

DAT 500 S

DuraiSundaramoorthi

```
knitr::opts_chunk$set(echo = TRUE)
```

2.3.1 Basic Commands

Basic Commands Chunk 1

```
x = c(1, 3, 2, 5)
x
```

```
## [1] 1 3 2 5
```

```
x = c(1, 6, 2)
x
```

```
## [1] 1 6 2
```

```
y = c(1, 4, 3)
y
```

```
## [1] 1 4 3
```

```
length(x)
```

```
## [1] 3
```

```
length(y)
```

```
## [1] 3
```

```
x+y
```

```
## [1] 2 10 5
```

```
ls()
```

```
## [1] "x" "y"
```

```
rm(x,y)

rm(list = ls())

ls()
```

```
## character(0)
```

Basic Commands Chunk 2

```
?matrix

x = matrix(data = c(1,2,3,4), nrow = 2, ncol = 2)
x
```

```
##      [,1] [,2]
## [1,]    1    3
## [2,]    2    4
```

```
x = matrix(c(1,2,3,4),2,2)
x
```

```
##      [,1] [,2]
## [1,]    1    3
## [2,]    2    4
```

```
x = matrix(c(1,2,3,4),2,2,byrow = TRUE)
x
```

```
##      [,1] [,2]
## [1,]    1    2
## [2,]    3    4
```

```
sqrt(x)
```

```
##      [,1]      [,2]
## [1,] 1.000000 1.414214
## [2,] 1.732051 2.000000
```

Basic Commands Chunk 3

```
set.seed(1)

x = rnorm(50)

x
```

```
## [1] -0.62645381  0.18364332 -0.83562861  1.59528080  0.32950777 -0.82046838
## [7]  0.48742905  0.73832471  0.57578135 -0.30538839  1.51178117  0.38984324
## [13] -0.62124058 -2.21469989  1.12493092 -0.04493361 -0.01619026  0.94383621
## [19]  0.82122120  0.59390132  0.91897737  0.78213630  0.07456498 -1.98935170
## [25]  0.61982575 -0.05612874 -0.15579551 -1.47075238 -0.47815006  0.41794156
## [31]  1.35867955 -0.10278773  0.38767161 -0.05380504 -1.37705956 -0.41499456
## [37] -0.39428995 -0.05931340  1.10002537  0.76317575 -0.16452360 -0.25336168
## [43]  0.69696338  0.55666320 -0.68875569 -0.70749516  0.36458196  0.76853292
## [49] -0.11234621  0.88110773
```

```
mean(x)
```

```
## [1] 0.1004483
```

```
sd(x)
```

```
## [1] 0.8313939
```

Basic Commands Chunk 4

```
y = x + rnorm(50, mean=50, sd=0.1)
y
```

```
## [1] 49.41336 50.12244 49.19848 51.48234 50.47281 49.37757 50.45071 50.63391
## [9] 50.63275 49.68111 51.75194 50.38592 49.44773 47.78810 51.05060 49.97395
## [17] 49.80331 51.09039 50.83655 50.81116 50.96653 50.71114 50.13564 47.91724
## [25] 50.49446 49.97302 49.79988 48.52936 49.52928 50.35899 51.30181 49.88369
## [33] 50.50548 49.79384 48.68234 49.61830 49.71202 49.91027 51.13703 50.78989
## [41] 49.78122 49.86743 50.81300 50.62668 49.46993 49.34835 50.23692 50.71121
## [49] 49.76519 50.83377
```

```
cor(x,y)
```

```
## [1] 0.9932279
```

Basic Commands Chunk 5

```
set.seed(3)
```

```
y = rnorm(100)
mean(y)
```

```
## [1] 0.01103557
```

```
var(y)
```

```
## [1] 0.7328675
```

```
sqrt(var(y))
```

```
## [1] 0.8560768
```

