**Financial Econometrics**

**Project #1**

Name: Xue Jin

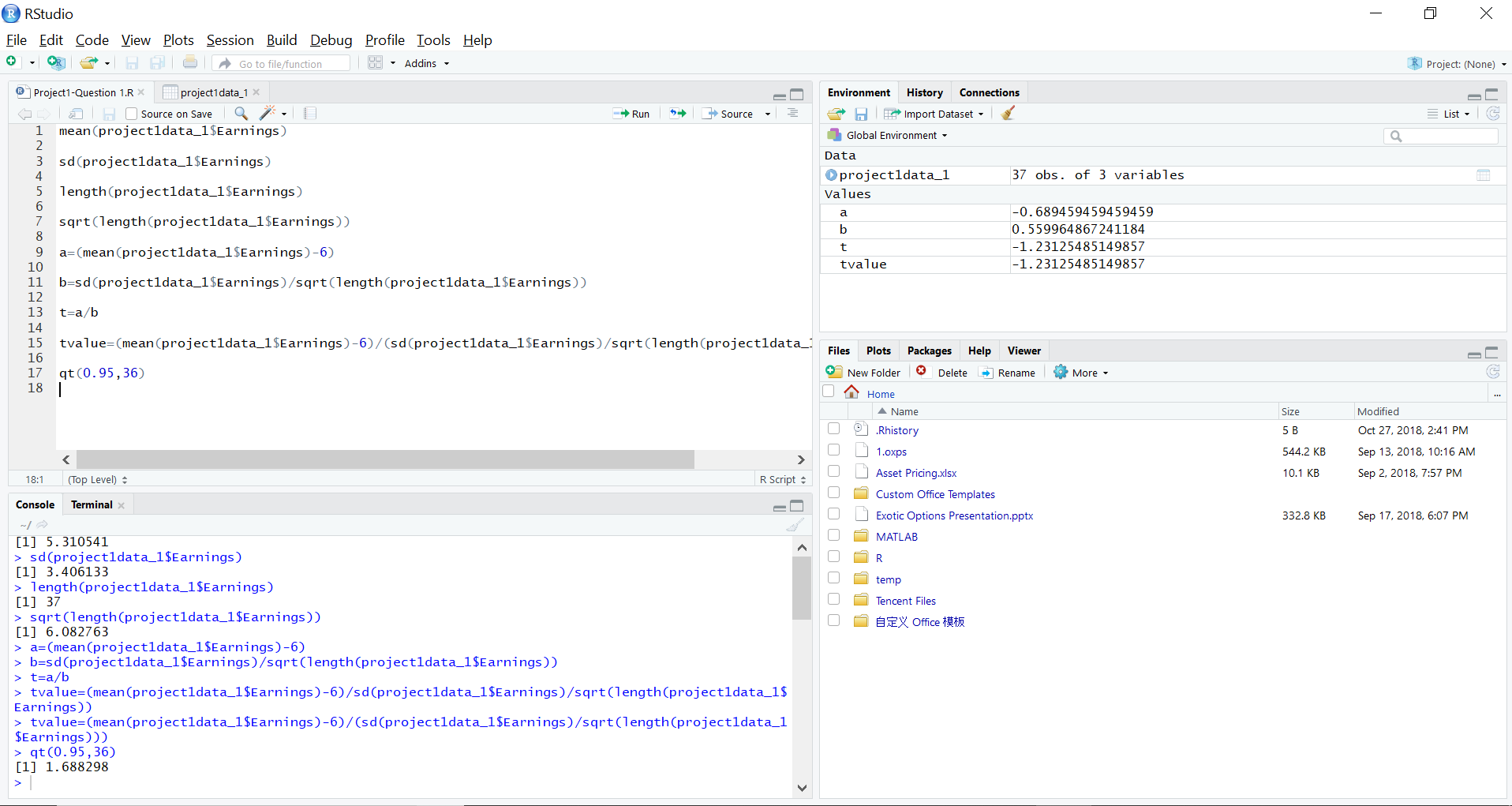
TU ID: 915697511

Notice: please submit the completed Word document though the link on Canvas before the deadline. Please name your document in the format of “Project 1 – Your Name”.

**Question 1:**

The dataset “project1data\_1” has information for a sample of stocks, which are randomly chosen from a large population (i.e. all the stocks in the stock market). Please use approach #1 to test if the population mean of stock earnings is greater than $6 at the 5% significance level.

