**Hướng dẫn sử dụng R cơ bản**

Họ và tên sinh viên: Hoàng Quang Thắng

Mã sinh viên: 20E1020011

**1: Cách cài đặt ngôn ngữ R:**

Bước 1: Truy cập vào đường link sau <https://cran.r-project.org/>

Bước 2: Chọn Download R for Windows

Bước 3: Chọn Base sau đó chọn Download R 3.5.1 for Windows

Bước 4: Sau khi hiện lên cửa sổ thì bấm next và chọn ổ đĩa để lưu,cuối cùng bấm finish để hoàn tất quá trình cài đặt.

**2: Cách cài đặt Rstudio:**

Bước 1:Truy cập vào đường link sau <https://rstudio.com/>

Bước 2: Chọn phiên bản Rstudio dành cho window sau đó tiến hành cài đặt.

Bước 3: Làm tương tự bước 4 ở trên để hoàn tất quá trình cài đặt.

**3: Cửa sổ làm việc trong phần mềm R:**

**+Phần mềm làm việc R có 4 cửa sổ chính:**

**-Cửa sổ thứ 1 gọi là:** script dùng để viết các câu lệnh.

**-Cửa sổ thứ 2** gọi là: console có thể dùng để viết câu lệnh trực tiếp vào đây.

**-Cửa sổ thứ 3** bao gồm 3 cửa sổ nhỏ:

\* environment dùng để chứa các bảng biểu hoặc các tham số mà chúng ta đưa ra.

\* history chứa lịch sử các câu lệnh.

\* connections dùng để liên kết với các file khác.

**-Cửa sổ thứ 4 bao gồm:**

\*file chứa các thư mục gốc.

\*plots chứa bảng biểu của các câu lệnh.

\*packages chứa những gói mà ta đã cài đặt.

\*help tìm kiếm lệnh và trợ giúp.

\*viewer

**4: Khởi động R và các phép tính cơ bản:**

-phép cộng: 1+2, 3+2...

-phép trừ: 5-3, 8-2…

-phép nhân: 4\*3, 6\*8…

-phép chia: 6/3, 8/4...

-phép mũ: 2\*\*4 hoặc 2^4…

-phần nguyên: 9%/%5, 8%/%3…

-phần dư: 9%%5, 9%%2…

-căn bậc hai: sqrt(4)=2…

-trị tuyệt đối: abs(-2)=2…

**5: Cấu trúc vector, matrix, list:**

**5.1: Cấu trúc vector:**

-Vector là kiểu dữ liệu chứa nhiều phần tử trong nó. Các phần tử trong cùng 1 vector phải cùng kiểu. Cách tạo một vector:

-tạo một vector:

x < - c(1,2,3)

x

-kết quả: ## [1] 1 2 3

-kiểu của 1 vector: class()

- Khi các phần tử trong 1 vector lưu các biến khác kiểu:

x < - c(1,2, “khanh”)

x

-kết quả: ## [1] "1" "2" "Khanh"

**5.2: Cấu trúc matrix:**

-Matrix là một dữ liệu có dạng ma trận kích thước mxn. Trong đó m là số dòng, n là số cột. Các dòng và các cột của một matrix phải cùng độ dài và kiểu dữ liệu. Cú pháp chung tạo một matrix:

mymatrix = matrix(vector, nrow=r, ncol=c, byrow=FALSE, dimnames=list(char\_vector\_rownames, char\_vector\_colnames))

**-Cách tạo 1 matrix:**

-Tạo matrix 3x4: y = matrix(1:12, nrow = 3, ncol = 4)

cells = c(12,3,45,2)

rnames = c("R1","R2")

cnames = c("C1","C2")

mymatrix1 = matrix(cells, nrow = 2, ncol = 2, byrow = TRUE,

dimnames = list(rnames,cnames))

mymatrix1

**-Kết quả:**

## C1 C2

## R1 12 3

## R2 45 2

**-Để lấy các phần tử là dòng cột trong matrix ta làm như sau**

-giá trị mymatrix3

## x y

## [1,] 1 2

## [2,] 2 3

## [3,] 3 4

## [4,] 4 5

## [5,] 5 6

**-phần tử dòng thứ nhất**

mymatrix3[,1]

## [1] 1 2 3 4 5

-**phần tử cột thứ hai**

mymatrix3[2,]

## x y

## 2 3

**-Lấy dòng thứ 2 và 3 của toàn bộ các cột**

mymatrix3[2:3,]

## x y

## [1,] 2 3

## [2,] 3 4

**5.3: Danh sách list:**

-List là một object trong R chứa nhiều phần tử có định dạng khác nhau như numeric, string, vector, và thậm chí là trong list có thể chứa các list khác. Một list có thể chứa cả hàm, data frame, array và matrix như là những phần tử của nó.

Để khởi tạo một list ta sử dụng hàm list:

newList = list(c(1,2,3), "Khanh", "Pham Đình", df, TRUE, 12.4,434.523)

newList

**-Kết quả:**

## [[1]]

## [1] 1 2 3

##

## [[2]]

## [1] "Khanh"

##

## [[3]]

## [1] "Pham Ðình"

##

## [[4]]

## STT Name Salary Birday newcol

## 1 1 Khanh 1238.09644474868 1993-07-09 90.92010

## 2 2 Linh Update 982.882232745103 1998-11-21 79.55551

## 3 3 Trang 1280.08912814394 1994-06-09 99.92532

## 4 4 Mai 1117.68697602614 1999-07-06 102.34751

## 5 5 Alex 804.399648627006 1992-05-19 106.31634

## 6 6 Mari 841.286513516871 1993-07-09 98.42687

## 7 7 Bayes 1067.10114307723 1998-11-21 86.84785

## 8 8 Jon 691.278532358373 1994-06-09 94.18305

## 9 9 Jame 842.396307641932 1999-07-06 98.85276

## 10 10 But 812.405301712449 1992-05-19 88.80000

## 11 11 Last 900 1990-12-01 111.37788

##

## [[5]]

## [1] TRUE

##

## [[6]]

## [1] 12.4

##

## [[7]]

## [1] 434.523

**-Truy cập vào phần tử thứ nhất của list:**

newList[1]

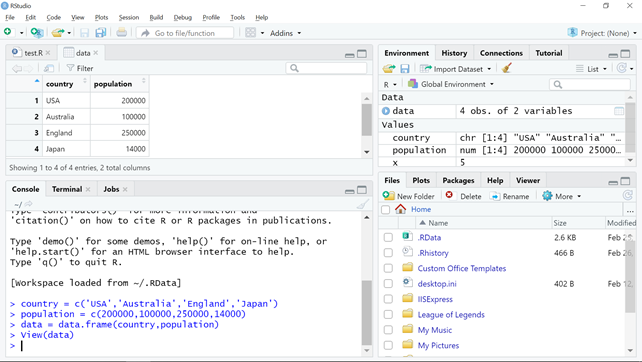
**-Thêm sửa xóa các phần tử trong list:**

newList < - list(“Name”,df)

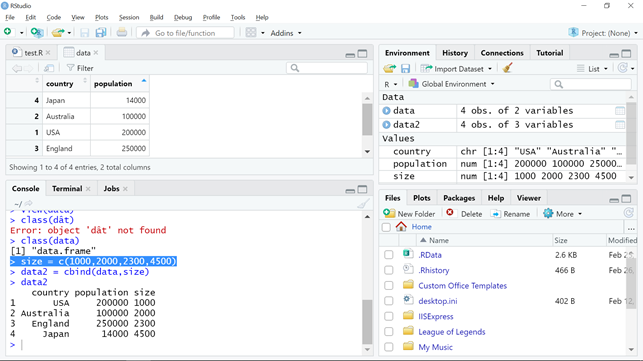
newList

**6: Cấu trúc Dataframe:**

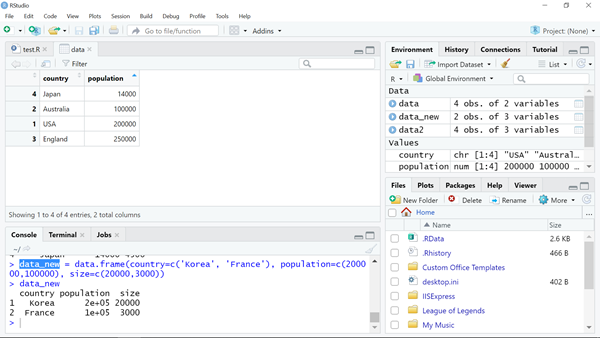
-Các dataframe được nhập vào và kết hợp với nhau:



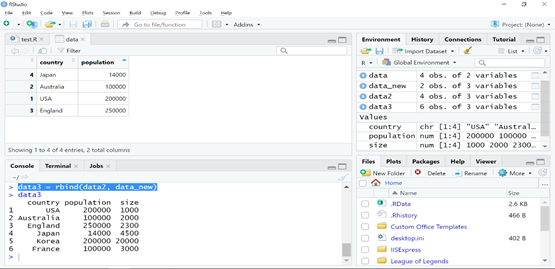
-Tạo ra 1 dataframe và kết hợp với dataframe có sẵn theo cột:



-Tạo cùng lúc nhiều dataframe cùng lúc:

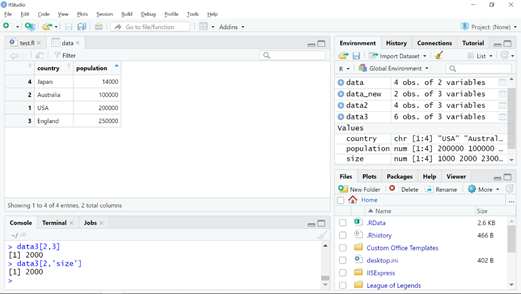


-Kết hợp nhiều dataframe theo hàng:

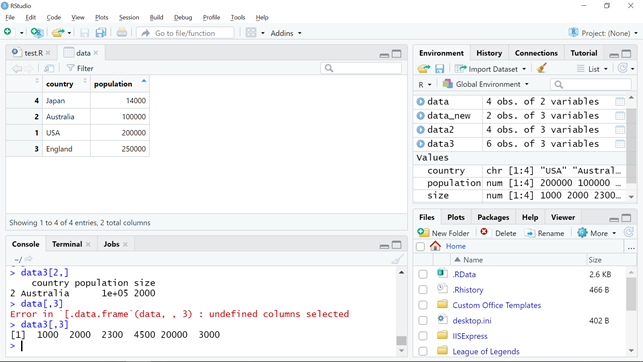


-Trích xuất dữ liệu từ dataframe:

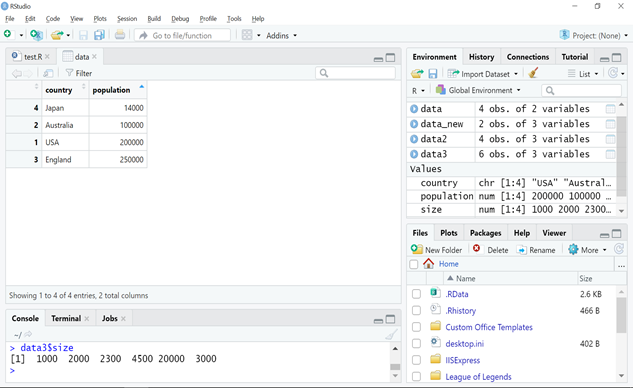
data3[2,3]= data3[2,“size”]: lấy dữ liệu từ hàng 2 cột 3



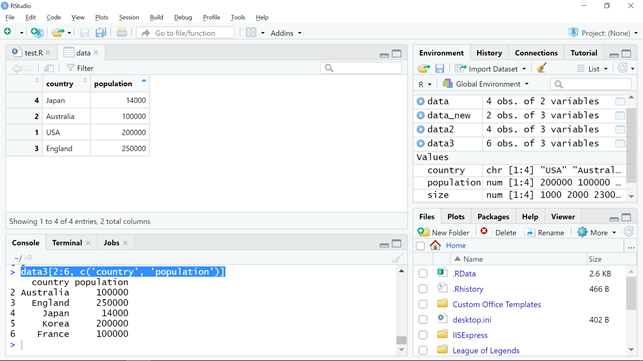
data3[2,] và data3[,3]: lấy cả hàng 2 hoặc cả cột 3



-data$size: cách khác để lấy dữ liệu từ cột 3(size)



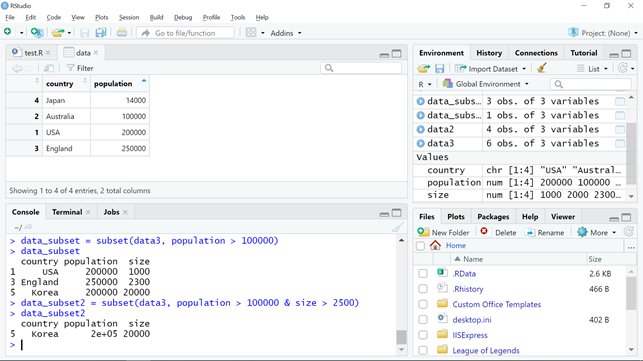
-data3[2:6, c('country', 'population')]: lấy dữ liệu từ hang 2 đến hang 6, cột “country” và “population”