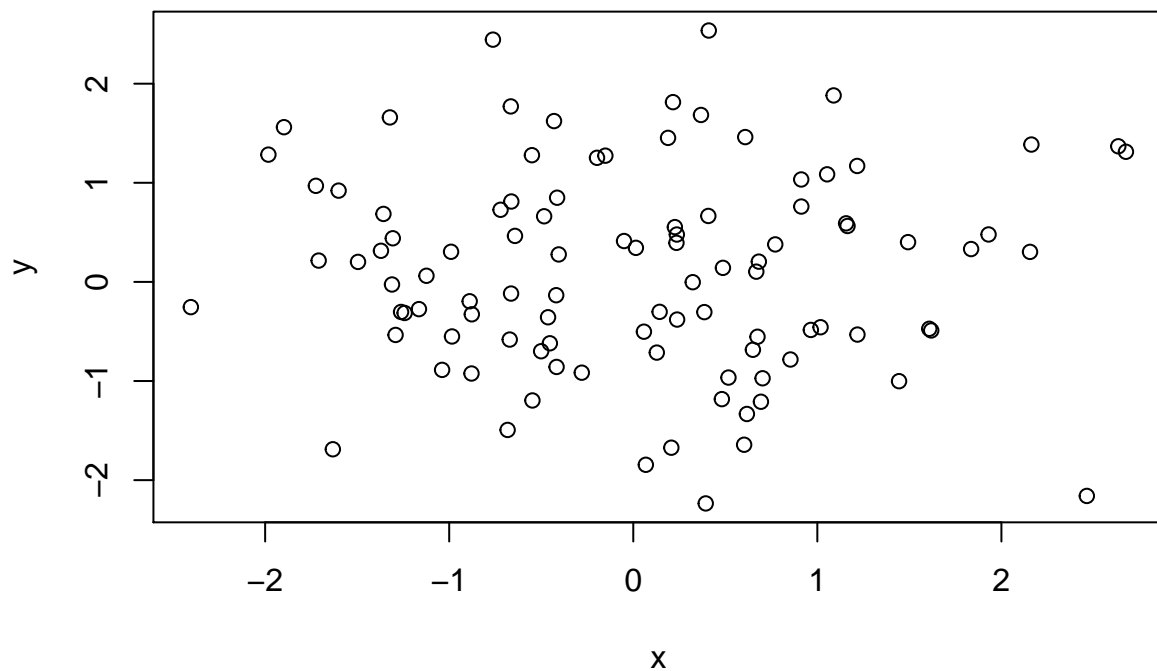
```
sd(y)
```

```
## [1] 0.8560768
```

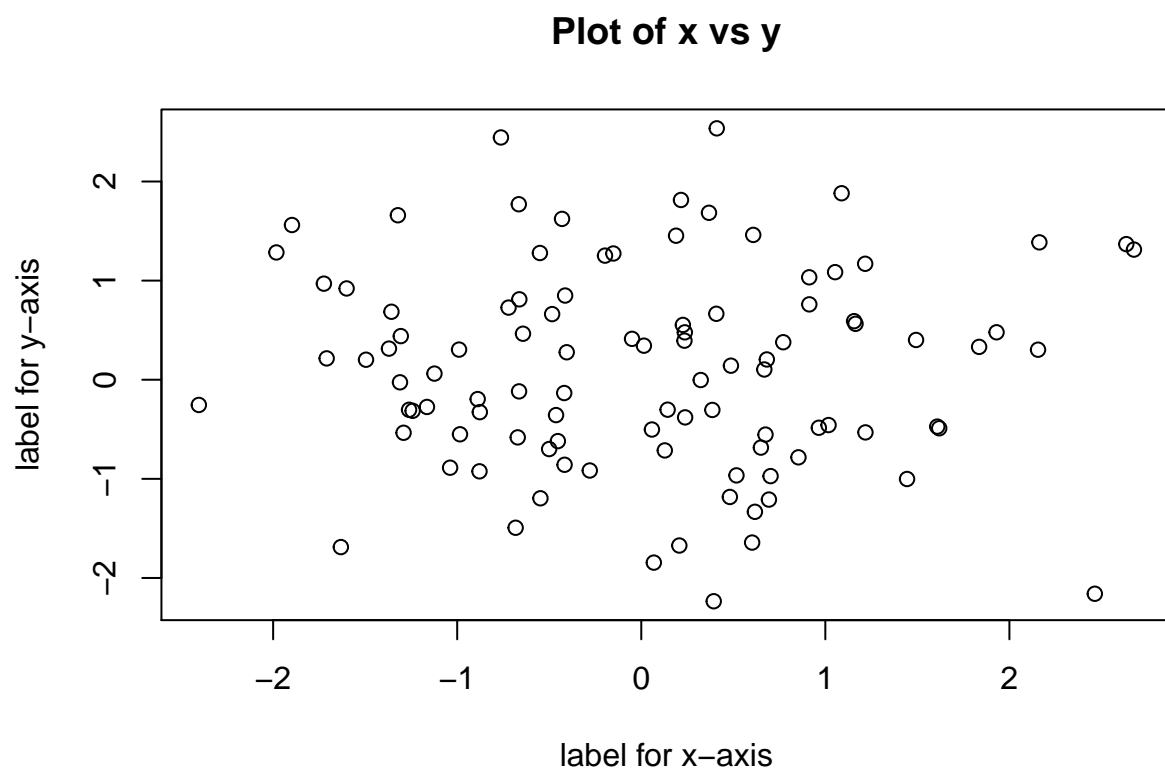
2.3.2 Graphics

Graphics Chunk 6

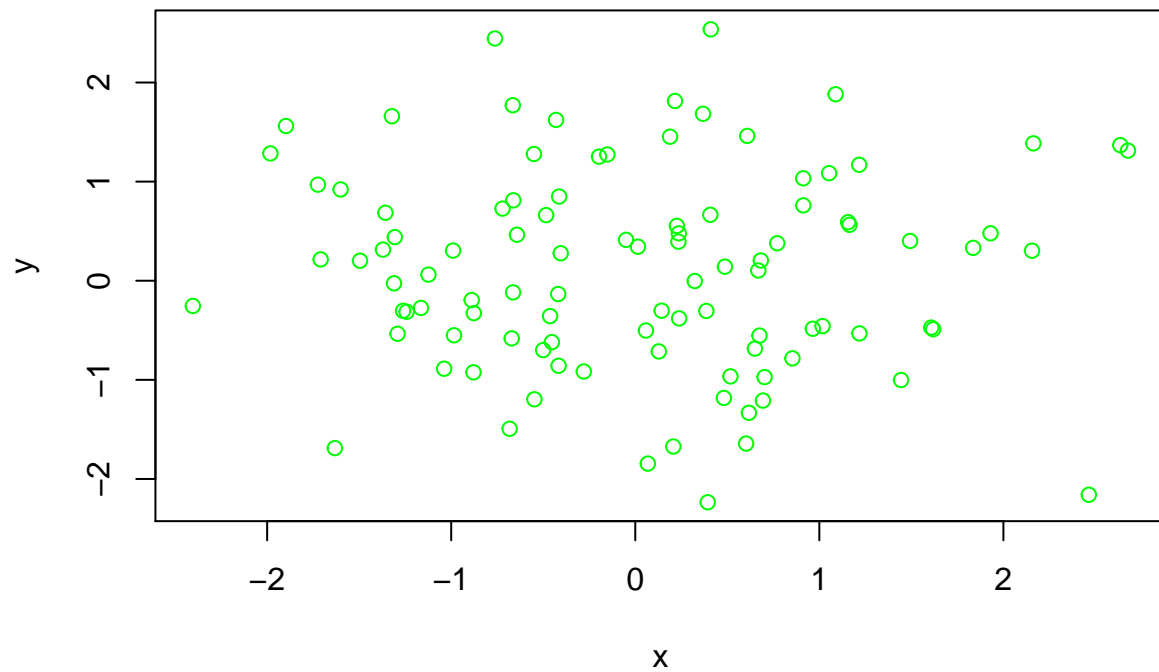
```
x = rnorm(100)  
y = rnorm(100)  
plot(x,y)
```



```
plot(x,y,xlab = "label for x-axis", ylab = "label for y-axis", main = "Plot of x vs y")
```



