

Math 169 Spring 2009 (CRN #30863) Statistics Project

- **Pick a topic** of quantitative study where an analysis of the numerical data involves finding the four measurements of central tendency : mean, median, mode, midrange , and the four measurements of variation : range, IQR, variance, standard deviation
- **Define the variable** that you will be measuring
- **Make a hypothesis (or hypotheses)** related to the data to be collected in your study
- Over the next few weeks, **gather data** (N=101) , making sure that your sample represents the population you are studying (no convenience sampling ... pick every 4th passerby at various places)
- **Find : mean, median, mode** (if one or two exist), & **midrange**
- **Find : range, IQR** (interquartile range), **variance, MAD, & standard deviation**
- A spreadsheet utility (like Microsoft [MS] Excel) is to be used in your analysis
 - Display all your raw data in the first column, titled , X
 - Sort the data in increasing order in a second column, titled , X , again
 - Use the summation function of Excel to compute \bar{X} , the mean
 - Find $X - \bar{X}$, in a third column, titled , $X - \bar{X}$
 - Find $(X - \bar{X})^2$, in a third column, titled , $(X - \bar{X})^2$
 - Use the summation function of Excel and the formulas outlined in Chapter 8, to compute the variance, v , and the standard deviation, s
- **Form a frequency chart** (using Excel) tallying & sorting your data into between 5-10 categories, making your range of values in each category, realistic groupings, for instance, for test scores :40-49, 50-59, 60-69, ...
- Using the above generated frequency chart, **construct the related histogram** (using Excel), being sure to title your graph & label your axes
- Using the above generated frequency chart, **construct the related line graph** (using Excel), being sure to title your graph & label your axes
- **Construct a related circle graph** of your choice (using Excel), being sure to title your graph & include any legends
- Use the 5-Number Summary, to **graph** (by hand or by computer, if you know how... you could use Adobe Illustrator) a **Box Plot**, using a straight edge, to ensure a professional appearance
- **Identify any outlier(s)**, if they exist, using the $[Q_1 - 1.5IQR, Q_3 + 1.5IQR]$ outlier test
- **If outliers are present, also graph** (by hand or by computer, if you know how... you could use Adobe Illustrator) a **Modified Box Plot**, again using a straight edge, to ensure a professional appearance
- In a separate report drafted in a word processing document (generated in say MS Word), **summarize your results with a 2-3 paragraph explanation of your findings**, noting values that you anticipated, and didn't anticipate (if any exist). Do not use terminology that insinuates that you have proven anything. Mention whether your results confirmed your hypothesis (or hypotheses), or not. Note how you would alter your study, if you were to come back to this topic again. What did you learn from this exercise? Did it help you to better understand statistics, and its role in society? Do you better understand how a spreadsheet utility like MS Excel can help in analyzing data, and presenting it in graphical form? What was your most and least favorite part of the project? Did you data collection help you to overcome public speaking fears? Did you properly collect your data?