Learning Objectives

- Describe three models of central tendency
- Compute three models of central tendency from raw data
- Compare three models' responses to outliers in data & sampling variability

3 Models of Central Tendency

1. Mode (*Mo*)
2. Median (*Mdn* or P_{50})
3. Mean ($\bar{X}$ or M, μ for population)

1. ***Mode*** is the most frequent observation
   - What is the mode of this data?

ID	DV
1	9
2	5
3	10
4	5
5	10
6	5
7	1
8	10
9	10
10	8

1. Mode is the most frequent observation
   - What is the mode of this data?
2. ***Median*** is P_{50} or the score that splits the distribution evenly
   - 50% observations higher & 50% lower

When *N* is even, get the **mean** of obs. directly above & below middle

$$Mdn = \frac{(9+8)}{2} = \mathbf{8.5}$$

ID	DV
3	10
5	10
8	10
9	10
1	9
10	8
2	5
4	5
6	5
7	1

1. Mode is the most frequent observation
   - What is the mode of this data?
2. Median is P_{50} or the score that splits the distribution evenly
   - 50% observations higher/50% lower
3. ***Mean*** or average:

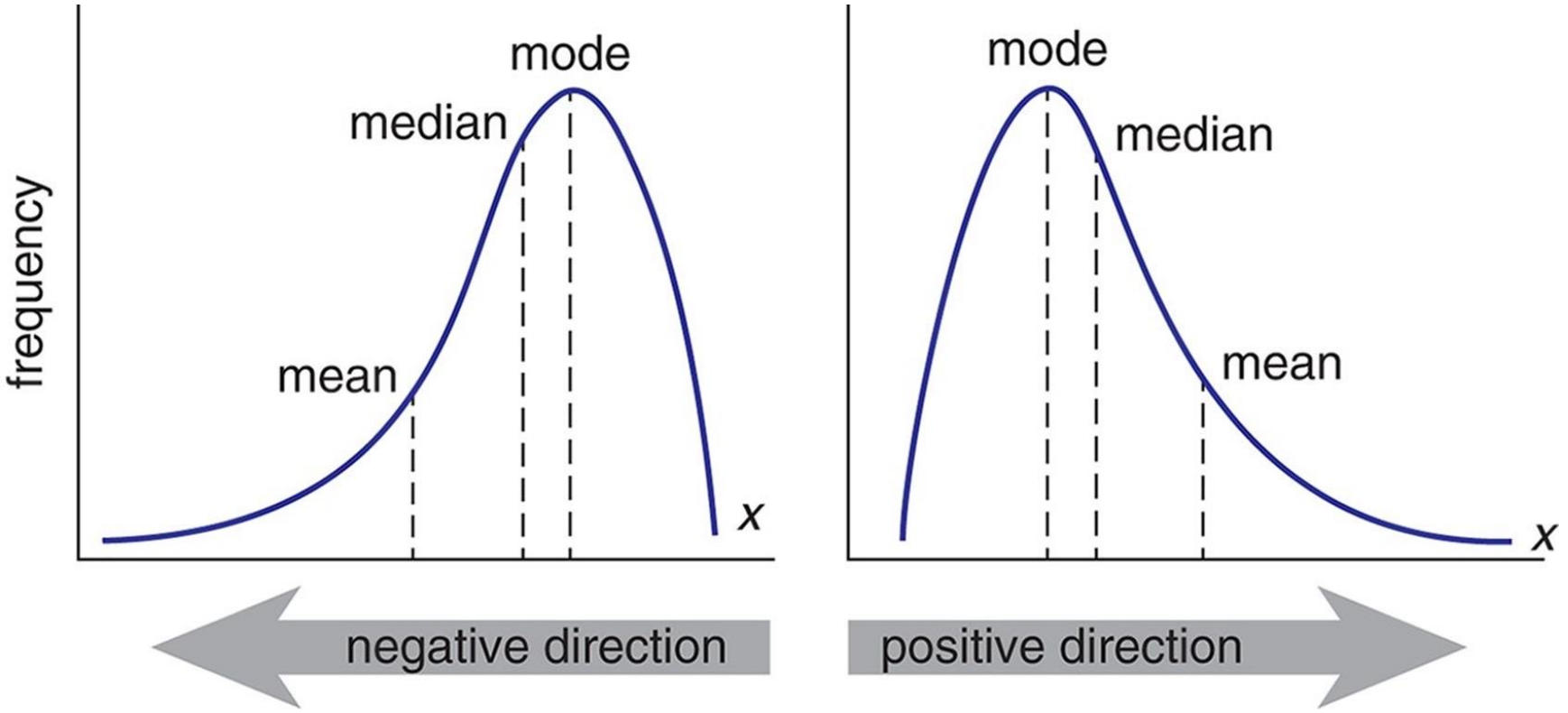
$$\frac{\Sigma X_i}{N}$$

$$\frac{(10 + 10 + 10 + 10 + 9 + 8 + 5 + 5 + 5 + 1)}{10} = 7.3$$

ID	DV
3	10
5	10
8	10
9	10
1	9
10	8
2	5
4	5
6	5
7	1

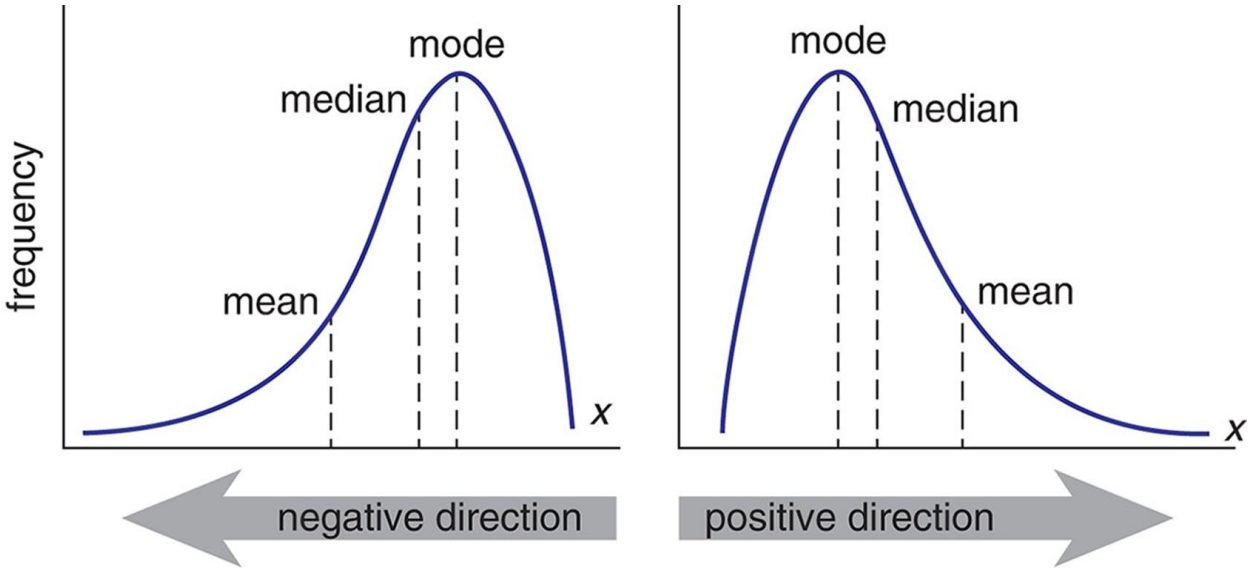
$\bar{X}$ cannot be used for ordinal data, but *Mdn* can!

