MOD 10: STANDARD DEVIATION (MEASURING SPREAD ABOUT THE MEAN)

(This is a modified version of a Los Medanos College activity).

Learning Goals

- Summarize the distribution of quantitative data in context. Describe the overall pattern (shape, center, and spread) and striking deviations from the pattern.
- Relate measures of center and spread to the shape of the distribution. Choose appropriate measures for different context.
- Describe the effect of outliers on a set of data.

Exploring the notion of variability (we have not defined it yet)

1) Below is a set of three distributions labeled A, B, and C. Assume the scales to be the same on each horizontal axis.



 a) Use your own visual sense to order this set of distributions from least amount of variability to most amount of variability.

2) Below is another set of three distributions labeled D, E, and F. Assume the scales to be the same on each horizontal axis.



 a) Use your own visual sense to order this set of distributions from least amount of variability to most amount of variability. 3) Below is another set of three distributions labeled G, H, and I. Assume the scales to be the same on each horizontal axis.



 a) Use your own visual sense to order this set of distributions from lest amount of variability to most amount of variability.

Now make up a mathematical method to measure the spread (total distance from the mean) or variability. Apply your mathematical measurement to each dotplot in numbers 2 and 3. Does your mathematical measurement for variability work? Explain why or why not.

4) (Wait for the mini-lecture to do this one) Explain how to find ADM and then find the ADM for the following set of data.



5) (Again ... wait for the mini-lecture to do this one). Explain how to find the standard deviation (or just write down the formula) and then find the standard deviation for the distribution below.



6) You may need scratch paper for this problem. Consider the distribution shown in the dotplot below.



Use this distribution to complete the left side of the table below. Then remove the outliers and complete the right side of the table.

With the Outliers	Without the Outliers
IQR =	IQR =
ADM =	ADM =
sd =	sd =

Which of the three measurements of spread about the center was affected the most by removing the outliers. Why?

7) The 1999 Consumer Reports new Car Buying Guide reported the number of seconds required for a variety of cars to accelerate form 0 to 30 mph. The cars were also classified into six categories by type. The following boxplots display the distributions of acceleration times for each type of car. (Note: the asterisks on the boxplot for the small type of cars, these denote outliers.)



- a) If we compare a *typical* car in each category, which type accelerates the fastest? What part of the boxplots did you compare to make your choice?
- b) If we compare the *typical range* of acceleration times for each car type, which type performs the most consistently? Which type has the most inconsistent performance? What part of the boxplots did you compare to make your choice?
- c) If the outliers were removed from the dataset of Small cars, which of the following measures of spread would be least affected? Overall range, interquartile range (the distance between the 1st and 3rd quartile marks), standard deviation? Explain.