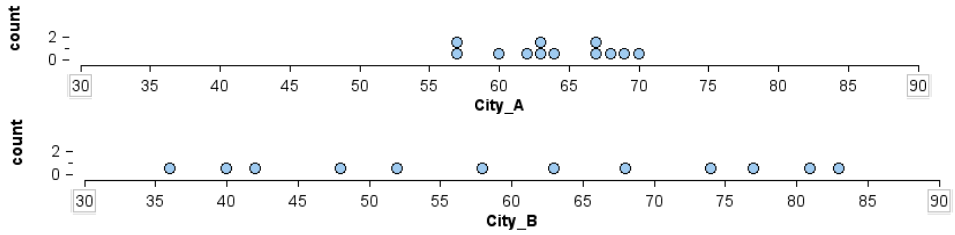
Learning Objective: Distinguish between graphs with large or small standard deviation using the concept of average deviation from the mean.

Warm-up:

1) The dot plots below show the average monthly high temperatures for New York City and San Francisco over a period of 10 years.



a) Is San Francisco City A or City B? How do you know?

b) One city has a median of 60.5°F; the other has a median of 63.5°F. Which is the median monthly high temperature in San Francisco? How do you know?

c) One city has an IQR of 6.5°F; the other has an IQR of 30.5°F. Which is the IQR for San Francisco? How do you know?

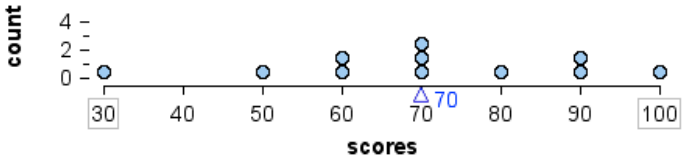
d) Here are the 5-number summaries for the two cities. Give intervals of typical average monthly high temperatures for the two cities.

City	Min	Q1	Q2	Q3	Max
A	57	61	63.5	67.5	70
B	36	45	60.5	75.5	83

e) Draw boxplots above the dot plots to summarize the monthly high temperatures.

The IQR is a way to measure variability relative to the median. How do we measure variability relative to the mean? That is the question we will investigate next. We will not start with a formula; instead we will work to build our intuition using graphs.

2) Here are exam scores for 11 students. The mean score is 70 points out of 100.



Which score varies the most from the mean? What is this score's distance from the mean?

Which score varies the least from the mean? What is this score's distance from the mean?

How far above the mean is the highest score?

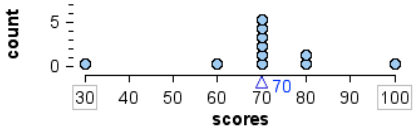
How many students had scores that vary 10 points from the mean?

Statisticians invented **standard deviation** to measure variability about the mean. ***The standard deviation is roughly the average distance that the data points vary from the mean.***

To estimate the standard deviation we will use the average distance from the mean. Fill in the table by finding the distance of each data point from the mean. What is the average distance that scores vary from the mean?

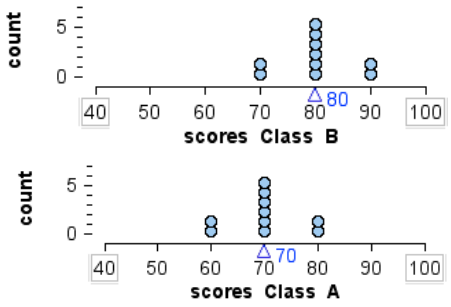
Data	30	50	60	60	70	70	70	80	90	90	100
Distance from mean		20						10			

3) Compare this graph of exam scores to the graph in #2. Here the mean is also 70. Do you think the average distance from the mean will be larger or smaller or the same? Why?

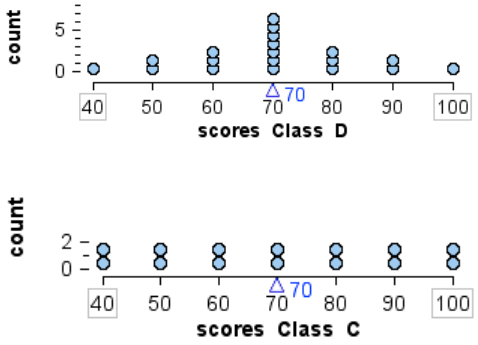


Group work: To develop your intuition about deviation from the mean, try to answer the following questions first without calculating anything, then check your intuition by calculating the average distance from the mean.

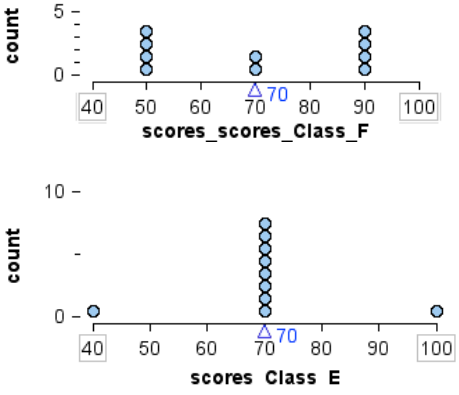
4) For each dot plot, the triangle marks the mean. Which class has the smaller average distance from the mean exam score? Or are the average distances from the mean equal? Why do you think so?



5) For each dot plot, the triangle marks the mean. Which class has the smaller average distance from the mean exam score? Or are the average distances from the mean equal? Why do you think so?



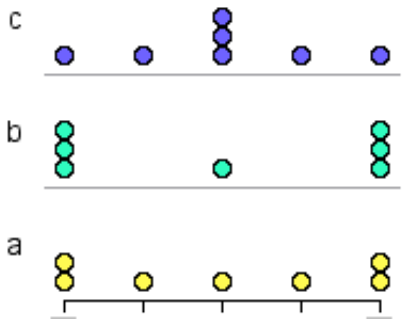
6) For each dot plot, the triangle marks the mean. Which class has the smaller average distance from the mean exam score? Or are the average distances from the mean equal? Why do you think so?



7) Draw a dot plot of 5 scores that has a mean of 70 and an average distance from the mean of zero. Is there more than one way to do this? Why or why not?

8) For the dot plots below we have removed the numerical values. We did this so that you can develop your intuition without doing any calculations. All three data sets have the same mean and are graphed on the same scale.

Which data set has the most variability about the mean? How do you know?



Which has the least? How do you know?

9) What have you learned so far about deviation from the mean? When comparing two graphs, what tips do you have for identifying the graph with the smaller deviation from the mean?

10) Now we want to think about what affects the average distance from the mean (ADM). Are the following statements (a) – (d) are always true?

- If a statement is always true, explain how you know.
- If a statement is not always true, use the calculations of the ADM for graphs (3)-(6) to provide an example that shows the statement is false (not always true).

Here are the ADM measurements for the data sets in (3)-(6).

Graph	#3	#4A	#4B	#5C	#5D	#6E	#6F
ADM	9.1	4	4	17.1	10.5	6	16

a) When comparing two data sets, the data set with the fewest data points always has the smaller average distance from the mean (ADM).

b) When comparing two data sets, the one with the smaller overall range always has the smaller ADM.

c) When comparing two dot plots of data, the dot plot with the shortest stacks of data (smallest frequencies) always has the smaller ADM.

d) When comparing two data sets, the data set with the smaller mean always has the smaller ADM.