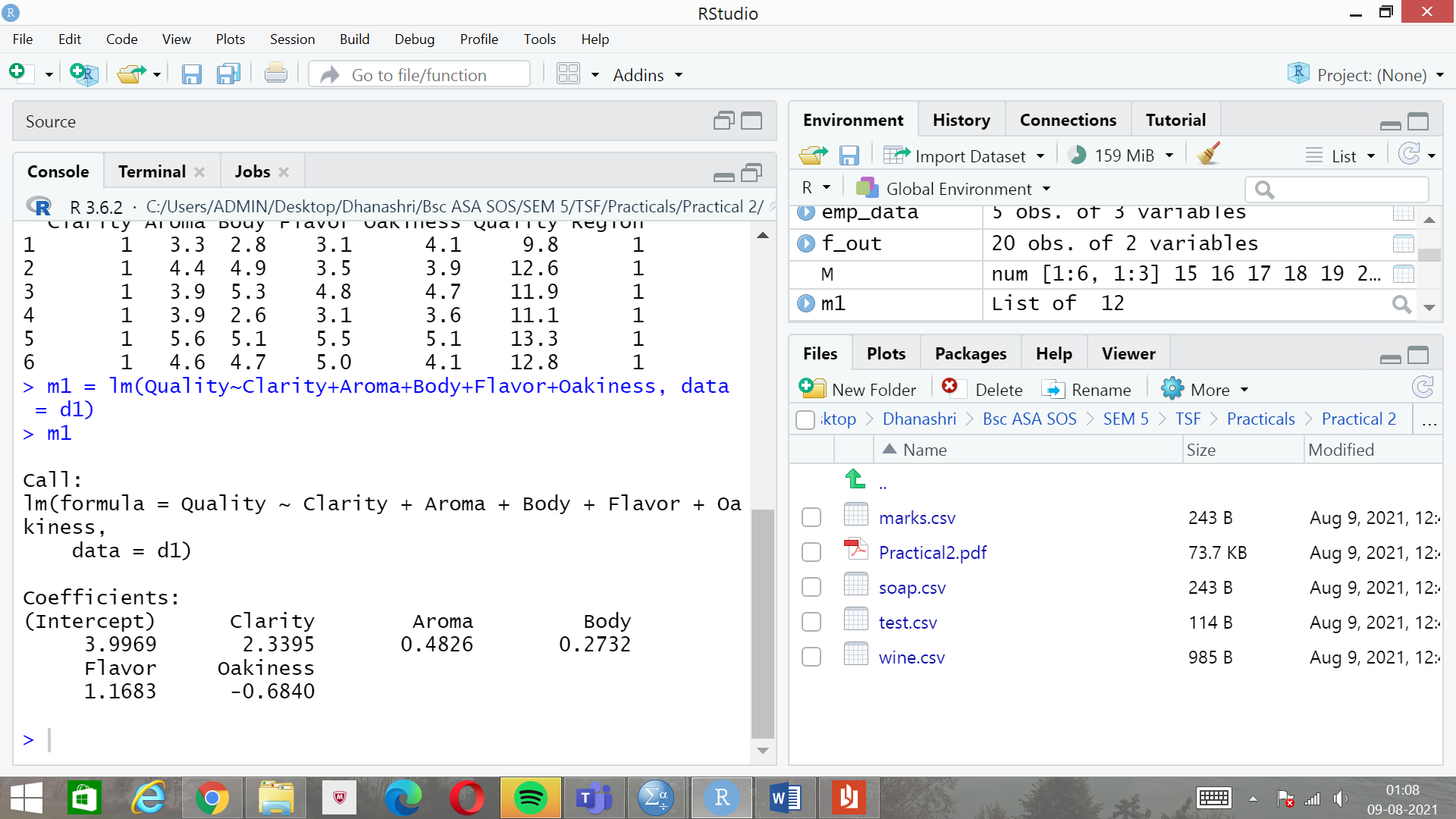
Q1)

1. Multiple Linear Regression

d1 = read.csv('wine.csv')

head(d1)

m1 = lm(Quality~Clarity+Aroma+Body+Flavor+Oakiness, data = d1)



1. Testing Overall Regression:

#Testing overall regression

#m2 = lm(Quality~ .,data = d1 )

A1=anova(m1)

SS=A1$'Sum Sq'

SSR=sum(SS[1:5])

MSR=SSR/5

MSE=A1$'Mean Sq'[6]

F1=MSR/MSE

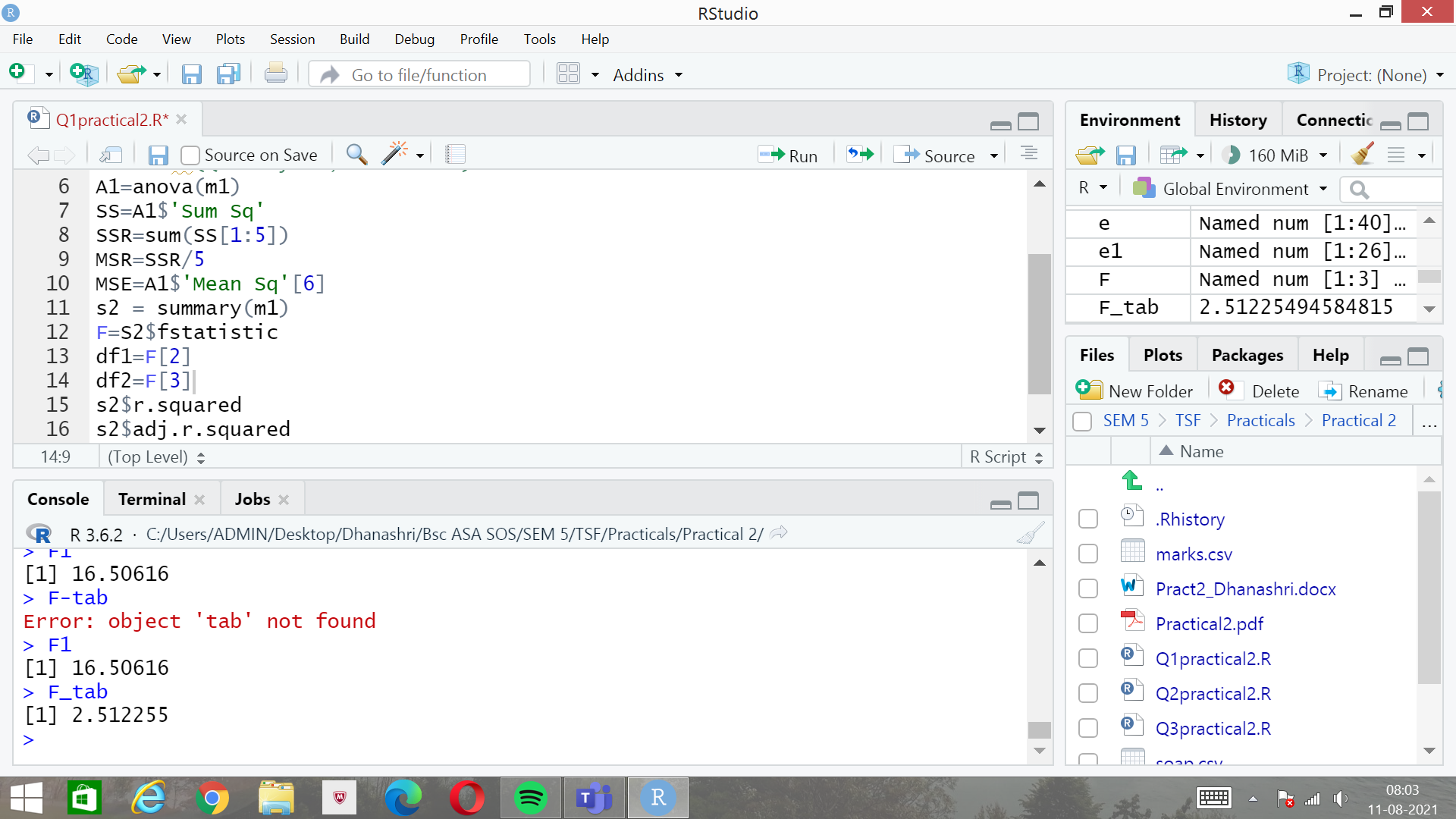
s2 = summary(m1)

F=S2$fstatistic

df1=F[2]

df2=F[3]

F\_tab=qf(0.95,df1,df2)



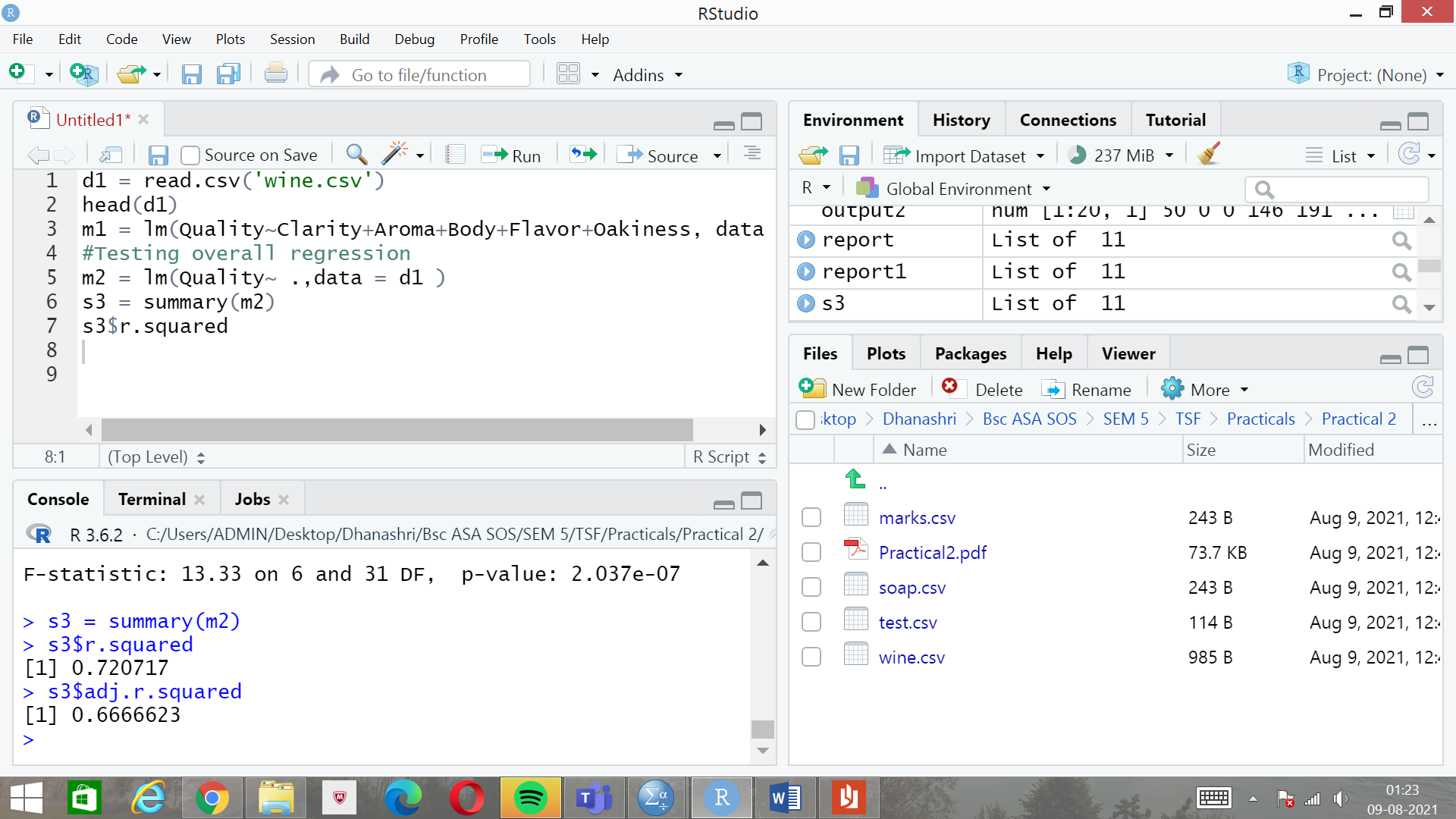
As the F-calculated is greater than F-tabulated, we reject the null hypothesis.

1. R-squared and adjusted R-squared:

s2 = summary(m1)

s2$r.squared

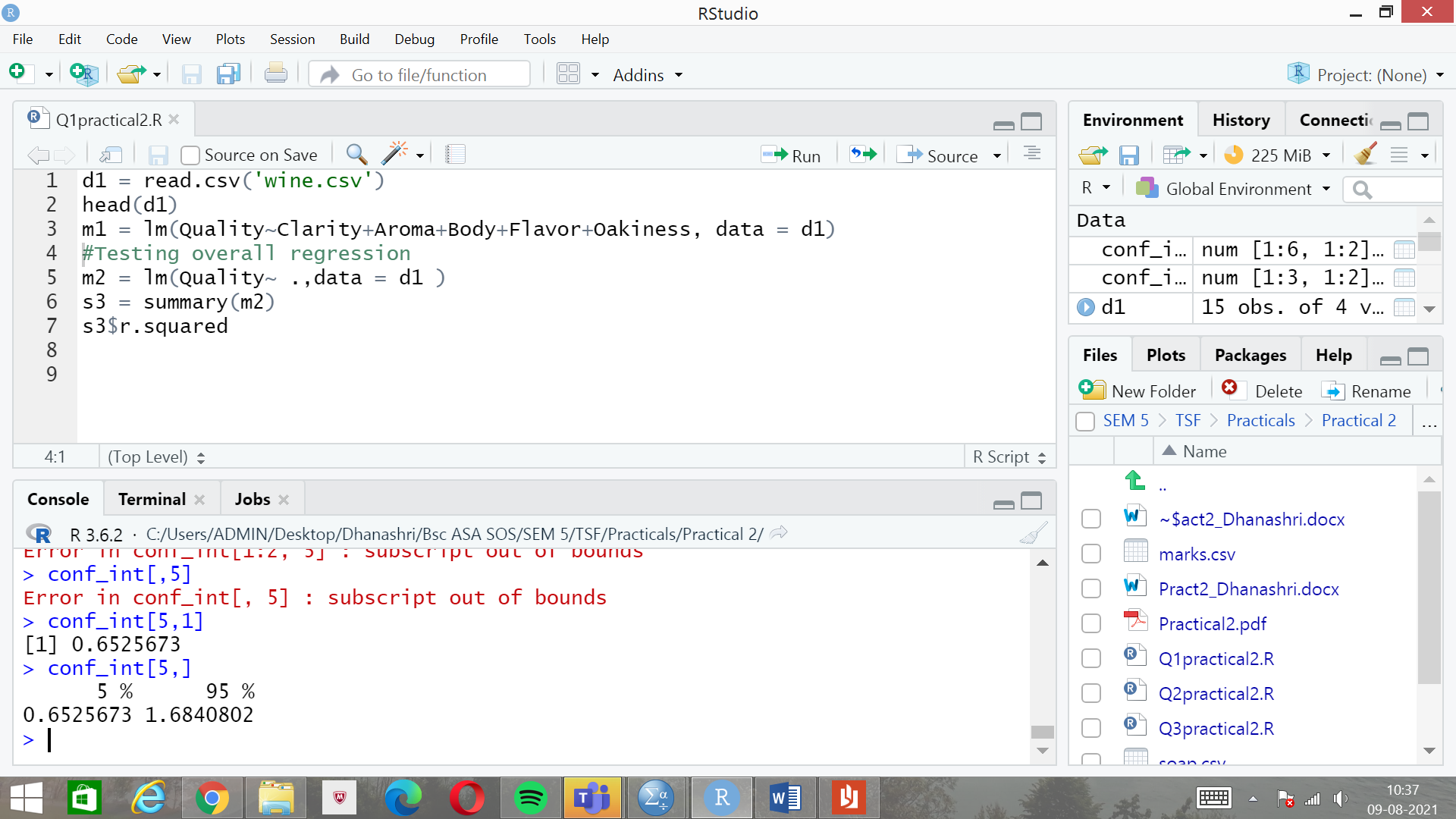
s2$adj.r.squared



Variation explained by R-squared is 66.66%.

1. 90% confidence interval of coefficients of flavour:

conf\_int = confint(m1,level = 0.90)



Q2)

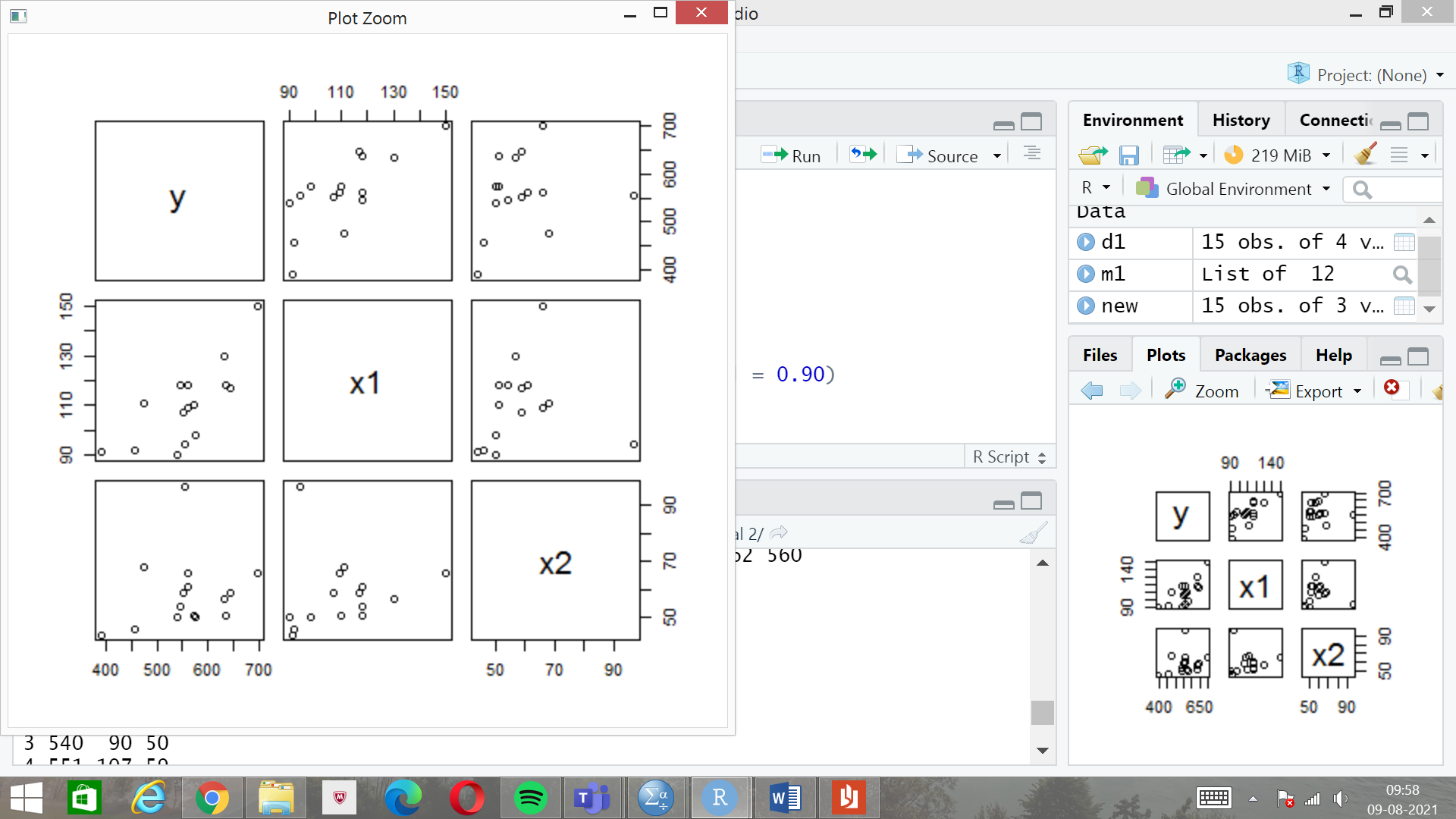
1. Multiple Scatterplot:

d2 = read.csv('soap.csv')

head(d2)

new = d1[,-1]

plot(new)



1. Multiple Linear Regression:

m3 = lm(y~x1+x2, data = new)

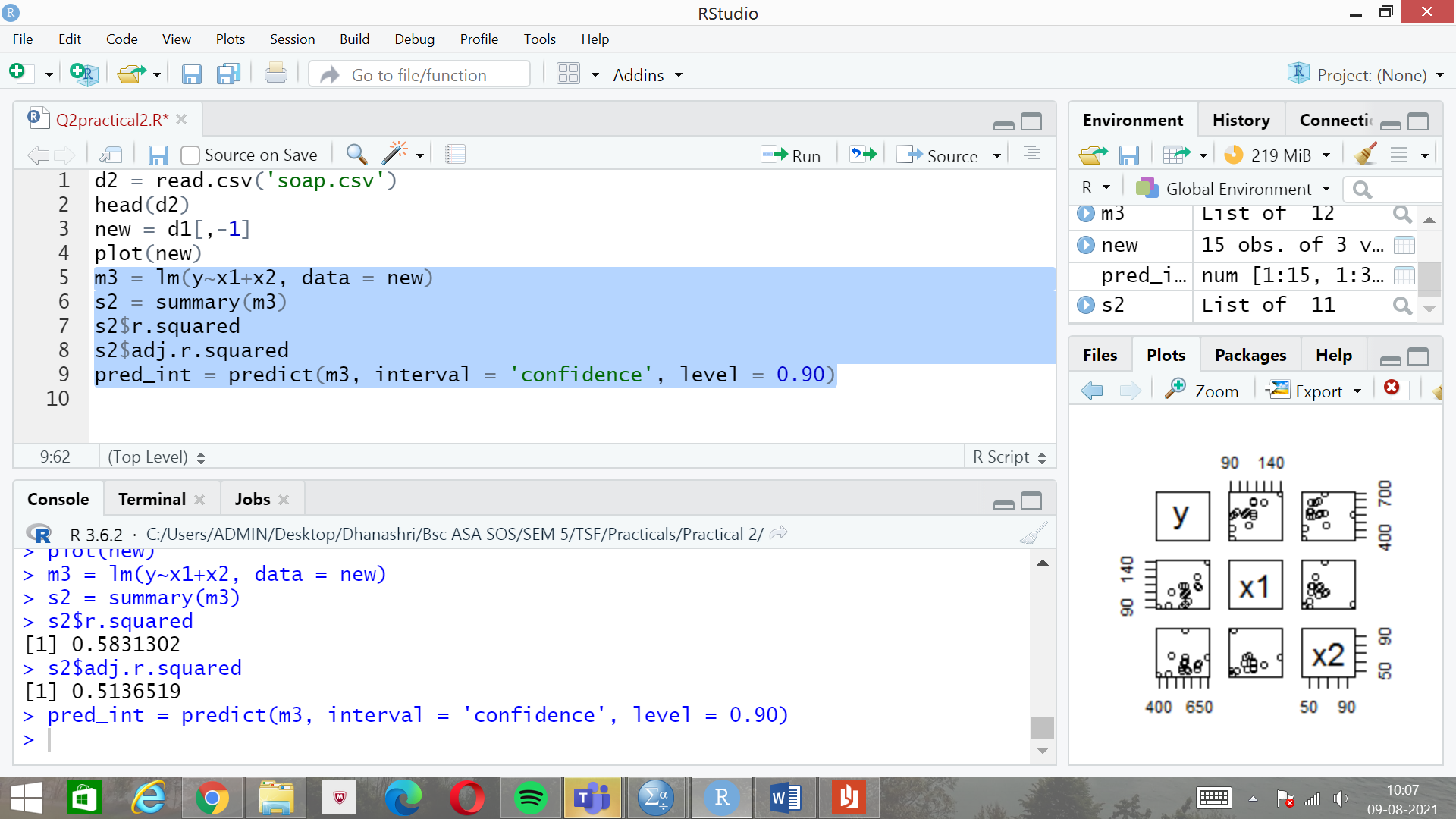


1. R-squared and Adjusted R-squared:

s2 = summary(m3)

s2$r.squared

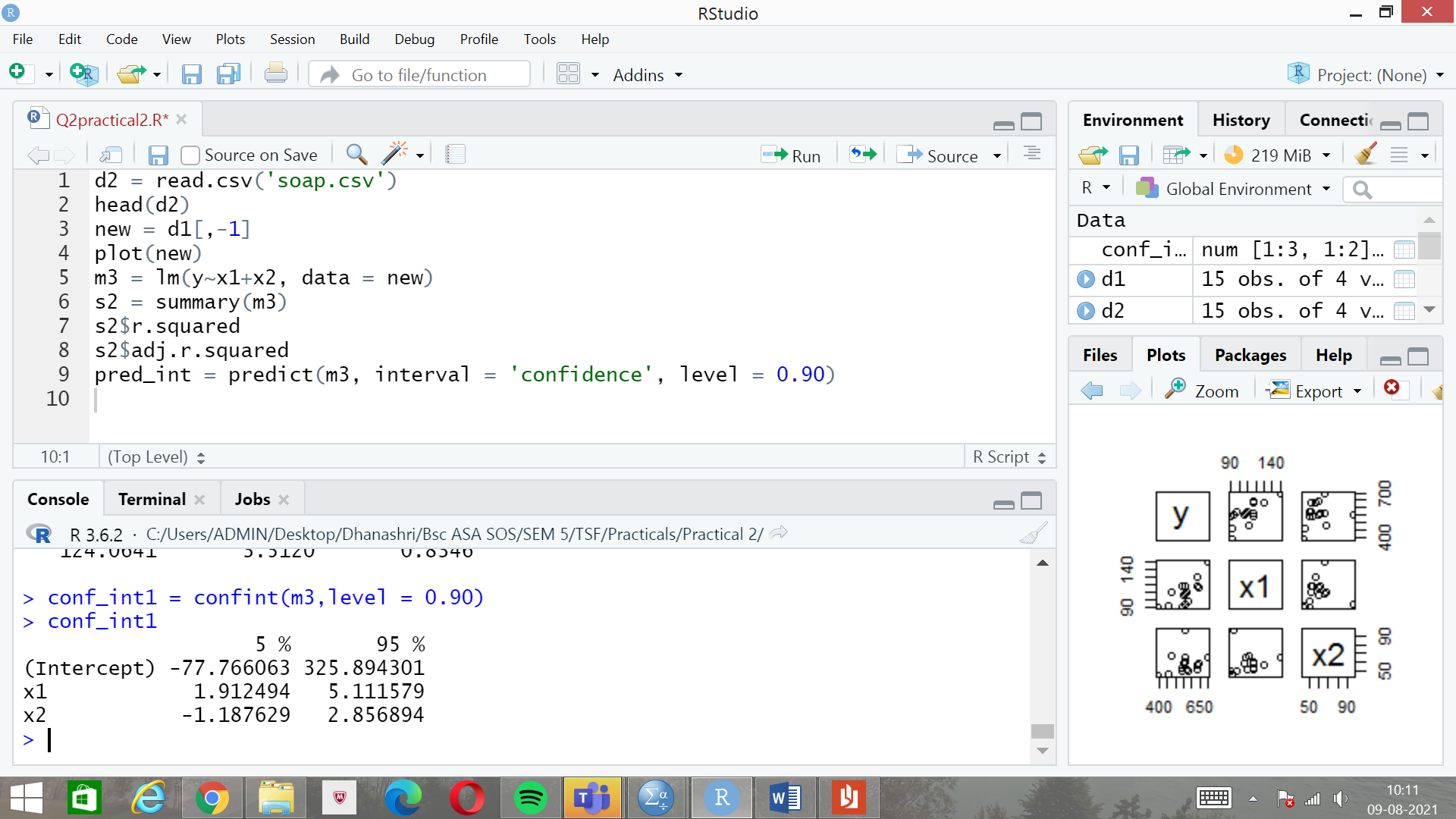
s2$adj.r.squared



Variation explained by R-squared is 51%

1. Confidence interval for the data:

conf\_int = confint(m3,interval = 0.90)

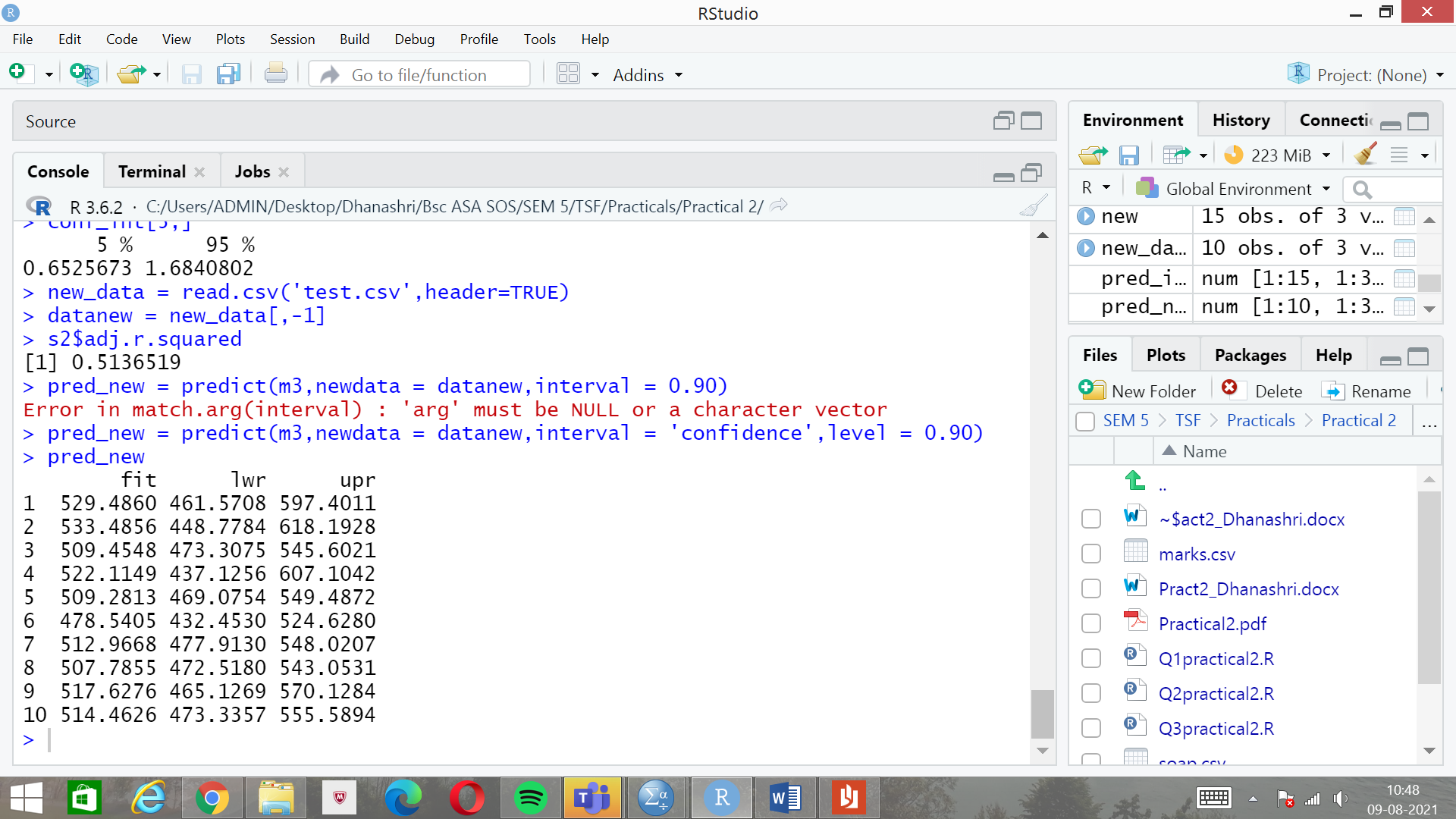


1. Predicting confidence interval:

new\_data = read.csv('test.csv',header=TRUE)

datanew = new\_data[,-1]

pred\_new = predict(m3,newdata = datanew,interval = 'confidence',level = 0.90)



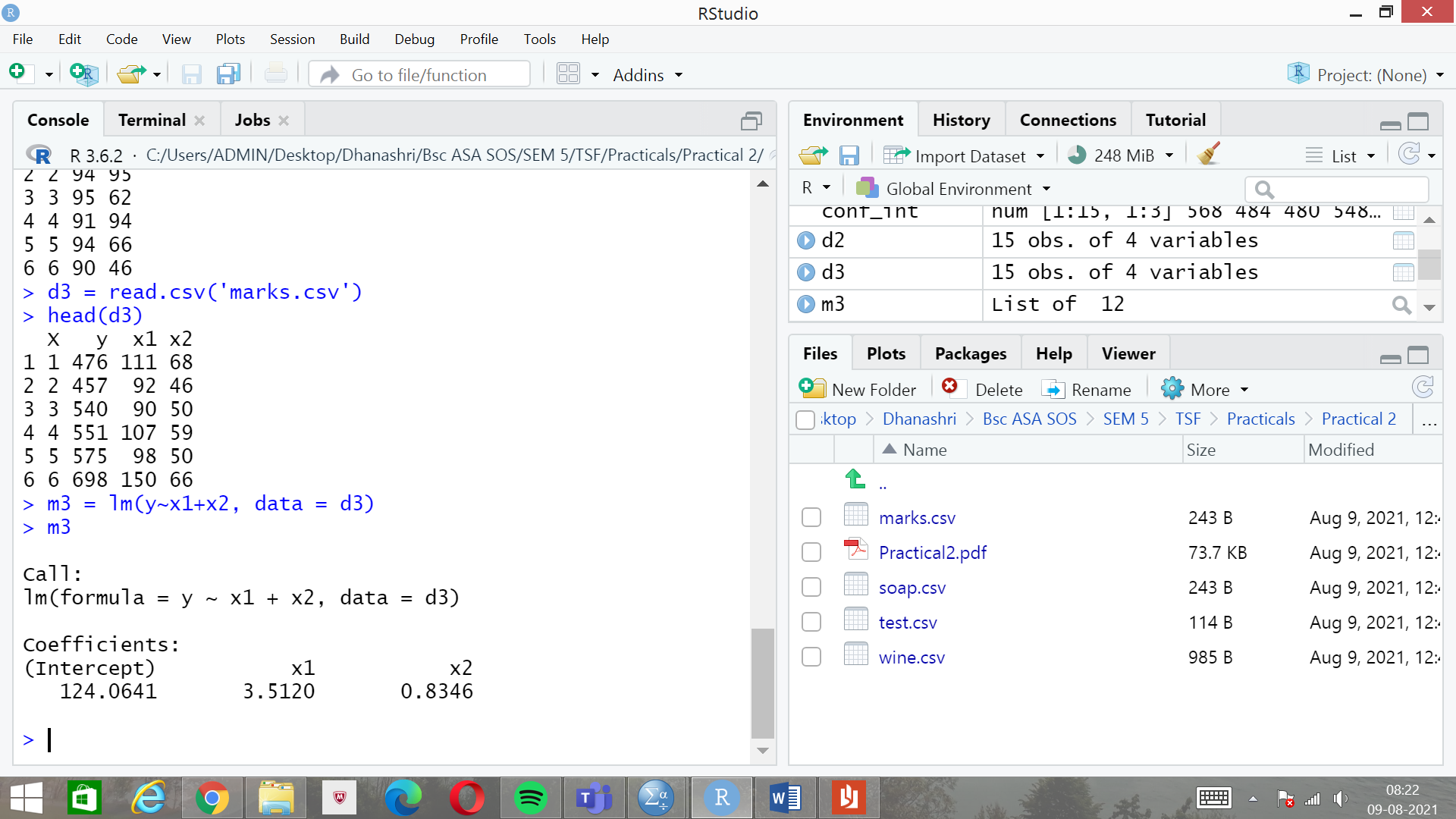
Q3)

1. Multiple Linear Regression:

d3 = read.csv('marks.csv')

head(d3)

m3 = lm(y~x1+x2, data = d3)

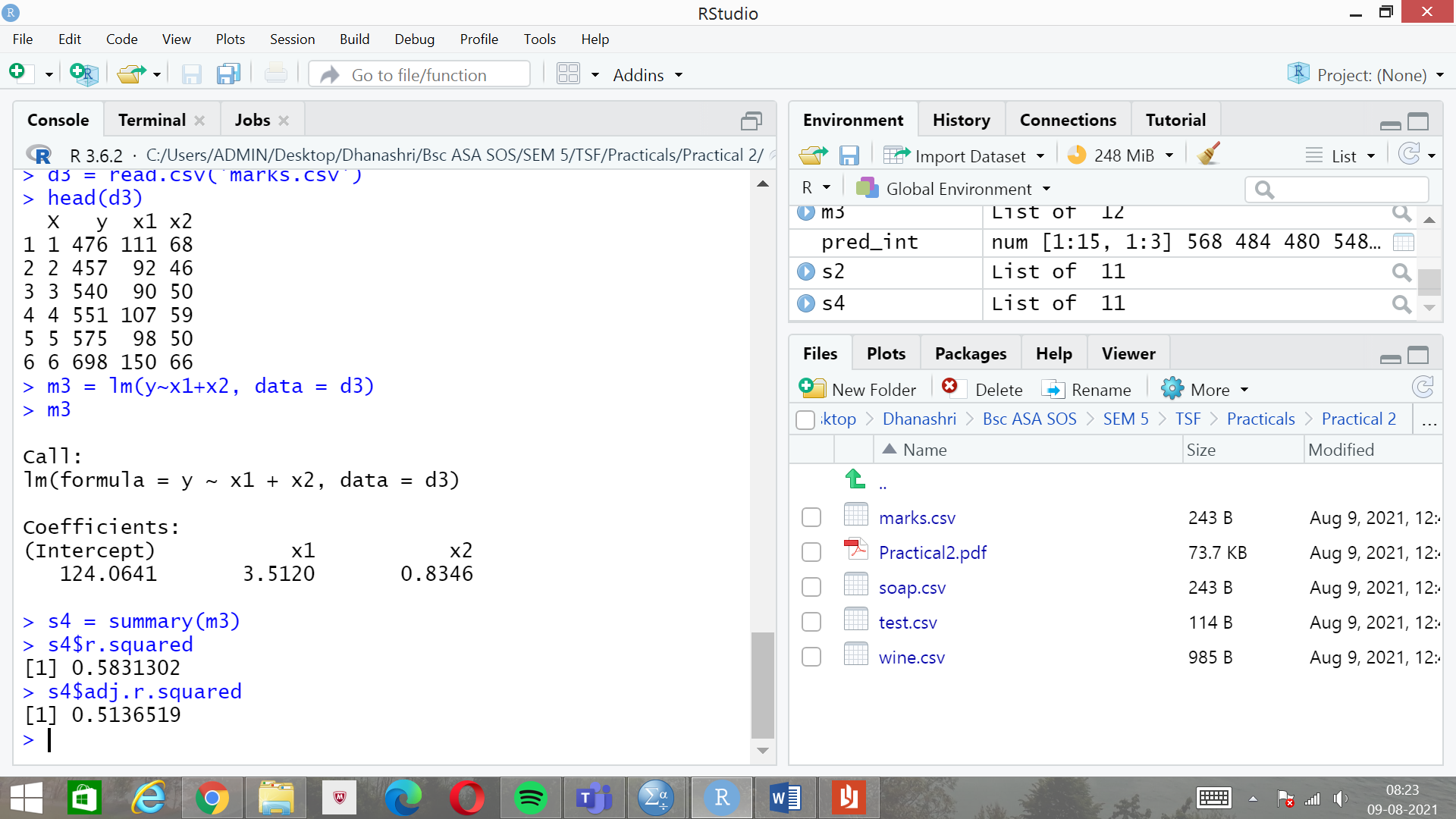


1. R-squared and Adjusted R-squared:

s4 = summary(m3)

s4$r.squared

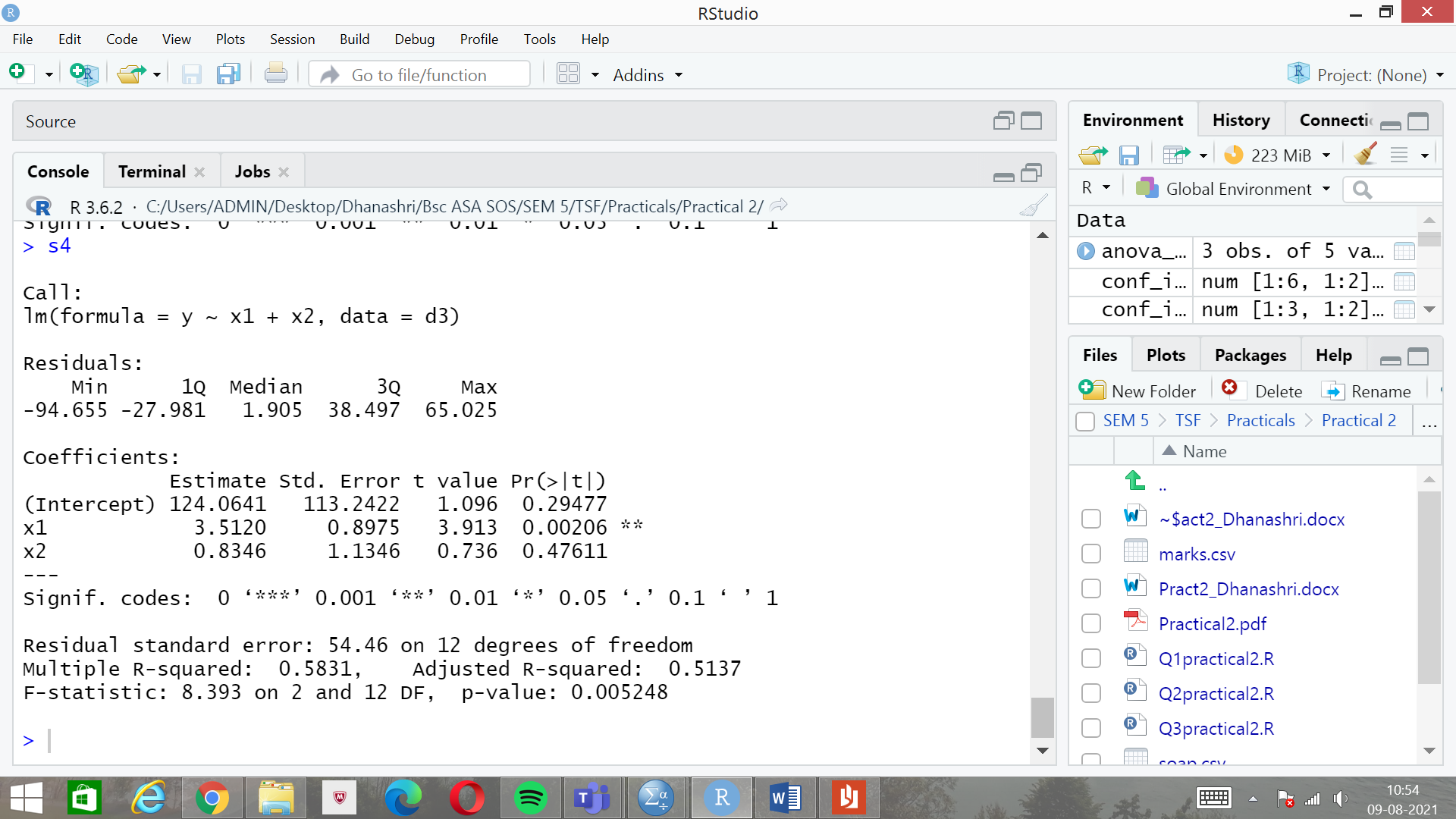
s4$adj.r.squared



Variation explained by R-squared is 51%.

1. T-test for testing individual coefficients:

s4 = summary(m3)



As p-value is less than significance level in case of coefficients of x1,we reject the null hypothesis (H01).In case of coefficients of x2, we accept the null hypothesis H02 because the p-value is greater than the significance level.