– Mean is more sensitive to skew & extreme scores



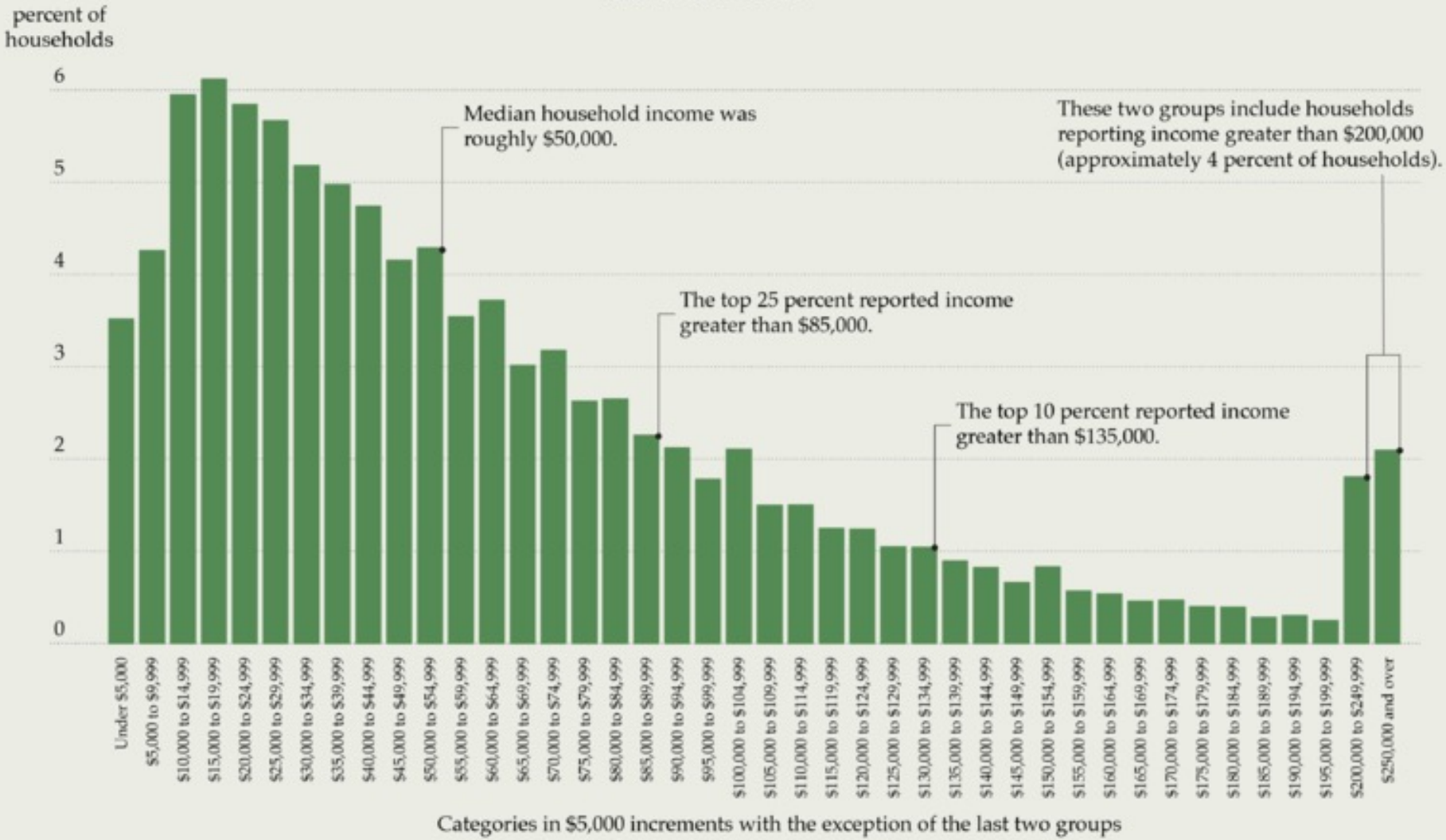
U.S. Household Net Worth

- 3 models of central tendency:
  - $692,100
  - $97,300
  - $20,000



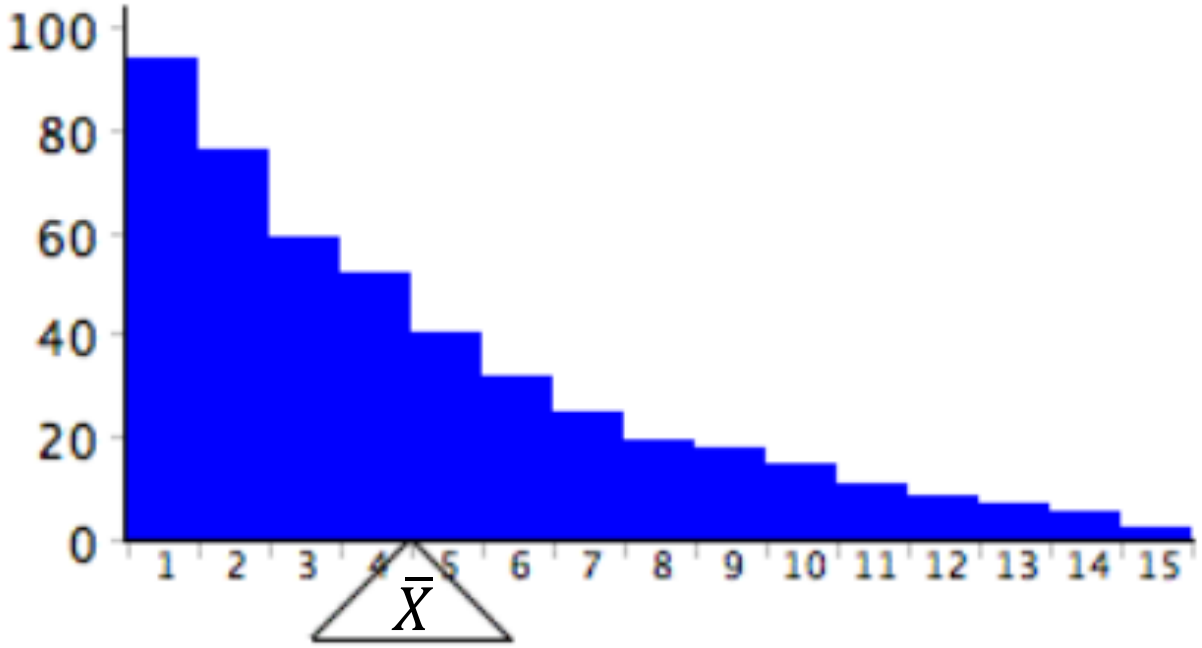
Distribution of annual household income in the United States

