**Financial Econometrics**

**Project #2**

Name: Xue Jin

TU ID: 915697511

Notice: please submit the completed Word document though the link on Canvas before the deadline. Please name your document in the format of “Project 2 – Your Name”.

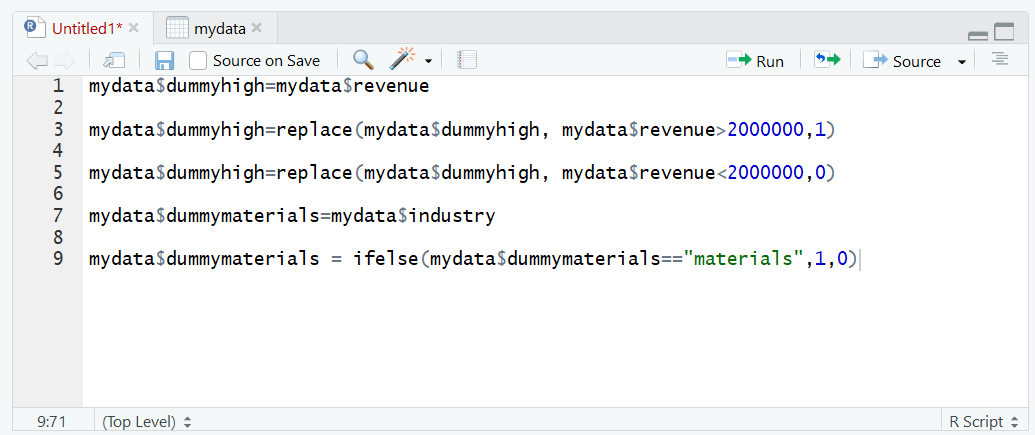
This project is to study how different factors can affect firms’ stock prices (y) with multiple linear regression models. There are 4 independent variables, including firms’ PE ratio (denoted as x1), firms’ dividend payment (denoted as x2), firms’ revenue amount (denoted as revenue), and firms’ industry (denoted as industry).

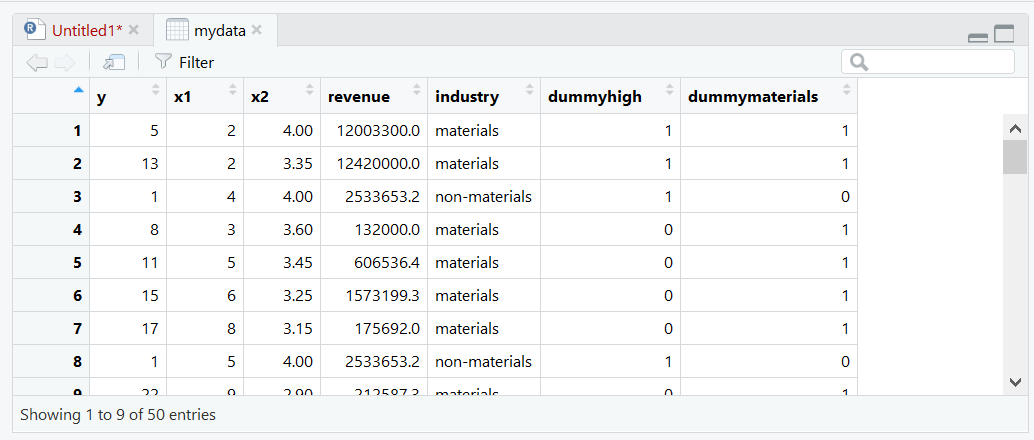
The dataset “project2data” is randomly created for you to practice and to understand how to do econometric analysis by using R.

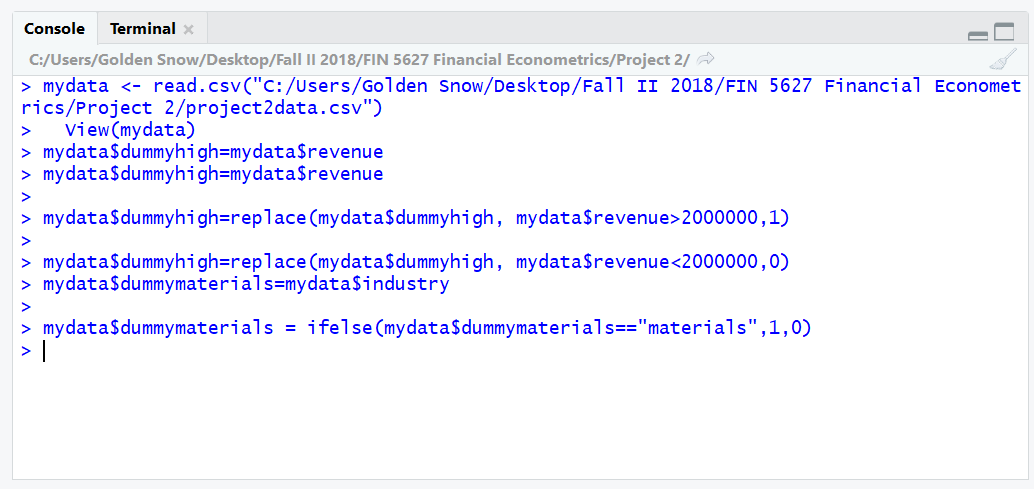
1. Create two dummy variables.