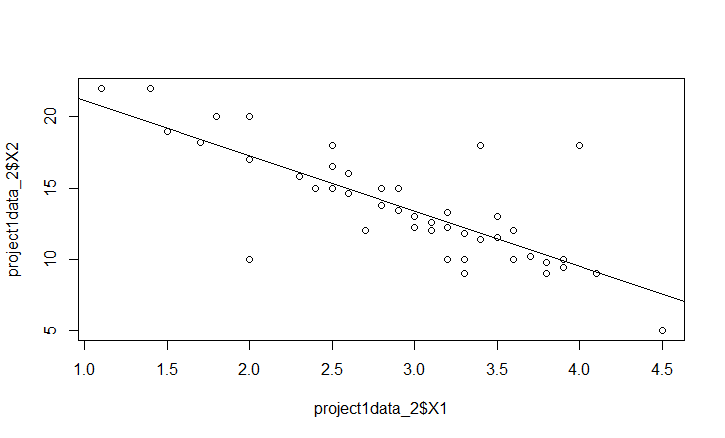
1. The alternative hypothesisis the population mean of stock earnings is greater than $6 at the 5%significance level.
2. The test-statistic is -1.231255
3. The critical value(s) is(are) 1.688298 (can be confirmed on T-chart)
4. We should not reject the null hypothesis, and conclude that the population mean of stock earnings is less than or equal $6 at the 5% significance level.
5. Please screenshot the R code you used for calculation, and insert the image below.



**Question 2:**

Use the dataset “project1data\_2” and a simple linear regression model to analyze how the time spent on watching television shows can affect the quiz score. The data is created for you to practice and to understand how to do econometric analysis by using R. There are two variables in the data set. The first column is the independent variable, which represents the number of hours spent on watching television shows. The second column is the dependent variable, which represents the quiz score. There is no variable name in the data set, so you have to specify the header. The default names for the variables are X1 and X2, which have to be in the form of upper case. Assume the significance level is 5%.

1. Draw the scatter plot and the fitted straight line, and insert the image below



1. The regression equation is X2=-3.88641\*X1+25.04628+2.182
2. Analyze the summary of regression results (i.e. state the estimated value of regression coefficients, determine if the slope coefficient is different from zero by using t-value or P-value, interpret the R-square, and conclude whether and how watching TV show affects quiz score)

Coefficients:

Estimate Std. Error t value

(Intercept) 25.0463 1.2384 20.225

X1 -3.8864 0.4146 -9.374

Slope coefficient t value=

H0: The slope coefficient is zero

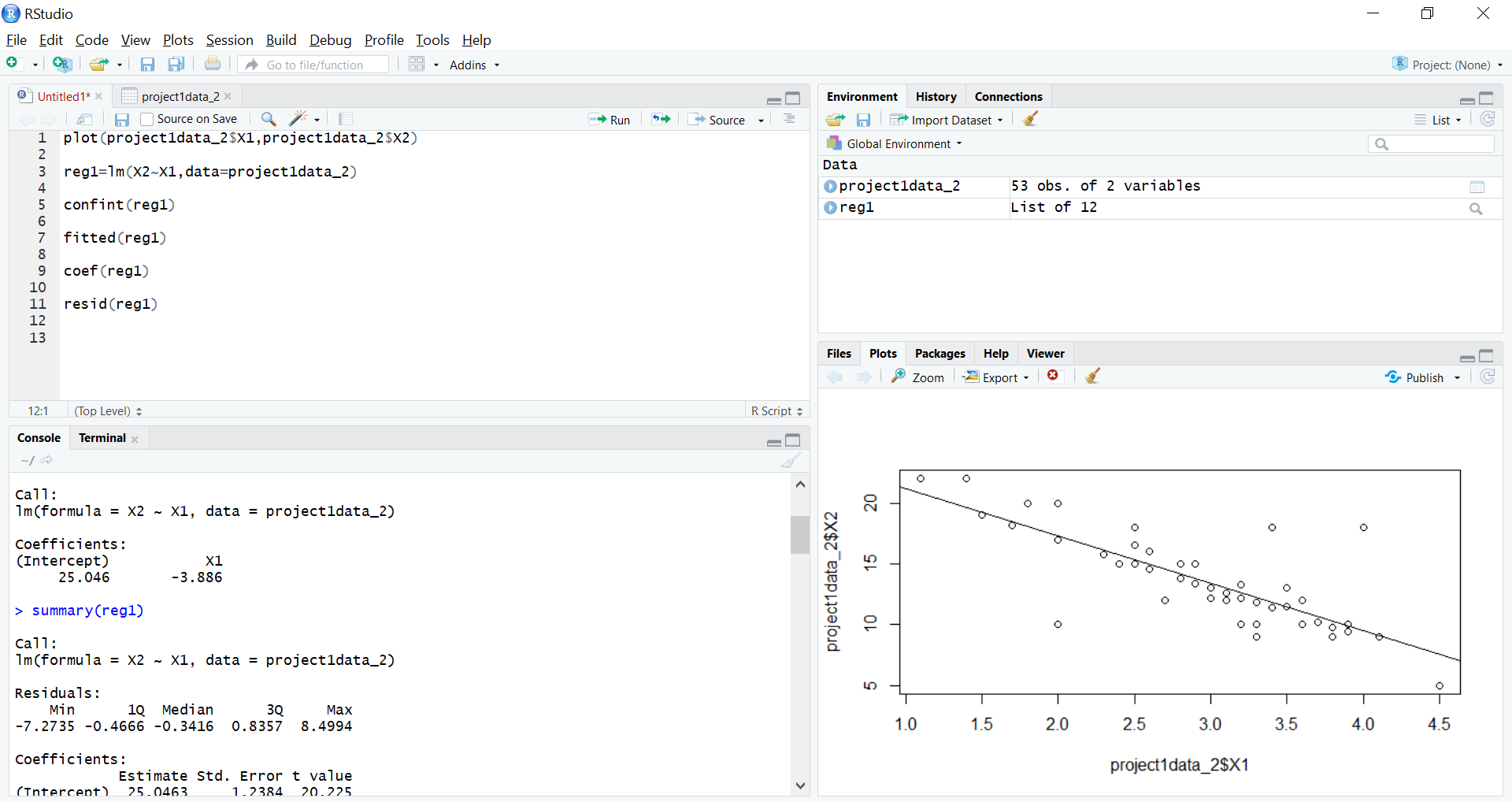
HA: The slope coefficient is different from zero

