Descriptive Statistics Solutions

1. Calculate the mean, median and mode of the following values {2, 4, 4, 4, 12, 12, 20, 110}.

>   mean=(2+4+4+4+12+12+20+110)/8=21
>
>   median=8 (halfway between 4 and 12, the middle two values)
>
>   mode= 4 (the most frequent value)

2. True or False: If a distribution is strongly positively skewed, the median will most likely be greater than the mean.

>   FALSE. When a distribution is strongly positively skewed, it likely has a long right tail that pulls the mean to be greater than the median.

3. True or False: If a data set contains an outlier, it should always be eliminated from the data.

>   FALSE. Outliers may represent an accurate observation in the data and should not automatically be removed.   Outliers may occur because of natural randomness even when the parent population is normally distributed.

4. True or False: IQR contains 75% of the observations in a distribution.

>   FALSE. The IQR contains the middle 50% of the observations (between the 75th and 25th percentiles.

5.   If there is an outlier in a data set, its presence most strongly affects which measure of central tendency?

>   An outlier will likely affect the mean more than the other measures of central tendency.