HW1

Load necessary libraries

```r
library(fields)
```

```
## Warning: package 'fields' was built under R version 3.2.5
## Loading required package: spam
## Warning: package 'spam' was built under R version 3.2.5
## Loading required package: grid
## Spam version 1.4-0 (2016-08-29) is loaded.
## Type 'help( Spam)' or 'demo( spam)' for a short introduction
## and overview of this package.
## Help for individual functions is also obtained by adding the
## suffix '.spam' to the function name, e.g. 'help( chol.spam)'.
##
## Attaching package: 'spam'
## The following objects are masked from 'package:base':
##
##     backsolve, forwardsolve
## Loading required package: maps
## Warning: package 'maps' was built under R version 3.2.5
```

```r
library(SpatialTools)
```

```
## # This research was partially supported under NSF Grant ATM-0534173
```

```r
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.2.5
```

```r
library(GGally)
```

```
## Warning: package 'GGally' was built under R version 3.2.5
## Warning: replacing previous import by 'utils::capture.output' when loading
## 'GGally'
## Warning: replacing previous import by 'utils::head' when loading 'GGally'
## Warning: replacing previous import by 'utils::installed.packages' when
## loading 'GGally'
## Warning: replacing previous import by 'utils::str' when loading 'GGally'
```

Problem 1: Explain whether each scenario is a classification or regression prob- lem, and indicate whether we are most interested in inference or pre- diction. Finally, provide n and p.

(a) We collect a set of data on the top 500 firms in the US. For each firm we record profit, number of employees, industry and the CEO salary. We are interested in understanding which factors affect CEO salary.

Regression, inference, n = 500, p = 4

(b) We are considering launching a new product and wish to know whether it will be a success or a failure. We collect data on 20 similar products that were previously launched. For each prod- uct we have recorded whether it was a success or failure, price charged for the product, marketing budget, competition price, and ten other variables.

Classification Prediction n = 20, p = 14,

(c) We are interesting in predicting the % change in the US dollar in relation to the weekly changes in the world stock markets. Hence we collect weekly data for all of 2012. For each week we record the % change in the dollar, the % change in the US market, the % change in the British market, and the % hange in the German market.

Problem 2

Complete Exercise 3 from section 2.4 of the textbook (p. 52).

3. We now revisit the bias-variance decomposition.

(a) Provide a sketch of typical (squared) bias, variance, training er- ror, test error, and Bayes (or irreducible) error curves, on a sin- gle plot, as we go from less flexible statistical learning methods towards more flexible approaches. The x-axis should representthe amount of flexibility in the method, and the y-axis should represent the values for each curve. There should be five curves. Make sure to label each one.
(b) Explain why each of the five curves has the shape displayed in part (a).

i. bias is inversely related to flexibility as higher flexibility creates a closer fit

ii. variance - increases monotonically because increases in flexibility yield overfitting

iii. training error - decreases monotonically because increases in flexibility yield a closer fit

iv. test error - concave up curve because increase in flexibility yields a closer fit before it overfits

v. Bayes (irreducible) error - defines the lower limit, the test error is bounded below by the irreducible error due to variance in the error (epsilon) in the output values (0 <= value). When the training error is lower than the irreducible error, overfitting has taken place. The Bayes error rate is defined for classification problems and is determined by the ratio of data points which lie at the 'wrong' side of the decision boundary, (0 <= value < 1).

Problem 3

Complete Exercise 7 from section 2.4 of the textbook (p. 53).

```
Obs <- seq(1,6,length=6)
X1 <- c(0,2,0,0,-1,1)
X2 <- c(3,0,1,1,0,1)
X3 <- c(0,0,3,2,1,1)
Y <- c('Red', 'Red', 'Red', 'Green', 'Green', 'Red')
df <- data.frame(Obs, X1, X2, X3, Y)
(df)
```

```
##   Obs X1 X2 X3     Y
## 1   1  0  3  0   Red
## 2   2  2  0  0   Red
## 3   3  0  1  3   Red
## 4   4  0  1  2 Green
## 5   5 -1  0  1 Green
## 6   6  1  1  1   Red
```

```
coords_mat <- data.matrix(df[2:4])
coords_mat
```

```
##      X1 X2 X3
## [1,]  0  3  0
## [2,]  2  0  0
## [3,]  0  1  3
```

```
## [4,]  0  1  2
## [5,] -1  0  1
## [6,]  1  1  1
```

a)

```
dist <- NULL
for(i in 1:nrow(coords_mat)) {
    dist[i] <- dist(rbind(coords_mat[i,], c(0,0,0)))
}
df$dist <- dist

df
```

```
##   Obs X1 X2 X3     Y     dist
## 1   1  0  3  0   Red 3.000000
## 2   2  2  0  0   Red 2.000000
## 3   3  0  1  3   Red 3.162278
## 4   4  0  1  2 Green 2.236068
## 5   5 -1  0  1 Green 1.414214
## 6   6  1  1  1   Red 1.732051
```

b) If the only datapoint we care about is the one nearest neighbor, then the prediction will be Green (Obs 5)

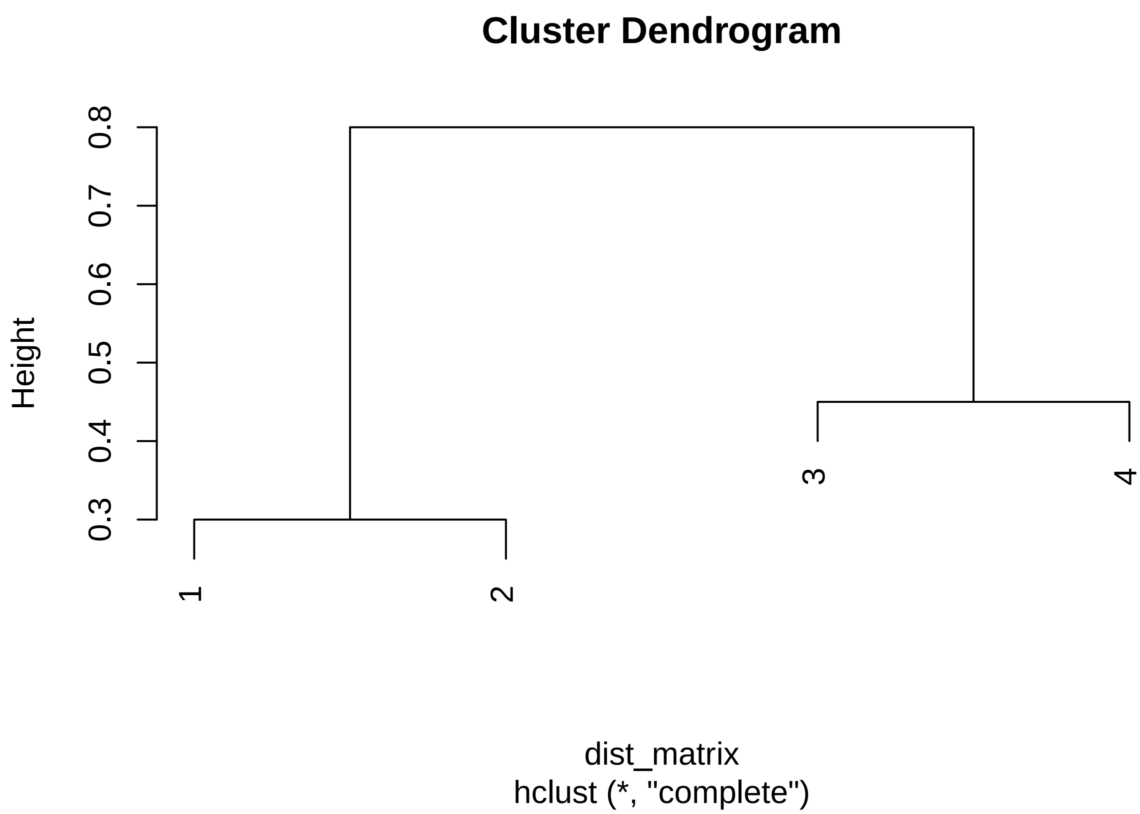
c) Obs 2, 5, 6 will be the closest 3 neighbors for [0,0,0], which corresponds with a Y of Red, Green and Red respectively; thus the prediction would be red.

d)

Problem 4: Exercise 1 (p. 413) a) K-Means Clustering: Prove equation 10.12 b)

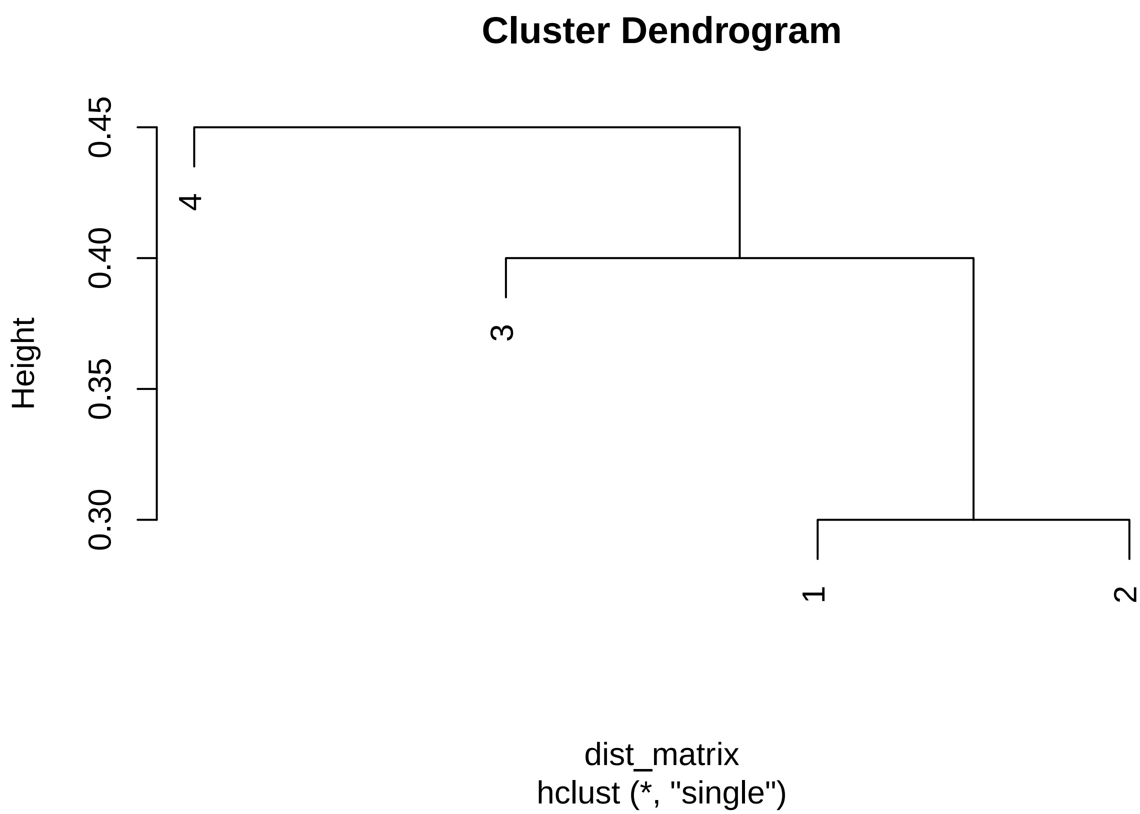
Problem 5: Exercise 2 (p. 413) For given Dissimilarity matrix, a) On the basis of this dissimilarity matrix, sketch the dendrogram that results from hierarchically clustering these four observa- tions using complete linkage. Be sure to indicate on the plot the height at which each fusion occurs, as well as the observations corresponding to each leaf in the dendrogram.

```
dist_matrix <- as.dist(matrix(c(0, 0.3, 0.4, 0.7,
                      0.3, 0, 0.5, 0.8,
                      0.4, 0.5, 0.0, 0.45,
                      0.7, 0.8, 0.45, 0.0), nrow=4))
plot(hclust(dist_matrix, method="complete"))
```



b)

