**Excel Project #1: Graphical Summaries of Data**

Introduction: In this project, you will be using Excel to create several different graphical representations of data. You will be provided with two datasets as Excel spreadsheets. Go to *D2L > Contents > Excel Projects > Project #1* to find the folder containing the datasets. You will find a file called '*EP1\_yourname.xls*'. Replace *yourname* with your own name and save the workbook. The workbook contains two spreadsheets that can be accessed through the tabs on the bottom of the workbook. Follow the instructions for each part and then upload the finished product to D2L before the due date. If you have any questions, email me at: [grant.moore@frontrange.edu](mailto:grant.moore@frontrange.edu) .

**Part 1: SAT Scores**

Instructions: In the first spreadsheet of the workbook you will find a sample of SAT scores. SAT scores are graded on a scale of 400 – 1600. High school graduates were randomly sampled and their scores recorded. Complete the following problems regarding this sample of data.

1. Using 8 classes, create the following:

a. A frequency distribution.

b. A relative frequency distribution

C. A cumulative relative frequency distribution

2. Using the distributions created in the first problem, create the following histograms:

a. A relative frequency histogram

b. A cumulative relative frequency histogram

c. Based on your answers for a – b, estimate:

i. The median. Save and label your answer in a cell.

*HINT*: Recall the median is the observation such that 50% of all observations in the sample are behind it! The cumulative relative frequency distribution tells you what percent of a distribution is below a class.

3. Create a pie chart based on the relative frequency distribution created in #1. Appropriately label and title the chart.

i. Based on the pie chart, what appears to be the most frequent range of scores on the SAT?

**Part 2: Exam Score vs. Time Studied Per Week**

Instructions: In the second spreadsheet of the workbook you will find a dataset involving paired observations. A survey was given in a statistics class; Students were asked how many hours they spent studying a week prior to taking the final exam. Their final exam scores were recorded against the answer they provided to survey.

3. Create a scatter plot of this data using the exam score as the dependent variable and the time studied per week as the independent variable (i.e., put exam score on the vertical axis and time studied per week on the horizontal axis). Make sure to label, scale and format your axes to fit the data. Give your graph a title.

a. Insert the line of best fit on the graph.

b. Insert R-squared on the graph.

4. Calculate the value of the correlation coefficient *r* and save it in the appropriate cell.

5. Determine whether or not the correlation calculated in the previous step is statistically signficant by following the steps:

a. Determine the number of observations in the dataset.

b. Find the critical value of the Pearson Correlation Coefficient from the table in the appendix of the book.

c. Compare the critical value against the calculate value of the correlation and make a decision about its statistical significane.

6. Using the regression line, make a prediction what exam score you expect to get if you only study 5 hours a week.