Andrew Lee

Professor Shimano

Business Analytics

Mini Project #2 Step by Step Instructions

1). Identify the variables to be used in project

2). Life expectancy is dependent variable, GDP per Capita, Population, and Schooling will be independent variables

3). Copy and Paste data from the “Country, Year, Life Expectancy, GDP per Capita, Population, and Schooling” columns into a new worksheet. This new worksheet will be the “MultipleRegressionMainAnalysis” sheet. Here, we will use the Data Analysis pack under the data section of the toolbar.

4). Upon pressing data analysis, choose “regression.” Input y range will be the Life expectancy column, and input x range will be the GDP per Capita, Population, and Schooling columns.

5). Check residuals, then press OK.

6). The summary output will show regression statistics and further data. Identify the Multiple R and R Square sections, then view the coefficients, t-stat, and p-value sections. Interpret and analyze. Schooling has a p value of 0, and a large coefficient. It is likely a factor in life expectancy.

7). From the data, develop a life expectancy equation using the coefficients: Life Expectancy = 43.52 + .000088GDP/Capita + .00000000083Population + 2.07Schooling

8). We have identified that population may not be significant. Run the regression analysis again. Life expectancy equation: Life Expectancy = 43.53 + .000088GDP/Capita + 2.07Schooling

9). Next, create another sheet, and copy over GDP and Life expectancy data. Run a regression analysis again. Look at multiple R and R Square data. Draw conclusions. Create a scatter plot from the data.

10). Lastly (since population isn’t too significant), create another sheet (“Life and Schooling”). Create a scatter plot and run regression analysis by pressing Data Analysis again. Analyze and note that schooling seems to play a role in life expectancy. It definitely has the most clear cut relationship (with Life expectancy) among the three independent variables.