

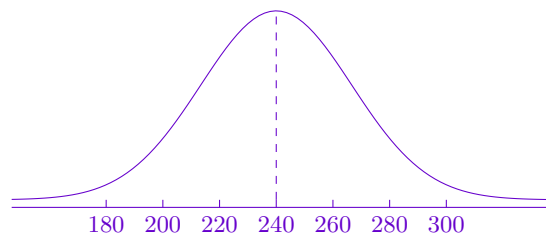
1. A data set has a mean that is much higher than the median. Which of the following is most likely **true**?
 - A. The distribution of values is symmetric.
 - B. The distribution of values is skewed left.
 - C. The distribution of values is skewed right.**
 - D. The distribution has a few high outliers.
2. The Clemson intramural basketball team has 15 players who are each a different height. The team trades its shortest player for a tall center who is now the tallest person on the team. Which of the statements is **false**?
 - A. The range of heights might be different.
 - B. The median height will remain the same.**
 - C. The mean height of the team will increase.
 - D. The standard deviation of the heights might be different.
3. **Explain** your reasoning for Question 2. Why did you choose your answer over the others?

Student answers will vary. Correct answers should include something about how the median will shift positions when the lowest observation is removed and a high observation is added.

4. The number of coffee shop customers on a given day at Central Perk follows a distribution that is roughly symmetric and unimodal with a mean of 240 customers and a standard deviation of 20 customers.
 - (a) Why is it appropriate to use the Empirical Rule here? (What do we know about the distribution that allows us to apply this rule?)

The distribution is symmetric and unimodal (bell-shaped), so the Empirical Rule applies.

- (b) Use the Empirical Rule to draw a sketch of this distribution, labeling the horizontal axis.



- (c) According to the Empirical Rule, on what percentage of days can the coffee shop expect between 200 and 280 customers?
- (d) The maximum occupancy of the coffee shop is 300 customers. According to the Empirical Rule, on what percentage of days can the coffee shop expect more than 300 customers?

We would only expect the coffee shop to have to turn customers away on 0.15% of days.

99.7% of the data fall between 180 and 300 (3 standard deviations), so $100\% - 99.7\% = 0.3\%$ falls outside this range. The graph is symmetric and we are interested in the upper tail, so $0.3\% \div 2 = 0.15\%$.

5. At an Amateur Rubik's Cube Competition, the solving times (in seconds) for each of ten randomly-selected participants are listed in the table below.

| | | | | |
|----|----|----|----|----|
| 28 | 32 | 33 | 35 | 37 |
| 39 | 42 | 46 | 51 | 59 |

Find the following statistics from your sample. For each one, be sure to **label** the values with the appropriate symbol, **show** your work, and include your **answer** with **units**.

- (a) Calculate the **mean** of the distribution.

$$\bar{x} = \frac{28 + 32 + 33 + 35 + 37 + 39 + 42 + 46 + 51 + 59}{10} = \frac{402}{10} = 40.2 \text{ seconds}$$

- (b) Find the **median** of the distribution.

$n = 10$ is even, so we average the $\frac{n}{2} = 5$ th and 6th values.

$$\text{Median} = \frac{37 + 39}{2} = 38 \text{ seconds}$$

- (c) Find the **range** of the distribution.

$$R = 59 - 28 = 31 \text{ seconds}$$

- (d) Find the **standard deviation** of the distribution. (A table like the one on p. 28 of the Lecture Notes is provided, but you do not have to use it as long as you show your work.)

| x | $(x - \bar{x})$ | $(x - \bar{x})^2$ |
|-------------|-----------------|-------------------|
| 28 | -12.2 | 148.84 |
| 32 | -8.2 | 67.24 |
| 33 | -7.2 | 51.84 |
| 35 | -5.2 | 27.04 |
| 37 | -3.2 | 10.24 |
| 39 | -1.2 | 1.44 |
| 42 | 1.8 | 3.24 |
| 46 | 5.8 | 33.64 |
| 51 | 10.8 | 116.64 |
| 59 | 18.8 | 353.44 |
| Sum: | 0 | 813.6 |

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1} = \frac{813.6}{9} = 90.4 \text{ seconds}^2$$

$$\text{If using Formula 4.5 (table not provided): } s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n - 1} = \frac{16,974 - \frac{(402)^2}{10}}{9} = 90.4 \text{ seconds}^2$$

$$s = \sqrt{90.4} \approx 9.51 \text{ seconds}$$