Title comes here

your name

1/15/2020

Get DataFrame

```r
# install.packages("MASS")
library(MASS)
library(help=MASS)
d1 = Cars93
dim(d1)
```

```
## [1] 93 27
```

```r
str(d1)
```

```
## 'data.frame':    93 obs. of  27 variables:
##  $ Manufacturer      : Factor w/ 32 levels "Acura","Audi",..: 1 1 2 2 3 4 4 4 4 5 ...
##  $ Model             : Factor w/ 93 levels "100","190E","240",..: 49 56 9 1 6 24 54 74 73 35 ...
##  $ Type              : Factor w/ 6 levels "Compact","Large",..: 4 3 1 3 3 3 2 2 3 2 ...
##  $ Min.Price         : num  12.9 29.2 25.9 30.8 23.7 14.2 19.9 22.6 26.3 33 ...
##  $ Price             : num  15.9 33.9 29.1 37.7 30 15.7 20.8 23.7 26.3 34.7 ...
##  $ Max.Price         : num  18.8 38.7 32.3 44.6 36.2 17.3 21.7 24.9 26.3 36.3 ...
##  $ MPG.city          : int  25 18 20 19 22 22 19 16 19 16 ...
##  $ MPG.highway       : int  31 25 26 26 30 31 28 25 27 25 ...
##  $ AirBags           : Factor w/ 3 levels "Driver & Passenger",..: 3 1 2 1 2 2 2 2 2 2 ...
##  $ DriveTrain        : Factor w/ 3 levels "4WD","Front",..: 2 2 2 2 3 2 2 3 2 2 ...
##  $ Cylinders         : Factor w/ 6 levels "3","4","5","6",..: 2 4 4 4 2 2 4 4 4 5 ...
##  $ EngineSize        : num  1.8 3.2 2.8 2.8 3.5 2.2 3.8 5.7 3.8 4.9 ...
##  $ Horsepower        : int  140 200 172 172 208 110 170 180 170 200 ...
##  $ RPM               : int  6300 5500 5500 5500 5700 5200 4800 4000 4800 4100 ...
##  $ Rev.per.mile      : int  2890 2335 2280 2535 2545 2565 1570 1320 1690 1510 ...
##  $ Man.trans.avail   : Factor w/ 2 levels "No","Yes": 2 2 2 2 2 1 1 1 1 1 ...
##  $ Fuel.tank.capacity: num  13.2 18 16.9 21.1 21.1 16.4 18 23 18.8 18 ...
##  $ Passengers        : int  5 5 5 6 4 6 6 6 5 6 ...
##  $ Length            : int  177 195 180 193 186 189 200 216 198 206 ...
##  $ Wheelbase         : int  102 115 102 106 109 105 111 116 108 114 ...
##  $ Width             : int  68 71 67 70 69 69 74 78 73 73 ...
##  $ Turn.circle       : int  37 38 37 37 39 41 42 45 41 43 ...
##  $ Rear.seat.room    : num  26.5 30 28 31 27 28 30.5 30.5 26.5 35 ...
##  $ Luggage.room      : int  11 15 14 17 13 16 17 21 14 18 ...
##  $ Weight            : int  2705 3560 3375 3405 3640 2880 3470 4105 3495 3620 ...
##  $ Origin            : Factor w/ 2 levels "USA","non-USA": 2 2 2 2 2 1 1 1 1 1 ...
##  $ Make              : Factor w/ 93 levels "Acura Integra",..: 1 2 4 3 5 6 7 9 8 10 ...
```

```r
n = nrow(d1)
n
```

```
## [1] 93
```