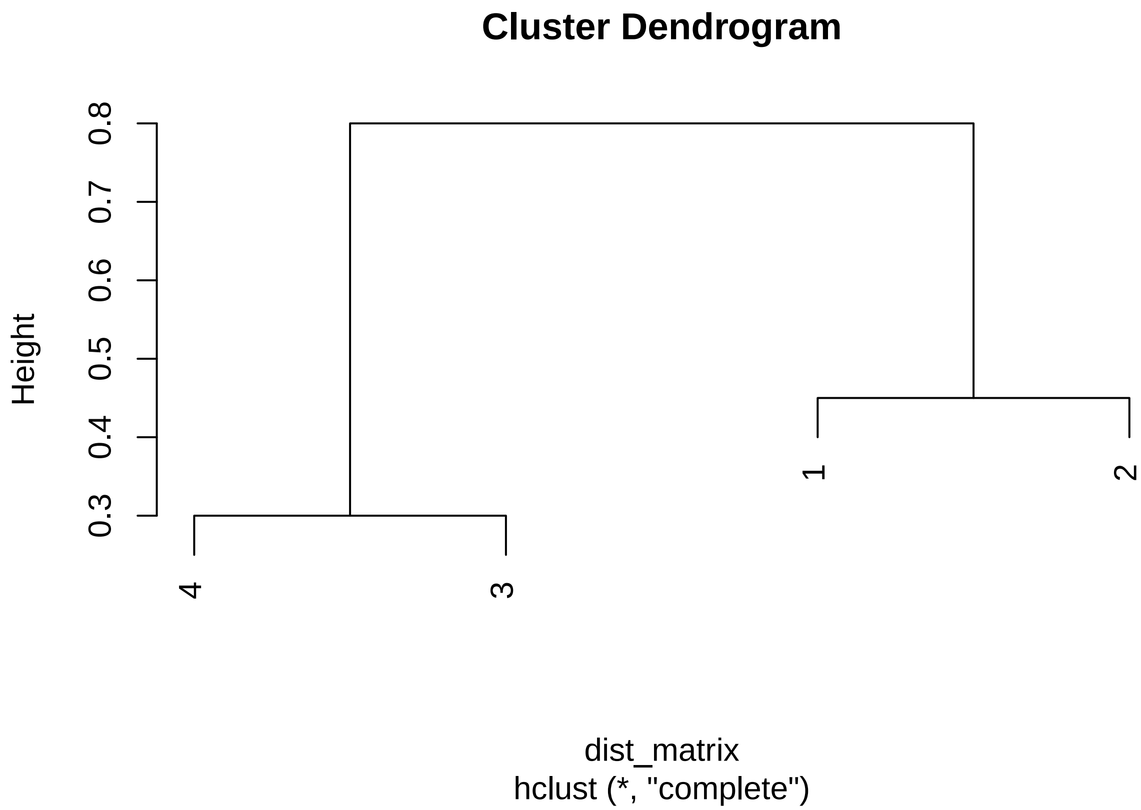
```r
plot(hclust(dist_matrix, method="single"))
```



c) Cluster1: Observations 1 and 2; Cluster2: Observations 3 and 4

d) Cluster1: Observations 1,2, and 3; Cluster2: Observation 4
e) The following dendrogram swaps positions of the two clusters without changing the meaning

```
plot(hclust(dist_matrix, method="complete"), labels=c(4,3,1,2))
```



Problem 6: Exercise 4 (p. 414) Suppose that for a particular data set, we perform hierarchical clus- tering using single linkage and using complete linkage. We obtain two dendrograms. (a) At a certain point on the single linkage dendrogram, the clus- ters {1,2,3} and {4,5} fuse. On the complete linkage dendro- gram, the clusters {1, 2, 3} and {4, 5} also fuse at a certain point. Which fusion will occur higher on the tree, or will they fuse at the same height, or is there not enough information to tell?

This question requires more information to answer and is dependent on both the organization of information as well as the dissimilarity measure (euclidian distance, correlation etc). Complete linkage joins on maximal intercluster dissimilarity, while single linknage joins on minimal intercluster dissimilarity; were these to be equal, then the two clusters in question would fuse at the same height. Otherwise, a dendrogram formed with complete linkage would fuse them at a greater height than a dendrogram formed with single linkage.

(b) At a certain point on the single linkage dendrogram, the clusters {5} and {6} fuse. On the complete linkage dendrogram, the clus- ters {5} and {6} also fuse at a certain point. Which fusion will occur higher on the tree, or will they fuse at the same height, or is there not enough information to tell? They would fuse at the same height because the choice of complete vs single linkage operates on observations between two clusters instead of the clustering of two individual observations that are not yet clustered.

Problem 7: Exercise 9 (p. 416)

```
data("USArrests")
names(USArrests)
```

```
## [1] "Murder"   "Assault"  "UrbanPop" "Rape"
```

```
dim(USArrests)
```

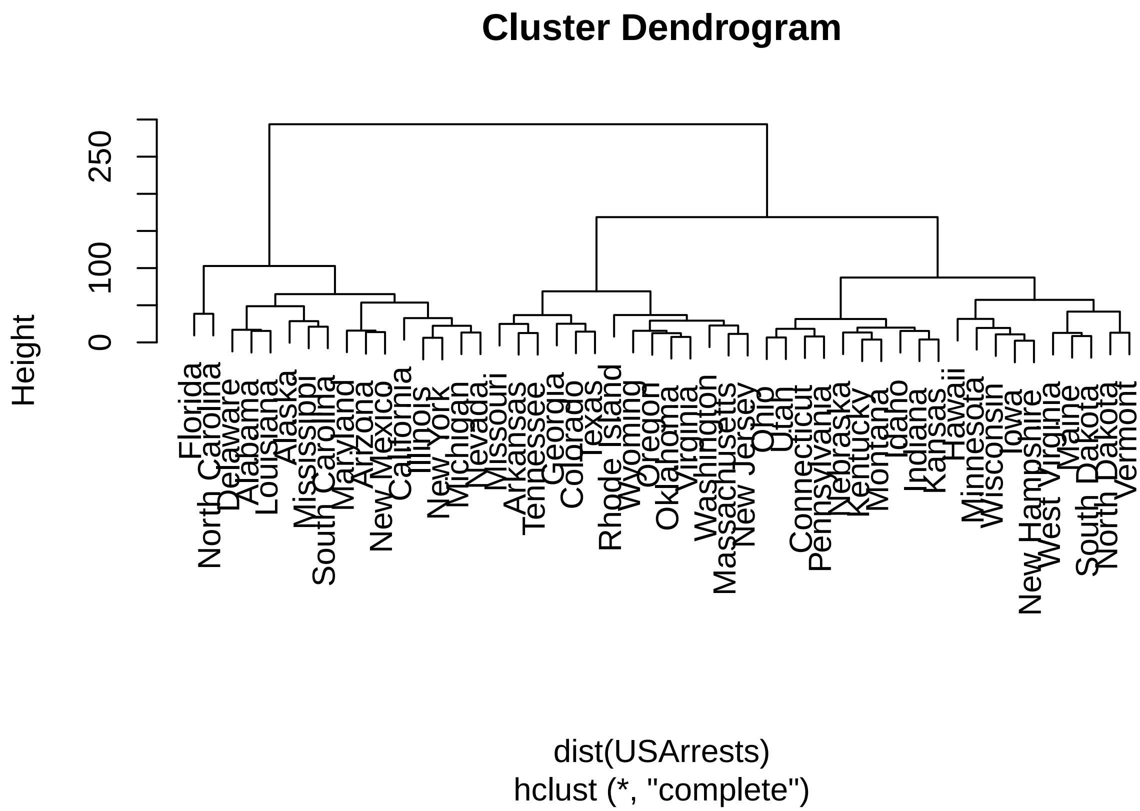
```
## [1] 50  4
```

```r
class(USArrests)
```

```
## [1] "data.frame"
```

a)

```r
cluster_USArrests <- hclust(dist(USArrests), method="complete")
plot(cluster_USArrests)
```



b)

