

R Graphics

An overview of possibilities

Kévin Cazelles and Nicolas Casajus Université du Québec à Rimouski

Outline

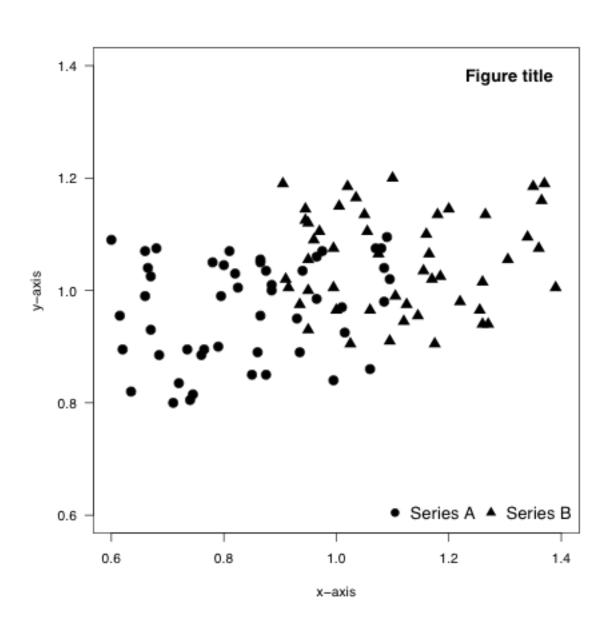
- ✓ Introduction
- ✓ The graphics package basis
- ✓ Composition and multi-panel plotting
- ✓ Graphics automation and exporting
- ✓ Resources
- ✓ Exercises

The importance of graphics

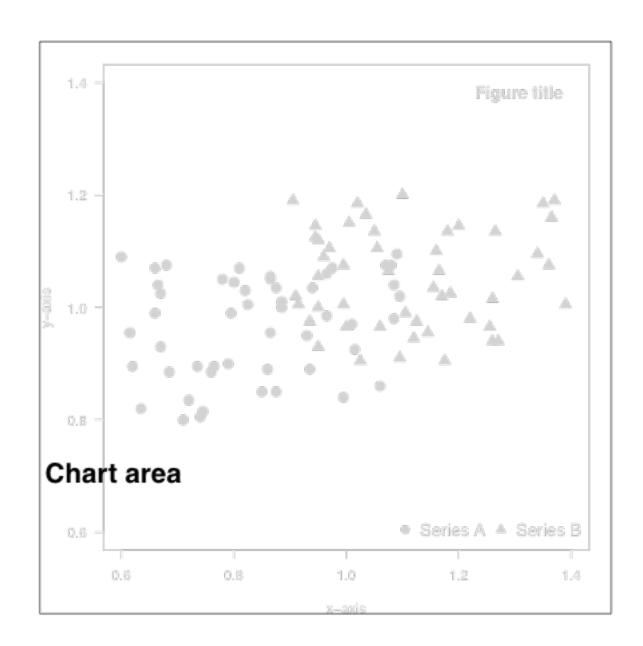
A picture is worth a thousand words

- ✓ Visual summary of data / information
- ✓ More efficient than table and text
- ✓ Useful for exploring data
 - trends, correlations, cycles, outliers, etc.
- ✓ Essential for presenting results
- ✓ But a bad graph can lie about data
- ✓ And sometimes a graphic is not the solution

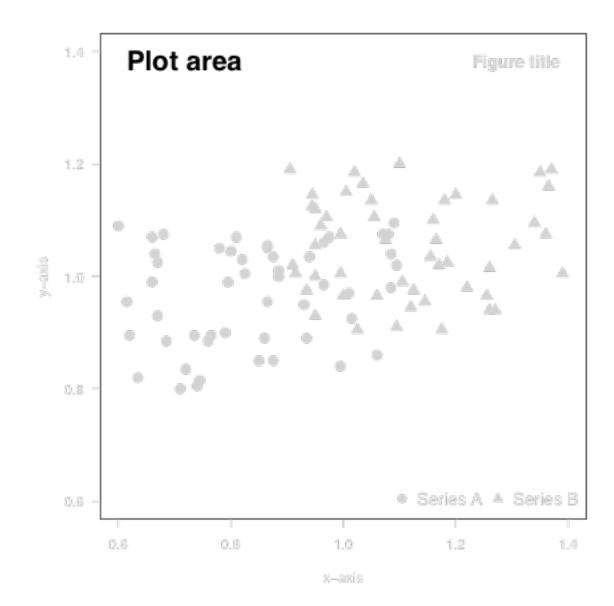




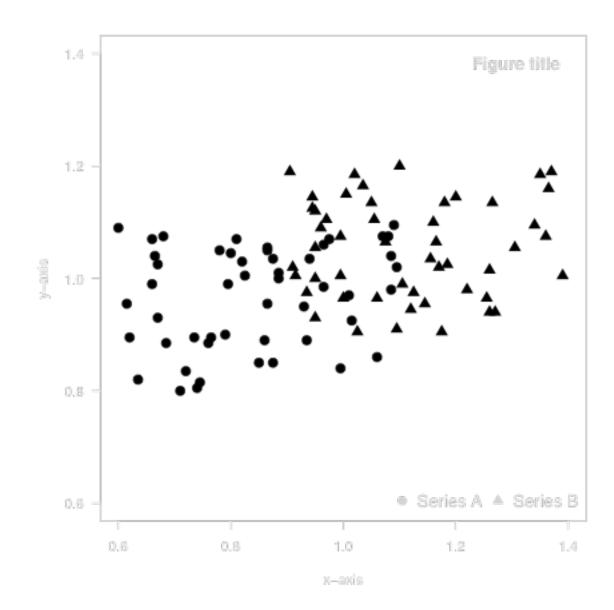
✓ Chart area



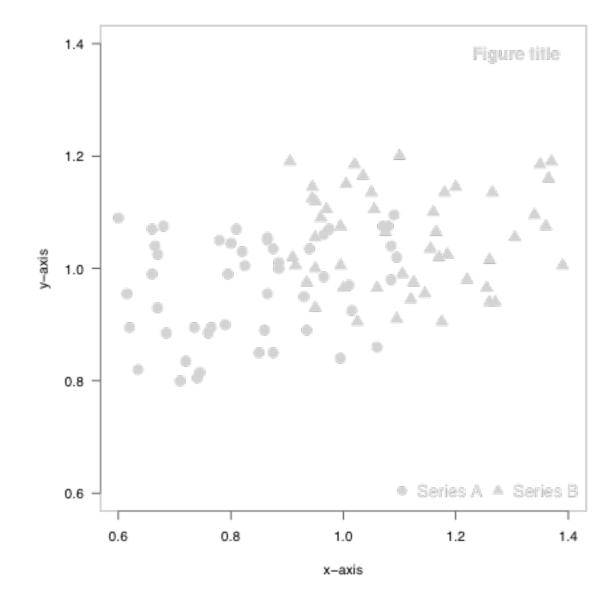
- ✓ Chart area
- ✓ Plot area



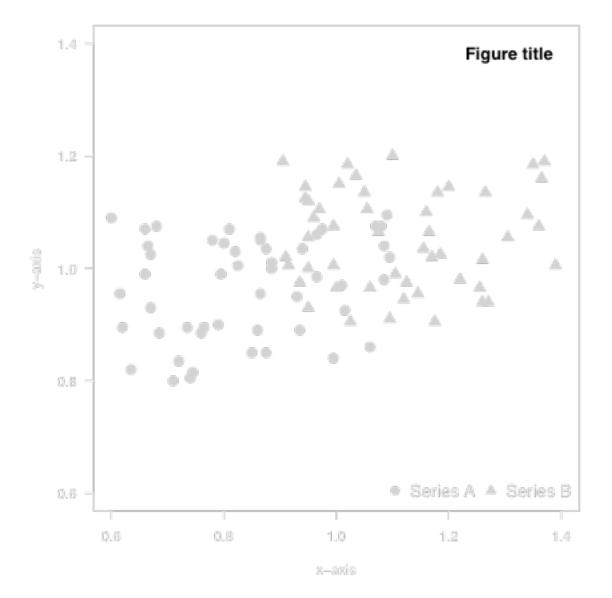
- ✓ Chart area
- ✓ Plot area
- ✓ Data representation



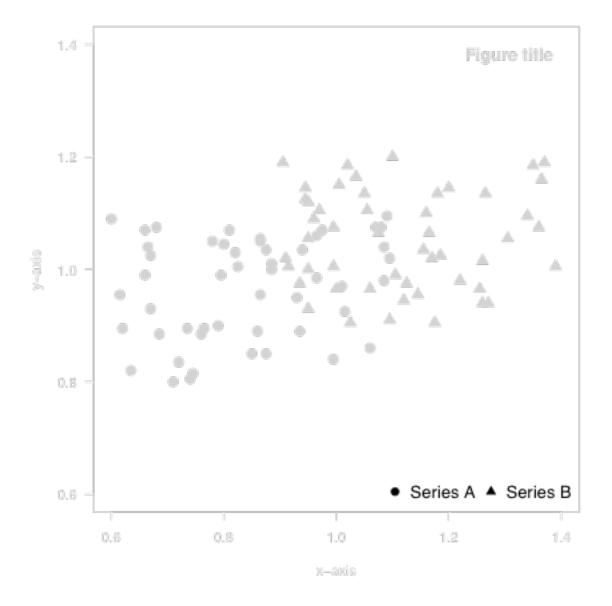
- ✓ Chart area
- ✓ Plot area
- ✓ Data representation
- ✓ Axis (scaling, labeling)



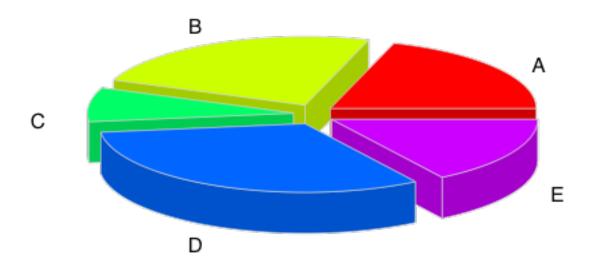
- ✓ Chart area
- ✓ Plot area
- ✓ Data representation
- ✓ Axis (scaling, labeling)
- ✓ Figure title



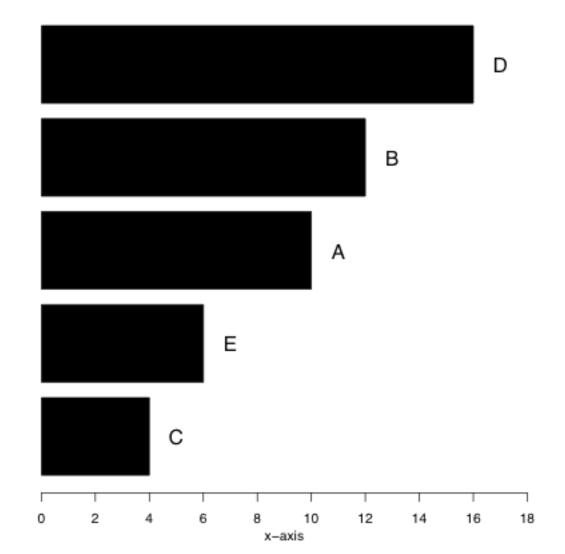
- ✓ Chart area
- ✓ Plot area
- ✓ Data representation
- ✓ Axis (scaling, labeling)
- ✓ Figure title
- ✓ Legend



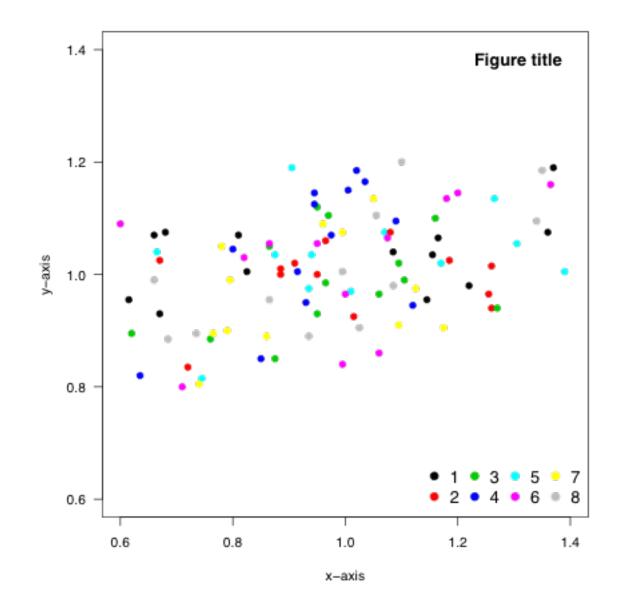
- ✓ Do not use pie chart
- ✓ Do not use 3D (never)
- ✓ Use consistent colors



