**Question 1:**  Linear regression: Read the regression part of the article at this site: http://reliawiki.org/index.php/Multiple\_Linear\_Regression\_Analysis. Run the example linear regression in R. Please compute the value of F0 (211.9) separately step-by-step, either in Excel or in R, and then arrive at the same result you obtained by running summary in the regression example

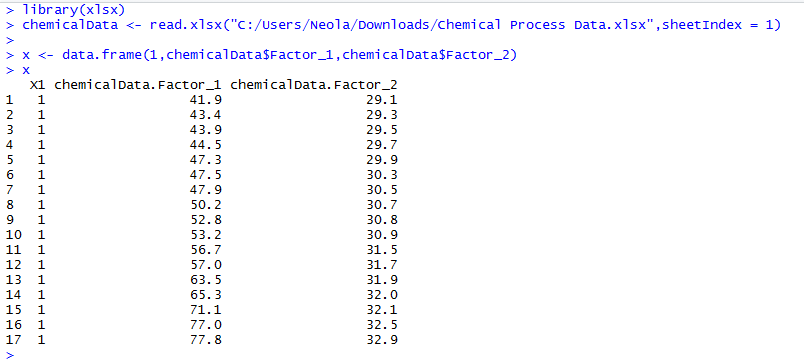
Output:

library(xlsx)

chemicalData <- read.xlsx("C:/Users/Neola/Downloads/Chemical Process Data.xlsx",sheetIndex = 1)

x <- data.frame(1,chemicalData$Factor\_1,chemicalData$Factor\_2)

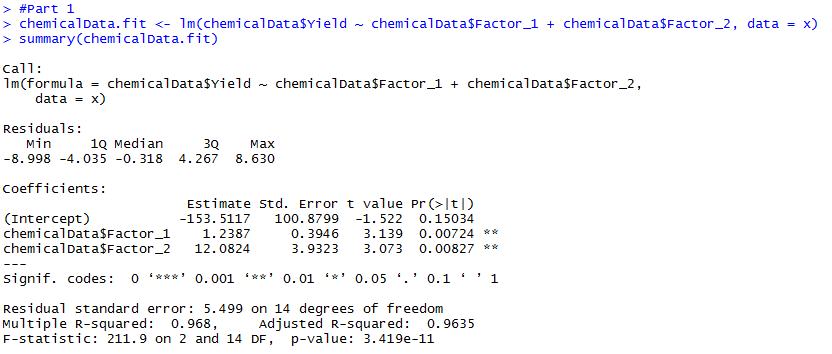
x



#Part 1

chemicalData.fit <- lm(chemicalData$Yield ~ chemicalData$Factor\_1 + chemicalData$Factor\_2, data = x)

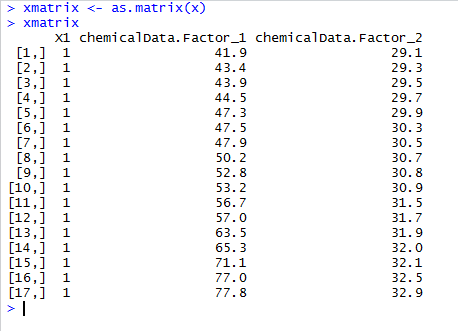
summary(chemicalData.fit)



#Part2

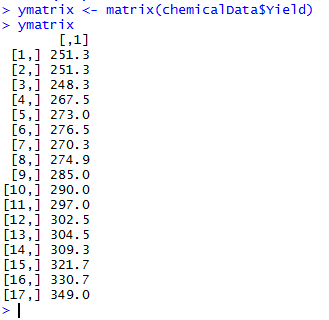
xmatrix <- as.matrix(x)

xmatrix



ymatrix <- matrix(chemicalData$Yield)

