Topic 1 Review Notes

1.

Quantitative Data: Data that contains numbers
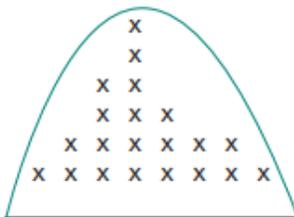
    Categorical Data: Data that lists categories not numbers

2.

Outlier: An outlier is a data point that is either much larger or much smaller than the other data points. Think of it as a data value that doesn't belong.
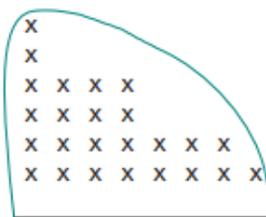
3.

Shapes of Typical Distributions of Graphical Displays of Data
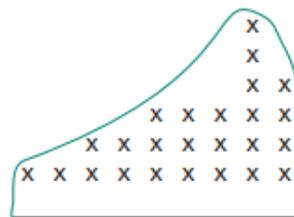

symmetric

- The left and right halves of the graph are mirror images of each other.
- The peak is in the middle, because there are many data values in the center.

skewed right

- The peak of the data is to the left side of the graph.
- There are only a few data points to the right side of the graph.

skewed left

- The peak of the data is to the right side of the graph.
- There are only a few data points to the left side of the graph.

A peak is usually the value with the greatest frequency, or one of the values with the greatest frequency, and is often surrounded by data values with other large numbers of data points.

4.

Histogram vs. Bar graphs
- Histograms have bars that touch (sometimes called bins)
- The bars in the histogram represent a range of data values. On the graph given the first bar represents 0-10 people the second bar represents 11-20 people and so on. (Bar graph bars usually represent categories)

5.

Discrete Data: Data points that can't have partial values. (When we graphed points these were the graphs where we **did not** connect the points.)

Continuous Data: Data points that can have partial values. (When we graphed points these were the graphs where we **did** connect the points)

6.

A **statistical question** is a question that anticipates an answer based on data that **varies.**
- You want to have more than one answer to your question.
- This should not be a question just about you and your preferences. It should include many people