

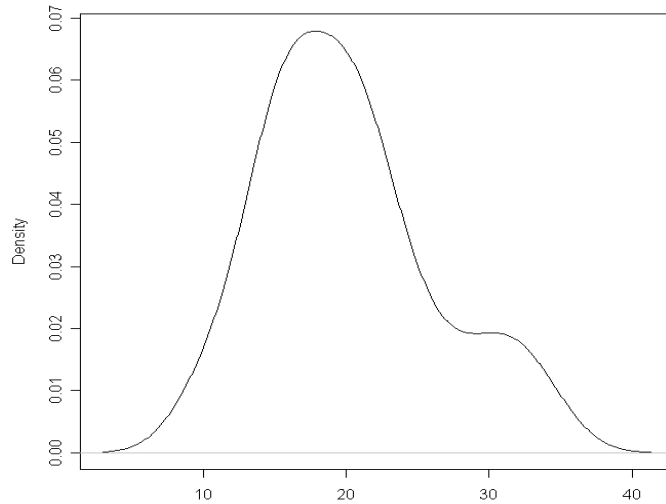


R course for beginners

Session 3: Plots and graphics

Plots and Graphics: from simple ones...

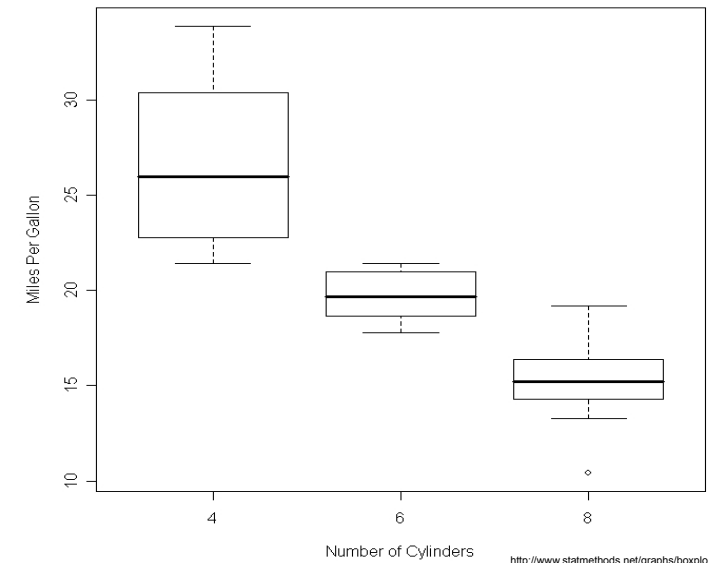
density.default(x = mtcars\$mpg)



<http://www.statmethods.net/graphs/density.html>

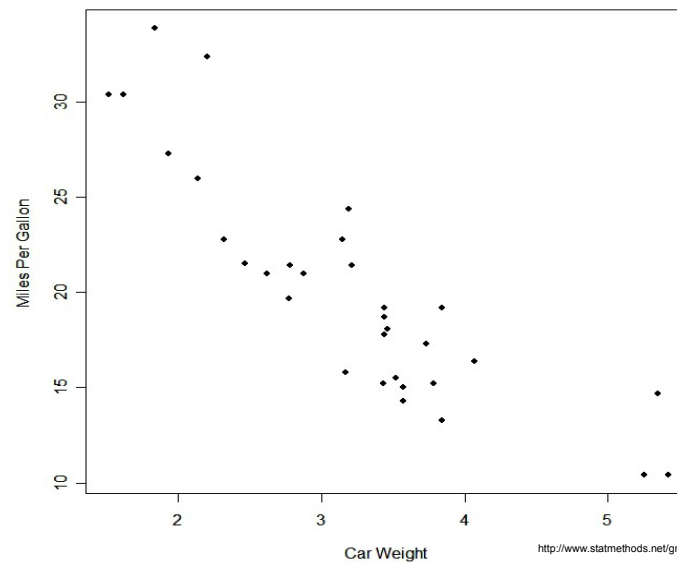
N = 32 Bandwidth = 2.477

Car Milage Data



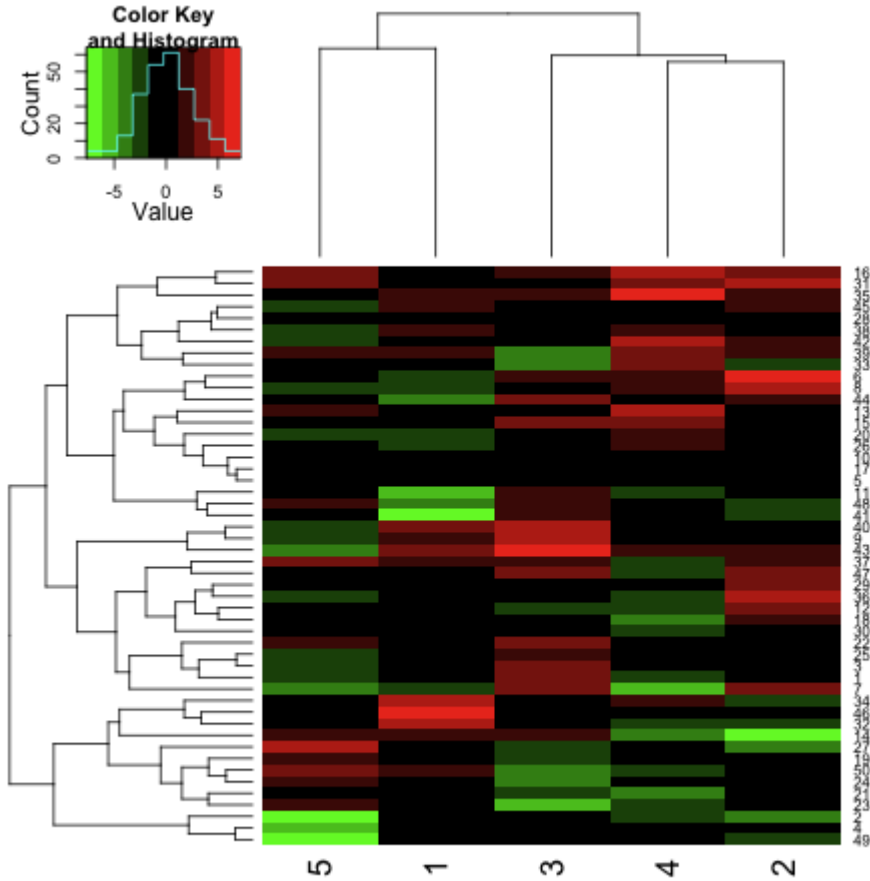
<http://www.statmethods.net/graphs/boxplot.html>

Scatterplot Example

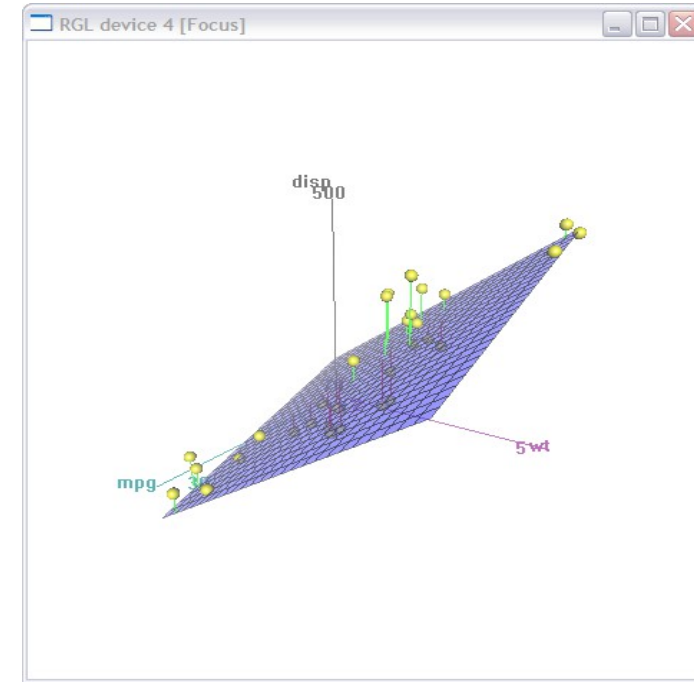


<http://www.statmethods.net/graphs/scatterplot.html>

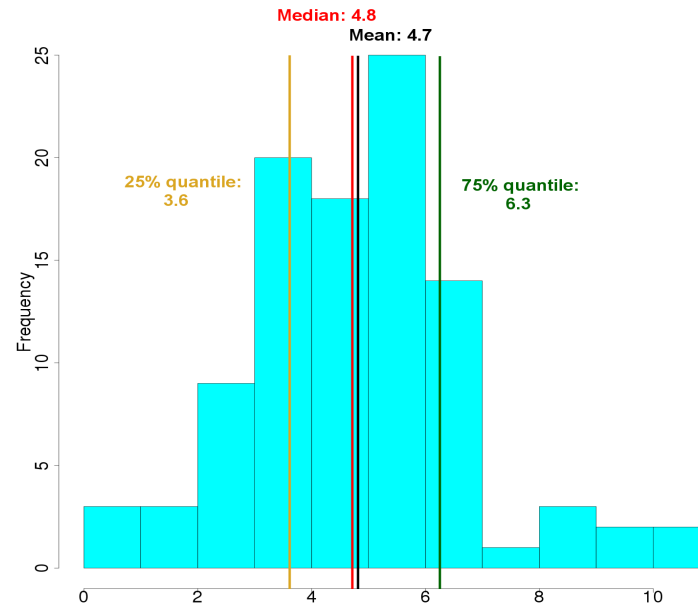
Plots and Graphics: ... to shiny graphics



http://3.bp.blogspot.com/_6rTMouB4cJY/S75cygN8YxI/AAAAAAAAAEw/8f6En3G0YyA/s1600/exampleHeatmap.png



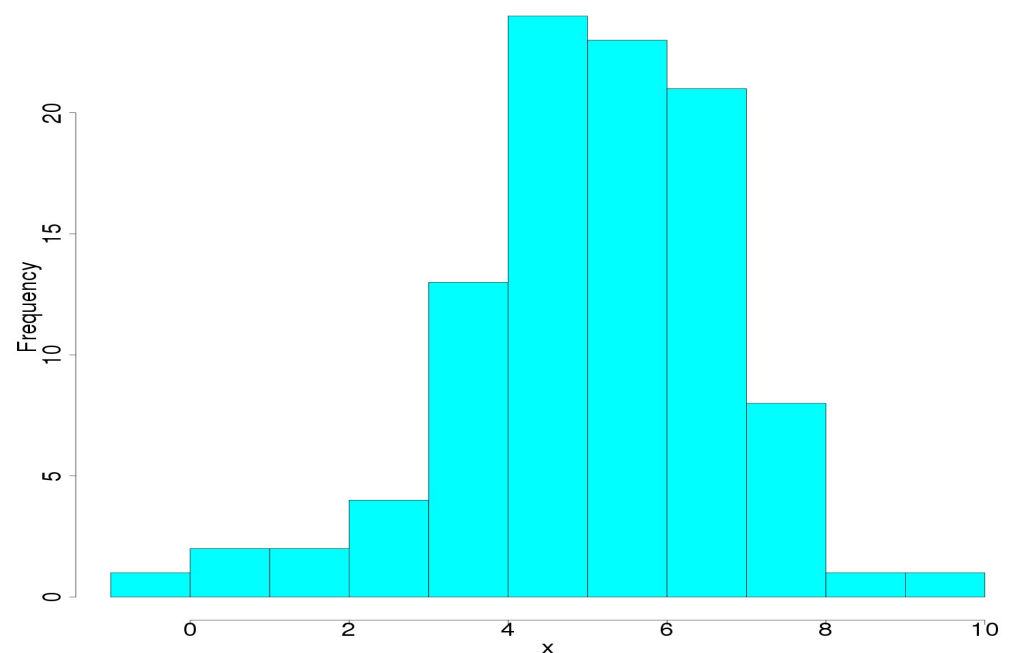
<http://www.statmethods.net/graphs/images/scatter3d.png>



Plots and Graphics: Recap

a) Histogram

- Graphical representation of distribution of continuous data



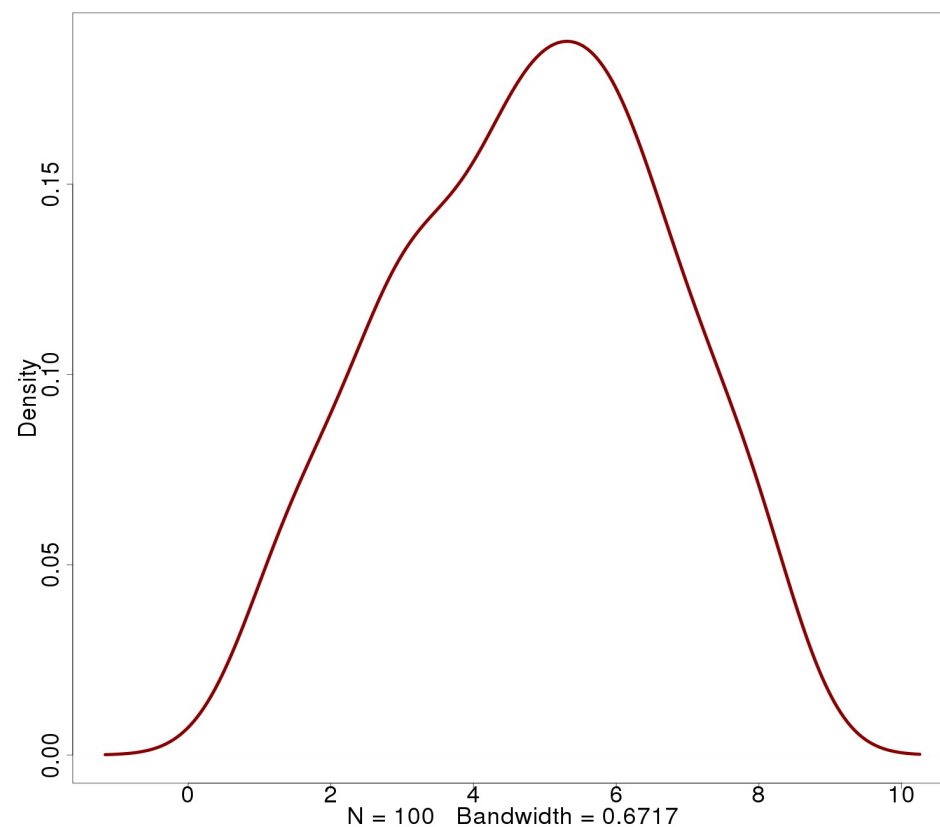
- Divides range of data in intervals, showing how many data points fall in each interval

Plots and Graphics: Recap

b) Density Plot

- Graphical representation of distribution of continuous data

- Smoothed line





Plots and Graphics: Recap

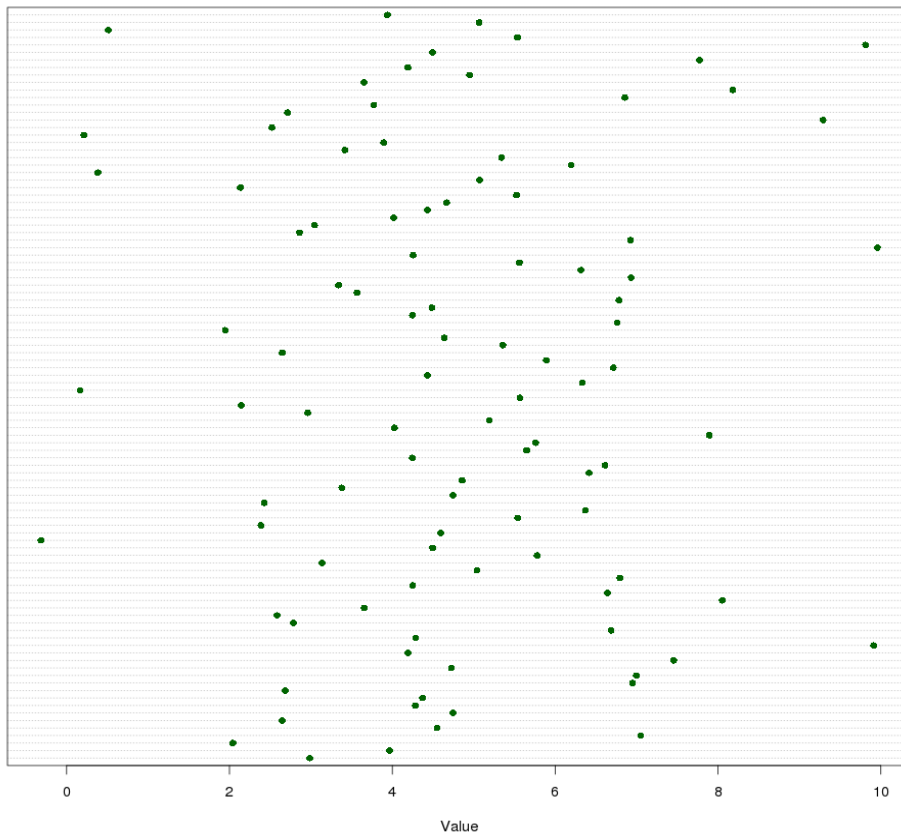
c) Dot Plot

- Plotting data points on fairly simple scale
- Two different version:
 - > Wilkinson: depict distribution
 - > Cleveland: depict quantitative values associated with categorical variables