SUBSETTING

These are different ways of subsetting a dataframe

```
d2 = d1[c(1:10),c(2,4,9)]
d2
```

```
##          Model Min.Price           AirBags
## 1       Integra      12.9              None
## 2        Legend      29.2 Driver & Passenger
## 3            90      25.9        Driver only
## 4           100      30.8 Driver & Passenger
## 5          535i      23.7        Driver only
## 6       Century      14.2        Driver only
## 7       LeSabre      19.9        Driver only
## 8     Roadmaster      22.6        Driver only
## 9       Riviera      26.3        Driver only
## 10      DeVille      33.0        Driver only
```

```
d2 = data.frame(manufacturer = d1$Manufacturer, price = d1$Price)
head(d2)
```

```
##   manufacturer price
## 1        Acura  15.9
## 2        Acura  33.9
## 3         Audi  29.1
## 4         Audi  37.7
## 5          BMW  30.0
## 6        Buick  15.7
```

```
d2 = subset(d1,select=c(Manufacturer,Price))
head(d2)
```

```
##   Manufacturer Price
## 1        Acura  15.9
## 2        Acura  33.9
## 3         Audi  29.1
## 4         Audi  37.7
## 5          BMW  30.0
## 6        Buick  15.7
```

```
# Ford cars
d2 = subset(d1,subset = Manufacturer == "Ford")
dim(d2)
```

```
## [1]  8 27
```

```
head(d2,8)
```

```
##    Manufacturer          Model    Type Min.Price Price Max.Price MPG.city
## 31         Ford        Festiva   Small       6.9   7.4       7.9       31
## 32         Ford         Escort   Small       8.4  10.1      11.9       23
## 33         Ford          Tempo Compact      10.4  11.3      12.2       22
## 34         Ford        Mustang  Sporty      10.8  15.9      21.0       22
## 35         Ford          Probe  Sporty      12.8  14.0      15.2       24
## 36         Ford        Aerostar     Van      14.5  19.9      25.3       15
## 37         Ford         Taurus Midsize      15.6  20.2      24.8       21
## 38         Ford Crown_Victoria   Large      20.1  20.9      21.7       18
##    MPG.highway     AirBags DriveTrain Cylinders EngineSize Horsepower   RPM
```

```
## 31          33      None      Front     4      1.3        63 5000
## 32          30      None      Front     4      1.8       127 6500
## 33          27      None      Front     4      2.3        96 4200
## 34          29 Driver only     Rear     4      2.3       105 4600
## 35          30 Driver only    Front     4      2.0       115 5500
## 36          20 Driver only      4WD     6      3.0       145 4800
## 37          30 Driver only    Front     6      3.0       140 4800
## 38          26 Driver only     Rear     8      4.6       190 4200
##    Rev.per.mile Man.trans.avail Fuel.tank.capacity Passengers Length Wheelbase
## 31         3150             Yes               10.0          4    141        90
## 32         2410             Yes               13.2          5    171        98
## 33         2805             Yes               15.9          5    177       100
## 34         2285             Yes               15.4          4    180       101
## 35         2340             Yes               15.5          4    179       103
## 36         2080             Yes               21.0          7    176       119
## 37         1885              No               16.0          5    192       106
## 38         1415              No               20.0          6    212       114
##    Width Turn.circle Rear.seat.room Luggage.room Weight Origin
## 31    63          33           26.0           12   1845    USA
## 32    67          36           28.0           12   2530    USA
## 33    68          39           27.5           13   2690    USA
## 34    68          40           24.0           12   2850    USA
## 35    70          38           23.0           18   2710    USA
## 36    72          45           30.0           NA   3735    USA
## 37    71          40           27.5           18   3325    USA
## 38    78          43           30.0           21   3950    USA
##                Make
## 31      Ford Festiva
## 32       Ford Escort
## 33        Ford Tempo
## 34      Ford Mustang
## 35        Ford Probe
## 36      Ford Aerostar
## 37        Ford Taurus
## 38 Ford Crown_Victoria
```

```r
# cars weighting > 4000
d2 = d1[d1$Weight>4000,]
d2
```

```
##     Manufacturer      Model   Type Min.Price Price Max.Price MPG.city MPG.highway
## 8          Buick Roadmaster  Large      22.6  23.7      24.9       16          25
## 17     Chevrolet      Astro    Van      14.7  16.6      18.6       15          20
## 52       Lincoln   Town_Car  Large      34.4  36.1      37.8       18          26
## 66        Nissan      Quest    Van      16.7  19.1      21.5       17          23
##              AirBags DriveTrain Cylinders EngineSize Horsepower  RPM
## 8        Driver only       Rear         6        5.7        180 4000
## 17              None        4WD         6        4.3        165 4000
## 52 Driver & Passenger       Rear         8        4.6        210 4600
## 66              None      Front         6        3.0        151 4800
##    Rev.per.mile Man.trans.avail Fuel.tank.capacity Passengers Length Wheelbase
## 8          1320              No                 23          6    216       116
## 17         1790              No                 27          8    194       111
## 52         1840              No                 20          6    219       117
## 66         2065              No                 20          7    190       112
```

```
##    Width Turn.circle Rear.seat.room Luggage.room Weight  Origin
## 8     78          45           30.5           21   4105     USA
## 17    78          42           33.5           NA   4025     USA
## 52    77          45           31.5           22   4055     USA
## 66    74          41           27.0           NA   4100 non-USA
##             Make
## 8  Buick Roadmaster
## 17  Chevrolet Astro
## 52 Lincoln Town_Car
## 66     Nissan Quest
```

```
# there are four cars
```