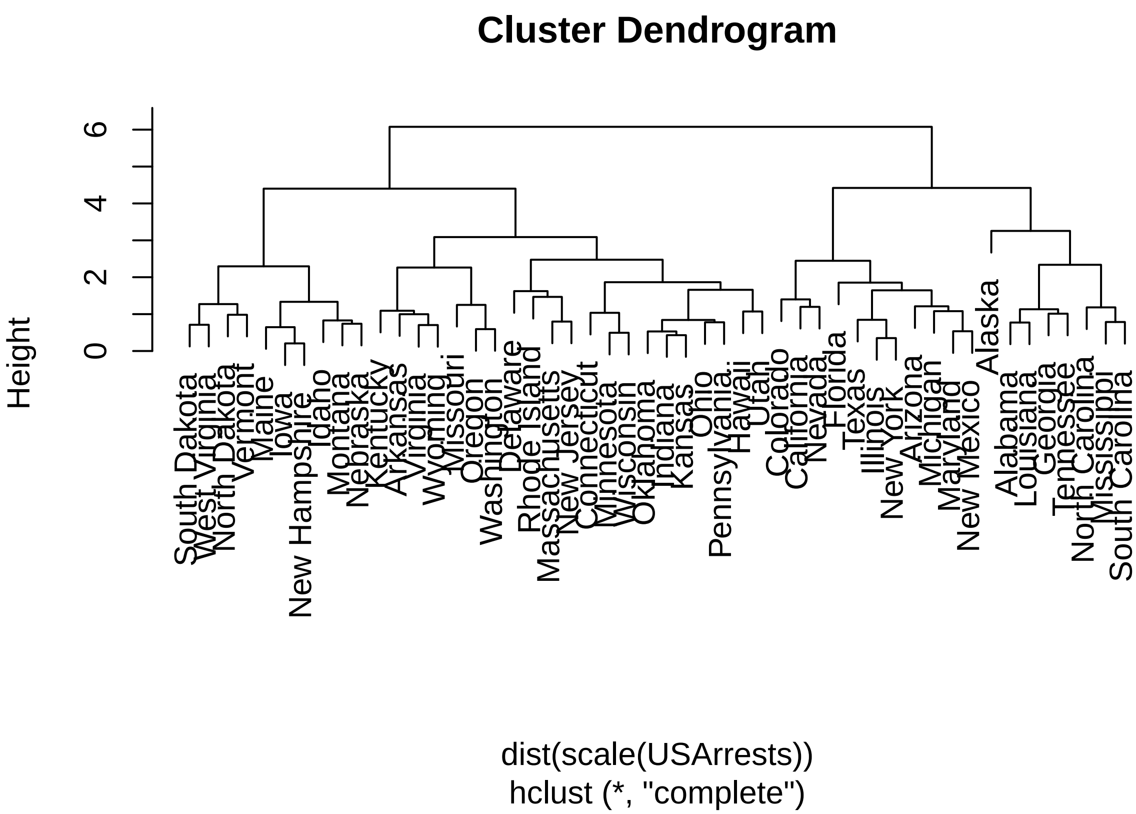
```r
cutree(cluster_USArrests, 3)
```

```
##        Alabama         Alaska        Arizona       Arkansas     California
##              1              1              1              2              1
##       Colorado    Connecticut       Delaware        Florida        Georgia
##              2              3              1              1              2
##         Hawaii          Idaho       Illinois        Indiana           Iowa
##              3              3              1              3              3
##         Kansas       Kentucky      Louisiana          Maine       Maryland
##              3              3              1              3              1
##  Massachusetts       Michigan      Minnesota    Mississippi       Missouri
##              2              1              3              1              2
##        Montana       Nebraska         Nevada  New Hampshire     New Jersey
##              3              3              1              3              2
##     New Mexico       New York North Carolina   North Dakota           Ohio
##              1              1              1              3              3
##       Oklahoma         Oregon   Pennsylvania   Rhode Island South Carolina
##              2              2              3              2              1
##   South Dakota      Tennessee          Texas           Utah        Vermont
```

```
##              3              2              2              3              3
##       Virginia     Washington  West Virginia      Wisconsin        Wyoming
##              2              2              3              3              2
```

c)

```r
cluster_USArrests_scaled = hclust(dist(scale(USArrests)), method="complete")
plot(cluster_USArrests_scaled)
```



d) Number of states in each cluster without scaling USArrests:

```r
table(cutree(cluster_USArrests, 3))
```

```
##
##  1  2  3
## 16 14 20
```

Number of states in each cluster after scaling USArrests:

```r
table(cutree(cluster_USArrests_scaled, 3))
```

```
##
##  1  2  3
##  8 11 31
```

The overall height and spread of the dendrogram was not dramatically altered after scaling the input dataset. The states that ended up in each 3 cluster did change. In general, scaling each variable vector to standardize variance makes sense. The variables in USArrests dataset have different units with different inherent variance. Units with a larger variance has a greater effect on euclidian distance, and thus have a greater influence on how clusters are formed.

Problem 8: Exercise 4 (p. 120)

4. I collect a set of data (n = 100 observations) containing a single predictor and a quantitative response. I then fit a linear regression model to the data, as well as a separate cubic regression, i.e. Y = 0 + 1X + 2X2 + 3X3 + .

(a) Suppose that the true relationship between X and Y is linear, i.e. Y = 0 + 1X + . Consider the training residual sum of squares (RSS) for the linear regression, and also the training RSS for the cubic regression. Would we expect one to be lower than the other, would we expect them to be the same, or is there not enough information to tell? Justify your answer. Adding more variables to the least squares equations always improves the fit to the training data; thus, the RSS to training data should decrease

(b) Answer (a) using test rather than training RSS. test RSS should decrease due to the overfitting and failing to generalize overfit model to test dataset

(c) Suppose that the truerelationshipbetweenXandYisnotlinear, but we don't know how far it is from linear. Consider the training RSS for the linear regression, and also the training RSS for the cubic regression. Would we expect one to be lower than the other, would we expect them to be the same, or is there not enough information to tell? Justify your answer. The increased flexibility from polynomial regression will lead to a better fit to training data over a linear regression.

(d) Answer (c) using test rather than training RSS. Since the true relationship is not known, there is not enough information to exactly tell whether test dataset RSS will be better with a polynomial fit;

