

ECON 0150 | Economic Data Analysis

The economist's data analysis pipeline.

Part 1.3 | Data Structures

Part 1.3 | Data Structures

...three main relationships between data points.

The way we understand data depends on the relationship between the data points.

	Cross-Sectional	Time-Series	Panel Data
Focus	Multiple units, one time point	One unit, many times	Multiple units, many time points
Shape	Wide format	Long format	Long format
Ex.	Household income, 2025	US GDP, 10 years	Household income, 10 years

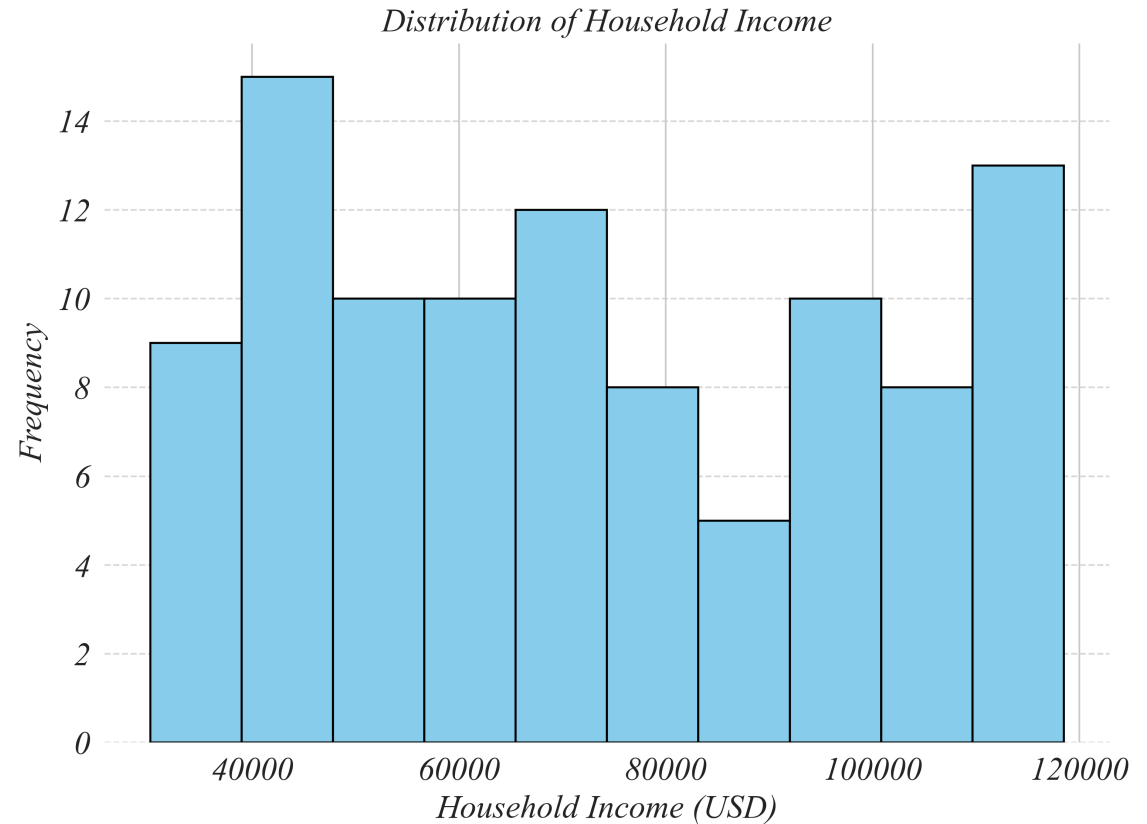
> *we'll introduce geographic data structures later.*

