Module

Data Analysis

Lesson

Data Analysis

Lecture

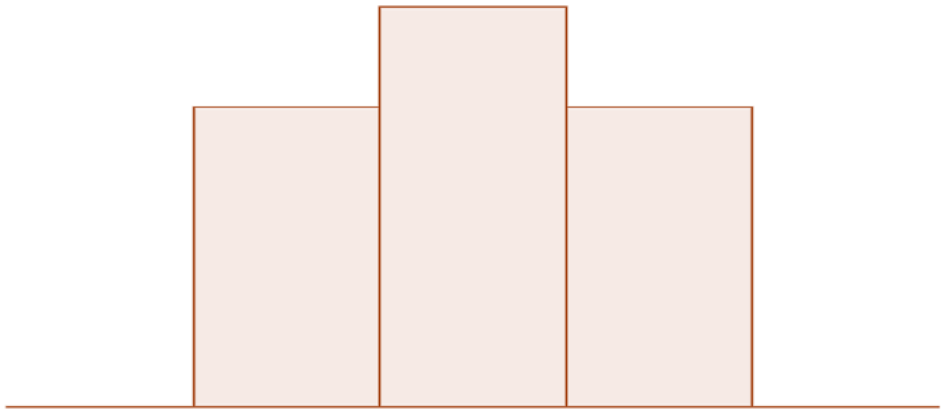
Mean and Standard Deviation

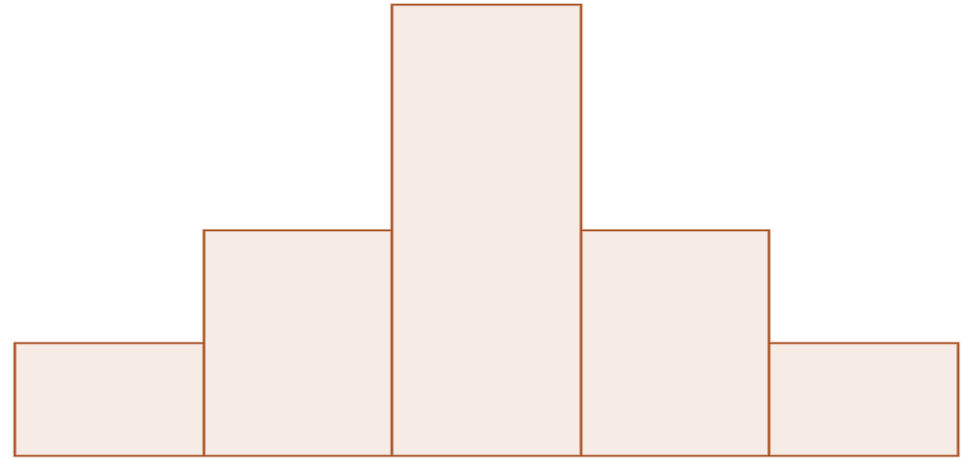
In this lecture, we'll look at mean and standard deviation as ways to analyze data

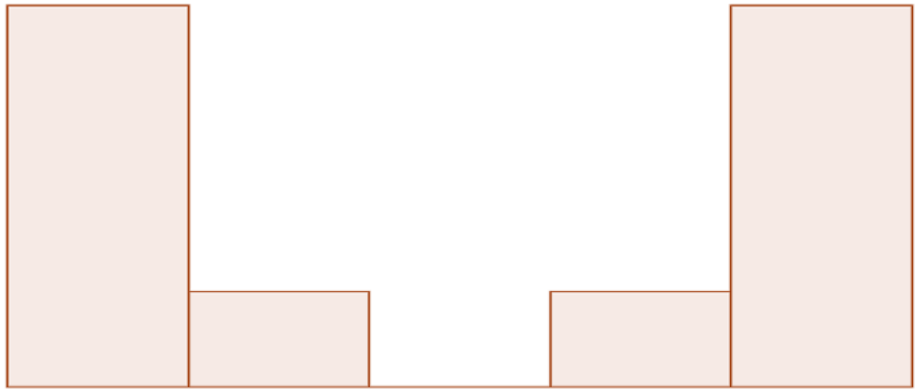
Source:

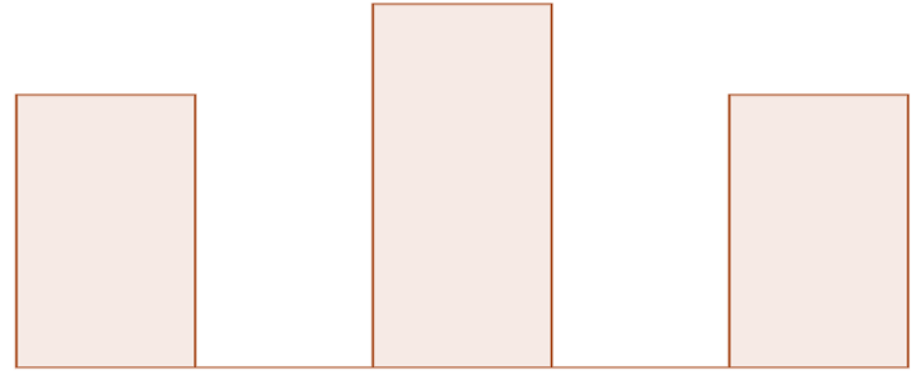
https://docs.google.com/document/d/1jfmgY-OqCCzM2dXP9oZcHbBG5s8Fv8Kjw-FXRii8-3I/edit

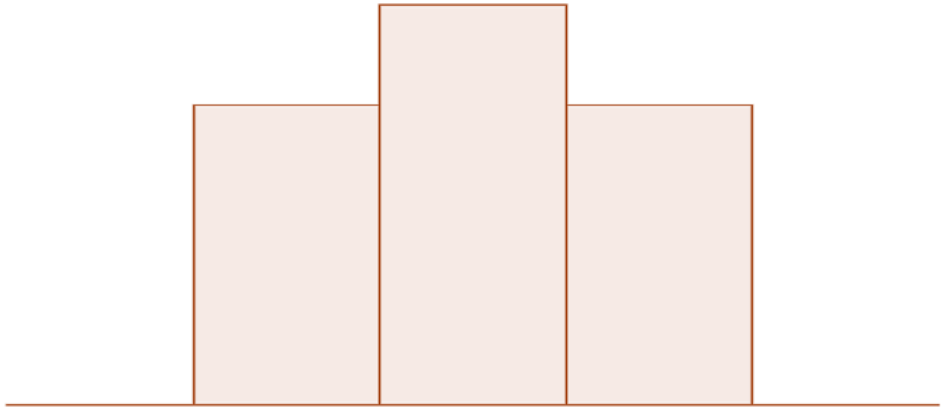
- Mean
  - Also commonly called average
- What does it tell us?
  - The central tendency of the data
- What doesn’t it tell us?
  - The actual distribution of data



