Making Data Speak

USING R SOFTWARE

1 November 2017

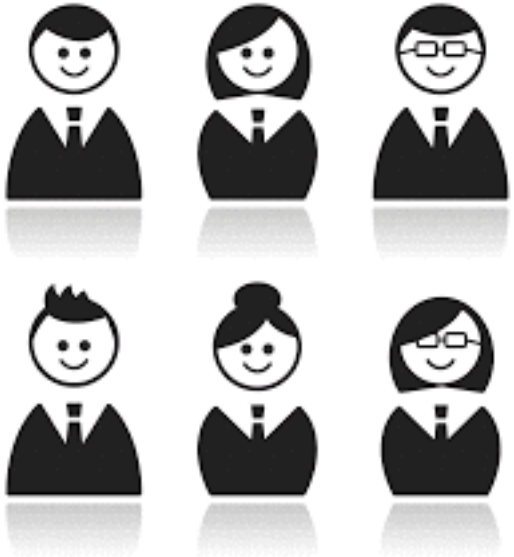
Dr. Norhaiza Ahmad
Department of Mathematical Sciences
Faculty of Science
Universiti Teknologi Malaysia

Welcome.

The title for this talk should really be Making Data Speak for FREE with R

http://science.utm.my/norhaiza/

Making Data Speak: Background



Data

- Researchers
- Non-Researchers

- Universities/Colleges
- Government Agencies
- Industries

Making Data Speak: Background

Academicians, Scientists, Engineers etc

Got a ton of data to analyze or include in your next paper?

Non-Researcher

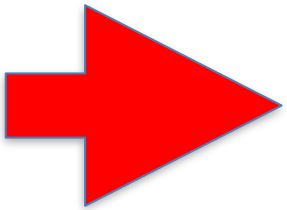
Librarians, Executives, Managers, Journalists etc

- Profile respondent/customers
- Use data for Decision Making
- Complement your words with visuals to tell your stories.

Making Data Speak: background

Express information contained in the data

Make Data Speak



Data

Big Data vs Small Data

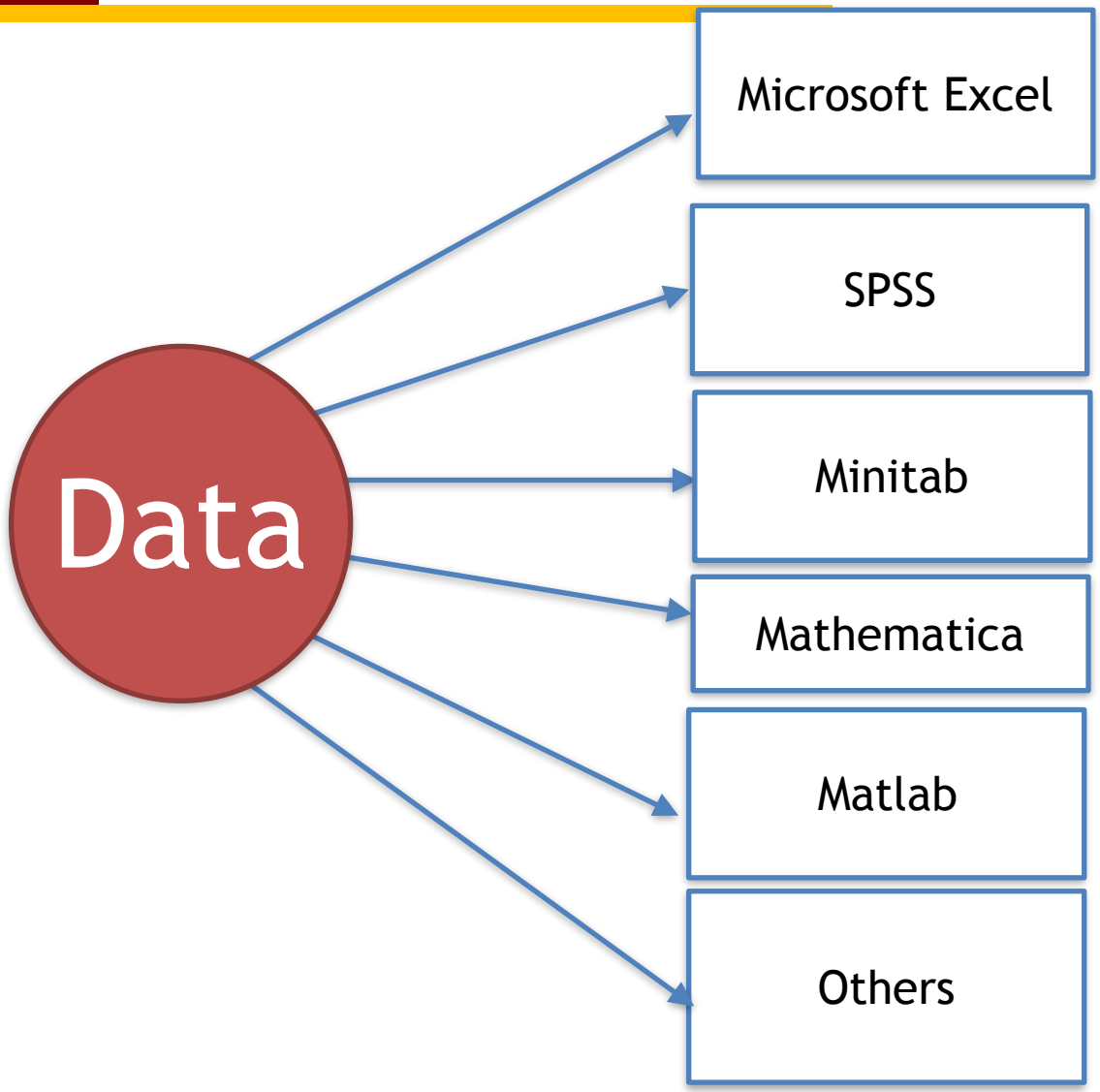
Decision Making

Knowledge Discovery

STATISTICAL

via DATA ANALYSIS

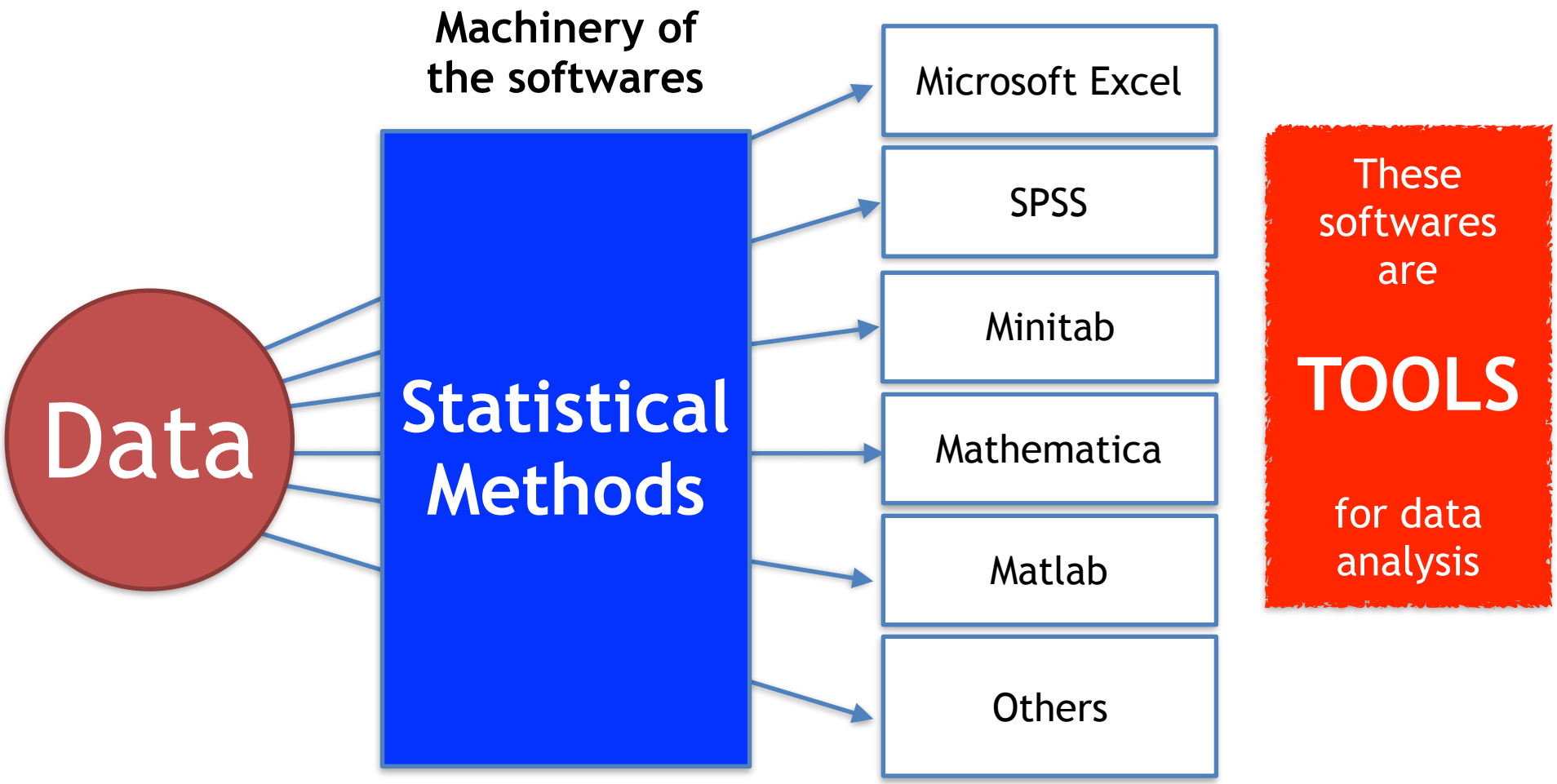
Making Data Speak: Softwares for Data Analysis



Common softwares used for data analysis

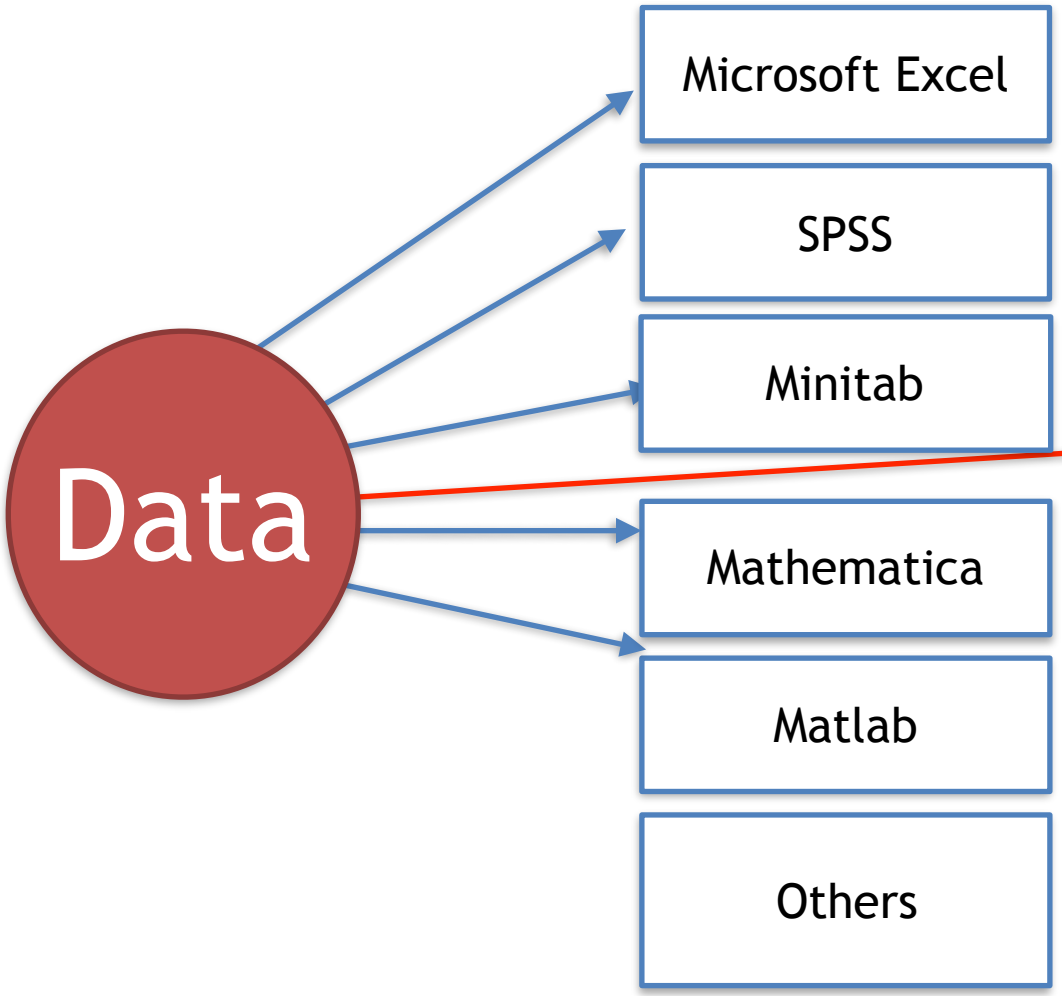
Useful & Relevant in its own right

Making Data Speak: Softwares for Data Analysis



AVOID ABUSE OF THESE SOFTWARE TOOLS BY UNDERSTANDING THE STATISTICAL METHODS BEHIND IT

AIM OF THIS TALK:



- Introduction to an alternative software for data analysis

- Focussing on data exploration with visualisation

Outline of This Talk

1. **Background**
2. **About R**
   - What is R software?
   - Why you should learn R
   - Anatomy of R
3. **Making Data Speak with R**
   - Types of data analysis
   - Visualization
     - Exploratory vs Confirmatory
   - Examples
4. Others: R for big data

   Tips for newbies to R

   AOB: bridging from other softwares

ABOUT R

What is R?

- A computer language, with orientation towards statistical applications

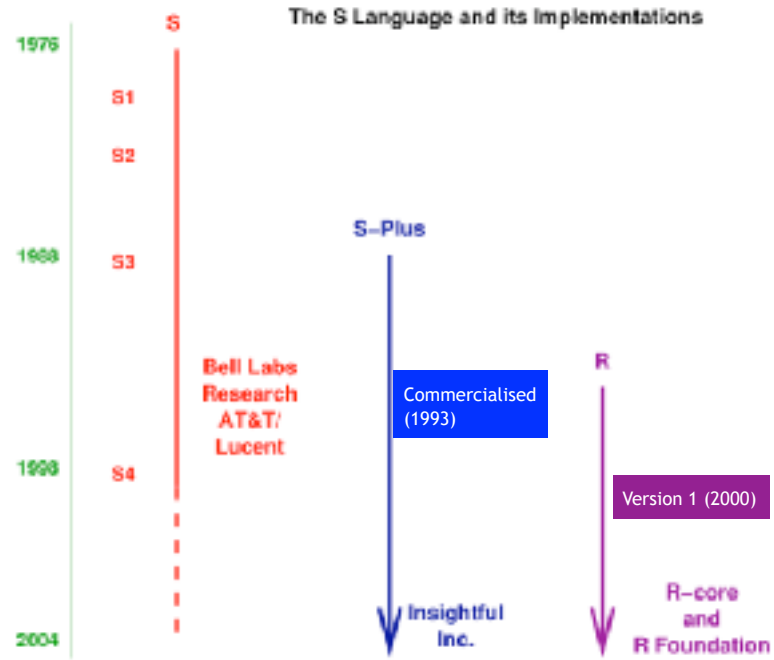
R is a dialect from a programming language called S.
S (language dev.1976)—> S Plus (commercial software license 1993)—>
R software (dev. 1991- R version 1.0.0 in 2000)

S was developed about 40 years ago for research in telecommunication industry

R was created by statisticians from University of Auckland

Ross Ihaka

Robert Gentleman



https://www.r-project.org/conferences/useR-2006/Slides/Chambers.pdf

What is R?