Cross-Sectional (Numerical) Data

Exercise: Summarize cross_sectional_numerical_dataset.csv

Summary

- *Focus: distribution*
 - *Histogram*
 - *Boxplots*
 - *Jitters*

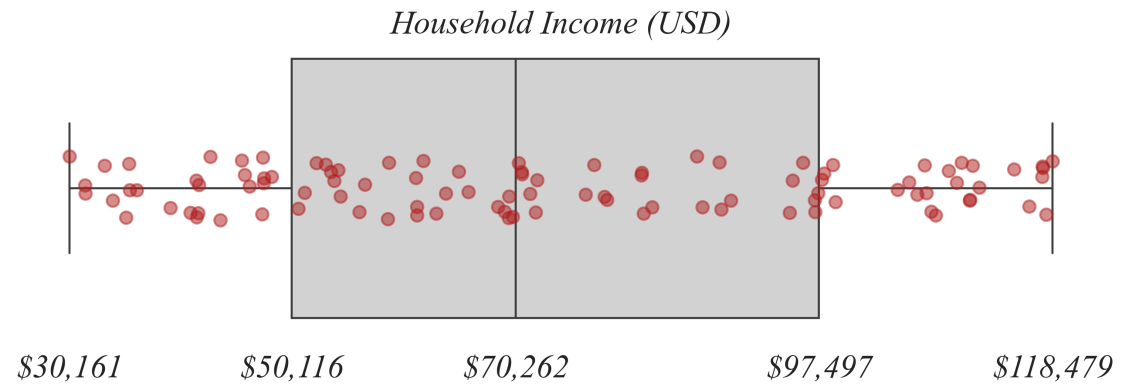


Cross-Sectional (Numerical) Data

Exercise: Summarize cross_sectional_numerical_dataset.csv

Summary

- *Focus: distribution*
 - *Histogram*
 - *Boxplots*
 - *Jitters*

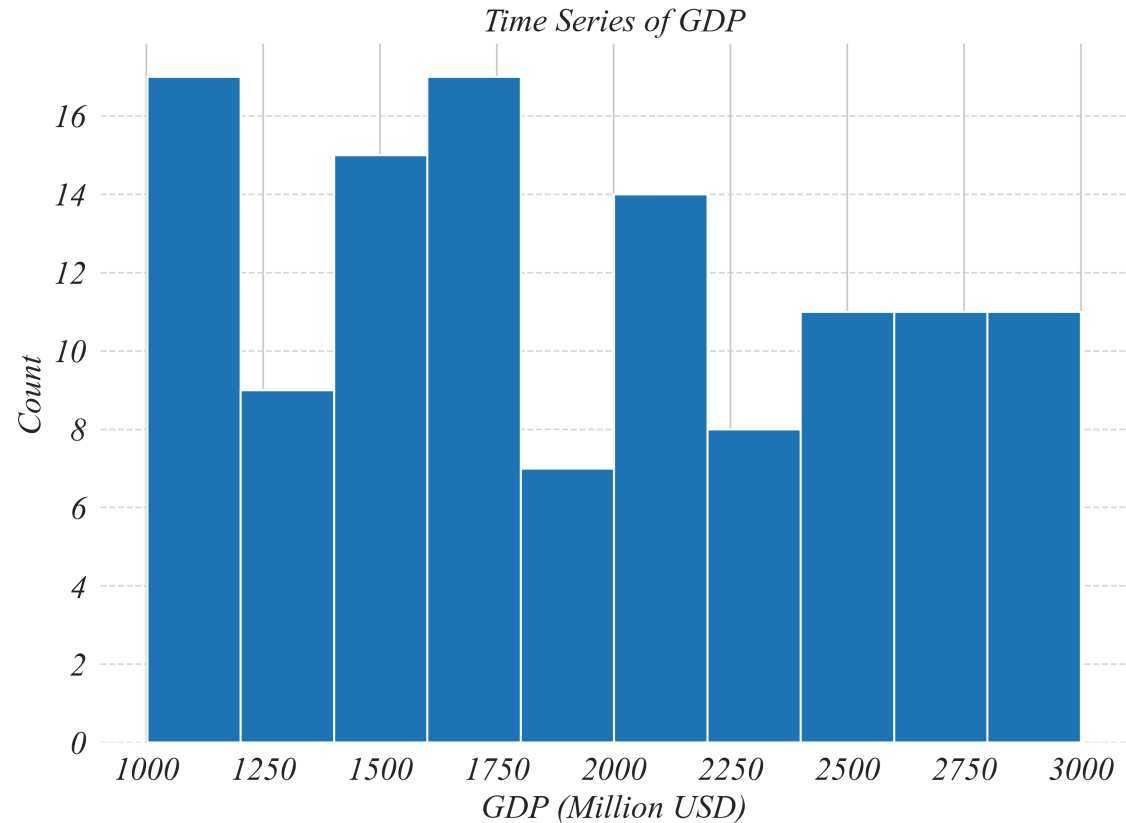


Time-Series (Numerical) Data

Exercise: Summarize time_series_numerical_dataset.csv

Summary

- *Show: distribution*
 - *Histogram*
 - *Boxplots*
 - *Jittered Scatters*
- *Focus: time-path*
 - *Linegraph*

