Exercise: Summarize ordinal_categorical_dataset.csv

Summary

- *Use a **Bar Plot***
- *Similar to nominal*
- *The axis is ordered*



Numerical Variables: Examples

... hands on examples for both types

Discrete: countable numbers with meaningful intervals

- *Number of Children in a Household.*

Continuous: quantities measurable on the reals.

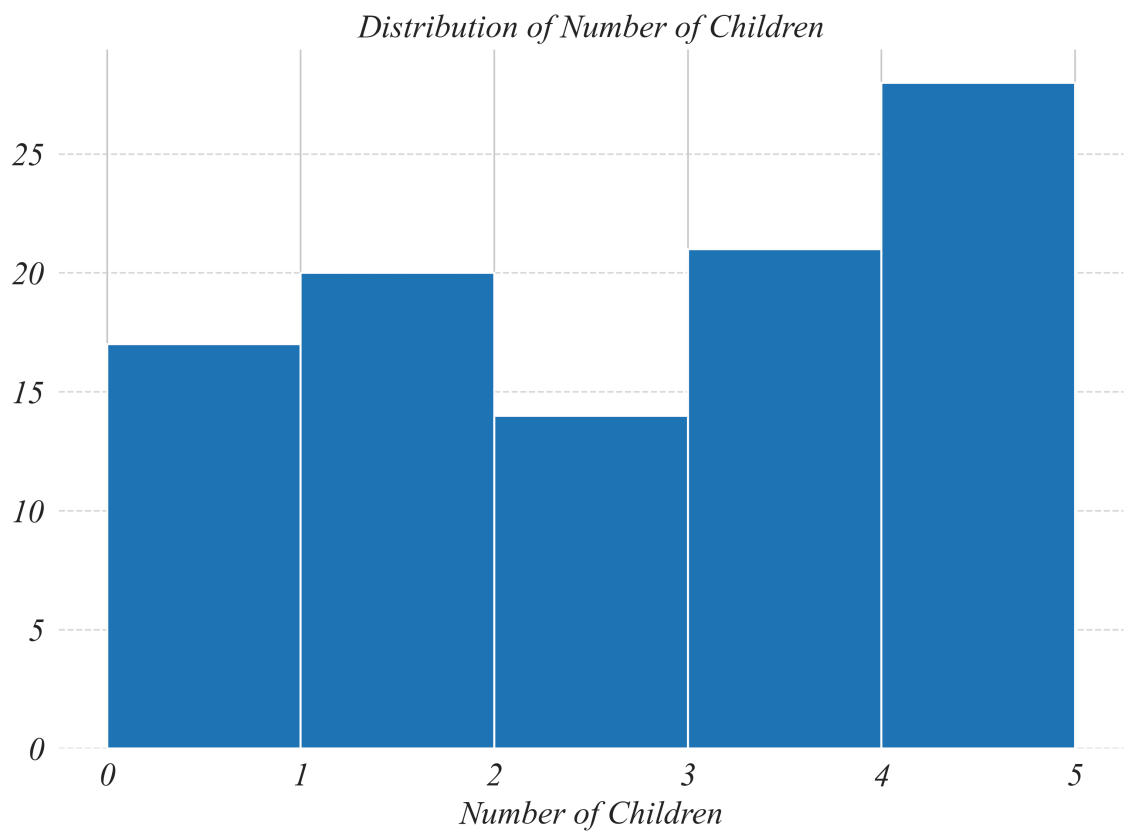
- *Household Income in USD.*

Discrete Numerical Variables

Exercise: Summarize discrete_numerical_dataset.csv

Summary

- *Use a **Histogram** or a **Summary Table***
- *Similar to Ordinal Cat.*

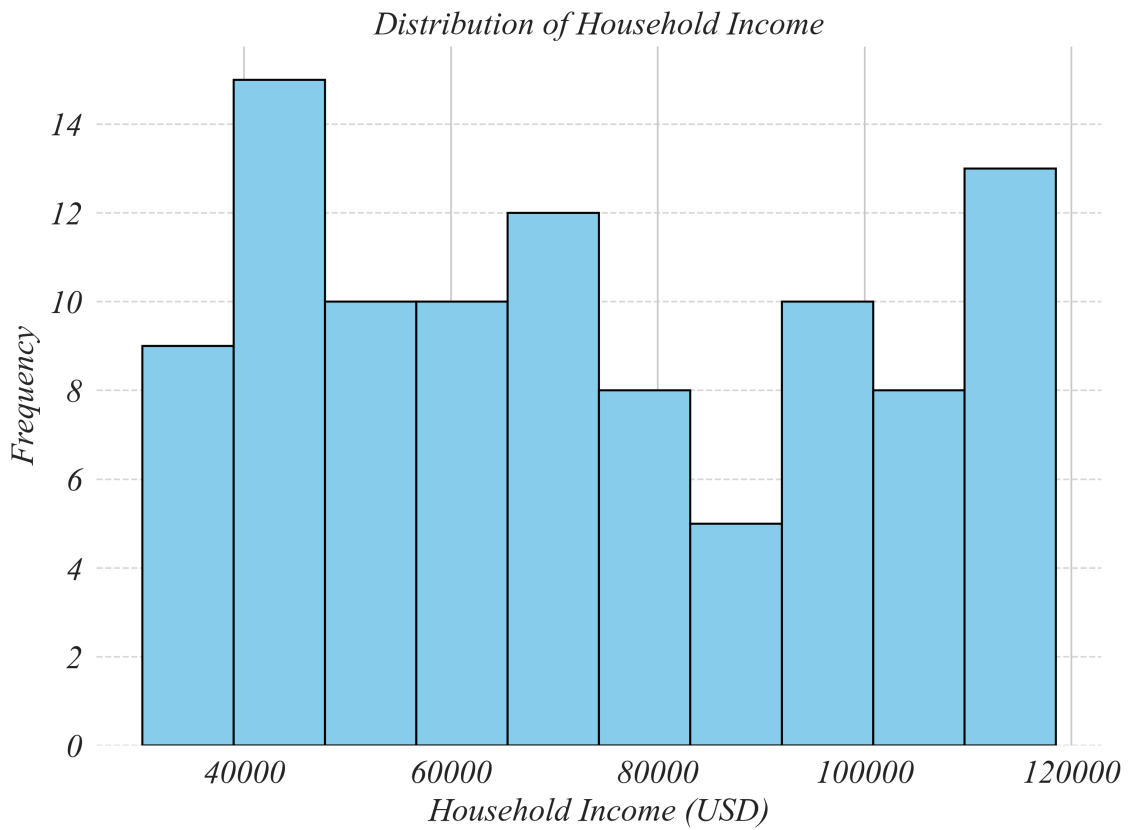


Continuous Numerical Variables

Exercise: Summarize continuous_numerical_dataset.csv

Summary

- *Similar to Discrete*
- *Use a Summary Table,* ***Histogram****, Boxplot, or Jitter*



Continuous Numerical Variables

Exercise: Summarize continuous_numerical_dataset.csv

Summary

- *Similar to Discrete*
- *Use a Summary Table, Histogram, **Boxplot**, or **Jitter***