

SOC510 Data Analysis Homework 2

Due December 7th (Thursday)

To get the full credit (30 points), submit a printed copy of the MS-Word (or any other typesetting program) document in which your answers are typed with a hardcopy of your R results to the Sociology Main Office (Fraser 716) by 4pm on December 7th (Thursday).

Using `soc510hw2.csv`, do the following data analysis:

Table 1: Variables of `soc510hw2.csv`

Variable	Remarks
female	– 0 male; 1 female
race	– 1 white; 2 black; 3 hispanic; 4 others
educ	– 1 Less than high school (LTHS); 2 High school graduate (HSG); 3 Some College (SC); 4 Bachelor degree (BA); 5 Graduate degree (Grad)
married	– 0 not-married; 1 married
forborn	– 0 US-born; 1 foreign born
pubst	– 0 working in private sector; 1 working in public sector
workhour	– usual working hours a week
union	– 0 non-union member; 1 union member
occ	– occupation
wage	– hourly wage

- Estimate the mean and standard deviation of wage
 - Estimate the mean and standard deviation of wage by race
 - Estimate the standard error of wage by race
 - Estimate the mean and standard deviation of wage by union membership
 - Estimate the standard error of wage by union membership
- Estimate the regression model in which wage is a dependent variable and usual working hours is an independent variable
 - Interpret the result (i.e., interpret intercept and slope)
 - Do the same analysis limiting your sample to men.
 - Do the same analysis limiting your sample to women.
 - Compare and interpret the results between men and women.
- Do t-test to see if there is a significant wage difference between married and not-married ($H_a: \mu_{\text{married}} \neq \mu_{\text{not-married}}$)
 - Interpret the result
 - Do the same test with only female workers and interpret the result.
 - Do the same test with only male workers and interpret the result.

4.
 - (a) Create a new variable called 'BA+' in which LTHS and HSG are coded 0 and SC, BA, and Grad are coded 1.
 - (b) Do t-test to see if the mean wage for workers with BA+ is higher than that for workers without bachelor degrees. ($H_a: \mu_{BA+} > \mu_{no-BA}$)
 - (c) Limit your target population to the private sector and do the same test. Interpret the result.

5.
 - (a) Create another new variable called 'minority' in which whites are coded 0 and all other races are coded 1.
 - (b) Do t-test to see if the proportion of BA+ for minority workers is lower than that for white workers. ($H_a: P_{white} > P_{minority}$)
 - (c) Limit your target population to the private sector and do the same test. Interpret the result.