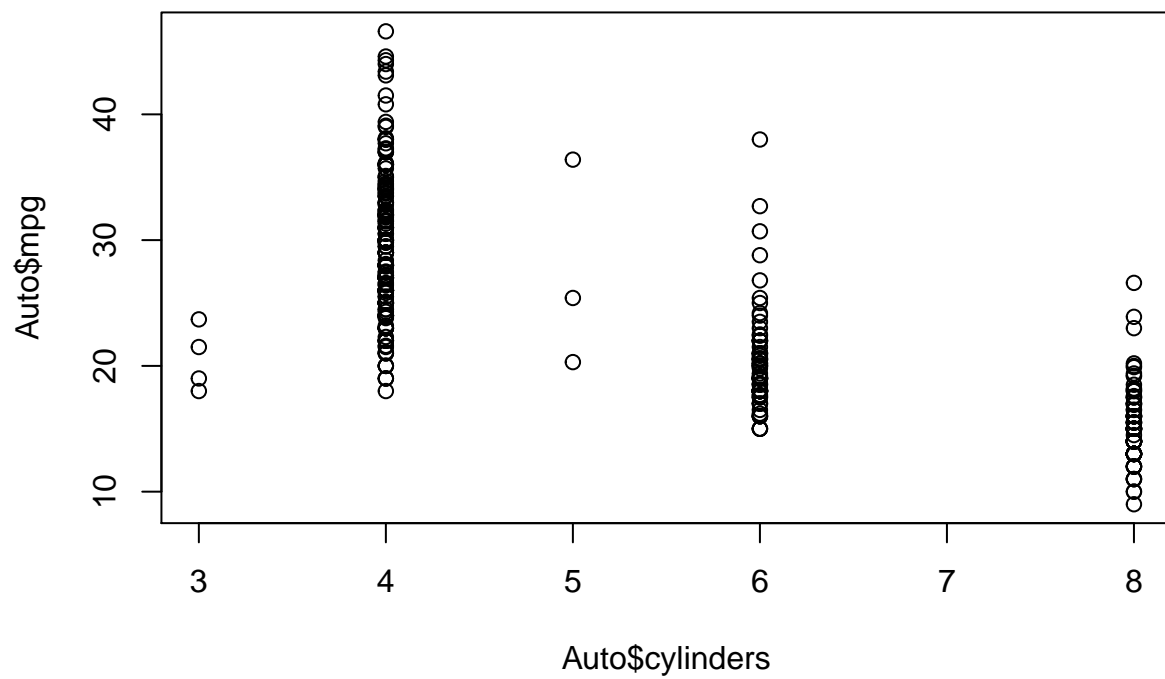
```
plot(x,y,col = "green")
```



2.3.5 Additional Graphical and Numerical Summaries

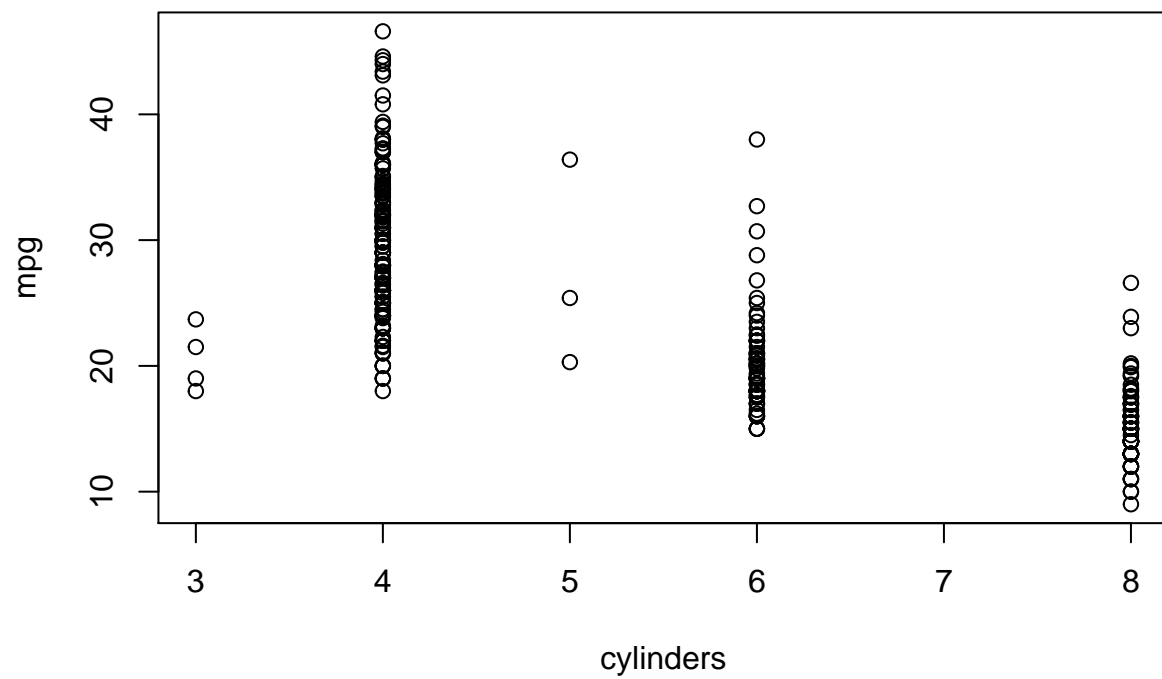
Graphics Chunk 7