Problem 9: Exercise 9 (p. 122). In parts (e) and (f), you need only try a few interactions and transformations.

```
Auto = na.omit(read.csv("Auto.csv", na.strings="?")) #massage out question marks and lists with missing
Auto
```

```
##      mpg cylinders displacement horsepower weight acceleration year origin
## 1   18.0         8        307.0        130   3504         12.0   70      1
## 2   15.0         8        350.0        165   3693         11.5   70      1
## 3   18.0         8        318.0        150   3436         11.0   70      1
## 4   16.0         8        304.0        150   3433         12.0   70      1
## 5   17.0         8        302.0        140   3449         10.5   70      1
## 6   15.0         8        429.0        198   4341         10.0   70      1
## 7   14.0         8        454.0        220   4354          9.0   70      1
## 8   14.0         8        440.0        215   4312          8.5   70      1
## 9   14.0         8        455.0        225   4425         10.0   70      1
## 10  15.0         8        390.0        190   3850          8.5   70      1
## 11  15.0         8        383.0        170   3563         10.0   70      1
## 12  14.0         8        340.0        160   3609          8.0   70      1
## 13  15.0         8        400.0        150   3761          9.5   70      1
## 14  14.0         8        455.0        225   3086         10.0   70      1
## 15  24.0         4        113.0         95   2372         15.0   70      3
## 16  22.0         6        198.0         95   2833         15.5   70      1
## 17  18.0         6        199.0         97   2774         15.5   70      1
## 18  21.0         6        200.0         85   2587         16.0   70      1
## 19  27.0         4         97.0         88   2130         14.5   70      3
## 20  26.0         4         97.0         46   1835         20.5   70      2
## 21  25.0         4        110.0         87   2672         17.5   70      2
## 22  24.0         4        107.0         90   2430         14.5   70      2
## 23  25.0         4        104.0         95   2375         17.5   70      2
## 24  26.0         4        121.0        113   2234         12.5   70      2
## 25  21.0         6        199.0         90   2648         15.0   70      1
## 26  10.0         8        360.0        215   4615         14.0   70      1
## 27  10.0         8        307.0        200   4376         15.0   70      1
## 28  11.0         8        318.0        210   4382         13.5   70      1
## 29   9.0         8        304.0        193   4732         18.5   70      1
```

```
## 30  27.0         4        97.0        88  2130         14.5  71     3
## 31  28.0         4       140.0        90  2264         15.5  71     1
## 32  25.0         4       113.0        95  2228         14.0  71     3
## 34  19.0         6       232.0       100  2634         13.0  71     1
## 35  16.0         6       225.0       105  3439         15.5  71     1
## 36  17.0         6       250.0       100  3329         15.5  71     1
## 37  19.0         6       250.0        88  3302         15.5  71     1
## 38  18.0         6       232.0       100  3288         15.5  71     1
## 39  14.0         8       350.0       165  4209         12.0  71     1
## 40  14.0         8       400.0       175  4464         11.5  71     1
## 41  14.0         8       351.0       153  4154         13.5  71     1
## 42  14.0         8       318.0       150  4096         13.0  71     1
## 43  12.0         8       383.0       180  4955         11.5  71     1
## 44  13.0         8       400.0       170  4746         12.0  71     1
## 45  13.0         8       400.0       175  5140         12.0  71     1
## 46  18.0         6       258.0       110  2962         13.5  71     1
## 47  22.0         4       140.0        72  2408         19.0  71     1
## 48  19.0         6       250.0       100  3282         15.0  71     1
## 49  18.0         6       250.0        88  3139         14.5  71     1
## 50  23.0         4       122.0        86  2220         14.0  71     1
## 51  28.0         4       116.0        90  2123         14.0  71     2
## 52  30.0         4        79.0        70  2074         19.5  71     2
## 53  30.0         4        88.0        76  2065         14.5  71     2
## 54  31.0         4        71.0        65  1773         19.0  71     3
## 55  35.0         4        72.0        69  1613         18.0  71     3
## 56  27.0         4        97.0        60  1834         19.0  71     2
## 57  26.0         4        91.0        70  1955         20.5  71     1
## 58  24.0         4       113.0        95  2278         15.5  72     3
## 59  25.0         4        97.5        80  2126         17.0  72     1
## 60  23.0         4        97.0        54  2254         23.5  72     2
## 61  20.0         4       140.0        90  2408         19.5  72     1
## 62  21.0         4       122.0        86  2226         16.5  72     1
## 63  13.0         8       350.0       165  4274         12.0  72     1
## 64  14.0         8       400.0       175  4385         12.0  72     1
## 65  15.0         8       318.0       150  4135         13.5  72     1
## 66  14.0         8       351.0       153  4129         13.0  72     1
## 67  17.0         8       304.0       150  3672         11.5  72     1
## 68  11.0         8       429.0       208  4633         11.0  72     1
## 69  13.0         8       350.0       155  4502         13.5  72     1
## 70  12.0         8       350.0       160  4456         13.5  72     1
## 71  13.0         8       400.0       190  4422         12.5  72     1
## 72  19.0         3        70.0        97  2330         13.5  72     3
## 73  15.0         8       304.0       150  3892         12.5  72     1
## 74  13.0         8       307.0       130  4098         14.0  72     1
## 75  13.0         8       302.0       140  4294         16.0  72     1
## 76  14.0         8       318.0       150  4077         14.0  72     1
## 77  18.0         4       121.0       112  2933         14.5  72     2
## 78  22.0         4       121.0        76  2511         18.0  72     2
## 79  21.0         4       120.0        87  2979         19.5  72     2
## 80  26.0         4        96.0        69  2189         18.0  72     2
## 81  22.0         4       122.0        86  2395         16.0  72     1
## 82  28.0         4        97.0        92  2288         17.0  72     3
## 83  23.0         4       120.0        97  2506         14.5  72     3
## 84  28.0         4        98.0        80  2164         15.0  72     1
```

```
## 85  27.0          4          97.0          88  2100           16.5  72       3
## 86  13.0          8         350.0         175  4100           13.0  73       1
## 87  14.0          8         304.0         150  3672           11.5  73       1
## 88  13.0          8         350.0         145  3988           13.0  73       1
## 89  14.0          8         302.0         137  4042           14.5  73       1
## 90  15.0          8         318.0         150  3777           12.5  73       1
## 91  12.0          8         429.0         198  4952           11.5  73       1
## 92  13.0          8         400.0         150  4464           12.0  73       1
## 93  13.0          8         351.0         158  4363           13.0  73       1
## 94  14.0          8         318.0         150  4237           14.5  73       1
## 95  13.0          8         440.0         215  4735           11.0  73       1
## 96  12.0          8         455.0         225  4951           11.0  73       1
## 97  13.0          8         360.0         175  3821           11.0  73       1
## 98  18.0          6         225.0         105  3121           16.5  73       1
## 99  16.0          6         250.0         100  3278           18.0  73       1
## 100 18.0          6         232.0         100  2945           16.0  73       1
## 101 18.0          6         250.0          88  3021           16.5  73       1
## 102 23.0          6         198.0          95  2904           16.0  73       1
## 103 26.0          4          97.0          46  1950           21.0  73       2
## 104 11.0          8         400.0         150  4997           14.0  73       1
## 105 12.0          8         400.0         167  4906           12.5  73       1
## 106 13.0          8         360.0         170  4654           13.0  73       1
## 107 12.0          8         350.0         180  4499           12.5  73       1
## 108 18.0          6         232.0         100  2789           15.0  73       1
## 109 20.0          4          97.0          88  2279           19.0  73       3
## 110 21.0          4         140.0          72  2401           19.5  73       1
## 111 22.0          4         108.0          94  2379           16.5  73       3
## 112 18.0          3          70.0          90  2124           13.5  73       3
## 113 19.0          4         122.0          85  2310           18.5  73       1
## 114 21.0          6         155.0         107  2472           14.0  73       1
## 115 26.0          4          98.0          90  2265           15.5  73       2
## 116 15.0          8         350.0         145  4082           13.0  73       1
## 117 16.0          8         400.0         230  4278            9.5  73       1
## 118 29.0          4          68.0          49  1867           19.5  73       2
## 119 24.0          4         116.0          75  2158           15.5  73       2
## 120 20.0          4         114.0          91  2582           14.0  73       2
## 121 19.0          4         121.0         112  2868           15.5  73       2
## 122 15.0          8         318.0         150  3399           11.0  73       1
## 123 24.0          4         121.0         110  2660           14.0  73       2
## 124 20.0          6         156.0         122  2807           13.5  73       3
## 125 11.0          8         350.0         180  3664           11.0  73       1
## 126 20.0          6         198.0          95  3102           16.5  74       1
## 128 19.0          6         232.0         100  2901           16.0  74       1
## 129 15.0          6         250.0         100  3336           17.0  74       1
## 130 31.0          4          79.0          67  1950           19.0  74       3
## 131 26.0          4         122.0          80  2451           16.5  74       1
## 132 32.0          4          71.0          65  1836           21.0  74       3
## 133 25.0          4         140.0          75  2542           17.0  74       1
## 134 16.0          6         250.0         100  3781           17.0  74       1
## 135 16.0          6         258.0         110  3632           18.0  74       1
## 136 18.0          6         225.0         105  3613           16.5  74       1
## 137 16.0          8         302.0         140  4141           14.0  74       1
## 138 13.0          8         350.0         150  4699           14.5  74       1
## 139 14.0          8         318.0         150  4457           13.5  74       1
```

```
## 140 14.0         8       302.0        140  4638         16.0  74       1
## 141 14.0         8       304.0        150  4257         15.5  74       1
## 142 29.0         4        98.0         83  2219         16.5  74       2
## 143 26.0         4        79.0         67  1963         15.5  74       2
## 144 26.0         4        97.0         78  2300         14.5  74       2
## 145 31.0         4        76.0         52  1649         16.5  74       3
## 146 32.0         4        83.0         61  2003         19.0  74       3
## 147 28.0         4        90.0         75  2125         14.5  74       1
## 148 24.0         4        90.0         75  2108         15.5  74       2
## 149 26.0         4       116.0         75  2246         14.0  74       2
## 150 24.0         4       120.0         97  2489         15.0  74       3
## 151 26.0         4       108.0         93  2391         15.5  74       3
## 152 31.0         4        79.0         67  2000         16.0  74       2
## 153 19.0         6       225.0         95  3264         16.0  75       1
## 154 18.0         6       250.0        105  3459         16.0  75       1
## 155 15.0         6       250.0         72  3432         21.0  75       1
## 156 15.0         6       250.0         72  3158         19.5  75       1
## 157 16.0         8       400.0        170  4668         11.5  75       1
## 158 15.0         8       350.0        145  4440         14.0  75       1
## 159 16.0         8       318.0        150  4498         14.5  75       1
## 160 14.0         8       351.0        148  4657         13.5  75       1
## 161 17.0         6       231.0        110  3907         21.0  75       1
## 162 16.0         6       250.0        105  3897         18.5  75       1
## 163 15.0         6       258.0        110  3730         19.0  75       1
## 164 18.0         6       225.0         95  3785         19.0  75       1
## 165 21.0         6       231.0        110  3039         15.0  75       1
## 166 20.0         8       262.0        110  3221         13.5  75       1
## 167 13.0         8       302.0        129  3169         12.0  75       1
## 168 29.0         4        97.0         75  2171         16.0  75       3
## 169 23.0         4       140.0         83  2639         17.0  75       1
## 170 20.0         6       232.0        100  2914         16.0  75       1
## 171 23.0         4       140.0         78  2592         18.5  75       1
## 172 24.0         4       134.0         96  2702         13.5  75       3
## 173 25.0         4        90.0         71  2223         16.5  75       2
## 174 24.0         4       119.0         97  2545         17.0  75       3
## 175 18.0         6       171.0         97  2984         14.5  75       1
## 176 29.0         4        90.0         70  1937         14.0  75       2
## 177 19.0         6       232.0         90  3211         17.0  75       1
## 178 23.0         4       115.0         95  2694         15.0  75       2
## 179 23.0         4       120.0         88  2957         17.0  75       2
## 180 22.0         4       121.0         98  2945         14.5  75       2
## 181 25.0         4       121.0        115  2671         13.5  75       2
## 182 33.0         4        91.0         53  1795         17.5  75       3
## 183 28.0         4       107.0         86  2464         15.5  76       2
## 184 25.0         4       116.0         81  2220         16.9  76       2
## 185 25.0         4       140.0         92  2572         14.9  76       1
## 186 26.0         4        98.0         79  2255         17.7  76       1
## 187 27.0         4       101.0         83  2202         15.3  76       2
## 188 17.5         8       305.0        140  4215         13.0  76       1
## 189 16.0         8       318.0        150  4190         13.0  76       1
## 190 15.5         8       304.0        120  3962         13.9  76       1
## 191 14.5         8       351.0        152  4215         12.8  76       1
## 192 22.0         6       225.0        100  3233         15.4  76       1
## 193 22.0         6       250.0        105  3353         14.5  76       1
```

```
## 194 24.0         6        200.0        81   3012         17.6   76       1
## 195 22.5         6        232.0        90   3085         17.6   76       1
## 196 29.0         4         85.0        52   2035         22.2   76       1
## 197 24.5         4         98.0        60   2164         22.1   76       1
## 198 29.0         4         90.0        70   1937         14.2   76       2
## 199 33.0         4         91.0        53   1795         17.4   76       3
## 200 20.0         6        225.0       100   3651         17.7   76       1
## 201 18.0         6        250.0        78   3574         21.0   76       1
## 202 18.5         6        250.0       110   3645         16.2   76       1
## 203 17.5         6        258.0        95   3193         17.8   76       1
## 204 29.5         4         97.0        71   1825         12.2   76       2
## 205 32.0         4         85.0        70   1990         17.0   76       3
## 206 28.0         4         97.0        75   2155         16.4   76       3
## 207 26.5         4        140.0        72   2565         13.6   76       1
## 208 20.0         4        130.0       102   3150         15.7   76       2
## 209 13.0         8        318.0       150   3940         13.2   76       1
## 210 19.0         4        120.0        88   3270         21.9   76       2
## 211 19.0         6        156.0       108   2930         15.5   76       3
## 212 16.5         6        168.0       120   3820         16.7   76       2
## 213 16.5         8        350.0       180   4380         12.1   76       1
## 214 13.0         8        350.0       145   4055         12.0   76       1
## 215 13.0         8        302.0       130   3870         15.0   76       1
## 216 13.0         8        318.0       150   3755         14.0   76       1
## 217 31.5         4         98.0        68   2045         18.5   77       3
## 218 30.0         4        111.0        80   2155         14.8   77       1
## 219 36.0         4         79.0        58   1825         18.6   77       2
## 220 25.5         4        122.0        96   2300         15.5   77       1
## 221 33.5         4         85.0        70   1945         16.8   77       3
## 222 17.5         8        305.0       145   3880         12.5   77       1
## 223 17.0         8        260.0       110   4060         19.0   77       1
## 224 15.5         8        318.0       145   4140         13.7   77       1
## 225 15.0         8        302.0       130   4295         14.9   77       1
## 226 17.5         6        250.0       110   3520         16.4   77       1
## 227 20.5         6        231.0       105   3425         16.9   77       1
## 228 19.0         6        225.0       100   3630         17.7   77       1
## 229 18.5         6        250.0        98   3525         19.0   77       1
## 230 16.0         8        400.0       180   4220         11.1   77       1
## 231 15.5         8        350.0       170   4165         11.4   77       1
## 232 15.5         8        400.0       190   4325         12.2   77       1
## 233 16.0         8        351.0       149   4335         14.5   77       1
## 234 29.0         4         97.0        78   1940         14.5   77       2
## 235 24.5         4        151.0        88   2740         16.0   77       1
## 236 26.0         4         97.0        75   2265         18.2   77       3
## 237 25.5         4        140.0        89   2755         15.8   77       1
## 238 30.5         4         98.0        63   2051         17.0   77       1
## 239 33.5         4         98.0        83   2075         15.9   77       1
## 240 30.0         4         97.0        67   1985         16.4   77       3
## 241 30.5         4         97.0        78   2190         14.1   77       2
## 242 22.0         6        146.0        97   2815         14.5   77       3
## 243 21.5         4        121.0       110   2600         12.8   77       2
## 244 21.5         3         80.0       110   2720         13.5   77       3
## 245 43.1         4         90.0        48   1985         21.5   78       2
## 246 36.1         4         98.0        66   1800         14.4   78       1
## 247 32.8         4         78.0        52   1985         19.4   78       3
```

```
