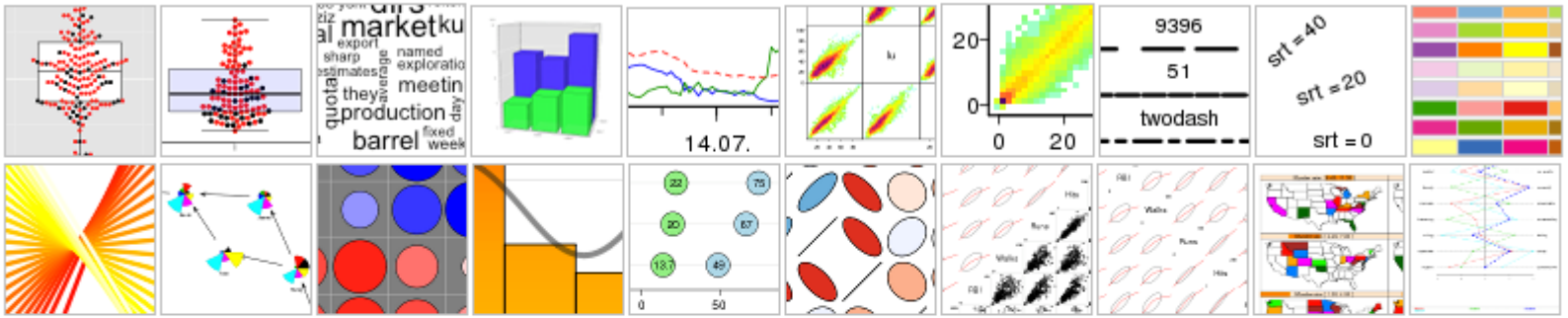
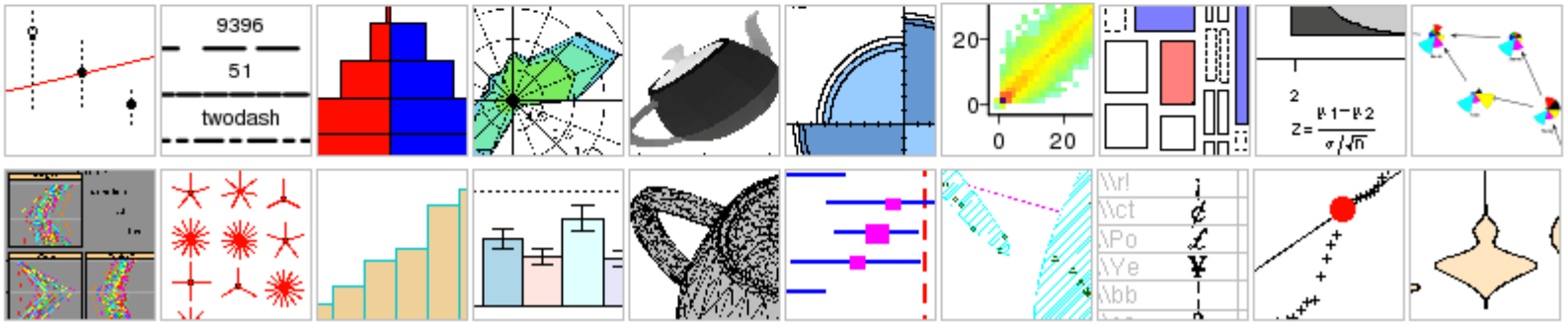
Special session: basic plotting

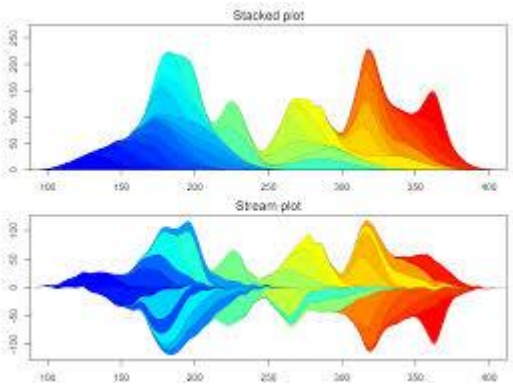
Samuel Brown & Marona Rovira

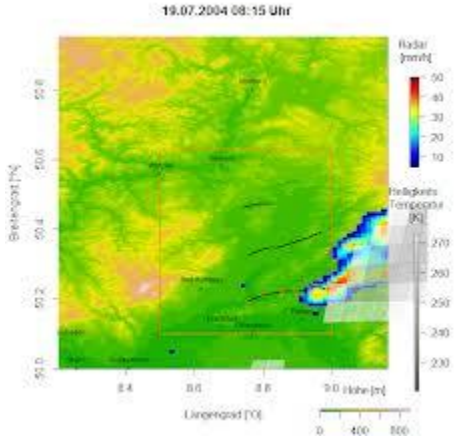
» Last entries ...

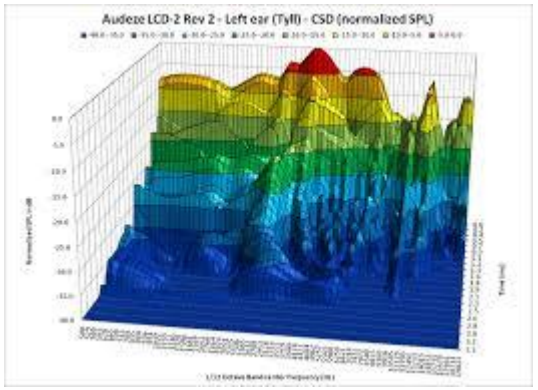


» Random entries









Analytic graphics

- Show comparisons
- Show causality, mechanisms, explanation, systematic structure.
- Show multivariate data (> 2 variables)
- Integrate the evidence (different visual tools)
- Describe and document

Base plotting system

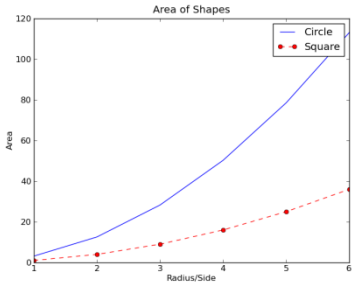
- First graphics system in R
- Add things one by one to the plot (e.g. text, lines, points, axis)

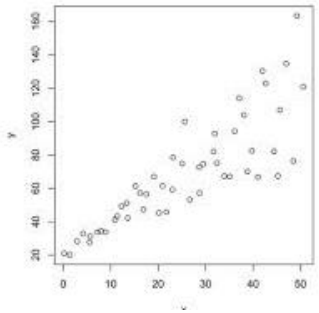
  - ✚ convenient
  - ✚ intuitive
  - ▬ must redraw to modify
  - ▬ easy to create ugly plots

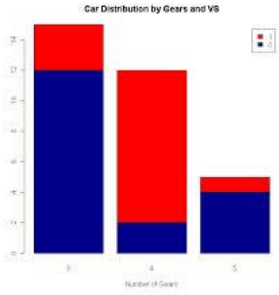
Base plotting functions

plot()

To make an scatterplot, or other type of plot depending on the class of the object being plotted.

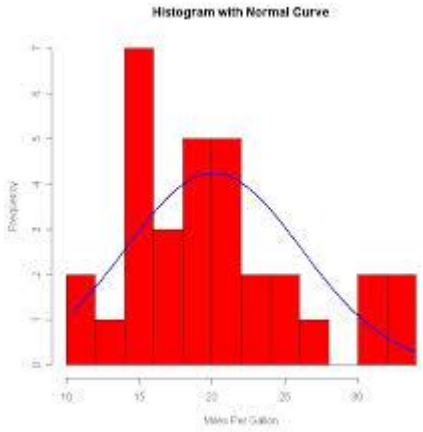






hist()

To make histograms (density plots).

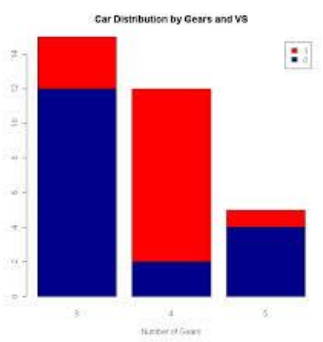


Base plotting functions

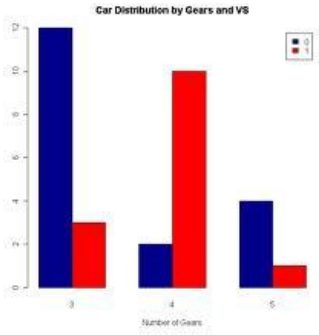
barplot()

Chart with rectangular bars with lengths proportional to the values that they represent.

Stacked bar plot

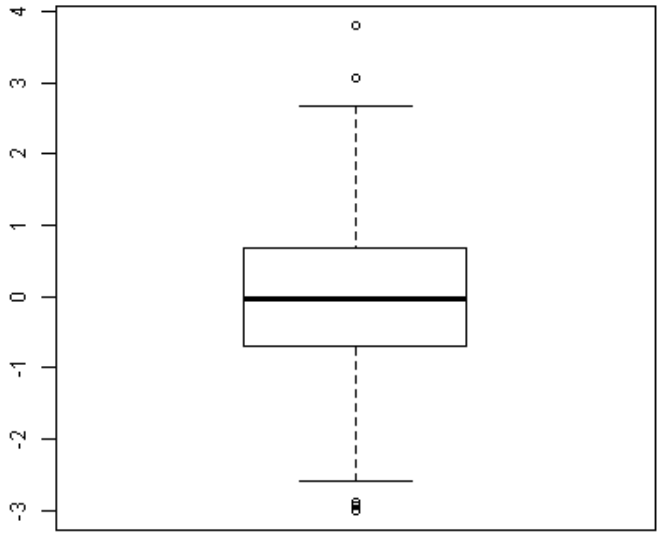


Grouped bar plot



boxplot()

Graphical depicting groups (they underlay statistical distribution) of numerical data through their quartiles.



Dataset Iris

Virginica – Setosa - Versicolor

Get the dataset and other useful things

```
read.csv()
read.table()

?read.table

head(iris)
names(iris)
```

Generic X-Y plotting

#Notation

- #x, y form

plot(iris$Sepal.Length, iris$Sepal.Width)

- #model form (response ~ dependant)

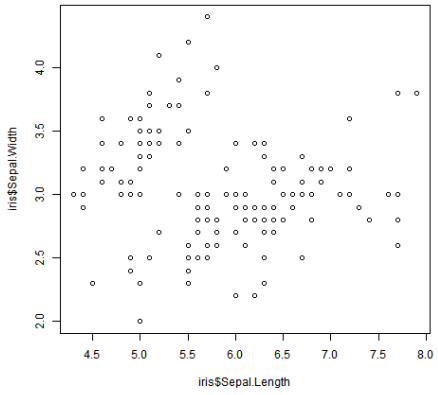
plot(iris$Sepal.Length ~ iris$Sepal.Width)

- #model form with seperate data parameter

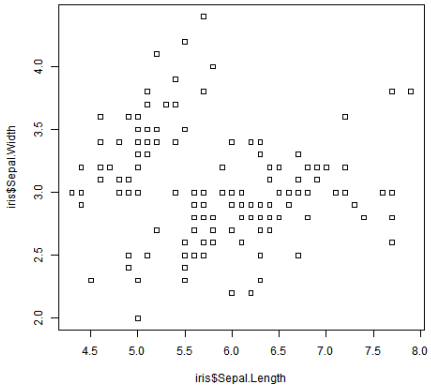
plot(Sepal.Length ~ Sepal.Width, data = iris)

“Building up” a plot

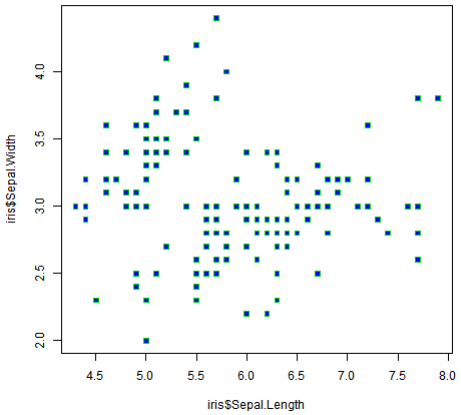
Default plot



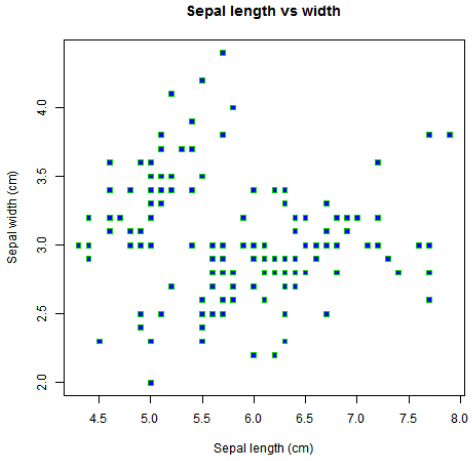
Change symbol



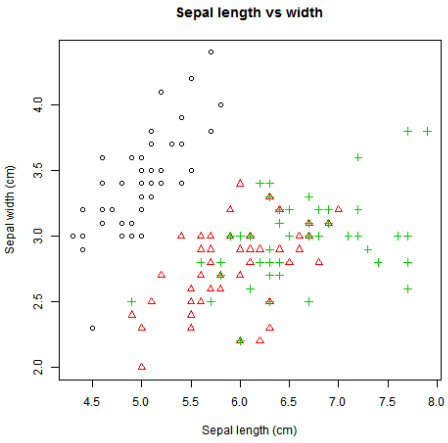
Change colour and background



Add main and axis title



A 3rd dimension (symbols and colours depending on the ‘species)



Important graphics parameters

pch : the plotting symbol (default is an open circle)
col : the plotting colour, specified as a number, string or hex code (**colors()** gives a vector of colours by name)
main : character string for the main title
xlab : character string for the x-axis label
ylab : character string for the y-axis label
lty : the line type (default is solid line)
lwd : the line width, specified as an integer multiple

Other parameters on the **par()** function

bg : the background colour
mar : the margin size
oma: the outer margin size (by default is zero for all sides)
mfrow : number of plots per row, column (filled row-wise)
mfcol : number of plots per row, column (filled column-wise)

Lines and legend

legend()

Add a legend

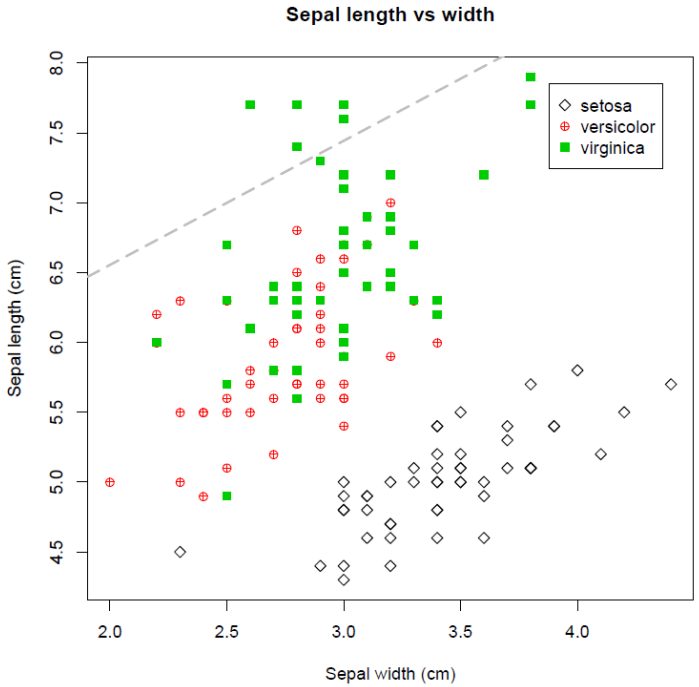
Legend (x, y, legend =, col=, pch=, ...)

locator(1) : get the coordinates for the legend

abline()

Add a regression line (y ~x)

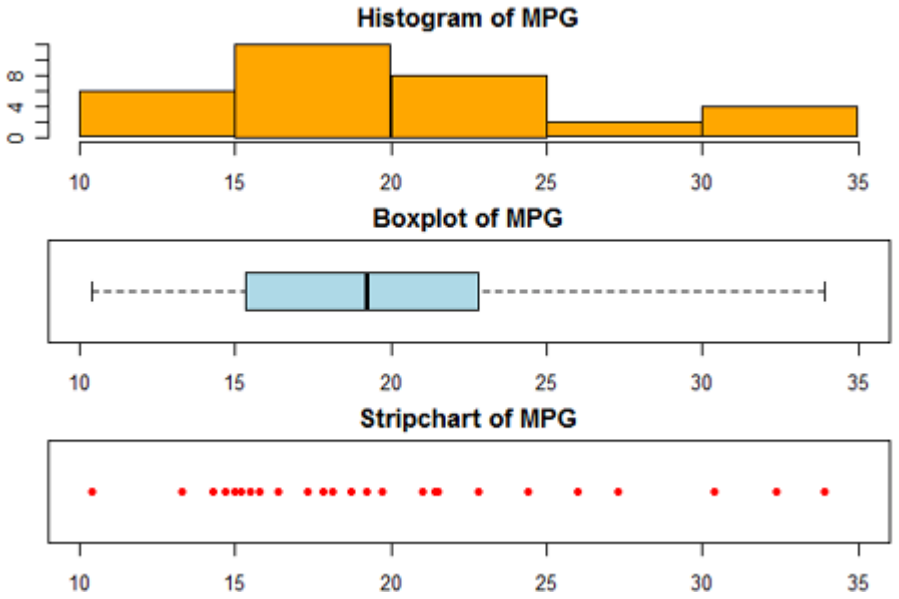
Abline(lm(y~x), col =)



