data(segmentationData)

df1 = segmentationData

View(df1)

table(df1$Class)

library("randomForest")

fit = randomForest(Class~. , data = df )

varImpPlot(fit)

fit1 = rpart(Class~. , data = df )

a = varImp(fit1)

a = a[order(a$Overall),]

View(a)

df$chidfren = ifelse(df$chidfren == "4+" , 4 , df$chidfren)

table(df$chidfren)

df$chidfren = ifelse(df$chidfren == 5 , 0 ,df$chidfren)

library(dplyr)

df = df%>%

select(-post\_area)

#Now we are done with preparing data , lets split the data.

set.seed(77)

s=sample(1:nrow(df),0.7\*nrow(df))

df\_trainval=df[s,]

df\_test=df[-s,]

str(df\_trainval)

rf = randomForest(as.factor(Revenue.Grid)~. , data = df\_trainval)

predicted = predict(rf, newdata = df\_test )

table(df\_test$Revenue.Grid , predicted)

importance(rf)

varImpPlot(rf)

plot(df$Investment.in.Equity, df$Revenue.Grid)

rf

summary(rf)