## 248 39.4           4          85.0         70   2070          18.6   78       3
## 249 36.1           4          91.0         60   1800          16.4   78       3
## 250 19.9           8         260.0        110   3365          15.5   78       1
## 251 19.4           8         318.0        140   3735          13.2   78       1
## 252 20.2           8         302.0        139   3570          12.8   78       1
## 253 19.2           6         231.0        105   3535          19.2   78       1
## 254 20.5           6         200.0         95   3155          18.2   78       1
## 255 20.2           6         200.0         85   2965          15.8   78       1
## 256 25.1           4         140.0         88   2720          15.4   78       1
## 257 20.5           6         225.0        100   3430          17.2   78       1
## 258 19.4           6         232.0         90   3210          17.2   78       1
## 259 20.6           6         231.0        105   3380          15.8   78       1
## 260 20.8           6         200.0         85   3070          16.7   78       1
## 261 18.6           6         225.0        110   3620          18.7   78       1
## 262 18.1           6         258.0        120   3410          15.1   78       1
## 263 19.2           8         305.0        145   3425          13.2   78       1
## 264 17.7           6         231.0        165   3445          13.4   78       1
## 265 18.1           8         302.0        139   3205          11.2   78       1
## 266 17.5           8         318.0        140   4080          13.7   78       1
## 267 30.0           4          98.0         68   2155          16.5   78       1
## 268 27.5           4         134.0         95   2560          14.2   78       3
## 269 27.2           4         119.0         97   2300          14.7   78       3
## 270 30.9           4         105.0         75   2230          14.5   78       1
## 271 21.1           4         134.0         95   2515          14.8   78       3
## 272 23.2           4         156.0        105   2745          16.7   78       1
## 273 23.8           4         151.0         85   2855          17.6   78       1
## 274 23.9           4         119.0         97   2405          14.9   78       3
## 275 20.3           5         131.0        103   2830          15.9   78       2
## 276 17.0           6         163.0        125   3140          13.6   78       2
## 277 21.6           4         121.0        115   2795          15.7   78       2
## 278 16.2           6         163.0        133   3410          15.8   78       2
## 279 31.5           4          89.0         71   1990          14.9   78       2
## 280 29.5           4          98.0         68   2135          16.6   78       3
## 281 21.5           6         231.0        115   3245          15.4   79       1
## 282 19.8           6         200.0         85   2990          18.2   79       1
## 283 22.3           4         140.0         88   2890          17.3   79       1
## 284 20.2           6         232.0         90   3265          18.2   79       1
## 285 20.6           6         225.0        110   3360          16.6   79       1
## 286 17.0           8         305.0        130   3840          15.4   79       1
## 287 17.6           8         302.0        129   3725          13.4   79       1
## 288 16.5           8         351.0        138   3955          13.2   79       1
## 289 18.2           8         318.0        135   3830          15.2   79       1
## 290 16.9           8         350.0        155   4360          14.9   79       1
## 291 15.5           8         351.0        142   4054          14.3   79       1
## 292 19.2           8         267.0        125   3605          15.0   79       1
## 293 18.5           8         360.0        150   3940          13.0   79       1
## 294 31.9           4          89.0         71   1925          14.0   79       2
## 295 34.1           4          86.0         65   1975          15.2   79       3
## 296 35.7           4          98.0         80   1915          14.4   79       1
## 297 27.4           4         121.0         80   2670          15.0   79       1
## 298 25.4           5         183.0         77   3530          20.1   79       2
## 299 23.0           8         350.0        125   3900          17.4   79       1
## 300 27.2           4         141.0         71   3190          24.8   79       2
## 301 23.9           8         260.0         90   3420          22.2   79       1
```

```
## 302 34.2         4        105.0         70   2200         13.2   79      1
## 303 34.5         4        105.0         70   2150         14.9   79      1
## 304 31.8         4         85.0         65   2020         19.2   79      3
## 305 37.3         4         91.0         69   2130         14.7   79      2
## 306 28.4         4        151.0         90   2670         16.0   79      1
## 307 28.8         6        173.0        115   2595         11.3   79      1
## 308 26.8         6        173.0        115   2700         12.9   79      1
## 309 33.5         4        151.0         90   2556         13.2   79      1
## 310 41.5         4         98.0         76   2144         14.7   80      2
## 311 38.1         4         89.0         60   1968         18.8   80      3
## 312 32.1         4         98.0         70   2120         15.5   80      1
## 313 37.2         4         86.0         65   2019         16.4   80      3
## 314 28.0         4        151.0         90   2678         16.5   80      1
## 315 26.4         4        140.0         88   2870         18.1   80      1
## 316 24.3         4        151.0         90   3003         20.1   80      1
## 317 19.1         6        225.0         90   3381         18.7   80      1
## 318 34.3         4         97.0         78   2188         15.8   80      2
## 319 29.8         4        134.0         90   2711         15.5   80      3
## 320 31.3         4        120.0         75   2542         17.5   80      3
## 321 37.0         4        119.0         92   2434         15.0   80      3
## 322 32.2         4        108.0         75   2265         15.2   80      3
## 323 46.6         4         86.0         65   2110         17.9   80      3
## 324 27.9         4        156.0        105   2800         14.4   80      1
## 325 40.8         4         85.0         65   2110         19.2   80      3
## 326 44.3         4         90.0         48   2085         21.7   80      2
## 327 43.4         4         90.0         48   2335         23.7   80      2
## 328 36.4         5        121.0         67   2950         19.9   80      2
## 329 30.0         4        146.0         67   3250         21.8   80      2
## 330 44.6         4         91.0         67   1850         13.8   80      3
## 332 33.8         4         97.0         67   2145         18.0   80      3
## 333 29.8         4         89.0         62   1845         15.3   80      2
## 334 32.7         6        168.0        132   2910         11.4   80      3
## 335 23.7         3         70.0        100   2420         12.5   80      3
## 336 35.0         4        122.0         88   2500         15.1   80      2
## 338 32.4         4        107.0         72   2290         17.0   80      3
## 339 27.2         4        135.0         84   2490         15.7   81      1
## 340 26.6         4        151.0         84   2635         16.4   81      1
## 341 25.8         4        156.0         92   2620         14.4   81      1
## 342 23.5         6        173.0        110   2725         12.6   81      1
## 343 30.0         4        135.0         84   2385         12.9   81      1
## 344 39.1         4         79.0         58   1755         16.9   81      3
## 345 39.0         4         86.0         64   1875         16.4   81      1
## 346 35.1         4         81.0         60   1760         16.1   81      3
## 347 32.3         4         97.0         67   2065         17.8   81      3
## 348 37.0         4         85.0         65   1975         19.4   81      3
## 349 37.7         4         89.0         62   2050         17.3   81      3
## 350 34.1         4         91.0         68   1985         16.0   81      3
## 351 34.7         4        105.0         63   2215         14.9   81      1
## 352 34.4         4         98.0         65   2045         16.2   81      1
## 353 29.9         4         98.0         65   2380         20.7   81      1
## 354 33.0         4        105.0         74   2190         14.2   81      2
## 356 33.7         4        107.0         75   2210         14.4   81      3
## 357 32.4         4        108.0         75   2350         16.8   81      3
## 358 32.9         4        119.0        100   2615         14.8   81      3
```

```
## 359 31.6         4        120.0         74   2635         18.3   81       3
## 360 28.1         4        141.0         80   3230         20.4   81       2
## 361 30.7         6        145.0         76   3160         19.6   81       2
## 362 25.4         6        168.0        116   2900         12.6   81       3
## 363 24.2         6        146.0        120   2930         13.8   81       3
## 364 22.4         6        231.0        110   3415         15.8   81       1
## 365 26.6         8        350.0        105   3725         19.0   81       1
## 366 20.2         6        200.0         88   3060         17.1   81       1
## 367 17.6         6        225.0         85   3465         16.6   81       1
## 368 28.0         4        112.0         88   2605         19.6   82       1
## 369 27.0         4        112.0         88   2640         18.6   82       1
## 370 34.0         4        112.0         88   2395         18.0   82       1
## 371 31.0         4        112.0         85   2575         16.2   82       1
## 372 29.0         4        135.0         84   2525         16.0   82       1
## 373 27.0         4        151.0         90   2735         18.0   82       1
## 374 24.0         4        140.0         92   2865         16.4   82       1
## 375 36.0         4        105.0         74   1980         15.3   82       2
## 376 37.0         4         91.0         68   2025         18.2   82       3
## 377 31.0         4         91.0         68   1970         17.6   82       3
## 378 38.0         4        105.0         63   2125         14.7   82       1
## 379 36.0         4         98.0         70   2125         17.3   82       1
## 380 36.0         4        120.0         88   2160         14.5   82       3
## 381 36.0         4        107.0         75   2205         14.5   82       3
## 382 34.0         4        108.0         70   2245         16.9   82       3
## 383 38.0         4         91.0         67   1965         15.0   82       3
## 384 32.0         4         91.0         67   1965         15.7   82       3
## 385 38.0         4         91.0         67   1995         16.2   82       3
## 386 25.0         6        181.0        110   2945         16.4   82       1
## 387 38.0         6        262.0         85   3015         17.0   82       1
## 388 26.0         4        156.0         92   2585         14.5   82       1
## 389 22.0         6        232.0        112   2835         14.7   82       1
## 390 32.0         4        144.0         96   2665         13.9   82       3
## 391 36.0         4        135.0         84   2370         13.0   82       1
## 392 27.0         4        151.0         90   2950         17.3   82       1
## 393 27.0         4        140.0         86   2790         15.6   82       1
## 394 44.0         4         97.0         52   2130         24.6   82       2
## 395 32.0         4        135.0         84   2295         11.6   82       1
## 396 28.0         4        120.0         79   2625         18.6   82       1
## 397 31.0         4        119.0         82   2720         19.4   82       1
##                                     name
## 1              chevrolet chevelle malibu
## 2                      buick skylark 320
## 3                     plymouth satellite
## 4                          amc rebel sst
## 5                            ford torino
## 6                       ford galaxie 500
## 7                       chevrolet impala
## 8                      plymouth fury iii
## 9                       pontiac catalina
## 10                    amc ambassador dpl
## 11                   dodge challenger se
## 12                    plymouth 'cuda 340
## 13                 chevrolet monte carlo
## 14               buick estate wagon (sw)
```

```
## 15                 toyota corona mark ii
## 16                       plymouth duster
## 17                            amc hornet
## 18                         ford maverick
## 19                          datsun pl510
## 20         volkswagen 1131 deluxe sedan
## 21                           peugeot 504
## 22                           audi 100 ls
## 23                              saab 99e
## 24                              bmw 2002
## 25                           amc gremlin
## 26                             ford f250
## 27                             chevy c20
## 28                            dodge d200
## 29                              hi 1200d
## 30                          datsun pl510
## 31                   chevrolet vega 2300
## 32                         toyota corona
## 34                           amc gremlin
## 35            plymouth satellite custom
## 36            chevrolet chevelle malibu
## 37                       ford torino 500
## 38                           amc matador
## 39                      chevrolet impala
## 40            pontiac catalina brougham
## 41                      ford galaxie 500
## 42                     plymouth fury iii
## 43                     dodge monaco (sw)
## 44             ford country squire (sw)
## 45                   pontiac safari (sw)
## 46           amc hornet sportabout (sw)
## 47                  chevrolet vega (sw)
## 48                      pontiac firebird
## 49                          ford mustang
## 50                    mercury capri 2000
## 51                             opel 1900
## 52                           peugeot 304
## 53                             fiat 124b
## 54                  toyota corolla 1200
## 55                           datsun 1200
## 56                 volkswagen model 111
## 57                      plymouth cricket
## 58                 toyota corona hardtop
## 59                    dodge colt hardtop
## 60                     volkswagen type 3
## 61                        chevrolet vega
## 62                   ford pinto runabout
## 63                      chevrolet impala
## 64                      pontiac catalina
## 65                     plymouth fury iii
## 66                      ford galaxie 500
## 67                    amc ambassador sst
## 68                       mercury marquis
## 69                  buick lesabre custom
```

```
## 70           oldsmobile delta 88 royale
## 71               chrysler newport royal
## 72                      mazda rx2 coupe
## 73                     amc matador (sw)
## 74     chevrolet chevelle concours (sw)
## 75               ford gran torino (sw)
## 76       plymouth satellite custom (sw)
## 77                      volvo 145e (sw)
## 78                volkswagen 411 (sw)
## 79                     peugeot 504 (sw)
## 80                      renault 12 (sw)
## 81                      ford pinto (sw)
## 82                      datsun 510 (sw)
## 83          toyouta corona mark ii (sw)
## 84                      dodge colt (sw)
## 85             toyota corolla 1600 (sw)
## 86                    buick century 350
## 87                          amc matador
## 88                     chevrolet malibu
## 89                     ford gran torino
## 90                  dodge coronet custom
## 91             mercury marquis brougham
## 92            chevrolet caprice classic
## 93                             ford ltd
## 94             plymouth fury gran sedan
## 95         chrysler new yorker brougham
## 96             buick electra 225 custom
## 97              amc ambassador brougham
## 98                     plymouth valiant
## 99                 chevrolet nova custom
## 100                          amc hornet
## 101                       ford maverick
## 102                     plymouth duster
## 103             volkswagen super beetle
## 104                    chevrolet impala
## 105                        ford country
## 106               plymouth custom suburb
## 107             oldsmobile vista cruiser
## 108                          amc gremlin
## 109                       toyota carina
## 110                      chevrolet vega
## 111                          datsun 610
## 112                           maxda rx3
## 113                          ford pinto
## 114                    mercury capri v6
## 115                 fiat 124 sport coupe
## 116              chevrolet monte carlo s
## 117                   pontiac grand prix
## 118                            fiat 128
## 119                          opel manta
## 120                          audi 100ls
## 121                         volvo 144ea
## 122                    dodge dart custom
## 123                           saab 99le
```

```
## 124                         toyota mark ii
## 125                       oldsmobile omega
## 126                        plymouth duster
## 128                             amc hornet
## 129                         chevrolet nova
## 130                            datsun b210
## 131                             ford pinto
## 132                    toyota corolla 1200
## 133                         chevrolet vega
## 134    chevrolet chevelle malibu classic
## 135                            amc matador
## 136           plymouth satellite sebring
## 137                       ford gran torino
## 138             buick century luxus (sw)
## 139            dodge coronet custom (sw)
## 140               ford gran torino (sw)
## 141                      amc matador (sw)
## 142                              audi fox
## 143                     volkswagen dasher
## 144                            opel manta
## 145                         toyota corona
## 146                            datsun 710
## 147                            dodge colt
## 148                              fiat 128
## 149                           fiat 124 tc
## 150                           honda civic
## 151                                subaru
## 152                             fiat x1.9
## 153               plymouth valiant custom
## 154                        chevrolet nova
## 155                       mercury monarch
## 156                         ford maverick
## 157                      pontiac catalina
## 158                     chevrolet bel air
## 159                    plymouth grand fury
## 160                              ford ltd
## 161                         buick century
## 162            chevroelt chevelle malibu
## 163                           amc matador
## 164                         plymouth fury
## 165                         buick skyhawk
## 166                   chevrolet monza 2+2
## 167                       ford mustang ii
## 168                        toyota corolla
## 169                            ford pinto
## 170                           amc gremlin
## 171                         pontiac astro
## 172                         toyota corona
## 173                     volkswagen dasher
## 174                            datsun 710
## 175                            ford pinto
## 176                     volkswagen rabbit
## 177                             amc pacer
## 178                            audi 100ls
```

```
## 179                          peugeot 504
## 180                          volvo 244dl
## 181                            saab 99le
## 182                     honda civic cvcc
## 183                             fiat 131
## 184                            opel 1900
## 185                             capri ii
## 186                           dodge colt
## 187                         renault 12tl
## 188    chevrolet chevelle malibu classic
## 189               dodge coronet brougham
## 190                          amc matador
## 191                     ford gran torino
## 192                     plymouth valiant
## 193                       chevrolet nova
## 194                        ford maverick
## 195                           amc hornet
## 196                   chevrolet chevette
## 197                      chevrolet woody
## 198                            vw rabbit
## 199                          honda civic
## 200                       dodge aspen se
## 201                    ford granada ghia
## 202                   pontiac ventura sj
## 203                        amc pacer d/l
## 204                    volkswagen rabbit
## 205                         datsun b-210
## 206                       toyota corolla
## 207                           ford pinto
## 208                            volvo 245
## 209           plymouth volare premier v8
## 210                          peugeot 504
## 211                       toyota mark ii
## 212                  mercedes-benz 280s
## 213                     cadillac seville
## 214                            chevy c10
## 215                            ford f108
## 216                           dodge d100
## 217                    honda accord cvcc
## 218              buick opel isuzu deluxe
## 219                        renault 5 gtl
## 220                    plymouth arrow gs
## 221                datsun f-10 hatchback
## 222              chevrolet caprice classic
## 223           oldsmobile cutlass supreme
## 224               dodge monaco brougham
## 225              mercury cougar brougham
## 226                  chevrolet concours
## 227                        buick skylark
## 228               plymouth volare custom
## 229                         ford granada
## 230                pontiac grand prix lj
## 231         chevrolet monte carlo landau
## 232                     chrysler cordoba
```