Multiple plots

?matrix
m <- matrix (c(1:3), ncol=1, byrow=TRUE)
layout (m)

All the plots that you want to include, e.g.:
histo()
boxplot()
plot()



Export the plots

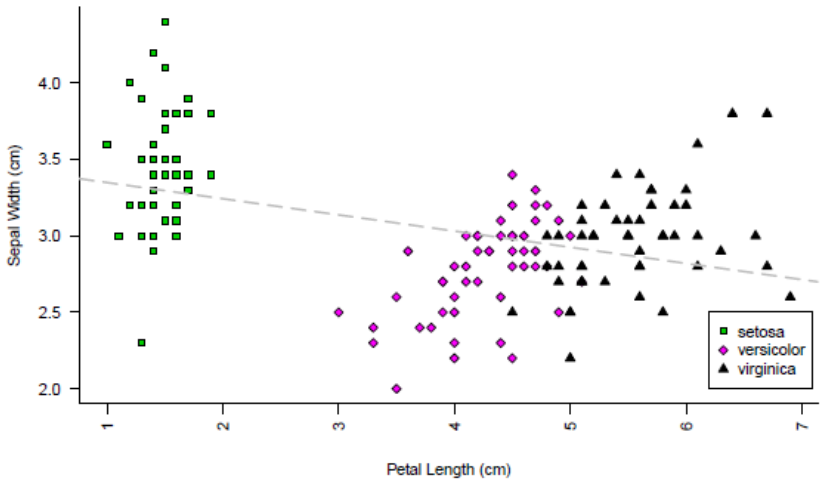
- Options: pdf(), jpeg(), png()…

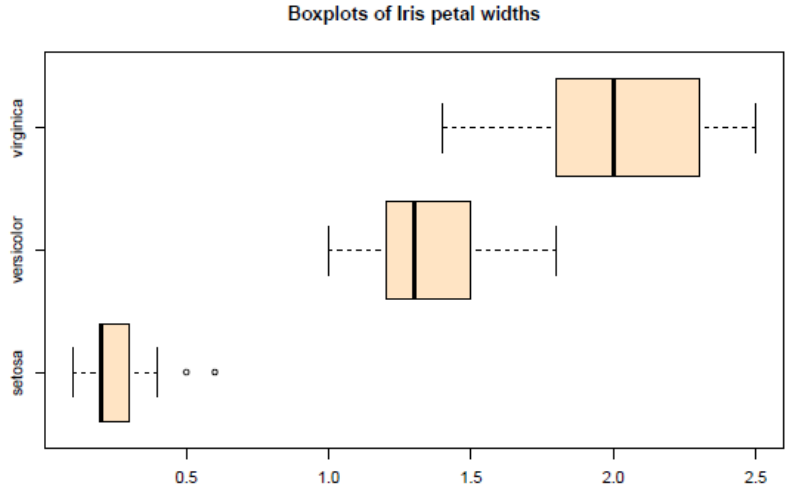
pdf(file =“name.pdf”, width, height…)

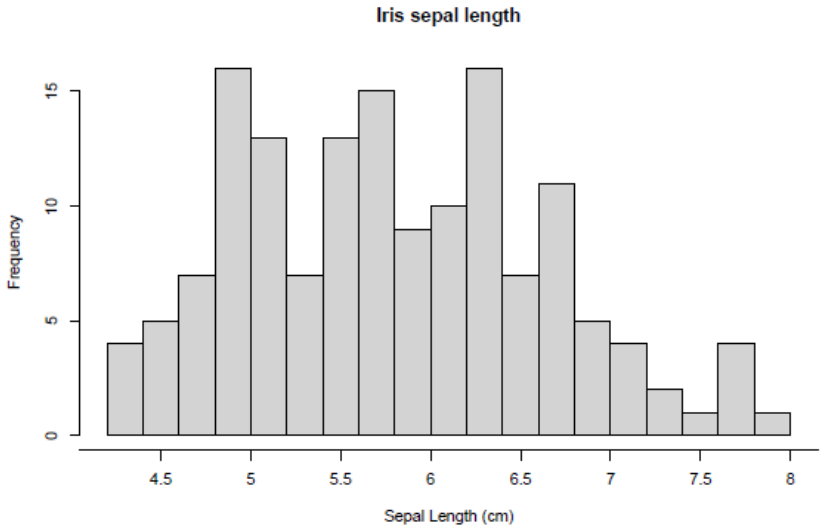
The plot that you want to include: histo(), boxplot(), or plot(), and legend, regression line, etc.

