



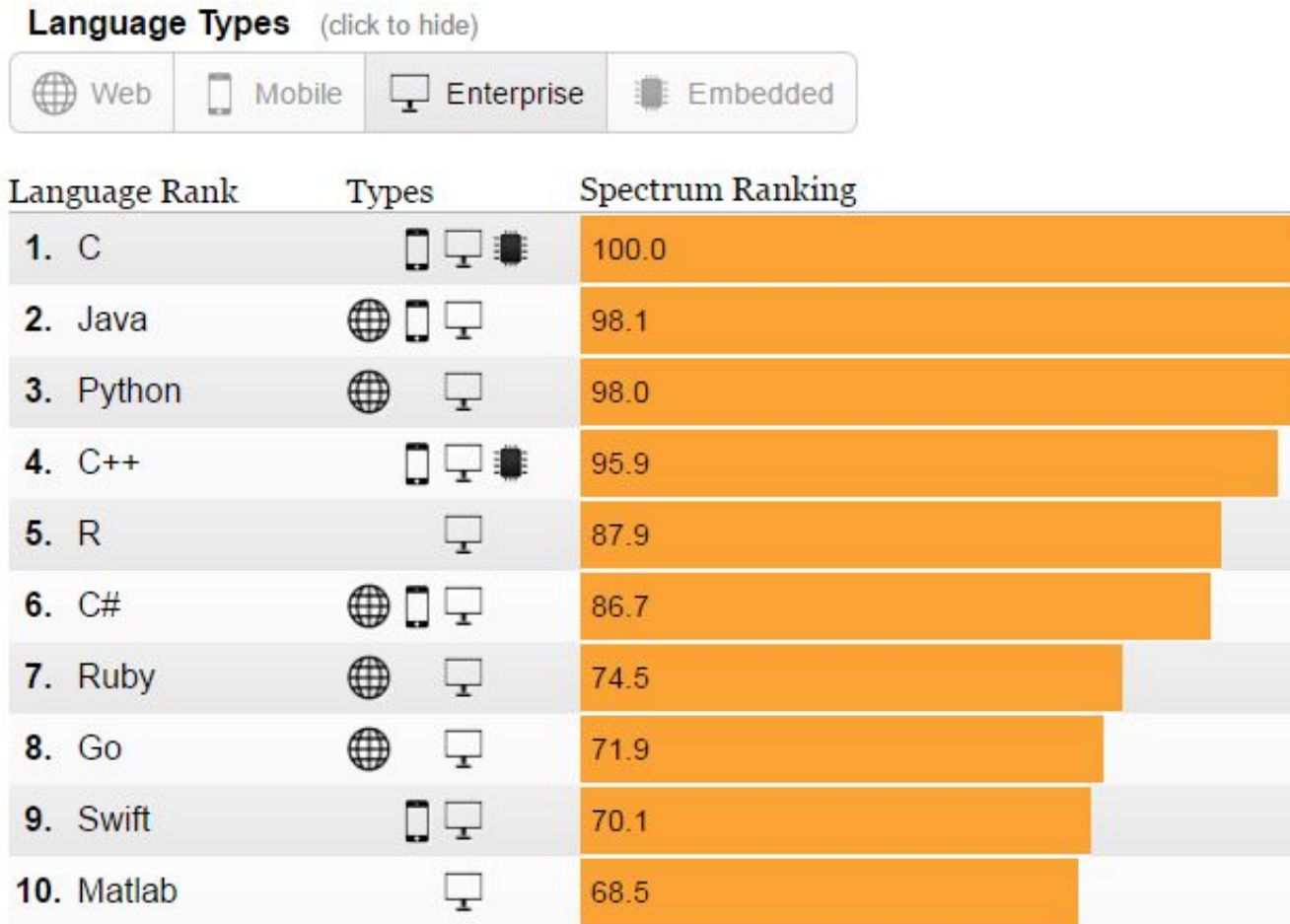
R in Finance:

Using R in a commercial and highly regulated environment

Gabriel Foix

Mirai Solutions GmbH
Tödistrasse 48
CH-8002 Zurich
www.mirai-solutions.com

- Financial industry needs
 - Talented **IT and data-analysis** experts
- R provides
 - One of the most popular **statistical programming** languages
 - Wide academic support and training possibilities
 - **Rmetrics**: <https://www.rmetrics.org/>
 - New Course Series in **DataCamp**: Applied Finance with R
 - **Coursera**: Introduction to Computational Finance and Financial Econometrics
 - Active community and conferences
 - **R/Finance 2016**: Applied Finance with R - <http://www.rinfinance.com/>
 - **R in Insurance** - <http://www.rininsurance.com/>