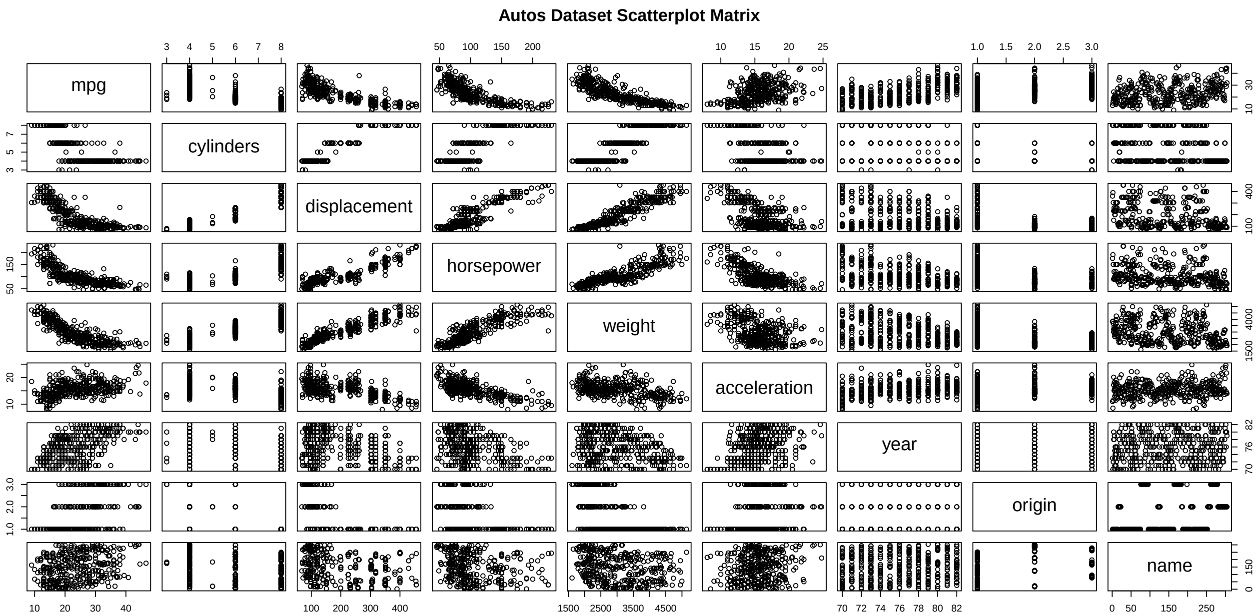
```
## 233                   ford thunderbird
## 234            volkswagen rabbit custom
## 235               pontiac sunbird coupe
## 236             toyota corolla liftback
## 237                 ford mustang ii 2+2
## 238                  chevrolet chevette
## 239                      dodge colt m/m
## 240                           subaru dl
## 241                   volkswagen dasher
## 242                          datsun 810
## 243                            bmw 320i
## 244                          mazda rx-4
## 245     volkswagen rabbit custom diesel
## 246                         ford fiesta
## 247                    mazda glc deluxe
## 248                      datsun b210 gx
## 249                    honda civic cvcc
## 250   oldsmobile cutlass salon brougham
## 251                      dodge diplomat
## 252                 mercury monarch ghia
## 253                  pontiac phoenix lj
## 254                    chevrolet malibu
## 255                ford fairmont (auto)
## 256                 ford fairmont (man)
## 257                     plymouth volare
## 258                         amc concord
## 259               buick century special
## 260                      mercury zephyr
## 261                         dodge aspen
## 262                     amc concord d/l
## 263        chevrolet monte carlo landau
## 264     buick regal sport coupe (turbo)
## 265                         ford futura
## 266                     dodge magnum xe
## 267                  chevrolet chevette
## 268                       toyota corona
## 269                          datsun 510
## 270                          dodge omni
## 271           toyota celica gt liftback
## 272                    plymouth sapporo
## 273              oldsmobile starfire sx
## 274                       datsun 200-sx
## 275                           audi 5000
## 276                         volvo 264gl
## 277                          saab 99gle
## 278                       peugeot 604sl
## 279                 volkswagen scirocco
## 280                     honda accord lx
## 281                   pontiac lemans v6
## 282                    mercury zephyr 6
## 283                     ford fairmont 4
## 284                    amc concord dl 6
## 285                       dodge aspen 6
## 286            chevrolet caprice classic
```

```
## 287                      ford ltd landau
## 288                mercury grand marquis
## 289                      dodge st. regis
## 290              buick estate wagon (sw)
## 291             ford country squire (sw)
## 292        chevrolet malibu classic (sw)
## 293 chrysler lebaron town @ country (sw)
## 294                     vw rabbit custom
## 295                     maxda glc deluxe
## 296          dodge colt hatchback custom
## 297                        amc spirit dl
## 298                   mercedes benz 300d
## 299                    cadillac eldorado
## 300                          peugeot 504
## 301    oldsmobile cutlass salon brougham
## 302                     plymouth horizon
## 303                 plymouth horizon tc3
## 304                           datsun 210
## 305                   fiat strada custom
## 306                buick skylark limited
## 307                   chevrolet citation
## 308            oldsmobile omega brougham
## 309                      pontiac phoenix
## 310                            vw rabbit
## 311                toyota corolla tercel
## 312                   chevrolet chevette
## 313                           datsun 310
## 314                   chevrolet citation
## 315                        ford fairmont
## 316                          amc concord
## 317                          dodge aspen
## 318                            audi 4000
## 319               toyota corona liftback
## 320                            mazda 626
## 321                 datsun 510 hatchback
## 322                       toyota corolla
## 323                            mazda glc
## 324                           dodge colt
## 325                           datsun 210
## 326                 vw rabbit c (diesel)
## 327                   vw dasher (diesel)
## 328                  audi 5000s (diesel)
## 329                   mercedes-benz 240d
## 330                  honda civic 1500 gl
## 332                            subaru dl
## 333                     vokswagen rabbit
## 334                        datsun 280-zx
## 335                        mazda rx-7 gs
## 336                    triumph tr7 coupe
## 338                         honda accord
## 339                     plymouth reliant
## 340                        buick skylark
## 341               dodge aries wagon (sw)
## 342                   chevrolet citation
```

```
## 343                       plymouth reliant
## 344                         toyota starlet
## 345                         plymouth champ
## 346                       honda civic 1300
## 347                                 subaru
## 348                         datsun 210 mpg
## 349                          toyota tercel
## 350                            mazda glc 4
## 351                     plymouth horizon 4
## 352                         ford escort 4w
## 353                         ford escort 2h
## 354                       volkswagen jetta
## 356                          honda prelude
## 357                         toyota corolla
## 358                           datsun 200sx
## 359                              mazda 626
## 360              peugeot 505s turbo diesel
## 361                           volvo diesel
## 362                        toyota cressida
## 363                      datsun 810 maxima
## 364                          buick century
## 365                  oldsmobile cutlass ls
## 366                        ford granada gl
## 367                 chrysler lebaron salon
## 368                     chevrolet cavalier
## 369               chevrolet cavalier wagon
## 370              chevrolet cavalier 2-door
## 371             pontiac j2000 se hatchback
## 372                         dodge aries se
## 373                        pontiac phoenix
## 374                   ford fairmont futura
## 375                    volkswagen rabbit l
## 376                     mazda glc custom l
## 377                       mazda glc custom
## 378                 plymouth horizon miser
## 379                         mercury lynx l
## 380                       nissan stanza xe
## 381                           honda accord
## 382                         toyota corolla
## 383                            honda civic
## 384                     honda civic (auto)
## 385                          datsun 310 gx
## 386                  buick century limited
## 387      oldsmobile cutlass ciera (diesel)
## 388             chrysler lebaron medallion
## 389                         ford granada l
## 390                       toyota celica gt
## 391                      dodge charger 2.2
## 392                       chevrolet camaro
## 393                        ford mustang gl
## 394                              vw pickup
## 395                          dodge rampage
## 396                            ford ranger
## 397                             chevy s-10
```