ymatrix



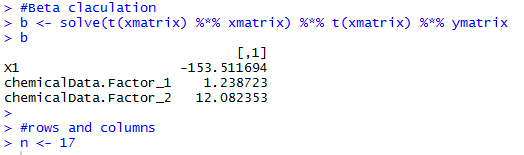
#Beta claculation

b <- solve(t(xmatrix) %\*% xmatrix) %\*% t(xmatrix) %\*% ymatrix

b

#rows and columns

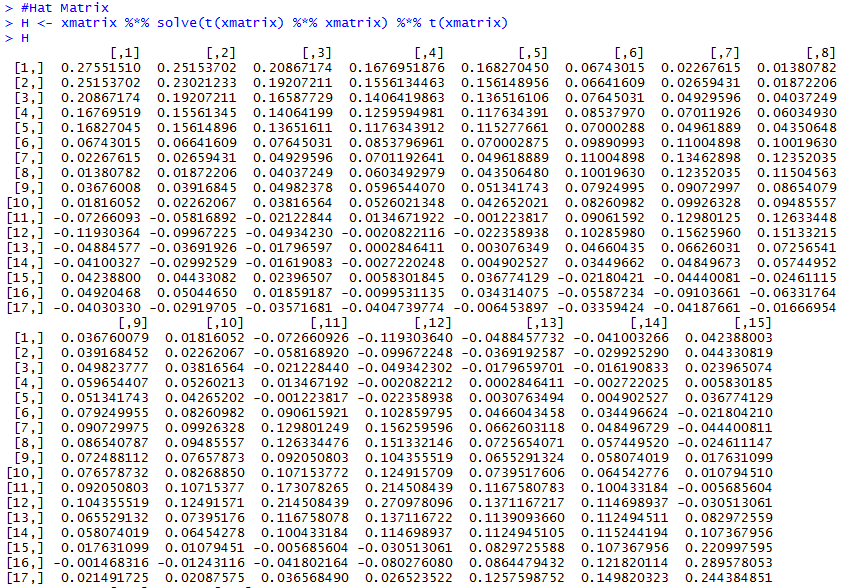
n <- 17

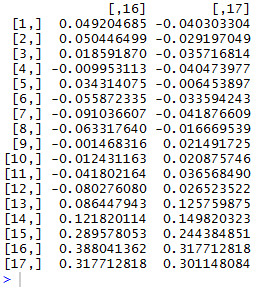


#Hat Matrix

H <- xmatrix %\*% solve(t(xmatrix) %\*% xmatrix) %\*% t(xmatrix)

H





#SSR calculation

J <- matrix(1, nrow = 17, ncol = 17)

JM <- J/17

mid <- H-JM

#SSR (Regression Sum of squares)

SSR <- t(ymatrix) %\*% mid %\*% ymatrix

SSR

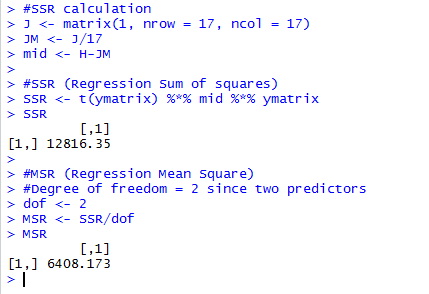
#MSR (Regression Mean Square)

#Degree of freedom = 2 since two predictors

dof <- 2

MSR <- SSR/dof

MSR



#SSE (Error Sum of squares)

I <- diag(n)

SSE <- t(ymatrix) %\*% (I-H) %\*% ymatrix

SSE

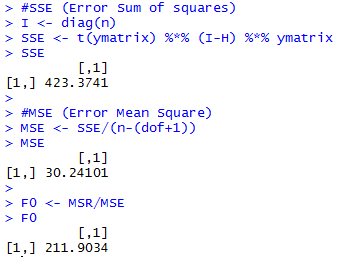
#MSE (Error Mean Square)

MSE <- SSE/(n-(dof+1))

MSE

F0 <- MSR/MSE

F0

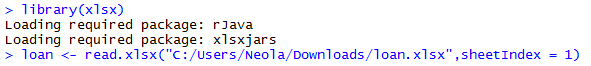


**Question 2:** Logistics regression: Run logistics regression on the loan example with the variable Decision as the dependent variable and the five categorical variables identified in the class (Res\_status, Occupation, Job\_status, Liab\_ref, Acc\_ref) as the independent variables. Show your prediction for input (owner, creative\_, governmen, f, given) and (rent, creative\_, governmen, f, given).