<http://spectrum.ieee.org/static/interactive-the-top-programming-languages-2016>

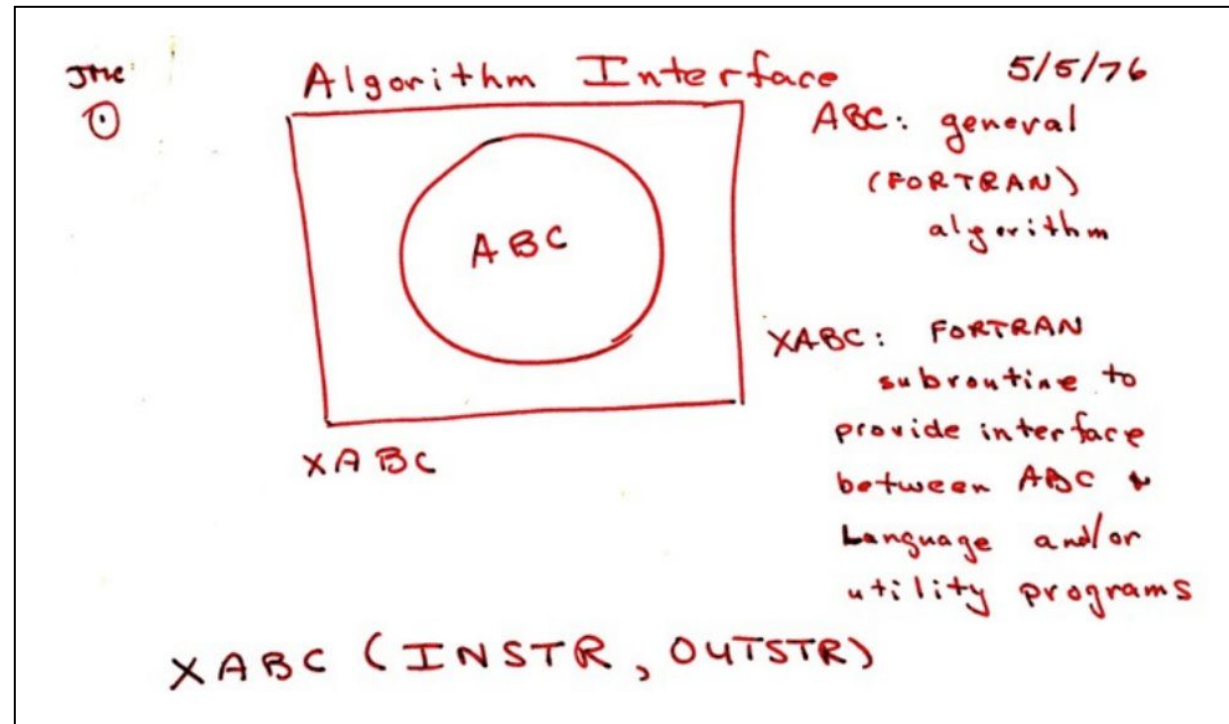
- Financial industry needs
 - Powerful and efficient **data manipulation** utilities
- R provides
 - Several advanced and intuitive packages, S3 and S4 class systems
 - `require(dplyr)`
 - `require(data.table)`
 - `require(reshape2)`, `require(plyr)`
 - `require(bigmemory)`
 - `require(R6)`
 - In-memory and file-based data-storage integration
 - `rhdf5` (Bioconductor)
 - `require(RH2)`

- Financial industry needs
 - **Domain-specific** packages
- R provides
 - Several packages for finance and insurance industry
 - require(**PerformanceAnalytics**)
 - require(**quantmod**), require(**copula**), require(**evd**)
 - require(**ChainLadder**)
 - require(**MRMR**), require(**actuar**)
 - require(**zoo**), require(**xts**), require(**seasonal**)
 - require(**lifecontingencies**)
 - require(**RQuantLib**)
 - **CRAN Task View**: Empirical Finance
 - *<https://cran.r-project.org/web/views/Finance.html>*

- Financial industry needs
 - Flexible and extensive **integration capabilities** to existing IT-environments

“The original concept for R was to create an interface language. [...] So the idea was to create a language that was really good at providing interfaces to other computing systems.”