2010 estimate

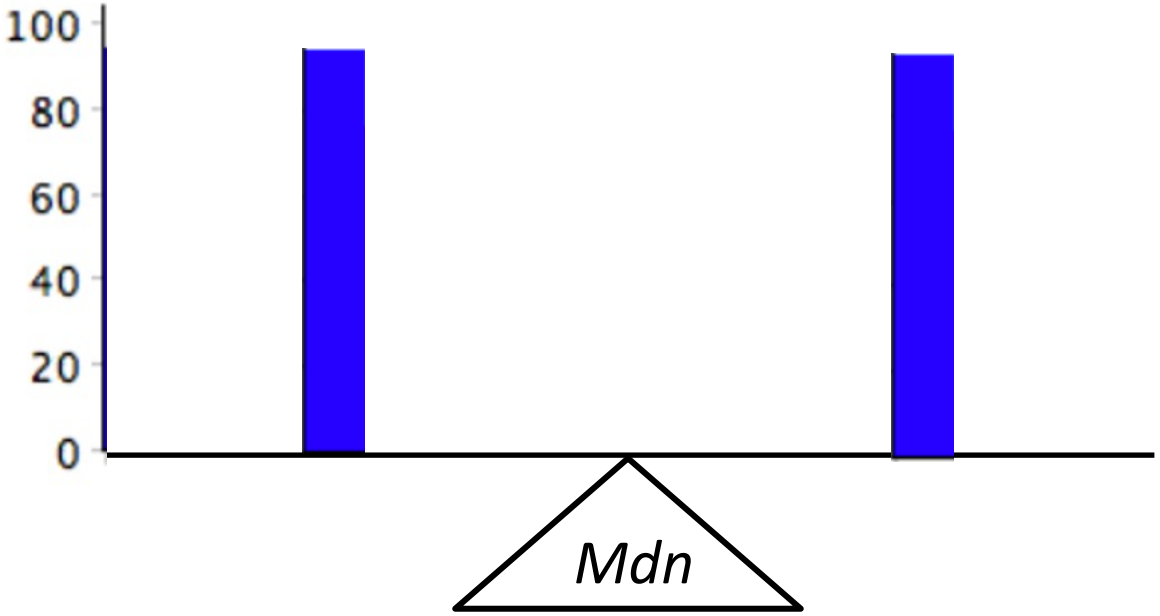


Source: U.S. Census Bureau, Current Population Survey, 2011 Annual Social and Economic Supplement

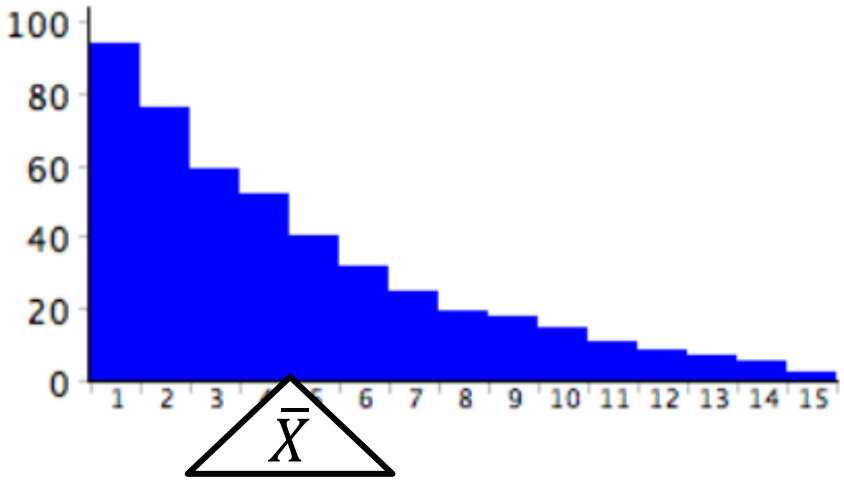
Mean is a fulcrum (balancing point)

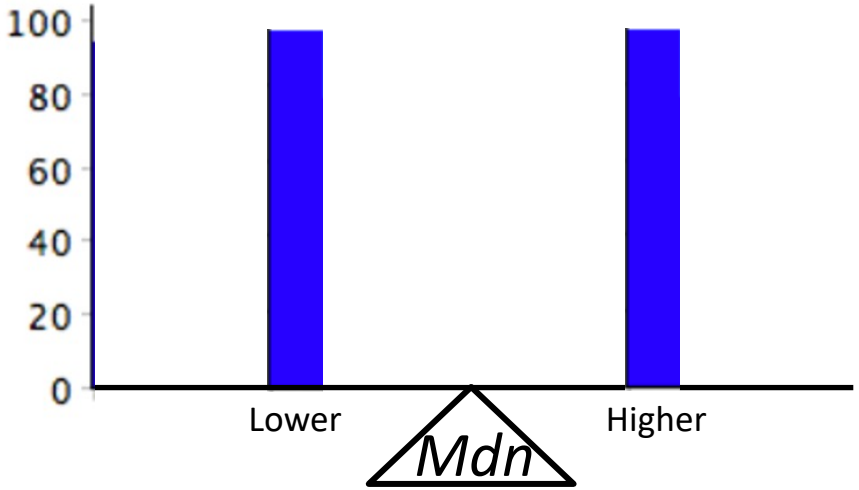


Median is also a fulcrum?!



