Team members:
Sirisha Lanka - PES1201700294
Malavikka R - PES1201700794
Kevin Arulraj - PES1201700659

Assignment 4

Problem Statement:
Perform linear regression on the Iris dataset and multiple regression on the Obesity dataset.

Dataset:
Iris dataset, Obesity dataset

Code:
#1 Linear regression

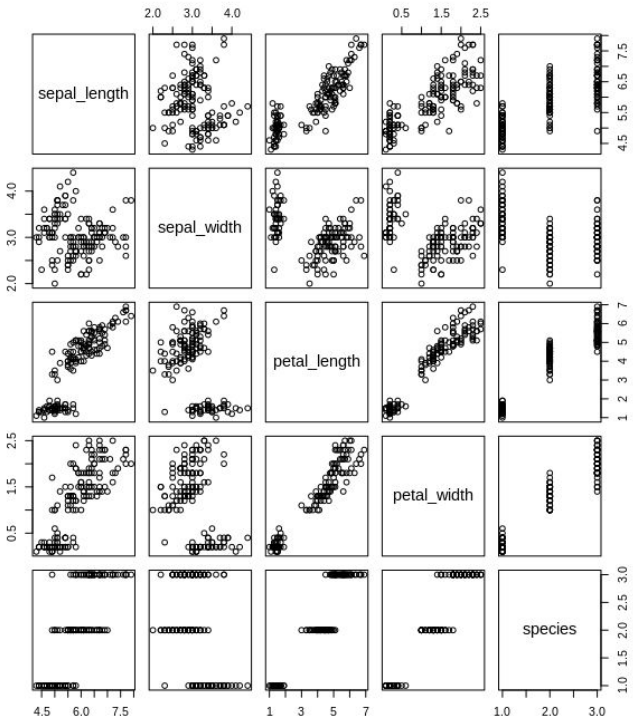
```
data <- read.csv("iris.data")
colnames(data) <- c("sepal_length" , "sepal_width" , "petal_length" , "petal_width" , "species")
summary(data)
plot(data)
cor(data$sepal_length,data$sepal_width)
fit <- lm(data$petal_length ~ data$petal_width)
fit
par(mfrow=c(2,2))
plot(fit)
residuals(fit)
summary(fit)
data2 <- subset(data , data$species != 'Iris-setosa')
plot(data2)
cor(data2$petal_length , data2$petal_width)
fit <- lm(data2$petal_length ~ data2$petal_width)
summary(fit)
```