a)

```r
pairs(Auto, main="Autos Dataset Scatterplot Matrix")
```



b)

```r
cor(data.matrix(Auto[,-length(Auto)]))
```

```
##                     mpg  cylinders displacement horsepower     weight
## mpg           1.0000000 -0.7776175   -0.8051269 -0.7784268 -0.8322442
## cylinders    -0.7776175  1.0000000    0.9508233  0.8429834  0.8975273
## displacement -0.8051269  0.9508233    1.0000000  0.8972570  0.9329944
## horsepower   -0.7784268  0.8429834    0.8972570  1.0000000  0.8645377
## weight       -0.8322442  0.8975273    0.9329944  0.8645377  1.0000000
## acceleration  0.4233285 -0.5046834   -0.5438005 -0.6891955 -0.4168392
## year          0.5805410 -0.3456474   -0.3698552 -0.4163615 -0.3091199
## origin        0.5652088 -0.5689316   -0.6145351 -0.4551715 -0.5850054
##              acceleration       year     origin
## mpg             0.4233285  0.5805410  0.5652088
## cylinders      -0.5046834 -0.3456474 -0.5689316
## displacement   -0.5438005 -0.3698552 -0.6145351
## horsepower     -0.6891955 -0.4163615 -0.4551715
## weight         -0.4168392 -0.3091199 -0.5850054
## acceleration    1.0000000  0.2903161  0.2127458
## year            0.2903161  1.0000000  0.1815277
## origin          0.2127458  0.1815277  1.0000000
```

c)

```r
df <- Auto[,-length(Auto)]
# help(lm)
model <- lm(mpg ~ ., data=df)
summary(model)
```

```
##
## Call:
## lm(formula = mpg ~ ., data = df)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.5903 -2.1565 -0.1169  1.8690 13.0604
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -17.218435   4.644294  -3.707  0.00024 ***
## cylinders     -0.493376   0.323282  -1.526  0.12780
## displacement   0.019896   0.007515   2.647  0.00844 **
## horsepower    -0.016951   0.013787  -1.230  0.21963
## weight        -0.006474   0.000652  -9.929  < 2e-16 ***
## acceleration   0.080576   0.098845   0.815  0.41548
## year           0.750773   0.050973  14.729  < 2e-16 ***
## origin         1.426141   0.278136   5.127 4.67e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.328 on 384 degrees of freedom
## Multiple R-squared:  0.8215, Adjusted R-squared:  0.8182
## F-statistic: 252.4 on 7 and 384 DF,  p-value: < 2.2e-16
```

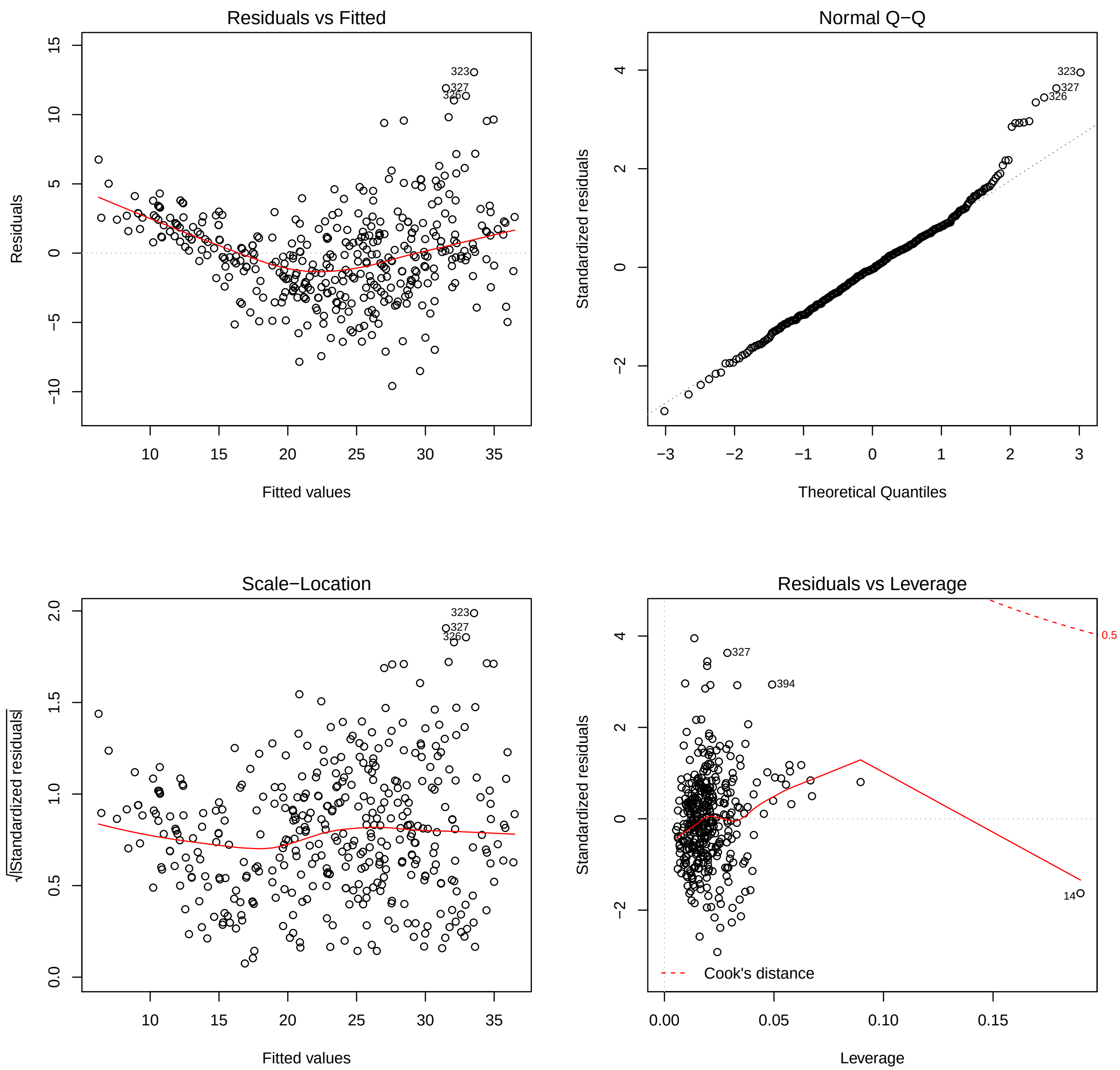
c.i) The model does predict a relationship between the predictors and response; R^2 at 0.903 is high, expressing that the model fits training data well. The F-statistic is high (considering the number of observations is much larger than our number of predictors) and p value is very low, thus the null-hypothesis can be rejected.

c.ii) Based on their low p values (highlighted by the 2-3 stars next to each p-value), displacement, weight, year and origin appear to have a statistically significant relationship to the response

c.iii) With each passing year, mpg increases by ~0.75 mpg

d)

```r
par(mfrow=c(2,2))
plot(model)
```




```r
par(mfrow=c(1,1)) # Change back to 1 x 1
```

The residuals vs fitted plot can be used to diagnose non-linear behavior in residuals. There appears to be a parabolic shape to the residuals curves, where non-linear relationship that was not explained by the model was left out in the residuals. The Normal Q-Q plot shows if residuals are normally distributed; there is some deviation in the points that have gained labels: 326, 327 and 323 The Scale-Location plot shows that the residuals are randomly spread and homoscedastic The residuals vs leverage plot shows that residual 14 has leverage, but all points are within cook's distance, meaning there aren't any particular residual that is highly influential to regression results

e) with some car knowledge that engine performance depends on an interaction of number of cylinders and total displacement, I believed it would be interesting to see whether this can influence our model:

```r
model_interaction <- lm(mpg~cylinders*displacement, data=df)
summary(model_interaction)
```

```
##
## Call:
```

```
## lm(formula = mpg ~ cylinders * displacement, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.0432  -2.4308  -0.2263   2.2048  20.9051
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)            48.22040    2.34712  20.545  < 2e-16 ***
## cylinders              -2.41838    0.53456  -4.524 8.08e-06 ***
## displacement           -0.13436    0.01615  -8.321 1.50e-15 ***
## cylinders:displacement  0.01182    0.00207   5.711 2.24e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.454 on 388 degrees of freedom
## Multiple R-squared:  0.6769, Adjusted R-squared:  0.6744
## F-statistic:   271 on 3 and 388 DF,  p-value: < 2.2e-16
```

Cylinders, displacement as well as their interaction proved to be very statistically significant in predicting mpg, which makes sense from a mechanical standpoint; more cylinders should allow for a more stable engine and higher RPMs, coupled with greater displacement fuel use can increase very quickly