Median is also a fulcrum?!





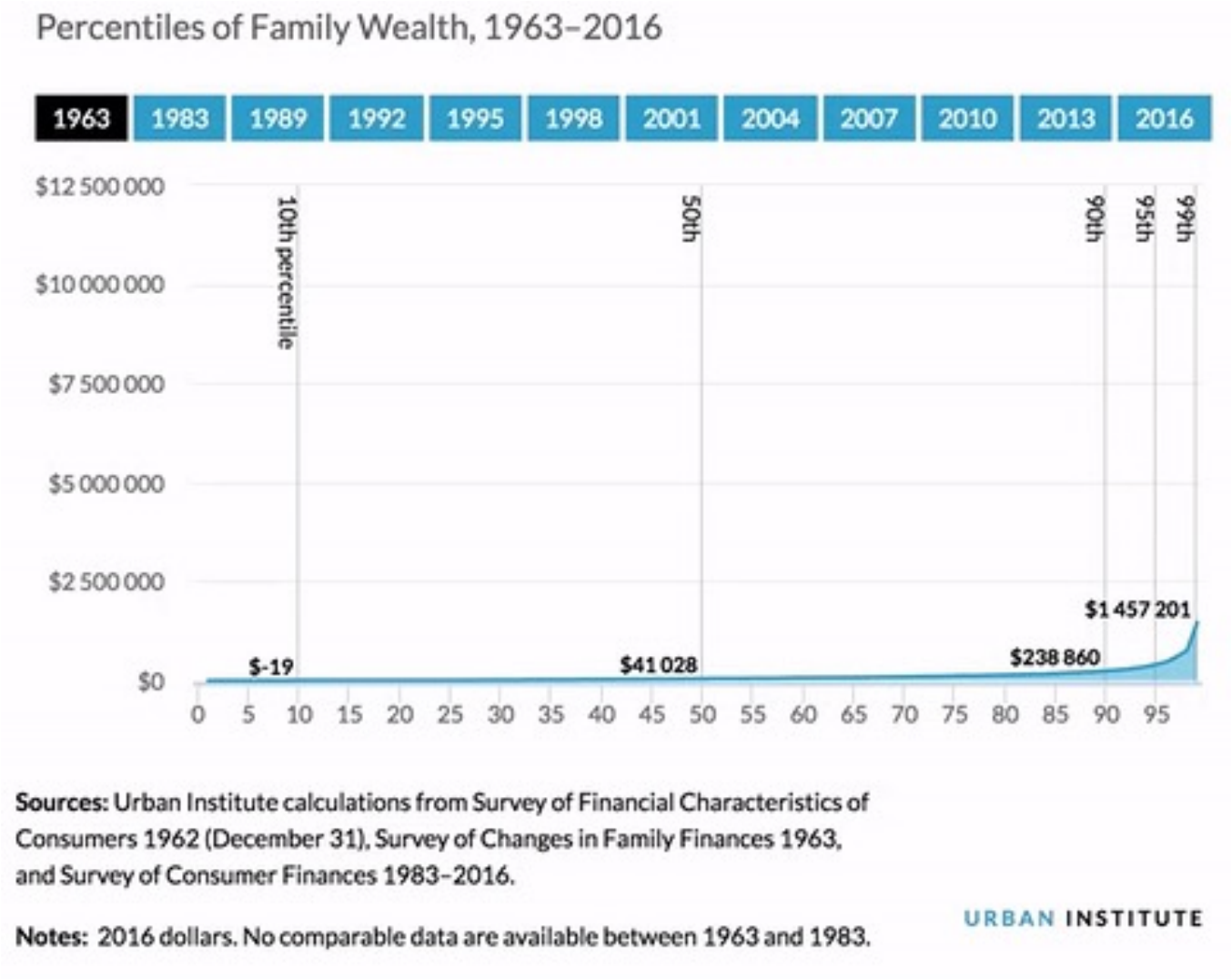
$\bar{X}$ balances the magnitude of scores above and below