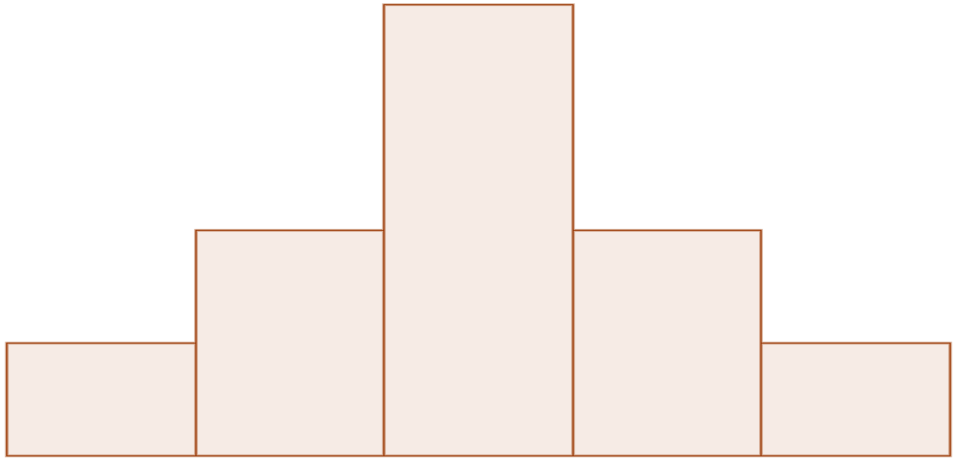




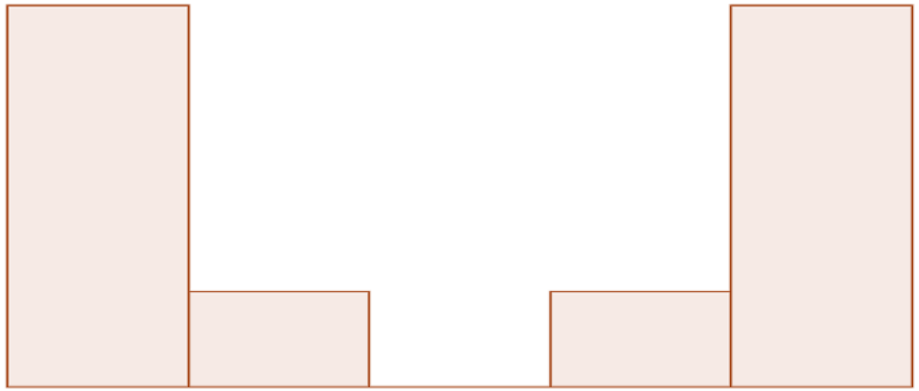




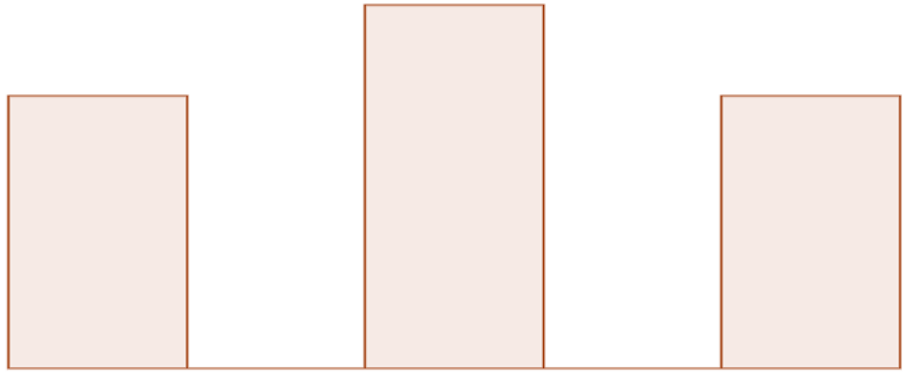
$\mu = 2$



$\mu = 2$



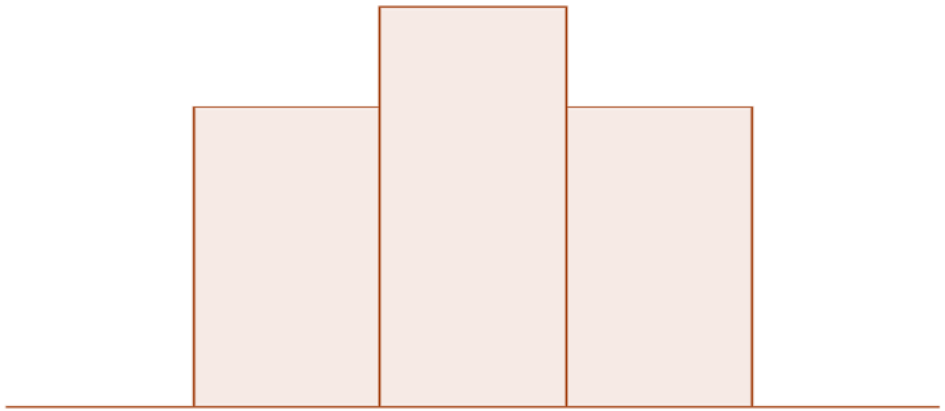
$\mu = 2$



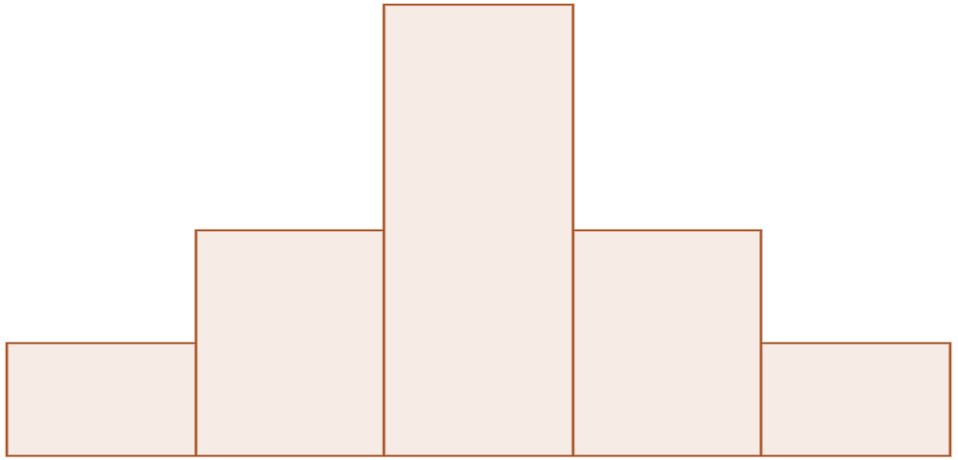
$\mu = 2$

- Mean salary in U.S.
  - $43,000
  - 130,000,000 data points
- CEOs make 262 times what the average worker makes
- Remove 100 CEO salaries from the dataset
  - What’s the new mean?

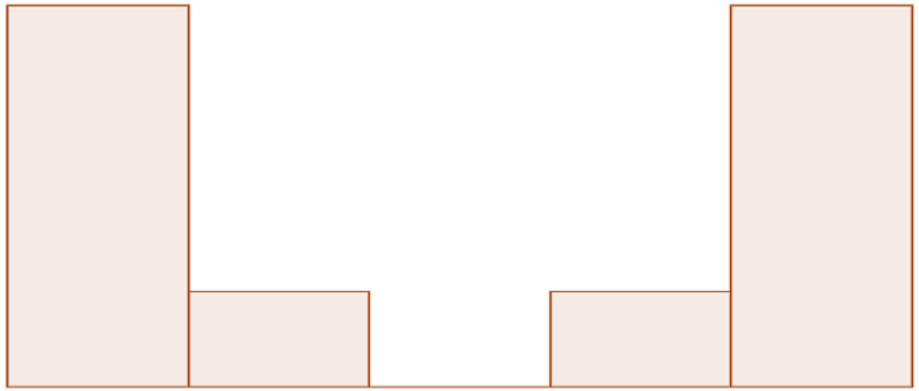
- Standard Deviation
