

Chapter 5: Describing Distributions Numerically

Histograms:

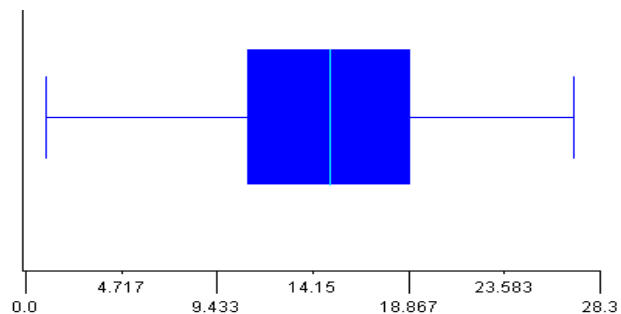
- ~ There are two ways to determine the center of the data.
 - One way is to use the **median** of the data. The median is the middle value that divides the histogram into equal areas. Thus arrange the data from least to greatest and the median would be the middle number.
 - The other method is to use the **mean** of the data. The mean is the sum of all the data and divided by the number of numbers added. ($\bar{y} = (\sum y) / n$)
- If the histogram of the data is skewed or bimodal the best way to describe the center would be to use the median. However, if the histogram is symmetric the mean could be used to describe the center, as well as the median of the data. The only time to use the mean of the data is when the histogram is nearly symmetric and there are no outliers
- ~ There are also three ways to describe the spread of the data.
 - One way to describe the spread of the data is to use **IQR** otherwise known as the **Interquartile Range**. The IQR of the data is the upper quartile minus the lower quartile. ($IQR = Q_3 - Q_1$)
 - Another method is to use the **standard deviation** of the data. The standard deviation is how far a data is from the center of the data. ($s = \sqrt{(\sum (y - \bar{y})^2 / (n - 1))}$)
 - The last method is the **range** of the data. The range is the maximum number of the data set minus the minimum of the data set. ($Range = Max - Min$)
- If the data is skewed the best measurement of spread is to use the IQR. The standard deviation should only be use if the data is symmetric and contains no outliers.

Box Plots:

- ~ A quick way to draw a box plot is using a **5-number summary**. A 5-number summary reports the medium, upper and lower quartile, as well as the maximum and minimum of a data set.

For Example:

Max	27
Q₃	19
Medium	15
Q₁	11
Min	1



- Using these numbers it can easily be recreated into a box plot. The “box” would be created using the Q₃, median, and the Q₁ by drawing a vertical line where the numbers of each are on the number line and then connecting the vertical lines at the top and bottom. To complete the box plot place a dot where the max and min are on the number line and then draw a line from the middle of the box to each point.
- When comparing two data sets, making a box plot of each makes it easy to compare.