library(readxl)

HospData <-read\_xlsx("E:/Simplilearn/Project/Project 7/hospitalcosts.xlsx")

View(HospData)

summary(HospData)

hist(HospData$AGE)

summary(as.factor(HospData$AGE))

aggregate(TOTCHG ~ AGE, FUN = sum, data = HospData)

max(aggregate(TOTCHG ~ AGE, FUN = sum, data = HospData))

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

CostForDiagnosis<- aggregate(TOTCHG~APRDRG, FUN = sum, data = HospData)

CostForDiagnosis

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

summary(as.factor(HospData$RACE))

head(HospData)

HospData<-(na.omit(HospData))

HospData

PatientRace<-as.factor(HospData$RACE)

Mod1<- aov(TOTCHG~RACE, data = HospData)

Mod1

summary(Mod1)

summary(HospData)

Mod2<-lm(TOTCHG~AGE+FEMALE, data=HospData)

Mod2

Fem<-as.factor(HospData$FEMALE)

summary(Mod2)

summary(Fem)

head(HospData)

Mod3<-lm(TOTCHG~AGE+FEMALE+RACE, data = HospData)

Mod3

summary(Mod3)

Mod4<-lm(TOTCHG ~ ., data=HospData)

Mod4

summary(Mod4)