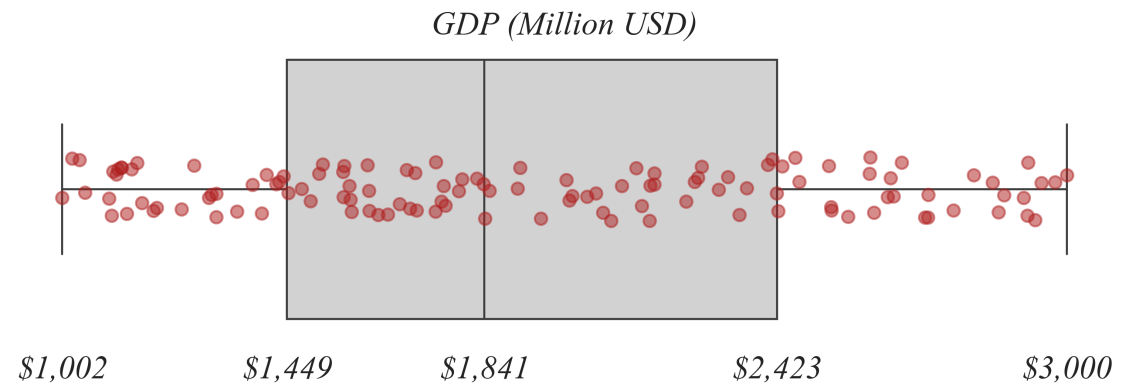


Time-Series (Numerical) Data

Exercise: Summarize time_series_numerical_dataset.csv

Summary

- *Show: distribution*
 - *Histogram*
 - *Boxplots*
 - *Jittered Scatters*
- *Focus: time-path*
 - *Linegraph*

