About R

Stat 133 by Gaston Sanchez

Statistical Software

Tool

Some of you may have used statistical software with a GUI (e.g. SPSS)

Some of you may be familiar with other programming languages (e.g. Python, Java)

Some of you may have only used Excel

We are going to use **R**


> 20 yrs old

Creators of R

Ross Ihaka
(New Zealand)
PhD at Berkeley

Robert Gentleman
(Canada)
PhD at UW

Auckland University, New Zealand

Ross Ihaka and Robert Gentleman 1996 paper

R: A Language for Data Analysis and Graphics

Ross Ihaka and Robert Gentleman

In this article we discuss our experience designing and implementing a statistical computing language. In developing this new language, we sought to combine what we felt were useful features from two existing computer languages We feel that the new language provides advantages in the areas of portability, computational efficiency, memory management, and scoping.

Key Words: Computer language; Statistical computing.

https://www.stat.auckland.ac.nz/~ihaka/downloads/R-paper.pdf

R is a free implementation of the **S** language


Bell Labs, Murray Hill, New Jersey

1970s inventions at Bell Labs AT&T

More and more computer related inventions as part of the personal computer revolution

1972 Dennis Ritchie developed the compiled **programming language C**

1970s Dennis Ritchie and Ken Thompson developed the **UNIX operating system**

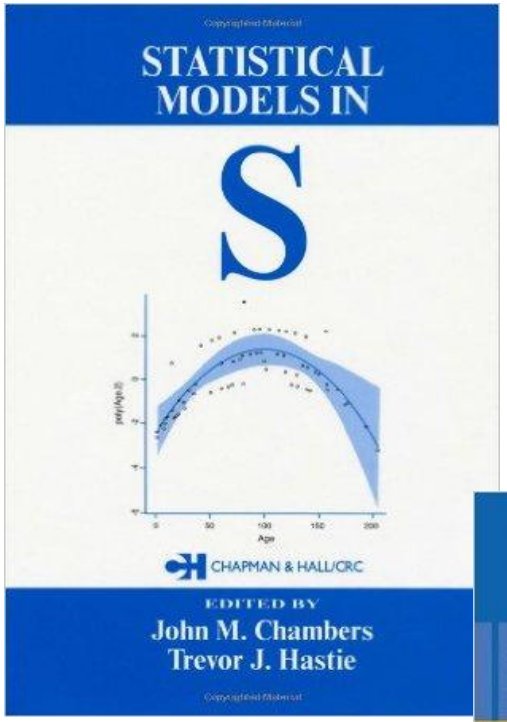
Aho, Weinberger and Kernighan developed **AWK**

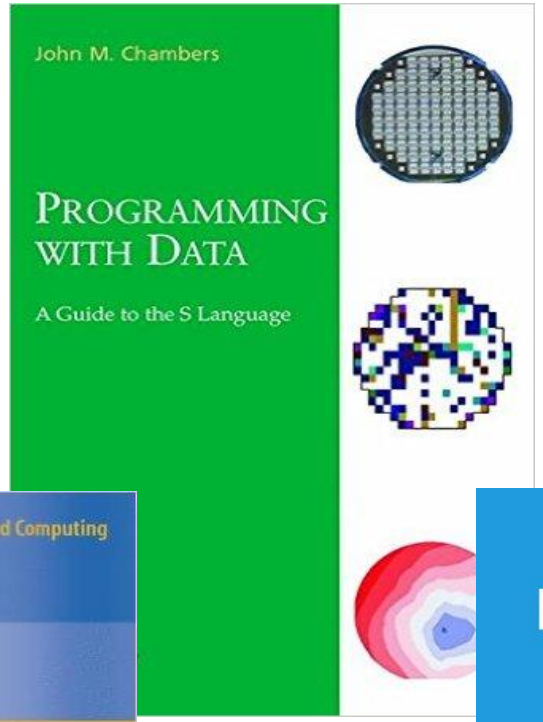
About S

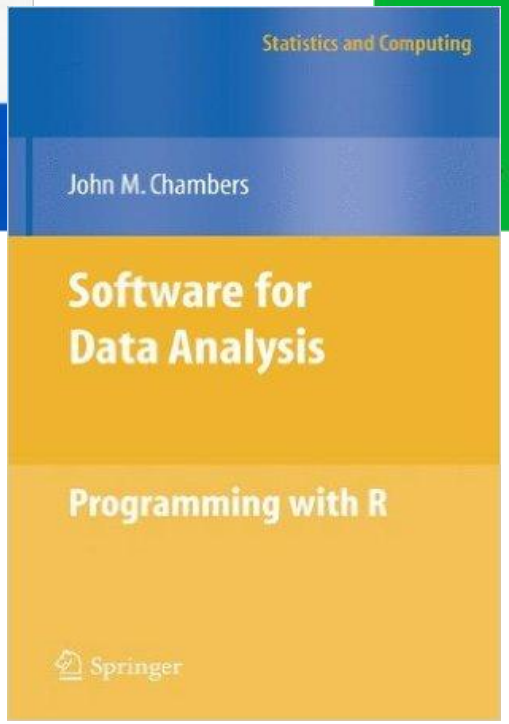
S is the statistics and graphics environment created by John Chambers and colleagues