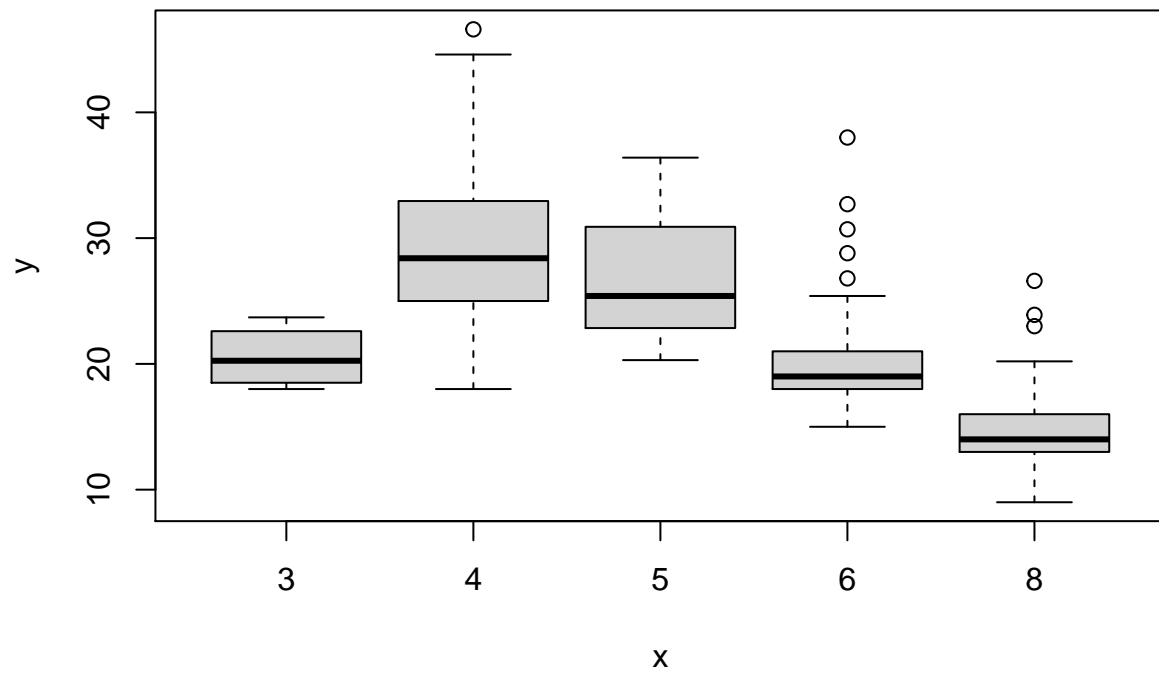
```
library(ISLR)
# plot(cylinders, mpg)
plot(Auto$cylinders, Auto$mpg)
```



Graphics Chunk 8

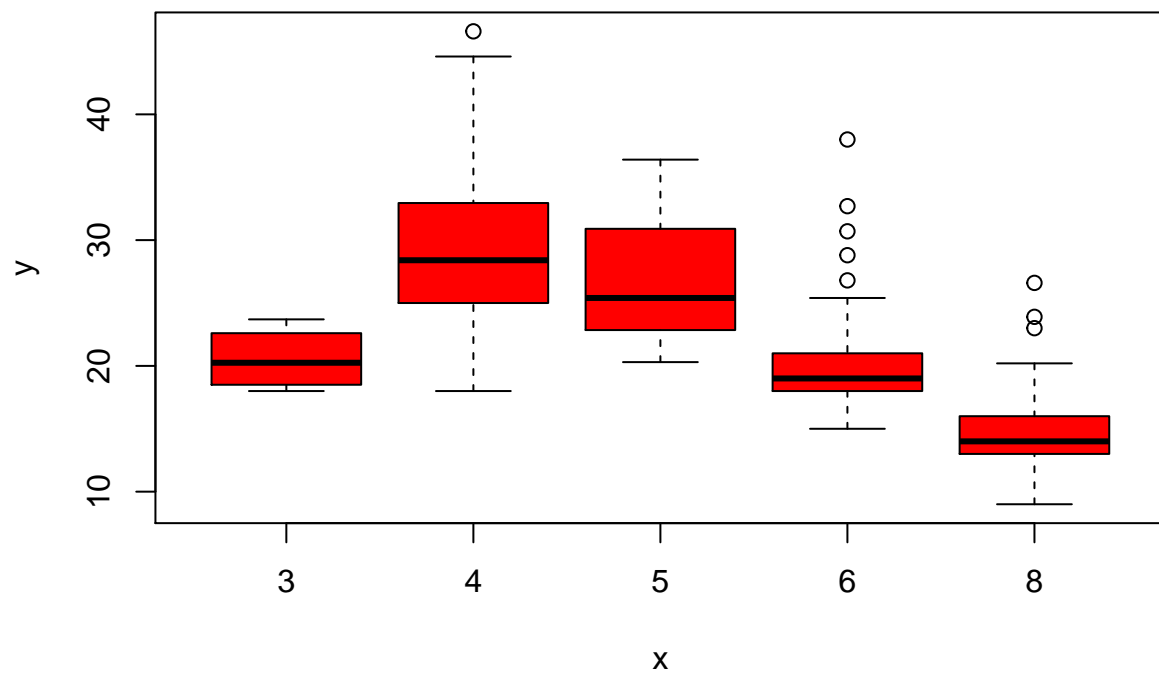
```
attach(Auto)
plot(cylinders,mpg)
```



