Pernah sekali saya ikut seleksi di salah satu Perusahaan Telco. Pada saat interview user, saya tidak mengira bakal ditanya langsung mengenai technical skill yang saya punya. Pada saat itu saya di suruh bawa laptop dan beliau suruh saya buka R studio. “boleh dikasih tunjuk script-script yang selama ini dikerjain ?”. Lho? Terkejut aku terheran – heran. To the point banget kan. Bagaimana jika anda menggunakan laptop yang ada di kantor mereka, bagi teman – teman yang menyimpan da di github atau google drive bagaimana jika tidak ada koneksi internet? Jadi pada kesempatan kali ini mari kita bahas sedikit tentang data set yang tersedia di R itu sendiri, barangkali dapat membantu sedikit yaa.

List Pre-loaded data

Untuk melihat beberapa data yang sudah di muat di R, anda cukup mengetik :

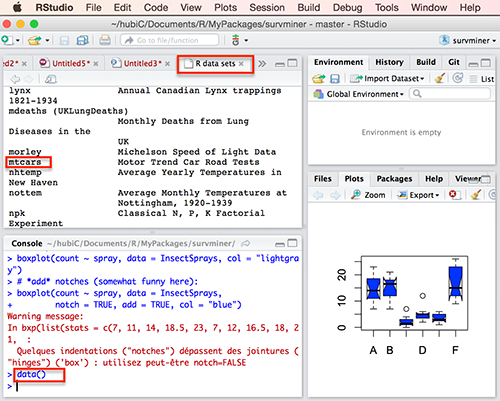
data()

**List of pre-loaded data**

To see the list of pre-loaded data, type the function **data**():

data()

The output is as follow:



**Loading a built-in R data**

Load and print mtcars data as follow:

# Loading

data(mtcars)

# Print the first 6 rows

**head**(mtcars, 6)

mpg cyl disp hp drat wt qsec vs am gear carb

Mazda RX4 21.0 6 160 110 3.90 2.620 16.46 0 1 4 4

Mazda RX4 Wag 21.0 6 160 110 3.90 2.875 17.02 0 1 4 4

Datsun 710 22.8 4 108 93 3.85 2.320 18.61 1 1 4 1

Hornet 4 Drive 21.4 6 258 110 3.08 3.215 19.44 1 0 3 1

Hornet Sportabout 18.7 8 360 175 3.15 3.440 17.02 0 0 3 2

Valiant 18.1 6 225 105 2.76 3.460 20.22 1 0 3 1

If you want learn more about mtcars data sets, type this:

?mtcars

**Most used R built-in data sets**

**mtcars: Motor Trend Car Road Tests**

The data was extracted from the 1974 Motor Trend US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models)

* View the content of *mtcars* data set:

# 1. Loading

data("mtcars")

# 2. Print

**head**(mtcars)

* It contains 32 observations and 11 variables:

# Number of rows (observations)

nrow(mtcars)

[1] 32

# Number of columns (variables)

ncol(mtcars)

[1] 11

* Description of variables:

1. mpg: Miles/(US) gallon
2. cyl: Number of cylinders
3. disp: Displacement (cu.in.)
4. hp: Gross horsepower
5. drat: Rear axle ratio
6. wt: Weight (1000 lbs)
7. qsec: 1/4 mile time
8. vs: V/S
9. am: Transmission (0 = automatic, 1 = manual)
10. gear: Number of forward gears
11. carb: Number of carburetors

If you want to learn more about *mtcars*, type this:

?mtcars

**iris**

**iris** data set gives the measurements in centimeters of the variables sepal length, sepal width, petal length and petal width, respectively, for 50 flowers from each of 3 species of iris. The species are Iris setosa, versicolor, and virginica.

data("iris")

**head**(iris)

Sepal.Length Sepal.Width Petal.Length Petal.Width Species

1 5.1 3.5 1.4 0.2 setosa

2 4.9 3.0 1.4 0.2 setosa

3 4.7 3.2 1.3 0.2 setosa

4 4.6 3.1 1.5 0.2 setosa

5 5.0 3.6 1.4 0.2 setosa

6 5.4 3.9 1.7 0.4 setosa

**ToothGrowth**

ToothGrowth data set contains the result from an experiment studying the effect of vitamin C on tooth growth in 60 Guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, (orange juice or ascorbic acid (a form of vitamin C and coded as VC).

data("ToothGrowth")

**head**(ToothGrowth)

len supp dose

1 4.2 VC 0.5

2 11.5 VC 0.5

3 7.3 VC 0.5

4 5.8 VC 0.5

5 6.4 VC 0.5

6 10.0 VC 0.5

1. len: Tooth length
2. supp: Supplement type (VC or OJ).
3. dose: numeric Dose in milligrams/day

**PlantGrowth**

Results obtained from an experiment to compare yields (as measured by dried weight of plants) obtained under a control and two different treatment condition.

data("PlantGrowth")

**head**(PlantGrowth)

weight group

1 4.17 ctrl

2 5.58 ctrl

3 5.18 ctrl

4 6.11 ctrl

5 4.50 ctrl

6 4.61 ctrl

**USArrests**

This data set contains statistics about violent crime rates by us state.

data("USArrests")

**head**(USArrests)

Murder Assault UrbanPop Rape

Alabama 13.2 236 58 21.2

Alaska 10.0 263 48 44.5

Arizona 8.1 294 80 31.0

Arkansas 8.8 190 50 19.5

California 9.0 276 91 40.6

Colorado 7.9 204 78 38.7

1. Murder: Murder arrests (per 100,000)
2. Assault: Assault arrests (per 100,000)
3. UrbanPop: Percent urban population
4. Rape: Rape arrests (per 100,000)