Plots and Graphics: Recap

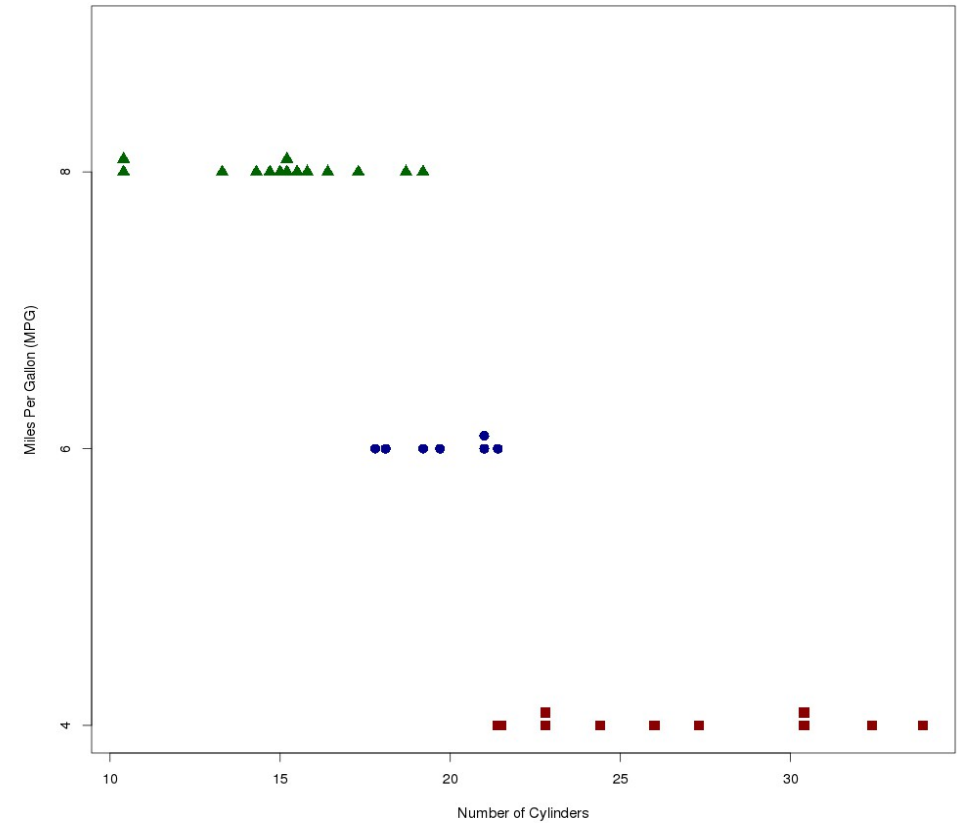
c) Dot Plot

- Plotting data points on fairly simple scale



es

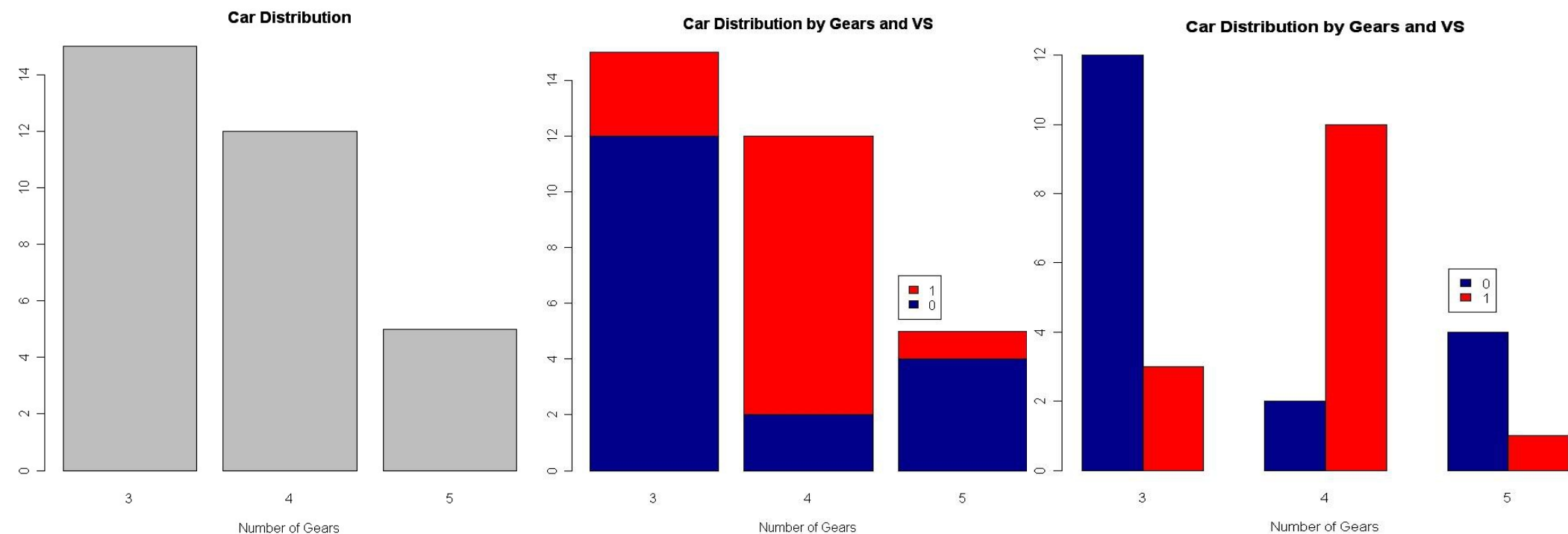
Miles Per Gallon vs. Number of Cylinders



Plots and Graphics: Recap

d) Bar Plot

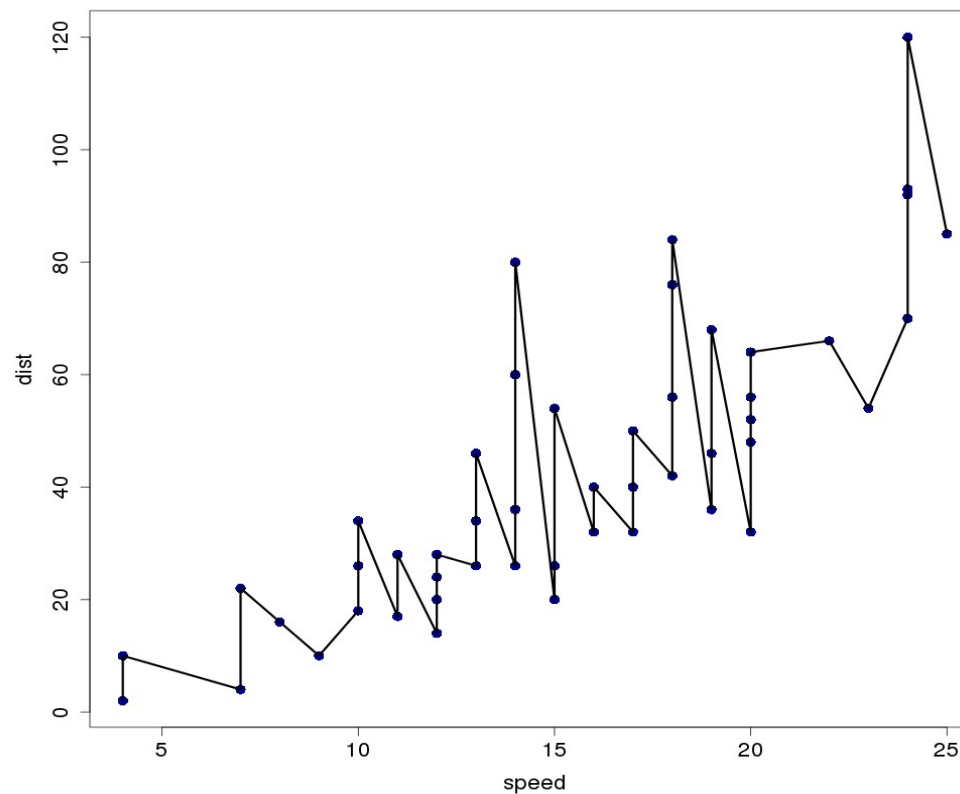
- Used for comparing categories
- One axis showing category, the other represents discrete value
- Can also be used to represent bars clustered in groups (grouped bar graphs) or divide bars into subgroups (stacked bar graphs)



Plots and Graphics: Recap

e) Line chart

- Information as series of data points, connected by line

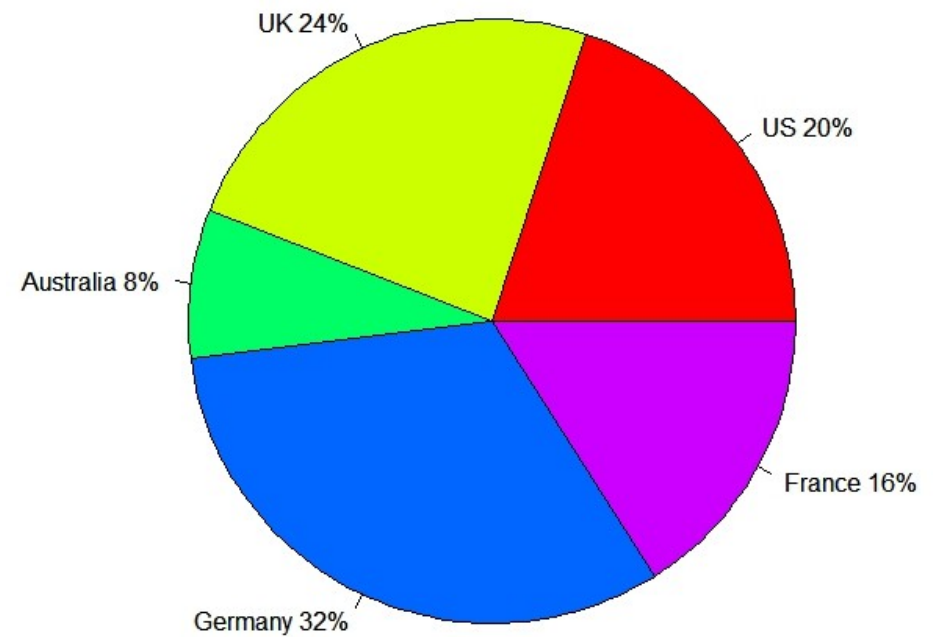


Plots and Graphics: Recap

f) Pie charts

- Circular chart representing proportion
- Arc length of sector \sim quantity

Pie Chart of Countries



<http://www.statmethods.net/graphs/pie.html>



Plots and Graphics: Recap

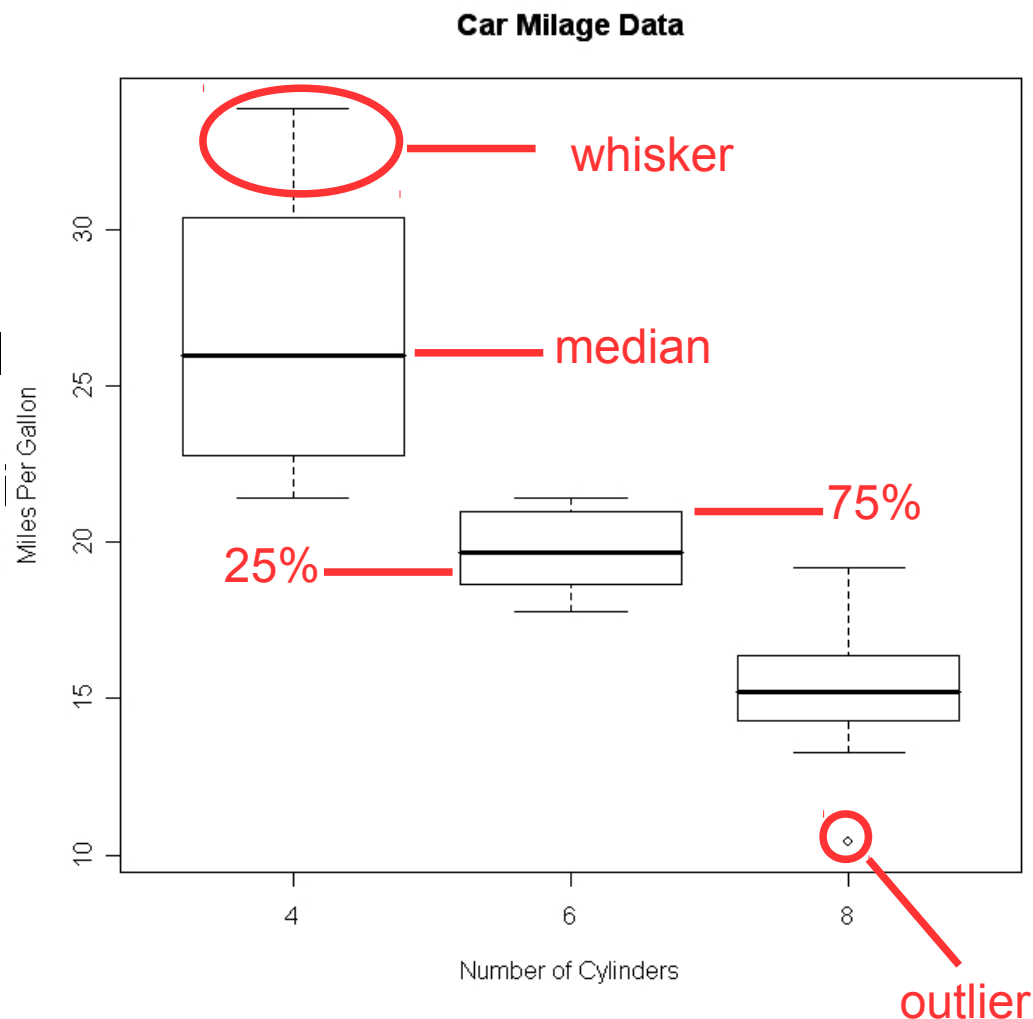
g) Boxplots

- Plot of quartiles
 - > upper line of box: 75%
 - > lower line of box: 25%
 - > thick line in middle: median
 - > lines extending from box vertically are called whiskers, indicating variability
 - > single plots outside the box: outliers

Plots and Graphics: Recap

g) Boxplots

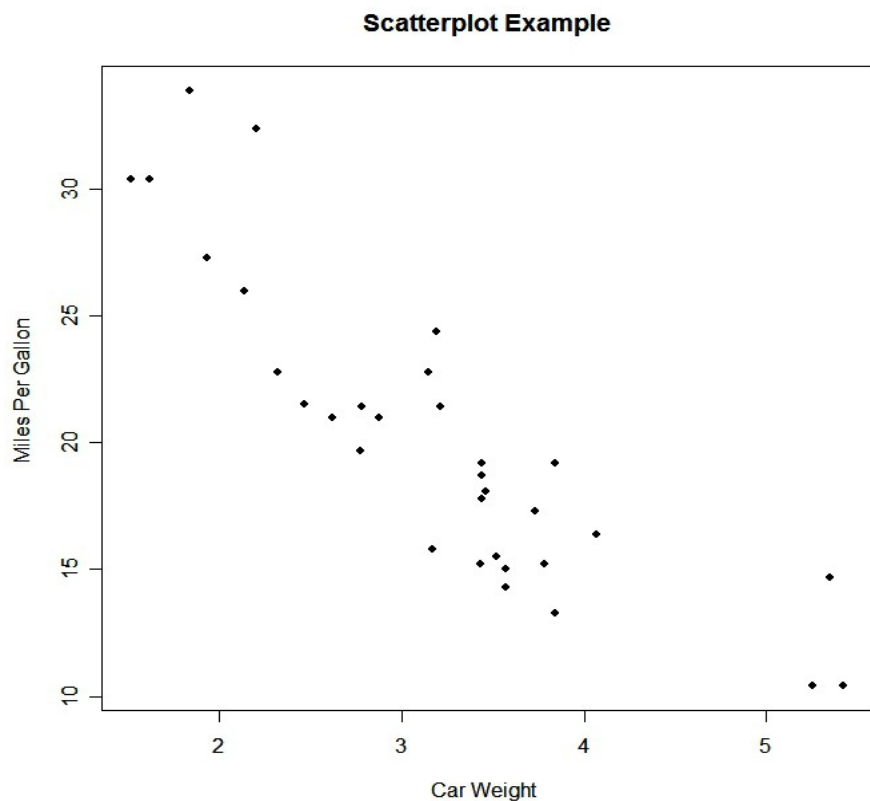
- Plot of quartiles
 - > upper line of box: 75%
 - > lower line of box: 25%
 - > thick line in middle: median
 - > lines extending from box vertical variability
 - > single plots outside the box: outlier



Plots and Graphics: Recap

h) Scatter plots

- Represents data as point collection
- Often used to show similarity of two data sets



<http://www.statmethods.net/graphs/scatterplot.html>



Plots and Graphics: Recap

- A good graphic supports arguments and displays data intrinsic structures
- Resemblance, order, proportion, expressiveness, effectiveness
- Figures should be immediately comprehensive
- Showing more isn't always the best



Plots and Graphics: Basic command plot()

Command: plot()