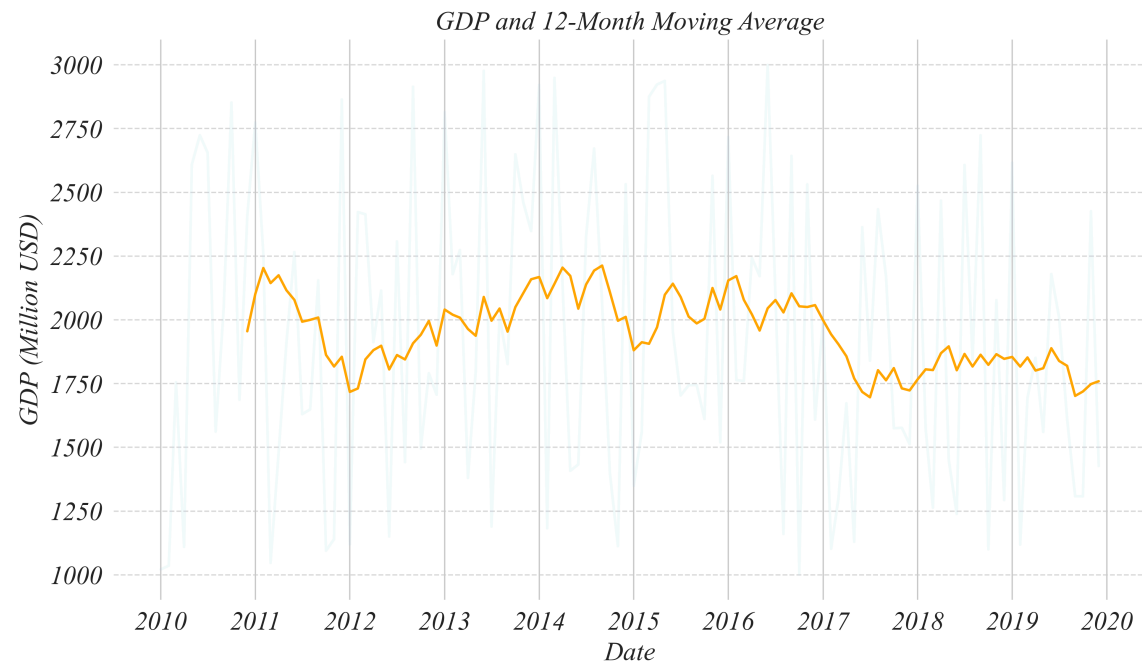


Time-Series (Numerical) Data

Exercise: Summarize time_series_numerical_dataset.csv

Summary

- *Show: distribution*
 - *Histogram*
 - *Boxplots*
 - *Jittered Scatters*
- *Focus: time-path*
 - *Linegraph*