COUNTING

```
# number of cars by DriveTrain
names(d1)
```

```
##  [1] "Manufacturer"      "Model"             "Type"
##  [4] "Min.Price"         "Price"             "Max.Price"
##  [7] "MPG.city"          "MPG.highway"       "AirBags"
## [10] "DriveTrain"        "Cylinders"         "EngineSize"
## [13] "Horsepower"        "RPM"               "Rev.per.mile"
## [16] "Man.trans.avail"   "Fuel.tank.capacity" "Passengers"
## [19] "Length"            "Wheelbase"         "Width"
## [22] "Turn.circle"       "Rear.seat.room"    "Luggage.room"
## [25] "Weight"            "Origin"            "Make"
```

```
# categories of DriveTrain
unique(d1$DriveTrain)
```

```
## [1] Front Rear  4WD
## Levels: 4WD Front Rear
```

```
# number of cars
table(d1$DriveTrain)
```

```
##
##   4WD Front  Rear
##    10    67    16
```

```
# proportions
prop.table(table(d1$DriveTrain))
```

```
##
##       4WD     Front      Rear
## 0.1075269 0.7204301 0.1720430
```

```
# count by two factors
table(d1$AirBags,d1$DriveTrain)
```

```
##
##                     4WD Front Rear
##   Driver & Passenger  0    11    5
##   Driver only         5    28   10
##   None                5    28    1
```

MEASURING

```r
# median weight per DriveTrain
tapply(d1$Weight,d1$DriveTrain,median)
```

```
##   4WD Front  Rear
##  3720  2910  3520
```

```r
# sorted
aux1 = tapply(d1$Weight,d1$DriveTrain,median)
sort1 = aux1[order(aux1)]
sort1
```

```
## Front  Rear   4WD
##  2910  3520  3720
```

SORTING

```r
d2 = subset(d1,select=c(Manufacturer,Price,Weight,Width))
head(d2)
```

```
##   Manufacturer Price Weight Width
## 1        Acura  15.9   2705    68
## 2        Acura  33.9   3560    71
## 3         Audi  29.1   3375    67
## 4         Audi  37.7   3405    70
## 5          BMW  30.0   3640    69
## 6        Buick  15.7   2880    69
```
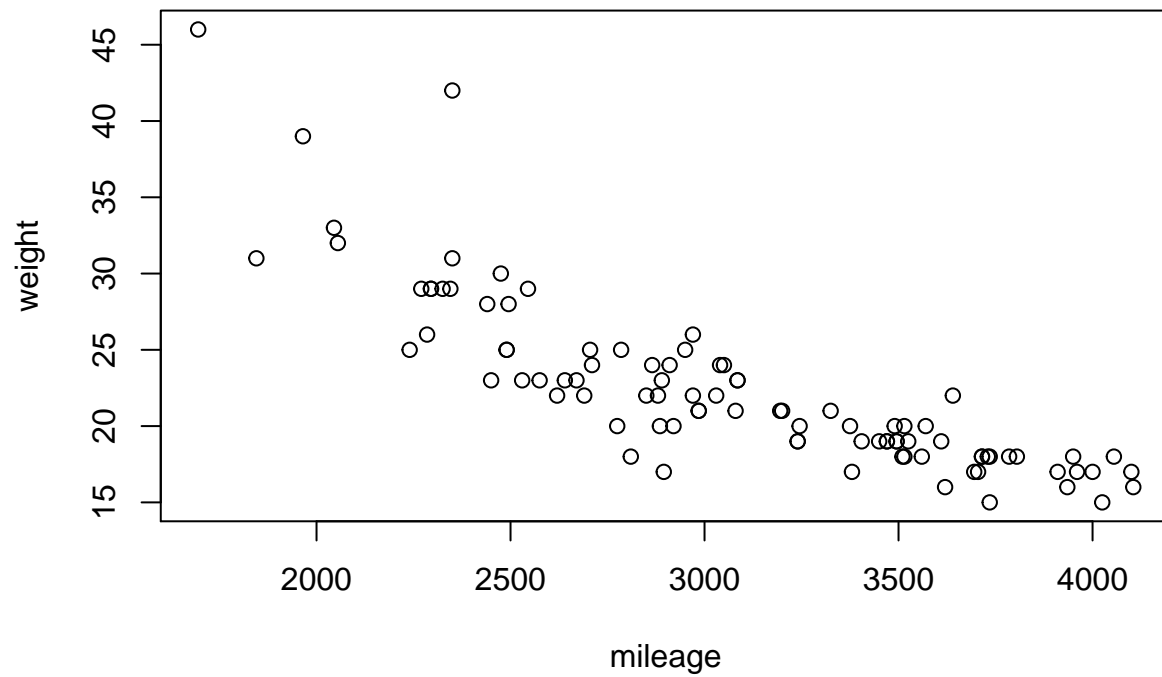
```r
# sort by Width
d3 = d2[order(d2$Width),]
head(d3)
```

```
##    Manufacturer Price Weight Width
## 80       Subaru   8.4   2045    60
## 31         Ford   7.4   1845    63
## 39          Geo   8.4   1695    63
## 44      Hyundai   8.0   2345    63
## 83       Suzuki   8.6   1965    63
## 88   Volkswagen   9.1   2240    63
```

Including Plots

You can also embed plots, for example a scatterplot of `Weight` vs `MPG.city`

Note that the `{echo = FALSE}` prevents printing the R code that generated the plot.