Name: _____

DATA 151: Class 9A
Warm-up: Quartiles and Box Plots

The **five-number summary** of a distribution consists of the smallest observation, the first quartile, the median, the third quartile, and the largest observation, written in order from smallest to largest.

$$Minimum \quad Q_1 \quad M \quad Q_3 \quad Maximum$$

The five-number summary is used to…

Write the five number summaries for Washington and Colorado (We did Oregon together last class)

State	Min	Q1	Median	Q3	Max
Oregon	13	35.8	62.5	116	245
Washington					
Colorado					

    A) How does the **mean** AQI value for Washington compare to its **median**?

    B) What does this tell you about the **shape** of the distribution?

Box Plots
The five-number summary divides the distribution roughly into quarters. This leads to a new way to display quantitative data, the **boxplot.**

       How to Make a Boxplot
         1.   Draw and label a number line that includes the range of the distribution.
         2.   Draw a central box from Q_1 to Q_3.
         3.   Note the median M inside the box.
         4.   Extend lines (whiskers) from the box out to the minimum and maximum values <u>that are not outliers</u>.

IQR: One way to describe the spread of the values would be to report the quartiles, the interquartile range, or the standard deviation of the data.

How to Calculate the Quartiles and the Interquartile Range

To calculate the **quartiles:**

1. Arrange the observations in increasing order and locate the overall median M.
2. The **first quartile Q_1** is the median of the observations located to the left of the overall median.
3. The **third quartile Q_3** is the median of the observations located to the right of the overall median.

The **interquartile range (IQR)** is defined as: $IQR = Q_3 - Q_1$

How to spot outliers: The 1.5*IQR rule

In addition to serving as a measure of spread, the interquartile range (IQR) is used as part of a rule of thumb for identifying outliers.

The 1.5 ´ IQR Rule for Outliers (These are sometimes called "fences")

- $< Q_1 - 1.5 \times IQR$
- $> Q_3 + 1.5 \times IQR$

Calculate the fences for Washington. Would the minimum and/or maximum be considered outlier(s)?

Sketch a side-by-side box and whisker plot for Oregon and Washington:
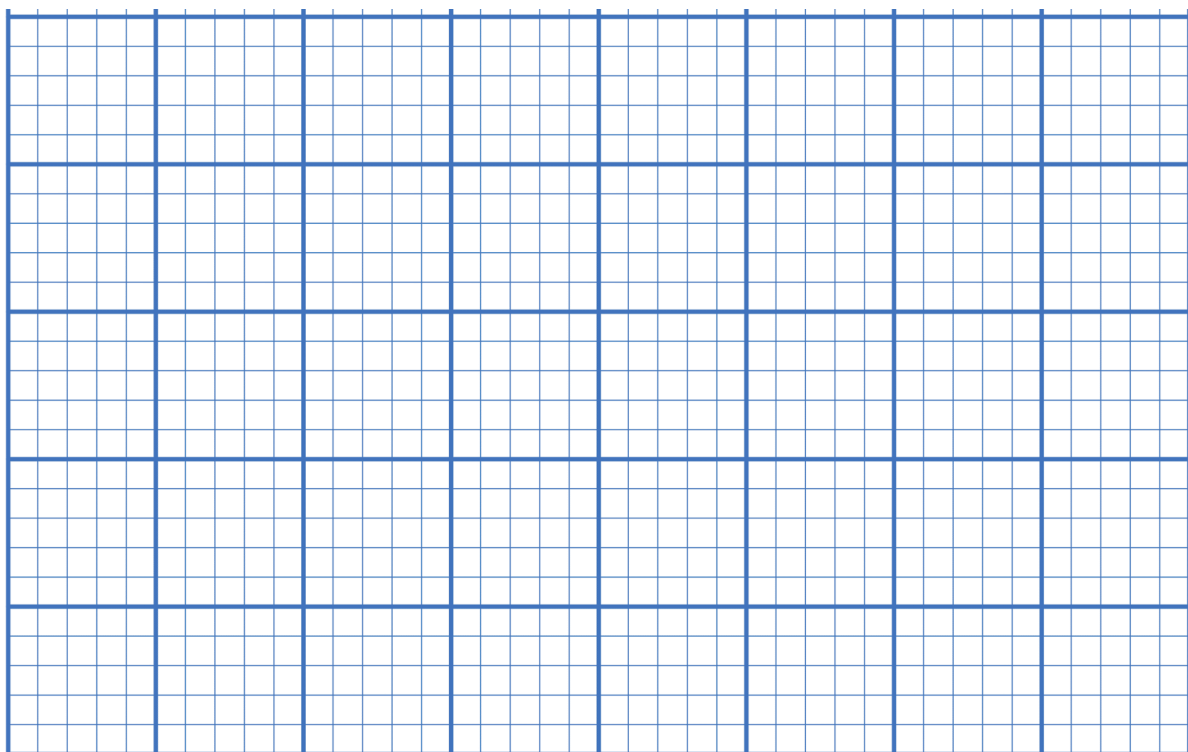
Hint: the second highest AQI value in Washington is 266 in Chehalis