- Open-sourced software - non-commercial

  Principle: open exchange, publicly accessible Community-oriented software

- Origin in academics:

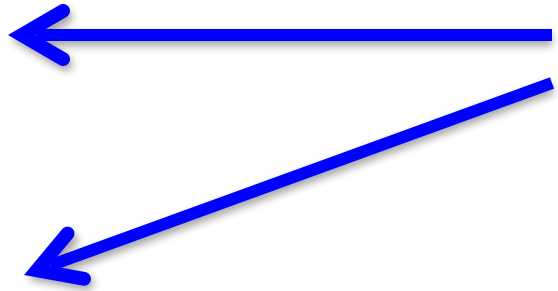
  solid foundation of core statistical and numerical algorithms and continues to grow to this end.

R Object Assignment

- R is object oriented program. #Each input/output can be assigned / stored to an object #case-sensitive

```
# use symbol '=' or '<-' for
  assignment
```

```
> x = 2
> x
> x <-2
> x
```



```
# CALL UP the r-
  object to display
  results
(if required)
```

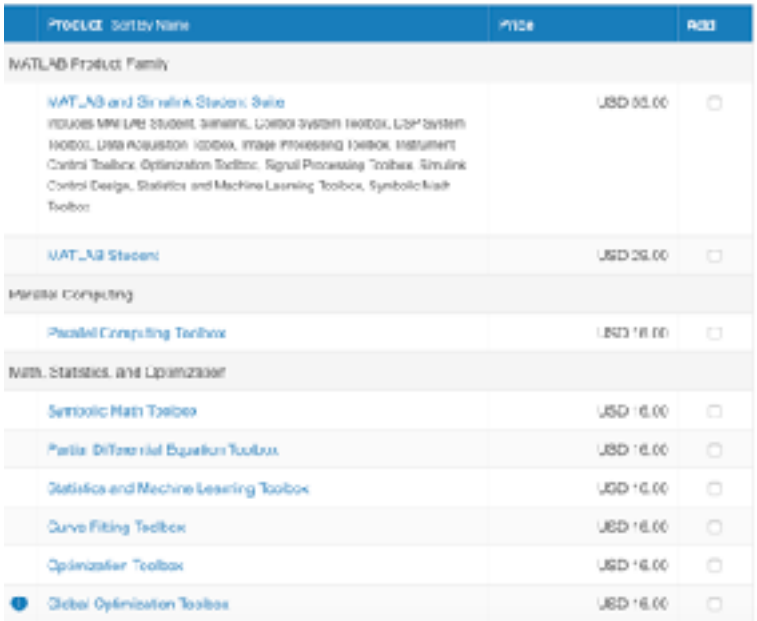
- Once an R object is assigned, it can be called upon at any time as long as it is saved. Here the number 2 is stored in an object called "x".
- In general, R objects are stored in an R workspace, also known as the global environment.

Comparative Cost of Softwares

- Microsoft Excel
- SPSS
- Minitab
- Mathematica
- Matlab
- Others
- R

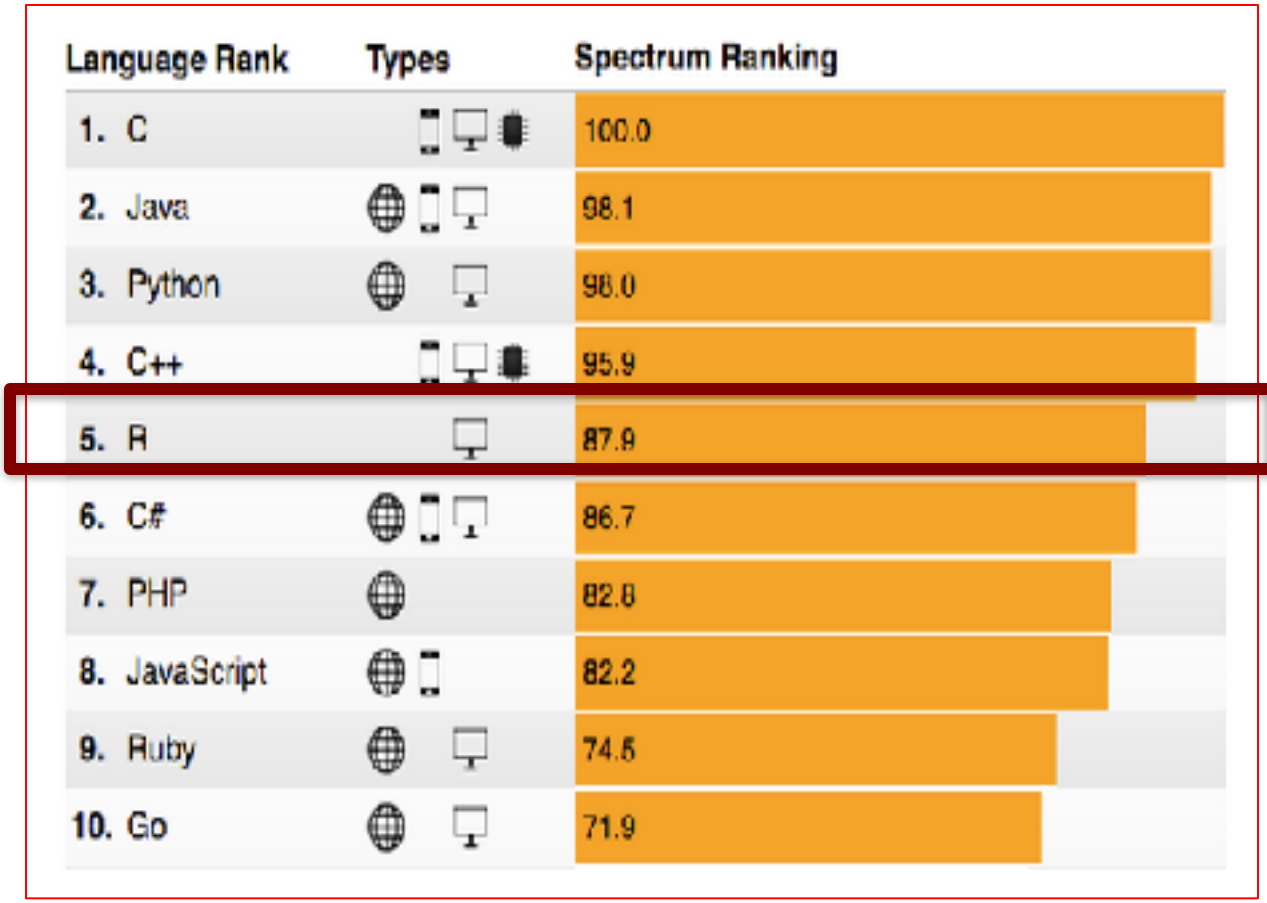
R= $0

IBM SPSS Statistics Base v24
SKU# D0EJ9LL
Excl. GST: RM 21,000.00
Incl. GST: **RM 22,260.00**



Why You Should Learn R?

IEEE's Spectrum 2016 Top Programming Languages



R tops rank for statistical/data-analysis programming languages

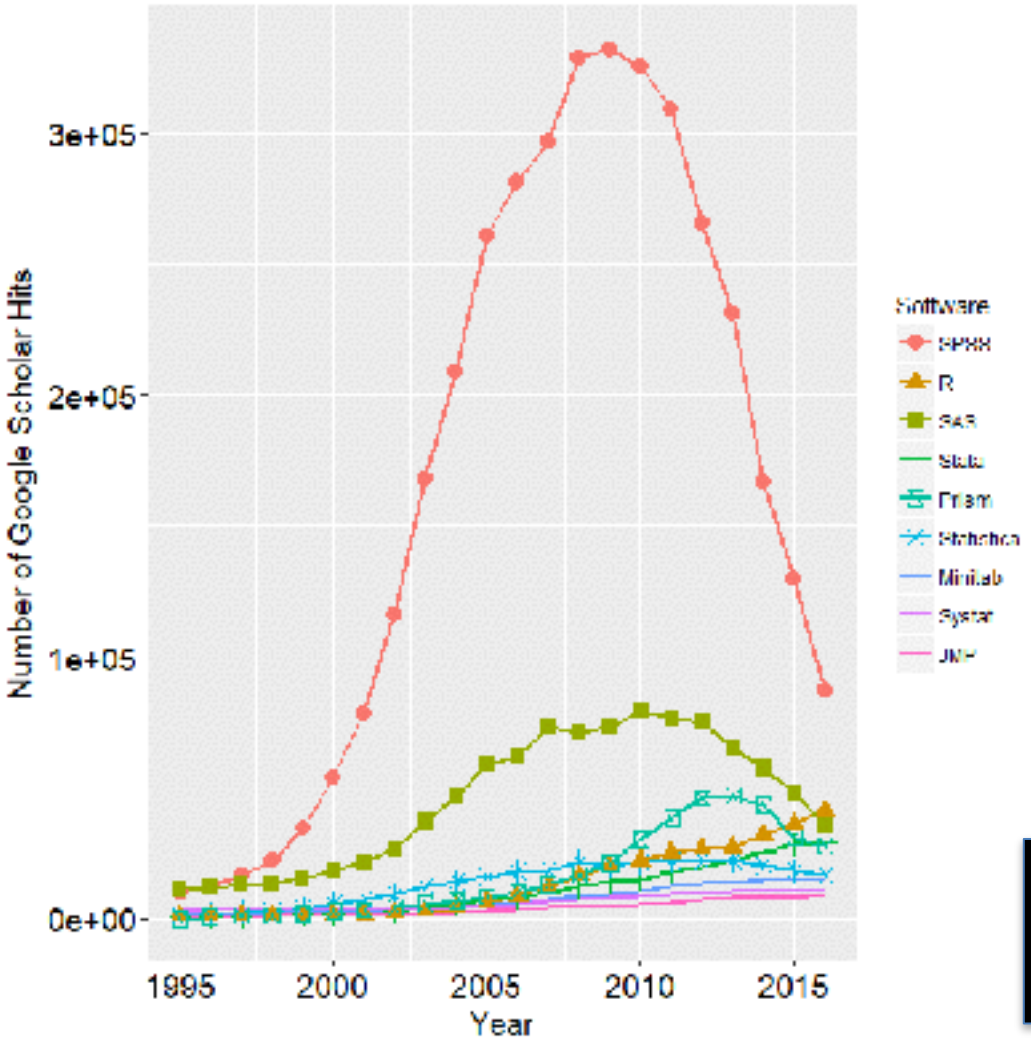
Note: Open source data-analysis languages eg. R, Go show huge gains in rankings

Vs

proprietary data-analysis languages: eg Matlab, SAS

http://spectrum.ieee.org/computing/software/the-2016-top-programming-languages

Why You Should Learn R



No. Of Google Scholar Hits

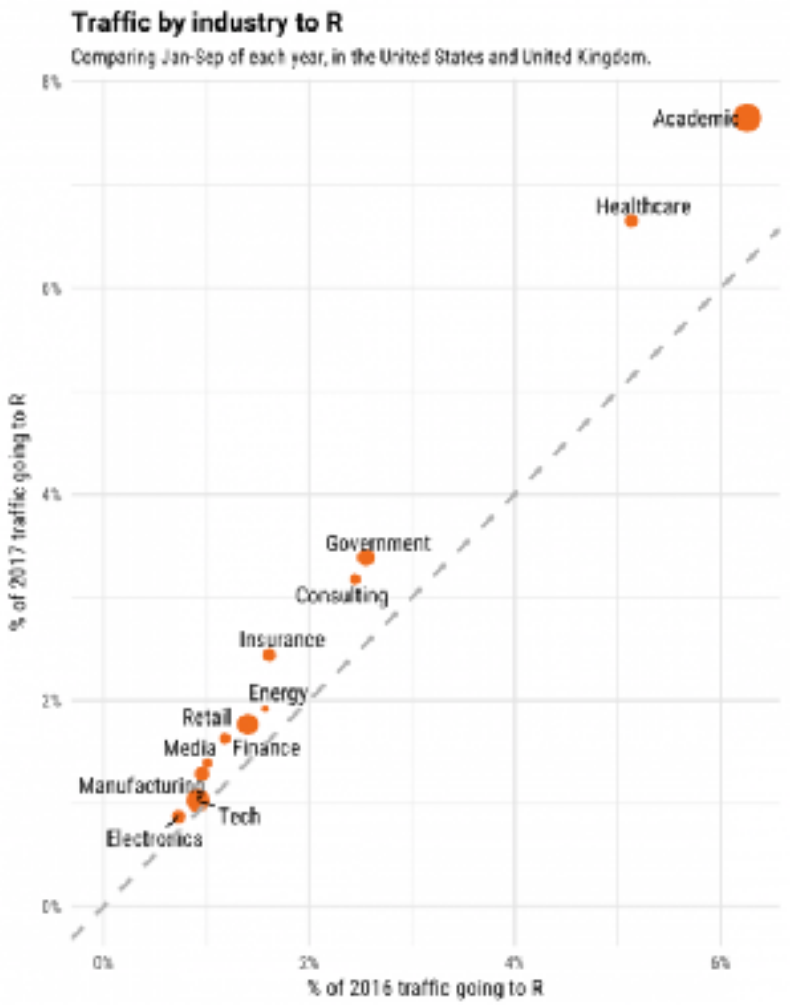
Usage of Various Softwares

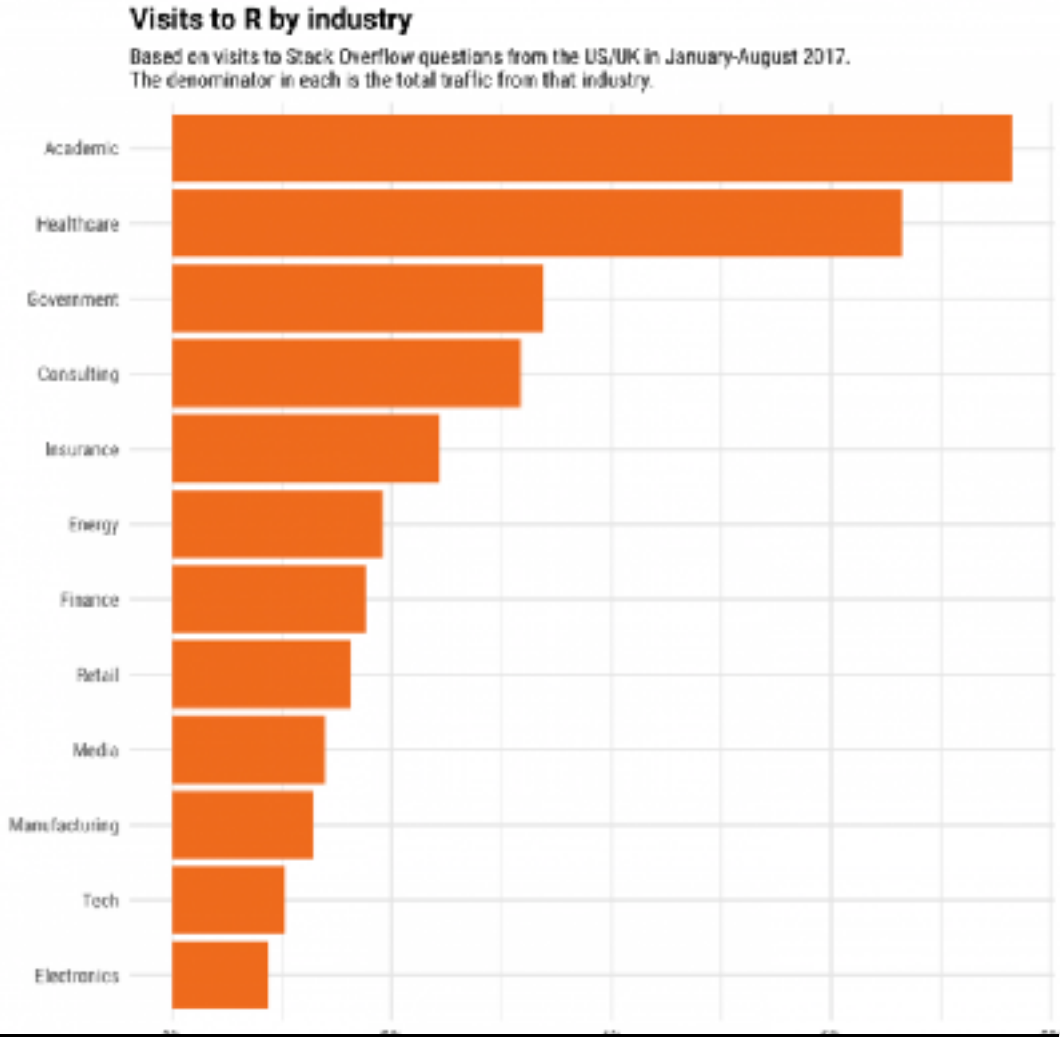
1995-2015

Increasing trend on R vs
Decreasing trend of other softwares

Why You Should Learn R?

