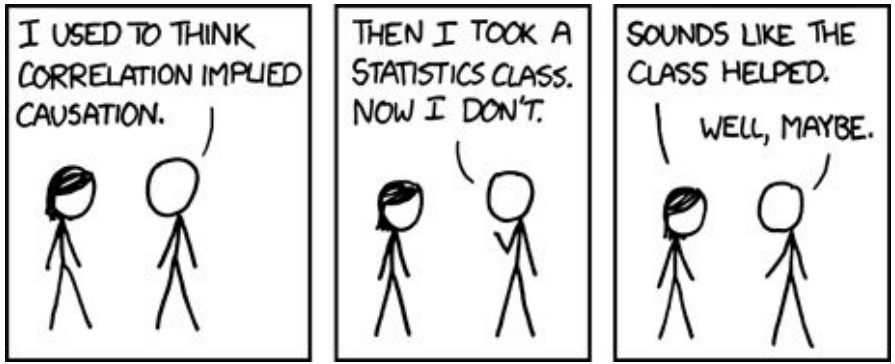
Economic Data Analysis

Taylor Weidman



Interpreting and Communicating Economic Data

- The primary goal of the class is to equip you with the ability to interpret and communicate with economic data.
- This starts with:
  - Developing a framework for thinking about problems involving uncertainty
  - Developing a skillset with spreadsheets and code
  - Communicating with figures, tables, and text
- The focus will be on understanding statistical tools and applying them to data, rather than on either the theoretical foundations of the tools or on a simple use of formulas.

Excel and Python

- We'll often use Excel and Python with data to demonstrate the concepts in this class.
- The university provides access to Excel.
- Python is open source and free to use.
- Computers are available on campus for all assignments.

Uses of Economic Data

- Describing the landscape of economics
  - Have incomes risen in the last year?
  - How has unemployment changed?
  - Has the racial wealth gap narrowed?
- Distinguishing between economic theories
  - Do voters with neighbors belonging to the same party vote more?
  - Does the gender of a Lyft driver impact rates of tipping?
  - Does cooperation increase in the repeated prisoner's dilemma in ‘easier’ settings?
- Guiding policy and expectations about the future

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

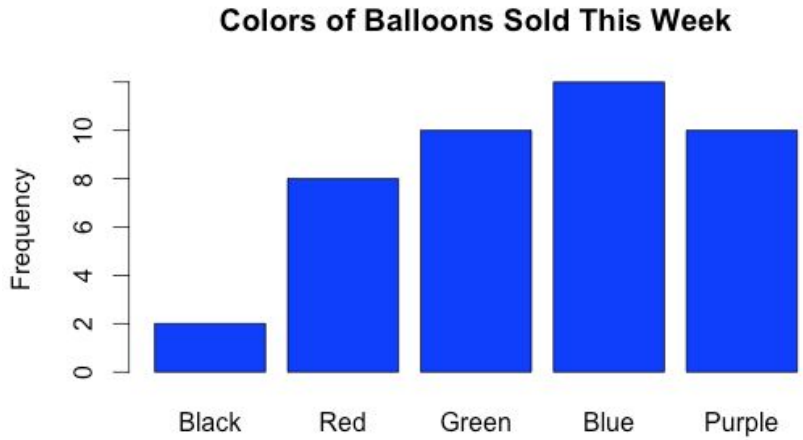
Numerical data is data that’s best recorded in numerical form, continuous or discrete.

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

Types of Data: Categorical Data

Categorical data is data recorded with values belonging to one or more groups. Categorical data can be recorded using ‘codes’ (eg. A, B, C) or using numbers with no inherent meaning.

