





Prof. Dr. Jan Kirenz

R

R Programming Language

R is a free software environment for statistical computing and graphics. It is widely used among statisticians and data scientists for developing statistical software and data analysis. In R, you can install and use packages, which are collections of functions, data, and documentation that extends the capabilities of base R.

 **RStudio** is a free and open-source integrated development environment (IDE) for R and makes R easier to use. It includes a code editor, debugging & visualization tools

 **Jupyter Notebook:** One important third-party tool for data science is the Jupyter Notebook, an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text (Perkel, 2018). To learn more about this topic, read this article about Jupyter and R.

Recommended reading and resources:

- Perkel, J. M. (2018). Why Jupyter is data scientists' computational notebook of choice. Nature, 563(7729), p. 145.
- Wickham, H., & Grolemund, G. (2016). R for data science: import, tidy, transform, visualize, and model data. O'Reilly Media, Inc.
- Quick-R: a quick online reference for data input, basic statistics and plots



The first step is to install R. You can download and install R from the Comprehensive R Archive Network (CRAN).



Next, go to RStudio Download. Click the Download RStudio Desktop button. Select the installation file for your system. Run the installation file.

Data Science Concepts, Ideas, and Codes

- Towards Data Science
- Kaggle
- Stack Overflow

Newsletter:

- Data Science Weekly
- Data Elixir
- Data Machina

Datasets:

- Awesome Datasets