At 5% significance level, the critical value is . The t-statistic value is -9.374 which is smaller than -2.007584, so we reject the hypothesis. The slope coefficient is different from zero.

Multiple R-squared: 0.6327

The R-squared is 0.6327 which means the 63% variation in X2 is explained by X1. So the quiz score has more than one half percent relation with hours spent on watching television. Thus, watching TV show affects quiz score and this two has negative relations.

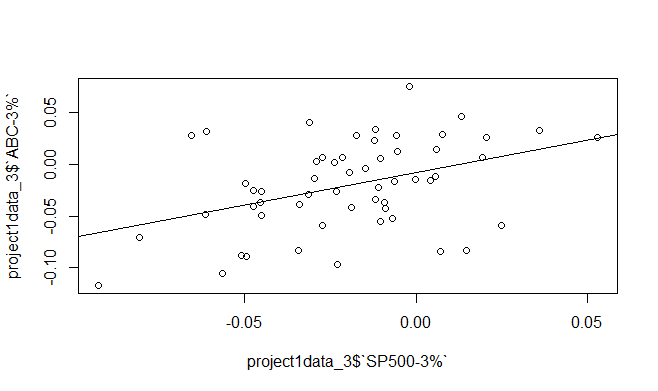
1. Please screenshot the R code you used for calculation, and insert the image below.



**Question 3:**

Use the dataset “project1data\_3” and a simple linear regression model to calculate the beta of stock ABC. The data is created for you to practice and to understand how to do econometric analysis by using R. Please notice that you need to convert the stock price into the stock return. You can refer to the Excel document “CAPM.xlsx” to see the conversion. Assume the significance level is 5%.

1. Draw the scatter plot and the fitted straight line, and insert the image below



X2: ABC-3% X1:SP500-3%

1. The regression equation is X2=0.62909\*X1-0.007879+0.03986
2. Analyze the summary of regression results (i.e. state the estimated value of slope coefficient, determine if the slope coefficient is different from zero by using t-value or P-value, interpret the R-squared, and conclude how does overall stock market affect stock ABC)

Coefficient

Estimate Std. Error t value

`SP500-3%` 0.629090 0.189056 3.328

Slope coefficient t value=

H0: The slope coefficient is zero

HA: The slope coefficient is different from zero

At 5% significance level, the critical value is . The t-statistic value is 3.328 which is greater than 2.004879, so we reject the hypothesis. The slope coefficient is different from zero.

Multiple R-squared: 0.1702

The R-squared is 0.1702 which means the 17.02% variation in X2 is explained by X1. So when X1 changes, X2 changes in a little range with positive relation.

1. Please screenshot the R code you used for calculation, and insert the image below.