The first dummy variable denotes the high revenue firm. If firm’s revenue is greater than 2 million, then the firm has high revenue, and its first dummy variable is equal to one. If firm’s revenue is less than 2 million, the firm has low revenue, and its first dummy variable is equal to zero.

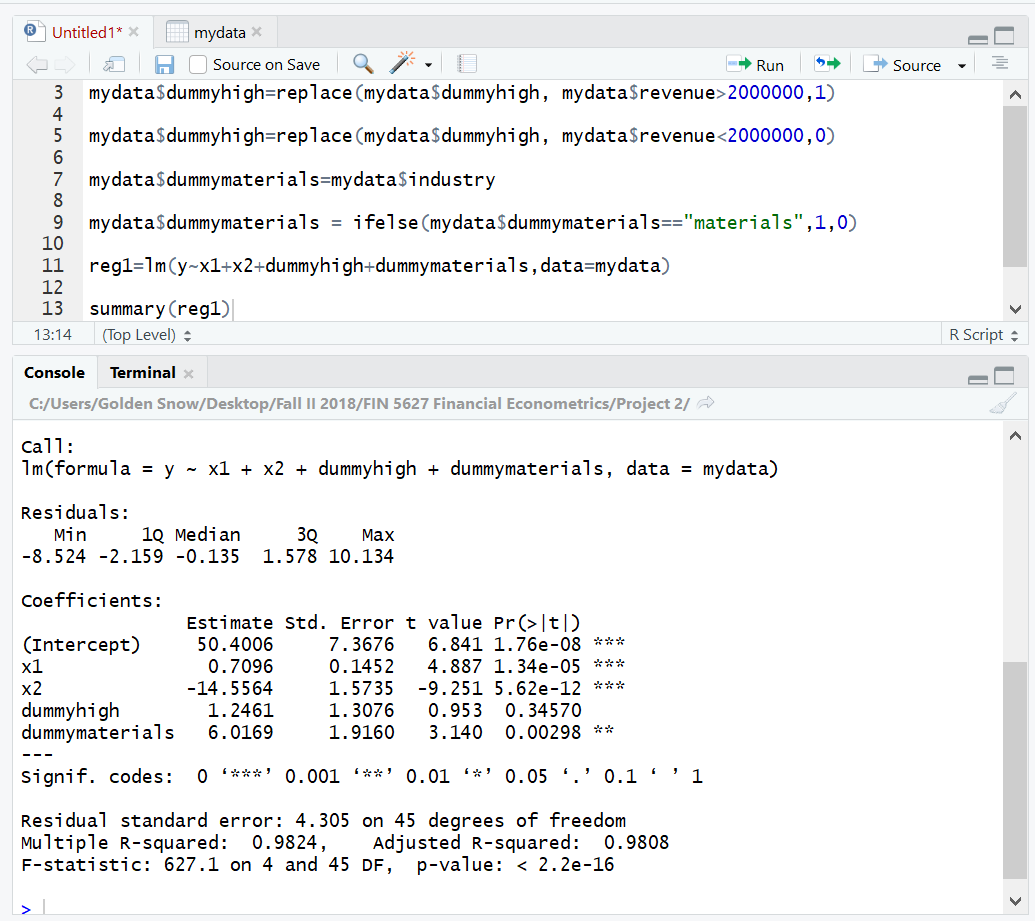
The second dummy variable denotes the materials firm. For the firm whose industry is materials, the second dummy variable is equal to one. For the firm whose industry is non-materials, the second dummy variable is equal to zero.







1. Run the multiple linear regression without dummy variables interaction (i.e. the four independent variables are x1, x2, and two newly created dummy variables). Please write the regression equation below.



1. Analyze the summary of regression results (i.e. state the estimated values of regression coefficients, interpret the use of R-squared, determine if each population regression coefficient is different from zero by using t-value or P-value. Interpret if the high revenue firms’ stock prices are different from low revenue firms, and if the materials firms’ stock prices are different from non-materials firms)

, , , , ,

The R squared is 0.9808 which is very close to 1. The regression model fits the observations good. The regression model accounts for more of the variance and the data points are closer to the regression line.