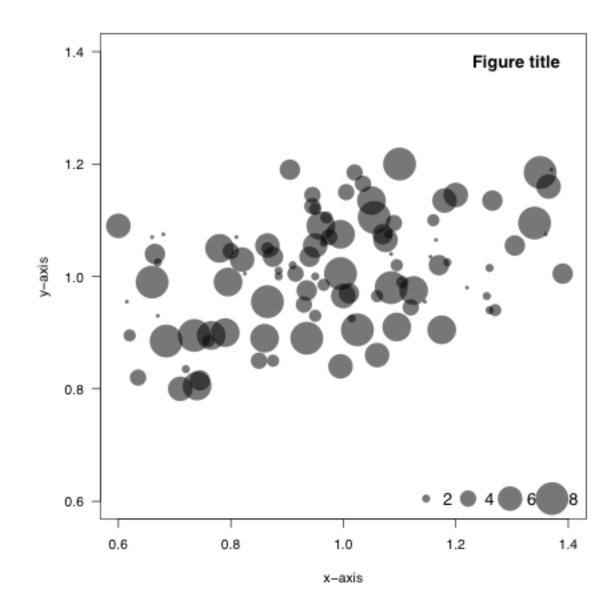
- ✓ Do not use pie chart
- ✓ Do not use 3D (never)
- ✓ Use consistent colors
- ✓ Do prefer this representation



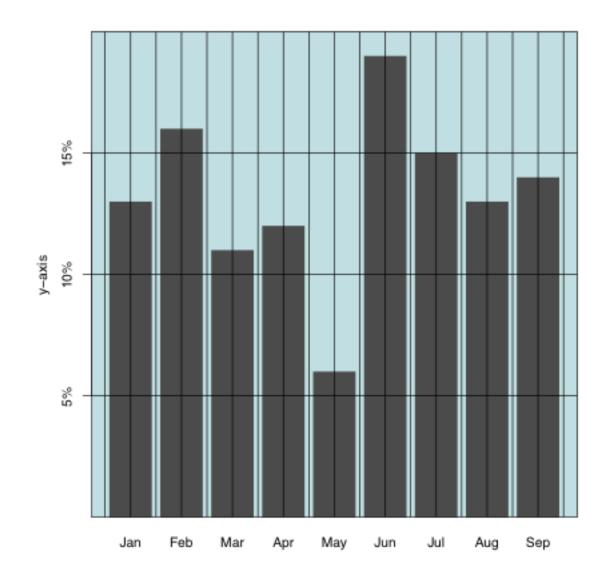
- ✓ Do not use more than 6 colors
- ✓ Do not use high contrast color



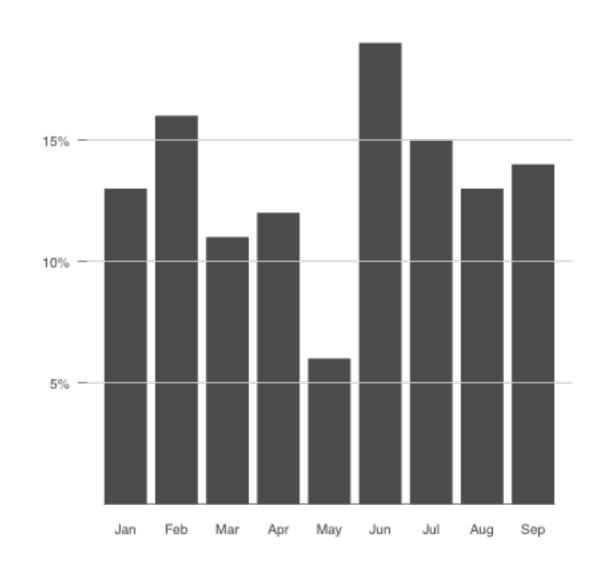
- ✓ Do not use more than 6 colors
- ✓ Do not use high contrast color
- ✓ Sometimes sizes and symbols are better



- ✓ Write textual informations horizontally
- ✓ Do not use distracting elements
- ✓ Do not add chart junk



- ✓ Write textual informations horizontally
- ✓ Do not use distracting elements
- ✓ Do not add chart junk
- ✓ Think about the Data-Ink ratio (Tufte, 1983)



Each element of a graph has to help understanding data

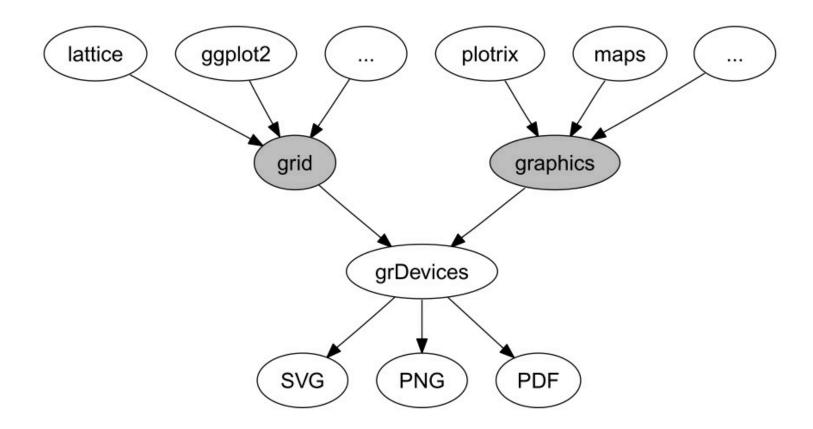
Choose the graphic that shows what you want to show

The R system



- ✓ Software environment for statistical computing and graphics
- ✓ Open-source, free and multiplatform
- ✓ Widely used in the scientific community
- ✓ Programming language
- ✓ Implementation of the S programming language
- ✓ The core system is extended through user-created packages
- ✓ You can do what you want with R

The R system



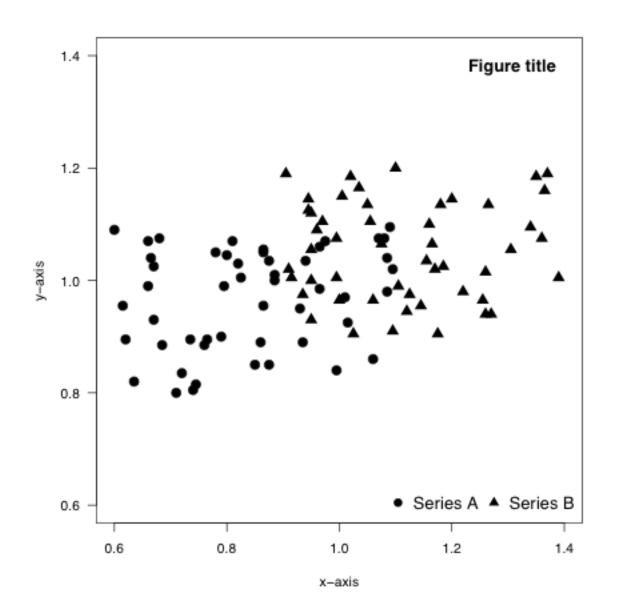
Murrell, P. (2015) The gridGraphics Package. The R Jounal.

Graphical packages - graphics

- ✓ Base package
- ✓ S-like plotting functions
- ✓ Contains the famous function plot()
- ✓ And a lot of well-known functions:

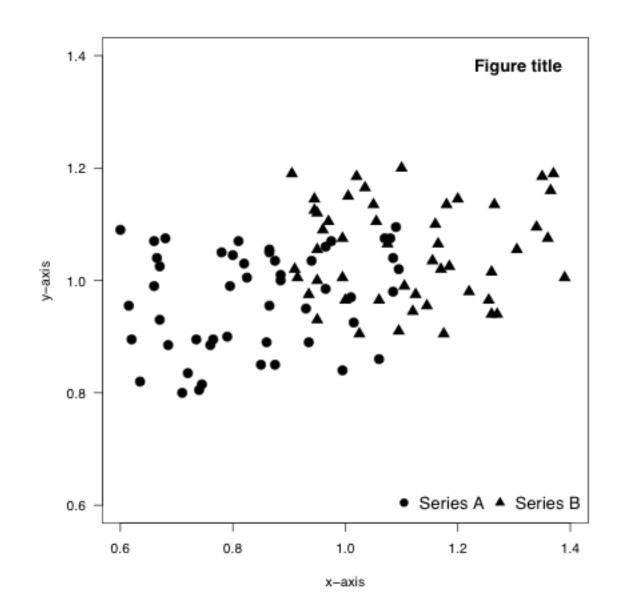
```
boxplot(), barplot(), hist(),
lines(), points(), legend(), etc.
```

```
plot(x, y, ...)
points(x, y, ...)
title(main, ...)
legend(legend, ...)
```



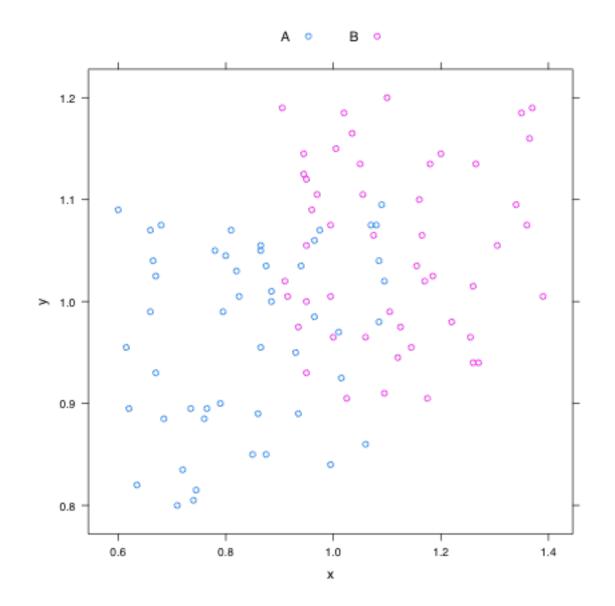
Graphical packages - grid

- ✓ An alternative set of graphical functions
- ✓ Well-suited for developers
- ✓ ggplot2 is based on this package



Graphical packages - lattice

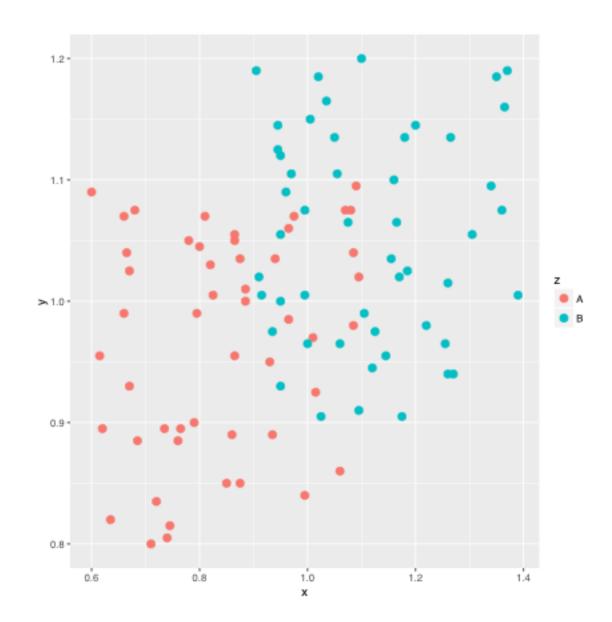
- ✓ Based on the grid package
- ✓ High-level system inspired by Trellis graphics
- ✓ Specialized on multivariate data
- ✓ and multipanel figures



Graphical packages - ggplot2

- ✓ Also based on the grid package
- ✓ A complete plotting system for R
- ✓ Based on the Grammar of Graphics
- ✓ But introduces its own syntax
- ✓ and requires a long time to master it

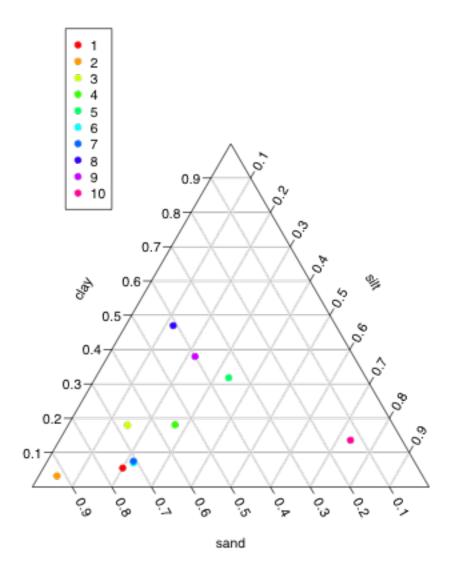
✓ See the QCBS workshop on ggplot2



Graphical packages - plotrix

- ✓ Based on the graphics package
- ✓ Contains a lots of specialized plots (i.e. polar plots)
- ✓ and various labeling, axis and color scaling functions

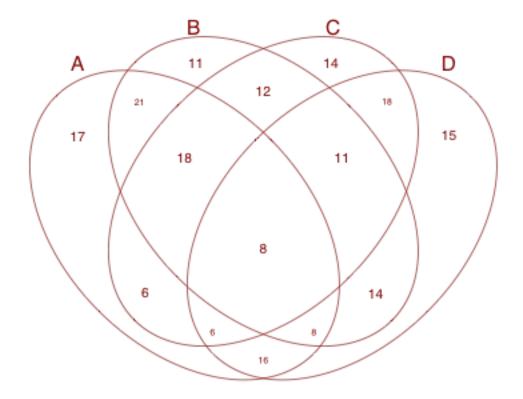
```
library(plotrix)
data(soils)
triax.plot(...)
```



Graphical packages - gplots

- ✓ Based on the graphics package
- ✓ Adds enhanced versions of standard plots (e.g. boxplot2)
- ✓ and some extra functions (e.g. Venn diagram)

```
library(gplots)
venn(...)
```



Graphical packages - others

- ✓ More than 80 others graphical packages
- ✓ For an overview see this R task view
- ✓ For a more exhaustive list see this post
- ✓ On this workshop we will only use the graphics package

- ✓ Default values of graphical parameters are stored in par()
- ✓ par() is an object
 - we can get the value of a parameter
- ✓ par() is also a function
 - we can change the default values

```
## How many graphical parameters?
length(par())
## [1] 72

## Let's get the default value of text color
par()$col
## [1] "black"

## Let's set 'red' for text color
par(col = 'red')

## Check
par()$col
## [1] "red"

## We're good!
```

✓ Important: when you change the value of one parameter, the new value affects all the graphs until the graphical window is closed

- ✓ Important: when you change the value of one parameter, the new value affects all the graphs until the graphical window is closed
- ✓ A recommendation:
 - Save the default par(): opar <- par()
 - Change the values: par(col='red')
 - Do the graph
 - Restaure the old par(): par(opar)

- ✓ Important: when you change the value of one parameter, the new value affects all the graphs until the graphical window is closed
- ✓ A recommendation:
 - Save the default par(): opar <- par()
 - Change the values: par(col = 'red')
 - Do the graph
 - Restaure the old par(): par(opar)
- ✓ Some graphical parameters can also be changed directly in plotting functions

High-level vs. low-level plotting functions

High-level plotting functions

- ✓ Open a new graphical window
- ✓ Or erase the content of the previous window
- ✓ Examples: plot(), boxplot(), barplot(), hist(), etc.

High-level vs. low-level plotting functions

High-level plotting functions

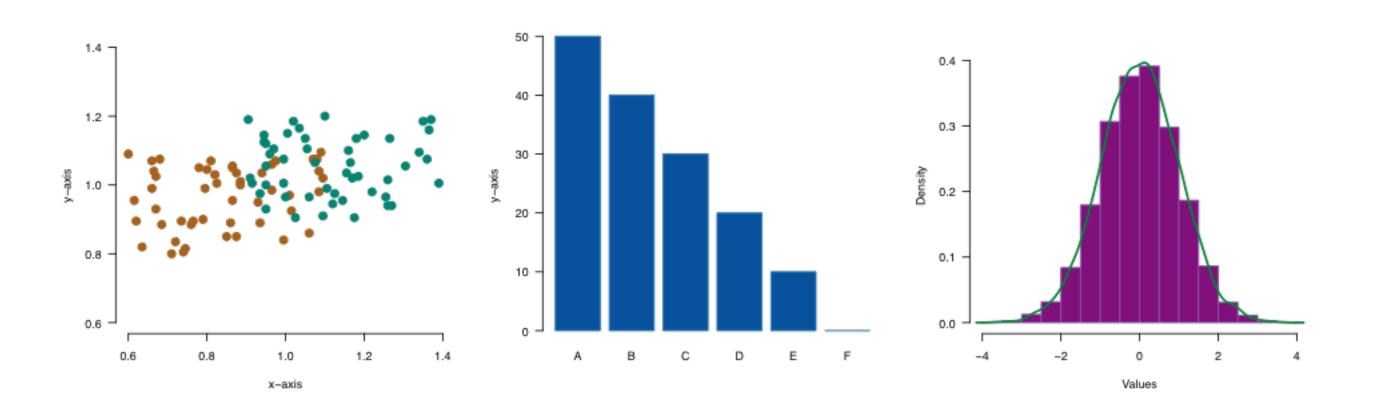
- ✓ Open a new graphical window
- ✓ Or erase the content of the previous window
- ✓ Examples: plot(), boxplot(), barplot(), hist(), etc.

Low-level plotting functions

- ✓ Work only when a graphical window is open
- ✓ Add content to the active window
- ✓ Examples: lines(), points(), axis(), legend(), etc.

High-level vs. low-level plotting functions

You only need to know one high-level plotting function: plot()



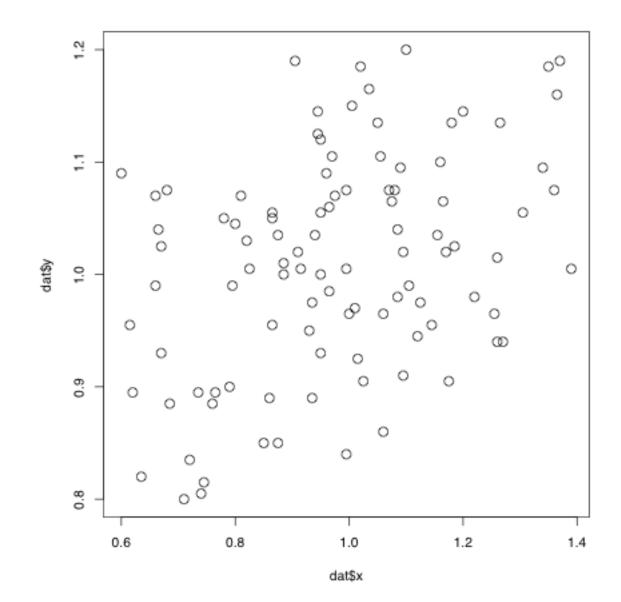
Let's take a look at the data

- ✓ Random data with no particular sense
- ✓ Three variables:
 - x and y: quantitative variables
 - z: qualitative variable (factor)

- ✓ The default plot
- ✓ Quite ugly, isn't it?

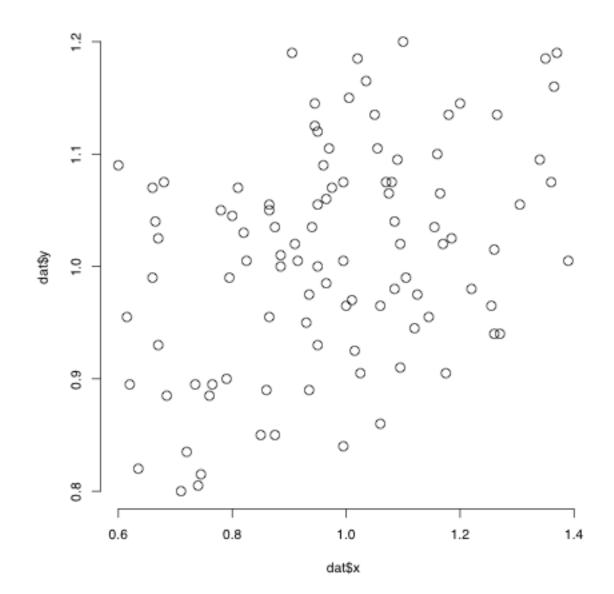
```
plot(x = dat$x, y = dat$y)
```

✓ Now we are going to remove each component of the graph to create an empty plot



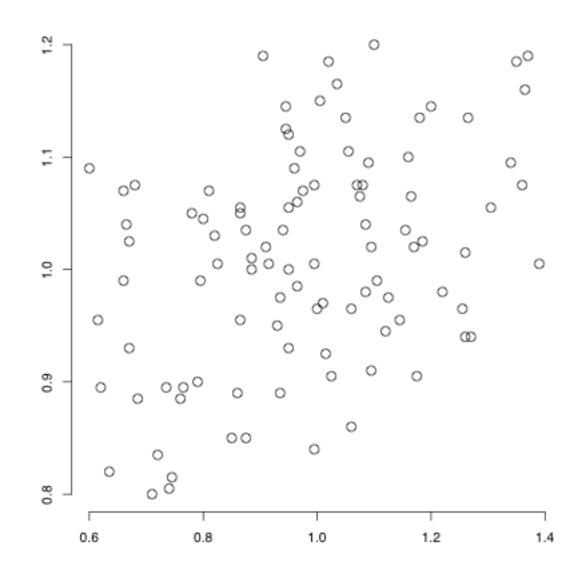
- ✓ First let's remove the box
- ✓ with the argument bty (default: 'o')

```
plot(x = datx, y = daty,
bty = 'n')
```



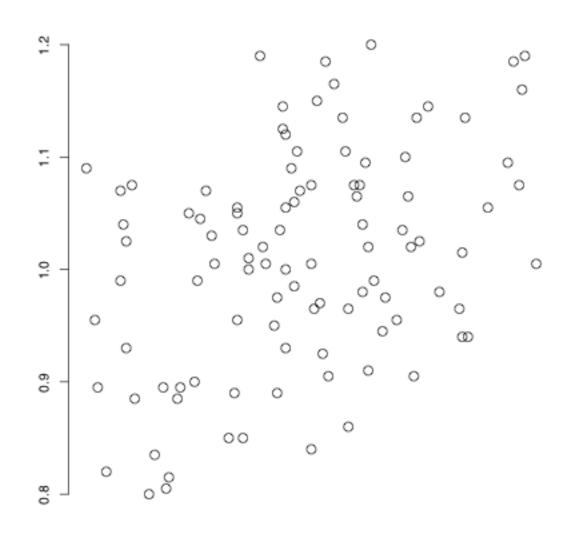
- ✓ Now let's remove the textual annotation
- ✓ with the argument ann (default: 'TRUE')

```
plot(x = dat$x, y = dat$y,
    bty = 'n',
    ann = FALSE)
```



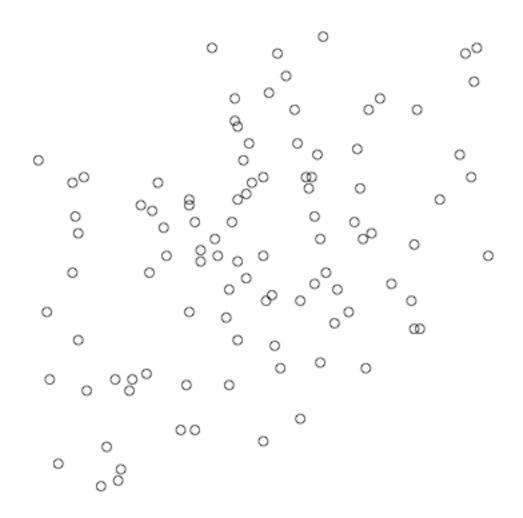
- ✓ Let's remove the x-axis
- ✓ with the argument xaxt (default: 's')

```
plot(x = dat$x, y = dat$y,
  bty = 'n',
  ann = FALSE,
  xaxt = 'n')
```



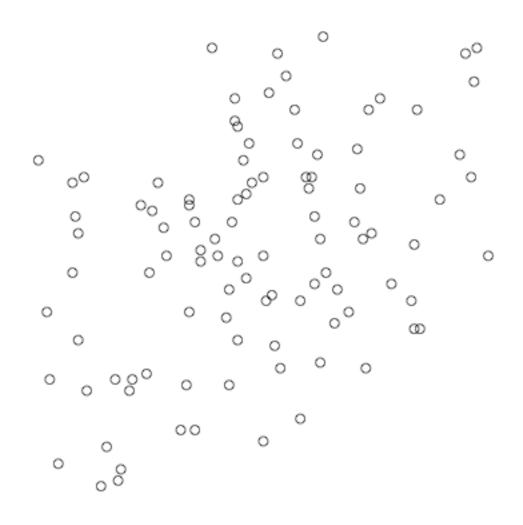
- ✓ And the y-axis
- ✓ with the argument yaxt (default: 's')

```
plot(x = dat$x, y = dat$y,
    bty = 'n',
    ann = FALSE,
    xaxt = 'n',
    yaxt = 'n')
```



- ✓ Using axes=FALSE is the same as:
- ✓ bty='n'+xaxt='n'+yaxt='n'

```
plot(x = dat$x, y = dat$y,
    ann = FALSE,
    axes = FALSE)
```



- ✓ Finally let's remove data
- ✓ with the argument type (default: 'p')

```
plot(x = dat$x, y = dat$y,
    ann = FALSE,
    axes = FALSE,
    type = 'n')
```

- ✓ Finally let's remove data
- ✓ with the argument type (default: 'p')

```
plot(x = dat$x, y = dat$y,
    ann = FALSE,
    axes = FALSE,
    type = 'n')
```

- ✓ In an empty plot, visual information is not displayed but the graph is defined in the window
- ✓ It is now possible to use low-level plotting
 functions such as points() or axis()

An empty plot, and now what?

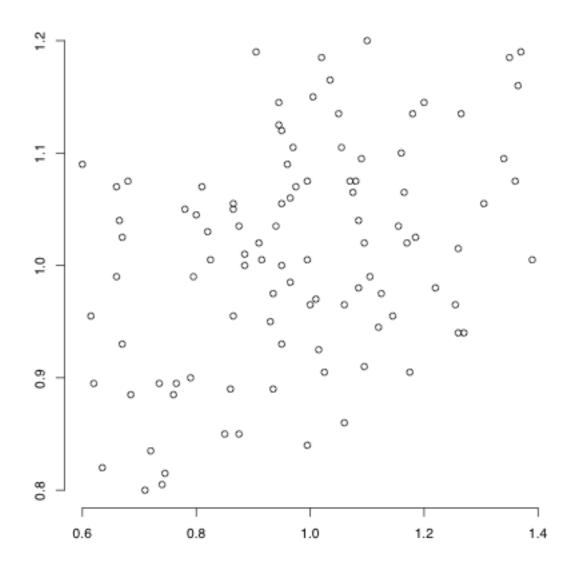
- ✓ Now we've got an empty plot
- ✓ We are going to add some informations to:
 - discover useful low-level plotting functions,
 - improve the quality of the default plot
- ✓ Let's go!

Adding points

- ✓ We will use the function points()
- ✓ It shares some arguments with the function plot()

```
## Empty plot
plot(x = dat$x, y = dat$y, ann = FALSE,
    bty = 'n', type = 'n')

## Adding points (default settings)
points(x = dat$x, y = dat$y)
```

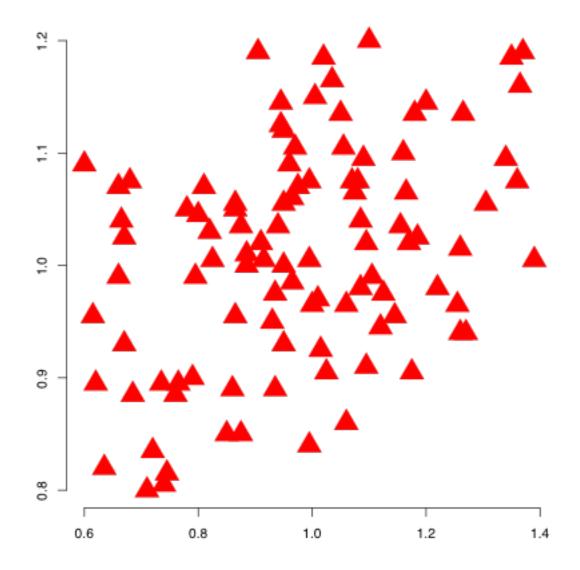


Adding points

- ✓ We can customize the points with:
 - cex, the size
 - col, the color
 - pch, the symbol

```
## Empty plot
plot(x = dat$x, y = dat$y, ann = FALSE,
    bty = 'n', type = 'n')

## Adding points (user settings)
points(x = dat$x, y = dat$y,
    cex = 3, col = 'red', pch = 17)
```

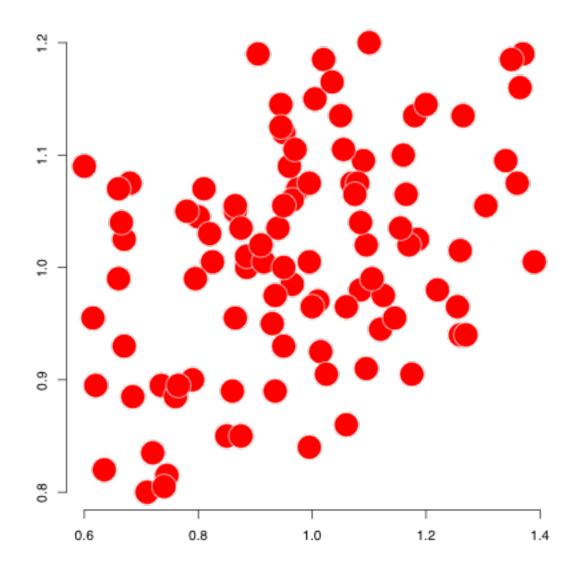


Adding points

- ✓ Some symbols have two colors:
 - col: the border color,
 - bg: the background color
- ✓ This is the case for pch = 21 to 25

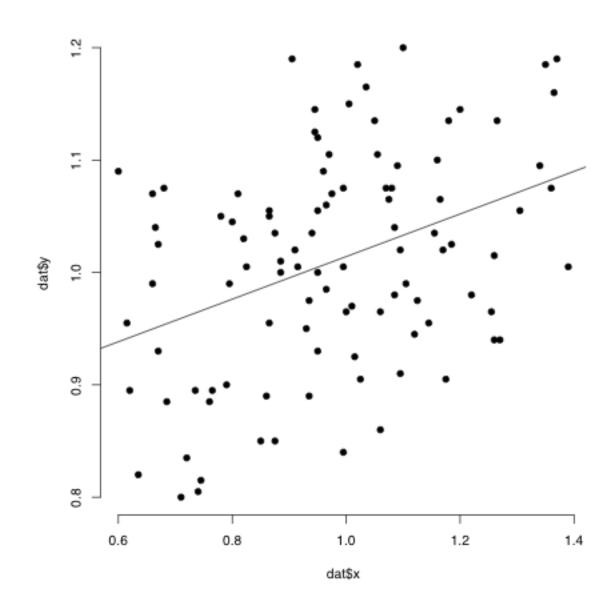
```
## Empty plot
plot(x = dat$x, y = dat$y, ann = FALSE,
    bty = 'n', type = 'n')

## Adding points (user settings)
points(x = dat$x, y = dat$y,
    cex = 4, pch = 21,
    col = 'white', bg = 'red')
```

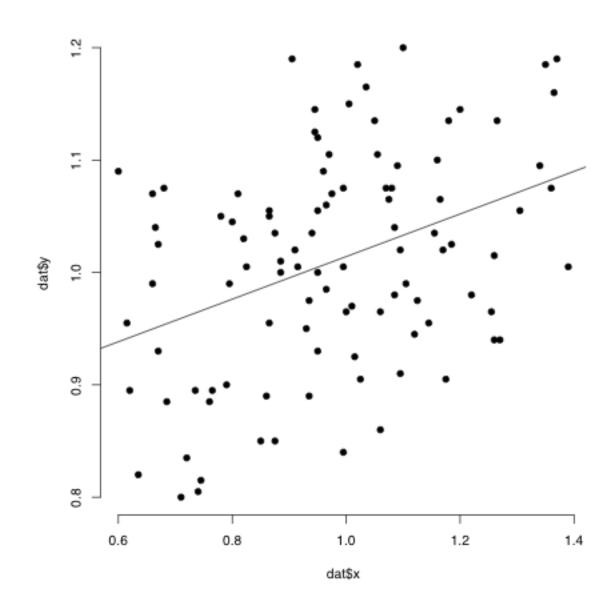


- ✓ Four functions allow to plot lines:
 - points()
 - lines()
 - abline()
 - segments()
- ✓ We will illustrate these functions with a linear regression

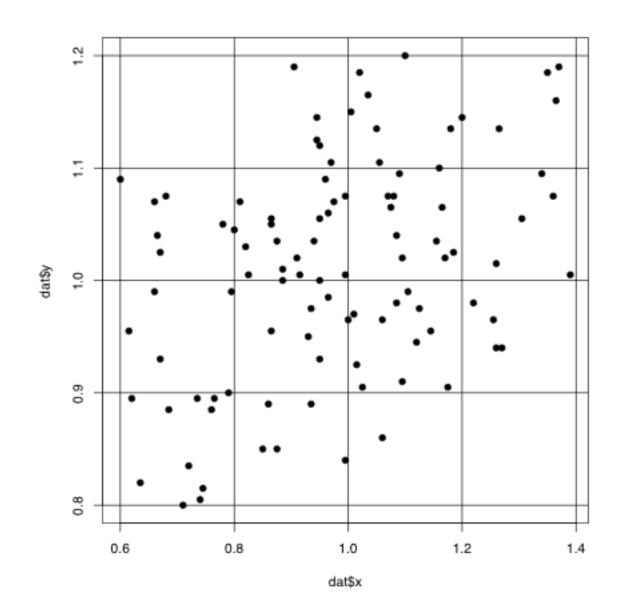
- ✓ First, let's try the function abline()
- ✓ with the first way



✓ The second way is to specify model parameters



✓ The function abline() allows to draw horizontal and vertical lines

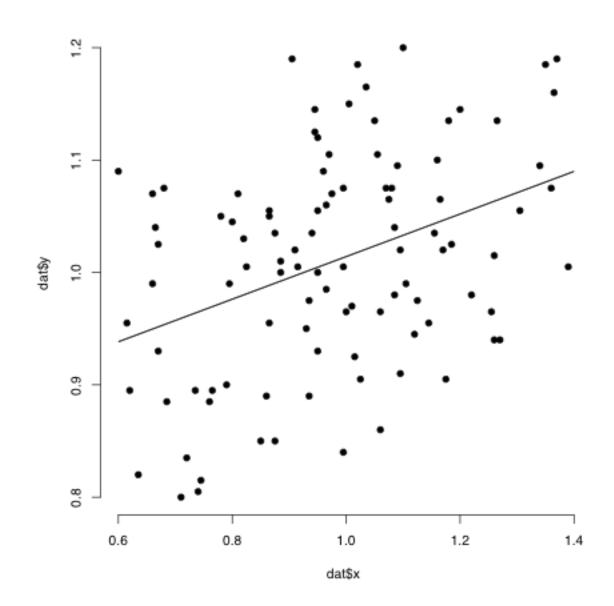


- ✓ Now take a look at the functions lines() and points()
- ✓ But first, we are going to predict the model on new data

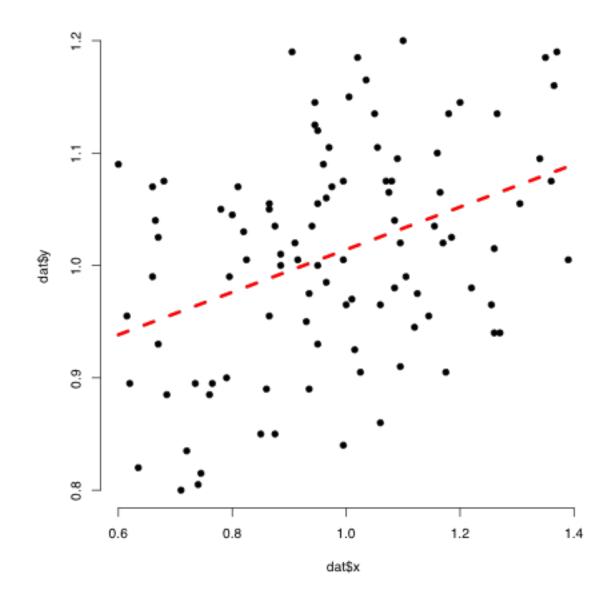
```
## New data frame
mat <- data.frame(x = seq(0.6, 1.4, by = 0.05))

## Model prediction
ypred <- predict(object = mod, newdata = mat)</pre>
```

✓ Let's add model regression with the functions lines() and points()



- ✓ We can customize the lines with:
 - lwd, the line width
 - col, the line color
 - lty, the line type



