3.1: Measures of Center:
One Number to Represent the Group

Home Values in a Subdivision

Data: $125k, $125k, $132k, $138k, $142k, $584k

- So, what is the "Typical price of a home in this subdivision?
- Who might need an average value?
- There is more than one way to find an average.

Mean

- The mean ($\bar{x}$ – read "x-bar") is the sum of the observations divided by the number of observations
- It is the center of mass

$$\bar{x} = \frac{\sum x}{n}$$

Example: Data: 24,12, 45, 33, 56, 62,14, 28

- $Mean(\bar{x}) = \frac{sum\ of\ the\ x's}{number\ of\ data\ points\ (n)}$
- $Mean = \frac{24+12+45+33+56+62+14+28}{8} = 34.25$

Median

- The median is the midpoint of the observations when they are ordered from the smallest to the largest (or from the largest to smallest)
- Order observations
- If the number of observations is:
  - Odd, then the median is the middle observation
  - Even, then the median is the average of the two middle observations

Example:

Find the sample size, *n*, and the median.

Order	Data
1	78
2	91
3	94
4	98
5	99
6	101
7	103
8	105
9	114

Order	Data
1	78
2	91
3	94
4	98
5	99
6	101
7	103
8	105
9	114
10	121

Try It:

CO_2 Pollution levels in 8 largest nations measured in metric tons per person:

2.3 1.1 19.7 9.8 1.8 1.2 0.7 0.2

Find the mean and Median.

Mean: **Median:**

Mode

- Value that occurs most often
- Highest bar in the histogram
- The mode is most often used with categorical data

Try it : Find the mean of these data sets

	Mean	Median	Mode
Set 1: 1, 2, 3, 4, 5			
Set 2: 1, 1, 2, 3, 102			
Set 3: 1, 1, 1, 1, 1, 2, 4, 69			

Think about this: Which measure of center best describes each data set?

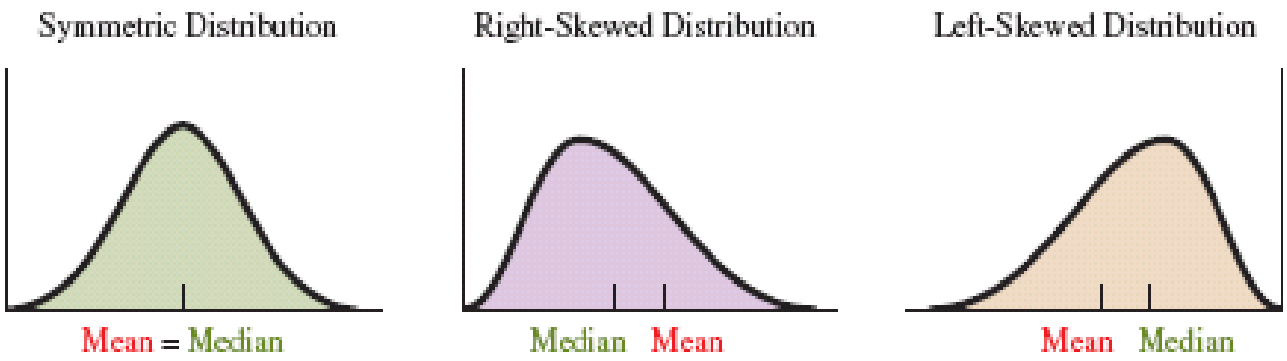
Set 1:

Set 2:

Set 3:

Comparing the Mean and Median

- **The mean and median of a symmetric distribution are close together.**
    - **For symmetric distributions, the mean is typically preferred because it takes the values of all observations into account**
- **In a skewed distribution, the mean is farther out in the long tail than the median**
    - **For skewed distributions the median is preferred because it is better representative of a typical observation**



- **A numerical summary measure is resistant if extreme observations (outliers) have little, if any, influence on its value**
    - **The Median is resistant to outliers**
    - **The Mean is not resistant to outliers**