

R for Analytics

Chris Campbell

John James

john@mango-solutions.com

Agenda



- This talk is about the R community
- Its about going down a corridor and opening doors
 - But not really going into them
- To disambiguate: Mango (and to check I'm online)



- R is an Environment
 - Language
 - Packages
 - Data
- Tools
 - Best update system available
 - 5000 packages each with 10-100 functions
 - Very active help community
- Projects

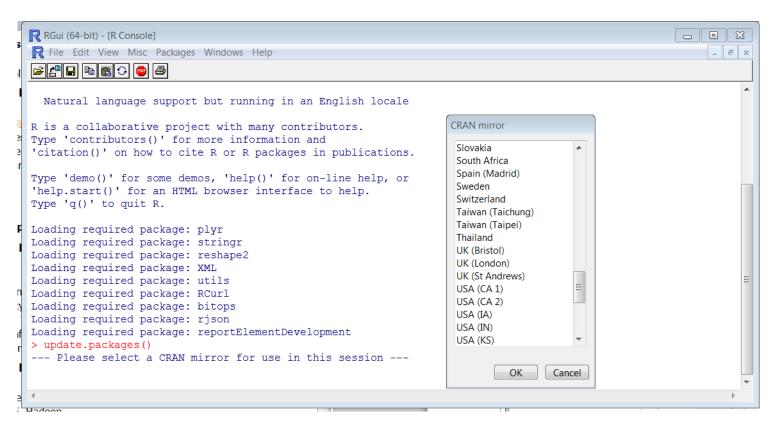


- Easy to use
 - Gui's (vanilla, RStudio, Eclipse, many more)
 - Intellisense, integrated help
- User community
 - Annual World Wide Conferences
 - Specialist Conference (R Finance, biometrics, ...)
 - Blogs
 - Local Communities (LondonR, ...)

Gui's (vanilla, RStudio, Eclipse, many more)

MANGOSOLUTIONS

data analysis that delivers



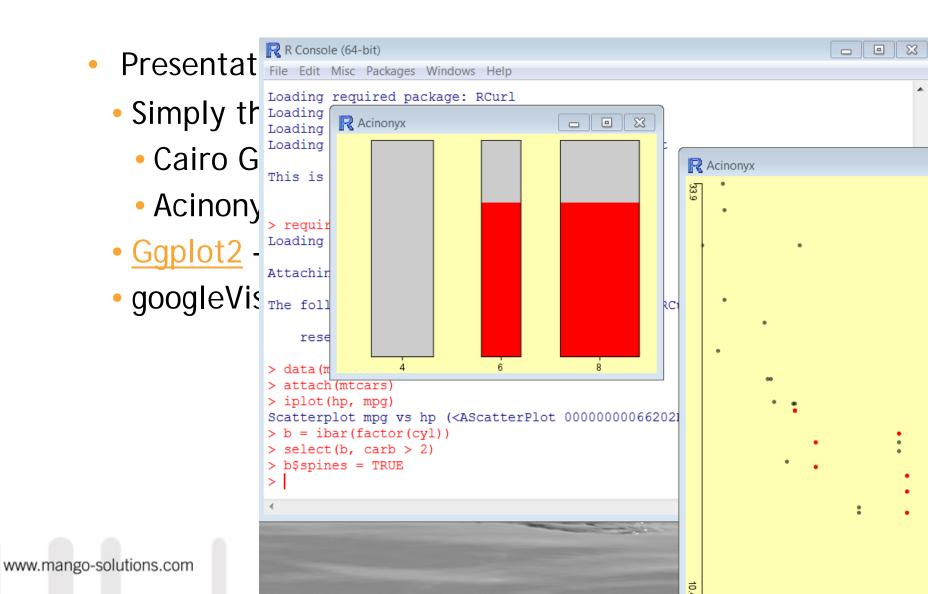


- Data
 - Everything web enabled
 - Large data h5, Hadoop
 - Parallel Processing (MPI, SMP ...)
- Integration with other programming languages
 - Rcpp
 - rJava
 - RdotNet
 - C
 - fortran



- Analysis
 - Every university
 - Every organisation
- Integration with other Systems
 - Oracle VM
 - RServe (java and .NET clients)
 - RdotNET
- Programming Support
 - Debugger
 - Documentation
 - Package Creation





The R environment



- an effective data handling and storage facility,
- a suite of operators for calculations on arrays, in particular matrices,
- a large, coherent, integrated collection of intermediate tools for data analysis,
- graphical facilities for data analysis and display either onscreen or on hardcopy, and
- a well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities.

Some Examples



- Graphics
 - GraphML
 - Interactive
- Parallel
 - Rmpi, parallel, smp ...
- Big Data
 - bigmemory, ff, rh5

What Next?



- China
- Bigger data, more parallelisation
- Professional Services
 - How to make money out of open source?
- julia
 - Revolution, Mango ...