



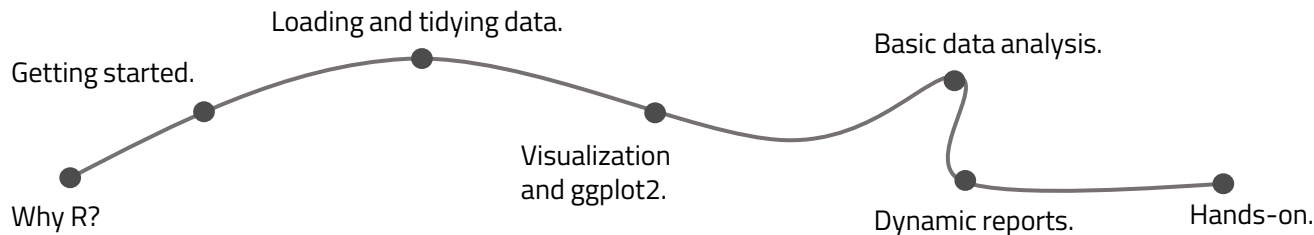
summeRworkshop

Thursday, Jul 25, 2019
10:00 - 16:00
UL-FRI, Lecture room 03



Data
Science
WORKSHOP

University of Ljubljana
Faculty of Computer and
Information Science



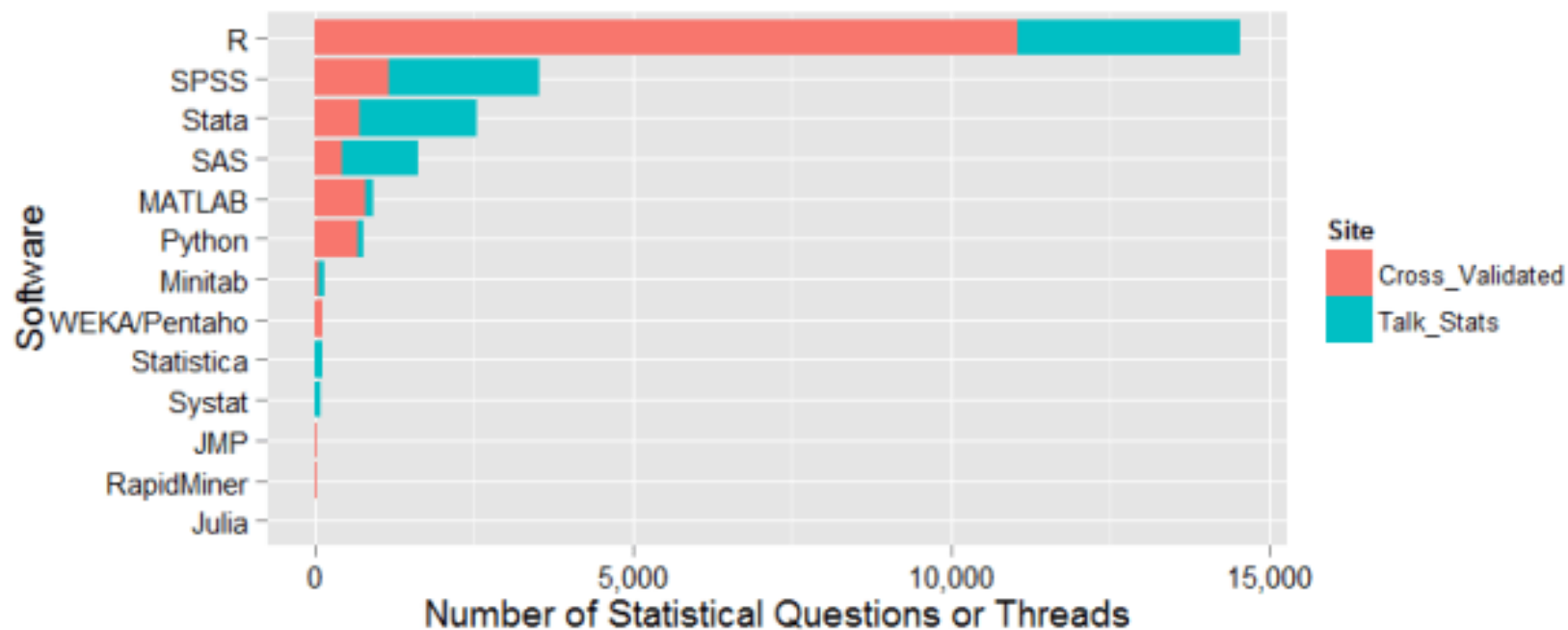
Gregor Pirš
Jure Demšar
Erik Štrumbelj

What is R?

- R is software for statistical computing and graphics.
- R is an interpreted and dynamically typed programming language.

Advantages of R

- It's free.
- Tailored to data manipulation and analysis.
- Large community.
- Extensive 3rd-party libraries (packages).
- Integration (MSSQL, SAS, Tableau, LaTeX...).



<https://r4stats.com/2015/10/19/using-discussion-forum-activity-to-estimate-analytics-software-market-share/>

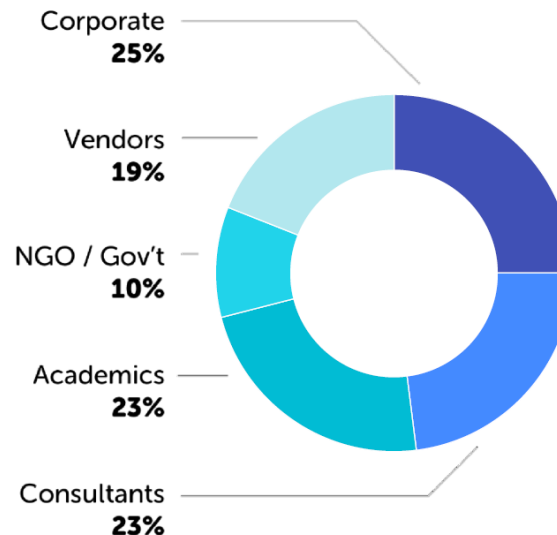
Disadvantages of R

- Relatively slow.
- Clunky object oriented programming.
- Less than ideal memory management.
- Bottom line: Not something that one would use for production-level software development.