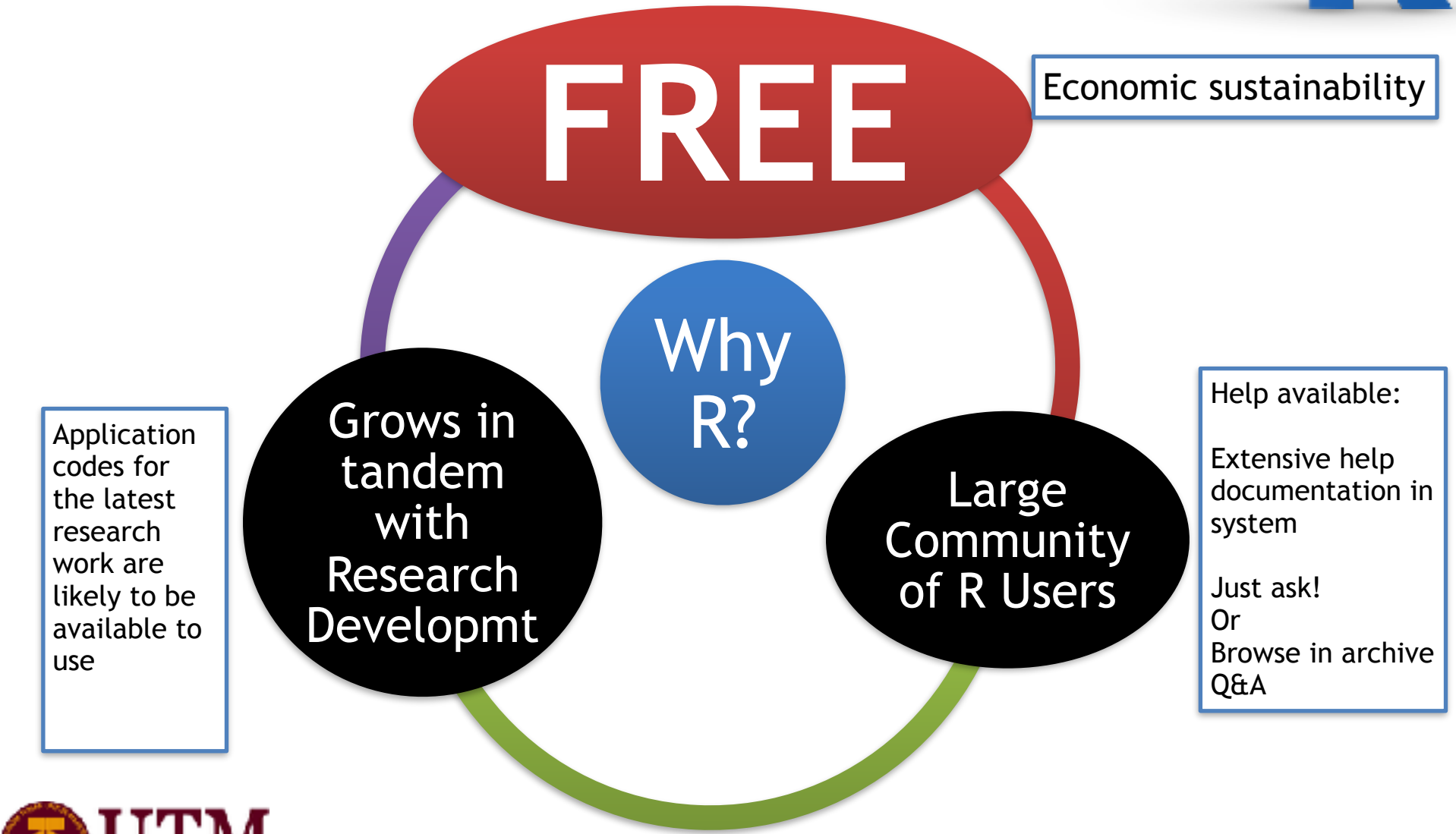
Traffic by Industry to R to Stack Overflow website



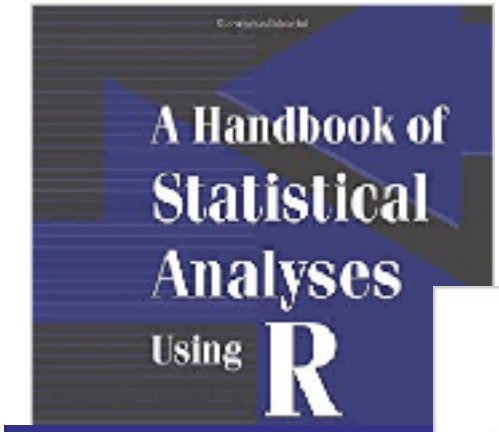


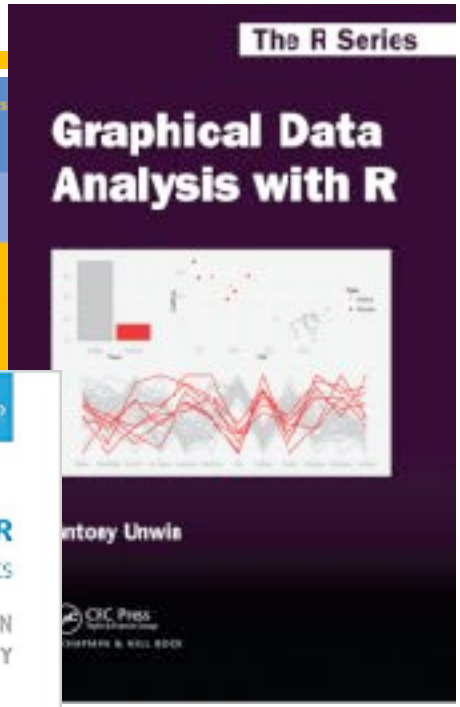
Stack Overflow is the largest, most trusted online community for developers to learn, share their programming knowledge.

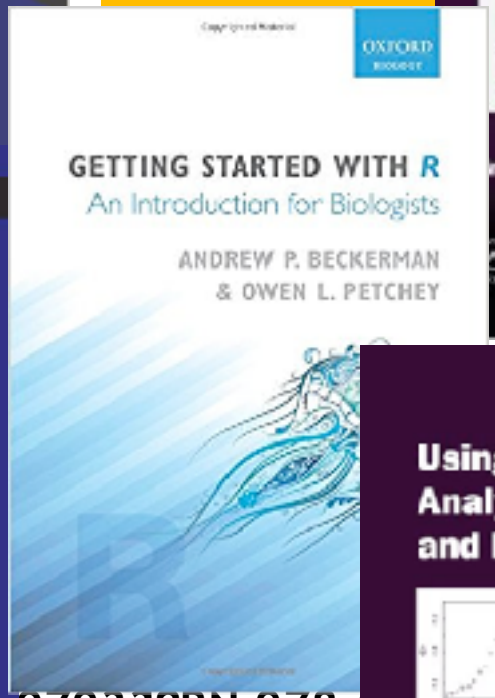
Why you should learn R?

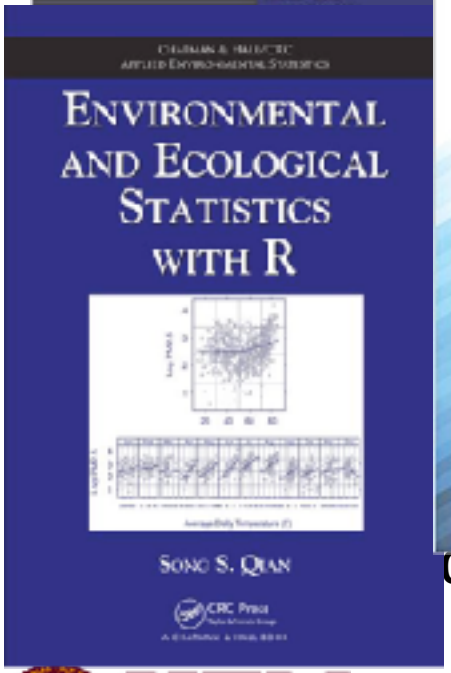


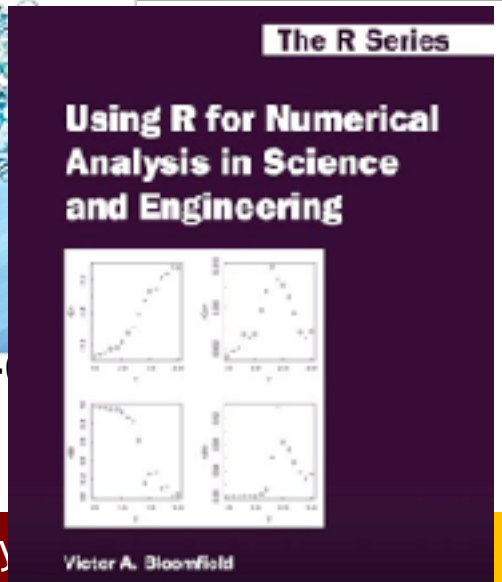
Books on R











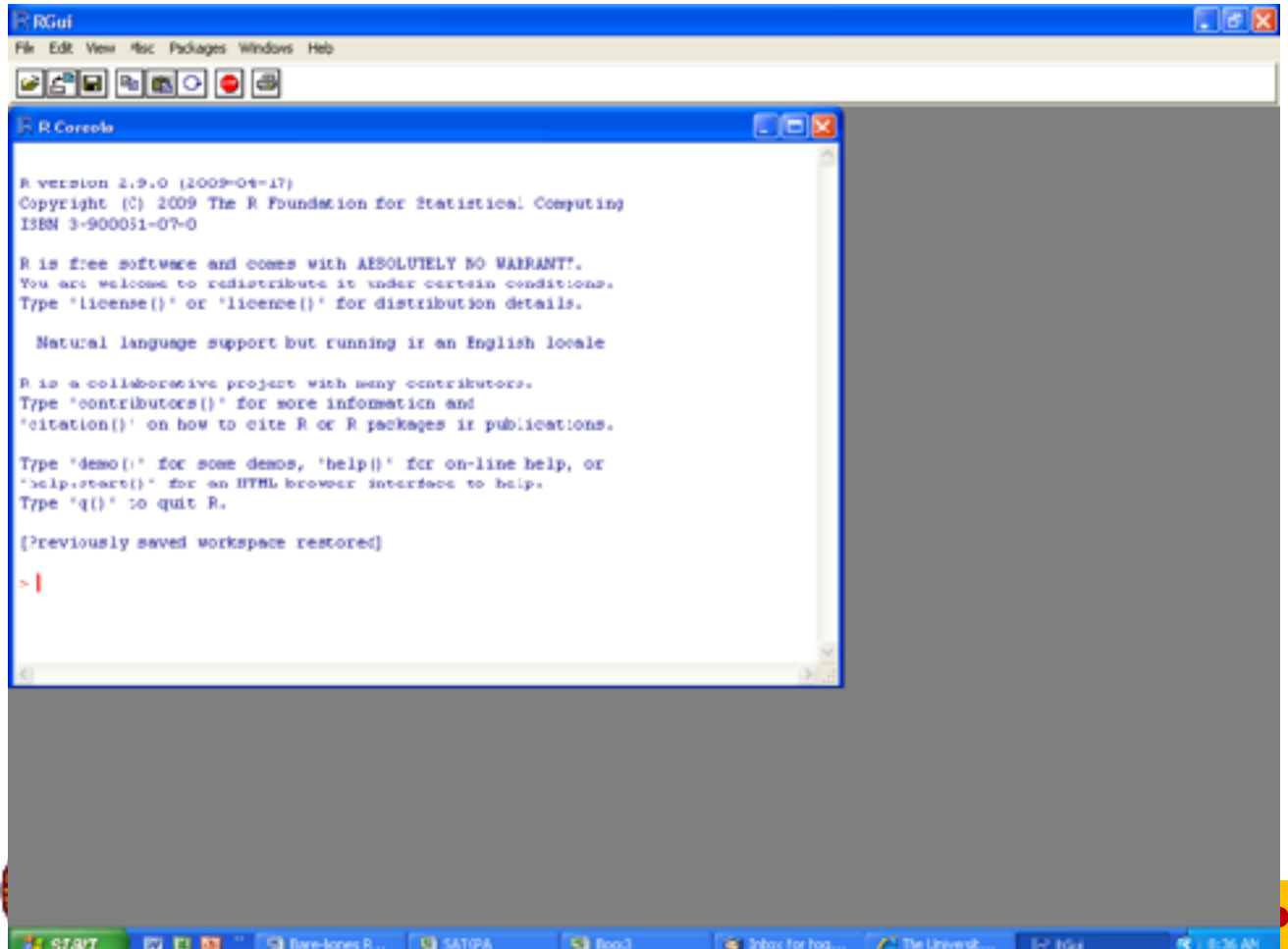
Plenty of online Resources!