- Assumes equal *intervals* separate values of variable

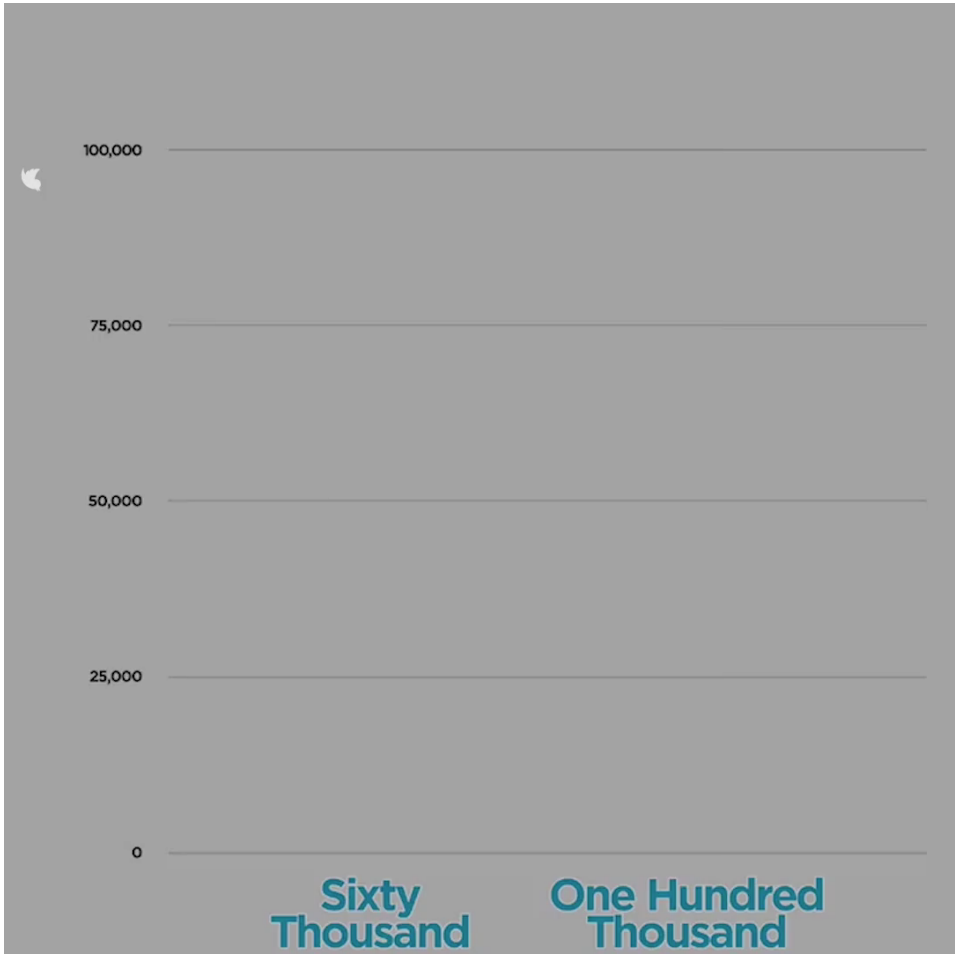
Mdn balances the number of scores above and below

