1. Kyle Nabors
2. PT1. Your Dataset
   a) Included in the Project submission 1. Apologies if the data was supposed to be included here.
3. PT2. Empirical Exercise
   a) Using the sum command with the detail option we find that the median of high in the sample is 67.
   b) |
      i) Using the sum command paired with an if statement we find that the earning if height is greater than or equal to 67 inches are $44,488.44.
      ii) Using the same command but instead checking for less than we find that the average earnings for people less than 67 inches are $49,987.88.
      iii) Using a t-test we can test the null hypothesis that the difference in mean of wage between short and tall workers is 0. Analysing the t-test we reject the null hypothesis that the difference in mean is 0 and accept the alternative hypothesis that the difference in the mean wage of short workers compared to tall workers is greater than 0. This means that short workers make more than tall workers.
   c) Using the scatter command, we can construct a scatterplot of wages and height. The plot consists of 23 horizontal lines. This is because the data for wage was collected using 23 income brackets instead of the actual value. So, everyone will fall into one of 23 categories leading to the lines we see in figure 1.
   d) |
      i) Using the regress command, we find that the estimated slope of the regression of earnings on height is 707.67.
      ii) Using estimated regression, we can predict the earnings of workers based on height. In doing this we find that a worker who is 67 inches tall will make $46,857, a worker who is 70 inches tall will make $48,987, and a worker who is 65 inches tall will make $48,978.
   e) By converting height to cm instead of inches we can run regressions on people's earnings based on their height in centimetres. Doing this we find the following information.
      i) The slope of the regression is 283.07.
      ii) The estimated intercept is -512.73.
      iii) The R^2 value is 0.0109
      iv) The standard error is 20.16
   f) Limiting the dataset to only females and running a regression on it we find the following information.
      i) The estimated slope is 511.22.
      ii) If a woman is 1 inch taller than the average, her earnings would be higher by $511.22.
   g) Now we can do the same regression for men and see how the data differs.
      i) The estimated slope is 1306.86.
      ii) If a male is one inch taller than the average his average income increases by $1306.86.
   h) I do think height is correlated with other factors that cause earnings. Just from the analysis above it seems to be correlated with gender. I do not think that the error term U_i has a conditional mean of zero given Height (X_i).

Appendix

Figure 1: