

Hướng dẫn mở đầu về R

Khoa Toán - Cơ - Tin học
Trường Đại học Khoa học Tự nhiên
Đại học Quốc gia Hà Nội

Ngày 5 tháng 9 năm 2023

Nội dung

- 1 Giới thiệu về phần mềm R
- 2 Hướng dẫn cài đặt R
- 3 Một số IDE cho R
- 4 Hướng dẫn cài đặt RStudio

- 1 Giới thiệu về phần mềm R
- 2 Hướng dẫn cài đặt R
- 3 Một số IDE cho R
- 4 Hướng dẫn cài đặt RStudio

Giới thiệu về phần mềm R

- R là một ngôn ngữ dành cho tính toán và đồ họa thống kê.
- R là một phần của dự án GNU Project
- R có nguồn gốc từ S, được phát triển tại Bell Laboratories bởi John Chambers và các cộng sự.
- R là phần mềm miễn phí, chạy được trên cả Windows, Linux, MacOS

Nội dung

- 1 Giới thiệu về phần mềm R
- 2 Hướng dẫn cài đặt R**
- 3 Một số IDE cho R
- 4 Hướng dẫn cài đặt RStudio

Hướng dẫn cài đặt R

- Truy cập vào trang web <https://www.r-project.org/>.
- Vào mục Download, chọn CRAN, sau đó chọn <https://cloud.r-project.org/>.
- Tùy vào hệ điều hành đang sử dụng, chúng ta chọn link thích hợp để tải về R và cài đặt

Nội dung

- 1 Giới thiệu về phần mềm R
- 2 Hướng dẫn cài đặt R
- 3 Một số IDE cho R**
- 4 Hướng dẫn cài đặt RStudio

Một số IDE cho R

IDE là viết tắt của cụm từ Integrated Development Environment, có nghĩa là môi trường phát triển tích hợp

IDE giúp cho việc viết code và xây dựng các chương trình trở nên nhanh chóng, dễ dàng hơn. Một số tính năng phổ biến của IDE bao gồm:

- Một trình soạn thảo code với các tính năng: Syntax highlighting, Code completion, ...
- Một trình biên dịch/thông dịch, giúp việc chạy code đơn giản hơn.
- Một trình gỡ lỗi (debugger) để sửa lỗi do quá trình code
- Một trình phân tích cú pháp (linter) để giúp nhanh chóng phát hiện các lỗi sai về cú pháp, phong cách lập trình, ...

Một số IDE cho R

Có rất nhiều IDE hỗ trợ ngôn ngữ R, dưới đây là một vài lựa chọn phổ biến:

- RStudio: IDE dành riêng cho R
- PyCharm/IntelliJ IDEA và plugin R Language for IntelliJ: Đẹp, rất nhiều tính năng, nhưng khá nặng.
- Datalore: IDE dùng trực tuyến, không cần cài đặt cho R, Python.

Nội dung

- 1 Giới thiệu về phần mềm R
- 2 Hướng dẫn cài đặt R
- 3 Một số IDE cho R
- 4 Hướng dẫn cài đặt RStudio**

Hướng dẫn cài đặt RStudio

- Vào đường link: <https://www.rstudio.com/products/rstudio/>
- Chọn RStudio Desktop và Download RStudio Desktop
- Tùy thuộc vào hệ điều hành, ta lựa chọn tải về phiên bản tương ứng

Cài đặt gói lệnh (Packages)

- RStudio được mở rộng một cách mạnh mẽ nhờ việc sử dụng các gói (packages) do người sử dụng cung cấp cho mỗi chức năng cụ thể hoặc mỗi lĩnh vực nghiên cứu cụ thể.
- Ví dụ:

Tên gói	Chức năng	Gói phụ thuộc
foreign	Nhập dữ liệu từ các phần mềm khác như: SPSS, SAS, STATA,...	stats (gói cơ bản)
qcc	Vẽ một số biểu đồ kiểm soát chất lượng (quality control charts), trong đó có biểu đồ pareto	mass
e1071	Tính một số hàm trong thống kê, chẳng hạn tính skewness và kurtosis	class
prob	Cung cấp những hàm tính xác suất trên không gian mẫu hữu hạn	
distriEx	Cung cấp những hàm tính xác suất và tính những đặc trưng quan trọng của một biến ngẫu nhiên	methods, distr, evd, actuar, startupmsg
distr	Cung cấp những hàm tính xác suất và tính những đặc trưng quan trọng của một số biến ngẫu nhiên thường gặp	methods, graphics, startupmsg, sfsmisc, SweaveListingUtils
BSDA	Cung cấp thêm nhiều hàm thống kê, đặc biệt là hàm cho dữ liệu thứ cấp	e1071, lattice
lmtest	Cung cấp những hàm kiểm định mô hình hồi qui tuyến tính	R(> 2.10.0), stats, zoo

