Suggested Answers for WarmUps for Lesson 05	
1.	Describe the characteristic of the mean which you consider most important. Why do you consider it most important? Be very specific.
Answer	There is no specific “correct” answer to this question. A knee-jerk sort of reaction might be to say that it is sensitive to extreme values or outliers. This is certainly important because it suggests that unless one is careful, the mean, particularly of a small data set or sample, might lead one to an incorrect idea about a central value. The fact that the deviations from the mean sum to zero is certainly important; it is from this characteristic that the mean as balance point is defined. Of course, the fact that the mean uses all the information in a data set is quite important. Considering that the median and the mode only look at certain aspects of the data set, I am inclined to give this characteristic my vote as most important.
2.	Describe the characteristic of the mode that you find most challenging. Why is it challenging? Be specific.
Answer	I personally find the fact that the mode is unstable to be its most challenging characteristic. If one only has a mode in a situation when other measures of center are possible, one knows very little that is reliable. There may be multiple modes; the mode may bear no relationship to the mean or median; the addition or subtraction of a single observation can change the mode dramatically.
3.	Suppose you run your own coffee shop. The mean cost of supplies you use per day is $170, median of $160 and mode of $150. Next month a convention of coffee drinkers will be in town and you expect your sales to increase by a factor of 2.5! If the convention lasts all month, what should you expect the mean, median and modal values of your daily supply costs to be next month? Explain how you arrived at your answers (and do remember that the question probably has to do with something we have read about or discussed in class.)
Answer	The mean, median and mode will each increase by a factor of 2.5. That is, $170*2.5=$425; $160*2.5=$400 and $150*2.5=$375
4.	You have been asked to calculate an Interquartile Range. You have used the location formula to calculate the location of the first quartile and it is 12.75. What does this number mean? What is it telling you to do?
Answer	The number means that the first quartile will be found at the 12.75^{th} observation in the data set. This number is telling you to start at the lowest value in an ordered array and count to the 12^{th} observation in order. Next, you look at the value of the 13^{th} observation, subtract the 12^{th} from the 13^{th}, multiply the difference by 0.75 and add this to the value of the 12^{th} observation. Suppose, for example, that the 12^{th} observation in a data set was equal to 14 and the 13^{th} value equal to 16. Subtract 14 from 16 and get 2. Multiply 2*0.75 and get 1.5. Add 1.5 to 14 and get 15.5. The first quartile, the value below which is 25% of the data in the set and above which is 75% of the data in the set, is 15.5.
5.	On page 43 of your workbook you will see a frequency distribution of Mean January Temperatures, grouped data. Write in order the steps you would follow to estimate the standard deviation of this data set. Please number your steps. (See section 4.7 in your text for help.)
Answer	1. Decide on a representative value for each class. 2. Calculate that value (the midpoint) for each class. 3. Sum the frequencies of all classes to find N. 4. Calculate relative frequencies by dividing the frequency of each class by N. 5. Estimate the mean of the data set by multiplying each class midpoint by its

	class relative frequency, then adding all the products together. 6. Subtract the estimated mean from each class midpoint. 7. Square the differences for each class (the deviations). 8. Multiply the squared deviation for each class by its relative frequency. 9. Sum the products of step 8 to get the variance. 10. Take the square root of the variance to get the standard deviation estimate.
6.	How is the range a measure of dispersion? How can it measure spread? Be specific.
Answer	The range is literally a distance, it is the space between and min and a max. Because it is reporting the distance from the smallest to the largest value, the range is reporting information about spread. It is not terribly valuable, but in its purest sense the range measures spread.
7.	Go to page 52 in your work book and find the two data sets in the bottom of the page. Plot the two data sets as instructed. In your response to this question, describe how they are different. Make specific reference to the relative values of the numbers. Calculate some measure of dispersion for each and compare them. What does dispersion measure?
Answer	Data set A is more "clustered" or "focused", that is, its values are closer to one another than are those of data set B. Specifically A has a range of 7 while B has a range of 27. Set A has one repeated value while B has none. The range is a measure of spread. Loosely, dispersion measures how much territory is covered by the observations in the data set. It can be interpreted as how similar items are to one another or how different; it can convey a sense of how crowded or sparsely populated a data space is.
8.	The standard deviation of major league batting averages has declined from 0.049 to 0.031 over the last 100 years, but the mean has remained at 0.260. In terms of this lesson what specifically does this mean?
Answer	Specifically, major league batters have been getting more and more similar over the years in terms of their batting ability. There used to be much greater differences in batting ability—some players could knock any pitch out of the park (batted very well) and some couldn't hit a barn if it was thrown at them (batted very poorly). Over the years the terrific batters were passed over for players who could not only hit the ball but could contribute something else to the club (fielder, catcher, infield). At the same time, the really poor batters (who were likely good pitchers, catchers and so on) worked at their batting and became more well-rounded players. These days there is much less variation in batting average among major league ball players.