- Gender
- Birthplace
- Religion
- Political affiliation



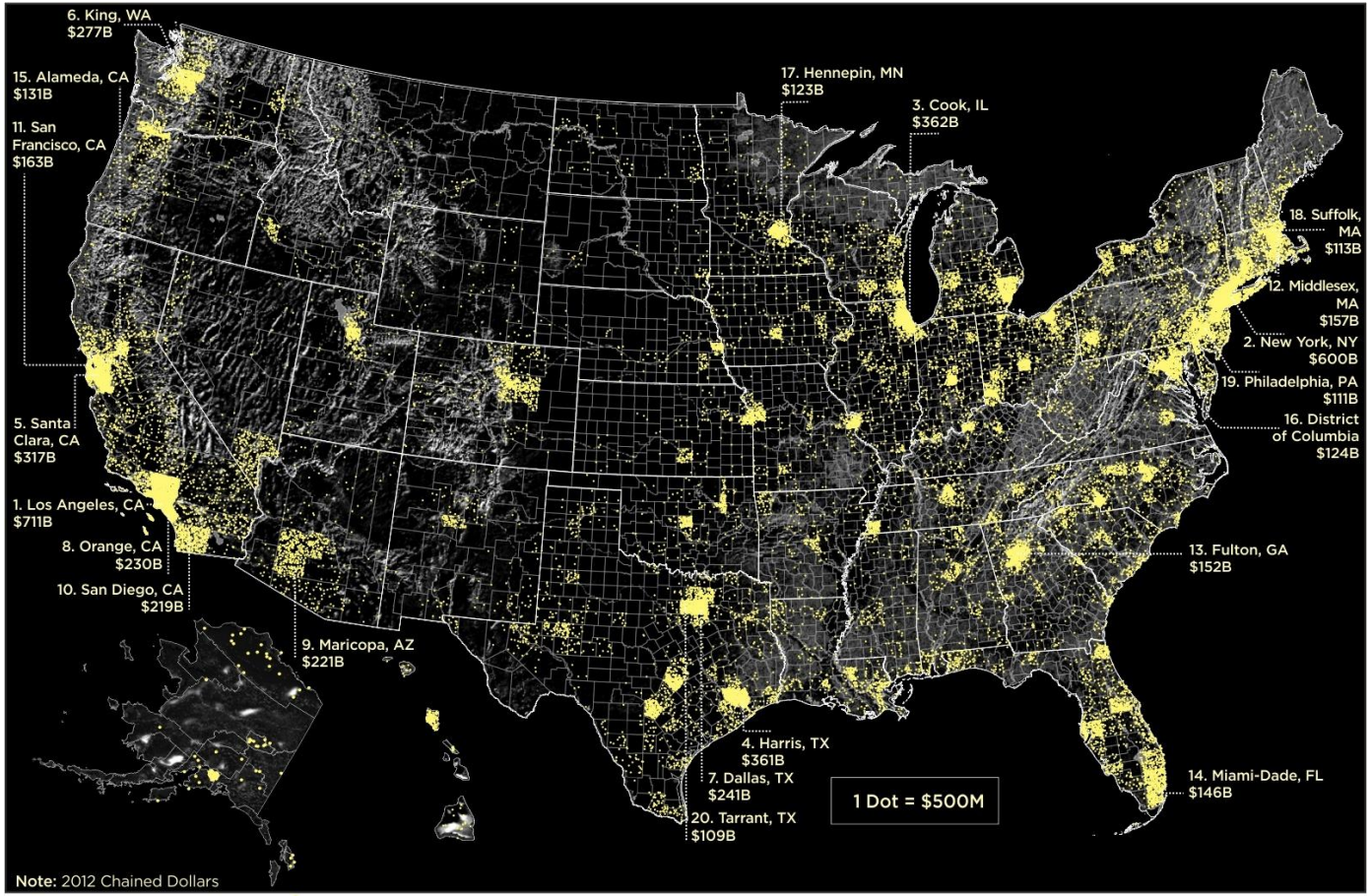
Types of Data: Cross-Section Data

Cross-section data are data on different individuals collected at a common point in time.

- A single year of census data
- GDP by country in a single year
- Unemployment rates by state for a particular year

America's Economic Output in 2018

GDP by County



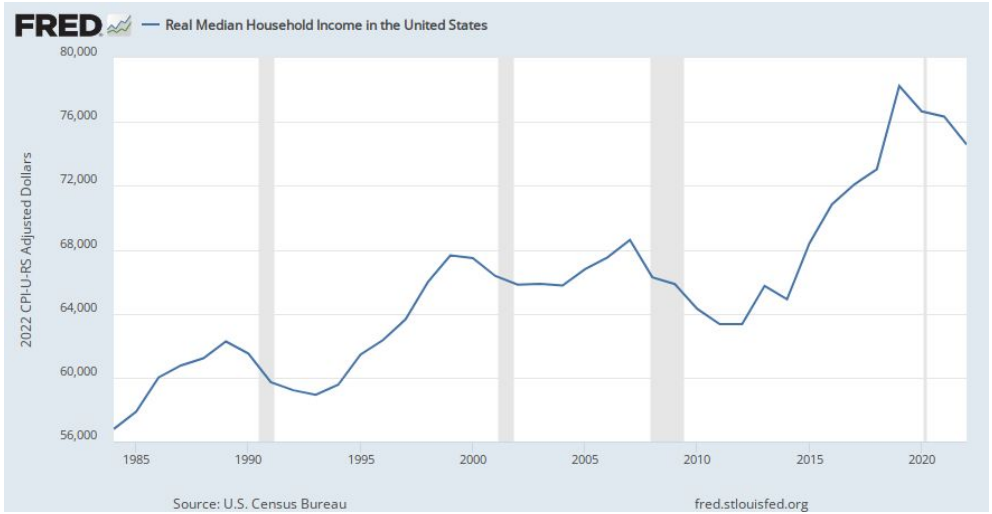
Article & Sources:
https://howmuch.net/articles/americas-economic-output-2018
Bureau of Economic Analysis - https://bea.gov

howmuch.net

Types of Data: Time-Series Data

Time series data is data measuring a phenomenon collected at different points in time.

