Using Dummy Variables

Using data the data set CPS-Econ317, answer the following questions. Note, this is a large data set and you may need to increase the amount of ram allocated to Stata.

(a) The data set contain information on the hourly wage (wage) and years of schooling (school) for individuals in 2006. What is the mean wage of the sample? What is the standard deviation of wages in the sample?

(b) Show a distribution of wages. Describe your results.

(c) What is the mean level of schooling in the sample? What is the standard deviation of schooling in the sample?

(d) Show the distribution of schooling. Describe your results.

(e) Show a scatter diagram examining the relationship between wages and schooling. Describe your graph.

(f) Estimate the following regression model: \( \text{Wage} = \beta_0 + \beta_1 \text{School} + u. \)

(g) Interpret your results.

(h) Show graphically the estimated regression in your scatter diagram.

(i) The variable race identifies four groups White, Black, Asian and Hispanic. What is the mean wage of each group?

(j) Show the distribution of wages for each group. Hint, use “by” function. Describe your results.

(k) What is the mean level of schooling of each group?

(l) Show the distribution of schooling for each group. Describe your results.

(m) Use the race variable to construct a set of dummy variables for race.

(n) Set up a regression model using your dummy variables to see if there is any difference in mean wage among the four groups.

(o) Estimate the regression model above. Interpret your results.

(p) Show graphically each regression in a single graph. Note there should be four regression lines.

(q) Set up a regression model using your dummy variables to examine the difference in mean wage among the four groups holding schooling constant.

(r) Estimate your regression model. Interpret your results. Do your result differ from part (o)? Explain.

(s) Show graphically each regression in a single graph. Note there should be four regression lines.