Should I learn R?

- A professional in the field of data analysis should be fluent in R.
- If analyzing data is important to your work and you'll learn only 1 programming language in your life, learn R.
- If you already know other programming languages or work in less-statistical areas (machine learning, engineering), R might not be the best first choice.

- 8th survey since 2007
- 67 questions
- 10,000+ invitations emailed & promoted by newsgroups, vendors and bloggers
- Respondents: 1,123 analytic professionals from 91 countries
- Data collected in first half of 2017

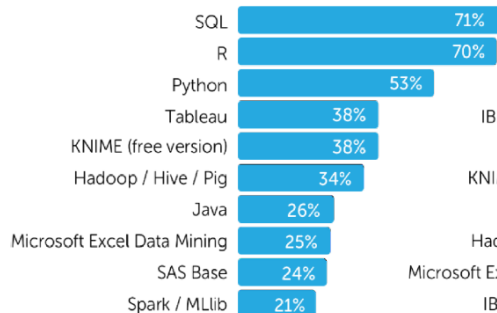


Most Data Scientists use Multiple Tools

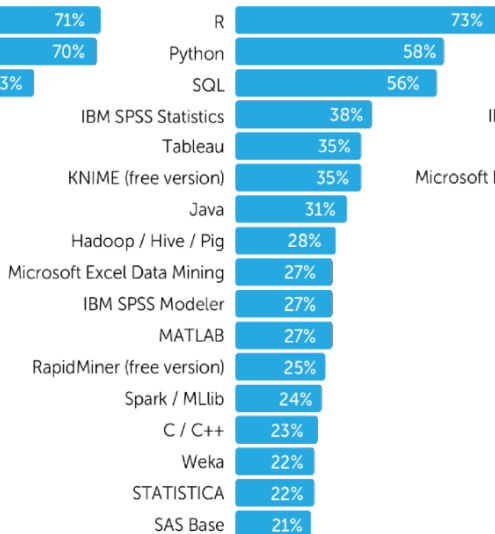


What data science / analytic tools, technologies, and languages did you use in the past year?

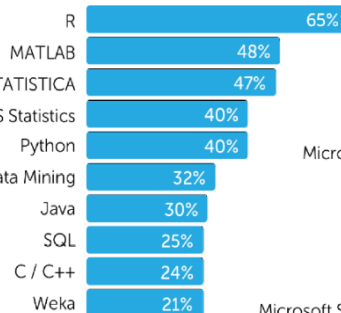
Corporate



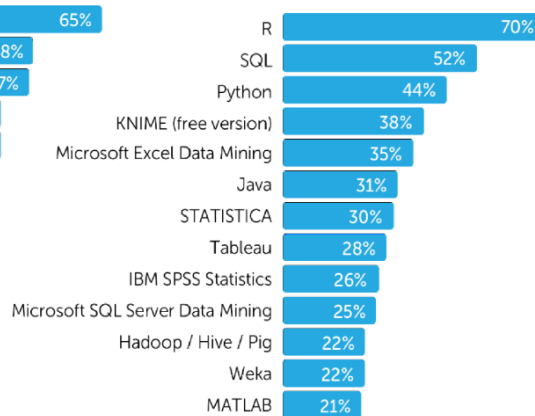
Consultants



Academics



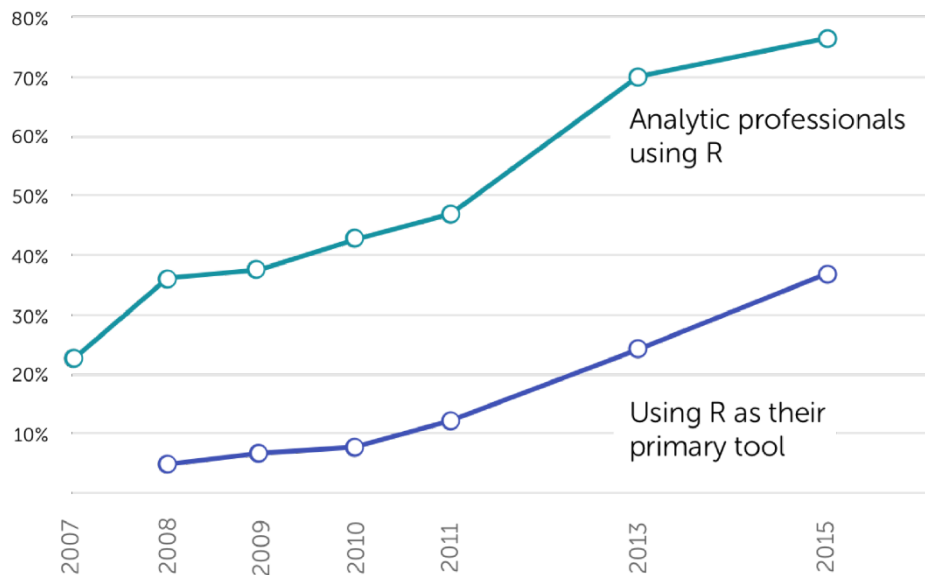
NGO / Gov't



A Decade of Dramatic Growth in the Use of R



The biggest change in tool adoption we've seen over the past decade has been the dramatic growth in the use of R



R in the job market