- ✓ To add a polygon, the function is polygon()
- ✓ A special, the rectangle can be drawn with rect()
- ✓ Let's predict again the model, but this time with the standard error

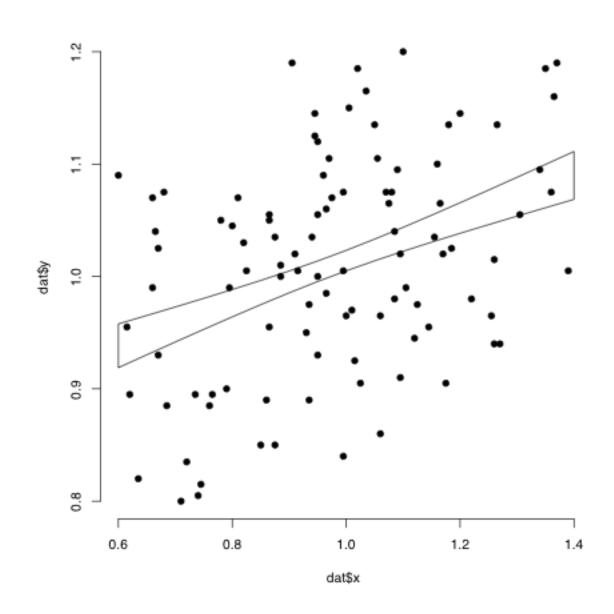
- ✓ We are going to add the error envelope with the function polygon()
- ✓ So, let's calculate the coordinates of this envelope

```
## Superior interval
xsup <- mat[ , 'x']
ysup <- ypred$fit + ypred$se.fit

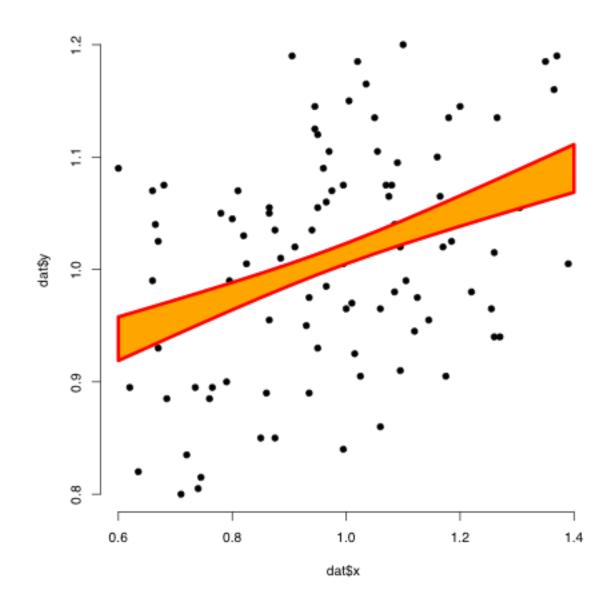
## Inferior interval
xinf <- mat[ , 'x']
yinf <- ypred$fit - ypred$se.fit

## Reverse sort of inf.
xinf <- mat[nrow(mat) : 1, 'x']
yinf <- yinf[length(yinf) : 1]</pre>
```

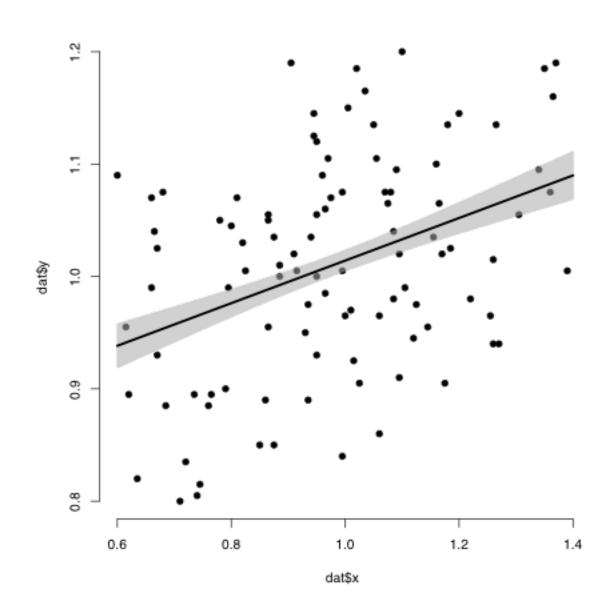
✓ Let's add model error envelope



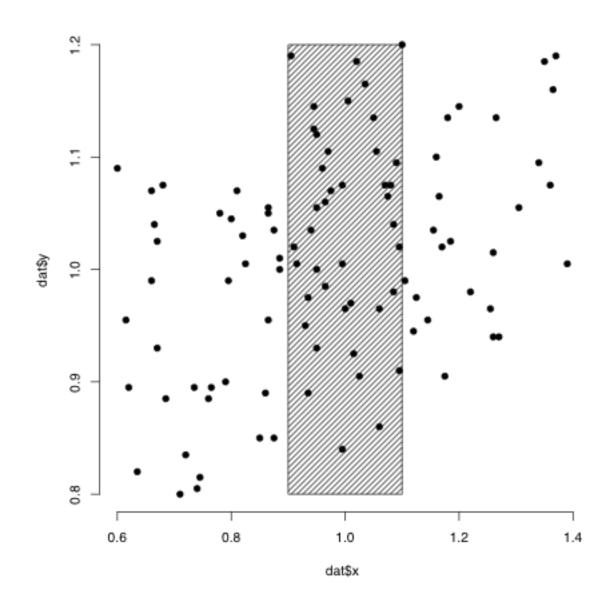
- ✓ We can customize the polygon with:
 - border, the border color
 - col, the background color
 - lwd, the border width



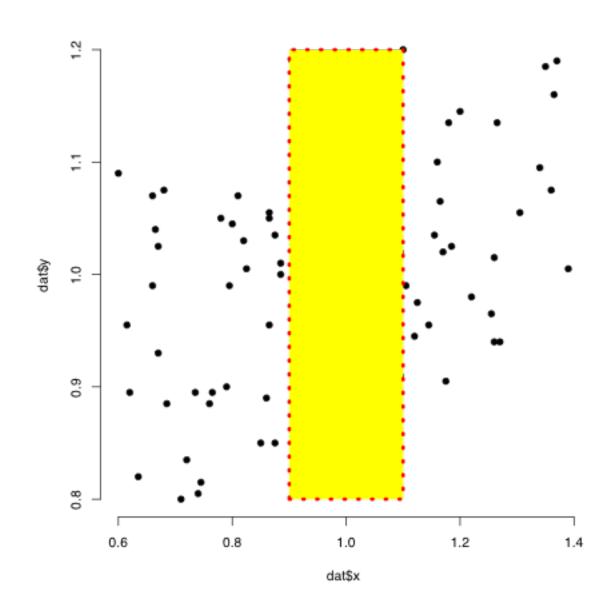
✓ Finally



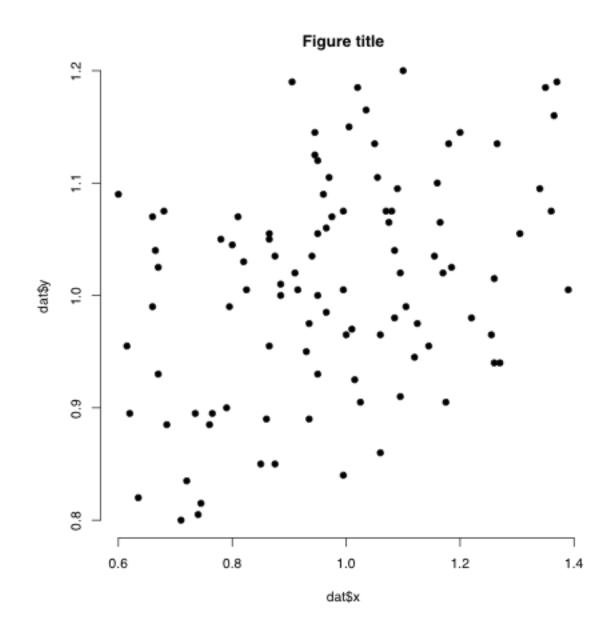
- ✓ The function rect() is appropriated
 when you want to add draw rectangle
- ✓ Here is an example



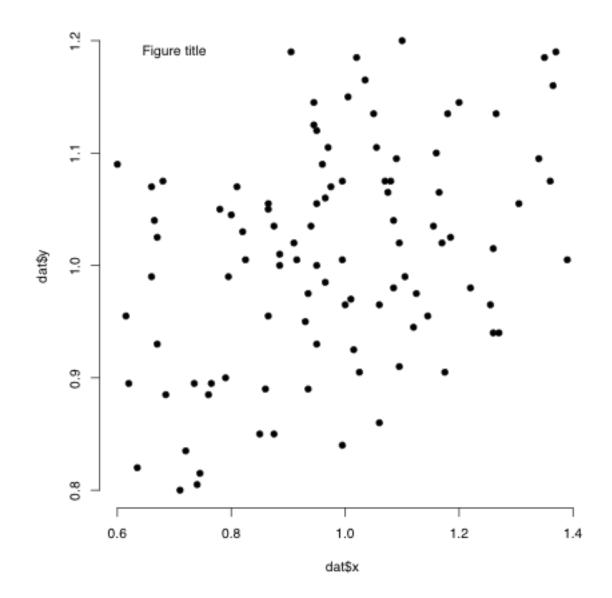
✓ You also can customize the rectangle



- ✓ Let's see now how to add text
- ✓ First let's add a main title with the function title()

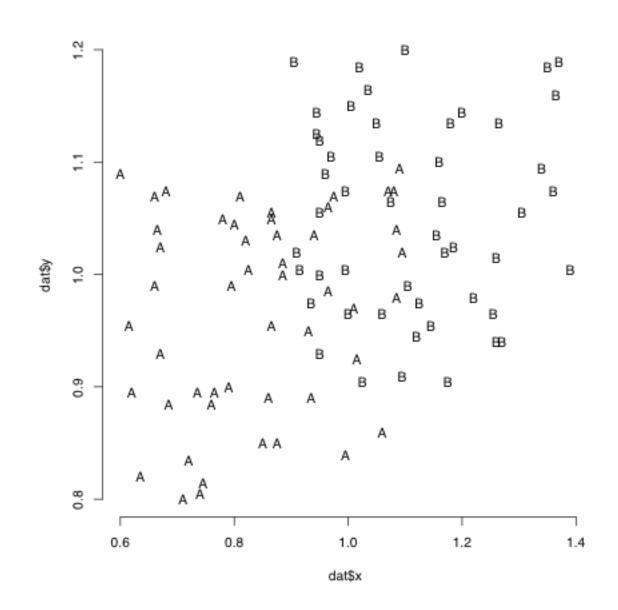


- ✓ What about adding text in the plot area?
- ✓ We will use the function text()
- ✓ Here is a first example

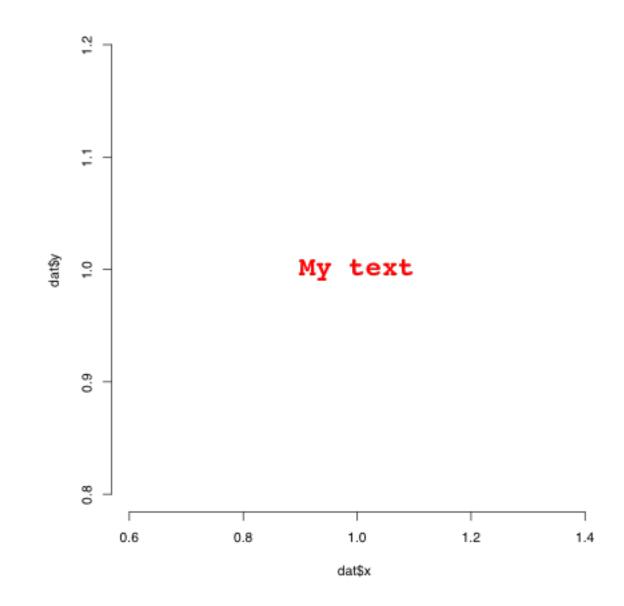


✓ And another example

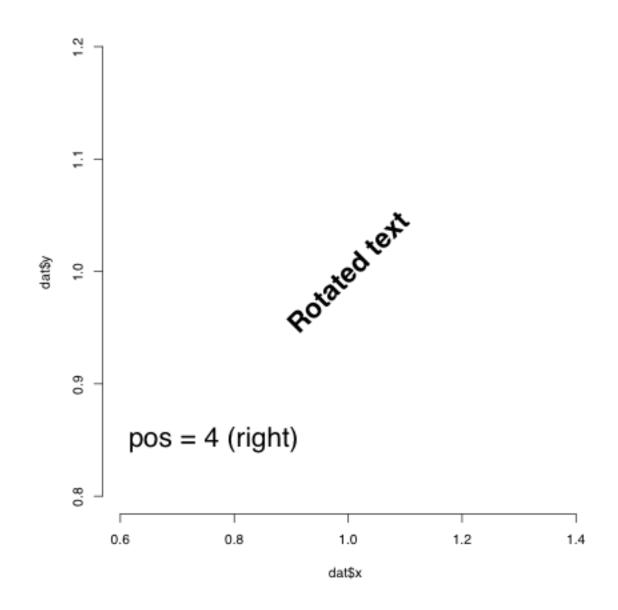
```
## Empty plot
plot(x = dat$x, y = dat$y, bty = 'n', type = 'n')
## Adding text
text(x = dat$x, y = dat$y, labels = dat$z)
```