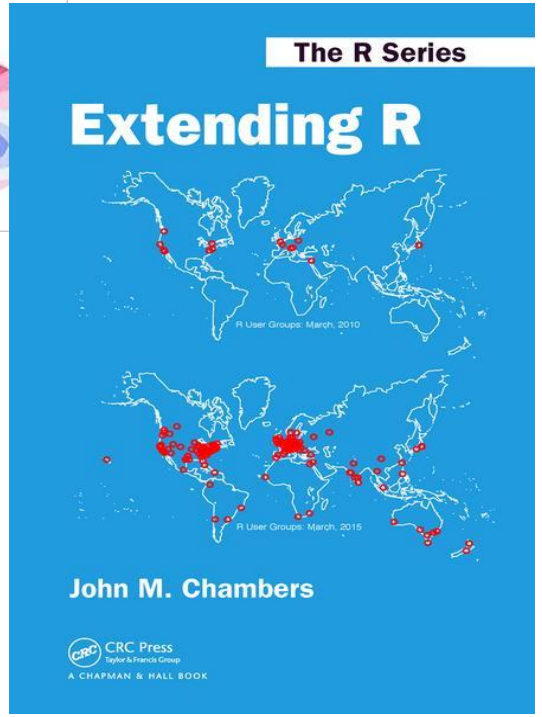
John Chambers
(main creator of S)
(PhD at Harvard)

Main books by John Chambers









History of S

May 5, 1976 at Bell Labs

Group of 5 researchers brainstormed designing a system for statistical computing

System designed to serve the needs of the statistics research group at Bell Labs

Preliminary version available by the end of 1976

History of S

No agreement on the suggested names

Although they all contained the letter “S”

Inspired by the recently designed language “C”, they reached an agreement with “S”

Portability was a main concern: it was decided to make a UNIX version of S, portable wherever UNIX was

History of S

Due to an antitrust case, AT&T was forbidden to enter the software industry

AT&T had to license any software pattents

1973 AT&T licensed UNIX to educational institutions

It also started to license S to universities and research laboratories worldwide

History of S

Building of S reflected many UNIX features

1980s saw a redesigning of S

1990s Ross Ihaka and Robert Gentleman (RR) designed a system compatible with S

RR joined the open-source movement

Self-managing group of volunteers took over the development of R: **“R Core group”**

Beginnings of R

August 1993: Binary copies of R shared on the s-news mailing list

Martin Maechler (ETH Zurich) encouraged *R&R* to release R source code as free software

June 1995: R released under GNU general license

Kurt Hornik of TU Wien established the main archive (CRAN) in Austria

mid-1997: R Core Group (developers) was established

Philosophy of S

S was designed as an environment for statistics and graphics

S was designed to blur the distinction between users and programmers

S was designed as a system for **interactive** data analysis

About R

R also follows the idea of **interactive** data analysis

Interactive: as having a dialogue with the computer

You type one or more commands, execute them, and get the results

i.e. ask questions, get answers

R

We will use **R** as our main computational-analytical tool for this course

www.r-project.org

https://www.r-project.org

The R Project for Statistical Computing

[Home]

Download

CRAN

R Project

About R
Logo
Contributors
What's New?
Reporting Bugs
Development Site
Conferences
Search

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To **download R**, please choose your preferred CRAN mirror.

If you have questions about R like how to download and install the software, or what the license terms are, please read our answers to frequently asked questions before you send an email.

News

- The **useR! 2017** conference will take place in Brussels, July 4 - 7, 2017, and details will be appear here in due course.
- **R version 3.3.1 (Bug in Your Hair)** has been released on Tuesday 2016-06-21.

R entails both:

Environment for Statistical Computing

Programming Language

R Foundation and Core Group

R Foundation for Statistical Computing: Vienna-based non-profit organization that oversees the R Project

R core group: group of developers with commit access to the R codebase

R contributors

The current R is the result of a collaborative effort with contributions from all over the world. R was initially written by Robert Gentleman and Ross Ihaka—also known as "R & R" of the Statistics Department of the University of Auckland. Since mid-1997 there has been a core group with write access to the R source

https://www.r-project.org/contributors.html

Why R?

Why R?

Allows custom analysis

High-level scripting language

Statistical programming language

Interactive exploratory data analysis

Why R?

Easy to replicate analysis

Sound numerical methods

Large community of contributors

Open Source

It’s Free

Why R?

