Data visualization (2)

Joaquin Cavieres

1. Base R Graphics

```
library("graphics")
library("datasets")
data(mtcars)
str(mtcars)
```

```
## 'data.frame':    32 obs. of  11 variables:
##  $ mpg : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
##  $ cyl : num  6 6 4 6 8 6 8 4 4 6 ...
##  $ disp: num  160 160 108 258 360 ...
##  $ hp  : num  110 110 93 110 175 105 245 62 95 123 ...
##  $ drat: num  3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
##  $ wt  : num  2.62 2.88 2.32 3.21 3.44 ...
##  $ qsec: num  16.5 17 18.6 19.4 17 ...
##  $ vs  : num  0 0 1 1 0 1 0 1 1 1 ...
##  $ am  : num  1 1 1 0 0 0 0 0 0 0 ...
##  $ gear: num  4 4 4 3 3 3 3 4 4 4 ...
##  $ carb: num  4 4 1 1 2 1 4 2 2 4 ...
```

It contains data about the design, performance and fuel economy of 32 automobiles from 1973 to 1974, extracted from the 1974 Motor Trend US magazine.

The `plot` function

The `plot` function is used to plot R objects and the arguments function are the following:

`plot(x,y,type,main,sub,xlab,ylab,asp,col,..)`

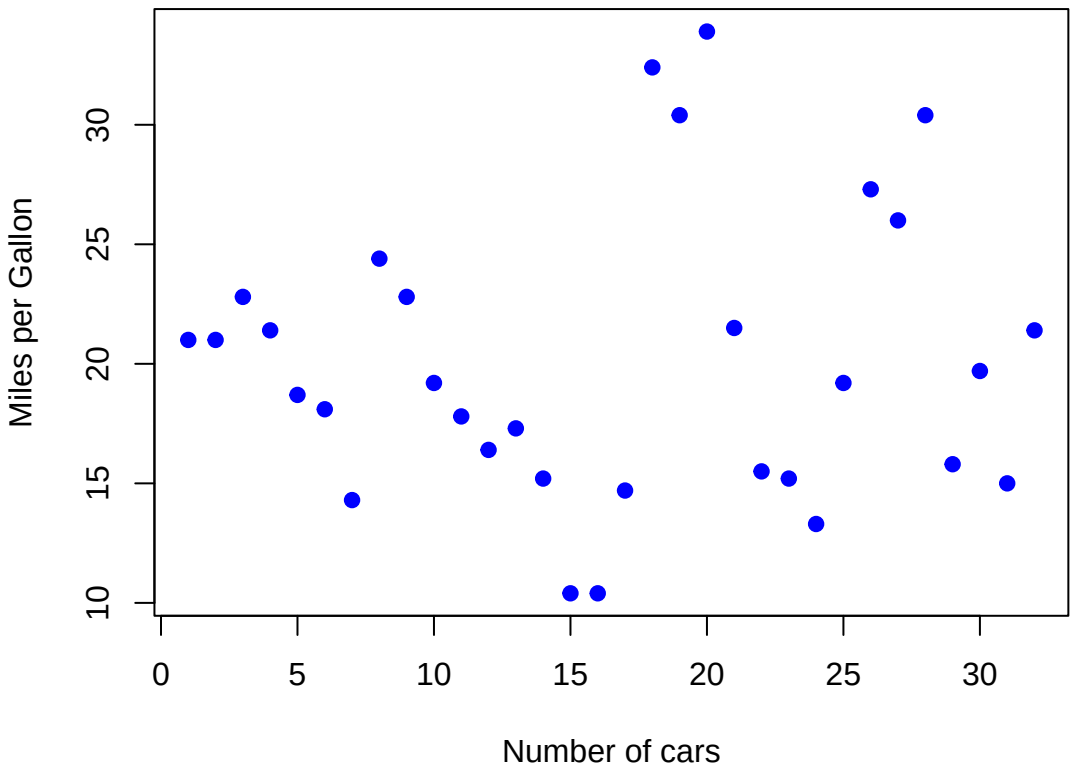
Where,

- x: The x coordinate of the plot, a single plotting structure, a function, or an R object

- y: The Y coordinate points in the plot (optional if x coordinate is a single structure)
- type: 'p' for points, 'l' for lines, 'b' for both, 'h' for high-density vertical lines, etc.
- main: Title of the plot
- sub: Subtitle of the plot
- xlab: Title for the x-axis
- ylab: Title for the y-axis
- asp : Aspect ratio(y/x)
- col: Color of the plot(points, lines, etc.)

For example,

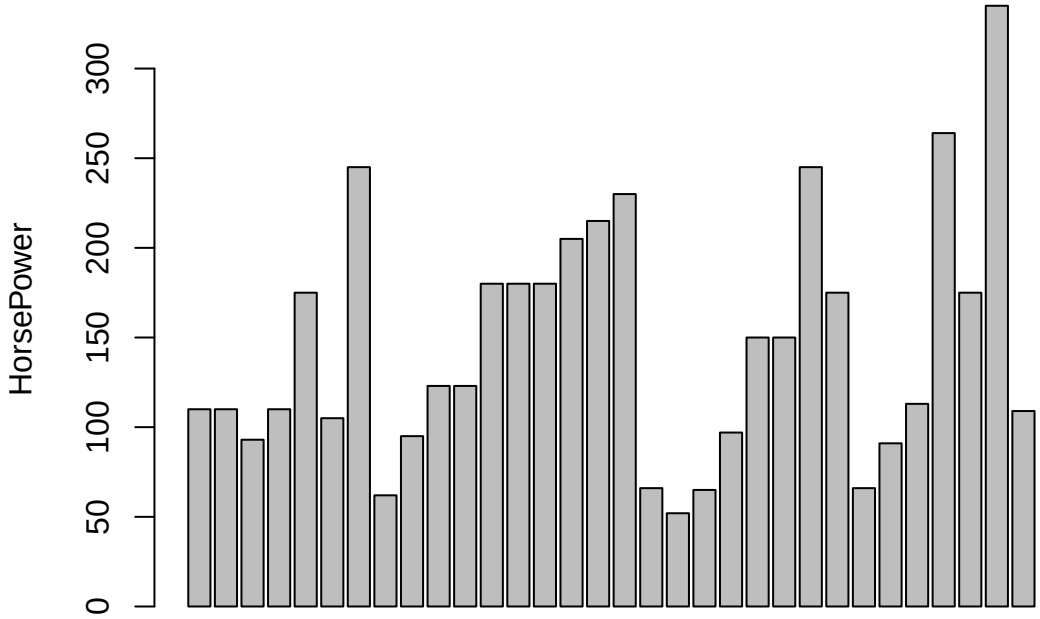
```
#To plot mpg(Miles per Gallon) vs Number of cars
plot(mtcars$mpg, xlab = "Number of cars", ylab = "Miles per Gallon", col = "blue", pch = 19)
```



Barplot

It is used to represent data in the form of rectangular bars, both in vertical and horizontal ways, and the length of the bar is proportional to the value of the variable.

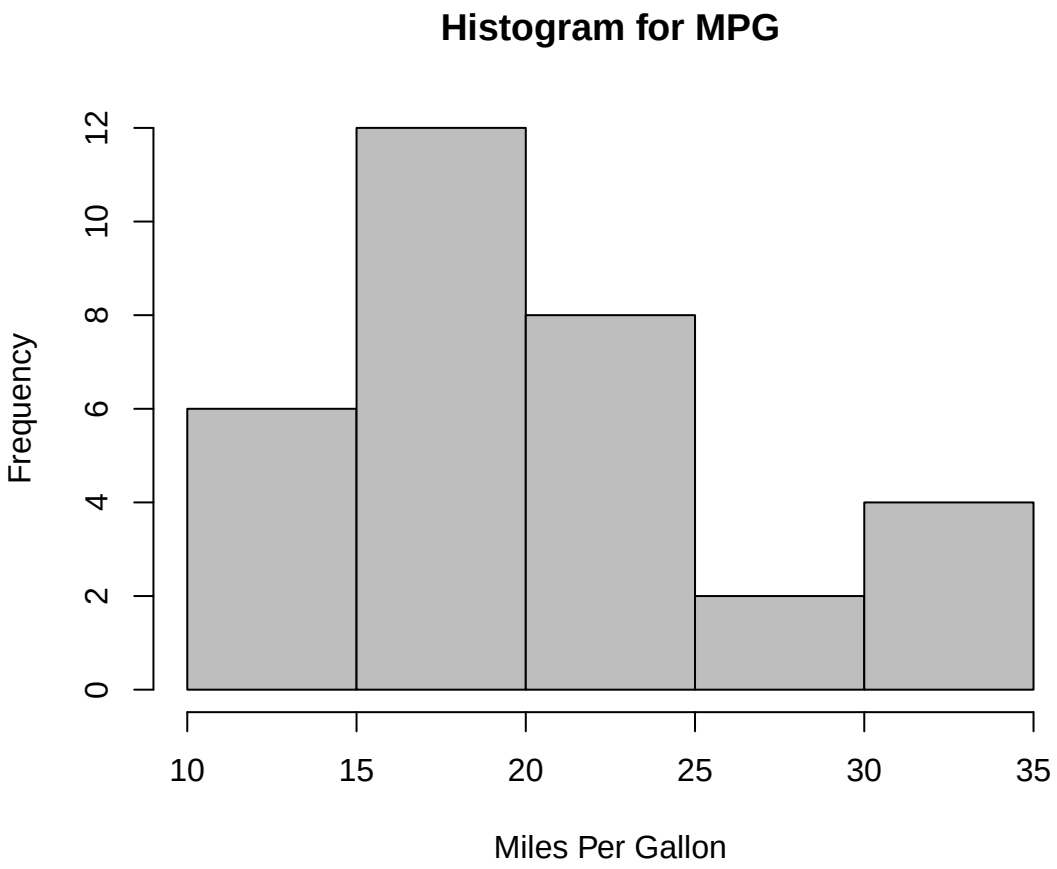
```
barplot(mtcars$hp, ylab = "HorsePower", col = "grey", horiz = FALSE)
```