How to download R?

http://www.r-project.org/

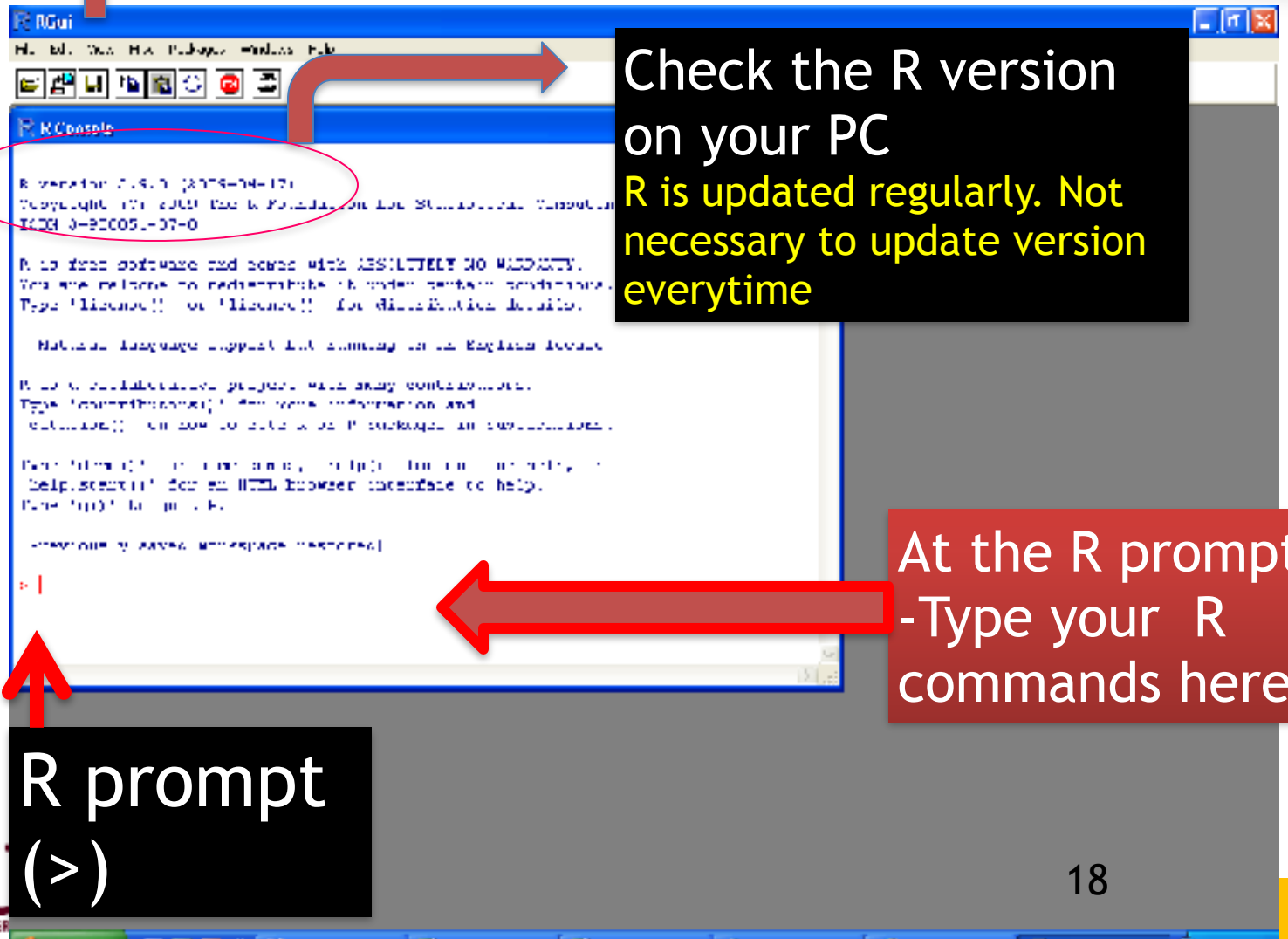
How to Run commands in R: R Console

- Displayed at the beginning of an Rsession
- Computations are performed on an R console



- R commands are typed and evaluated here.

How to Run commands in R: R Console



How to use R: as calculator

- Write your R commands after each prompt
- Hit **Enter** to execute command

```
> 6+3
> 6-3
> (22+ 18+35)/3
> 31 %% 7
```

- Other operators: +, -, *, /

How to use R: using R functions

- Use R functions for more *complicated* operations

Example of an R function

```
> mean(c(20,10,30))
> par(mfrow=c(1,1))
> plot(AirPassengers,type="l")
> library(ggmap)

> qmap(location = "Universiti Teknologi Malaysia")
```

- Example of R functions: `mean; par; plot; library; qmap`

How R works: Anatomy of R

- R has many inbuilt **functions, source codes, datasets & Help documentation**
- These are contained in **'Packages'** developed by R-team and the community

Package 'base'

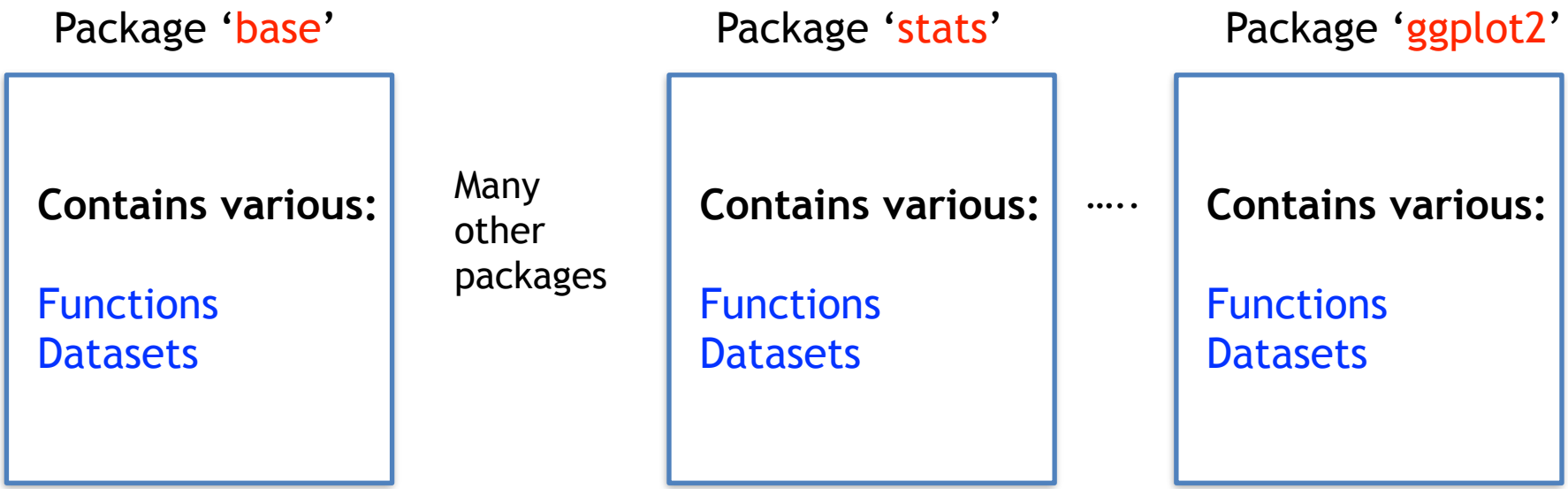
Contains various:

Functions
Datasets

Base Package
- Core package
- Automatically installed when you download R

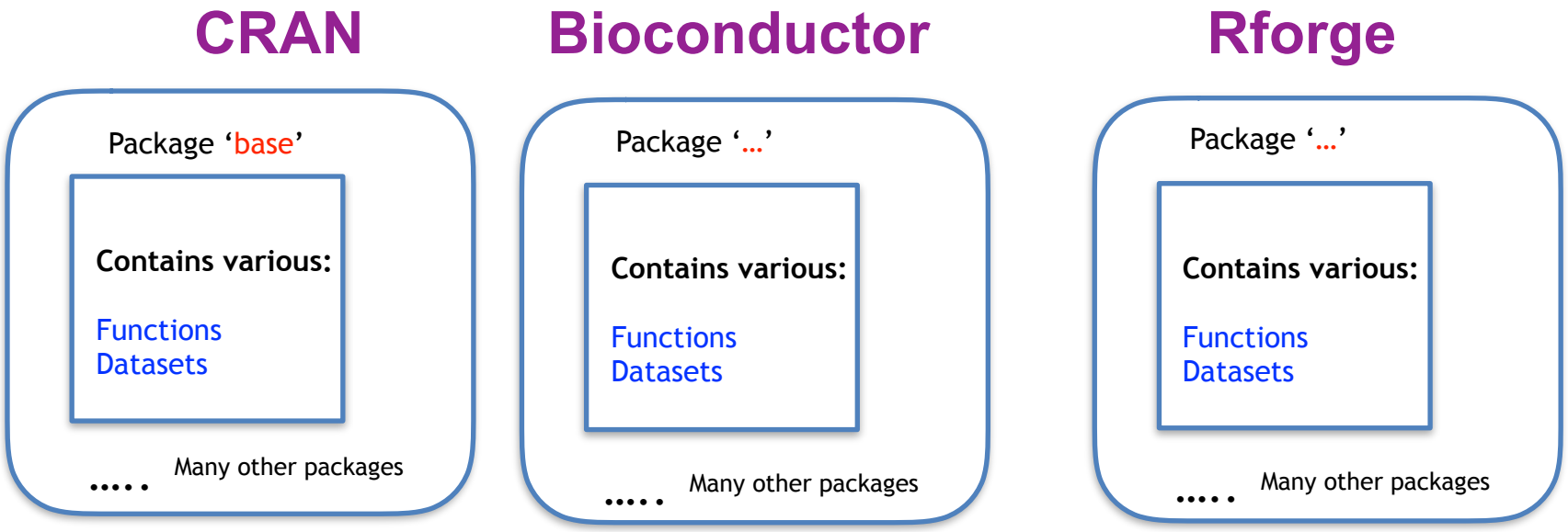
How R works: Anatomy of R

- R has many inbuilt **functions, source codes, datasets & Help documentation**
- These are contained in **'Packages'** developed by R-team and the community



How R works: Anatomy of R

- R **Packages** are stored in certain **Repositories**:



R Official/Default

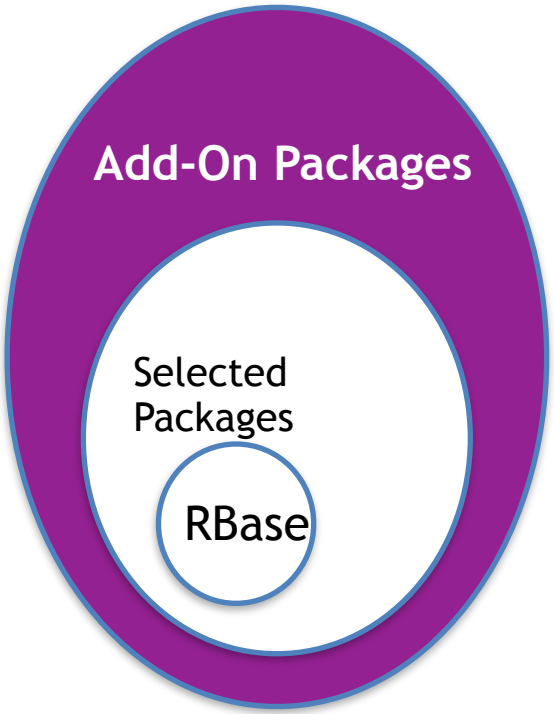
R specific to bioinformatics

Include development versions of packages

Other repository: **GitHub**

Not R but Repository for many open sourced projects

How R works: Anatomy of R



When you download and install R, you are downloading and installing the basic installation i.e Rbase and selected packages (from CRAN).

Packages in R= BRAIN of R.

Which type of R User are you?

CASUAL USER

- Point & Click User
- Not bothered with programming codes
- Similar to SPSS, Minitab etc

NON-CASUAL USER

- New/Experienced Programmer type user
- Need to modify the codes to suit own needs

Choice of Interface-R Users

NON-CASUAL USER:

NOVICE TO EXPERIENCE USER

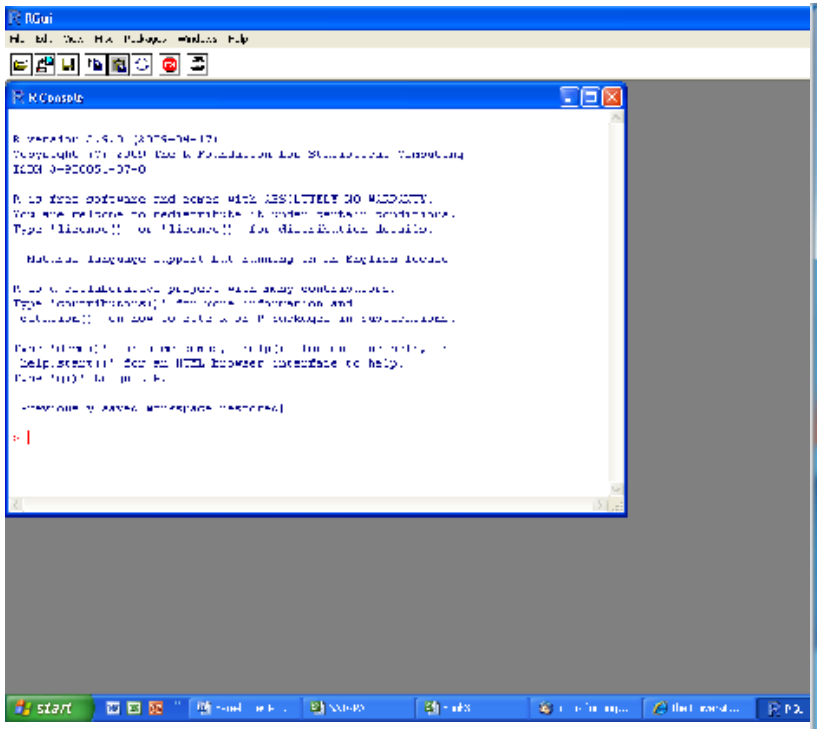
BACK-END

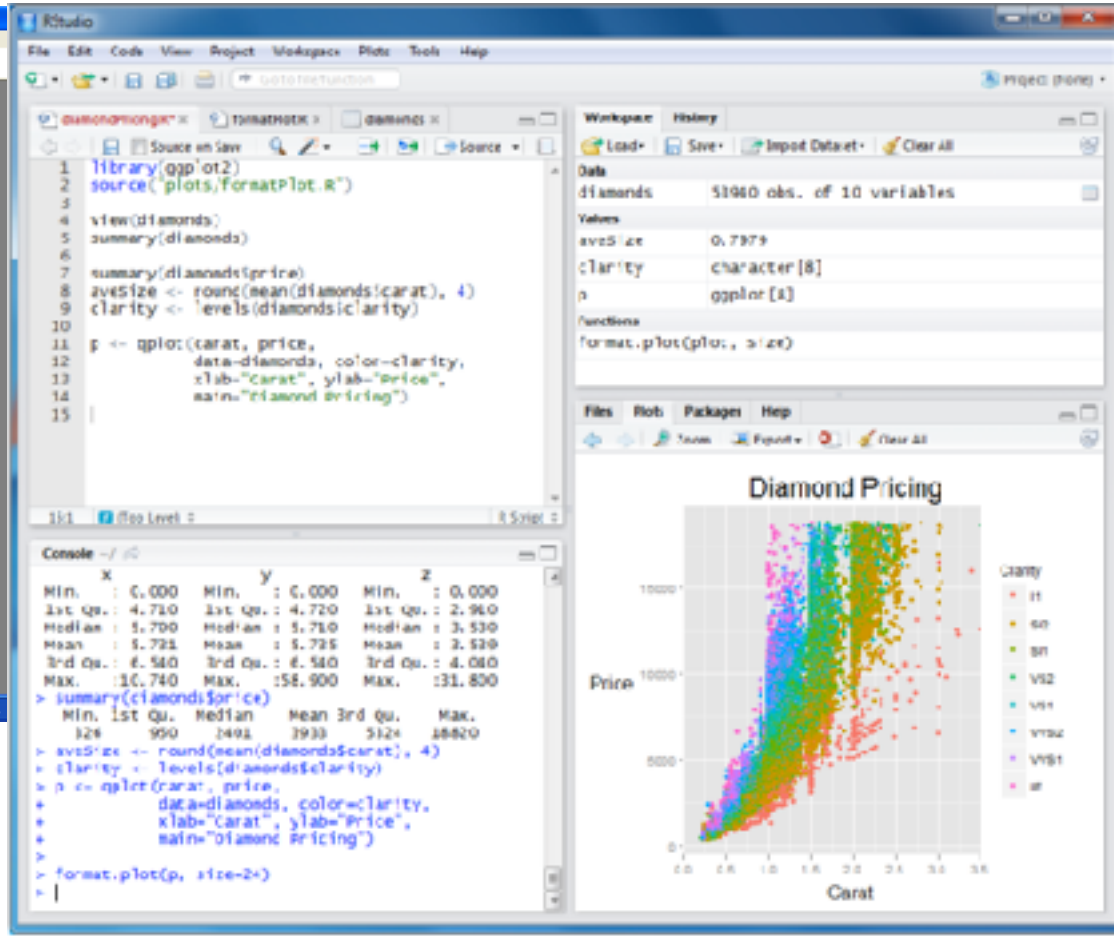
rbase

FRONT-END

rstudio

R Interface: R base vs R Studio





Rbase

RStudio

Choice of Interface-R Users

NON-CASUAL USER:

NOVICE TO EXPERIENCE USER

BACK-END

rbase

FRONT-END

rstudio

CASUAL USER

FRONT-END

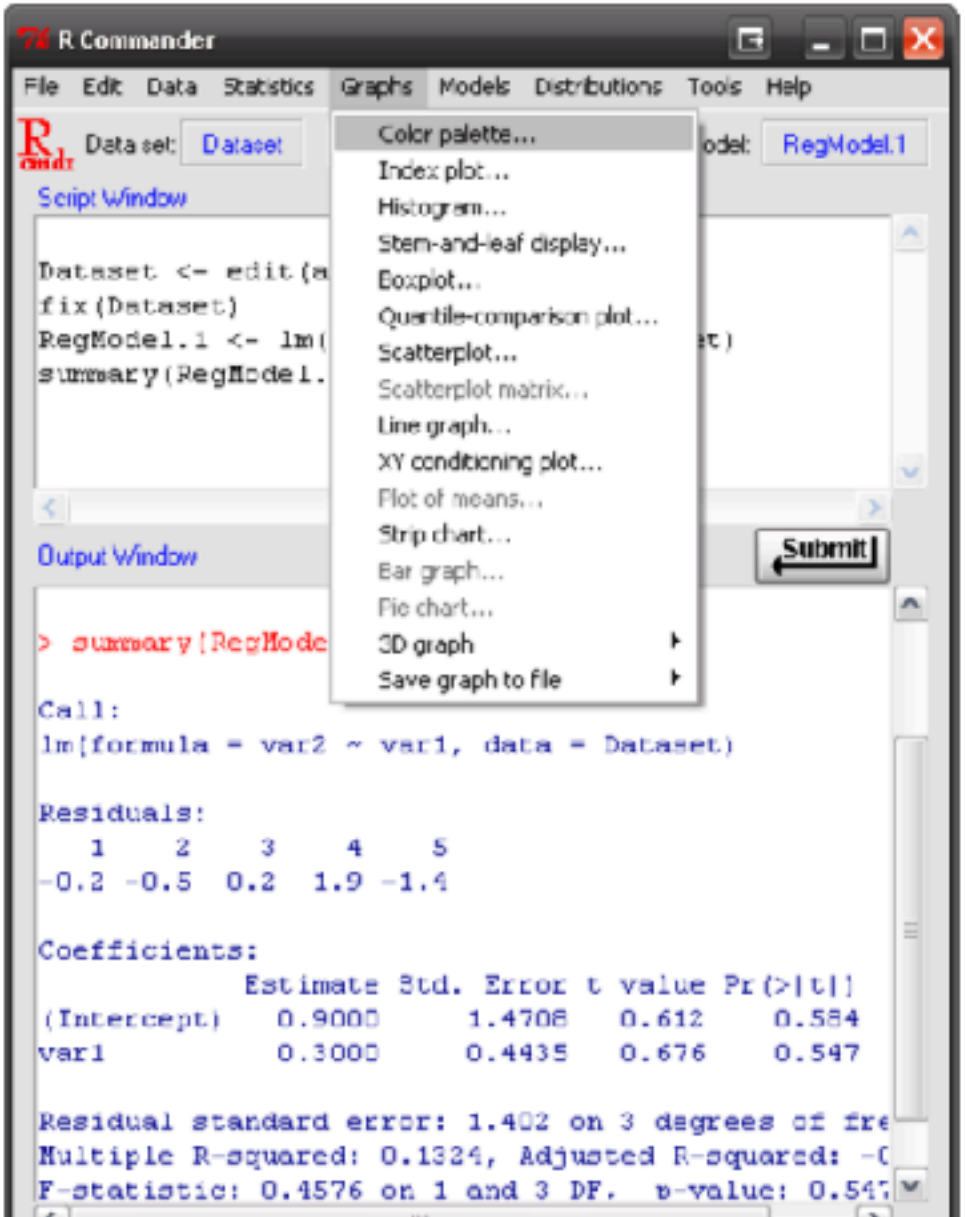
Rcommander

Which R-User are you?

R Interface: R Commander

R commander

- Point and click
- Dropdown functions
- Integrated Development Environment (IDE) for R-
- GUI based version of base R



Selections of R interface

BACK-END
rbase

- uses command-lines
- Basic platform to write program and run code in R

FRONT-END
rstudio

- Integrated Development Environment (IDE) for R
- RStudio makes life much easier for R coding but it is not a must-have to use R power.

FRONT-END
Rcommander

- Point and click
- Integrated Development Environment (IDE) for R-
- GUI based version of base R

RStudio/Rcommander Require Pre-Installation of Rbase

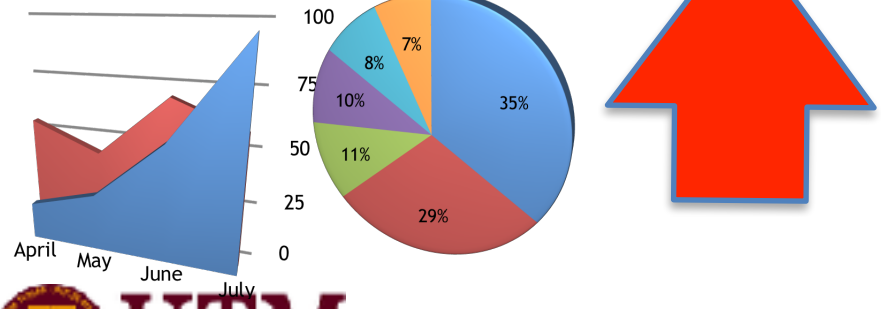
MAKING DATA SPEAK IN R:

VISUALIZATION

Types of Data Analysis

STATISTICAL DATA ANALYSIS

- VISUALIZATION
- INFERENCE

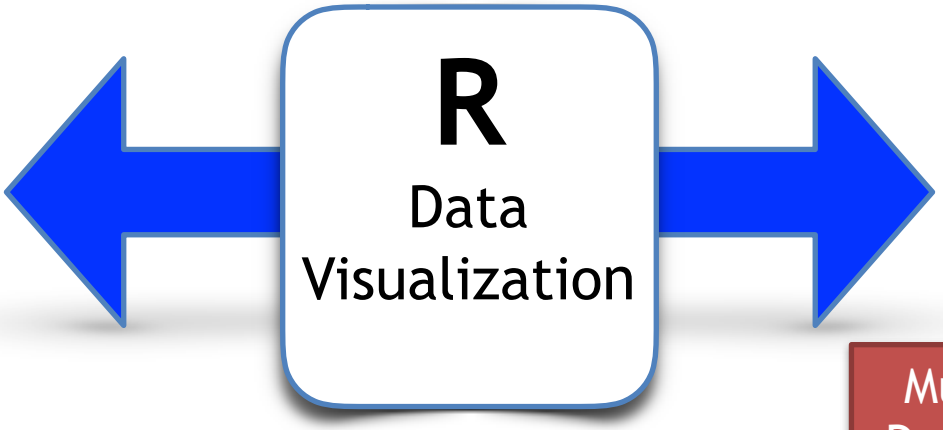


Hypothesis Testing
Confidence Interval
Modelling
Etc

Make Data Speak: Visualization

- Data visualization is an art of how to turn numbers into useful knowledge.
- A form of exploratory data analysis/data mining/ finding patterns etc
- Basic Presentation Types:
  - Comparison
  - Composition
  - Distribution
  - Relationship
- R Software offers various **inbuilt functions** and **packages** to build visualizations and present data.

Types of Visualisation in R



- Histogram
- Bar/Line Chart
- Scatter Plot
- Box Plot

Posh visuals

- 3D Graphs
- Heatmap
- Map Visualization

Multivariate Analysis/
Data Mining Techniques

- Regression models
- Cluster analysis
- Decision Trees
- Time Series Analysis
- Multidimensional Scaling

Simple Visual: IRIS data

- Iris flower data set is a collection of data to quantify the morphologic variation of Iris flowers.
- multivariate data set introduced by the British statistician and biologist Ronald Fisher in his 1936 paper *The use of multiple measurements in taxonomic problems*

Iris setosa

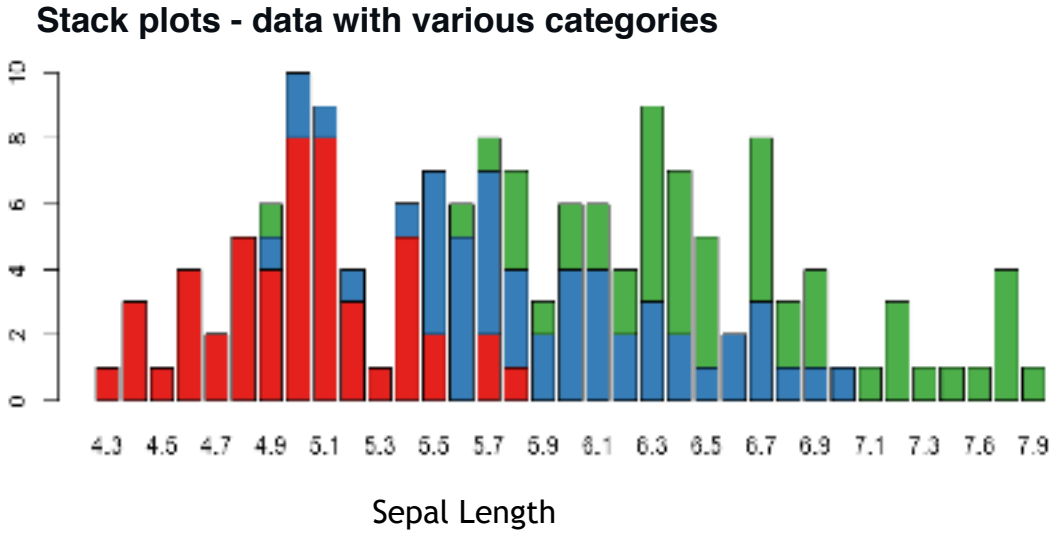
Iris versicolor

Iris virginica

50 samples from each of three species: Iris setosa, Iris versicolor and Iris virginica. Four components of the flowers' features were measured from each sample: length & width of the sepals and petals respectively.

```
> 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
..
..
49           5.3         3.7          1.5         0.2     setosa
50           5.0         3.3          1.4         0.2     setosa
51           7.0         3.2          4.7         1.4 versicolor
52           6.4         3.2          4.5         1.5 versicolor
..
..
99           5.1         2.5          3.0         1.1 versicolor
100          5.7         2.8          4.1         1.3 versicolor
101          6.3         3.3          6.0         2.5  virginica
102          5.8         2.7          5.1         1.9  virginica
..
149          6.2         3.4          5.4         2.3  virginica
150          5.9         3.0          5.1         1.8  virginica
```

Simple Visual: Iris -Stack & Box Plot




Iris setosa


Iris versicolor


Iris virginica

Sepal Length differs between Iris species. Setosa tends to have short sepal and Virginia tend to have longer sepal.

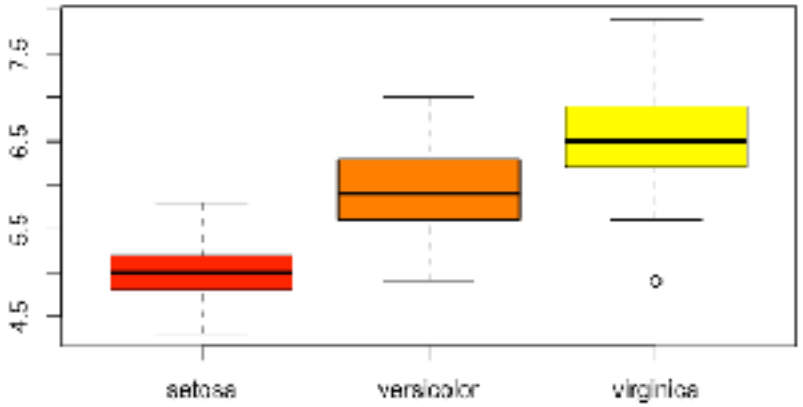
```
> iris
> barplot(table(iris$Species,iris$Sepal.Length))

> iris
> barplot(table(iris$Species,iris$Sepal.Length),col  = brewer.pal(3,"Set1"))
```

Simple Visual: Iris -Stack & Box Plot



Box plots - useful to show spread of data. Shows 5 statistically significant numbers- min, 25th percentile, median, 75th percentile and the max.