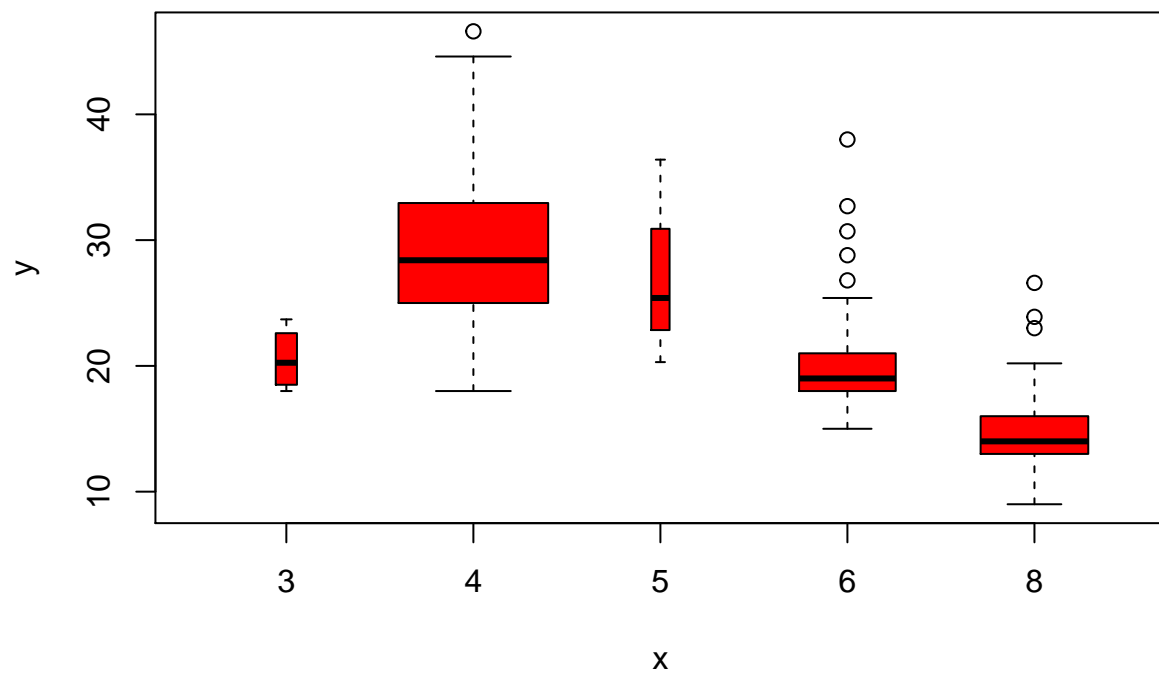


Graphics Chunk 10

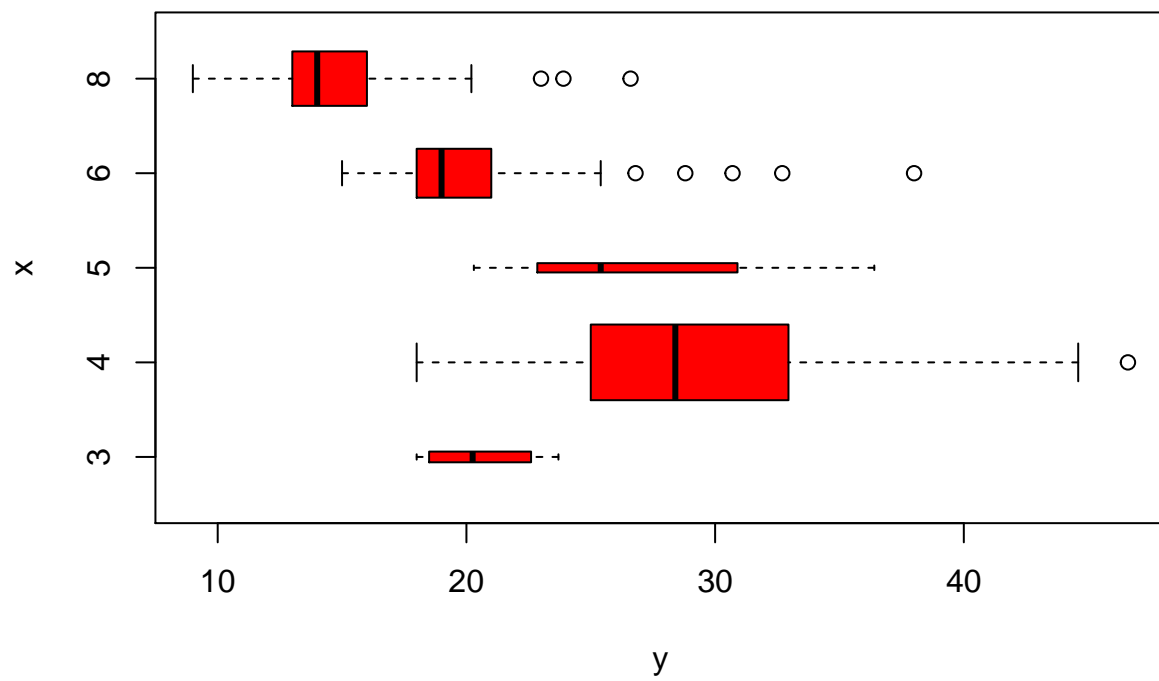
```
plot(cylinders,mpg, col = "red")
```



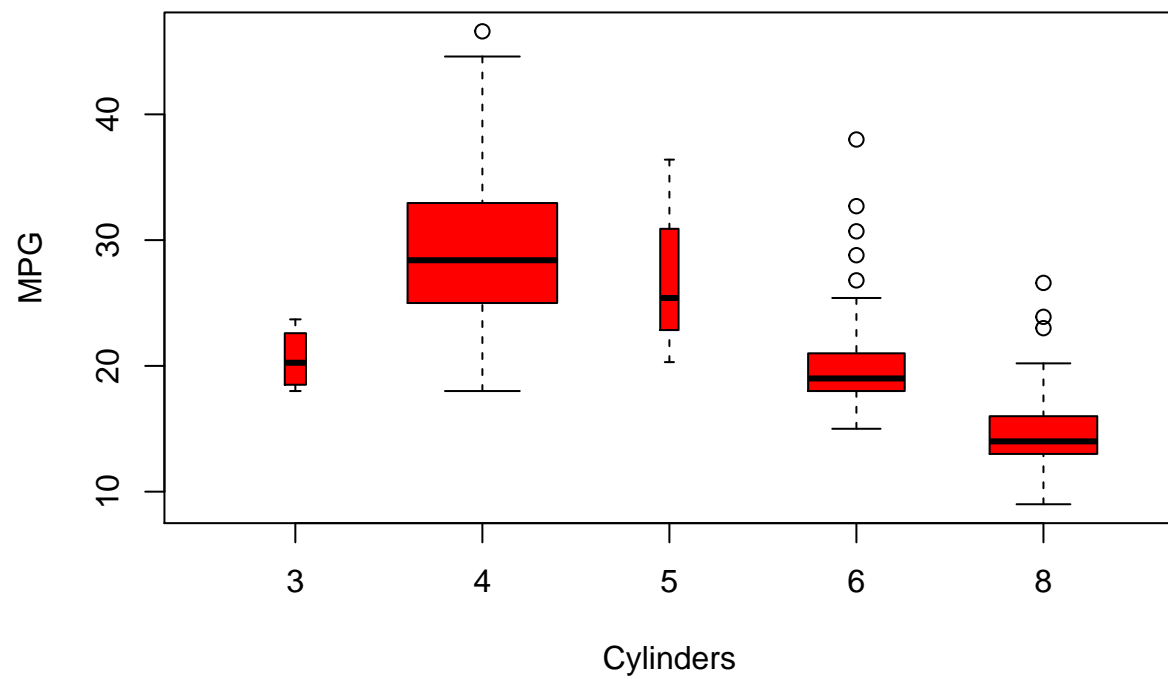
```
plot(cylinders,mpg, col = "red", varwidth = T)
```



```
plot(cylinders,mpg, col = "red", varwidth = T, horizontal = T)
```

