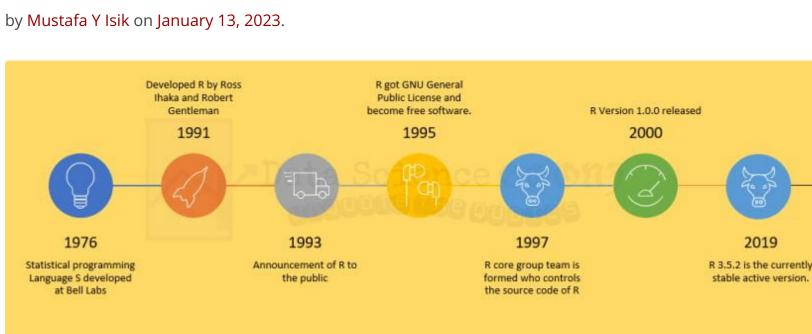
SB You Log In

Mustafa Isik CSE301 – History of R

HOME



Wait, what is R?

> total <- 10

ware."

It is widely used among statisticians and data scientists for developing statistical software and data analysis. "Who are my ancestors?" ~ What was there before R

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

The development of R can be traced back to the early 1970s, when John Chambers, Rick Becker, Trevor Hastie, Allan Wilks and others at the Bell Labs, began work on a statistical software package called S, which was developed as a tool for data analysis and statistical modeling, and it quickly gained popularity among statisticians and researchers as it became publicly available in the early 1980's. One thing about S is that it was proprietary software, which meant that it was only available to those who could afford to purchase a license, at a time when most couldn't afford to do so.

```
> make.counter <-</pre>
        function(total = 0)
          function() {
            total <<- total + 1
           total
   > counter <- make.counter()
   > counter()
    [1] 1
   > counter()
    [1] 2
   > counter()
    [1] 3
     Figure 1: A simple function demonstrating how the scoping rules in R differ from those of S.
"Well then, who is my father" ~ How did R Come to Exist
```

In the early 1990s, Ross Ihaka and Robert Gentleman, both computer science professors at the University of Auckland, decided to create an open-source version of S. Like many other cre-

ations, it was created based on a need; both Ross and Robert were having problems in their teaching labs where they used Macintosh computers and as they started working on the project, it started to turn their situation into a mission-and-vision type of statement where they shifted to make a statistical software package that was available to everyone, regardless of their financial means. The duo began work on R in August 1993, and the first version of the software was released in 1995. Ross Ihaka retells the story of the early days in one his articles: "The initial work on R by Robert Gentleman and I produced what looked like a potentially useful piece of software and we began preparing it for use in our teaching laboratory. We were heart-

ened enough by our progress to place some binary copies of R at Statlib and make a small announcement on the s-news mailing list in August of 1993. A number of people picked up our binaries and offered feedback. The most persistent of these was Martin Machler of ETH Zurich, who encouraged us to release the R source code as free soft-

Ross Ihaka

Robert Gentleman



was designed to be a high-level programming language, with a syntax similar to that of S. It also

included many of the same statistical and graphical functions as S, making it easy for users of S

to transition to R. The visualization and implementation quality would get even better in the fu-

ture thanks to certain GUI's being developed for R. "Richard Stallman's GNU would like a word" ~ Creating and Embracing the Community with Open Source In 1995, R was officially recognized as a GNU project, which meant that the source code for the

software was made freely available under the GNU General Public License. This could be seen as the reason behind its adoption among academics, researchers, and data scientists through-

out the years. R's open-source nature made it easier to distribute and modify. This meant that anyone can use R for any purpose, including commercial projects. Additionally, the open-source becomes an advantage of R since it allows for a large and active community of developers who continuously improve the software and add new features. In 1997, the Comprehensive R Archive Network (CRAN) was founded by Kurt Hornik and Fritz Leisch. The purpose of CRAN was to host built-in and ready-to-execute files, user-made packages, and their respective documentations. The creation and diverse spreading of CRAN would

#20

"If it isn't the 21st Century" ~ Getting on the Road for New Journeys

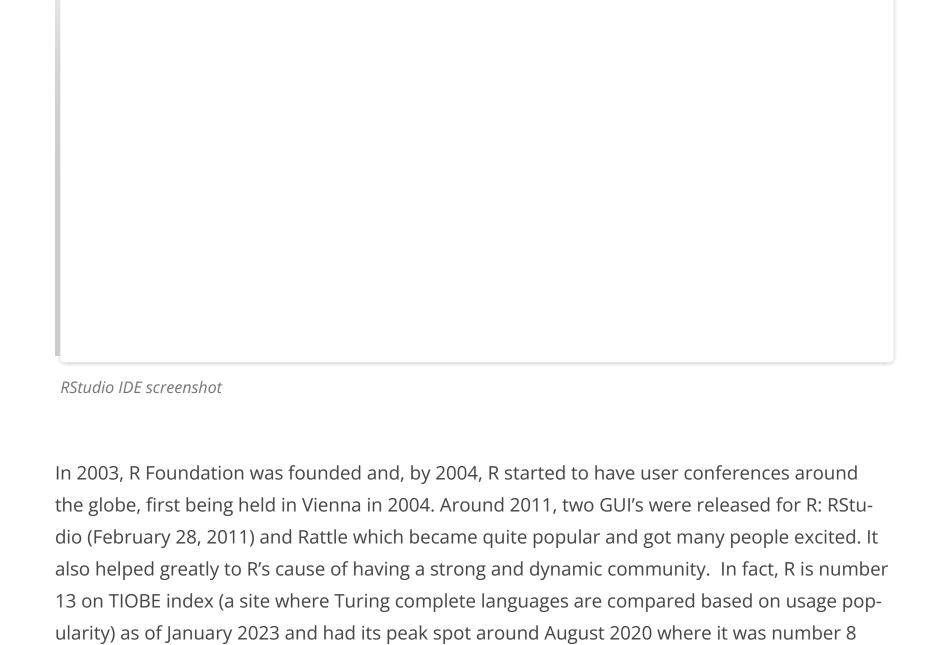
later go onto be a huge success.

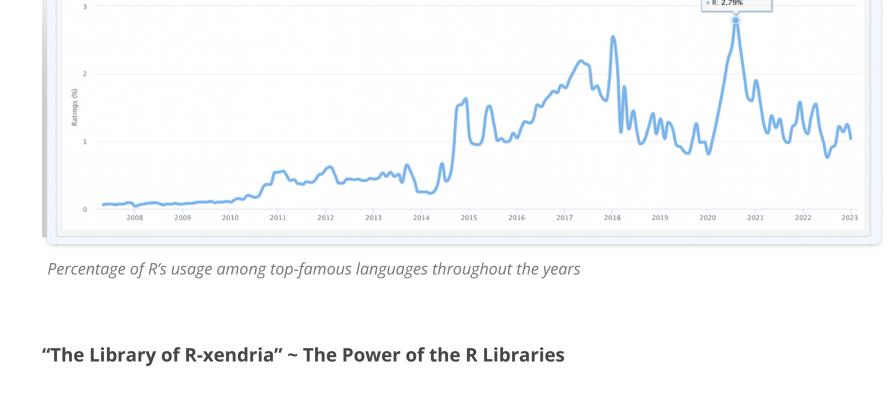


eling. However, as the popularity of R grew, it began to be adopted by a larger scale of users,

gramming languages -especially for data analysis- and is heavily used in a variety of industries, including finance, healthcare, and marketing. "Where is my GUI's, who took my IDE's" ~ How the Graphical User Interfaces Improved R and Kept it Popular to This Day

including data scientists, researchers, and businesses. Today, R is one of the most popular pro-





One of the main reasons for R's popularity has been and continues to be so is its extensive li-

brary of packages. R has a large and active community of developers who have created a wide

range of packages for various tasks, such as data visualization, machine learning, and natural

language processing. These packages are freely available and can be easily installed and used in

TIOBE Index for R

R. This makes it easy for users to perform complex tasks without having to write their own code or even go onto make customizations and additions on top of the existing ones. Number of R packages ever published on CRAN

"This data is enormous!" ~ Big Data and R

8000 -

keting,

this quote:

among the most popular programming languages.

★ Highest Position (since 2007): #8 in Aug 2020

¥ Lowest Position (since 2007): #73 in Dec 2008

6000 -5000 4000 -1000 -

"Am I being replaced?" ~ The Competitive Market of Data Analytics and R's Place in it R also had current competitors of SAS and Python. Their large communities and usage cases

make them potential substitutes for each other depending on a person/company's work prefer-

ences and style choices. The big difference would be that R is free (unlike SAS where you have

to pay up); R and its community focus fully on statistics/data (unlike Python where community

doesn't necessarily do the same since Python is a general-purpose language) as can be seen in

In recent years, R has also been used in big data and data science projects. With the increasing

availability and usage expansion of big data, R has become an important tool for data scientists

who need to analyze large amounts of data. R's extensive library of packages for data analysis,

visualization, and its ability to handle big data integration with platforms like Hadoop make it an

ideal tool for such projects. This helped R establish itself in industries of finance, retail, and mar-

"Another point, which I repeatedly make to students, is that R is free and will continue to exist. Nothing can make it go away. Once you learn it, you are no longer subject to price increases (e.g., from zero, when, as a grad student, you use your advisor's copy of SAS, to several hundred dollars or more after you leave). You can take it with you wherever you go. The investment in learning thus has a long-term payoff" (Baron, 2006).

Fundamental comparison between qualities of Python and R

environment for statistical computing and graphics. The development was started by Ross Ihaka and Robert Gentleman in the late 20th century and later turned into a venture for an opensource alternative to the proprietary software S. Since then, R kept on growing in popularity, particularly among statisticians, data scientists and lots of different industries of businesses around the world. Its large library of packages and open-source nature make it a powerful tool for data analysis and visualization, and its popularity continues to grow in academia and various business environments despite potential competitors. **Works Cited**

-ChatGPT was used substantially in the making of this post thanks to Prof. McKenna who al-

CHATGPT: Optimizing Language Models for Dialogue – Openai.com. openai.com/blog/chatgpt/.

To sum up our discussion, R is a powerful and versatile programming language and software

-"The R Project for Statistical Computing." R, www.r-project.org/. -Burns, Patrick (27 February 2007). "R Relative to Statistical Packages: Comment 1 on Technical Re-

Conclusion

port Number 1 (Version 1.0) Strategically using General Purpose Statistics Packages: A Look at Stata, SAS and SPSS -"Tiobe Index." TIOBE, 16 January 2023, www.tiobe.com/tiobe-index.

lowed us to turn this assignment into more of an experiment.

-Smith, David. 16 Years of R History, 16 Mar. 2016, blog.revolutionanalytics.com/2016/03/16years-of-r-history.html.

-Kumar, Mohair. R Overview and History, 12 Aug. 2020, medium.com/@ArtisOne/r-overview-and-

-Today Is the 20th Anniversary of the Release of R 1.0.0., Twitter, 29 Feb. 2020, twitter.com/_r_foun-

Provided by Stony Brook University

history-75ecb036d0df. -Python vs R, The Basics, towardsdatascience.com/python-vs-r-the-basics-d754c45c1596.

dation/status/1233671896144793600. -lhaka, Ross. *R : Past and Future History A Draft of a Paper for Interface '98*, Statistics Department

The University of Auckland Auckland, New Zealand, www.stat.auckland.ac.nz/~ihaka/downloads/Interface98.pdf.

Search **Recent Comments** No comments to show.

Search

Q