**Using t-value and p-value (1% significance level)**

The test statistic t=6.841

The 1% critical two-tailed t-values are two values with 45 degrees of freedom, which are 2.689585 and -2.689585. 6.841>2.689585, **we reject , and conclude that is different from zero.**

p-value=1.76e-08<1%, the p-value is less than the significance level, so **we reject the null hypothesis.**

The test statistic t=4.887

The 1% critical two-tailed t-values are two values with 45 degrees of freedom, which are 2.689585 and -2.689585. 4.887>2.689585, **we reject , and conclude that is different from zero.**

p-value=1.34e-05<1%, the p-value is less than the significance level, so **we reject the null hypothesis.**

The test statistic t= -9.251

The 1% critical two-tailed t-values are two values with 45 degrees of freedom, which are 2.689585 and -2.689585. -9.251<-2.689585, **we reject , and conclude that is different from zero.**

p-value=5.62e-12<1%, the p-value is less than the significance level, so **we reject the null hypothesis.**

The test statistic t=0.953

The 1% critical two-tailed t-values are two values with 45 degrees of freedom, which are 2.689585 and -2.689585. -2.689585<0.953<2.689585, **we fail to reject , and conclude that is not different from zero.**

p-value=0.34570>1%, the p-value is greater than the significance level, so **we fail to reject the null hypothesis.**

The test statistic t=3.140

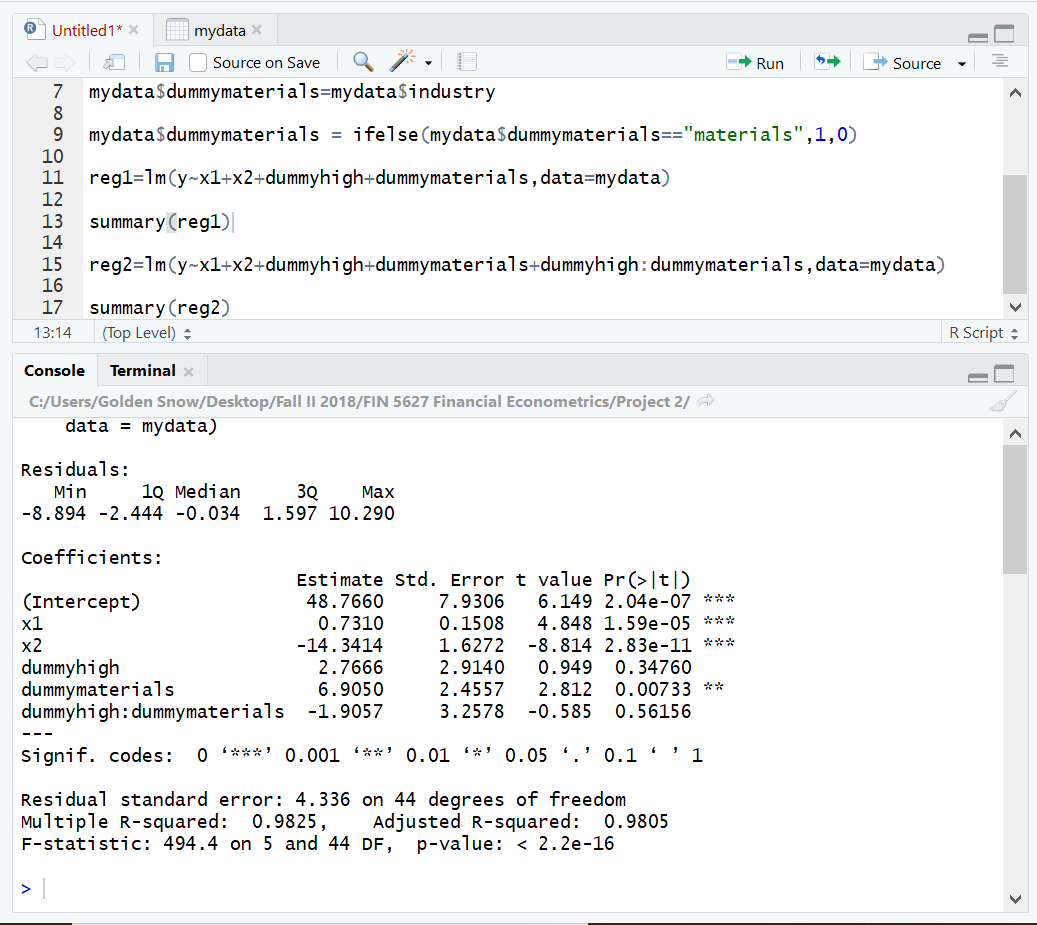
The 1% critical two-tailed t-values are two values with 45 degrees of freedom, which are 2.689585 and -2.689585. 3.140>2.689585, **we reject , and conclude that is different from zero.**

p-value=0.00298<1%, the p-value is less than the significance level, so **we reject the null hypothesis.**

**Since is not different from zero, the high revenue firms’ stock prices are not different from low revenue firms.**

**Since is different from zero, the materials firms’ stock prices are different from non-materials firms.**

1. Run the multiple linear regression with dummy variables interaction (i.e. the five independent variables are x1, x2, two newly created dummy variables and their interaction term). Please write the regression equation below.



1. Interpret if the dummy variables interaction exists

The test statistic t= -0.585

The 1% critical two-tailed t-values are two values with 44 degrees of freedom, which are 2.692278 and -2.692278. -2.692278<-0.585<2.689585, **we fail to reject , and conclude that is not different from zero.**

p-value=0.56156>1%, the p-value is greater than the significance level, so **we fail to reject the null hypothesis, and conclude that is not different from zero and the dummy variables interaction doesn’t exist.**