- Assumes only that variable can be *ranked* lower or higher

Median is robust to skew



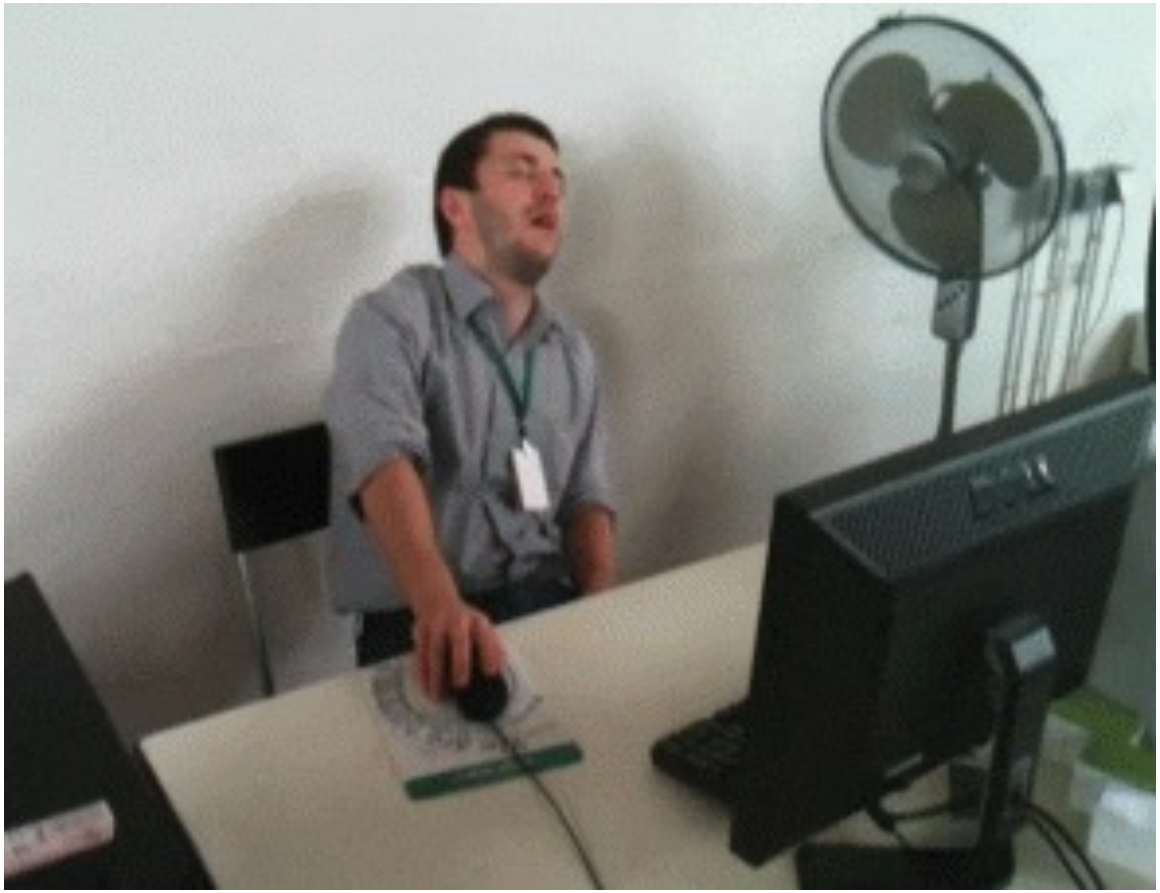
Bill Gates is an ***Outlier***



Outliers Skew Distributions

- ***Outliers*** are highly atypical observations
  - As with wealth/income, the mean is *highly sensitive* to outliers

- How can we deal with outliers?
  - Use the median as a model instead
  - Compare mean with and without outliers
    - Does the pattern of the data change?
  - Determine the source of the outlier
    - Is it a measurement error? Or a legitimate observation?



“As preregistered, we excluded all RTs greater than 10,000ms…”

Robustness in the face of outliers & sampling variability

Robustness to skew & outliers

- Mode > Median > Mean

Robustness to sampling variability

- Mean > Median > Mode

Mean is preferred to maximize consistency (then deal w/outliers later)

When sample size increases, robustness increases!

Which graph is labeled correctly?