**Data-Analyse (EBE)**

E.J. Roos

Daan van Noort 2001786

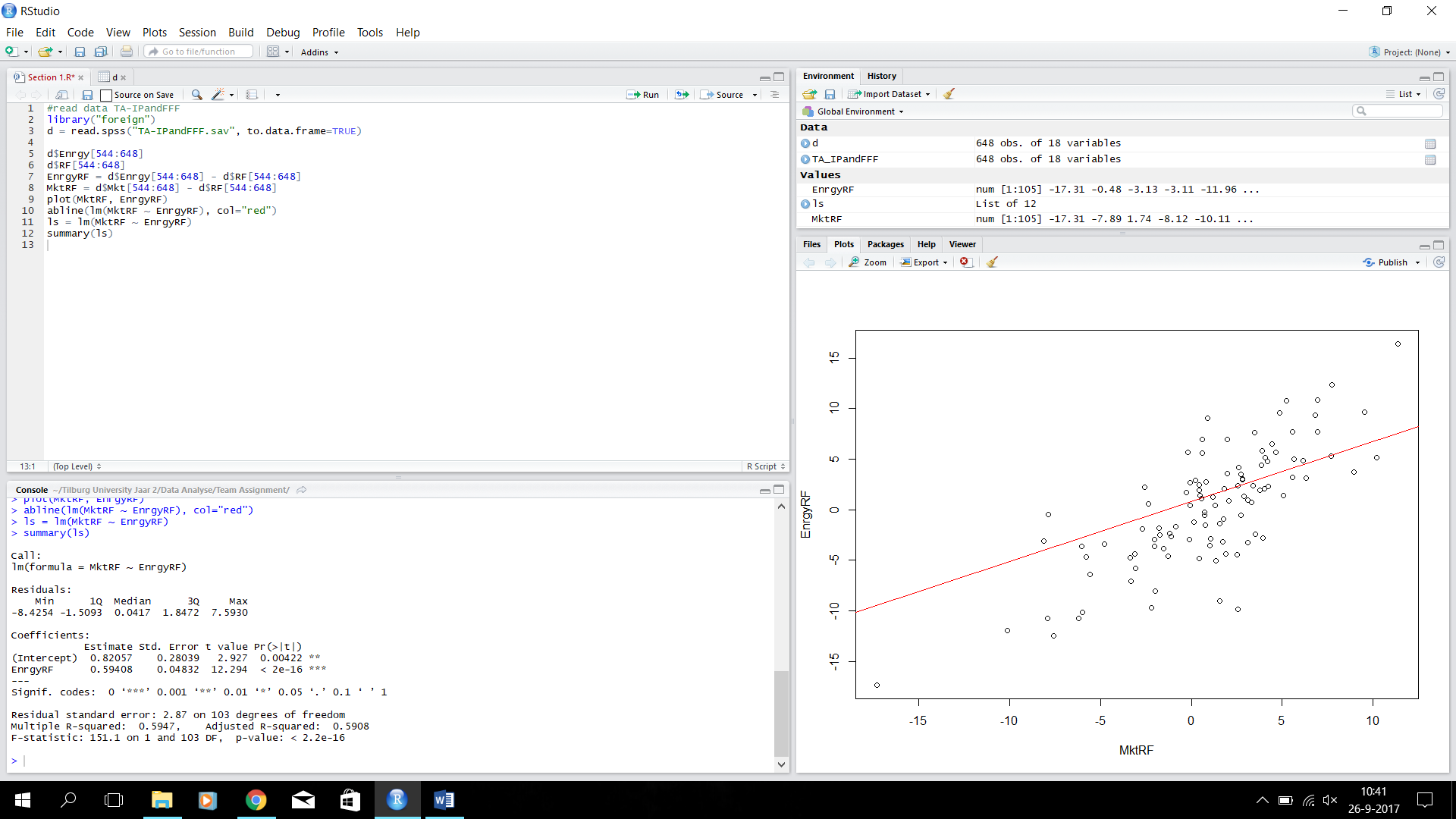
Yannick Kiffen 2001237

Bryan Hellings 2001393

Lianne van Krieken 2002733

Energy 4

**Data Analyse Team Assignment**

***Section 1:***

1. As you can see in our graph on the right, you can say that the linear market model suits our data set. It’s clear that there is some sort of linear correlation between EnrgyRF and MktRF. Therefore we came to the conclusion that we can use the linear market model for our data.
2. We made a red regression line in the plot using the following command:

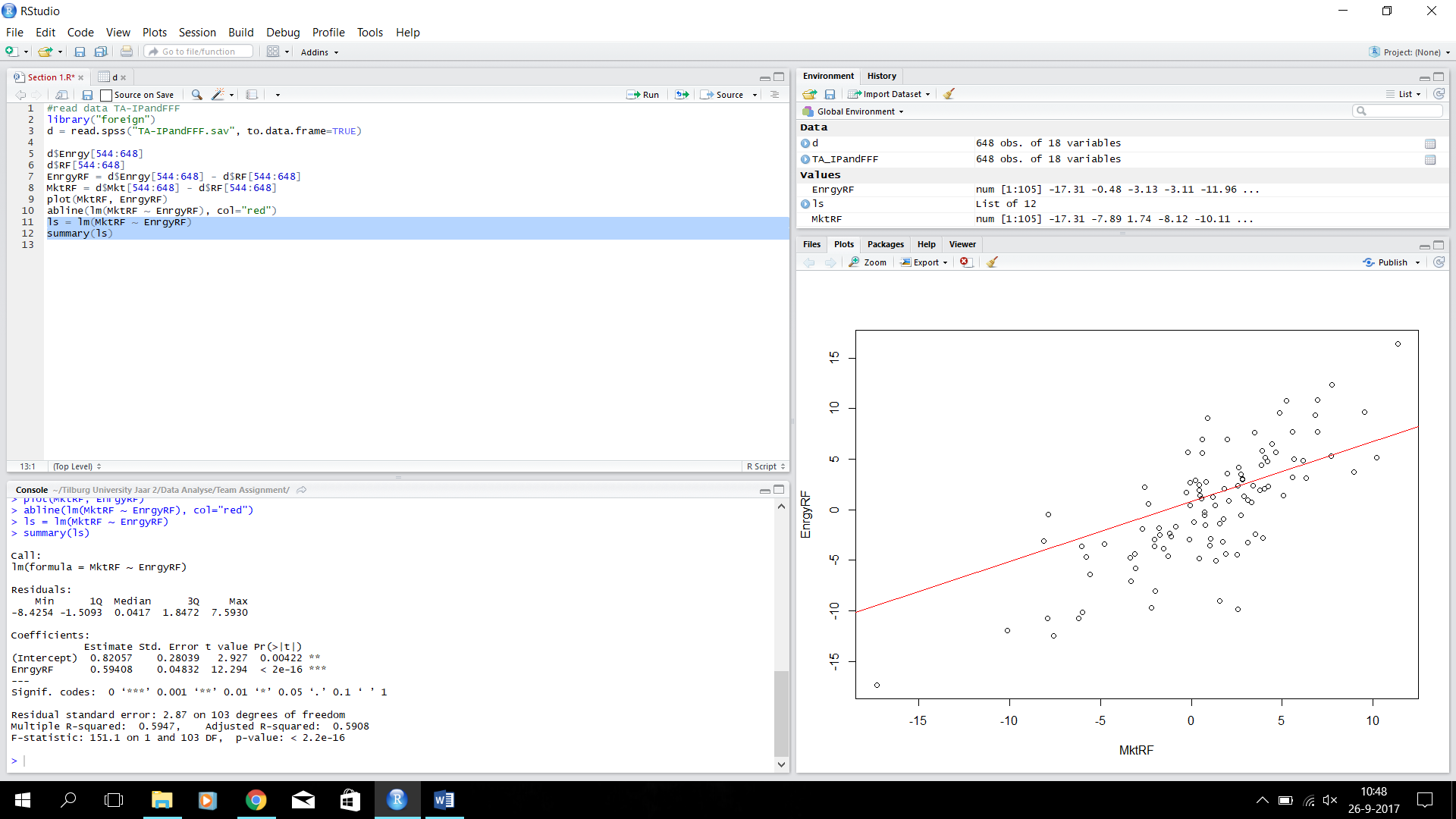
abline(lm(MktRF~EnrgyRF),col="red")

Now that we have the regression line, we estimated an equation that would fit this line. We used the data that was presented in the output after forming these commands:

ls = lm(MktRF ~ EnrgyRF)

summary(ls)

This was the output:



With this we could form this regression line equation: *EnrgyRF = 0.82057 + 0.59408 \* MktRF*

1. R shows us what the p-value is, and our α = 0.10. The p-value = 2.2 \* 10-16. This means that:

P-value < α so we can reject H0.

1. We used the t-test command to get an 90% confidence interval:

t.test(x=MktRF, mu=0.82057, conf=0.90)

This gave us an output:

data: MktRF

t = 0.37503, df = 104, p-value = 0.7084

alternative hypothesis: true mean is not equal to 0.82057

90 percent confidence interval:

0.258150 1.711374

sample estimates:

mean of x

0.9847619

As you can see, our interval is (0.258150, 1.711374). Our estimate (=0.82057) lies in this interval, so we can be at least 90% certain that our estimate is correct.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **H0** | **H1** | **Type Test** | **Val** | **RR** | **Conclusion** |
| **Exposure Market > 1?** | Β1 < 1 | Β1 > 1 | T-Test | -8.401 | >1.289825 | -8.401 < 1.289825, so we do not reject H0­, which means that the exposure to the market is not larger than 1 |

1. T = (B1-1)/SB1

Val = (0.59408-1)/0.04832 = -8.400662252

Reject H0 if t > t0.10, 103

We calculated the p-value using a command in R:

t.test(x=EnrgyRF, mu=0.59408, conf=0.90)

This gave us the following output:

data: EnrgyRF

t = -0.559, df = 104, p-value = 0.5774

alternative hypothesis: true mean is not equal to 0.59408

90 percent confidence interval:

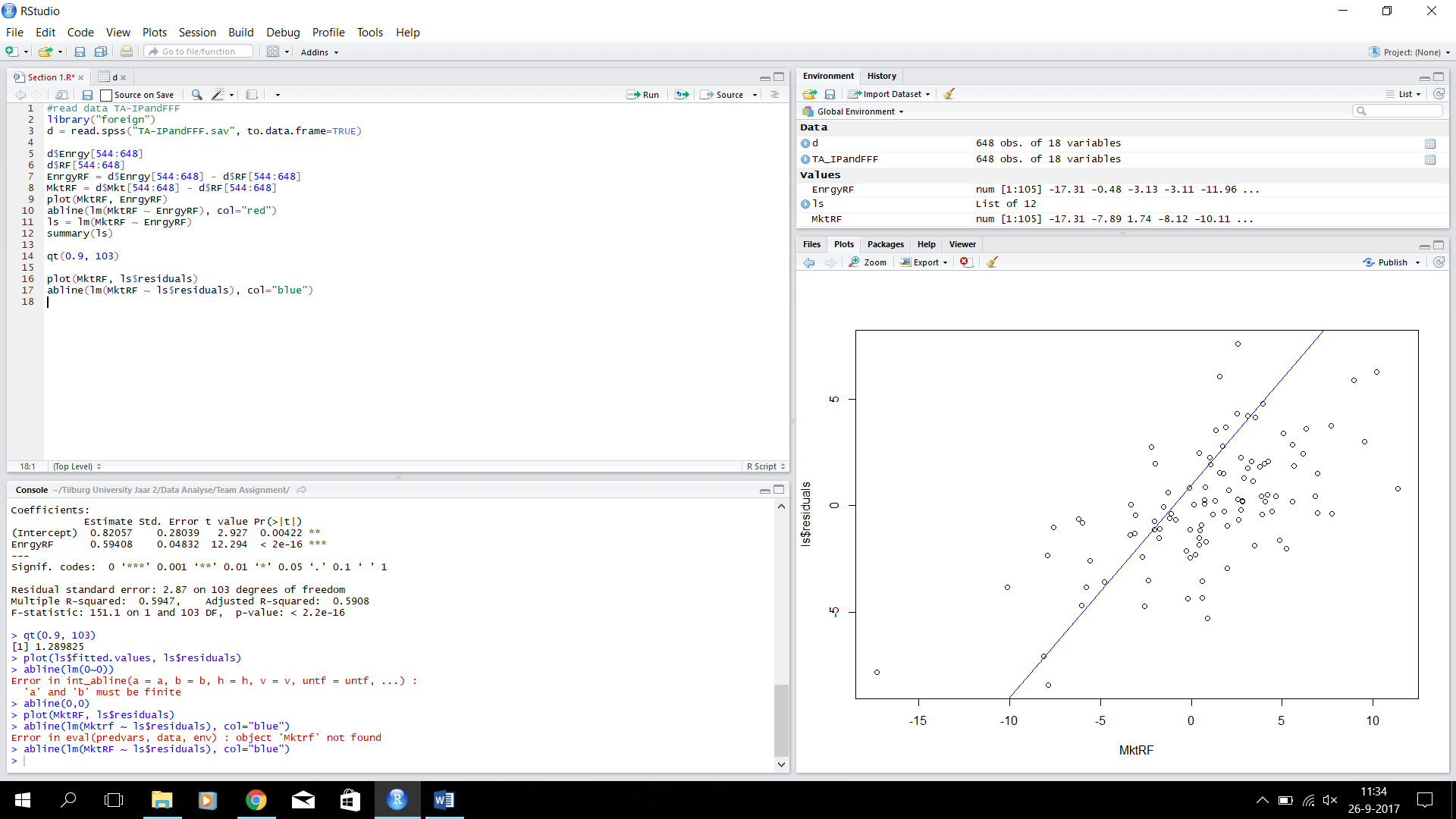
-0.6668448 1.2196067

sample estimates:

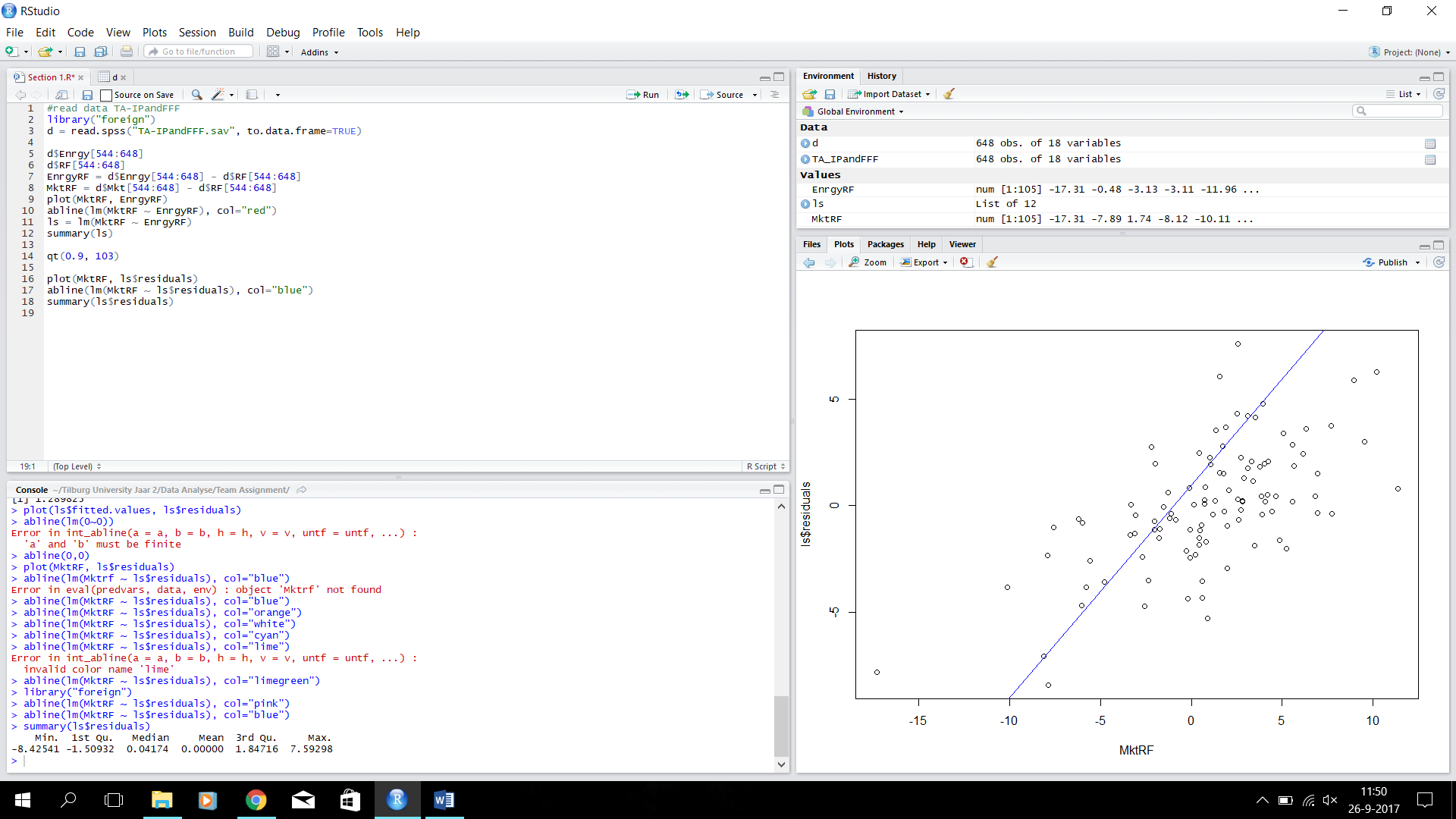
mean of x

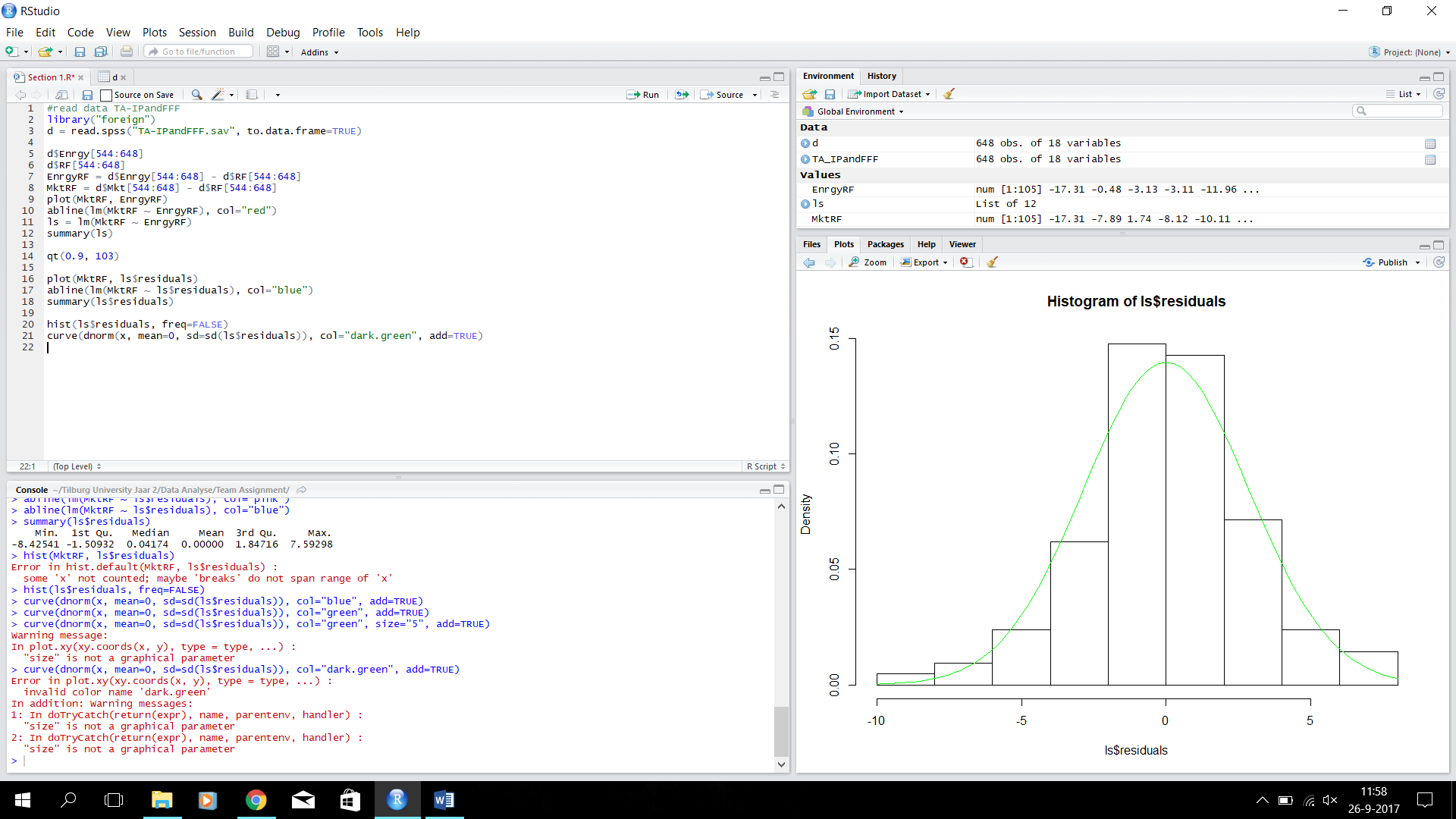
0.276381

So we can see that our p-value = 0.5774. This is bigger than alfa = 0.10, so we do not reject H0.



1. Linearity: our plot has a linear form to it. Because the mean of the residuals is equal to zero, as you can see in our output.



Homoskedasticity: in our plot the vertical variation of the residual points is more or less constant.

Normality: as you can see in our histogram on the right, the residuals have a normal distribution.