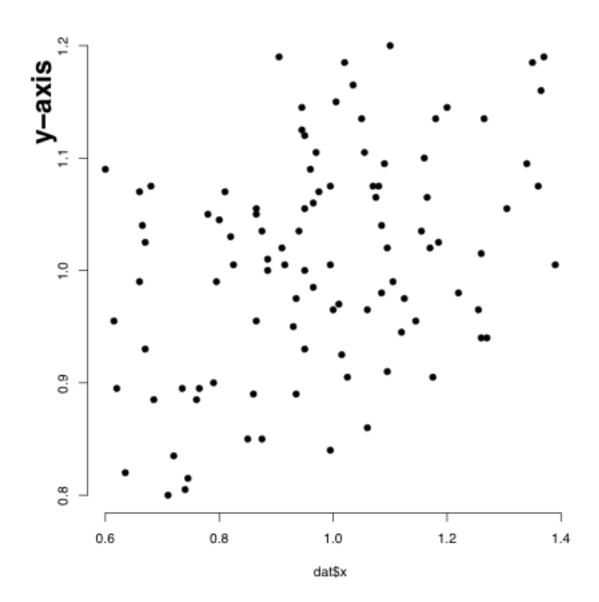
- ✓ Let's customize a little the text with:
 - cex, the size
 - col, the color
 - font, the font (bold, italic, etc.)
 - family, the typeface



- ✓ Finally, let's customize the orientation and position
 - srt, the rotation angle
 - pos, the position from coordinates



- ✓ But the function text() can't add text outside the plot area
- ✓ except if par(xpd=TRUE)
- ✓ We will use the function mtext()



Back to colors

✓ R has some predefine colors palettes

```
## Basic colors
palette()
## [1] "black" "red" "green3" "blue" "cyan" "magenta" "yellow"
## [8] "gray"

## and 657 others colors
colors()[1:10]
## [1] "white" "aliceblue" "antiquewhite" "antiquewhite1"
## [5] "antiquewhite2" "antiquewhite3" "antiquewhite4" "aquamarine"
## [9] "aquamarine1" "aquamarine2"
```

Back to colors

- ✓ You can define colors in the RGB system with rgb()
- ✓ You specify value between 0 to 1 for each primary colors
- ✓ For example:

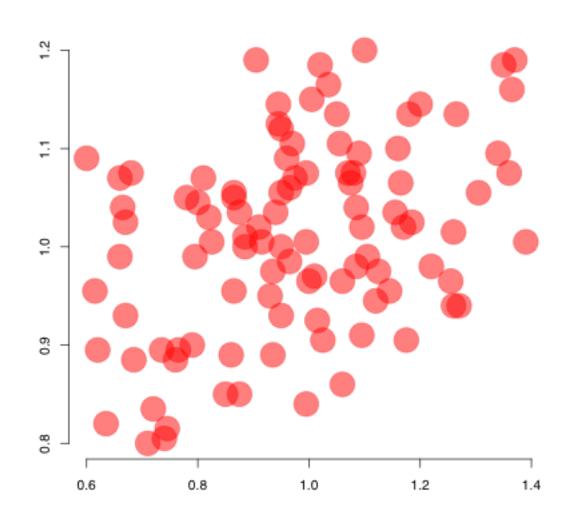
Back to colors

- ✓ You can define colors in the RGB system with rgb()
- ✓ You specify value between 0 to 1 for each primary colors
- ✓ For example:

- ✓ The alpha argument controls for opacity (default = 1)
- ✓ Its values vary from 0 (transparent) to 1 (opaque)
- ✓ For example

Back to colors

✓ Let's see an application



Back to colors

- ✓ You can also define colors in the hexadecimal system
- ✓ Each primary colors is define by two values varying from 0 to 9 and A to F
- ✓ For example:

```
red <- '#FF0000'
yellow <- '#FFFF00'
gray1 <- '#888888'
```

Back to colors

- ✓ You can also define colors in the hexadecimal system
- ✓ Each primary colors is define by two values varying from 0 to 9 and A to F
- ✓ For example:

```
red <- '#FF0000'
yellow <- '#FFFF00'
gray1 <- '#888888'
```

- ✓ To add transparency, we have to add two hexadecimal character at the end
- ✓ For example,

```
## Transparent red
redp <- '#FF000088'</pre>
```

Adding axis

- ✓ The function axis() allows to add axis
- ✓ Here is an example of usage

```
## Empty plot
plot(x = dat$x, y = dat$y, pch = 19)

## Adding top-axis
axis(side = 3, at = seq(0.6, 1.4, by = 0.1),
    labels = seq(0.6, 1.4, by = 0.1), las = 1)

## Adding right-axis
axis(side = 4, at = seq(0.8, 1.2, by = 0.1),
    labels = format(seq(0.8, 1.2, by = 0.1)),
    las = 2)
```

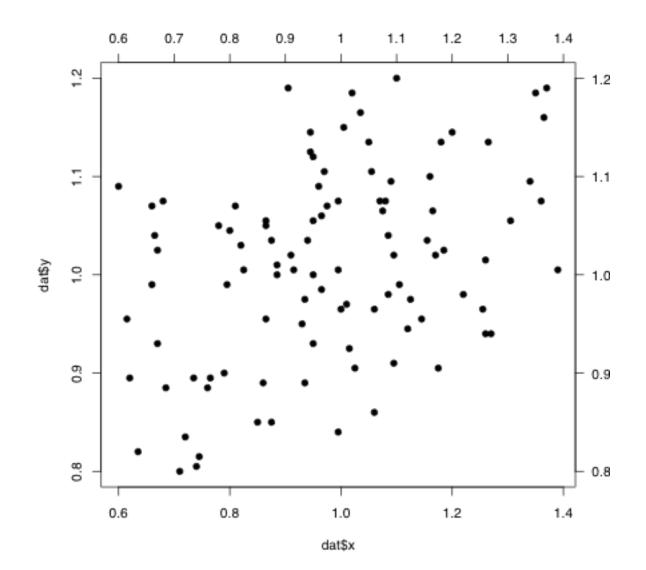
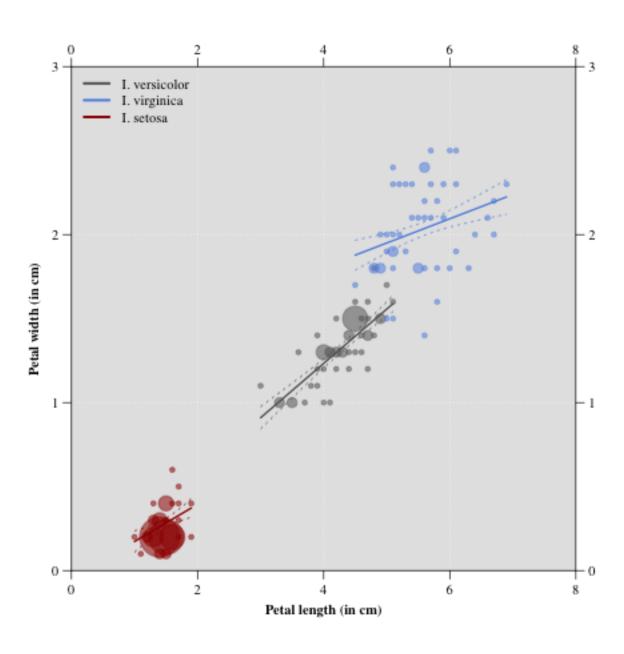


Figure margins

- ✓ To change the figure margins you have to change the values of the parameter mar in the par()
- ✓ The order is the follow: bottom, left, top
 and right
- ✓ For example:

```
par(mar = c(4, 4, 4, 4))
```

Exercise 1



Exercise 1

- ✓ Objective: reproduce this figure
- ✓ Using the dataset iris.txt (Dropbox)

- ✓ The size of the bubble is proportional to n
- ✓ Lines represent regression model and standard error

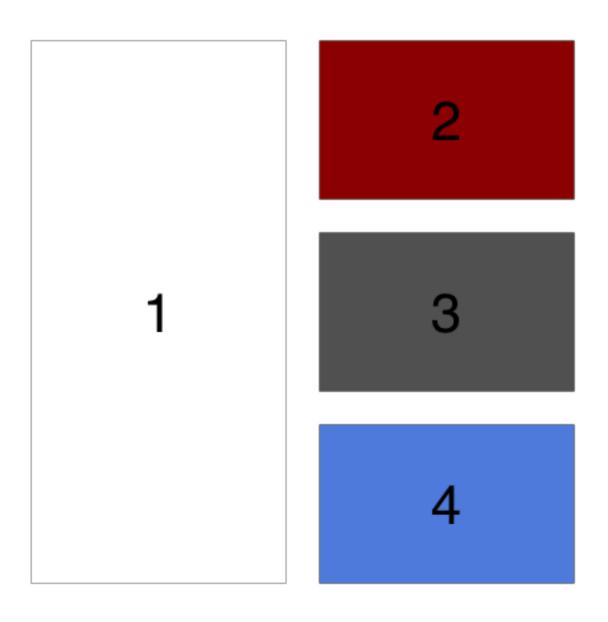
Dividing the output device

Dividing the output device

```
✓ mfrow and mfcol in par()

✓ split.screen()

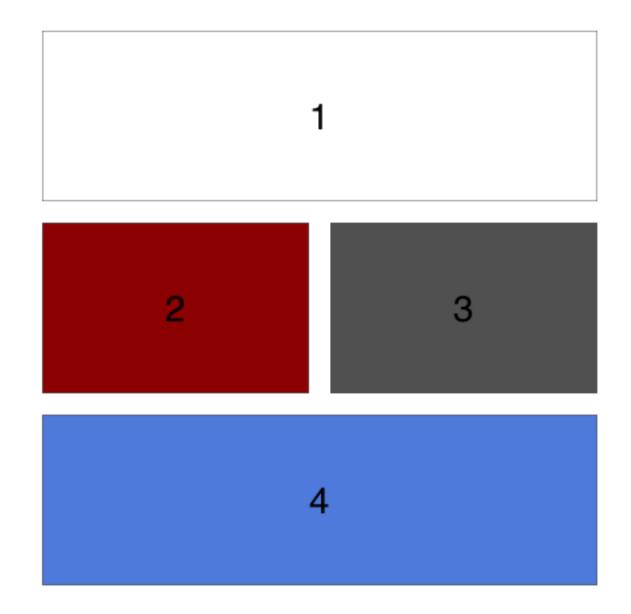
split.screen(c(1, 2))
split.screen(c(3, 1), screen = 2)
```



Dividing the output device

```
✓ mfrow and mfcol in par()
✓ split.screen()
✓ layout()
```

mat_lay <- matrix(c(1,2,4,1,3,4),nrow=3)
layout(mat_lay)</pre>

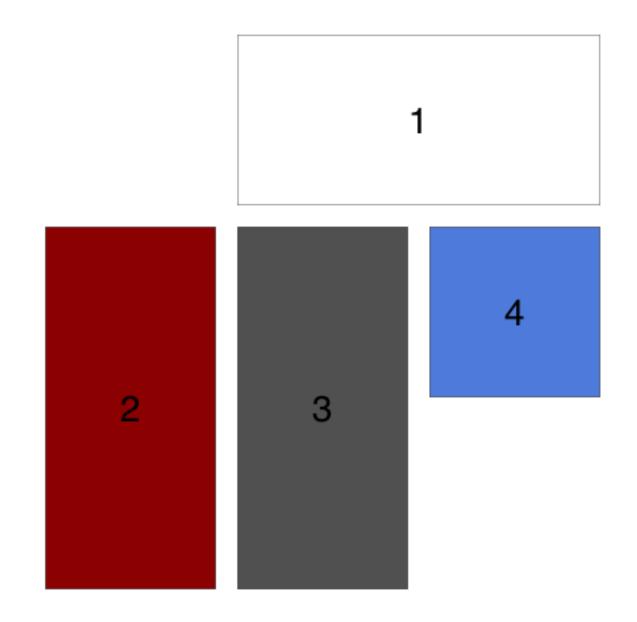


```
mat_lay <- matrix(c(1,2,4,1,3,4), nrow=3)
layout(mat_lay)</pre>
```