Cài đặt gói lệnh (Packages)

Một gói mới có thể cài từ gói được tải về máy hoặc cài trực tiếp từ trên mạng. Các bước thực hiện chi tiết cho từng cách cài đặt này được cho như sau:

Cài đặt gói mới từ gói được tải về máy

- Tại trang chủ CRAN của R, chọn mục Packages trong phần tiêu đề.
- Chọn sắp xếp các gói theo tên Table of available packages, sorted by name để thuận tiện cho việc tìm kiếm.
- Chọn đến gói đang cần để tải về máy.
- Ví dụ, bạn tải file **BSDA_1.2.1.tar.gz** và các file của gói lệnh phụ thuộc về và lưu tại thư mục **E:/Teaching/Introduction to data analysis/Lectures 2023-2024/Cài đặt**.

Cài đặt gói mới từ gói được tải về máy

Gói BSDA có thể được cài theo một trong hai cách sau:

- Tại cửa sổ dòng lệnh của R, gõ lệnh:

```
>setwd("E:/Teaching/Introduction to data analysis/Lectures  
2023-2024/Cài đặt")  
>install.packages("BSDA_1.2.1.tar.gz")
```

Sau đó `install.package` các gói lệnh phụ thuộc.
- Hoặc tại cửa sổ góc thứ tư trong RStudio, trên thanh công cụ chọn **Packages** → **Install** → **Browse** và chọn thư mục **E:/Teaching/Introduction to data analysis/Lectures 2023-2024/Cài đặt** và tệp **BSDA_1.2.1.tar.gz** và các tệp của gói phụ thuộc để cài.

Cài đặt gói mới từ gói được tải về máy

Khi một gói phụ thuộc vào nhiều gói khác, để cài được gói này ta phải tải hết các gói phụ thuộc và cài đặt các gói này trước. Ví dụ để cài được gói BSDA, ta phải cài hai gói e1071 và lattice. Do đó việc cài đặt một gói từ những gói đã tải về thường mất nhiều thời gian, nhất là khi cài những gói phụ thuộc vào nhiều gói khác. Cài đặt một gói mới trực tiếp từ trên mạng trong nhiều trường hợp thường nhanh và thuận tiện hơn.

Cài đặt một gói mới trực tiếp từ trên mạng

Muốn cài đặt một gói mới, ví dụ gói BSDA theo cách này, ta có thể thực hiện theo cách sau:

- Cài bằng hàm `install.packages`: Tại cửa sổ dòng lệnh của R, gõ lệnh `> install.packages("BSDA")`
- Chọn một bản sao (mirror) của CRAN, chẳng hạn VietNam và R sẽ tự động cài BSDA cùng với tất cả các gói mà BSDA phụ thuộc vào.

Chú ý: Trước khi sử dụng những hàm trong một gói mới, ta phải gọi gói này ra qua hàm **library**, chẳng hạn **>library(BSDA)** hoặc vào **Packages** trong Rstudio và chọn đến gói cần dùng.

- Để tìm hiểu cách dùng cũng như các tham số bên trong một hàm, ta dùng hàm `help`. Chẳng hạn để tìm hiểu về hàm `seq` ta gõ lệnh:
`> help(seq)`
hoặc
`> ?seq`
- Tuy nhiên hàm **help** chỉ được thực hiện khi tên hàm được nhớ chính xác. Trong trường hợp ta không nhớ chính xác tên hàm thì có thể dùng hàm **apropos**. Hàm **apropos** liệt kê tất cả những hàm (trong gói cơ bản hoặc gói đã gọi ra) chứa cụm kí tự đang được tìm kiếm. Ví dụ để tìm tất cả những hàm kiểm định (chứa cụm từ "eq"), ta có thể thực hiện:
`> apropos(".eq")`