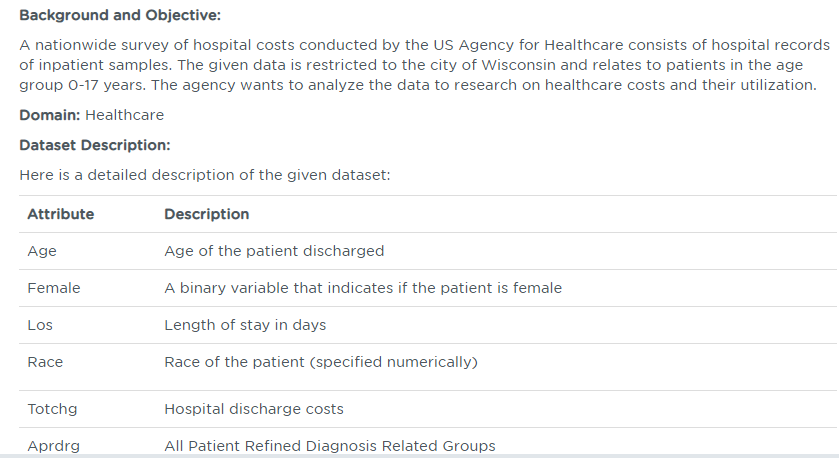
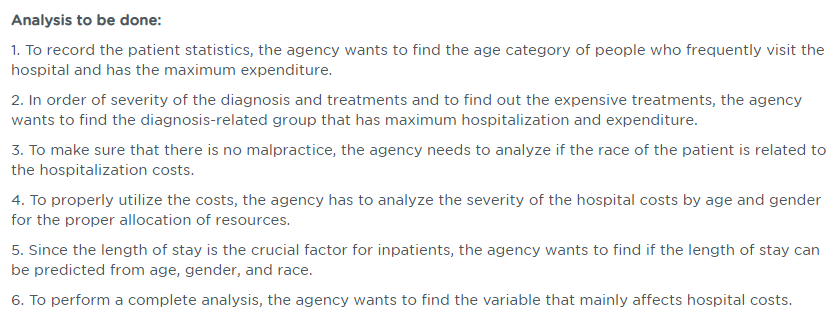
**HEALTHCARE COST ANALYSIS**





**SOLUTION (CODE)**

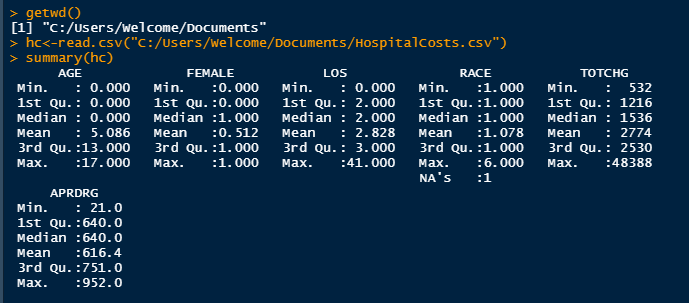
1. **Import the csv file and check summary**

getwd()

hc<-read.csv("C:/Users/Welcome/Documents/HospitalCosts.csv")

View(hc)

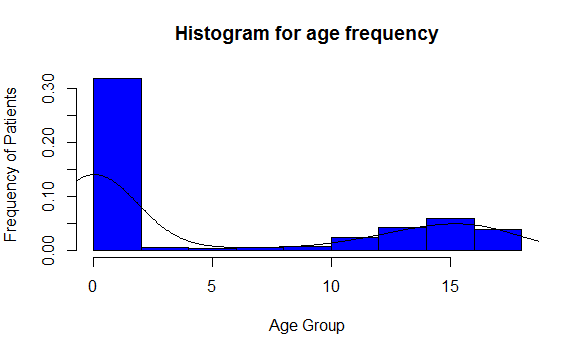
summary(hc)



**2) create histogram for age category of people frequently visit the hospital**

hist(hc$AGE, main = "Histogram for age frequency", xlab = "Age Group", ylab = "Frequency of Patients", prob= TRUE, col = "blue")

lines(density(hc$AGE))



summary(as.factor(hc$AGE))

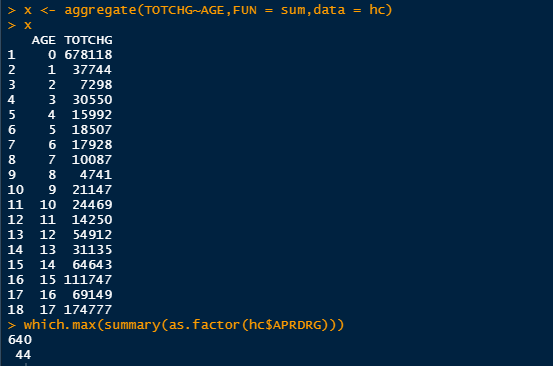
C:\Users\Welcome\Desktop\R PROGRAMMING SCREENSHOT\C3.PNG

x <- aggregate(TOTCHG~AGE,FUN = sum,data = hc)

x

max(x)

which.max(summary(as.factor(hc$APRDRG)))

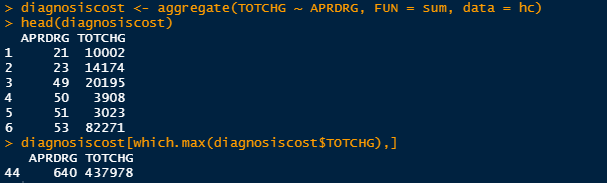


1. **Diagnosis related group that has max hospitalization expenditure**

diagnosiscost <- aggregate(TOTCHG ~ APRDRG, FUN = sum, data = hc)

head(diagnosiscost)

diagnosiscost[which.max(diagnosiscost$TOTCHG),]