Show the spread (of Sepal Length) across various categories of Species



Sepal Length

```
> boxplot(iris$Sepal.Length~iris$Species)
```

```
> boxplot(iris$Sepal.Length~iris$Species,col=heat.colors(3))
```

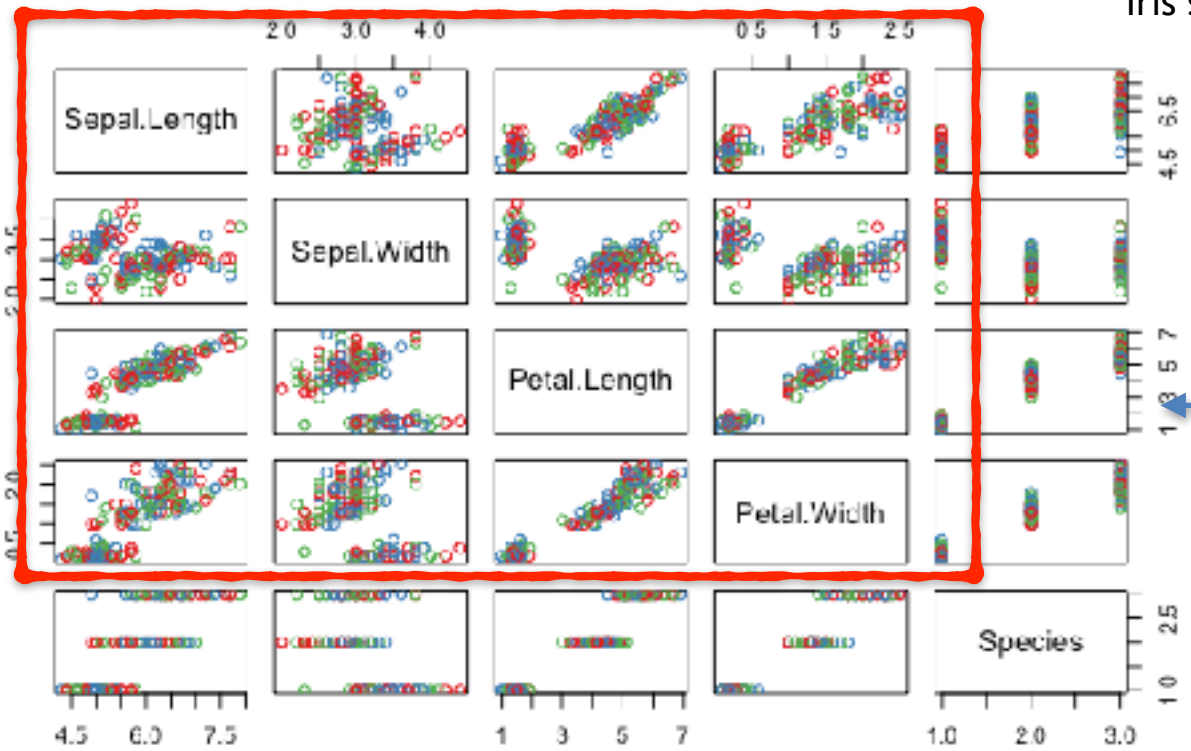
Simple Visual: Iris - Scatter Plot

Scatter Plot Matrix: visualize multiple variables across each other.


Iris setosa


Iris versicolor


Iris virginica



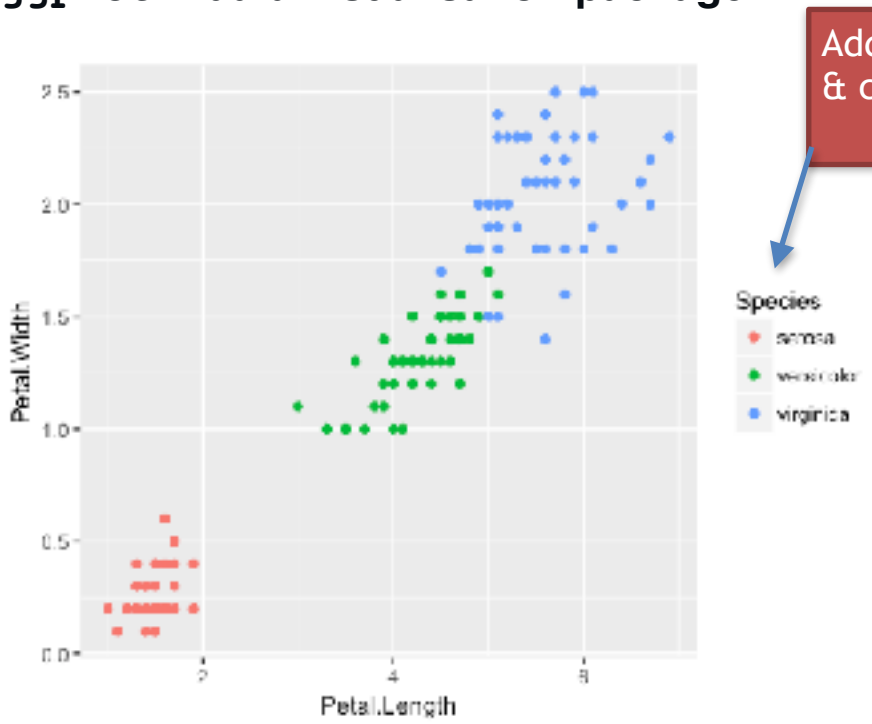
Plots shows some degree of relationship between the variables.

Eg.
Petal Length vs Petal Width
Sepal Length vs Petal Length

Detect relationships between variables

Simple+Advanced Visual: Iris - package **ggplot2**

ggplot2: data visualisation package in R



Plot of Petal Width vs Petal Length across species with legends and colour.
Indicates linear relationship between petal's width & length.


Iris setosa

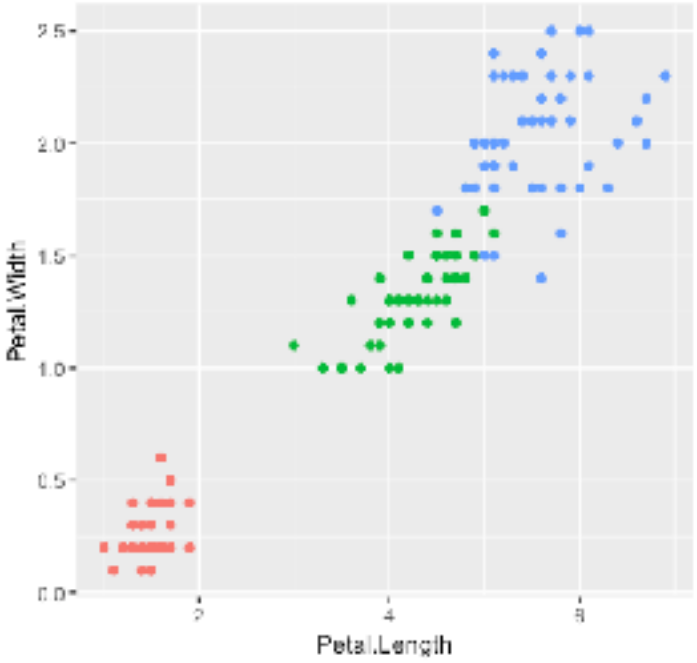

Iris versicolor


Iris virginica

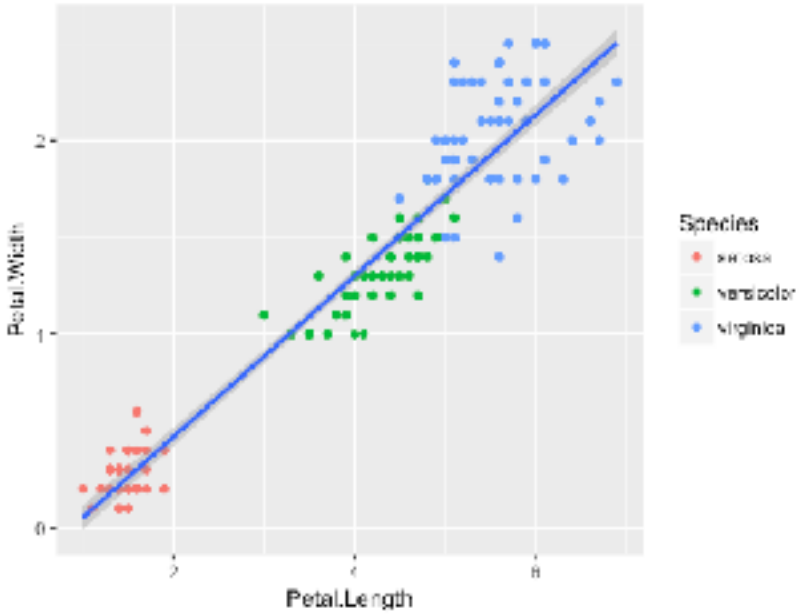
```
>library(ggplot2)
> p1 <- ggplot(data = iris, aes(x = Petal.Length, y = Petal.Width)); p1 #setgraph paper
> p2 <- p1 + geom_point(aes(color = Species));p2 #use geom to specify what to plot
```

Simple+Advanced Visual: Iris - package **ggplot2**

ggplot2: data visualisation package in R



Add legend & colour

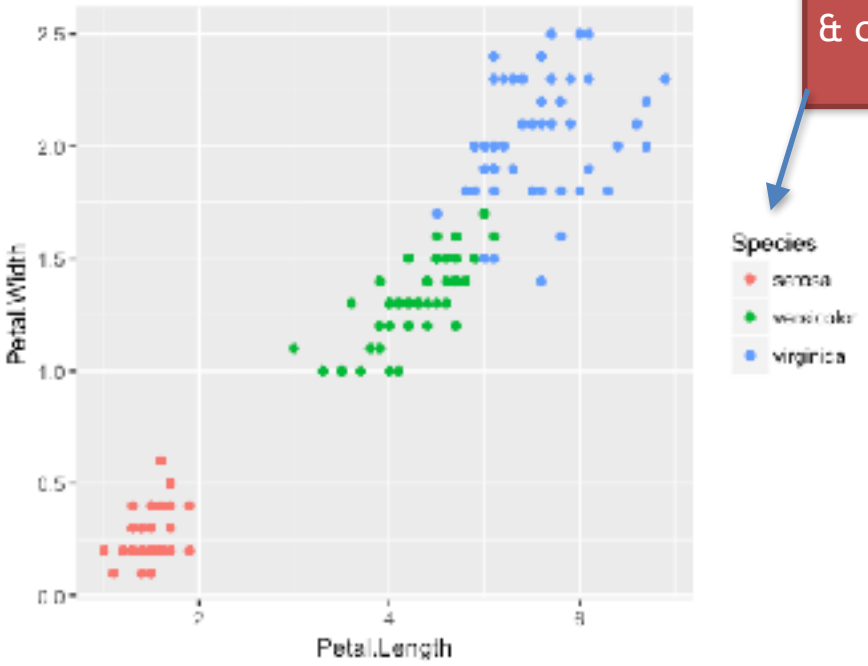


Fit a **linear regression line** to the plot

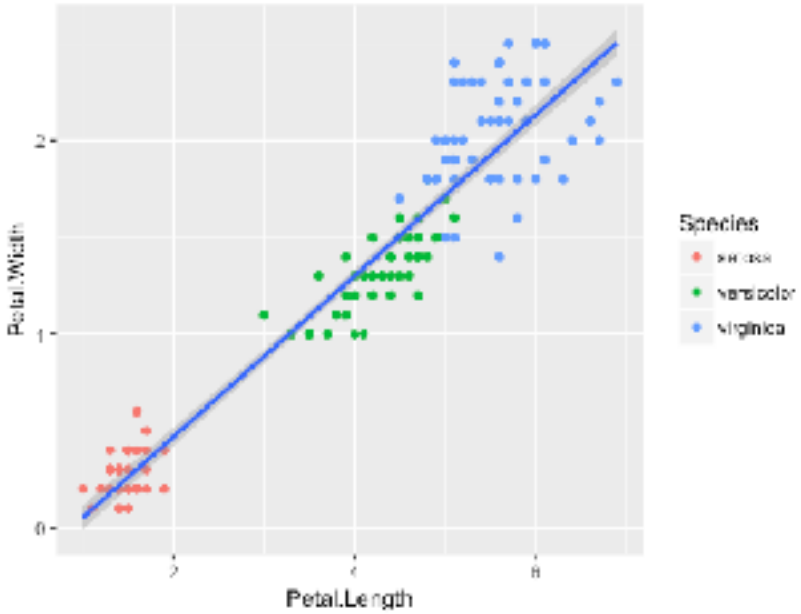
```
>library(ggplot2)
> p1 <- ggplot(data = iris, aes(x = Petal.Length, y = Petal.Width)); p1 #setgraph paper
> p2 <- p1 + geom_point(aes(color = Species));p2 #use geom to specify what to plot
> p3 <- p2 + geom_smooth(method='lm');p3 #add a linear regression model to fit the data
> p4 <- p3 + xlab("Petal Length (cm)") + ylab("Petal Width (cm)") + ggtitle("Petal Length versus Petal Width"); p4 #create/modify title
```

Simple+Advanced Visual: Iris - package `ggplot2`

`ggplot2`: data visualisation package in R



Add legend & colour



Plot of Petal Width vs Petal Length across species with legends and colour.
Indicates linear relationship between petal's width & length.

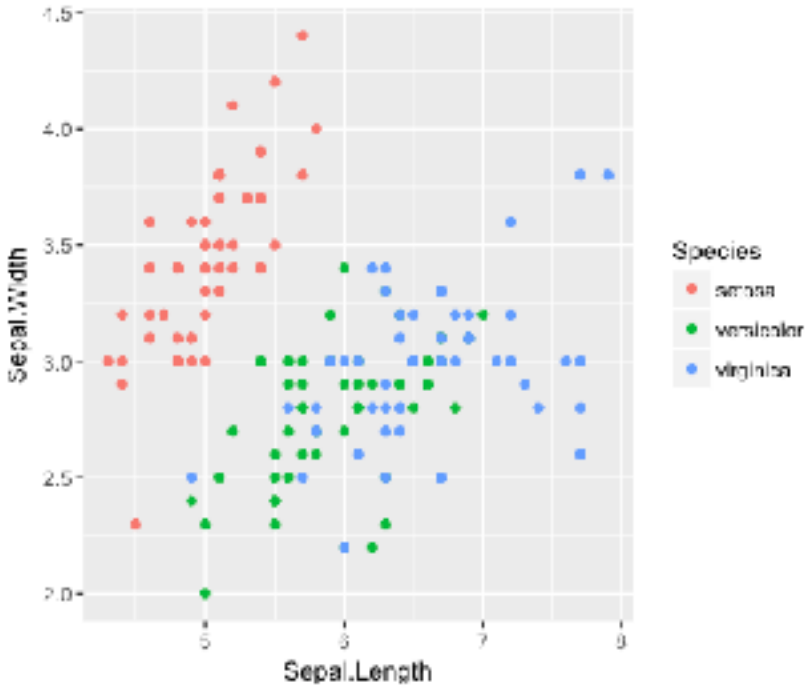
Fit a **linear regression line** to the plot


Iris setosa


Iris versicolor


Iris virginica

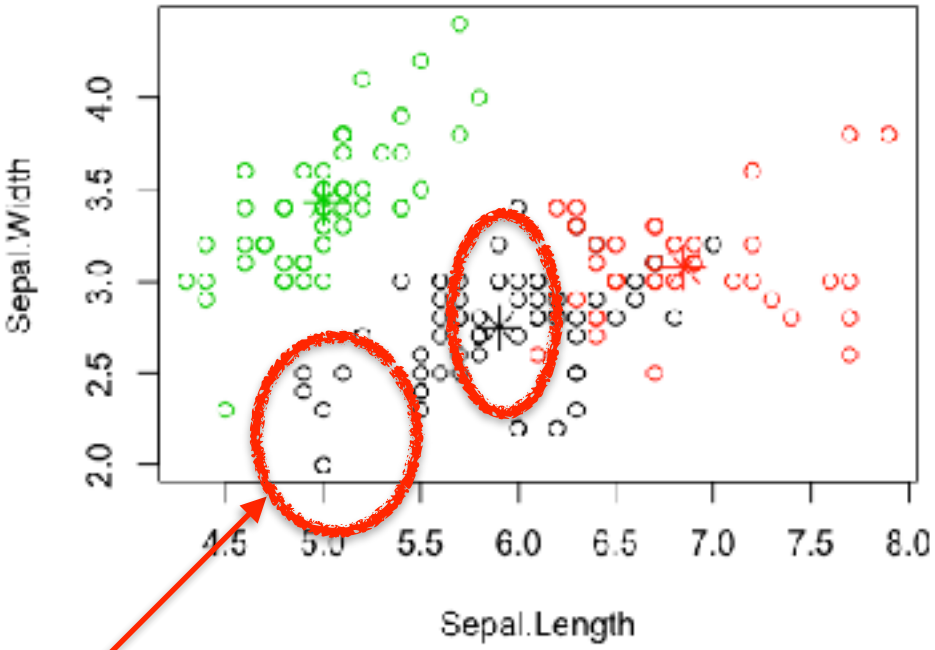
Advanced Visual: Iris - package **cluster**



Plot of Sepal Width vs Sepal Length across species.

