Math 217

Agenda #1

1. Introduce myself.
2. Check roster.
3. Discuss syllabus.
4. Reminder, no class Tuesday 5/1. Students, please use this time to read chapter 1 and start working the chapter 1 study problems. Also get hold of the textbook and an appropriate calculator; bring both to class on Weds 5/2 (meet 1 PM in computer lab).
5. Day 1 survey – collect data from students.
6. Section 1.1: Displaying Distributions with Graphs
   1. In a data set, we have individuals and variables. The data for an individual make up a case.
   2. A variable describes some characteristics of an individual. Variables are either quantitative or categorical. The distribution of a variable tells what values the variable actually takes (in a given data set) and how often it takes each of those values.
   3. To display the distribution of a categorical variable, we often use a pie chart or bar graph.
   4. To display the distribution of a quantitative variable, we often use a histogram (best generated by software) or stemplot (best generated by hand).
   5. To display the distribution of observations which are taken over time, make a time plot with time on the horizontal axis. Look for trends and seasonal variation.
   6. An outlier is an observation that lies outside the overall pattern of a distribution. These are sometimes the most interesting aspects of a distribution.
   7. Examine a quantitative variable’s distribution according to five aspects:
      1. Shape
      2. Center
      3. Spread
      4. Peaks (number of major peaks)
      5. Outliers
   8. The shape of a distribution includes the number of major peaks (unimodal, bimodal, etc.) and may also be one of the following:
      1. Symmetric
      2. Skewed left (long left tail)
      3. Skewed right (long right tail)
7. Section 1.2: Describing Distributions with Numbers
   1. Use numerical measures to quantify the center and spread of a distribution.
   2. Measures of center:
      1. Mean () – not appropriate for strongly skewed distributions
      2. Median (*M* or )
   3. Measures of spread:
      1. If center is measured by median, spread may be measured by:
         1. Quartiles (Q1, Q3) and five-number summary (see also boxplot)
         2. Interquartile range (Q3 minus Q1)
      2. If center is measured by mean, spread may be measured by:
         1. Variance (s2), the “adjusted average squared deviation from the mean”
         2. Standard deviation (s), the square root of the variance. See p.51, “properties of the standard deviation”.
   4. What is a resistant measure? Which of the measures discussed in this section are resistant, and why?
   5. Linear transformations () are important in some situations, for example, when converting from one unit of measurement to another. What effect would such a transformation have on a measure of center? On a measure of spread?
   6. Don’t forget to start with a visual depiction of the distribution. Numerical measures are meant to supplement, not replace, visual representations.