Data Science in Biomedicine Series 1

Introduction to R programming languages

History/Origin of R

R is a programming language and free software developed by Ross Ihaka and Robert Gentleman in 1993. R possesses an extensive catalog of statistical and graphical methods. It includes machine learning algorithms, linear regression, time series, statistical inference to name a few. Most of the R libraries are written in R, but for heavy computational tasks, C, C++ and Fortran codes are preferred. R is not only entrusted by academics, but many large companies also use R programming language, including Uber, Google, Airbnb, Facebook and so on.

What is R used for?

1. Data Science

2. R is Good for Business

3. R is a gateway to Lucrative Career

4. Open-source

5. Popularity

6. Robust Visualization Library

7. With R, you can develop amazing Web-Apps

8. R enjoys a vast Community Support

9. A go-to language for Statistics and Data Science

10. R is being used in almost every industry

Why use R as a Bioinformatician?

R is rapidly becoming the most important scripting language for both experimental- and computational biologists. It is well designed, efficient, widely adopted and has a very large base of contributors who add new functionality for all modern aspects of data analysis and visualization. Moreover it is free and open source. However, R's great power and expressivity can at first be difficult to approach without guidance, especially for those who are new to programming.

Is R difficult?

Introducing R and RStudio IDE

Fundamental operations and concepts in R

Some general purpose functions in R

Packages in R

Objects in R

Genomic Data Visualization in R

Assignments/Assessments

References

https://data-flair.training/blogs/why-learn-r/

https://a-little-book-of-r-for-bioinformatics.readthedocs.io/en/latest/

https://bioinformatics.ca/workshops/2016-introduction-r/