

In these two sections, we discuss the shapes, center, and variability of distributions of quantitative variables.

Activity 1. Groups. In the first two sections of Chapter 6, we learn to summarize quantitative data by various measures of center and spread. You probably already know much of this. Take about 5 minutes to talk with your group about all the definitions/concepts in the orange boxes in section 6A. Are any unfamiliar to you?

Activity 2: Individual and then in groups. In Jim's government class last semester, his grade was based entirely on five tests. He made 85, 92, 78, 20, and 88. Compute his "average" grade in two ways – as the mean and as the median. What was the effect of his really low grade on each of these?

Activity 3. Groups. Look at all the pictures of distributions on pages 375-377. Do you understand the meanings of all the labels? Discuss.

Activity 4. Groups. In 6B, p. 382. Do you think you will be happier standing in line at Big Bank or Best Bank? What is the difference?

Activity 5: Groups. Look over the information in orange boxes on p. 383 and the calculations. What does Example 1 tell you about the drawback of using the range to describe the variability. For the quartiles, do you agree that the numbers identified as the lower quartile and upper quartile are correct, according to the definitions?

Activity 6: Groups. Here are some data for commuting times (in minutes) for North Carolina workers. Compute the median and quartiles of these data. (Look carefully before you start computing. There are two places you need to rearrange before you start computing.)

10 5 10 10 10 12 15 25 20 20 30 30 40 40 60

Activity 7: Groups. After reading page 384, make a boxplot for the North Carolina commuting times.

Discussion: What is different about computing medians and quartiles in Example 2 on page 385? Answer: We have to choose a value BETWEEN two data points because that's where the middle is.

Discussion: We will NOT compute the standard deviation by hand. We will ONLY do it in Excel. However, we will talk about pages 386 and 387 enough for you to see why we interpret the standard deviation as "the average amount that a score in the distribution differs from the mean."

Rule of thumb: Estimate the standard deviation: Often a reasonable estimate of the standard deviation is the range divided by 4.

Using Excel to do various computations.

Obviously you can type data into Excel. I have also loaded some data files we will use for our course on this website:

<http://www.austincc.edu/mparker/software/data/Chapter01/1332data/index.php>

Sort data: You can select a column and sort the data in the column from smallest to largest. That is very convenient if you want to find any value in the five-number summary by hand.

Assume the data is in cells c3 to c23. Here are various useful commands:

Mean	=average(c3:c23)	
Median	=median(c3:c23)	
Mode	=mode(c3:c23)	The mode doesn't necessarily exist and this will only give one mode even if there is more than one mode in the data.
Max	=max(c3:c23)	
Minimum	=min(c3:c23)	
Quartile 1	=quartile(c3:c23,1)	
Quartile 3	=quartile(c3:c23,3)	
Standard deviation	=stdev(c3:c23)	
Range	=max(c3:c23) – min(c3:c23)	

In the next class period, we will discuss Sections 5C and 5D. Mostly we will emphasize reading graphs rather than producing the graphs.

HW 13 and Quiz 13. Due Wed. April 24 at the beginning of class.

6A 1, 2, 5, 9, 11, 15, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39

6B 1, 3, 7, 9, 11, 15, 17, 19, 21, 23

5C 3-6, 15, 17, 23, 29, 33, 39, 41 Spreadsheet: 27

5D 2, 3, 5, 6, 7, 13, 17, 21, 23, 27

Quiz 13

6A 16, 36

6B none

5C 34, spreadsheet: 58

5D 14, 16, 30, 32

Test 4 will cover sections 5A – 5D and 6A-6B. It will be available Thur. Apr. 25 – Wed. May 1 in the NRG Testing Center. It will not include spreadsheet work, of course.

Sections 5E and 6D will be the new material covered on the Final Exam, which will also include spreadsheet work.