hw4alex

```
head(Auto)
```

```
##   mpg cylinders displacement horsepower weight acceleration year origin
## 1  18         8          307        130   3504         12.0   70      1
## 2  15         8          350        165   3693         11.5   70      1
## 3  18         8          318        150   3436         11.0   70      1
## 4  16         8          304        150   3433         12.0   70      1
## 5  17         8          302        140   3449         10.5   70      1
## 6  15         8          429        198   4341         10.0   70      1
##                        name
## 1 chevrolet chevelle malibu
## 2         buick skylark 320
## 3        plymouth satellite
## 4             amc rebel sst
## 5               ford torino
## 6          ford galaxie 500
```

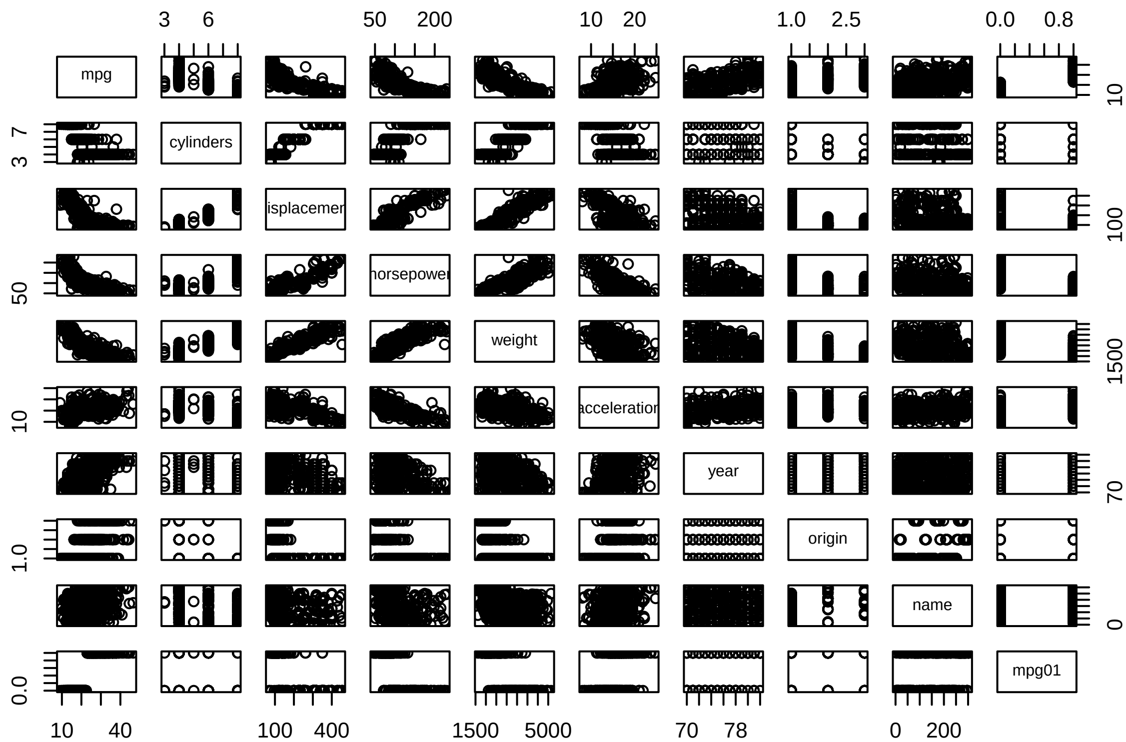
Question 1

```
cars = transform(Auto, mpg01= ifelse(mpg>median(Auto$mpg), 1, 0))
head(cars)
```

```
##   mpg cylinders displacement horsepower weight acceleration year origin
## 1  18         8          307        130   3504         12.0   70      1
## 2  15         8          350        165   3693         11.5   70      1
## 3  18         8          318        150   3436         11.0   70      1
## 4  16         8          304        150   3433         12.0   70      1
## 5  17         8          302        140   3449         10.5   70      1
## 6  15         8          429        198   4341         10.0   70      1
##                        name mpg01
## 1 chevrolet chevelle malibu     0
## 2         buick skylark 320     0
## 3        plymouth satellite     0
## 4             amc rebel sst     0
## 5               ford torino     0
## 6          ford galaxie 500     0
```

Question 2: From the pairs of scatter plots we can compare an attribute to the MPG column in order to find the most useful feature for predicting mpg01.The cylinders attribute does not give us much information, because it is a discreet variable. Comparing the displacement column to the MPG shows us that there is a negative correlation with the two, which makes sense as a bigger engine will be heavier and require more gas to move it. Similarly with horsepower, the more a car has the lower MPG it will most likely have. Weight also seems to be an important attribute when compared to MPG: there is a clear negative correlation between the two. Acceleration does not seem to correlate with MPG in anyway. From the year column, we can tell that newer cars to tend have slightly better MPG than older cars. Origin and Name attributes are discreet variables and are not very helpful in predicting the MPG of a vehicle.