  - Tells us how “spread out” the data is



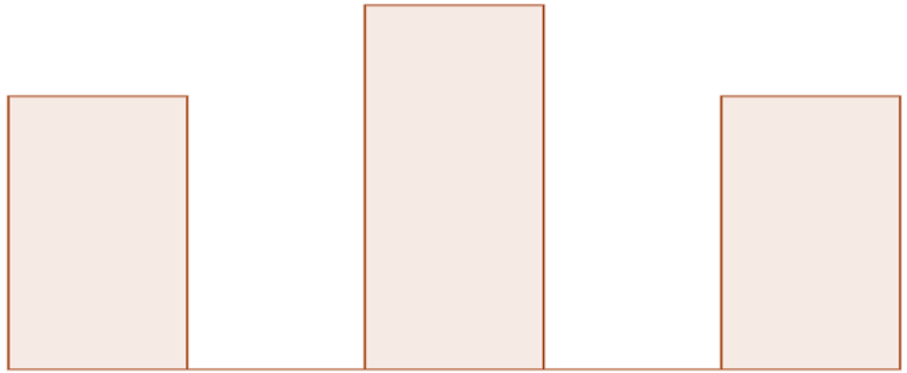
μ = 2
σ = 0.77



μ = 2
σ = 1



μ = 2
σ = 1.84



μ = 2
σ = 1.55

- Counts
  - Count the number of times each shoe size appears

- Let’s go to the movies!
  - Analyze the data for the movie attendance (we’ll use counts)
  - What kind of movie is it?

- Recap
  - Spreadsheets can be really useful for data analysis
  - Mean tells us about central tendency
  - Standard deviation tells us about the spread of the distribution