Lab 2A - All About Distributions

**Directions: Record your responses to the lab questions in the spaces provided.**

Write down the names of the 4 variables that contain the point-totals, or *scores*, for each personality color.

Write down the names of the variables that tell us an observation's *birth gender* and whether they participated in playing *sports*.

How many variables are in the data set?

How many observations are in the data set?

Which values came up the most frequently? About how many people in your class had a score similar to yours?

What, would you say, was a *typical* score for a person in your class for your predominant color? How does your own score for this color compare?

Use a similar line of code to calculate the median value of *your* predominant color.

Are the mean and median roughly the same? If not, use the dotPlot you made in the last slide to describe why.

Do males and females differ in their typical scores for your predominant color? Answer this statistical question using your dotPlot.

Assign the mean values a name. Then place the name into the diff() function to calculate the difference. How much more/less did one birth gender score over the other for your predominant color?

We often refer to this as the variability of the data.

Look at the spread of the dotPlot you made for your predominant color then fill in the blank:

Which birth gender, if either, seem to have values that are more spread out from the center?

The mean absolute deviation finds how far away, on average, the data are from the mean.

Based on the MAD, which birth gender has more variability for your predominant color's scores?

Does this match the answer you gave for the last question in the previous slide?

Perform an analysis that produces *numerical summaries* and *graphs*.

Then, write a few sentence interpretations that addresses this statistical question and considers the *shape*, *center* and *spread* of the distributions of the graphs you create.