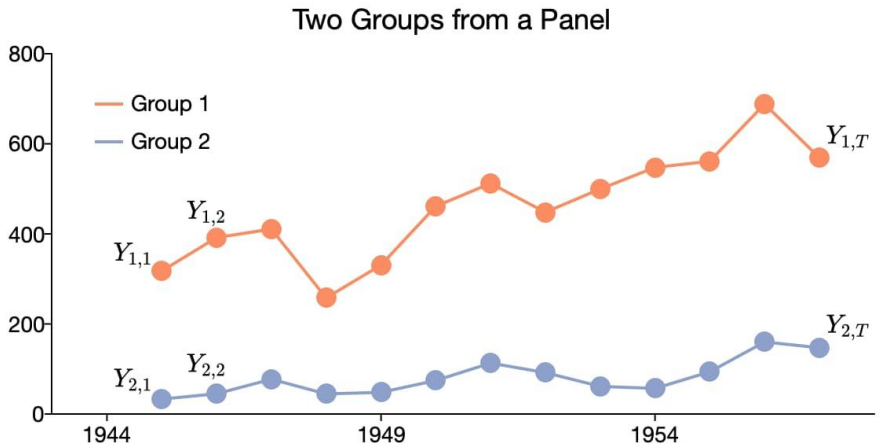
- The population of Pittsburgh between 1950 and 2020
- GDP of the US after 1980
- Median US household income



Types of Data: Panel Data

Panel data observes data on multiple individuals at multiple points through time.

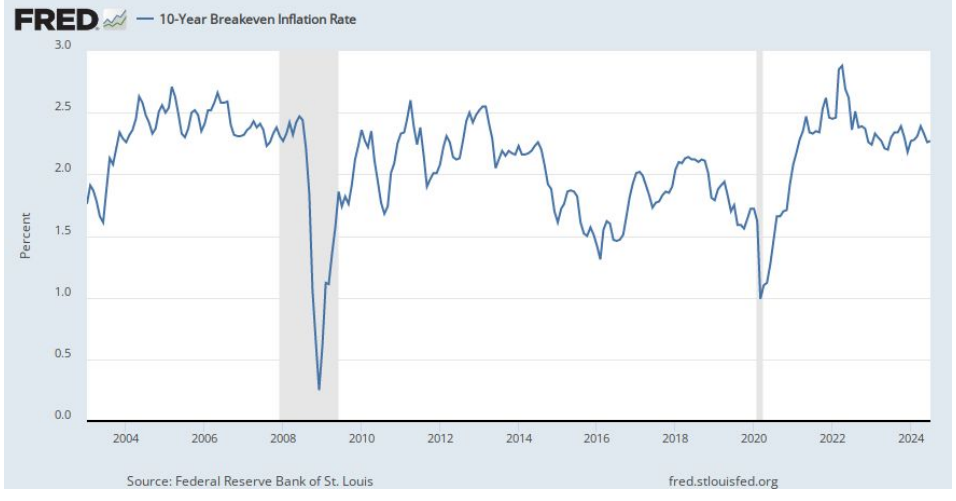
- Earnings of Pitt Econ graduates who graduated in 1995
- Life expectancy by county between 1946 and 1976
- Turnout decisions of PA voters in presidential elections, 2000 to 2024



Types of Data: Univariate Data

Univariate data is a single data series containing observations of only one variable. This can be either cross sectional, time series, or panel.

- Turnout decisions of PA voters in presidential elections, 2000 to 2024
- Earnings of high-school graduates in 2008
- Inflation rate after 2004



Types of Data: Multivariate Data

Multivariate data is composed of more than one (potentially) related data series. We're often interested in the relationship between multiple variables.

- Education (x) and earnings (y) for high school graduates
- Inflation (x) and unemployment (y) rates over time
- Education (x_1), gender (x_2), and income (y) for Allegheny County residents

Economic Data Analysis

The basic steps of analyzing economic data.

1. Data summary
    - Statistics: mean, median, variance, covariance, correlation
    - Visualizations: scatterplots, histograms, maps, timeseries
2. Statistical inference
    - Basic idea: draw conclusions about an unobservable relationship
    - We typically cannot reach definitive conclusions about a population since we typically only observe a sample
3. Interpretation
4. Communication