

Getting starting with R

Alexandre Courtiol

Leibniz Institute of Zoo and Wildlife Research

June 2018

Who am I?

- evolutionary biologist / statistician
- studies in France (Montpellier), postdoc in the UK (Sheffield)
- senior researcher at Leibniz IZW / Lecturer at Freie University
- experience with **R**:
 - 2003: starting to study **R** (still ongoing)
 - 2008: starting to use **R** most days
 - 2010: starting to teach **R**
 - 2016: starting to develop **R** packages

1 What is R?

Getting started with R

1 What is R?

What is R?

R is a programming language and software environment for statistical computing & graphics.

Key points about **R**:

- free for all
- open source (explore: <https://github.com/wch/r-source>)

What is R?

R is a programming language and software environment for statistical computing & graphics.

Key points about **R**:

- free for all
- open source (explore: <https://github.com/wch/r-source>)
- polyvalent:
 - from laptop to most advanced supercomputers
 - local or remote
 - Windows, MacOS, linux or many other Unix-based systems

What is R?

R is a programming language and software environment for statistical computing & graphics.

Key points about **R**:

- free for all
- open source (explore: <https://github.com/wch/r-source>)
- polyvalent:
 - from laptop to most advanced supercomputers
 - local or remote
 - Windows, MacOS, linux or many other Unix-based systems
- rich (12542 packages just on CRAN)
- cutting edge (check updates for today: <http://dirk.eddelbuettel.com/cranberries/cran/updated/>)

What is R?

R is a programming language and software environment for statistical computing & graphics.

Key points about **R**:

- free for all
- open source (explore: <https://github.com/wch/r-source>)
- polyvalent:
 - from laptop to most advanced supercomputers
 - local or remote
 - Windows, MacOS, linux or many other Unix-based systems
- rich (12542 packages just on CRAN)
- cutting edge (check updates for today: <http://dirk.eddelbuettel.com/cranberries/cran/updated/>)
- used by millions
- **R** is the best software environment for statistical computing, but it is far from perfect!

Is R good for you?

Good for:

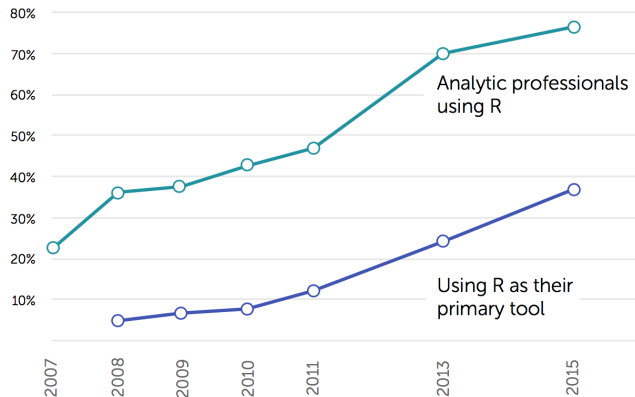
- data manipulation
- plots, including GIS
- analysing big (incl. seq.), medium and small data
- simulating data

Not optimal for:

- beginners
- data entry
- formal algebra

Who uses R?

RISE OF R USAGE

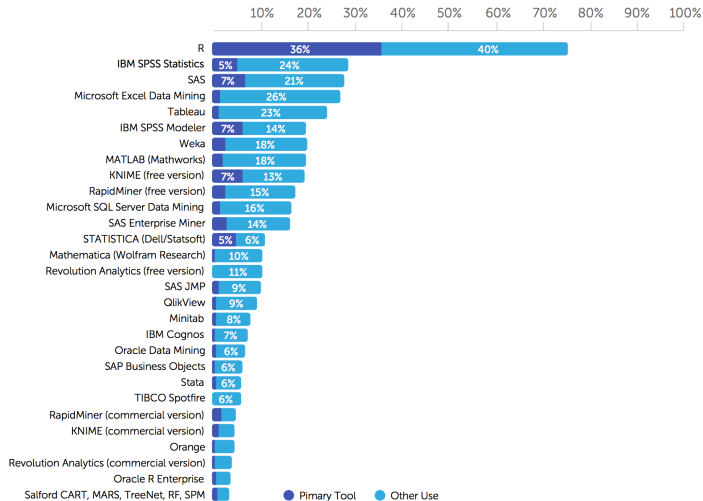


Rexer Analytics

[1220 analytic professionals from 72 countries participated in this survey]

What else?

TOOL USE



A short history of R

- early 1980s: **S** version 1: genesis of an interactive statistical system (Becker & Chambers at Bell Labs)
- mid 1980s: **S** version 2: portable version (thanks to UNIX)
- 1988: **S** version 3 (**S3**): object-based, C-based (very much like R)
- 1991: **S3** includes a large statistical modeling toolbox
- 1993: the replication of **S** as the **R** project starts (Ihaka & Gentleman at University of Auckland)
- 29/02/2000: **R** version 1.0.0 (R core)

Due to the ownership of **S** by a phone company, **S** could not be free, nor open source. Hence **R**!