John Chambers
Keynote speech
UseR 2014



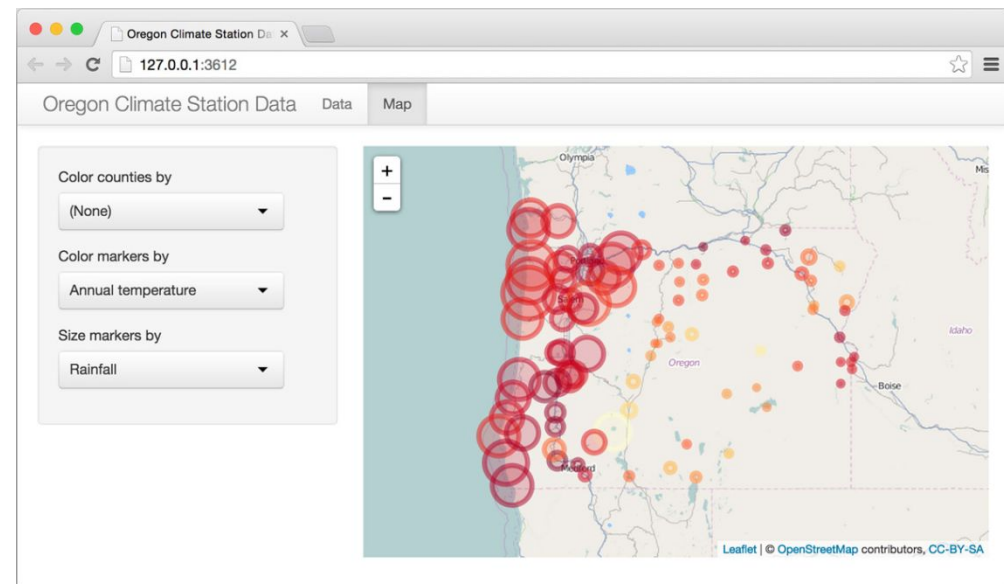
<http://datascience.la/john-chambers-user-2014-keynote/>

- Reading different data formats
 - Tabular
 - `read.table(file = "myfile")`
 - `require(XLConnect)`
 - `require(foreign)`
 - Hierarchical (XML, HTML, JSON)
 - `require(jsonlite)`
 - Relational
 - `require(DBI)`
 - `require(RJDBC)`
 - Distributed:
 - `require(SparkR)`
 - `require(sparklyr)`
- Packages connecting to financial data providers
 - `require(quantmod)` (Yahoo Finance, Oanda, Google Finance)
 - `require(Quandl)`
 - `require(Rblpapi)` (Bloomberg)
 - `require(rdatastream)` (Thomson Reuters)

- Financial industry needs
 - Computational power to handle
 - *Granular, **large** data sets / simulations*
 - *Speed to build **reactive** applications / prototypes*
- R provides
 - Running **C++** from R
 - **Rcpp** family of packages - <http://www.rcpp.org/>
 - `require(rTRNG)`
 - Profiling
 - `require(profvis)`, `require(microbenchmark)`
 - Big data interface with **Spark**
 - `require(SparkR)`
 - `require(sparklyr)`

- Financial industry needs
 - powerful and clear **data visualization**
 - **quick prototyping** of models

- R provides
 - `require(ggplot2)`
 - `require(ggvis)`
 - `require(sp)`
 - `require(Shiny)`
 - `require(manipulate)`
 - `require(plotly)`
 - `require(animation)`



www.htmlwidgets.org

- Financial industry needs
 - To meet **regulatory and audit requirements**
 - *Data traceability*
 - *Extensive testing*
 - *Documentation*
- R provides
 - `require(packrat)`
 - `require(RUnit)`
 - `require(testthat)`
 - `require(covr)`
 - `require(roxygen2)`
 - `require(rmarkdown)`
 - `require(knitr)`

As code increases...

- Source control
- Testing (unit, component, regression etc.)
- Automated build servers
- Continuous integration and deployment
- Package and version management

As teams grow...

- Development methodology (e.g. SCRUM)
- Issue-tracking and project management
- Training and on-boarding
- Software engineering - clean modular code
- Knowledge sharing – team collaboration



As computation scales...

- Shared computational servers
- C++ for low-level computation - Rcpp
- Customized execution environments
- Virtualization – infrastructure as code
- Parallelization – Spark
- Data storage, persistence

As usage diversifies...

- Integration with Excel, SAS, Python
- Reporting platforms – Shiny, Spotfire
- Consistent cross-platform experience
- Alternative development platforms: R Notebooks, Shiny

Integrating R in an Enterprise Environment

