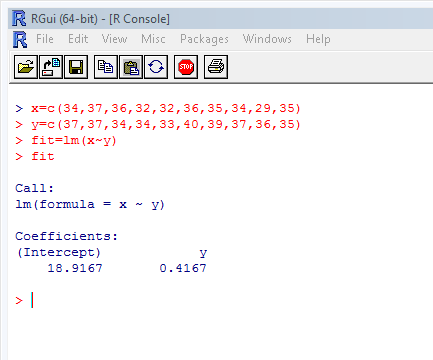
Statistics for Engineers (LAB) L13+L14  
Regression

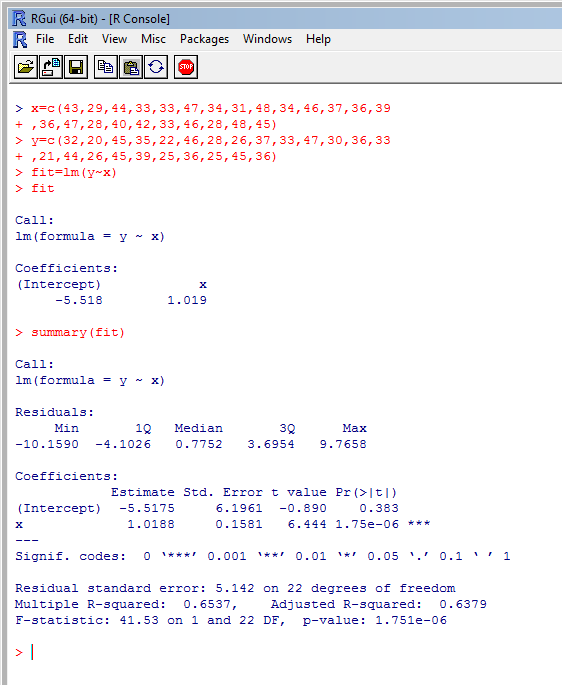
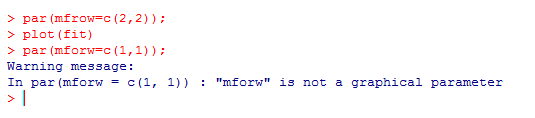
Prashant Singhai

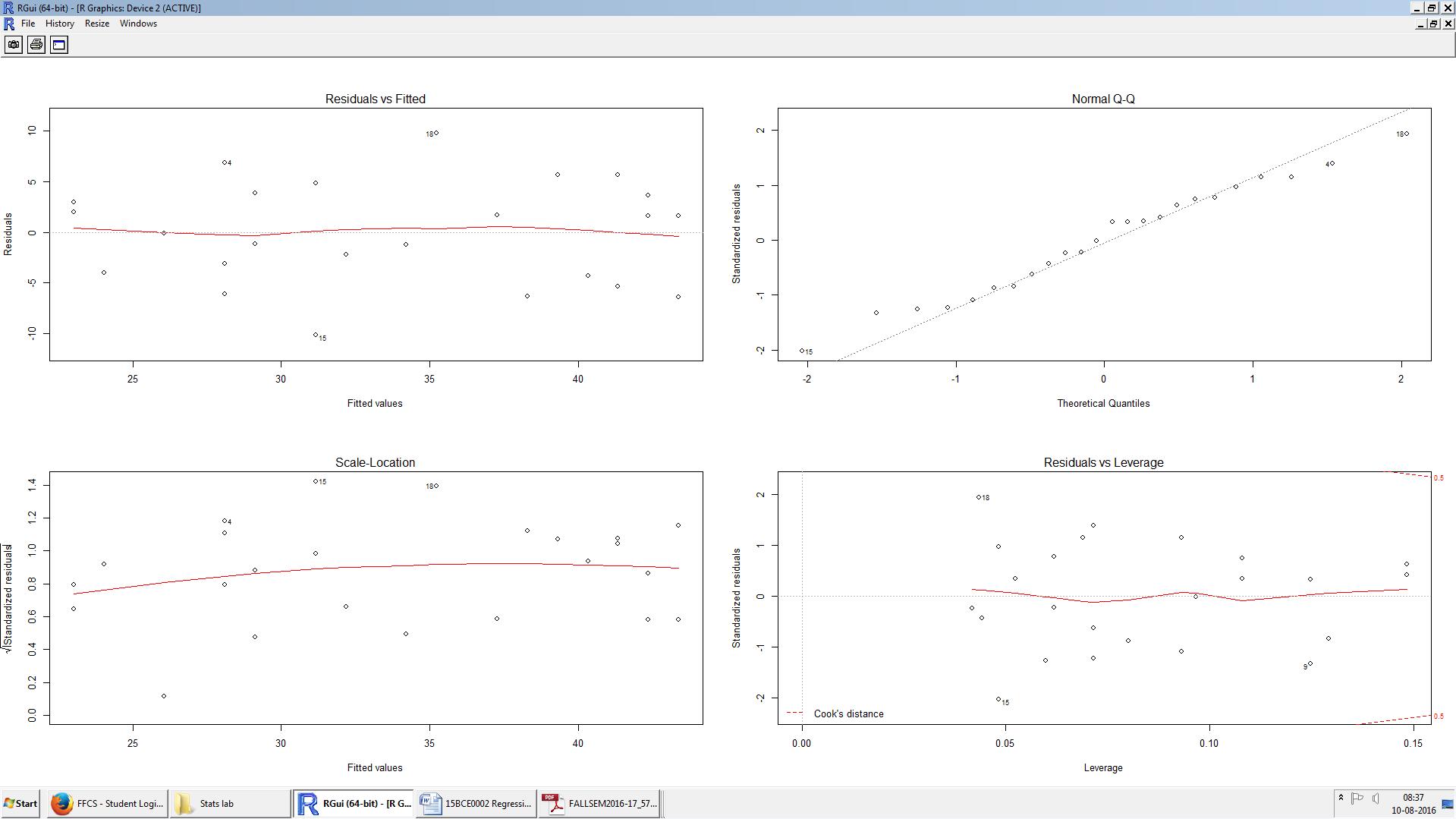
15BCE0531

Problem 1

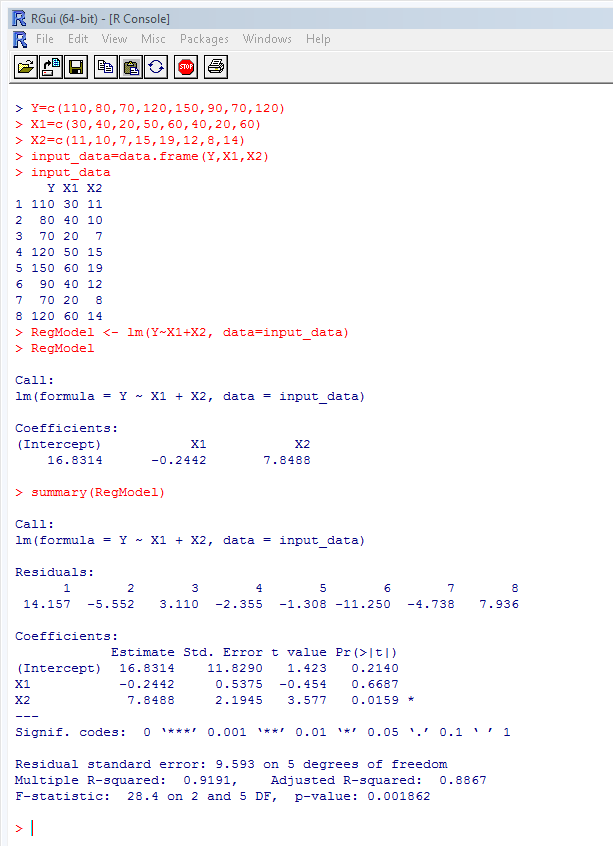


Problem 2





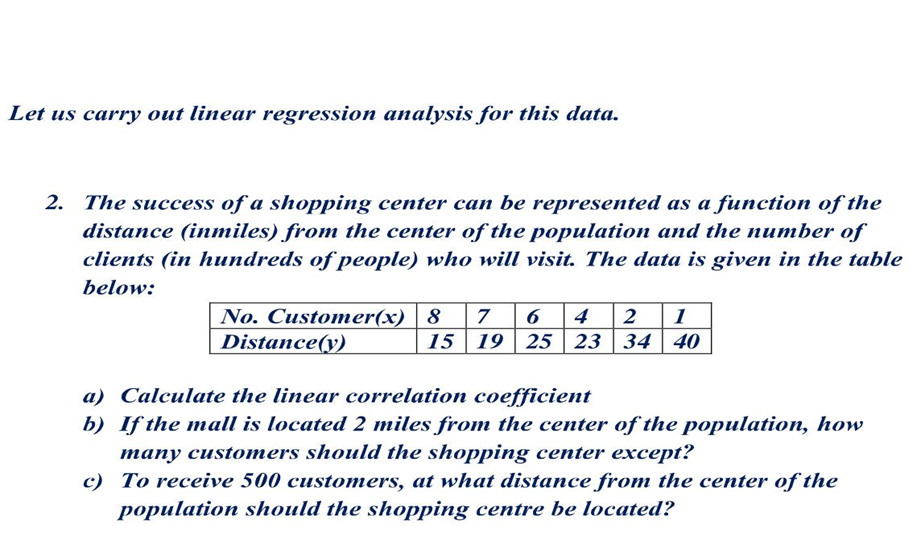
Problem 3



Problem 4



**Challenging Task:**

  
  
  
  
Solution:  
  
  
  
  
>distance=c(15,19,25,23,34,40)

* regnode<-lm(cust\_no~distance)
* regnode

Call:

lm(formula = cust\_no ~ distance)

Coefficients: (Intercept) distance

12.0530 -0.2841

* regnode2<-lm(cust\_no~distance)
* regnode2<-lm(distance~cust\_no)
* regnode2

Call:

lm(formula = distance ~ cust\_no)

Coefficients: (Intercept) cust\_no

40.831 -3.178

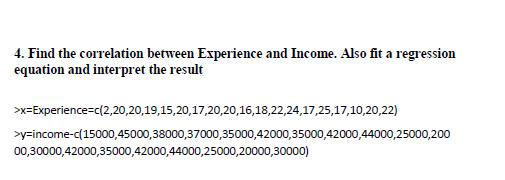
* cust\_no1=-0.2841\*2+12.0530
* cust\_no1

[1] 11.4848

> distance1=-3.178\*5+40.831

* distance1 [1] 24.941
* correlation=sqrt(0.2841\*3.178)
* correlation

[1] 0.9501946



**Answer:**

**Code:**

* data=read.csv("D:\\MAT\\lab\\exp&income.csv")
* data

Experience Income

1. 2 15000
2. 20 45000
3. 20 38000
4. 19 37000
5. 15 35000
6. 20 42000
7. 17 35000
8. 20 42000
9. 20 44000
10. 16 25000
11. 18 20000
12. 22 30000
13. 24 42000
14. 17 35000
15. 25 42000
16. 17 44000
17. 10 25000
18. 20 20000
19. 22 30000

* exp=c(2,20,20,19,15,20,17,20,20,16,18,22,24,17,25,17,10,20,22)

income=c(15000,45000,38000,37000,35000,42000,35000,42000,44000,25000,20000,30000,42000,

35000,42000,44000,25000,20000,30000)

>cor(income,exp) [1] 0.5899288

* model<-lm(income~exp)
* model

Call:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| lm(formula = income ~ exp) | | | |  |
| Coefficients: | |  |  |  |
| (Intercept) |  | exp |  |  |
| 14750 |  | 1063 |  |  |
| >summary.lm(model) | | | |  |
| Call: |  |  |  |  |
| lm(formula = income ~ exp) | | | |  |
| Residuals: |  |  |  |  |
| Min 1Q Median | | | 3Q | Max |
| -16014 -4319 | | 1986 | 5144 11175 | |
| Coefficients: | |  |  |  |
| Estimate Std. Error t value Pr(>|t|) | | | | |
| (Intercept) | 14750 | | 6631 | 2.224 0.03996 \* |
| exp | 1063 | | 353 3.012 0.00785 \*\* | |
| --- |  |  |  |  |
| Signif. codes: | | 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 | | |
| Residual standard error: 7715 on 17 degrees of freedom | | | | |
| Multiple R-squared: 0.348, | | | | Adjusted R-squared: 0.3097 |

F-statistic: 9.074 on 1 and 17 DF, p-value: 0.007845

**Interpretation:-**

The regression equation is income=14750+1063\* exp. Since the p- value of the test is 3.012 is greater than 0.05 we reject the hypothesis. Therefore the model we found is significant. The Multiple R-squared is the coefficient of determination. It provides a measure of how well future outcomes are likely to be predicted by the model. In this case the R square value is 0.348. Therefore 34.80% of data is well predicted.