- cluster "setosa" can be easily separated from the other clusters
- clusters "versicolor" and "virginica" are to a small degree overlapped with each other.

K-means Clustering: distance based technique in R



Kmeans clustering- Identify **clusters** based on their similarity (distance measure) and their **centers**.

The clusters here are slightly different from the plot before.

Note that some black points close to the green center (asterisk) are actually closer to the black center in the four dimensional space.

Reveal different characteristics to the data when different techniques are used

Advanced Visual: Iris - package `party`

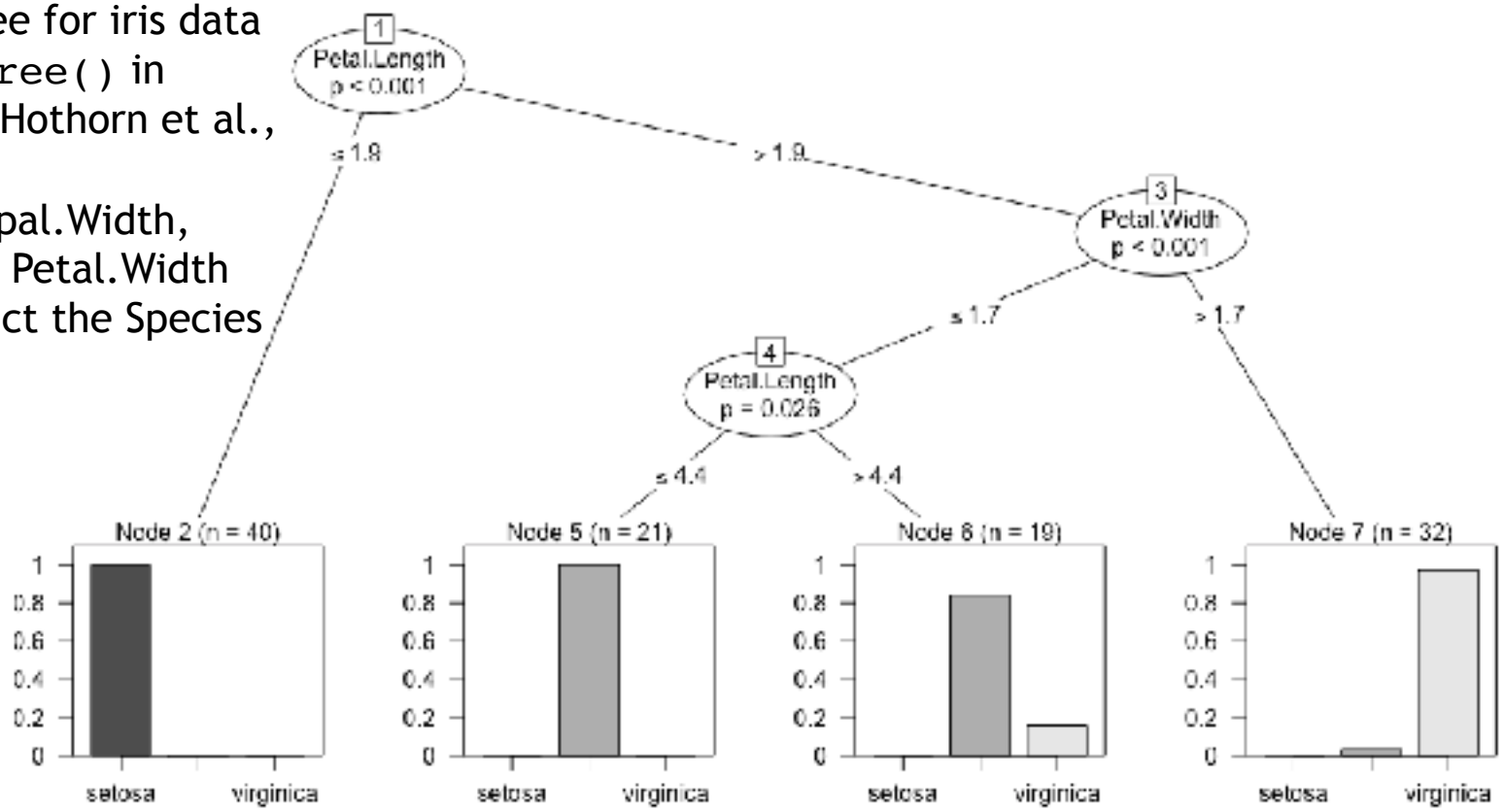
Classification using Decision Trees

Discriminant Analysis Functions for Predictive Modelling

iris data is split into two subsets: training (70%) and test (30%).

Species is the target variable and all other variables are independent variables

Build decision tree for iris data with function `ctree()` in package `party` [Hothorn et al., 2015].
Sepal.Length, Sepal.Width, Petal.Length and Petal.Width are used to predict the Species of flowers.



Reveal different characteristics to the data when different techniques are used

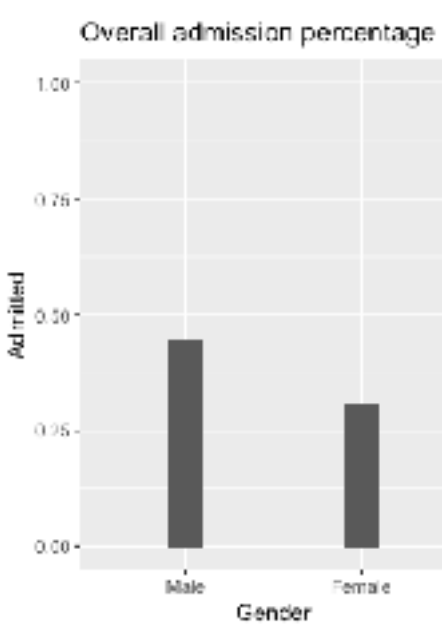
Simple Visual: Student Admissions data

Aggregate data on applicants to postgraduate school at Berkeley for the six largest departments classified by admission and sex.

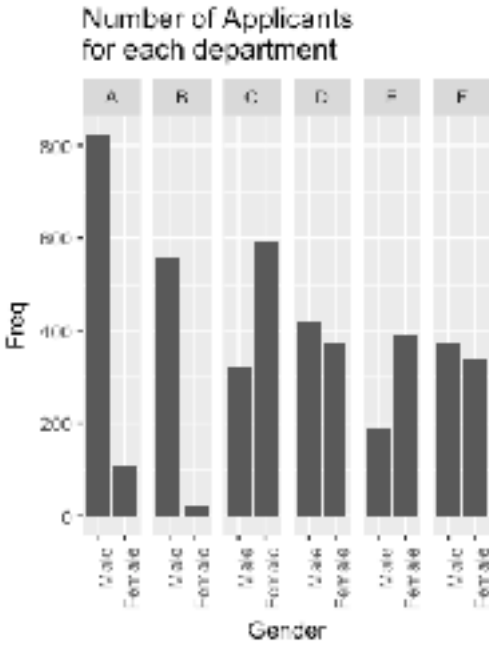
Admission Levels: Admitted/Rejected
Gender: Male/Female
Department: A-F

```
> UCBdt
       Admit Gender Dept Freq
1   Admitted   Male    A  512
2   Rejected   Male    A  313
3   Admitted Female    A   89
4   Rejected Female    A   19
5   Admitted   Male    B  353
6   Rejected   Male    B  207
7   Admitted Female    B   17
8   Rejected Female    B    8
9   Admitted   Male    C  120
10  Rejected   Male    C  205
11  Admitted Female    C  202
12  Rejected Female    C  391
13  Admitted   Male    D  138
14  Rejected   Male    D  279
15  Admitted Female    D  131
16  Rejected Female    D  244
17  Admitted   Male    E   53
18  Rejected   Male    E  138
19  Admitted Female    E   94
20  Rejected Female    E  299
21  Admitted   Male    F   22
22  Rejected   Male    F  351
23  Admitted Female    F   24
24  Rejected Female    F  317
```

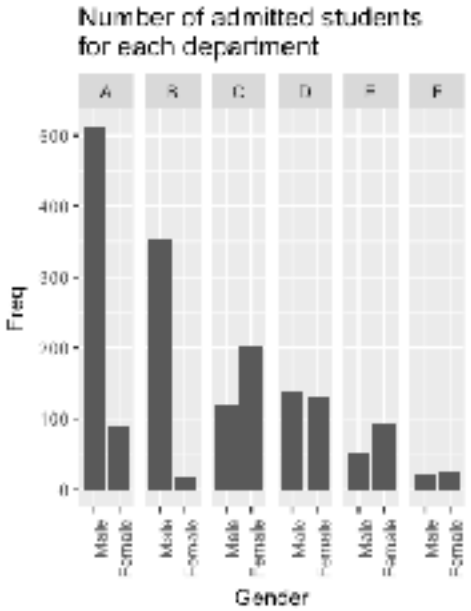
Simple Visual: Student Admissions -package **plyr**



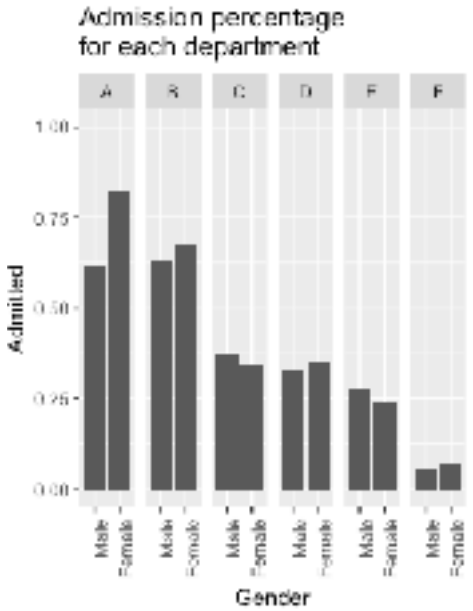
More males than females admitted to the university



Highest number of applicants for department A compared to the rest.



Highest number of admission for department A compared to the rest. Lowest number of admission admission for department F
Dept. A & B discriminate gender for admission.

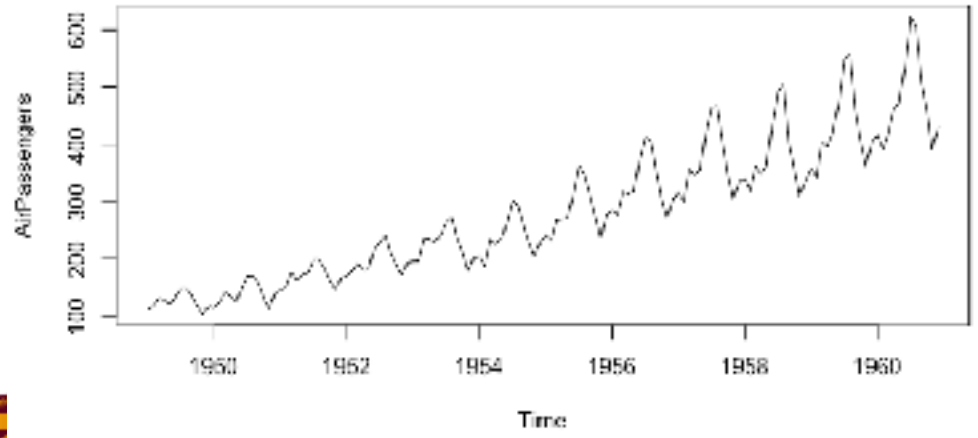


Simple Visual: Line Charts -time series data

Monthly Airline Passenger Numbers 1949-1960

```
> AirPassengers
     Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
1949 112 118 132 129 121 135 148 148 136 119 104 118
1950 115 126 141 135 125 149 170 170 158 133 114 140
1951 145 150 178 163 172 178 199 199 184 162 146 166
1952 171 180 193 181 183 218 230 242 209 191 172 194
1953 196 196 236 235 229 243 264 272 237 211 180 201
1954 204 188 235 227 234 264 302 293 259 229 203 229
1955 242 233 267 269 270 315 364 347 312 274 237 278
1956 284 277 317 313 318 374 413 405 355 306 271 306
1957 315 301 356 348 355 422 465 467 404 347 305 336
1958 340 318 362 348 363 435 491 505 404 359 310 337
1959 360 342 406 396 420 472 548 559 463 407 362 405
1960 417 391 419 461 472 535 622 606 508 461 390 432
```

```
>plot(AirPassengers,type="l")  #Simple Line Plot
```



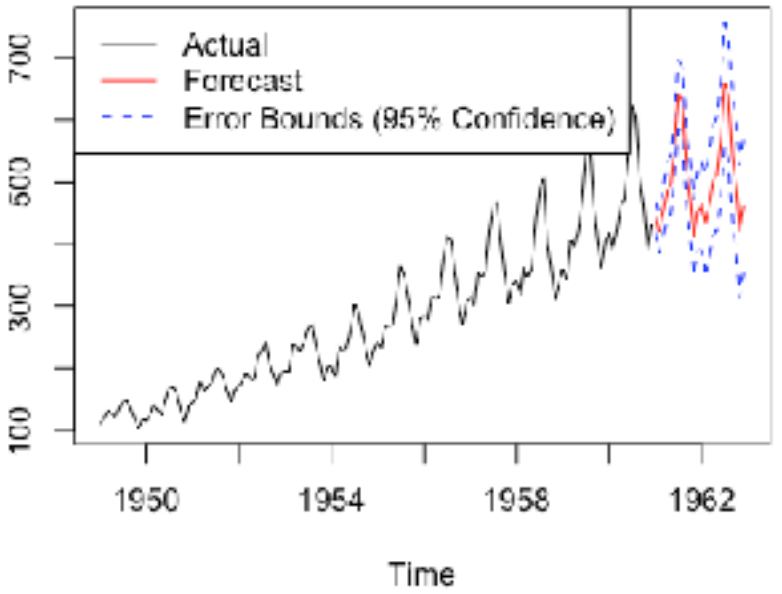
Line Charts :analyse trend spread over a time period. - or for comparing relative changes in quantities across some variable (eg.time).

Show increase in air passengers over a given time period.