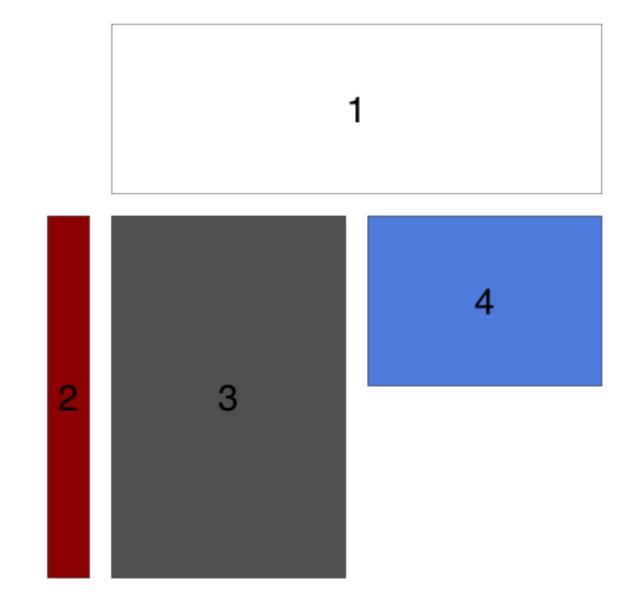
```
## [,1] [,2]
## [1,] 1 1
## [2,] 2 3
## [3,] 4 4
```

```
mat_lay <- matrix(c(0,2,2,1,3,3,1,4,0), nrow=3)
layout(mat_lay)</pre>
```

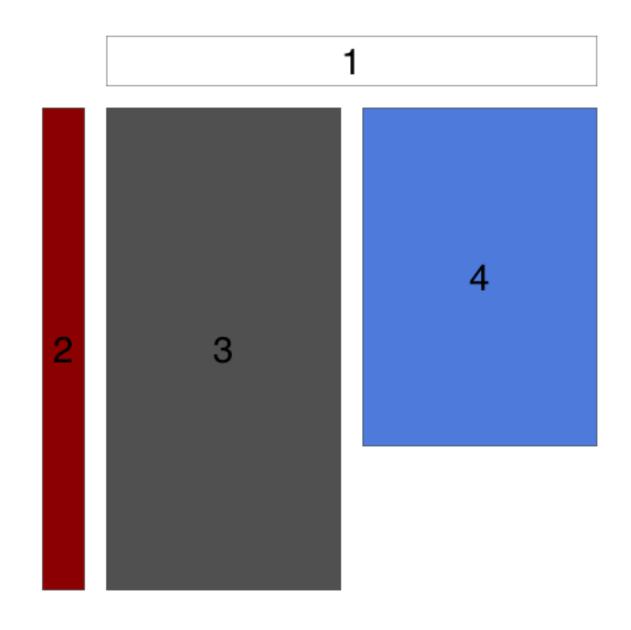
```
## [,1] [,2] [,3]
## [1,] 0 1 1
## [2,] 2 3 4
## [3,] 2 3 0
```



```
mat_lay <- matrix(c(0,2,2,1,3,3,1,4,0),nrow=3)
layout(mat_lay, widths=c(.25,1,1))</pre>
```



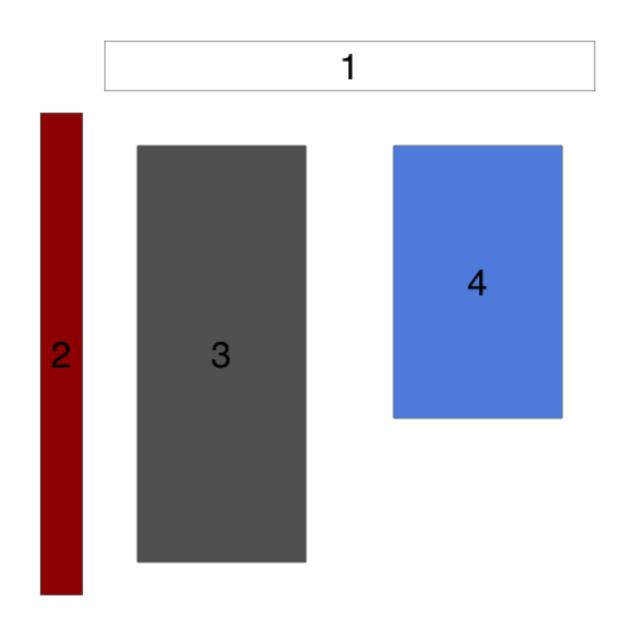
```
mat_lay <- matrix(c(0,2,2,1,3,3,1,4,0),nrow=3)
layout(mat_lay, widths=c(.25,1,1),
heights=c(.25,1,.25))</pre>
```



Combining 'layout()' and 'mar'

```
mat_lay <- matrix(c(0,2,2,1,3,3,1,4,0),nrow=3)
layout(mat_lay, widths=c(.25,1,1),
  heights=c(.2,1,.4))

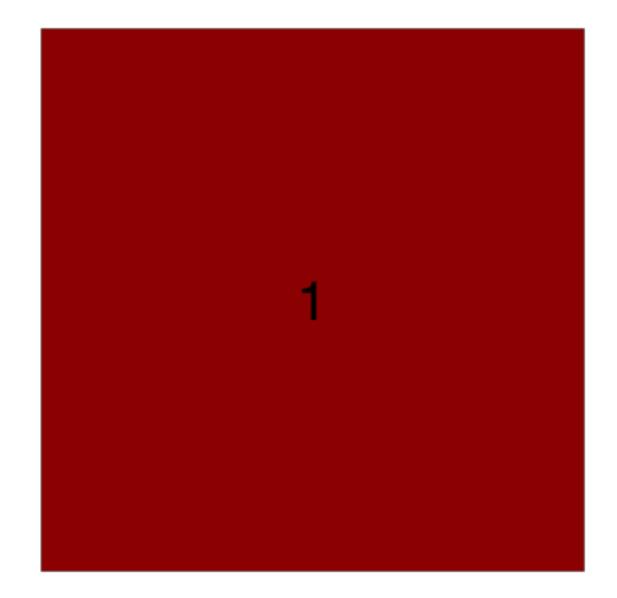
for (i in 1:4) {
  if (i<3) par(mar=rep(1,4)) else par(mar=rep(4,4))
  eplot()
  fillIt(col=i)
  text(0,0, labels=i, cex=4)
}</pre>
```



Embedded plots

- ✓ You must call new=TRUE and specifying fig in par():
- 1. create your first plot;

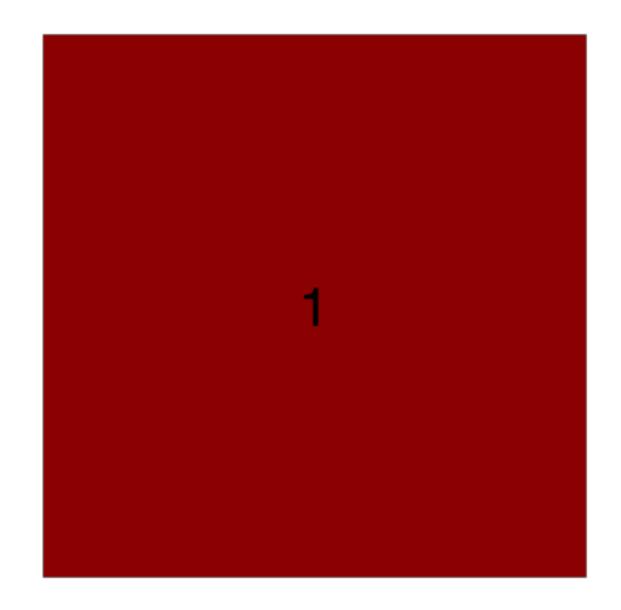
plot(...)



Embedded plots

- ✓ You must call new=TRUE and specifying fig in par():
- 1. create your first plot;
- 2. use par();

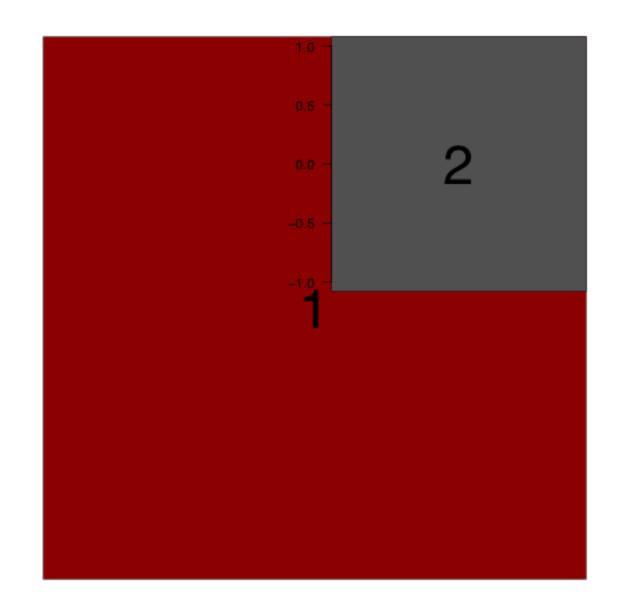
```
plot(...)
par(new=TRUE, fig=c(0.5,1,0.5,1))
```



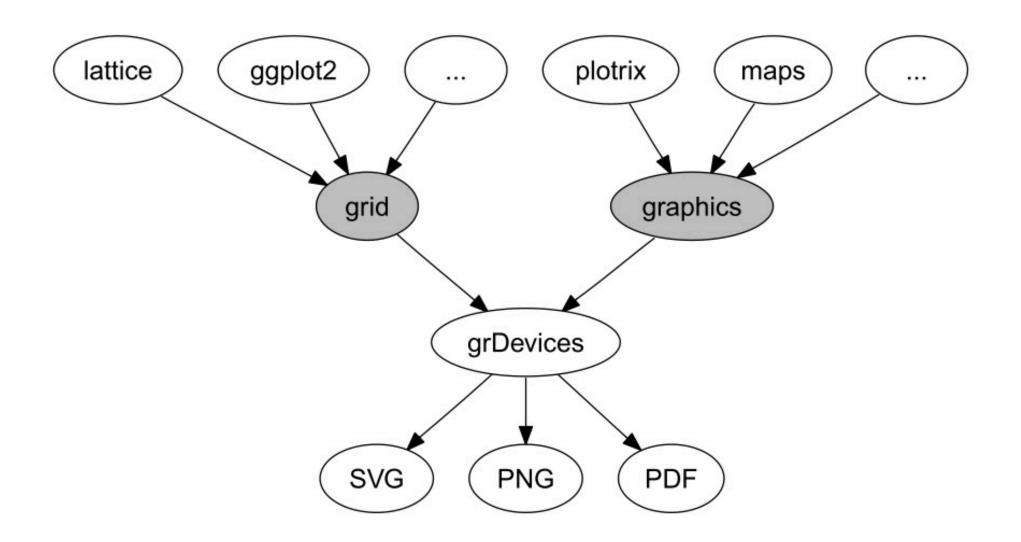
Embedded plots

- ✓ You must call new=TRUE and specifying fig in par():
- 1. create your first plot;
- 2. use par();
- 3. add your embedded plot;

```
plot(...)
par(new=TRUE, fig=c(0.5,1,0.5,1))
plot(...)
```



grDevices



Murrell, P. (2015) The gridGraphics Package. The R Jounal.