**Solution:**

library(xlsx)

loan <- read.xlsx("C:/Users/Neola/Downloads/loan.xlsx",sheetIndex = 1)



Res\_status <- as.factor(loan$Res\_status)

Occupation <- as.factor(loan$Occupation)

Job\_status <- as.factor(loan$Job\_status)

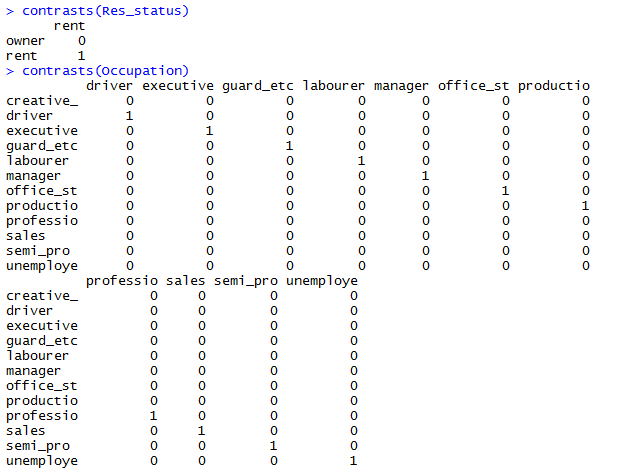
Liab\_ref <- as.factor(loan$Liab\_ref)

Acc\_ref <- as.factor(loan$Acc\_ref)

Decision <- as.factor(loan$Decision)

contrasts(Res\_status)

contrasts(Occupation)

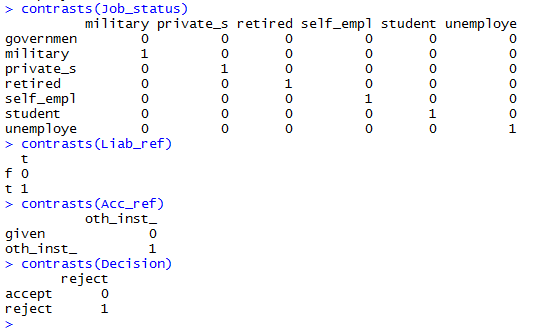


contrasts(Job\_status)

contrasts(Liab\_ref)

contrasts(Acc\_ref)

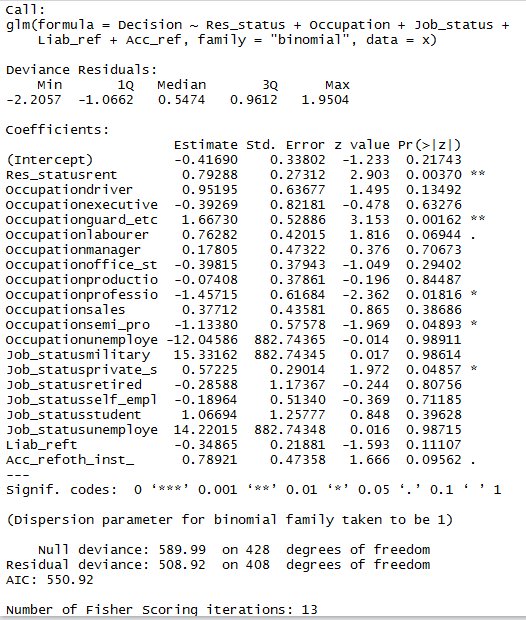
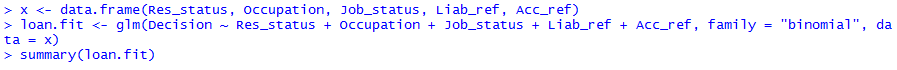
contrasts(Decision)



x <- data.frame(Res\_status, Occupation, Job\_status, Liab\_ref, Acc\_ref)

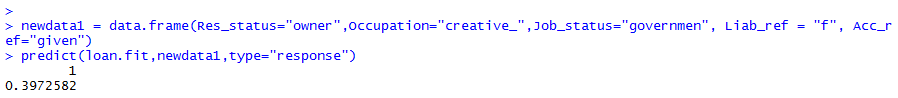
loan.fit <- glm(Decision ~ Res\_status + Occupation + Job\_status + Liab\_ref + Acc\_ref, family = "binomial", data = x)

summary(loan.fit)



newdata1 = data.frame(Res\_status="owner",Occupation="creative\_",Job\_status="governmen", Liab\_ref = "f", Acc\_ref="given")

predict(loan.fit,newdata1,type="response")



newdata2 = data.frame(Res\_status="rent",Occupation="creative\_",Job\_status="governmen", Liab\_ref = "f", Acc\_ref="given")

predict(loan.fit,newdata2,type="response")