```r
pairs(cars)
```



Question #3

```r
train = cars[1:300,]
test = cars[301:392,]
```

Question #4

```r
n_0 <- length(which(train$mpg01 == 0))
n_1 <- length(which(train$mpg01 == 1))
p_0 <- n_0/300
p_1 <- n_1/300
X0 <- train[train$mpg01 == 0, 3:5]
X1 <- train[train$mpg01 == 1, 3:5]
mean_0 <- colMeans(X0)
mean_1 <- colMeans(X1)
s_0 <- cov(X0)
s_1 <- cov(X1)
s_pooled <- ((n_0-1)*s_0 + (n_1-1)*s_1) / (n_0+n_1-2)
s_inv <- solve(s_pooled)
alpha_0 <- -0.5* t(mean_0) %*% s_inv %*% mean_0
alpha_1 <- -0.5* t(mean_1) %*% s_inv %*% mean_1
beta_0 <- s_inv %*% mean_0
beta_1 <- s_inv %*% mean_1
```

```r
prediction <- c()
d_0vec <- c()
d_1vec <- c()
label <- c("0", "1")

for (i in 1:nrow(test)){
  y <- t(test[i, 3:5])
  d_0 <- alpha_0 + t(beta_0) %*% y
  d_1 <- alpha_1 + t(beta_1) %*% y
  d_vec <- c(d_0, d_1)
  prediction <- append(prediction, label[which.max( d_vec )])
  d_0vec <- append(d_0vec, d_0)
  d_1vec <- append(d_1vec, d_1)
}

test$prediction <- prediction

error <- length(which(test$mpg01 != test$prediction)) / 92

test
```

```
##      mpg cylinders displacement horsepower weight acceleration year origin
## 303 34.5         4          105         70   2150         14.9   79      1
## 304 31.8         4           85         65   2020         19.2   79      3
## 305 37.3         4           91         69   2130         14.7   79      2
## 306 28.4         4          151         90   2670         16.0   79      1
## 307 28.8         6          173        115   2595         11.3   79      1
## 308 26.8         6          173        115   2700         12.9   79      1
## 309 33.5         4          151         90   2556         13.2   79      1
## 310 41.5         4           98         76   2144         14.7   80      2
## 311 38.1         4           89         60   1968         18.8   80      3
## 312 32.1         4           98         70   2120         15.5   80      1
## 313 37.2         4           86         65   2019         16.4   80      3
## 314 28.0         4          151         90   2678         16.5   80      1
## 315 26.4         4          140         88   2870         18.1   80      1
## 316 24.3         4          151         90   3003         20.1   80      1
## 317 19.1         6          225         90   3381         18.7   80      1
## 318 34.3         4           97         78   2188         15.8   80      2
## 319 29.8         4          134         90   2711         15.5   80      3
## 320 31.3         4          120         75   2542         17.5   80      3
## 321 37.0         4          119         92   2434         15.0   80      3
## 322 32.2         4          108         75   2265         15.2   80      3
## 323 46.6         4           86         65   2110         17.9   80      3
## 324 27.9         4          156        105   2800         14.4   80      1
## 325 40.8         4           85         65   2110         19.2   80      3
## 326 44.3         4           90         48   2085         21.7   80      2
## 327 43.4         4           90         48   2335         23.7   80      2
## 328 36.4         5          121         67   2950         19.9   80      2
## 329 30.0         4          146         67   3250         21.8   80      2
## 330 44.6         4           91         67   1850         13.8   80      3
## 332 33.8         4           97         67   2145         18.0   80      3
## 333 29.8         4           89         62   1845         15.3   80      2
## 334 32.7         6          168        132   2910         11.4   80      3
```

```
## 335 23.7         3          70        100   2420         12.5   80       3
## 336 35.0         4         122         88   2500         15.1   80       2
## 338 32.4         4         107         72   2290         17.0   80       3
## 339 27.2         4         135         84   2490         15.7   81       1
## 340 26.6         4         151         84   2635         16.4   81       1
## 341 25.8         4         156         92   2620         14.4   81       1
## 342 23.5         6         173        110   2725         12.6   81       1
## 343 30.0         4         135         84   2385         12.9   81       1
## 344 39.1         4          79         58   1755         16.9   81       3
## 345 39.0         4          86         64   1875         16.4   81       1
## 346 35.1         4          81         60   1760         16.1   81       3
## 347 32.3         4          97         67   2065         17.8   81       3
## 348 37.0         4          85         65   1975         19.4   81       3
## 349 37.7         4          89         62   2050         17.3   81       