R Commands for Numeric Summaries of Data and Boxplots

1. Mean, Median, Standard deviation, and Quartiles: The command sequence is **Statistics | Summaries | Numerical Summaries…** A dialog box will pop up. Select the variable or variables you want to summarize from the drop down menu. You may also obtain separate summaries for groups in this dialog box by clicking on the "Summarize by groups" button. Note that median = 50^{th} percentile = 50% in this output.
    a. Exercise: Import the 4^{th}-6^{th} graders Popularity Study data from the course website. Use R Commander to computer the mean, median, standard deviation, etc. of the **Sports variable** for girls and boys. This variable represents how the student ranked sports as to its importance for popularity by assigning a ranking of 1, 2, 3 or 4 where 1= most important for popularity, 4=least important for popularity.

	Mean	Median	Standard Deviation	Sample Size
Boys				
Girls				

    b. Comment on your results from part (a). What similarities or differences do you observe for Sports ranking between boys and girls? Is this finding unexpected?

2. Z-scores or standardized scores: To obtain a z-score for every value in your data set, select **Data | Manage variables in active data set | Standardize variables**. Every data value in your data set will be converted to a z-score = (data value – mean of data set)/ (standard deviation of data set). The z-scores are stored in a new column in your data set. You can view the z-scores by clicking the "View data set" button at the top of the R Commander window.
    a. Import the "Income per capita by country" data set from the course website. Calculate the z-score for income per capita. Do any countries have |z| >3? List them _____ What does it mean for a country to have a z-score whose absolute value is over 3?

    b. Is "income per capita" skewed or symmetric?

    c. Calculate the mean and standard deviation of income per capita. Calculate the interval $(\bar{x} - 3 * s, \bar{x} + 3 * s)$=_____ According to Chebychev's Rule, at least what percent of the incomes will fall in this interval? _____ What percent falls in the interval according to the Empirical Rule?_____. Verify that the Empirical Rule does NOT give the correct answer for these data. Why? _____ Does Chebychev's work?_____

3. Boxplots: Select **Graphs | Boxplot**. Select the variable you want to plot in the drop down menu of the resulting dialog box. If you want separate boxplots of the variable for two groups (for example, scores by gender), click the "Plot by groups" button and select the categorical variable you'd like to group on.