11.

# (a.)

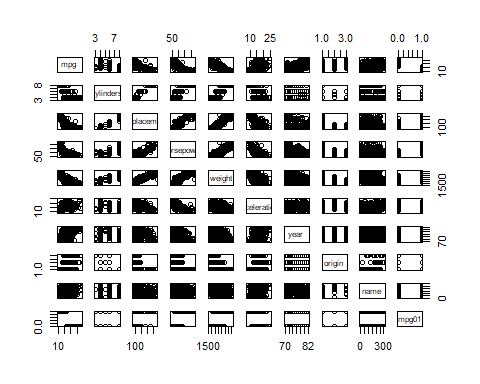
library(ISLR)  
summary(Auto)

## mpg cylinders displacement horsepower   
## Min. : 9.00 Min. :3.000 Min. : 68.0 Min. : 46.0   
## 1st Qu.:17.00 1st Qu.:4.000 1st Qu.:105.0 1st Qu.: 75.0   
## Median :22.75 Median :4.000 Median :151.0 Median : 93.5   
## Mean :23.45 Mean :5.472 Mean :194.4 Mean :104.5   
## 3rd Qu.:29.00 3rd Qu.:8.000 3rd Qu.:275.8 3rd Qu.:126.0   
## Max. :46.60 Max. :8.000 Max. :455.0 Max. :230.0   
##   
## weight acceleration year origin   
## Min. :1613 Min. : 8.00 Min. :70.00 Min. :1.000   
## 1st Qu.:2225 1st Qu.:13.78 1st Qu.:73.00 1st Qu.:1.000   
## Median :2804 Median :15.50 Median :76.00 Median :1.000   
## Mean :2978 Mean :15.54 Mean :75.98 Mean :1.577   
## 3rd Qu.:3615 3rd Qu.:17.02 3rd Qu.:79.00 3rd Qu.:2.000   
## Max. :5140 Max. :24.80 Max. :82.00 Max. :3.000   
##   
## name   
## amc matador : 5   
## ford pinto : 5   
## toyota corolla : 5   
## amc gremlin : 4   
## amc hornet : 4   
## chevrolet chevette: 4   
## (Other) :365

attach(Auto)  
mpg01=rep(1,length(mpg))  
mpg01[mpg>median(mpg)]=0  
data=data.frame(Auto,mpg01)

# (b.)

plot(data)



cylinders, weight, displacement, horsepower seems negatively correlated.

# (c.)

traset=Auto[c(1:(nrow(Auto)\*0.8)),]  
tesset=Auto[-c(1:(nrow(Auto)\*0.8)),]  
te=mpg01[-c(1:(nrow(Auto)\*0.8))]

# (d.)

library(MASS)  
attach(data)

## The following object is masked \_by\_ .GlobalEnv:  
##   
## mpg01

## The following objects are masked from Auto:  
##   
## acceleration, cylinders, displacement, horsepower, mpg, name,  
## origin, weight, year

mo=lda(mpg01 ~ cylinders + weight + displacement + horsepower, data = Auto)  
mop=predict(mo,tesset)  
mean(mop$class!=te)

## [1] 0.08860759

# (e.)

qmo=qda(mpg01 ~ cylinders + weight + displacement + horsepower, data = Auto)  
qmop=predict(qmo,tesset)  
mean(qmop$class!=te)

## [1] 0.1392405

# (f.)

logm=glm(mpg01 ~ cylinders + weight + displacement + horsepower, data = Auto)  
logmop=predict(logm,tesset)  
clas=rep(0,length(te))  
clas[logmop>0.5]=1  
mean(clas!=te)

## [1] 0.08860759

# (g.)

library(class)  
train=cbind(cylinders, weight, displacement, horsepower)[c(1:(nrow(Auto)\*0.8)),]  
test=cbind(cylinders, weight, displacement, horsepower)[314:392,]  
tmg01=mpg01[c(1:(nrow(Auto)\*0.8))]  
set.seed(1)  
km=knn(train,test,tmg01,k=1)  
mean(te!=km)

## [1] 0.2151899

# km2=knn(train,test,tmg01,k=100) mean(te!=km2)

## [1] 0.2151899

km3=knn(train,test,tmg01,k=150)  
mean(te!=km3)

## [1] 0.1898734

kmlast=knn(train,test,tmg01,k=200)  
mean(te!=kmlast)

## [1] 0.164557

After trying different k, we discover that in K=200, it performs the best on the data.