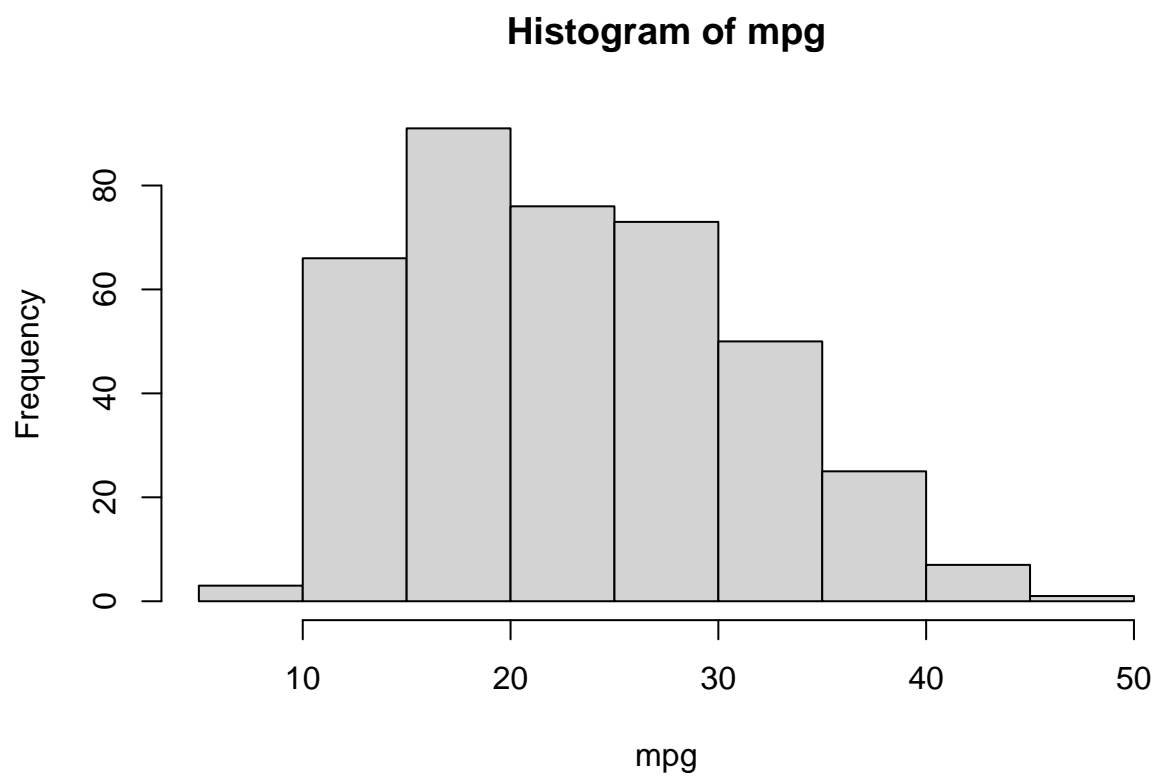


```
plot(cylinders,mpg, col = "red", varwidth = T, xlab = "Cylinders", ylab = "MPG")
```



