Notes: Describing Distributions Numerically

When describing distributions, we need to discuss __________, __________, and __________. How we measure the center and spread of a distribution depends on its __________. The center of a distribution is a “typical” value. If the shape is unimodal and symmetric, a “typical” value is in the __________. If the shape is skewed, however, a “typical” value is not necessarily in the middle.

For __________ distributions, use the __________ to determine the __________ of the distribution and the __________ to describe the __________ of the distribution.

The median:
- is the __________ data value (when the data have been __________) that divides the histogram into two equal __________
- has the same __________ as the data
- is __________ to outliers (extreme data values)

The range:
- is the difference between the __________ value and the __________ value
- is a __________, NOT an __________
- is __________ to outliers

The interquartile range (IQR):
- contains the __________ of the data
- is the difference between the __________ and __________ quartiles
- is a __________, NOT an __________
- is __________ to outliers

The __________ gives: __________.

A graphical display of the five-number summary is called a __________.

How many hours, on average, do you spend watching TV per week? ______ Collect data from the entire class and record the values in order from smallest to largest. Calculate the five-number summary:
Construct both a histogram and a boxplot (using the same scale). Compare the displays.

**Average Number of Hours per Week Spent Watching TV**

---

For __________________ distributions, use the _______________ to determine the _______________ of the distribution and the _______________ to describe the _______________ of the distribution.

The mean:

- is the arithmetic _______________ of the data values
- is the _______________ of a histogram
- has the same _______________ as the data
- is _______________ to outliers
- is given by the formula _______________.

The standard deviation:

- measures the “typical” distance each data value is from the _______________
- Because some values are above the mean and some are below the mean, finding the sum is not useful (positives cancel out negatives); therefore we first _______________ the deviations, then calculate an _______________ _______________. This is called the _______________. This statistics does not have the same units as the data, since we squared the deviations. Therefore, the final step is to take the _______________ of the variance, which gives us the _______________ _______________.
- is given by the formula _______________
- is _______________ to outliers, since its calculation involves the _______________.

Find the mean and standard deviation of the average number of hours spent watching TV per week for this class.