3
## 350 34.1         4          91         68   1985         16.0   81       3
## 351 34.7         4         105         63   2215         14.9   81       1
## 352 34.4         4          98         65   2045         16.2   81       1
## 353 29.9         4          98         65   2380         20.7   81       1
## 354 33.0         4         105         74   2190         14.2   81       2
## 356 33.7         4         107         75   2210         14.4   81       3
## 357 32.4         4         108         75   2350         16.8   81       3
## 358 32.9         4         119        100   2615         14.8   81       3
## 359 31.6         4         120         74   2635         18.3   81       3
## 360 28.1         4         141         80   3230         20.4   81       2
## 361 30.7         6         145         76   3160         19.6   81       2
## 362 25.4         6         168        116   2900         12.6   81       3
## 363 24.2         6         146        120   2930         13.8   81       3
## 364 22.4         6         231        110   3415         15.8   81       1
## 365 26.6         8         350        105   3725         19.0   81       1
## 366 20.2         6         200         88   3060         17.1   81       1
## 367 17.6         6         225         85   3465         16.6   81       1
## 368 28.0         4         112         88   2605         19.6   82       1
## 369 27.0         4         112         88   2640         18.6   82       1
## 370 34.0         4         112         88   2395         18.0   82       1
## 371 31.0         4         112         85   2575         16.2   82       1
## 372 29.0         4         135         84   2525         16.0   82       1
## 373 27.0         4         151         90   2735         18.0   82       1
## 374 24.0         4         140         92   2865         16.4   82       1
## 375 36.0         4         105         74   1980         15.3   82       2
## 376 37.0         4          91         68   2025         18.2   82       3
## 377 31.0         4          91         68   1970         17.6   82       3
## 378 38.0         4         105         63   2125         14.7   82       1
## 379 36.0         4          98         70   2125         17.3   82       1
## 380 36.0         4         120         88   2160         14.5   82       3
## 381 36.0         4         107         75   2205         14.5   82       3
## 382 34.0         4         108         70   2245         16.9   82       3
## 383 38.0         4          91         67   1965         15.0   82       3
## 384 32.0         4          91         67   1965         15.7   82       3
## 385 38.0         4          91         67   1995         16.2   82       3
## 386 25.0         6         181        110   2945         16.4   82       1
## 387 38.0         6         262         85   3015         17.0   82       1
## 388 26.0         4         156         92   2585         14.5   82       1
## 389 22.0         6         232        112   2835         14.7   82       1
## 390 32.0         4         144         96   2665         13.9   82       3
```

```
## 391 36.0         4          135         84   2370          13.0   82       1
## 392 27.0         4          151         90   2950          17.3   82       1
## 393 27.0         4          140         86   2790          15.6   82       1
## 394 44.0         4           97         52   2130          24.6   82       2
## 395 32.0         4          135         84   2295          11.6   82       1
## 396 28.0         4          120         79   2625          18.6   82       1
## 397 31.0         4          119         82   2720          19.4   82       1
##                                  name mpg01 prediction
## 303              plymouth horizon tc3     1          1
## 304                        datsun 210     1          1
## 305                fiat strada custom     1          1
## 306             buick skylark limited     1          1
## 307                chevrolet citation     1          1
## 308          oldsmobile omega brougham     1          1
## 309                   pontiac phoenix     1          1
## 310                         vw rabbit     1          1
## 311             toyota corolla tercel     1          1
## 312                chevrolet chevette     1          1
## 313                        datsun 310     1          1
## 314                chevrolet citation     1          1
## 315                     ford fairmont     1          1
## 316                       amc concord     1          1
## 317                       dodge aspen     0          0
## 