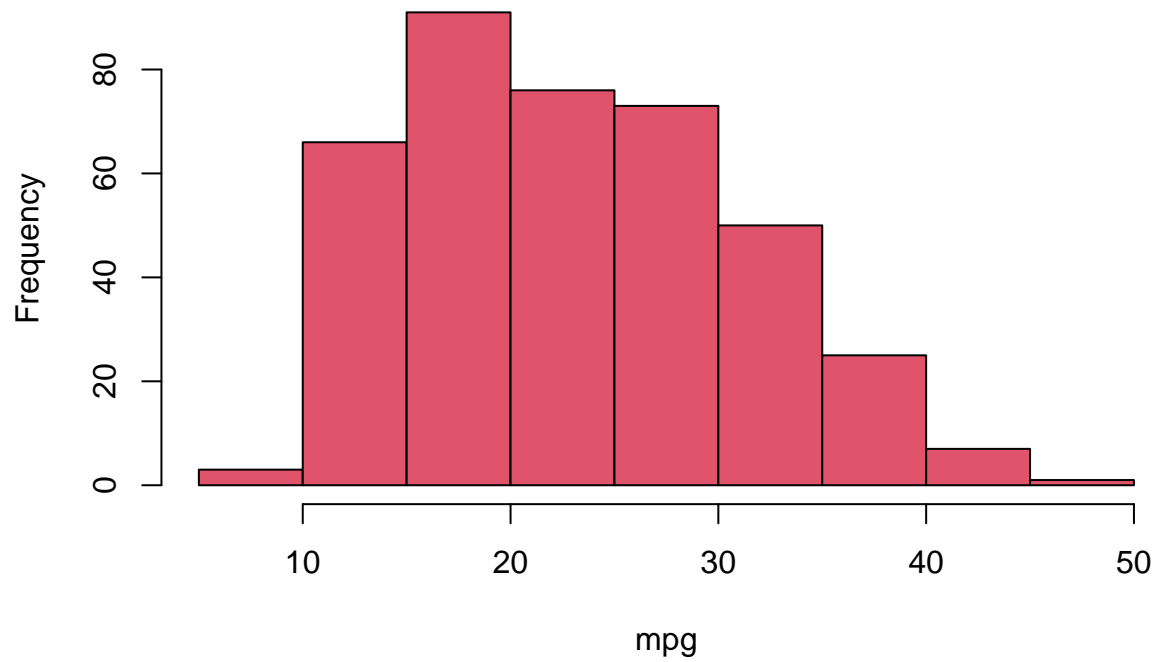
Graphics Chunk 11

```
hist(mpg)
```

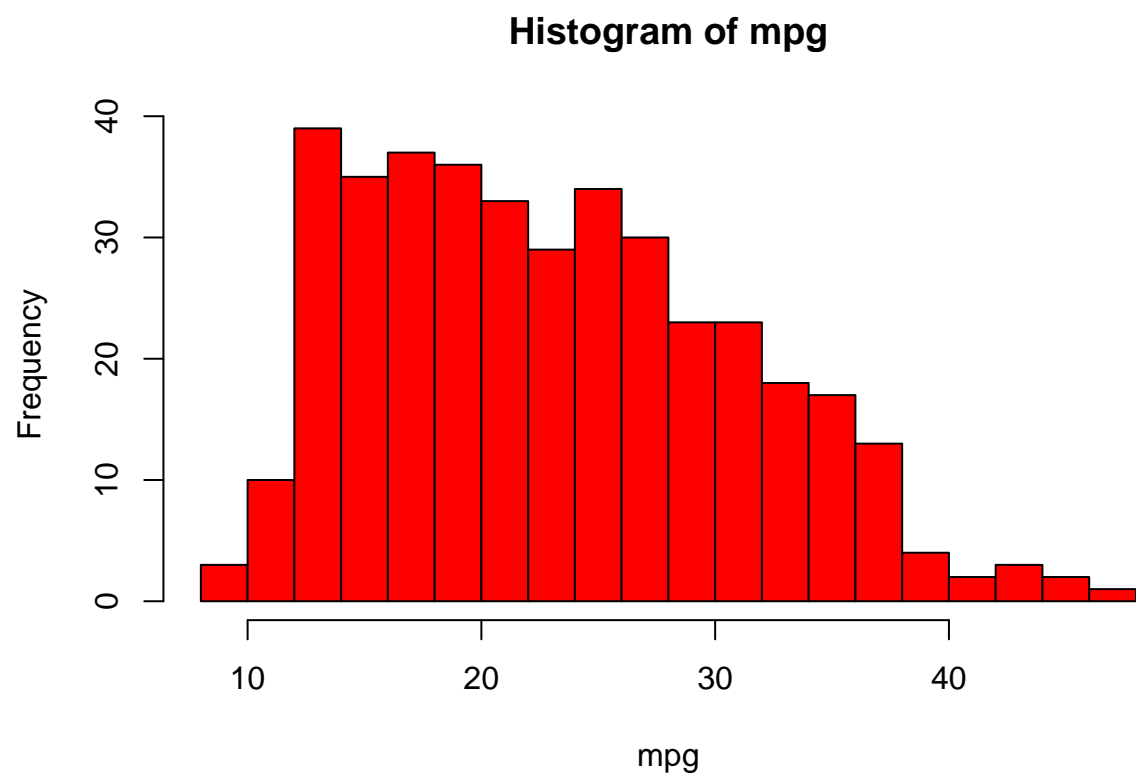


```
hist(mpg,col = 2)
```

Histogram of mpg

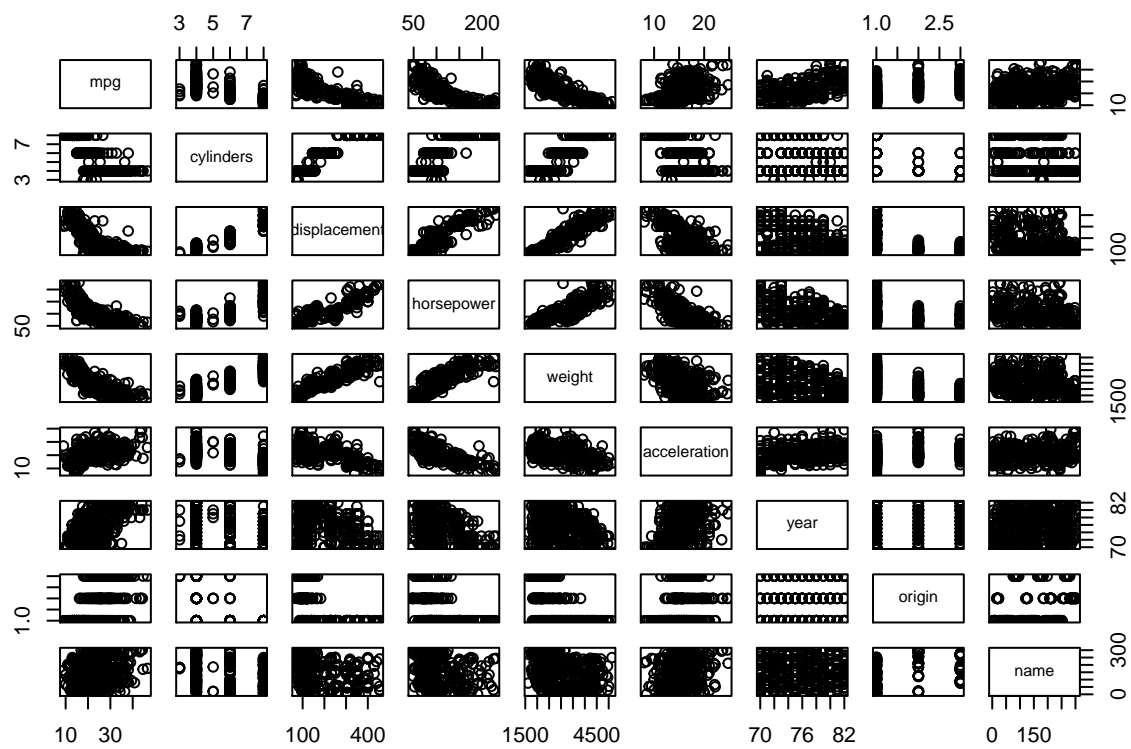


```
hist(mpg,col = "red", breaks = 15)
```