grDevices

```
options('device')
```

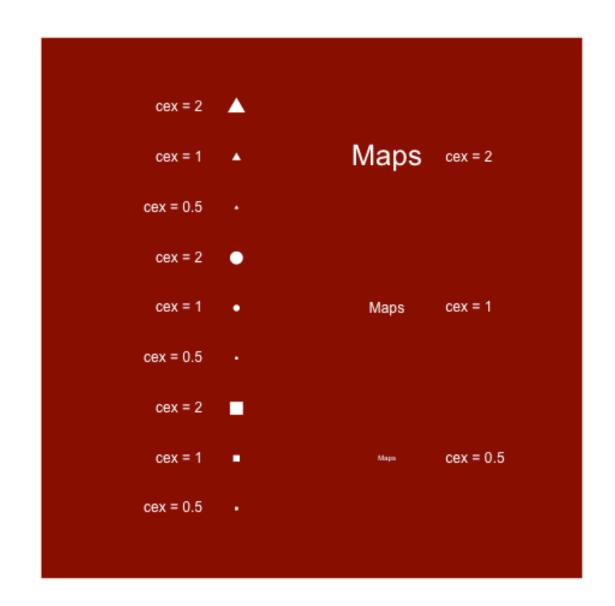
- ✓ Devices available :
 - Quartz
 - X11
 - pdf, jpeg, svg, ...
 - in add-on package:
 - <u>rgl package</u> (OpenGL website)
 - Internet browsers googleVis

✓ bmp(), jpeg(), png(), tiff()

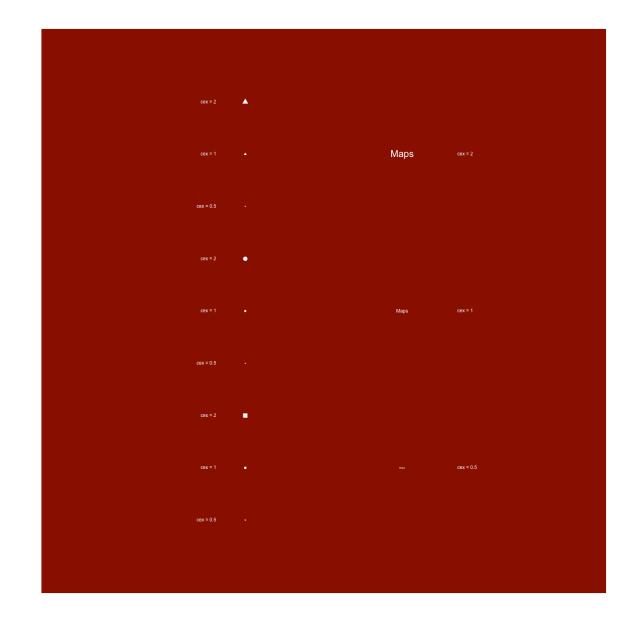
```
?jpeg
```

✓ use them:

```
png(filename, width=480, height=480)
...
dev.off()
```

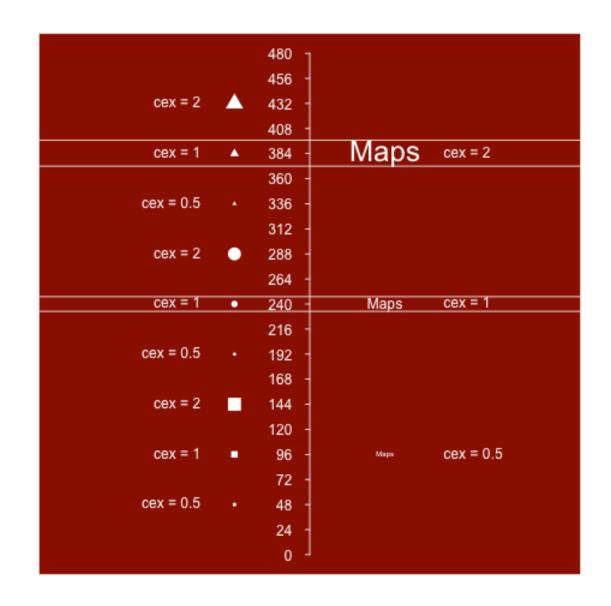


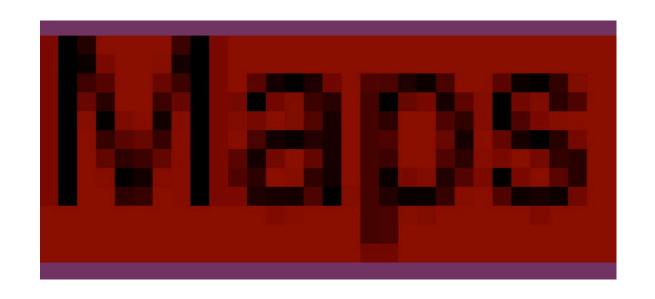
```
png(filename, width=1440, height=1440)
...
dev.off()
```



- ✓ pixel (px) = small colored square;
- √ width=480 + height=480 = grid of 480x480px;
- ✓ point (pt) = unit of length (measures the height of a font);
- ✓ 1pt = 1/72 inch;
- ✓ pointsize of plotted text = how many points your font will use (size of the text);
- ✓ resolution res (in px per inch, ppi) links pixel and size;
- ✓ res determines how many pixels = 1pt;
- ✓ if res=72 then one point will equal exactly one pixel.
- \checkmark res=72 + width=480 + height=480 = 6.667x6.667in => 16.9*16.9cm
- √ res=300 + width=480 + height=480 = 1.6x1.6in => 4.06cmx4.06cm
- √ font 12 points => 0.42cm

```
jpeg(filename, res=72,
  pointsize=12, width=480, height=480)
...
dev.off()
```



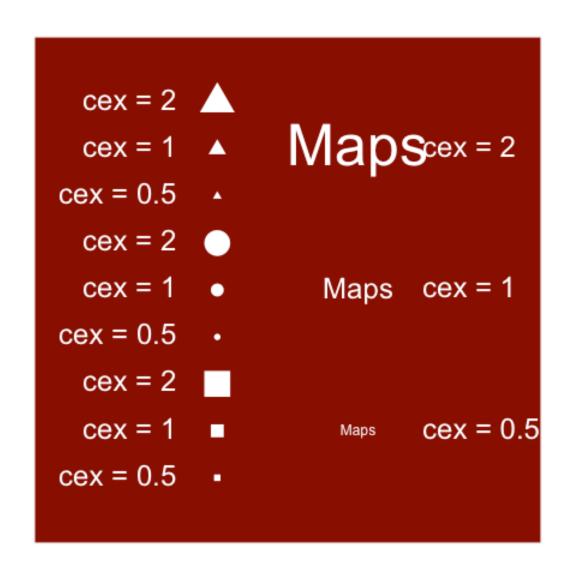




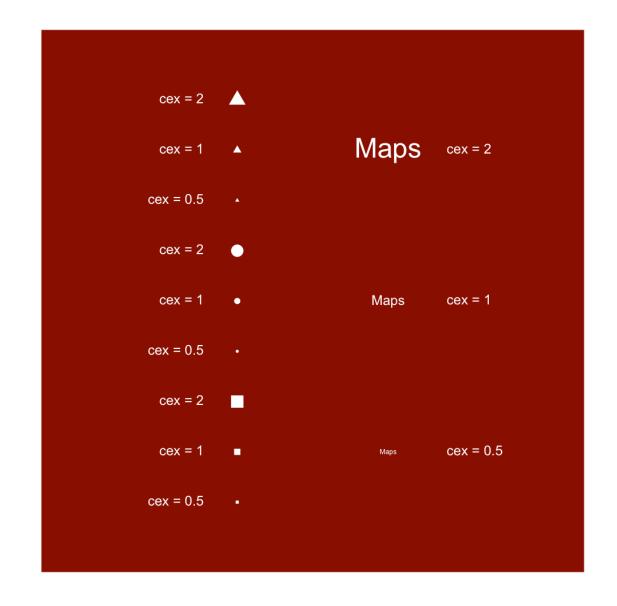
- ✓ pointsize=12
- √ res=72
- √ cex=1

- ✓ pointsize=12
- √ res=72
- ✓ cex=2

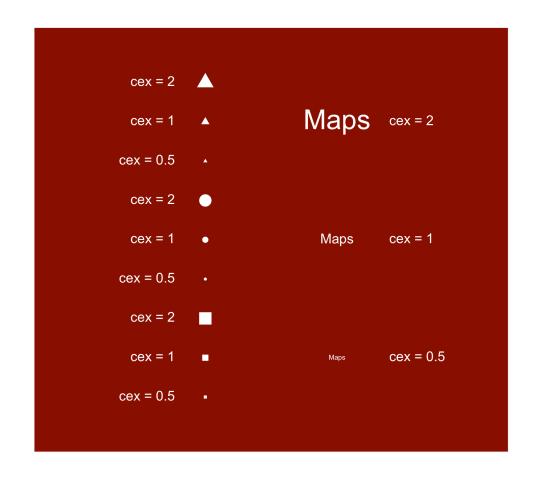
```
png(filename, res=144)
...
dev.off()
```



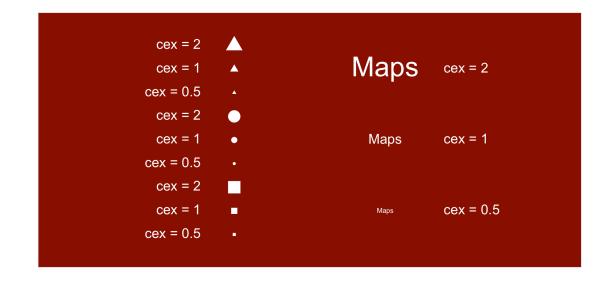
```
png(filename, res=144,
  height=7, width=7, unit="in")
...
dev.off()
```

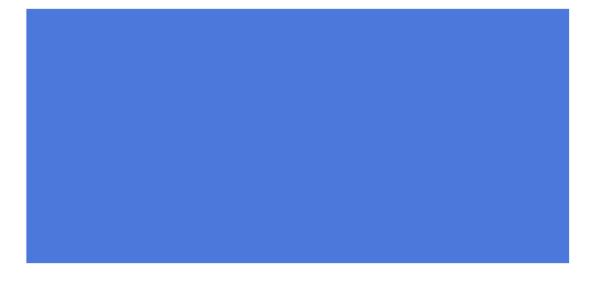


```
png(filename, res=288,
   height=7, width=7, unit="in")
...
dev.off()
```

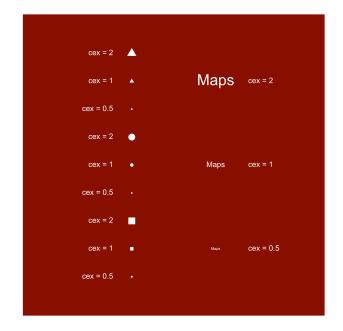


```
png(filename, res=288,
  height=7, width=7, unit="in")
...
dev.off()
```





```
png(filename, res=288,
  height=2*7, width=7, unit="in")
...
dev.off()
```





Exporting figures as vector files

```
/ pdf();

/ Cairo :

- cairo_pdf()

- cairo_ps()

- svg()
```

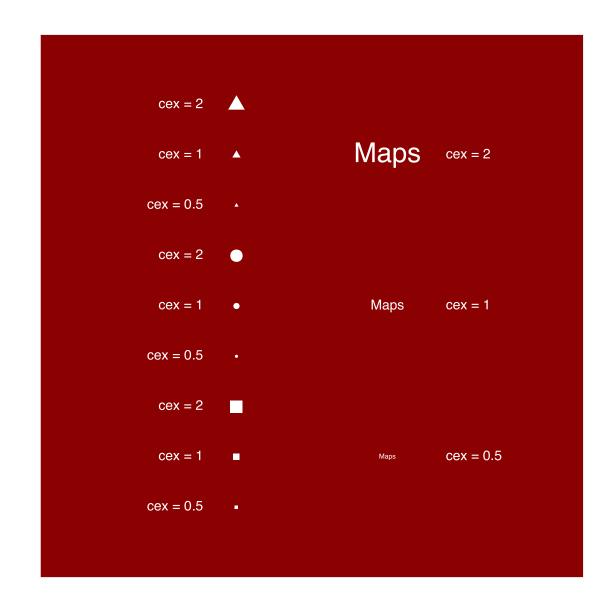
Exporting figures as vector files

```
pdf(fil, pointsize=12,
  height=2*7, width=7)
...
dev.off()
```



Exporting figures as vector files

```
svg(filename, pointsize=12,
  height=2*7, width=7)
...
dev.off()
```



Resources

- ✓ CRAN task view fro graphs
- ✓ more packages indexed
- ✓ ggplot2 website
- ✓ Color: An interesting blog post

Three challenges

- 1. "refaire une figure du samedi"
- 2. a fig quite complicated...
- 3. Code your own boxplot function.