318                         audi 4000     1          1
## 319            toyota corona liftback     1          1
## 320                         mazda 626     1          1
## 321              datsun 510 hatchback     1          1
## 322                    toyota corolla     1          1
## 323                         mazda glc     1          1
## 324                        dodge colt     1          1
## 325                        datsun 210     1          1
## 326              vw rabbit c (diesel)     1          1
## 327                vw dasher (diesel)     1          1
## 328               audi 5000s (diesel)     1          1
## 329                mercedes-benz 240d     1          0
## 330               honda civic 1500 gl     1          1
## 332                         subaru dl     1          1
## 333                  vokswagen rabbit     1          1
## 334                     datsun 280-zx     1          1
## 335                     mazda rx-7 gs     1          1
## 336                  triumph tr7 coupe     1          1
## 338                      honda accord     1          1
## 339                  plymouth reliant     1          1
## 340                     buick skylark     1          1
## 341            dodge aries wagon (sw)     1          1
## 342                chevrolet citation     1          1
## 343                  plymouth reliant     1          1
## 344                    toyota starlet     1          1
## 345                    plymouth champ     1          1
## 346                  honda civic 1300     1          1
## 347                            subaru     1          1
## 348                    datsun 210 mpg     1          1
## 349                     toyota tercel     1          1
## 350                       mazda glc 4     1          1
```

```
## 351                 plymouth horizon 4     1            1
## 352                     ford escort 4w     1            1
## 353                     ford escort 2h     1            1
## 354                   volkswagen jetta     1            1
## 356                      honda prelude     1            1
## 357                     toyota corolla     1            1
## 358                       datsun 200sx     1            1
## 359                          mazda 626     1            1
## 360          peugeot 505s turbo diesel     1            0
## 361                       volvo diesel     1            0
## 362                    toyota cressida     1            1
## 363                  datsun 810 maxima     1            1
## 364                      buick century     0            0
## 365              oldsmobile cutlass ls     1            0
## 366                    ford granada gl     0            0
## 367             chrysler lebaron salon     0            0
## 368                 chevrolet cavalier     1            1
## 369           chevrolet cavalier wagon     1            1
## 370          chevrolet cavalier 2-door     1            1
## 371         pontiac j2000 se hatchback     1            1
## 372                     dodge aries se     1            1
## 373                    pontiac phoenix     1            1
## 374               ford fairmont futura     1            1
## 375                volkswagen rabbit l     1            1
## 376                 mazda glc custom l     1            1
## 377                   mazda glc custom     1            1
## 378             plymouth horizon miser     1            1
## 379                     mercury lynx l     1            1
## 380                   nissan stanza xe     1            1
## 381                       honda accord     1            1
## 382                     toyota corolla     1            1
## 383                        honda civic     1            1
## 384                 honda civic (auto)     1            1
## 385                      datsun 310 gx     1            1
## 386              buick century limited     1            1
## 387 oldsmobile cutlass ciera (diesel)     1            0
## 388         chrysler lebaron medallion     1            1
## 389                     ford granada l     0            0
## 390                   toyota celica gt     1            1
## 391                  dodge charger 2.2     1            1
## 392                   chevrolet camaro     1            1
## 393                    ford mustang gl     1            1
## 394                          vw pickup     1            1
## 395                      dodge rampage     1            1
## 396                        ford ranger     1            1
## 397                         chevy s-10     1            1
```

The test error rate is 0.0543.