As the Spanish say:

- Bueno
- Bonito
- Barato

(good, beautiful, and inexpensive)

Learning R

Learning a programming language

Chinese proverb:

- I hear and I forget
- I see and I remember
- I do and I understand

Learning a programming language

Learning R (or any other programming language)

You'll get frustrated

It takes time to become fluent

Lots of trials and errors

Be patient

Practice, practice, practice

More resources

R website

https://www.r-project.org

Technical manuals

https://cran.r-project.org/manuals.html

Contributed documentation

https://cran.r-project.org/other-docs.html

More resources

Tasks Views

https://cran.r-project.org/web/views

R journal

https://journal.r-project.org

More resources

Stackoverflow R questions

http://stackoverflow.com/questions/tagged/r

R bloggers

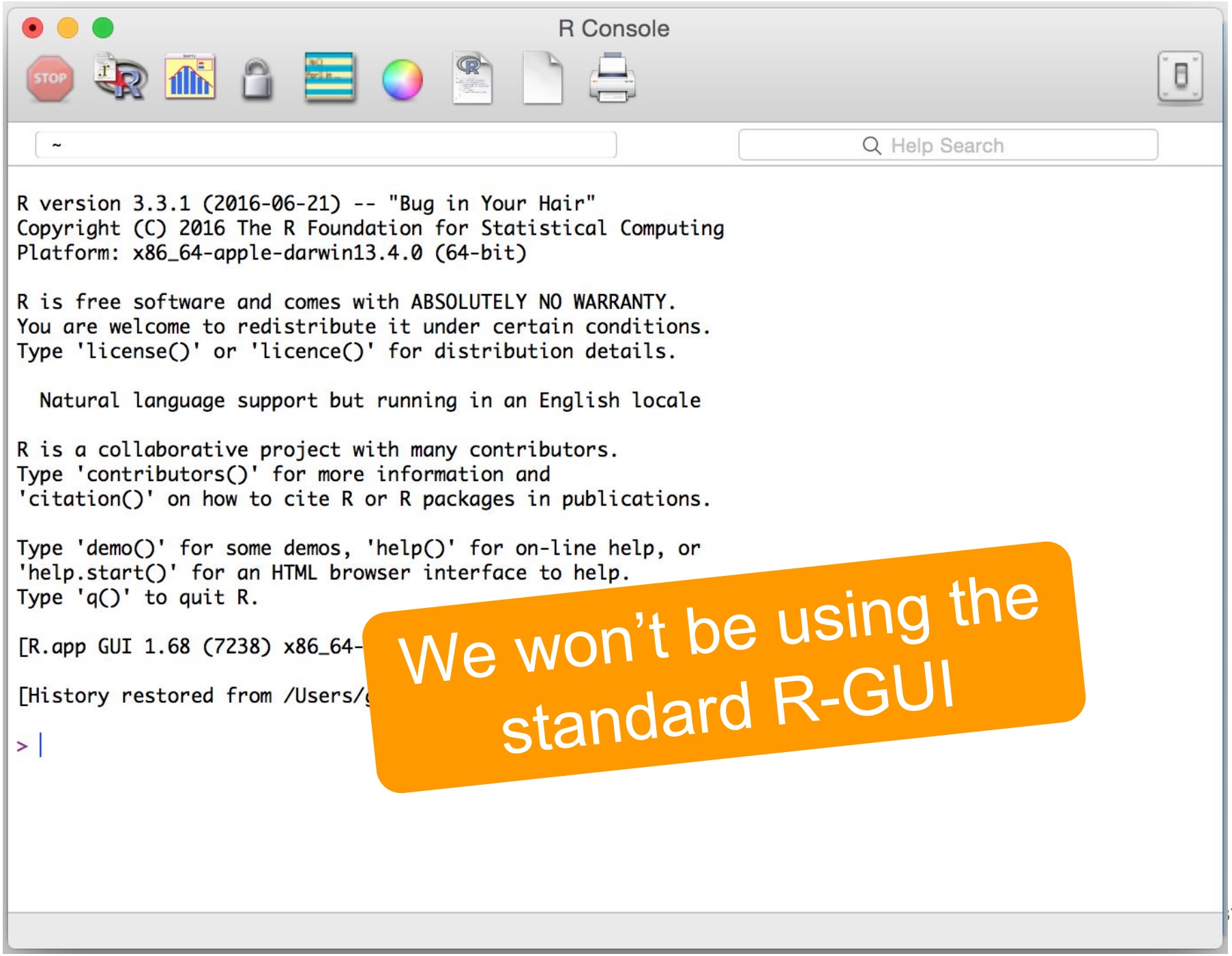
https://www.r-bloggers.com/

Hundreds of blogs

Many youtube videos

Thousands of online tutorials

RStudio


R Console
STOP
~
Help Search
R version 3.3.1 (2016-06-21) -- "Bug in Your Hair"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.4.0 (64-bit)
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
Natural language support but running in an English locale
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
[R.app GUI 1.68 (7238) x86_64-
[History restored from /Users/
>
We won't be using the standard R-GUI

RStudio

Integrated Development Environment (IDE)

RStudio

Learn how to customize RStudio

Learn keyboard shortcuts

Take a look at the menu options

Get familiar with the pane layout