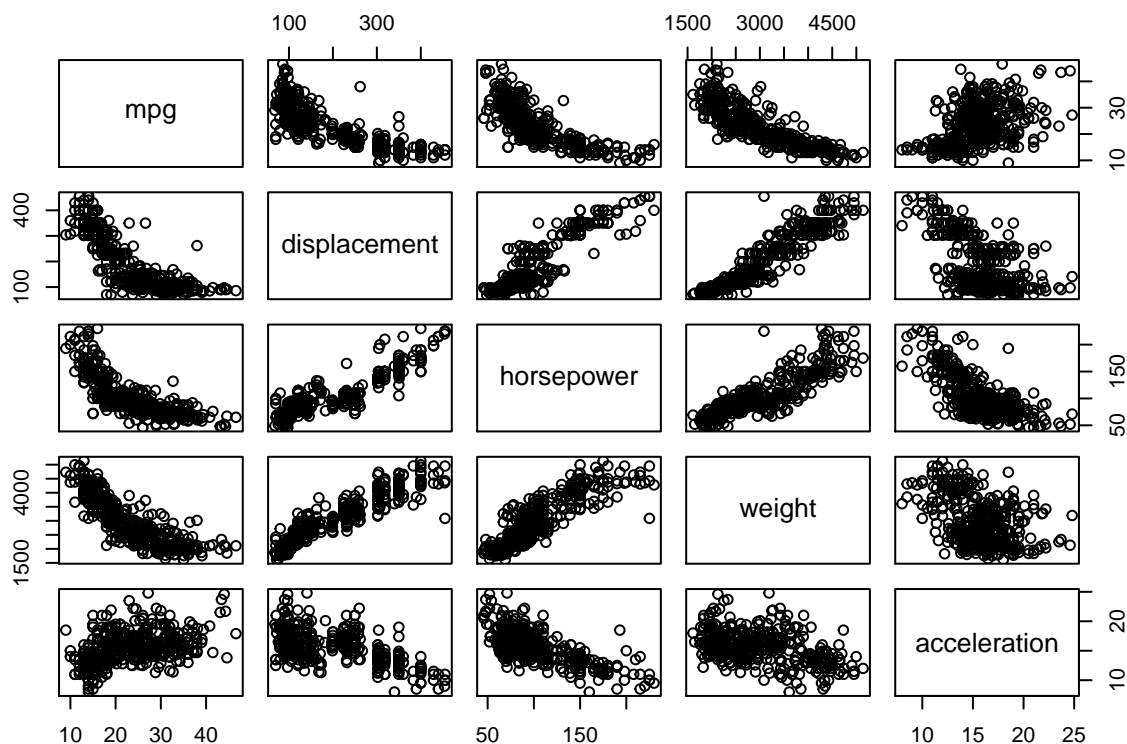


Graphics Chunk 12

```
pairs(Auto)
```

```
pairs(~mpg + displacement + horsepower + weight + acceleration, Auto)
```



Graphics Chunk 13

```
plot(horsepower,mpg)
identify(horsepower,mpg,name)
```

Graphics Chunk 14

```
summary(Auto)
```

```
##      mpg      cylinders  displacement  horsepower      weight
##  Min.   : 9.00   Min.   :3.000   Min.   : 68.0   Min.   : 46.0   Min.   :1613
##  1st Qu.:17.00   1st Qu.:4.000   1st Qu.:105.0   1st Qu.: 75.0   1st Qu.:2225
##  Median :22.75   Median :4.000   Median :151.0   Median : 93.5   Median :2804
##  Mean   :23.45   Mean   :5.472   Mean   :194.4   Mean   :104.5   Mean   :2978
##  3rd Qu.:29.00   3rd Qu.:8.000   3rd Qu.:275.8   3rd Qu.:126.0   3rd Qu.:3615
##  Max.   :46.60   Max.   :8.000   Max.   :455.0   Max.   :230.0   Max.   :5140
##
##  acceleration      year      origin      name
##  Min.   : 8.00   Min.   :70.00   Min.   :1.000   amc matador      : 5
##  1st Qu.:13.78   1st Qu.:73.00   1st Qu.:1.000   ford pinto       : 5
##  Median :15.50   Median :76.00   Median :1.000   toyota corolla   : 5
```

```
## Mean :15.54 Mean :75.98 Mean :1.577 amc gremlin : 4
## 3rd Qu.:17.02 3rd Qu.:79.00 3rd Qu.:2.000 amc hornet : 4
## Max. :24.80 Max. :82.00 Max. :3.000 chevrolet chevette: 4
## (Other) :365
```

```
summary(mpg)
```

```
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 9.00 17.00 22.75 23.45 29.00 46.60
```

2.3.3 Indexing Data

Indexing Chunk 15

```
A = matrix(1:16,4,4)
A
```

```
## [,1] [,2] [,3] [,4]
## [1,] 1 5 9 13
## [2,] 2 6 10 14
## [3,] 3 7 11 15
## [4,] 4 8 12 16
```

```
A[2,3]
```

```
## [1] 10
```

```
A[c(1,3),c(2,4)]
```

```
## [,1] [,2]
## [1,] 5 13
## [2,] 7 15
```

```
A[1:3,2:4]
```

```
## [,1] [,2] [,3]
## [1,] 5 9 13
## [2,] 6 10 14
## [3,] 7 11 15
```

```
A[1:2,]
```

```
## [,1] [,2] [,3] [,4]
## [1,] 1 5 9 13
## [2,] 2 6 10 14
```

```
A[,1:2]
```

```
##      [,1] [,2]
## [1,]    1    5
## [2,]    2    6
## [3,]    3    7
## [4,]    4    8
```

```
A[1,]
```

```
## [1]  1  5  9 13
```

```
A[-c(1,3),]
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    2    6   10   14
## [2,]    4    8   12   16
```

```
dim(A)
```

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

2.3.4 Loading Data

Loading Data Chunk 16

```
Auto = read.table("C:/C/DAT500S/Data/Auto.data")
fix(Auto)
```

```
Auto = read.table("C:/C/DAT500S/Data/Auto.data",header = T, na.strings = "?")
fix(Auto)
```

```
Auto = read.csv("C:/C/DAT500S/Data/Auto.csv",header = T, na.strings = "?")
fix(Auto)
```

Loading Data Chunk 17

```
dim(Auto)
```

```
## [1] 392  9
```

```
Auto[1:4,]
```

```
##   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
##                                     name
## 1 chevrolet chevelle malibu
## 2          buick skylark 320
## 3          plymouth satellite
## 4          amc rebel sst
```

```
Auto = na.omit(Auto)
dim(Auto)
```

```
## [1] 392  9
```

```
names(Auto)
```

```
## [1] "mpg"          "cylinders"    "displacement" "horsepower"   "weight"
## [6] "acceleration" "year"         "origin"       "name"
```