1. **Analysing relation between race of patient and hospitalization cost**

summary(as.factor(hc$RACE))

head(hc)

hc<-na.omit(hc)

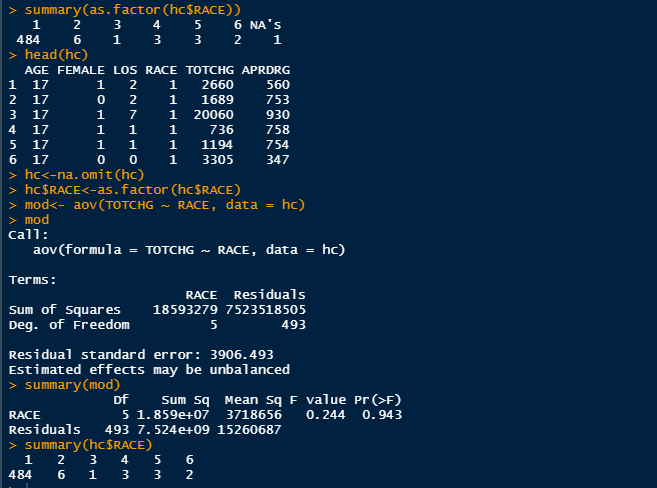
hc$RACE<-as.factor(hc$RACE)

mod<- aov(TOTCHG ~ RACE, data = hc)

mod

summary(mod)

summary(hc$RACE)



1. **Analyzing the severity of hospital cost by age and gender**

model1 <- lm(TOTCHG ~ AGE + FEMALE, data = hc)

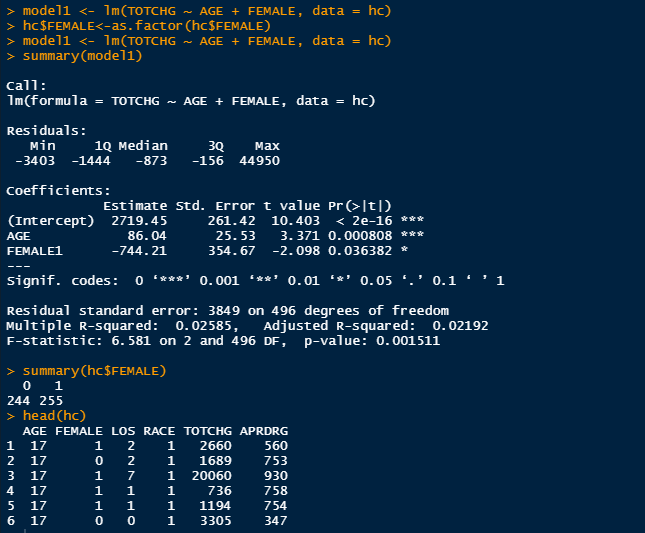
hc$FEMALE<-as.factor(hc$FEMALE)

model1 <- lm(TOTCHG ~ AGE + FEMALE, data = hc)

summary(model1)

summary(hc$FEMALE)

head(hc)

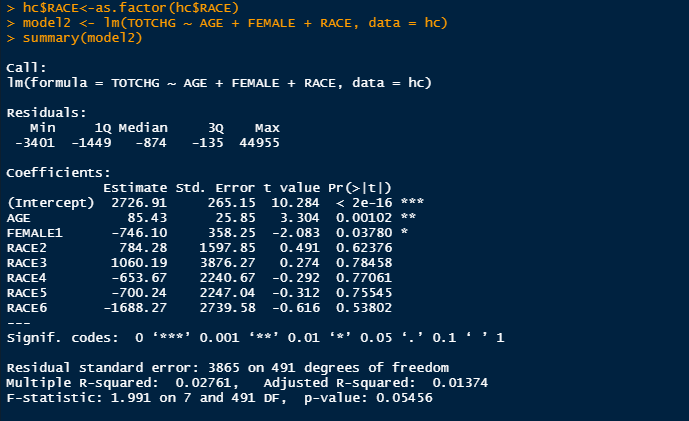


1. **Prediction of length of stay from age, gender and race**

hc$RACE<-as.factor(hc$RACE)

model2 <- lm(TOTCHG ~ AGE + FEMALE + RACE, data = hc)

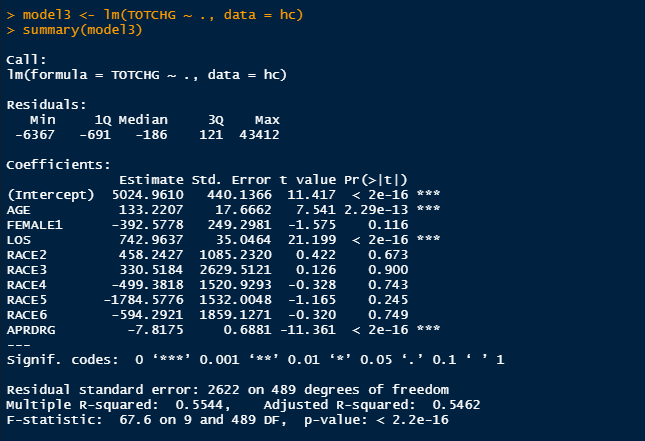
summary(model2)



1. **Factors affecting hospital cost**

model3 <- lm(TOTCHG ~ ., data = hc)

summary(model3)



Code screenshot