---> check **?plot**

---> Lots of parameters can be set, e.g. which type of plot you'd like to have

---> Plot() can be used to plot x versus y as well as plotting a whole data set



Plots and Graphics: Basic command plot()

Command: plot()

---> check **?plot**

---> Lots of parameters can be set, e.g. which type of plot you'd like to have

---> Plot() can be used to plot x versus y as well as plotting a whole data set

Example:

load data set mtcars, plot mpg versus wt

plot(mtcars\$mpg, mtcars\$wt)



Plots and Graphics: Basic command plot()

Command: plot()

---> check **?plot**

---> Lots of parameters can be set, e.g. which type of plot you'd like to have

---> Plot() can be used to plot x versus y as well as plotting a whole data set

Example:

load data set mtcars, plot mpg vs wt

plot(mtcars\$mpg, mtcars\$wt)

In R, a „\$“ will allow you to access an entry of a list

In this example, mpg and wt are the columns you'd like to access in the data set mtcars



Plots and Graphics: Basic command plot()

Command: plot()

---> check **?plot**

---> Lots of parameters can be set, e.g. which type of plot you'd like to have

Example:

Now, you'd like to add a regression line to show the correlation of mpg and wt

Command: lm()

Lm stands for „linear model“ and needs as input a formula

We'd like to see the correlation of mpg and wt, so our formula is

`wt~mpg`



Plots and Graphics: Basic command plot()

Command: plot()

---> check **?plot**

---> Lots of parameters can be set, e.g. which type of plot you'd like to have

Example:

Now, you'd like to add a regression line to show the correlation of mpg and wt

Command: lm()

Lm stands for „linear model“ and needs as input a formula

We'd like to see the correlation of mpg and wt, so our formula is

„~“ stands for „depends on“

wt~mpg



Plots and Graphics: Basic command plot()

Command: plot()

---> check **?plot**

---> Lots of parameters can be set, e.g. which type of plot you'd like to have

Example:

Now, you'd like to add a regression line to show the correlation of mpg and wt

Command: abline()

Adds a straight line in the current plot

So, our command for plotting a regression line is:



Plots and Graphics: Basic command plot()

Command: plot()

---> check **?plot**

---> Lots of parameters can be set, e.g. which type of plot you'd like to have

Example:

Now, you'd like to add a regression line to show the correlation of mpg and wt

Command: abline()

Adds a straight line in the current plot

abline(lm(mtcars\$wt~mtcars\$mpg))



Plots and Graphics: Basic command plot()

Command: plot()

---> check **?plot**

---> Lots of parameters can be set, e.g. which type of plot you'd like to have

Example:

Add a title to your plot

Command: title()

Remember: a string must be given in quotation marks

e.g.

title(„Example session 3 – title added“)



Plots and Graphics: Basic command plot()

There are many graphical parameters you can set...

... line width

... line type

... dot type

... colors

... titles

... adjusting the text size...

... adjusting the plotting areas...



Plots and Graphics: Basic command plot()

There are many graphical parameters you can set...

... line width

Command: **lwd=x**

X stands for a digit

Example: adjust the line width of abline to 5

abline(lm(mtcars\$wt~mtcars\$mpg), lwd=5)

Plots and Graphics: Basic command plot()

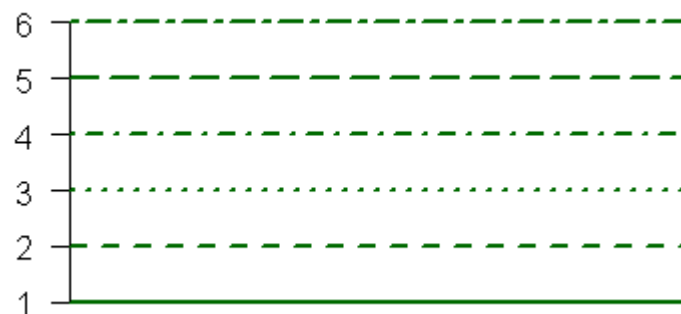
There are many graphical parameters you can set...

... line type

Command: **lty=x**

X stands for a digit

Example: get abline to be of type 3



<http://www.statmethods.net/advgraphs/parameters.html>

Plots and Graphics: Basic command plot()

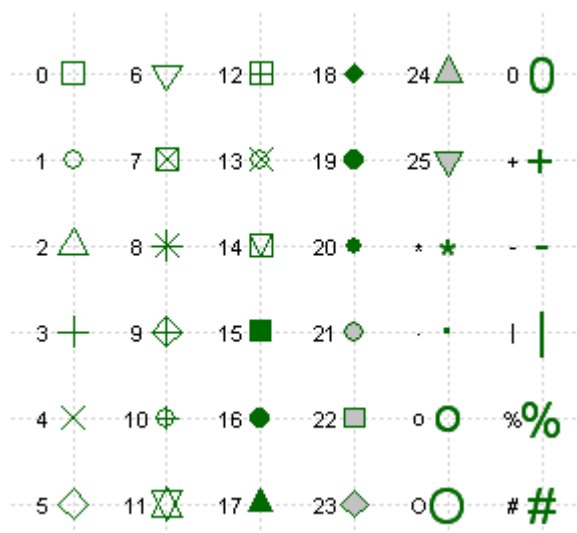
There are many graphical parameters you can set...

... dot type

Command: **pch=x**

X stands for a digit

Example: get dots to be of type 3



<http://www.statmethods.net/advgraphs/parameters.html>



Plots and Graphics: Basic command plot()

There are many graphical parameters you can set...

... colors

Command: **col="x"**

X stands for a name

Example: set regression line as darkgreen

abline(lm(mtcars\$wt~mtcars\$mpg), col="darkgreen")



Plots and Graphics: Basic command plot()

There are many graphical parameters you can set...

... colors

Command: **col="x"**

X stands for a name

There are many colors available – just search for them, e.g. searching the web for „R colors“



Plots and Graphics: Basic command plot()

There are many graphical parameters you can set...

... titles

Commands: **main="x"; xlab="x"; ylab="x"**

X stands for a name

Main will determine the main title, xlab the name of the x-axis and ylab the name of the y-axis

Example: choose some titles and give them to your plot



Plots and Graphics: Basic command plot()

There are many graphical parameters you can set...

... adjusting font size for titles

Command: **cex=x**

X is a digit

Adjustment of main title: `cex.main=x`

Adjustment of x/y title: `cex.lab=x`

Adjustment of axis annotation: `cex.axis=x`

Check **?cex**

Plots and Graphics: Basic command plot()

There are many graphical parameters you can set...

