1. complete.cases(wis)



2. 598

3.

summary(wis$CellSize)

Min. 1st Qu. Median Mean 3rd Qu. Max.

1.000 1.000 1.000 3.138 5.000 10.000

נקח גבול בין אחד ל3.138 כי זה בין המינמום לממוצע

נקח גבול בין 3.138 ל5 כי זה בין הממוצע לqu השלישי

> for(x in wis$CellSize){i=i+1}

> for(x in wis$CellSize){

+ i=i+1

+ if(x<=3.138){wis[i,"CellSize"]="Low"}

+ else if(x>3.138 & x<=5){wis[i,"CellSize"]="Mid"}

+ else if(x>5){wis[i,"CellSize"]="High"}}

zoo::na.aggregate(wis$Class)

4. wis$Class <-as.factor(wis$Class)

5.

library(caret)

inTrain<-createDataPartition(y=wis$Class,p=0.75,list=FALSE)

tr <-wis[inTrain,]

te <-wis[-inTrain,]

6. מספר שורות של training שווה ל 1048

ומספר השורות של testing שווה 348

7.

library(party)

df\_ctree = ctree(Class~.,data-training)

plot(df\_ctree)

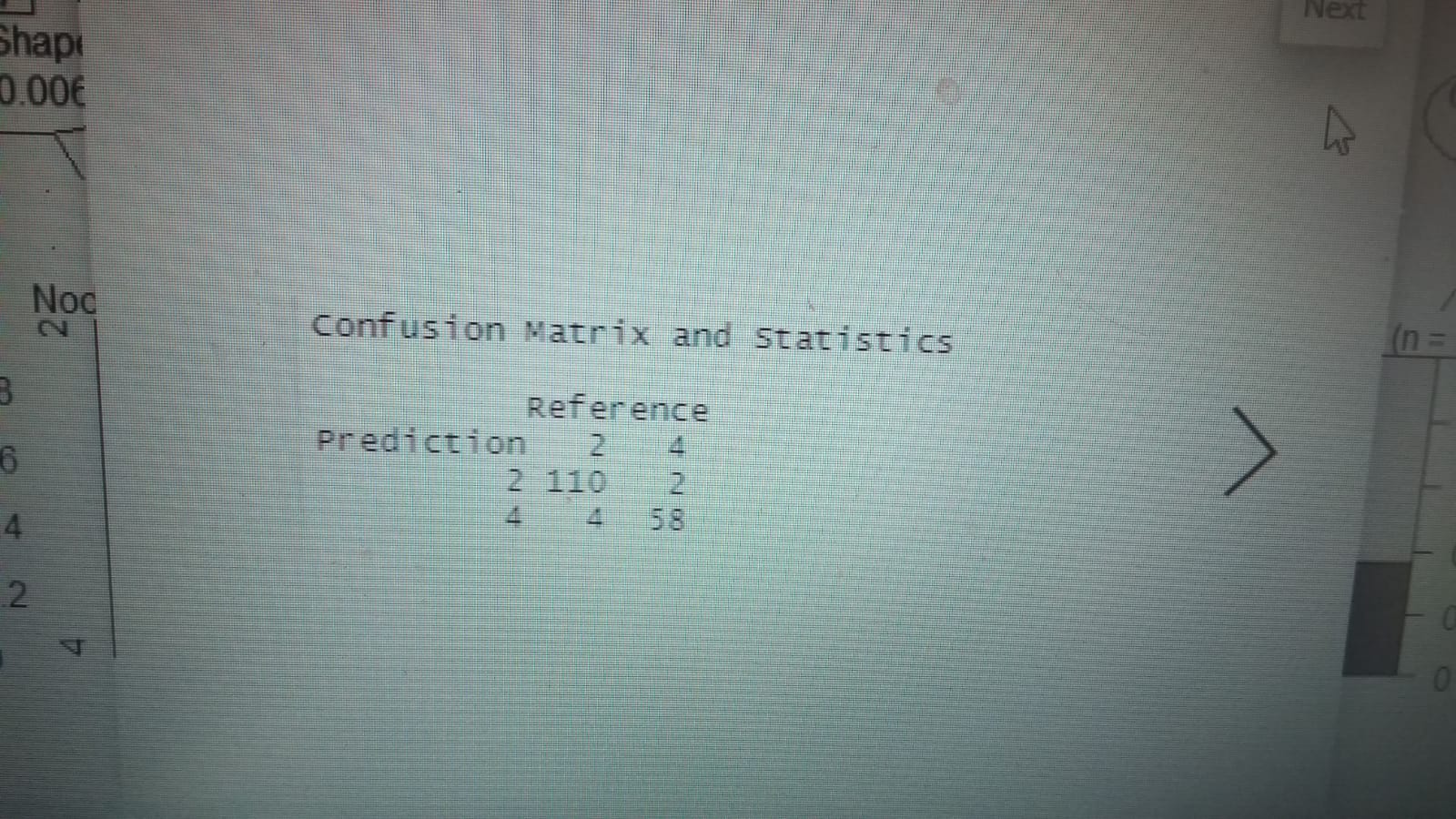
8.

Predictions = predict(df\_ctree,newdata=testing)

Plot(predictions)

Cm = confutionMatrix(Predictions,reference=testing$Class)

Cm

רמת הדיוק הינה 0.965

חלק 2

1.

lab <-read.csv("lab.csv")

2.

lab.features <-lab

> lab.features$Class<-NULL

> res <-kmeans(lab.features,2)

2.1.

res$cluster

[1] 2 2 1 2 1 1 2 2 1 2 2 2 1 1 2 2 2 2 2 1 1 1 1 1 2 1 1 1 2 1 1 2

2.2

table(lab$Class,res$cluster)

1 2

0 1 15

1 15 1

2.3 נרשמו בקלסטר הלא נכון