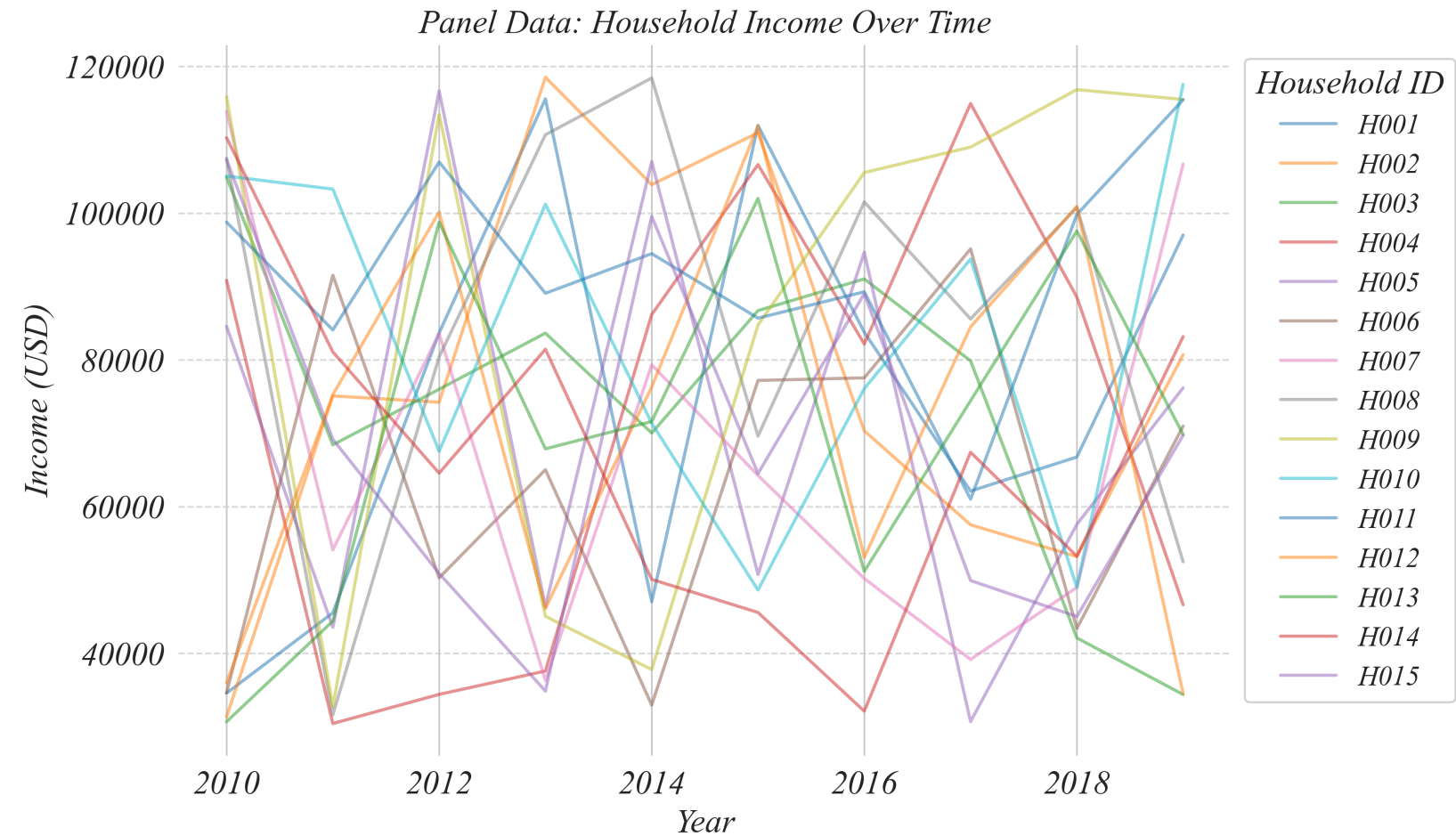


Panel (Numerical) Data

Exercise: Summarize panel_numerical_dataset.csv

Panel (Numerical) Data

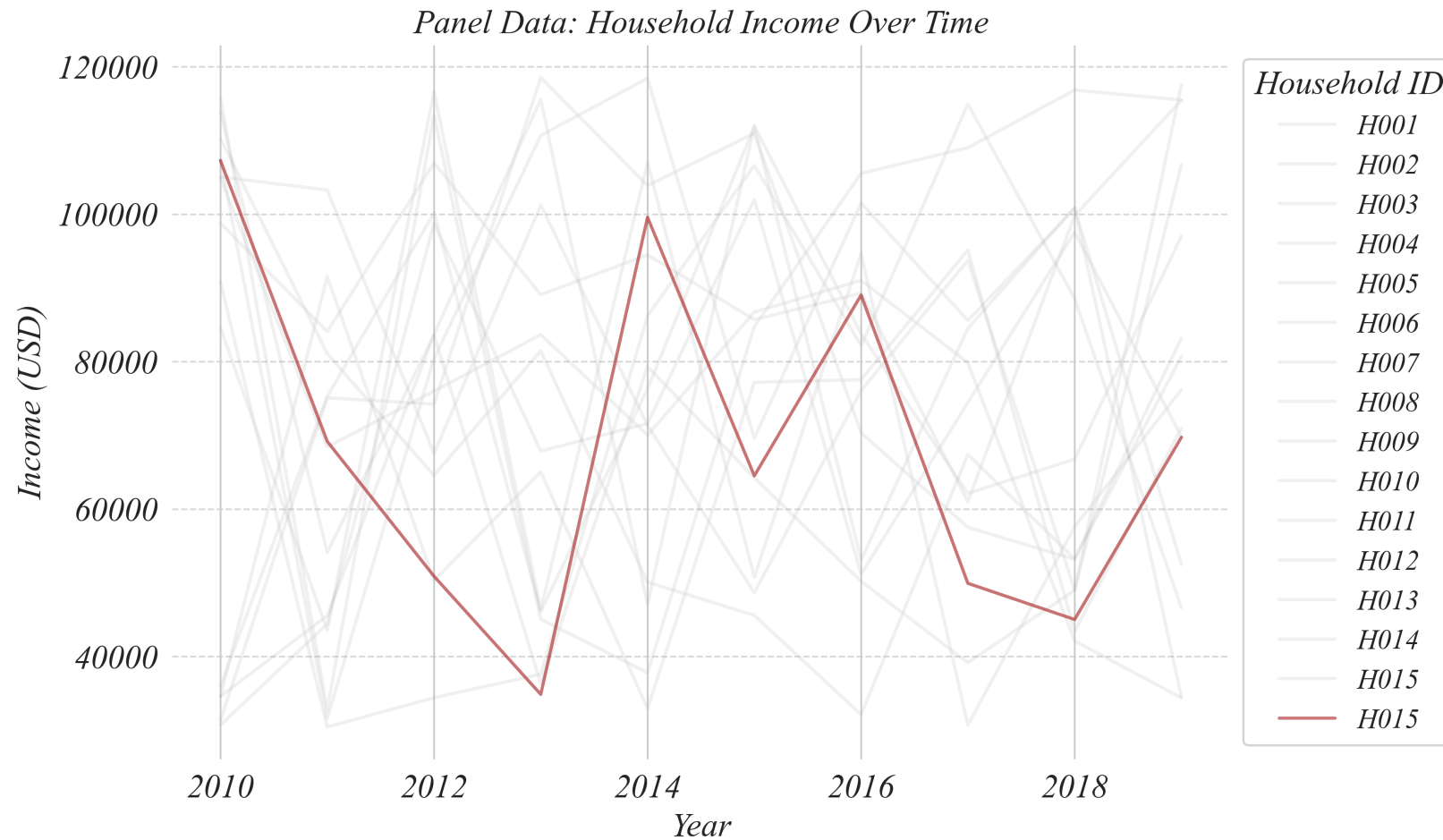
Exercise: Summarize panel_numerical_dataset.csv