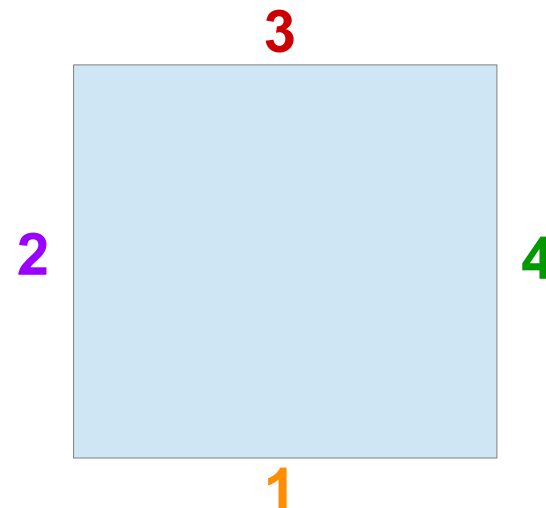
... adjusting margin of plotting area

Command: **par()**

Par stands for „parameter“

As example, the margins can be set using

`Par(mar=c(1,2,3,4))`



Plots and Graphics: Basic command plot()

There are many graphical parameters you can set...

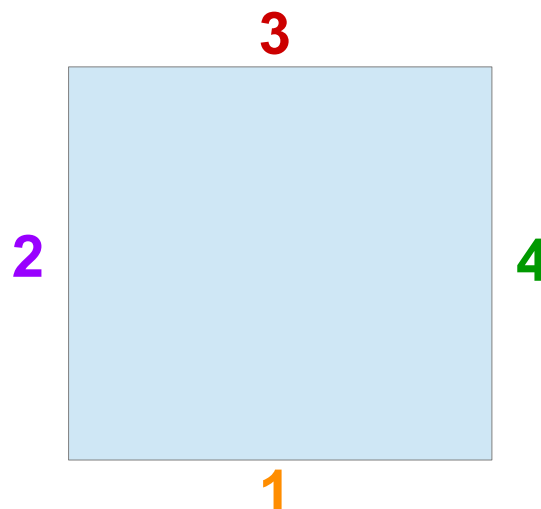
... adjusting outer plotting area

Command: **par()**

Par stands for „parameter“

As example, the outer plotting area can be set using

`Par(oma=c(1,2,3,4))`





Plots and Graphics: Basic command plot()

There are many graphical parameters you can set...

... we just had some examples!

More parameters could be font families, multiple plots...



Plots and Graphics: Histogram

Command: hist()

Creates a simply histogram

Example:

hist(mtcars\$mpg)

Task: try to create a histogram for mtcars, cyl, with blue bars



Plots and Graphics: Histogram

Command: hist()

Creates a simply histogram

Example:

hist(mtcars\$mpg)

Task: try to create a histogram for mtcars, cyl, with blue bars

hist(mtcars\$cyl, col="blue")



Plots and Graphics: Histogram

Command: hist()

Creates a simply histogram

Example:

hist(mtcars\$mpg)

Task: try to create a histogram for mtcars, cyl, with blue bars and a line denoting the mean

Ablines can be used to plot x versus y, plot a horizontal line or a vertical line



Plots and Graphics: Histogram

Command: hist()

Creates a simply histogram

Example:

hist(mtcars\$mpg)

Task: try to create a histogram for mtcars, cyl, with blue bars and a line denoting the mean

Ablines can be used to plot x versus y, plot a horizontal line or a vertical line

---> check **?abline**



Plots and Graphics: Histogram

Command: hist()

Creates a simply histogram

Example:

hist(mtcars\$mpg)

Task: try to create a histogram for mtcars, cyl, with blue bars and a line denoting the mean (size should be 5)

Ablines can be used to plot x versus y, plot a horizontal line or a vertical line

---> **abline(v=mean(mtcars\$cyl), lwd=5)**



Plots and Graphics: Histogram

Command: `hist()`

Creates a simply histogram

Example:

`hist(mtcars$mpg)`

Task: try to create a histogram for `mtcars`, `cyl`, with blue bars and a line denoting the mean (size should be 5), write „mean“ to this line

Command: `text(x,y,““)`

X, y stand for x and y coordinates where text should be written
„“ is a place holder for your string



Plots and Graphics: Histogram

Command: hist()

Creates a simply histogram

Example:

hist(mtcars\$mpg)

Task: try to create a histogram for mtcars, cyl, with blue bars and a line denoting the mean (size should be 5), write „mean“ to this line

Command: text(x,y,““)

---> text(6,14, „mean value“)



Plots and Graphics: Histogram

Command: hist()

Creates a simply histogram

Example:

hist(mtcars\$mpg)

Task: try to create a histogram for mtcars, cyl, with blue bars and a line denoting the mean (size should be 5), write „mean“ to this line

All commands will then be:

```
hist(mtcars$cyl, col="blue")  
abline(v=mean(mtcars$cyl), lwd=5)  
text(6,14,"mean")
```



Plots and Graphics: Density plot

Command: `density(x)`

X stands for the data you're supplying

Density() does not create a plot, so you'll have to plot the result of density

Example: create a density plot of mtcars, mpg



Plots and Graphics: Density plot

Command: density(x)

X stands for the data you're supplying

Density() does not create a plot, so you'll have to plot the result of density

Example: create a density plot of mtcars, mpg

plot(density(mtcars\$mpg))

OR:

d<-density(mtcars\$mpg)

plot(d)



Plots and Graphics: Dot plot

Command: dotchart()

Will draw a Cleveland dot plot

Example:

```
dotchart(mtcars$mpg, labels=rownames(mtcars), main="Miles per  
gallon")
```



Plots and Graphics: Bar plot

Command: `barplot(x)`

X stands for the data you're supplying

Example: create a bar plot of mtcars, distribution of gears

`barplot(mtcars$gear)`



Plots and Graphics: Bar plot

Command: `barplot(x)`

X stands for the data you're supplying

Example: create a bar plot of `mtcars`, distribution of gears

You'd like to have the counts plotted, not every data point!

Command: `table(x)`

```
table                                package:base                                R Documentation
Cross Tabulation and Table Creation
Description:
' table ' uses the cross-classifying factors to build a contingency
table of the counts at each combination of factor levels.
```



Plots and Graphics: Bar plot

Command: `barplot(x)`

X stands for the data you're supplying

Example: create a bar plot of mtcars, distribution of gears

You'd like to have the counts plotted, not every data point!

Command: `table(x)`

... will give us the counts!

```
table                                package:base                                R Documentation
Cross Tabulation and Table Creation
Description:
' table ' uses the cross-classifying factors to build a contingency
table of the counts at each combination of factor levels.
```



Plots and Graphics: Bar plot

Command: `barplot(x)`

X stands for the data you're supplying

Example: create a bar plot of mtcars, distribution of gears

You'd like to have the counts plotted, not every data point!

