Measures of Center

The Mode

The mode is the most commonly occurring value in a distribution.

To find the most commonly occurring value, or the mode, it is helpful to list the values in a frequency table.

Example: Average hours of sleep per night for 20 high school students.
7.5 7 7 6.5 6.5 6 7.5 5.5 7 9 8.5 8 7.5 6 7.5 8.5 6.5 8 8 6

Hours of Sleep	5.5	6	6.5	7	7.5	8	8.5	9
Frequency	1	3	3	3	4	3	2	1

The Mean

The center as balancing point

Arithmetic average of a set of observations

Suppose we have n observations: $x_1, x_2, \ldots, x_n$.

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$\bar{x}$ is also called the sample mean.

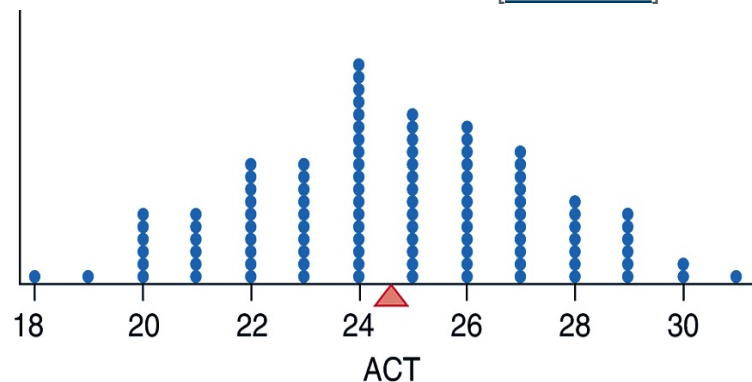**Example**: Average hours of sleep per night for 20 high school students.

7.5 7 7 6.5 6.5 6 7.5 5.5 7 9 8.5 8 7.5 6 7.5 8.5 6.5 8 8 6

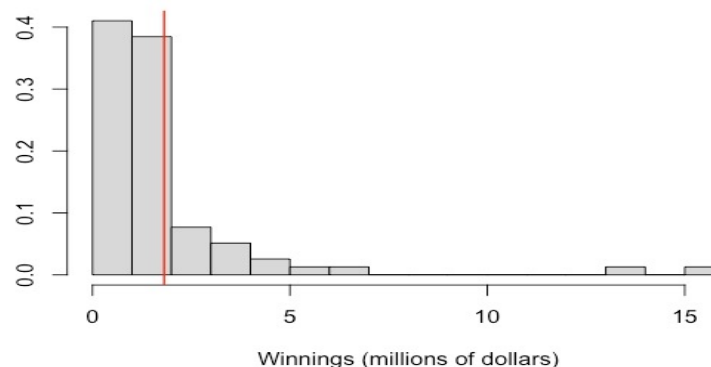| Hours of Sleep	5.5	6	6.5	7	7.5	8	8.5	9
Frequency	1	3	3	3	4	3	2	1

Measuring the Center

ACT scores for one class of statistics students
[StatCrunch]

The winnings of the 78 top-ranked tennis players in 2018 season
[StatCrunch]

The Median

The center as the middle point

▪ The midpoint of ranked values (sorted from smallest to largest).

▪ About 50% of the observations are below the median and about 50% are above it.

Steps:

1. Sort all data values in increasing order.
2. Consider whether the number of observations n is odd or even.
   - If n is odd, the median is the center observation in the ordered list, which sits at the $\frac{n+1}{2}$ position.
   - If n is even, the median is the mean of the two center observations in the ordered list which sits in between the $\frac{n}{2}$ and $\frac{n}{2}+1$ spots.

Calculate the Median -- Example

Example: Average hours of sleep per night for 20 high school students.

7.5 7 7 6.5 6.5 6 7.5 5.5 7 9 8.5 8 7.5 6 7.5 8.5 6.5 8 8 6

Mean VS. Median

- The mean describes the center as an average value, in which the **actual values** of the data points play an important role.

- The median, on the other hand, locates the middle value as the center, and the **order** of the data is the key to finding it.
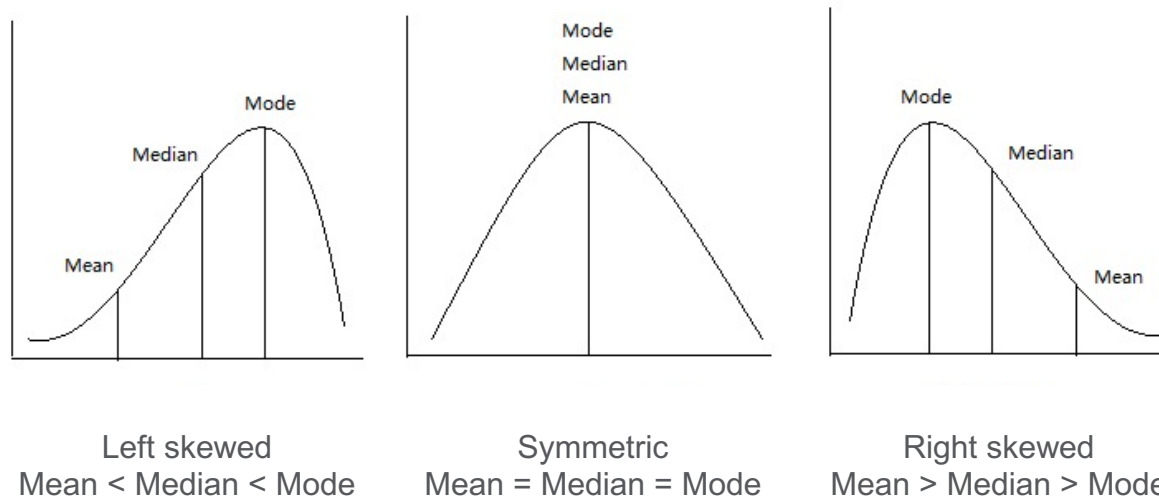
Example:

64 65 66 68 70 71 **73**

64 65 66 68 70 71 **730**

The mean is very **sensitive** to outliers, while the median is **resistant** to outliers.

Comparing Measures of Center

Left skewed	Symmetric	Right skewed
Mean < Median < Mode	Mean = Median = Mode	Mean > Median > Mode

Note:
- The long tail of skewed distributions and outliers affect the mean much more than the median.
- The median is a good measure of a typical value for skewed distributions.

Summarize Measures of Center

- The three main numerical measures for the center of a distribution are the **mode**, **mean**, and the **median**. The mode is the most frequently occurring value. The mean is the average value, while the median is the middle value.

- The mean is very **sensitive** to outliers (as it factors in their magnitude), while the median is **resistant** to outliers.

- The mean is an appropriate measure of center for **symmetric distributions with no outliers**. The median should be used to describe the center when the distribution is **skewed** or has **outliers**.