> not so great

Panel (Numerical) Data

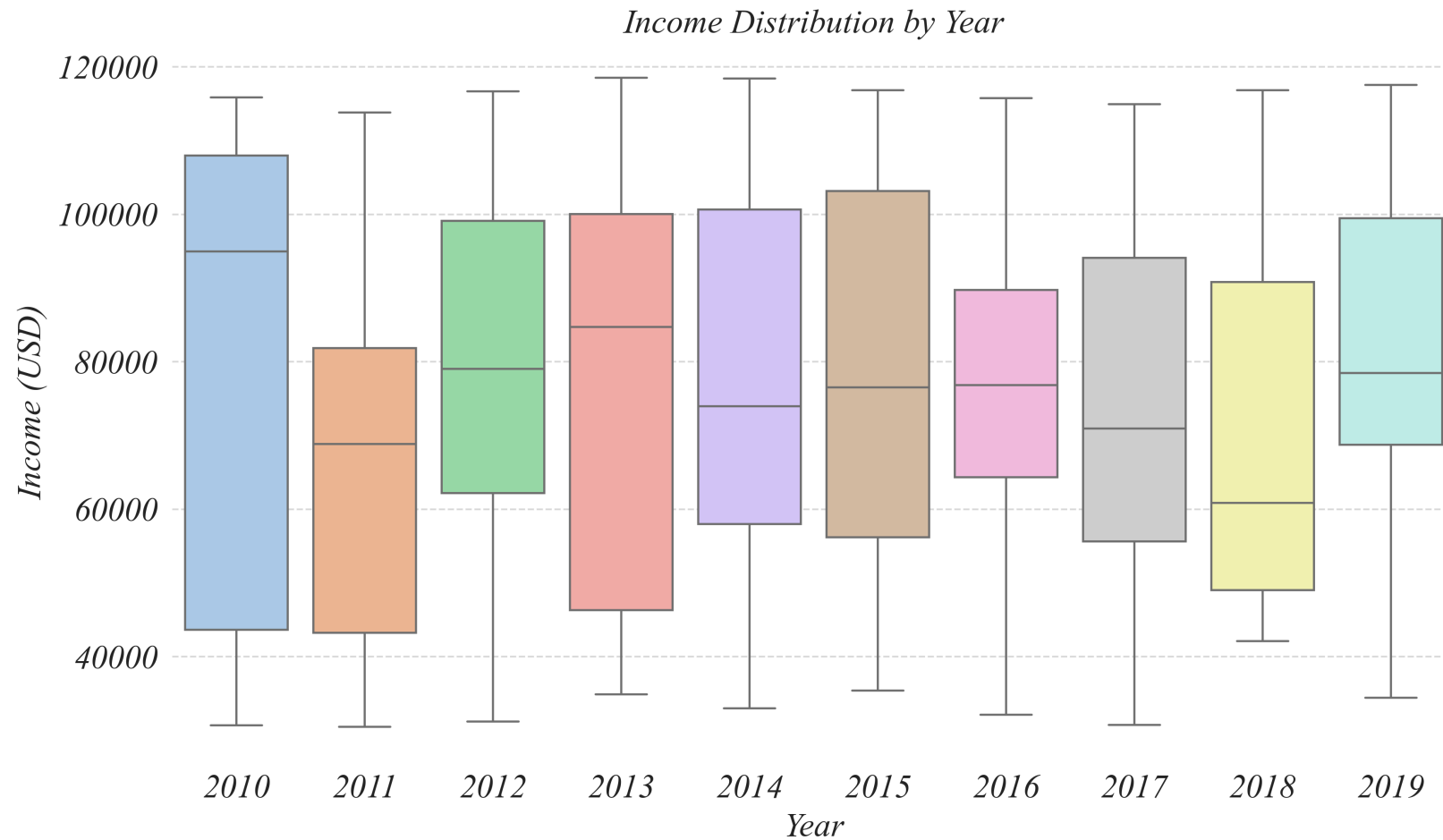
Exercise: Summarize panel_numerical_dataset.csv



> a little better

Panel (Numerical) Data

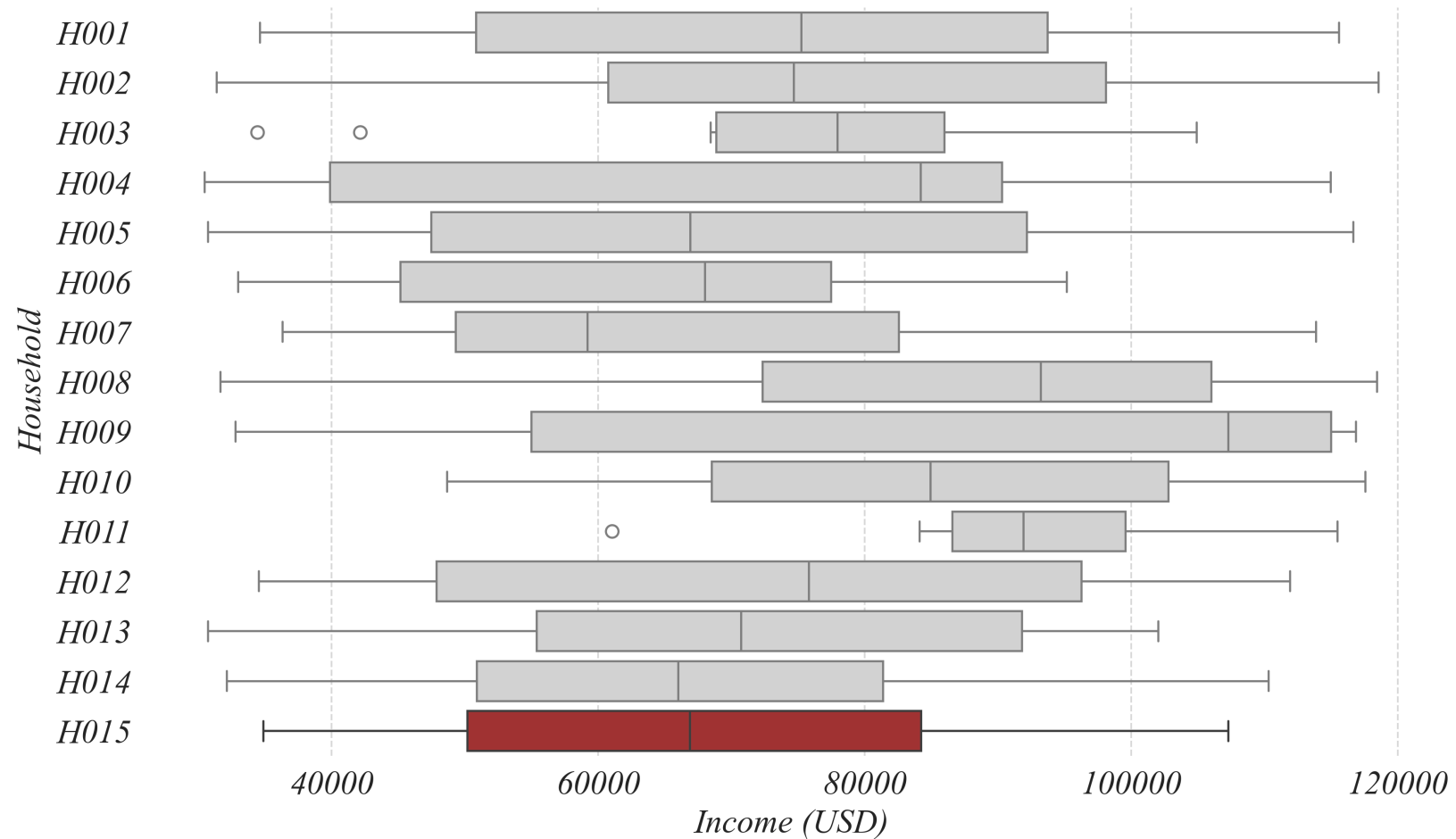
Exercise: Summarize panel_numerical_dataset.csv



> *shows cross sectional variation through time*

Panel (Numerical) Data

Exercise: Summarize panel_numerical_dataset.csv



> *shows variation within the household*