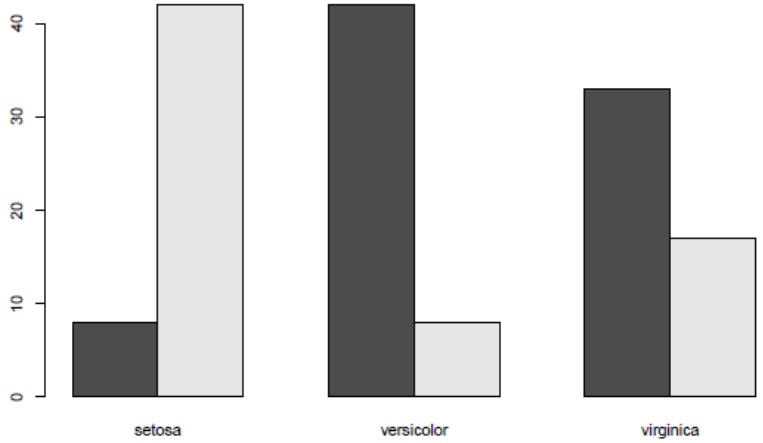
dev.off()

Final exercise

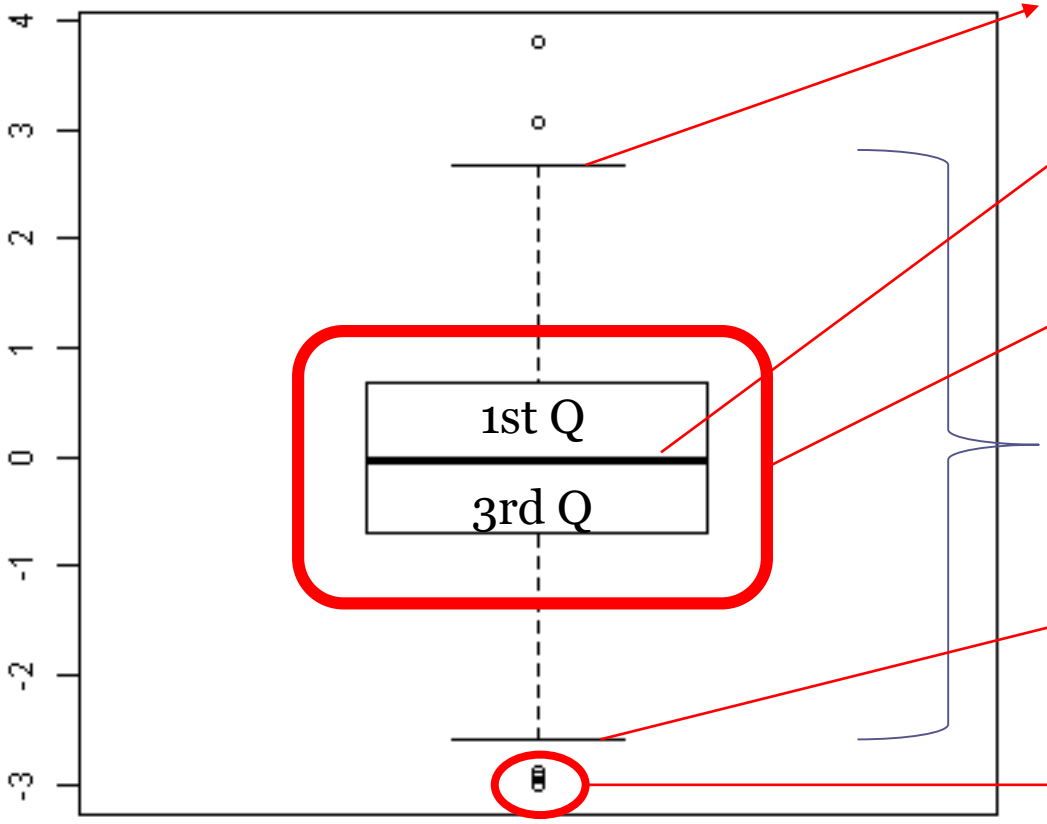








Boxplot interpretation



Maximum

2nd Q = median

Degree of dispersion, likely range of variation

Variability outside the upper and lower quartiles

Minimum

Outliers = individual points