Next I wondered whether there were strong interactions between year and origin, i.e. whether some countries of origin made great improvements overtime or vice veresa

```r
model_interaction <- lm(mpg~year*origin, data=df)
summary(model_interaction)
```

```
##
## Call:
## lm(formula = mpg ~ year * origin, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.3141  -3.7120  -0.6513   3.3621  15.5859
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -83.3809    12.0000  -6.948 1.57e-11 ***
## year          1.3089     0.1576   8.305 1.68e-15 ***
## origin       17.3752     6.8325   2.543   0.0114 *
## year:origin  -0.1663     0.0889  -1.871   0.0621 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.199 on 388 degrees of freedom
## Multiple R-squared:  0.5596, Adjusted R-squared:  0.5562
## F-statistic: 164.4 on 3 and 388 DF,  p-value: < 2.2e-16
```

However, here the significance is rather weak

Finally, knowing that acceleration can have a very high penalty on fuel usage, and that horsepower is a function of RPM:

```r
model_interaction <- lm(mpg~acceleration*horsepower, data=df)
summary(model_interaction)
```

```
##
## Call:
## lm(formula = mpg ~ acceleration * horsepower, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.3442  -2.7324  -0.4049   2.4210  15.8840
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             33.512440   3.420187   9.798  < 2e-16 ***
## acceleration             0.800296   0.211899   3.777 0.000184 ***
## horsepower               0.017590   0.027425   0.641 0.521664
## acceleration:horsepower -0.015698   0.002003  -7.838 4.45e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.426 on 388 degrees of freedom
## Multiple R-squared:  0.6809, Adjusted R-squared:  0.6784
## F-statistic: 275.9 on 3 and 388 DF,  p-value: < 2.2e-16
```

It turns out that a fast moving, high revving powerhouse consumes a ton of fuel: the interaction is statistically significacnt.

f)

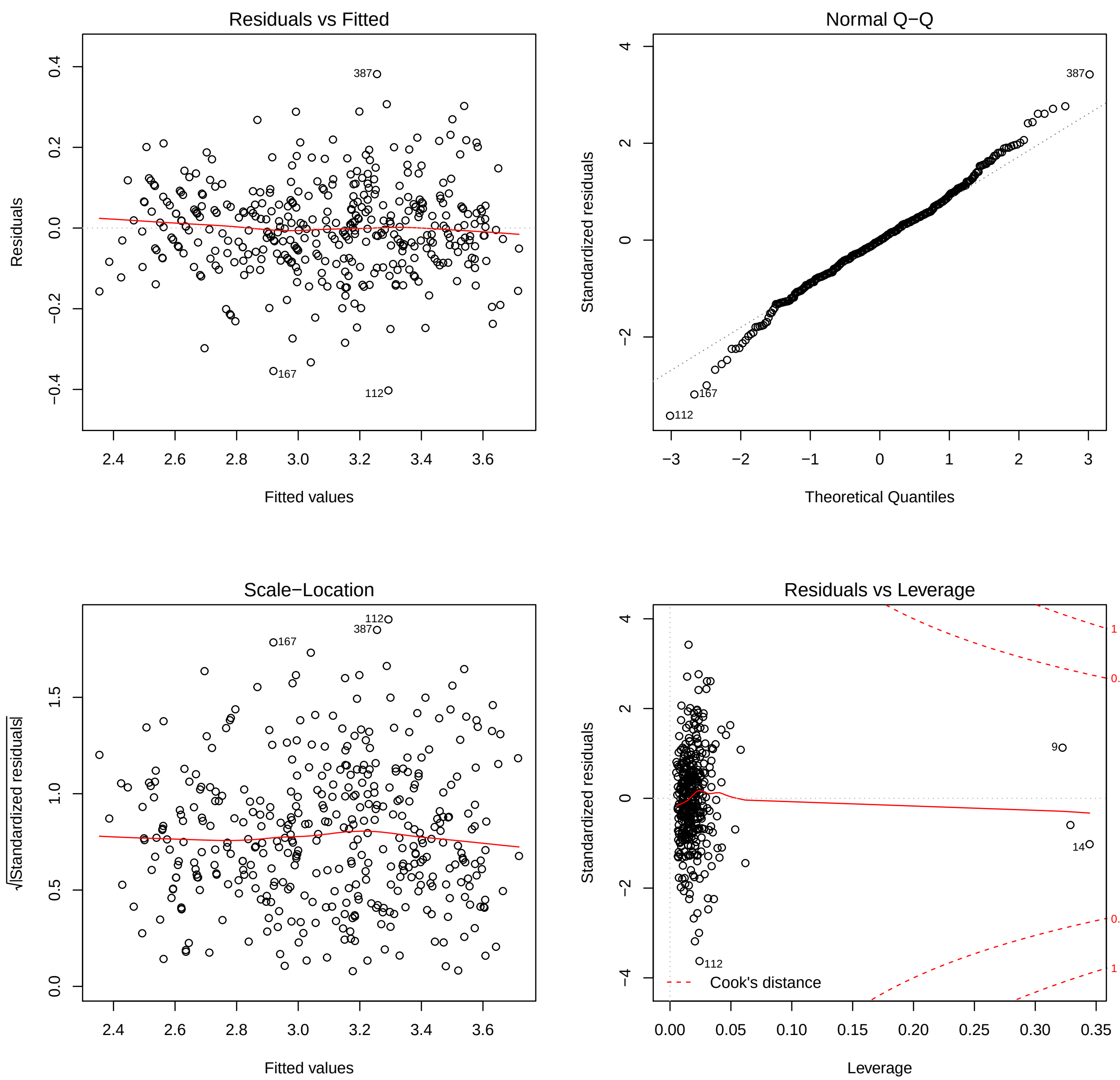
```r
model_transform <- lm(log(mpg)~cylinders+exp(displacement)+log(horsepower)+log(weight)+log(acceleration)
summary(model_transform)
```

```
##
## Call:
## lm(formula = log(mpg) ~ cylinders + exp(displacement) + log(horsepower) +
##     log(weight) + log(acceleration) + I(year^2) + origin, data = Auto)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.40259 -0.07022 -0.00125  0.06176  0.38161
##
## Coefficients:
##                     Estimate  Std. Error t value Pr(>|t|)
## (Intercept)        8.278e+00   3.461e-01  23.921  < 2e-16 ***
## cylinders         -1.515e-02   7.799e-03  -1.942  0.05284 .
## exp(displacement)  1.922e-199  0.000e+00     Inf  < 2e-16 ***
## log(horsepower)   -3.069e-01   5.805e-02  -5.287 2.09e-07 ***
## log(weight)       -5.475e-01   6.743e-02  -8.119 6.46e-15 ***
## log(acceleration) -1.846e-01   5.754e-02  -3.208  0.00145 **
## I(year^2)          1.959e-04   1.127e-05  17.385  < 2e-16 ***
## origin             2.401e-02   9.074e-03   2.646  0.00847 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1124 on 384 degrees of freedom
## Multiple R-squared:  0.8927, Adjusted R-squared:  0.8908
## F-statistic: 456.5 on 7 and 384 DF,  p-value: < 2.2e-16
```

Using a log transform of mpg proved a much better fit on its own; making additional transforms on predictor vectors, especially exp() on displacement and log() on horsepower, weight and acceleration further improved adjusted R^2 and reduced RSE

And now looking at the diagnostic plots:

```r
par(mfrow=c(2,2))
plot(model_transform)
```




```r
par(mfrow=c(1,1)) # Change back to 1 x 1
```

We see that with the aforementioned transforms on our response and predictor variables, the non-linear patterns on residuals greatly decreased (on the residuals vs fitted plot). Deviation of residuals from the straight line on Normal Q-Q plot also decreased, signifying more normally distributed residuals

Problem 10: Exercise 14 (p. 125) a)

```r
set.seed(1)
x1 = runif(100)
x2 = 0.5 * x1 + rnorm(100)/10
```

```r
y = 2 + 2*x1 + 0.3*x2 + rnorm(100)
```

Form of the linear model:

$$Y = 2 + 2X_1 + 0.3X_2 + \varepsilon$$

$$B_0 = 2$$

$$B_1 = 2$$

$$B_3 = 0.3$$

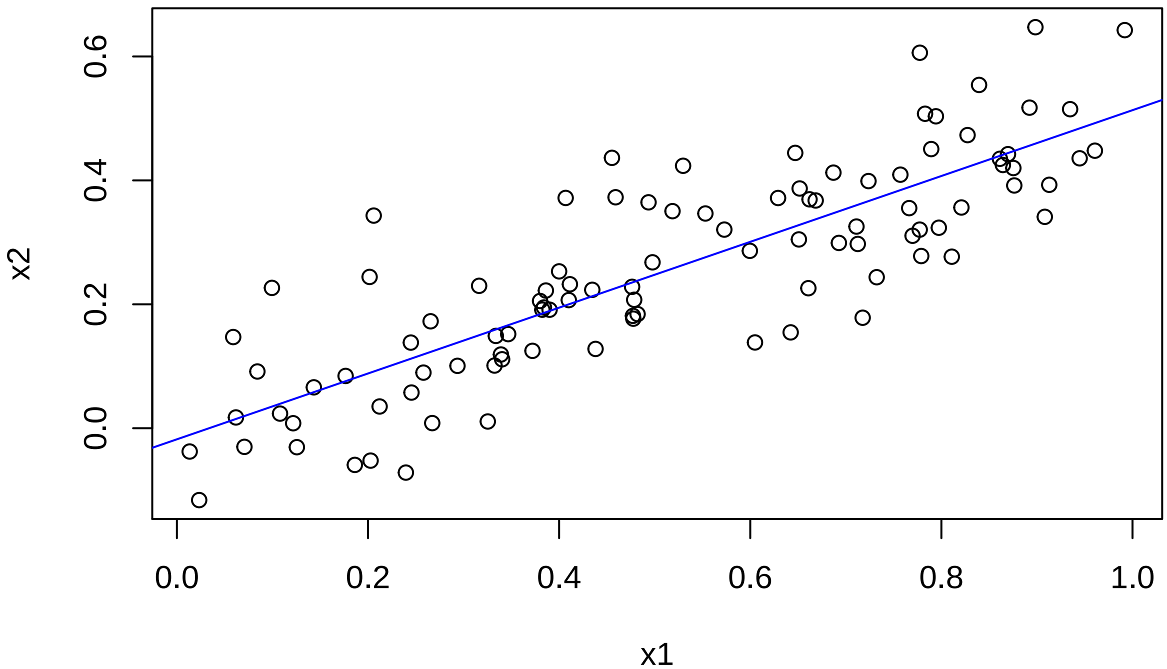
Figure 1: problem 10a

b)

```r
print(cor(x1, x2))
```

```
## [1] 0.8351212
```

```r
model = lm(x2~x1)
plot(x1, x2)
abline(model, col='blue')
```



c)