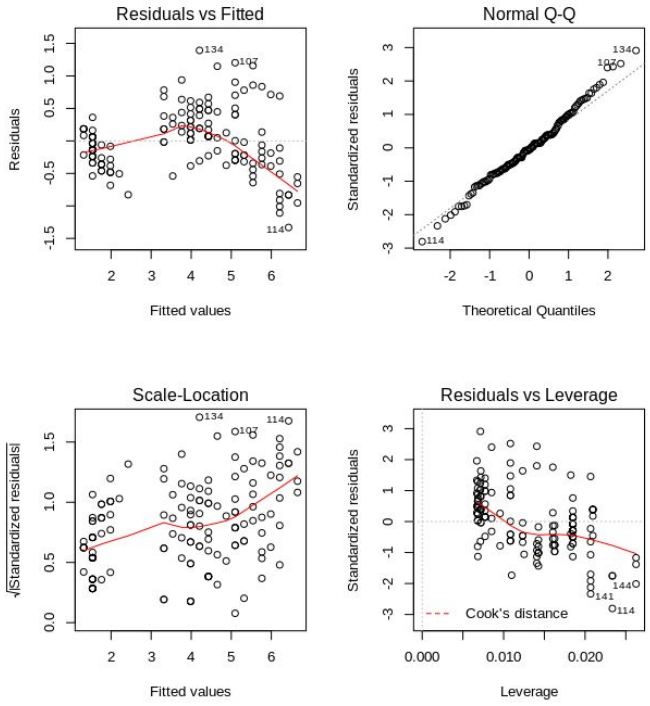
#2 Multiple regression

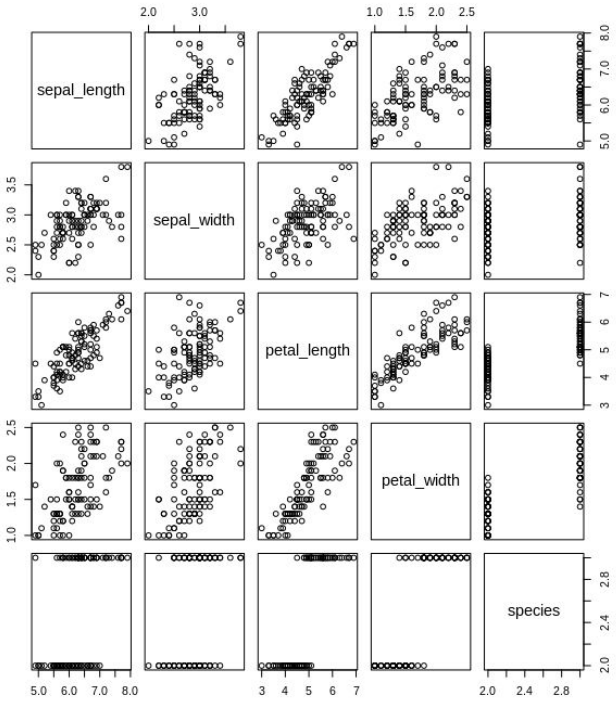
```
data=read.csv("D:\\R-3.6.1\\obese.csv")
summary(data)
model=lm(data$obese~ data$ ï..sex +data$sbp + data$dbp + data$scl +data$age +data$bmi)
plot(model)
plot(data)
summary(model)
resid(model)
fitted(model)
par(mfrow=c(2,2))
model=lm(data$obese ~ data$sbp +data$dbp)
summary(model)
plot(model)
```

Results:

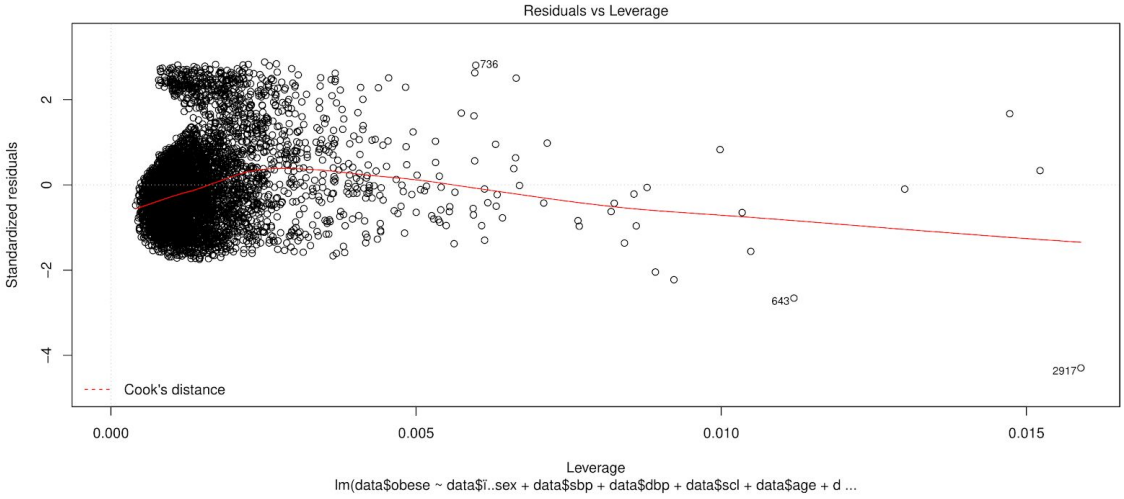
#1

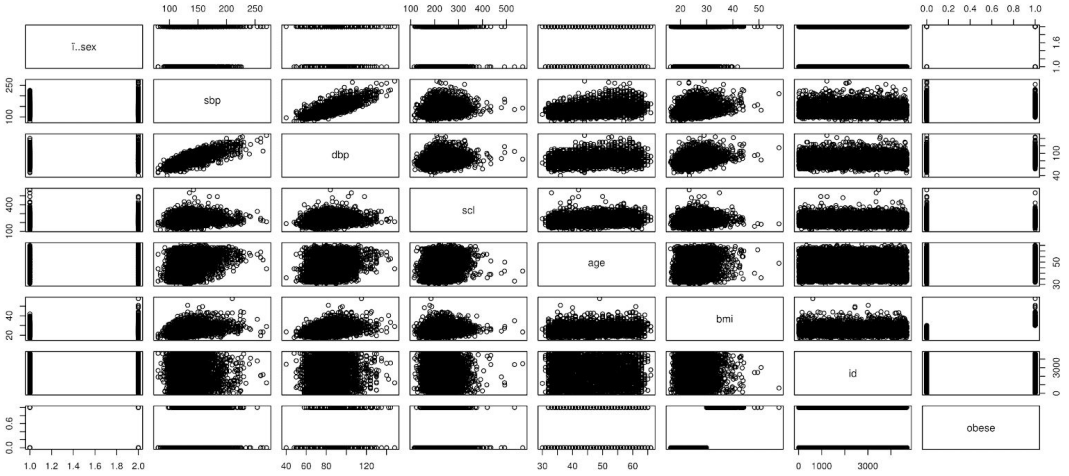


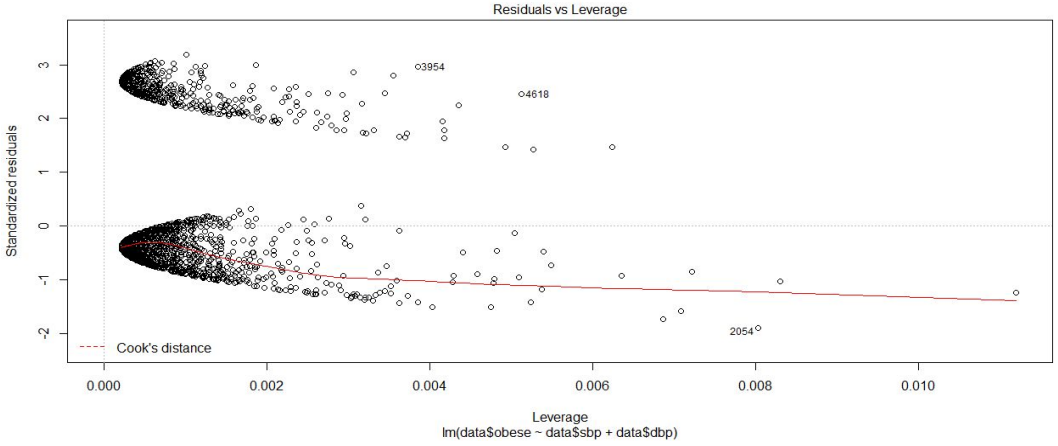

Residuals vs Fitted
Residuals
Fitted values
Normal Q-Q
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Fitted values
Residuals vs Leverage
Standardized residuals
Cook's distance
Leverage


sepal_length
sepal_width
petal_length
petal_width
species

#2


Residuals vs Leverage
Standardized residuals
Cook's distance
736
643
2917
0.000
0.005
0.010
0.015
Leverage
lm(data$obese ~ data$i..sex + data$sbp + data$dbp + data$scl + data$age + d ...


i..sex
sbp
dbp
scl
age
bmi
id
obese


Residuals vs Leverage
Standardized residuals
Cook's distance
3954
4618
2054
0.000
0.002
0.004
0.006
0.008
0.010
Leverage
lm(data$obese ~ data$sbp + data$dbp)