MET CS 555

Hao Wu

27/10/2021

Assignment 4

Question 1

图表, 散点图

描述已自动生成

Scatter plot form linear relationship, and positive relation. The correlation between prestige score and years of education is 0.8501769.

Question 2

图表, 散点图

描述已自动生成

The coefficient of constant is -10.732, and the coefficient of years of education is 5.361. Both P-value is less than 0.05. so we can get the equation

*Prestige Score = 5.361 \* Years of Education – 10.732*

There have some outliers between 40 – 50, but it won’t affect our final result.

Question 3

The coefficient of constant is -6.7943, and the coefficient of years of education is 4.1866, the coefficient of Income is 0.0013, and the coefficient of percentage of women is -0.0089. All of them are statistic significant on 95 confidence intervals except percentage of women. Equation below:

*Prestige Score = -6.7943 + 4.1866 \* Years of Education + 0.0013 \* Income + -0.0089 \* Percentage of Women*

Question 4

For income, the P-value is less than 0.05 so that we saw it is statistic significant at 95 percentage of confidence intervals. The coefficient of income is 0.000287. Confidence Intervals is (0.00233,0.00345)

For percentage of women, the P-value is greater than 0.05 which is 0.236, so we can’t reject Null hypothesis which is H0:βpercentage of women = 0; therefore, we can’t say it has significant relationship between prestige score.

Question 5

Residual Income

图表, 散点图

描述已自动生成

Residual Percentage of women

图表, 散点图

描述已自动生成

In the residual plot for income, we can saw less outlier exist, therefore it is fit of the model reasonable; however, if we take a look at residual plot for percentage of women, we will find out the residual point is separate, it will significant influence the final result, and let us know the week relationship between percentage of women and prestige score.