Histogram

It is used to divide values into groups of continuous ranges measured against the frequency range of the variable.

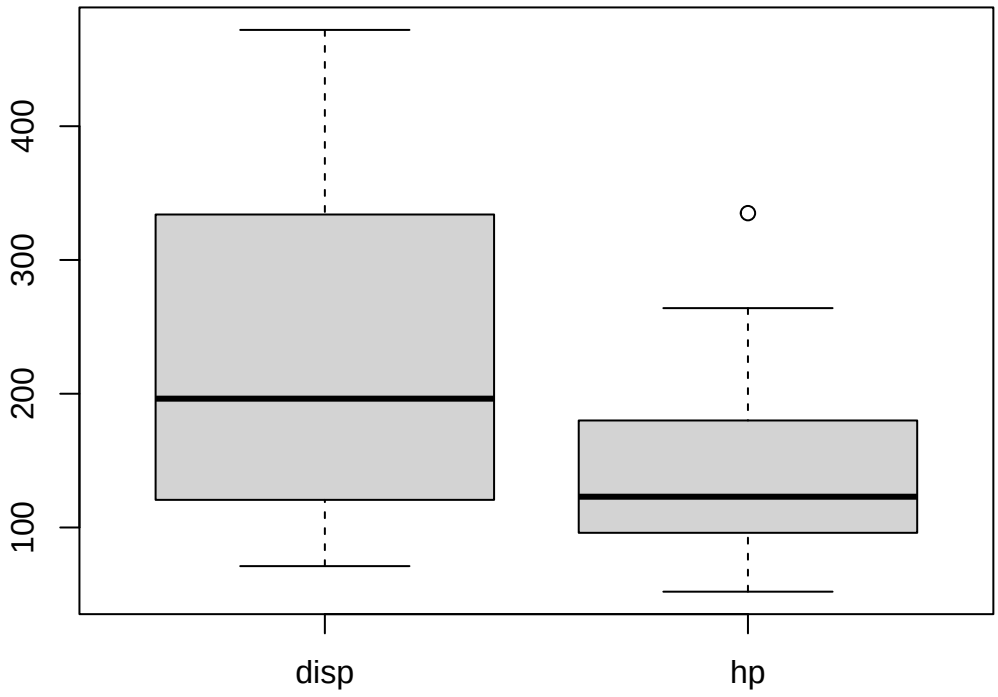
```
# Histogram for mpg (Miles per Gallon)
hist(mtcars$mpg,xlab = "Miles Per Gallon", main = "Histogram for MPG", col = "grey")
```



Boxplot

It is used to represent descriptive statistics of each variable in a dataset. It represents the minimum, first quartile, median, third quartile, and the maximum values of a variable.

```
# Boxplots for disp (Displacement) and hp (Horse Power)
boxplot(mtcars[,3:4])
```


400
300
200
100
disp
hp