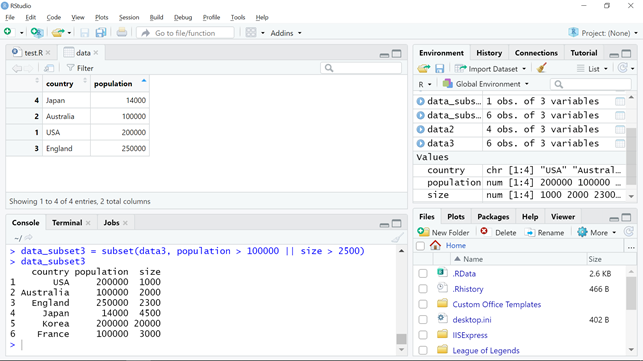
Lấy dữ liệu có các điều kiện nhất định:

data\_subset thì điều kiện là population > 100000

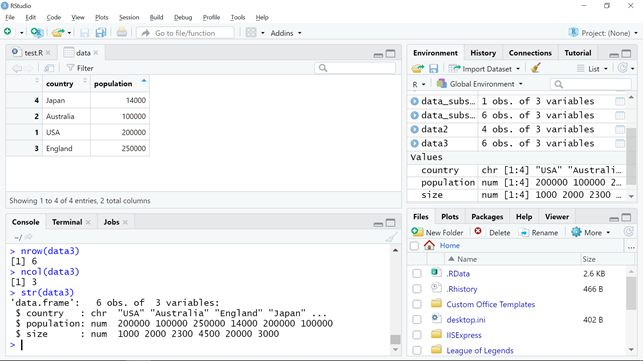
data\_subset2 thì điều kiện là population > 100000 và size > 2500



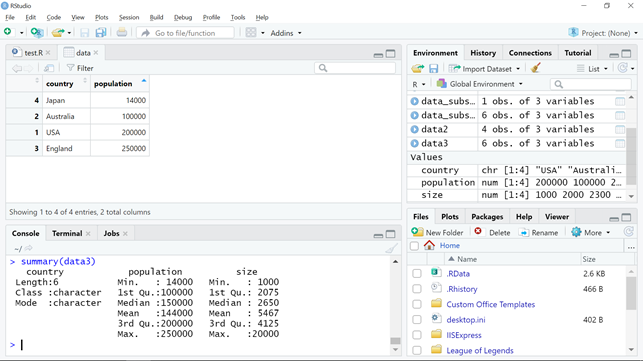
Lấy dữ liệu có population > 100000 hoặc size > 2500:



Số hàng, cột và thông tin cơ bản của dữ liệu data3:



Thống kê các thông tin của dữ liệu data3:



**7: Import text file,csv, excel vào R:**

Để import file Excel vào R, chúng ta có thể sử dụng package readxl.

library(readxl)

library(dplyr)

-Giả sử, chúng ta đang có 1 file Excel datasets.xlsx có sẵn trong thư mục package readxl trong đường dẫn sau:

path = readxl\_example("datasets.xlsx")

path

-File Excel datasets.xlsx có 4 sheets: iris, mtcars, chickwts, quakes.

-Bây giờ để import dữ liệu iris (sheet đầu tiên trong file Excel này) vào R, chúng ta sẽ sử dụng hàm read\_excel() với câu lệnh như sau:

iris = read\_excel(path,sheet = 1)

iris %>%

as.data.frame %>%

head

**-Để import dữ liệu từ file excel vào R thì ta dùng hàm sau:**

write.xlsx() trong package xlsx.

library(xlsx)

write.xlsx(iris

file = paste0(getwd(),"/","iris\_data.xlsx")

row.names = F

)

**8: Export số liệu từ R ra text file, csv, excel**

**-Export to txt file:**

write.table(“tên của dataframe”, “đường dẫn của file”, sep = “\t”)

VD: write.table(who, "C:/Users/RED/Documents/mydata.txt", sep="\t", row.names=FALSE)

**-Write csv in R:**

write.csv(“tên của dataframe”, “đường dẫn của file”)

write.table(“tên của dataframe”, “đường dẫn của file”,sep = “,”)

VD: write.csv(who, file = "C:/Users/RED/Documents/mydata2.csv", row.names=FALSE, na="None")

**-Export to excel:**

library(‘xlsx’)

write.xlsx(“tên của dataframe”, “đường dẫn của file)