Advanced Visual: Line Charts -time series data

Air Passengers 1949-1960

Air Passengers
Actual vs Forecasted Time Series

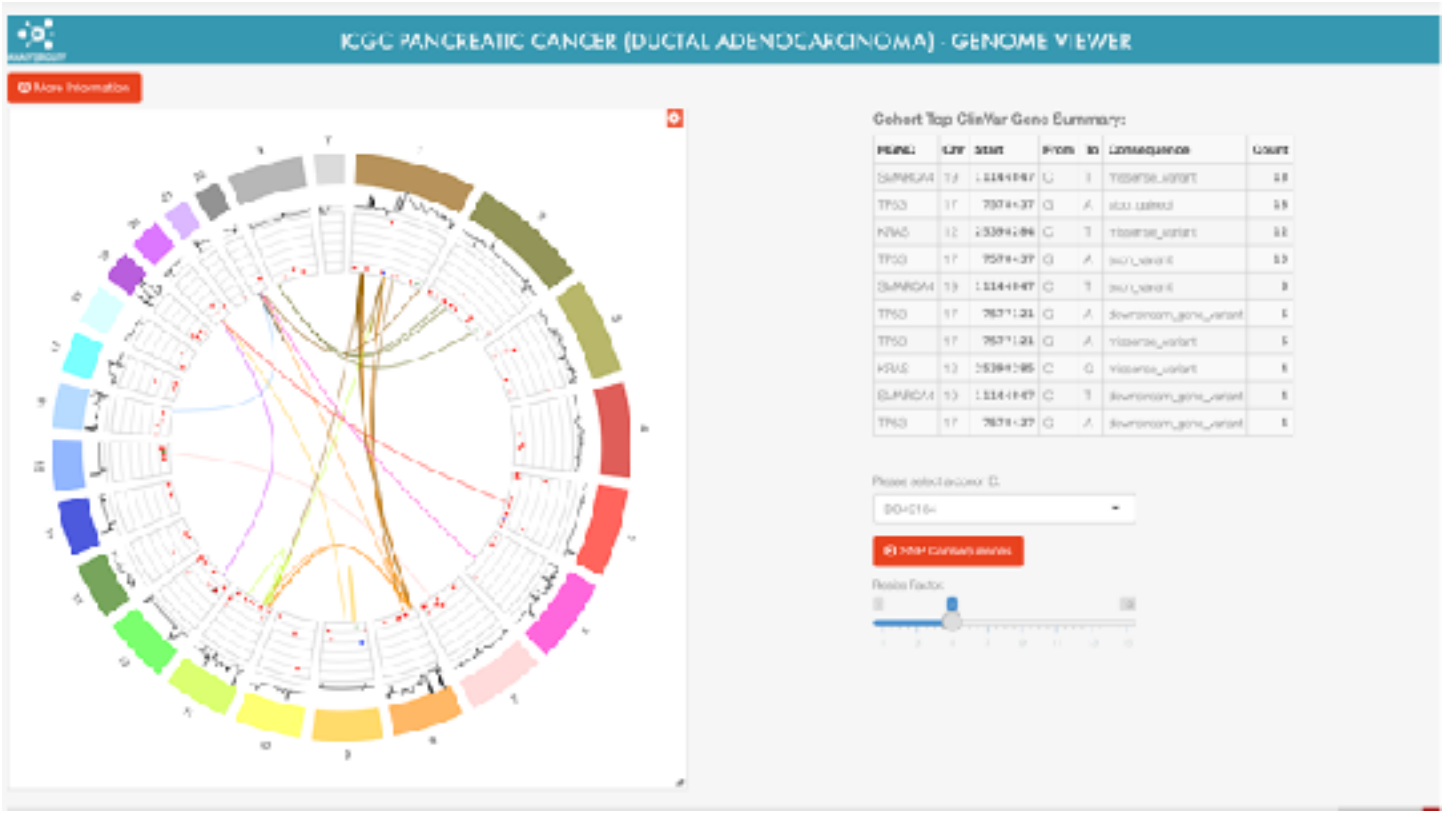


red solid line: forecasted values
blue dotted lines: error bounds at a confidence level of 95%.

Geographic/Spatial Visual: package **ggmap**


```
library(ggmap)
qmap(location = "Universiti Teknologi Malaysia") #qmap=quick map plot #search like google map
qmap(location = "Universiti Teknologi Malaysia",zoom=14)
qmap(location = "Universiti Teknologi Malaysia",zoom=14,maptype="hybrid")
```

Impressive Interactive Visualisation in R



Shiny by RStudio

R for Big Data

Big Data (5Vs): Volume (Tall & Wide); Variety (of secondary sources data); Velocity (Real-Time data), Value, Veracity

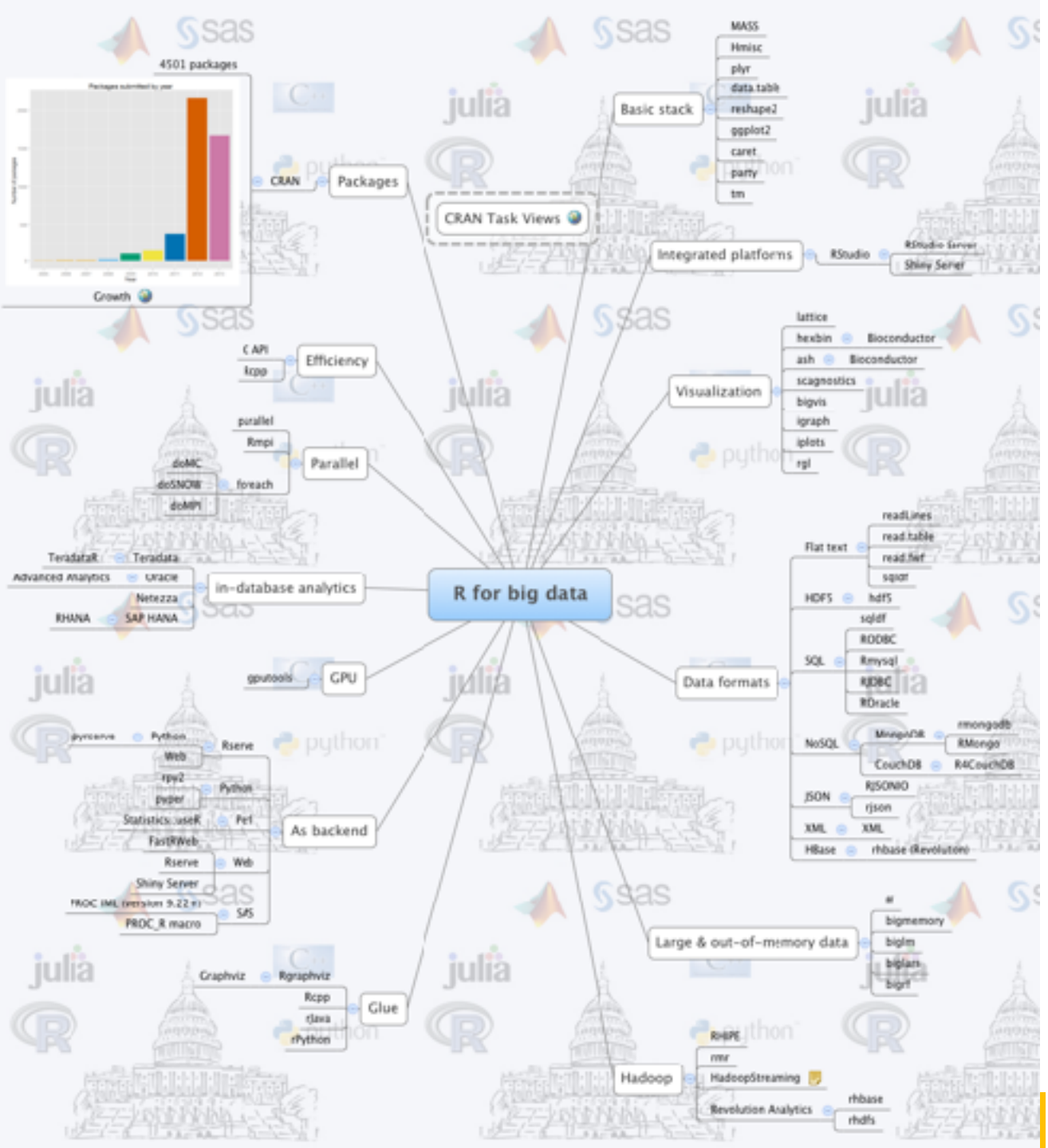
Challenge: Data could not load into memory (Data>RAM); Takes longer time to analyse data; Messy visualisations

Strategies in handling big data in R:

(1) Change memory size allocation in your R
(2) Use dimension reduction statistical methods to reduce dimension of data
(3) Take a subset of data: extract data to work with
(4) Increase machine memory
(5) Store big data in data warehouse
(6) Use R packages in R for Application Programming Interface (API) to data warehouse eg. data manipulation package `dplyr`
(7) Integrate with higher performing programming languages like C++ or Java

https://www.rstudio.com/resources/webinars/working-with-big-data-in-r/
http://www.columbia.edu/~sjm2186/EPIC_R/EPIC_R_BigData.pdf
https://www.r-bloggers.com/five-ways-to-handle-big-data-in-r/

R Packages in handling big data



https://www.datasciencecentral.com/profiles/blogs/r-for-big-data-in-one-picture

Additional Notes

Bridging R with other Softwares

- SAS, SPSS (version 16 onwards), Stata, Statistica, JMP
- Call C/Fortran from R
- Etc

Other softwares for visualisation

- Tableau, Plotly, DataHero, Chart.js, Raw, Dygraphs, ZingChart, InstantAtlas, Timeline, Exhibit, Modest Maps etc

Tips to New R Users

- Copy codes to get started

http://www.creativebloq.com/design-tools/data-visualization-712402

Tips for aspiring R newbies

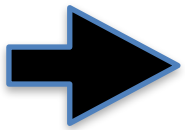
- Download R
- Find simple tutorial for Getting Started with R:
  - Books, internet etc
- Sign-up to any R workshop near you:

http://science.utm.my/mathematics/r-workshop/

- Try out new packages and Copy & Run the R codes
- Modify codes and tailor to your needs

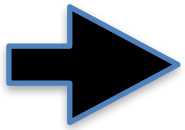
CONCLUSION

- If you can analyse & express your data effectively (BIG or SMALL)



your skills will be in high demand in a lot of places.

- R software is a tool for statistical analysis
  - **FREE**
  - **modeling / statistics tool with good visualization capability**
  - **Capable of LOW to HIGH end analysis**
  - Alternative to other costly softwares
  - **No programming skills - no problem!** Use various packages as templates



References

- Nawsher Khan, Ibrar Yaqoob, Ibrahim Abaker Targio Hashem, et al., "Big Data: Survey, Technologies, Opportunities, and Challenges," *The Scientific World Journal*, vol. 2014, Article ID 712826, 18 pages, 2014. doi:10.1155/2014/712826
- https://www.analyticsvidhya.com/blog/2015/07/guide-data-visualization-r/
- https://www.cio.com/article/3153389/it-industry/how-small-data-became-bigger-than-big-data.html
- https://www.analyticsvidhya.com/blog/2015/07/guide-data-visualization-r/
- http://www.creativebloq.com/design-tools/data-visualization-712402
- https://www.r-project.org/conferences/useR-2006/Slides/Chambers.pdf
- http://www.dataversity.net/big-data-small-data/

Rcodes used in slides

```
# Talk: Making data Speak: Using R Software
# Speaker: Dr. Haiza
# Codes used in slides


library(RColorBrewer)
#--------------------
iris
barplot(table(iris$Species,iris$Sepal.Length),col  = brewer.pal(3,"Set1")) #Stacked Plot

boxplot(iris$Sepal.Length~iris$Species,col=heat.colors(3))

#------------
# Advanced- Use package ggplot2   : IRIS data

library(ggplot2)
p1 <- ggplot(data = iris, aes(x = Petal.Length, y = Petal.Width)); p1 #setgraph paper
p2 <- p1 + geom_point(aes(color = Species));p2 #use geom to specify what to plot
p3 <- p2 + geom_smooth(method='lm');p3 #add a linear regression model to fit the data
p4 <- p3 + xlab("Petal Length (cm)") + ylab("Petal Width (cm)") + ggtitle("Petal Length versus Petal Width"); p4 #creat

#--------------------
# Advanced - Use package cluster   : Clustering iris flowers

sL1 <- ggplot(data = iris, aes(x = Sepal.Length, y = Sepal.Width)); sL1 #setgraph paper
sL2 <- sL1 + geom_point(aes(color = Species));sL2 #use geom to specify what to plot


iris2 <- iris
iris2$Species <- NULL
(kmeans.result <- kmeans(iris2, 3))
plot(iris2[c("Sepal.Length", "Sepal.Width")], col = kmeans.result$cluster) # plot cluster centers
points(kmeans.result$centers[,c("Sepal.Length", "Sepal.Width")], col = 1:3, pch = 8, cex=2)
```

Rcodes used in slides

```
#--------------------
# Advanced - Use package party   : Classify iris flowers - Decision Trees

library(party)
myFormula <- Species ~ Sepal.Length + Sepal.Width + Petal.Length + Petal.Width
iris_ctree <- ctree(myFormula, data=trainData)
table(predict(iris_ctree), trainData$Species)
print(iris_ctree)
plot(iris_ctree)
#-------------
# Advanced- Use package plyr   : Students Admission
library(plyr)
library(datasets)
UCBdt <- as.data.frame(UCBAdmissions)
overall <- ddply(UCBdt, .(Gender), function(gender) {
  temp <- c(sum(gender[gender$Admit == "Admitted", "Freq"]),
            sum(gender[gender$Admit == "Rejected", "Freq"])) / sum(gender$Freq)
  names(temp) <- c("Admitted", "Rejected")
  temp
})
departmentwise <- ddply(UCBdt, .(Gender,Dept), function(gender) {
  temp <- gender$Freq / sum(gender$Freq)
  names(temp) <- c("Admitted", "Rejected")
  temp
})

# A barplot for overall admission percentage for each gender.
p1 <- ggplot(data = overall, aes(x = Gender, y = Admitted, width = 0.2))
p1 <- p1 + geom_bar(stat = "identity") + ggtitle("Overall admission percentage") + ylim(0,1) ;p1

# A 1x6 panel of barplots, each of which represents the
# admission percentage for a department
p2 <- ggplot(data = UCBdt[UCBdt$Admit == "Admitted", ], aes(x = Gender, y = Freq))
p2 <- p2 + geom_bar(stat = "identity") + facet_grid(. ~ Dept) + ggtitle("Number of admitted students\nfor each department
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) ;p2

# A 1x6 panel of barplots, each of which represents the
# number of admitted students for a department
p3 <- ggplot(data = departmentwise, aes(x = Gender, y = Admitted))
p3 <- p3 + geom_bar(stat = "identity") + facet_grid(. ~ Dept) + ylim(0,1) + ggtitle("Admission percentage\nfor each depar
  theme(axis.text.x = element_text(angle = 90, hjust = 1));p3
# A 1x6 panel of barplots, each of which represents the
# number of applicants for a department
p4 <- ggplot(data = UCBdt, aes(x = Gender, y = Freq))
p4 <- p4 + geom_bar(stat = "identity") + facet_grid(. ~ Dept) + ggtitle("Number of Applicants\nfor each department") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)); p4
#--------------------
```

Rcodes used in slides

```
#--------------------

AirPassengers
par(mfrow=c(1,1))
plot(AirPassengers,type="l")  #Simple Line Plot


fit <- arima(AirPassengers, order=c(1,0,0), list(order=c(2,1,0), period=12))
fore <- predict(fit, n.ahead=24)
# error bounds at 95% confidence level
U <- fore$pred + 2*fore$se
L <- fore$pred - 2*fore$se
ts.plot(AirPassengers, fore$pred, U, L, col=c(1,2,4,4), lty = c(1,1,2,2))
legend("topleft", c("Actual", "Forecast", "Error Bounds (95% Confidence)"),col=c(1,2,4), lty=c(1,1,2))


#--------------------
# Advanced - - Use package ggmap   : Spatial Data - map


library(ggmap)
qmap(location = "Universiti Teknologi Malaysia") #qmap=quick map plot #search like google map
qmap(location = "Universiti Teknologi Malaysia",zoom=14)
qmap
```

Thank you

http://science.utm.my/norhaiza/

Department of Mathematical Sciences . Faculty of Science