Command: `table(x)`

---> look at **`table(mtcars$gear)`**

---> **`barplot(table(mtcars$gear))`**



Plots and Graphics: Bar plot

Command: `barplot(x)`

X stands for the data you're supplying

Example: create a bar plot of `mtcars`, distribution of gears, add title and axis
Description

Remember `xlab`, `ylab`, `main`



Plots and Graphics: Bar plot

Command: `barplot(x)`

X stands for the data you're supplying

Example: create a bar plot of mtcars, distribution of gears, add title and axis
Description

`barplot(table(mtcars$gear), main="Distribution of gears", xlab="Number of gears", ylab="Sum of cars")`



Plots and Graphics: Line chart

Command: `plot(x,y)`

The basic command for plotting x versus y

Command: `lines(x,y)`

X, y are the coordinates to be joined by line

Example: look at data set „cars“, which is given as demo data set in R



Plots and Graphics: Line chart

Command: `plot(x,y)`

The basic command for plotting x versus y

Command: `lines(x,y)`

X, y are the coordinates to be joined by line

Example: look at data set „cars“, which is given as demo data set in R

Plot cars

---> **`plot(cars)`**



Plots and Graphics: Line chart

Command: `plot(x,y)`

The basic command for plotting x versus y

Command: `lines(x,y)`

X, y are the coordinates to be joined by line

Example: look at data set „cars“, which is given as demo data set in R

Plot cars, add a line connecting the points

---> **`plot(cars)`**

`lines(cars)`



Plots and Graphics: Pie Chart

Command: pie()

Creates a simple pie chart.

Example: determine slices, give them names and plot the pie chart with colors blue, red, green and yellow

e.g.

```
slices <- c(10,20,30,40)  
labels <- c(„First“, „Second“, „Third“, „Fourth“)  
pie(slices, labels=labels, main=“Pie chart test“, col=c(„blue“, „red“, „green“, „yellow“))
```



Plots and Graphics: Pie Chart

Command: pie()

Creates a simple pie chart.

Example: determine slices, give them names and plot the pie chart with colors blue, red, green and yellow, add a legend

Command: legend(position, label,)

legend("topright", labels, fill=c("blue", "green", "red", "yellow"))



Plots and Graphics: Box plot

Command: `boxplot(formula, x)`

X stands for the data you're supplying, formula for the relationship you'd like to be plotted

Example: create a box plot of mtcars miles per gallon vs. cylinders, add title and axis description; boxes should be green

Remember: ~ stands for „depends on“
Setting a color can be done with „col“



Plots and Graphics: Box plot

Command: `boxplot(formula, x)`

X stands for the data you're supplying, formula for the relationship you'd like to be plotted

Example: create a box plot of mtcars miles per gallon vs. cylinders, add title and axis description; boxes should be green

```
boxplot(mtcars$mpg~mtcars$cyl,data=mtcars, main="Car Milage Data",  
xlab="Number of Cylinders", ylab="Miles Per Gallon", col="green")
```



Plots and Graphics: Scatter plot

Command: stripchart()

Example: create a box plot of mtcars miles per gallon vs. cylinders, add title and axis description; use different point types and colors

e.g.

```
stripchart(mtcars$mpg~mtcars$cyl, method="stack", pch=c(15,16,17),  
col=c("darkred","darkblue","darkgreen"), xlab="Number of Cylinders",  
ylab="Miles Per Gallon (MPG)", main="Miles Per Gallon vs. Number of  
Cylinders")
```



Plots and Graphics: Advanced plotting

For many plots, special libraries exist

With those libraries, very fancy figures can be created

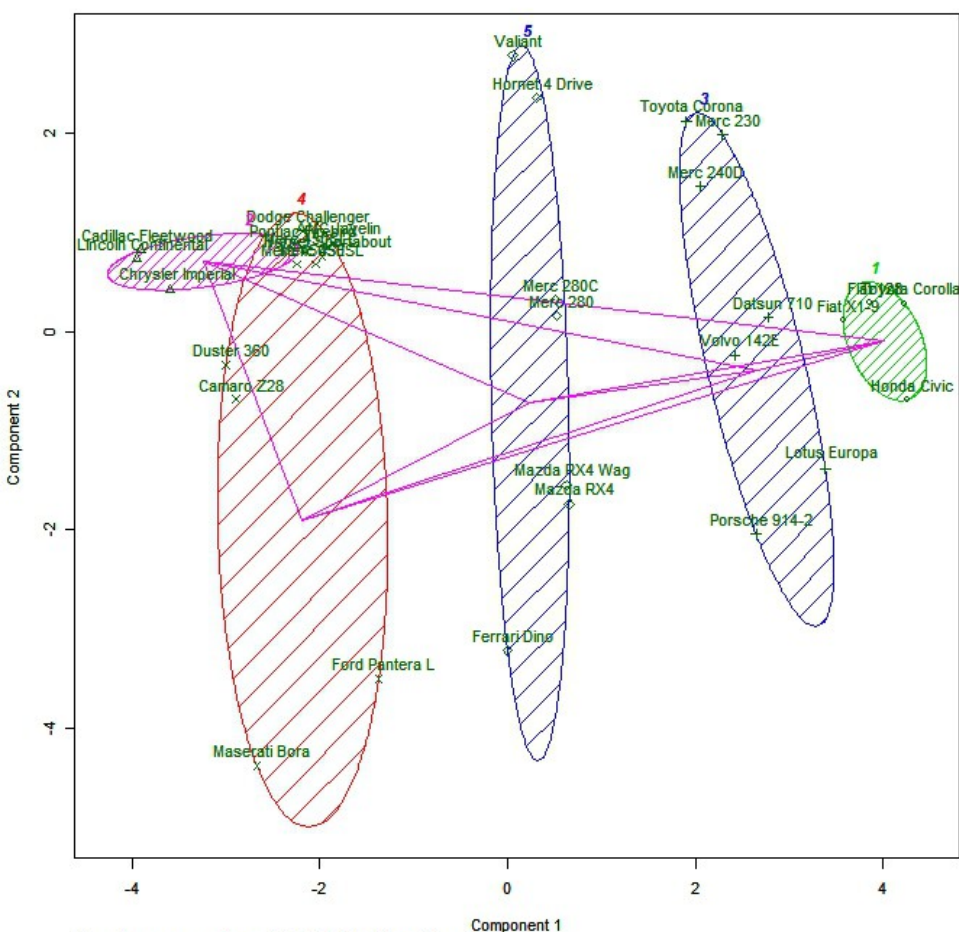
Plots and Graphics: Advanced plotting

For many plots, special libraries exist

With those libraries, very fancy figures can be created

<http://www.statmethods.net/advstats/images/clusters.jpg>

CLUSPLOT of MTCARS Data



http://studio-pubs-static.s3.amazonaws.com/2107_4eb1adc1e4d44b93b6fde7eb801519fe.html





Plots and Graphics: Advanced plotting

Example: correlation plot using corrplot package

- 1) Install package
- 2) load package
- 3) calculate correlation (example again: mtcars)
- 4) plot in various ways



Plots and Graphics: Saving plots

Command: `png(filename, width=x, height=x, pointsize=x, ...)`

---> check `?png`

Your plot will be saved as png file

Example: `png(„cars_line_plot.png“, width=1500, height=1000)`
`plot(cars)`
`dev.off()`

Command: `dev.off()`

Terminates graphic functions

Plot can also be saved as pdf, bmp, jpeg, tiff...

Plots and Graphics: What we did not cover

Heatmaps (afternoon session)

3D plots

Venn diagrams

Networks

...

..... many more plots available!

