**ĐẠI HỌC HUẾ**



# KHOA KỸ THUẬT VÀ CÔNG NGHỆ

🙠🙟🕮🙝🙢

****

(MẪU BÌA PHỤ)

**BÁO CÁO ĐỒ ÁN**

**NĂM HỌC 2021-2022**

**Giảng viên hướng dẫn: Hồ Quốc Dũng**

**Lớp: KHDL & TTNT**

**Sinh viên thực hiện: Hoàng Nữ Thu Phương**

*(ký tên và ghi rõ họ tên)*

|  |
| --- |
| Số phách  *(Do hội đồng chấm thi ghi)* |

**Thừa Thiên Huế, ngày 30 tháng 05 năm 2022**

ĐẠI HỌC HUẾ

**KHOA KỸ THUẬT VÀ CÔNG NGHỆ**

🙠🙟🕮🙝🙢

**PHIẾU ĐÁNH GIÁ ĐỒ ÁN/TIỂU LUẬN/BÀI TẬP LỚN**

**Học kỳ II, năm học 2021 - 2022**

|  |  |
| --- | --- |
| **Cán bộ chấm thi 1** | **Cán bộ chấm thi 2** |
| **Nhận xét:**  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  **Điểm đánh giá của CBCT1:**  Bằng số: .........................................  Bằng chữ: ....................................... | **Nhận xét:**  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  **Điểm đánh giá của CBCT2:**  Bằng số: .........................................  Bằng chữ: ....................................... |

Điểm kết luận: ...........................................................................................................

Bằng số:.....................................................................................................................

Bằng chữ: .................................................................................................................

Thừa Thiên Huế, ngày tháng năm 2022

|  |  |
| --- | --- |
| **Cán bộ chấm thi 1**  (Ký và ghi rõ họ và tên) | **Cán bộ chấm thi 2**  (Ký và ghi rõ họ và tên) |

**TÓM TẮT NỘI DUNG CHƯƠNG**

**Chương 1: Introduction**

    Nói một cách ngắn gọn, R là một phần mềm sử dụng cho phân tích thống kê và đồ thị. Thật ra

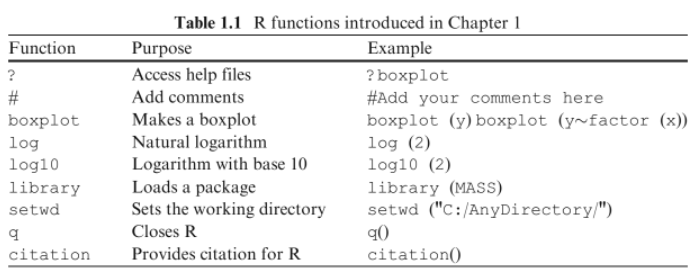
về bản chất, R là ngôn ngữ máy tính đa năng, có thể sử dụng cho nhiều mục tiêu khác nhau,

từ tính toán đơn giản, toán học giả trí, tính toán ma trận, đến các phân tích thống kê phức tạp.

Vì là một ngôn ngữ cho nên người ta có thể sử dụng R để phát triển các thành phần mềm

chuyên môn cho một vẫn đề tính toán cá biệt.

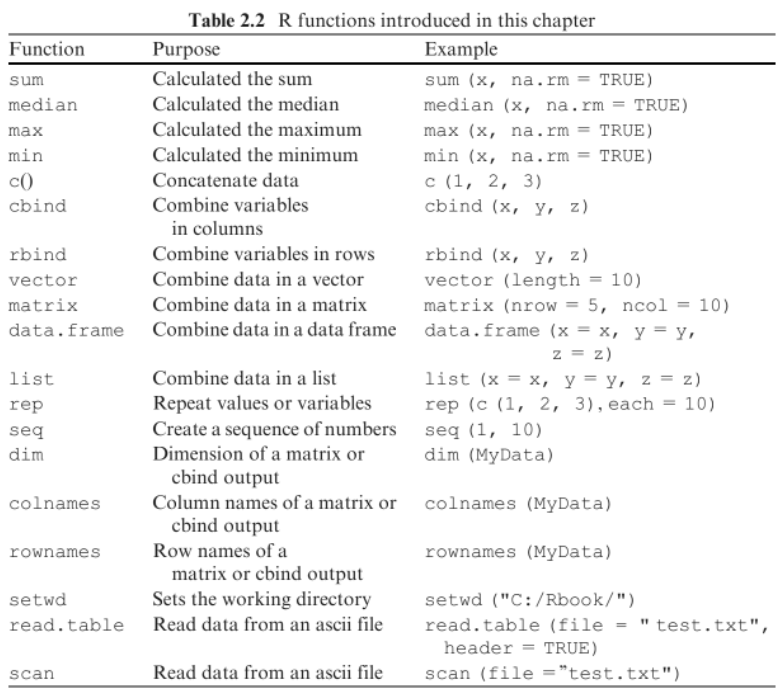
   Trong chương đầu này, Giới thiệu về ngôn ngữ lập trình R, hướng dẫn cách tải, cài đặt R và cách sử dụng ngôn ngữ này



|  |  |  |
| --- | --- | --- |
| **FUNCTION** | **PURPOSE** | **EXAMPLE** |
| **?** | Truy cập các tệp trợ giúp | ?print  ?install.packegs |
| **#** | Cho phép thêm nhận xét, ghi chú | # thu phuong xinh dep  # thu phuong dang yeu |
| **boxplot** | Tạo một boxplot | boxplot(x, ...)  boxplot(count ~ spray, data = InsectSprays,  notch = TRUE, add = TRUE, col = "blue") |
| **log** | Logarit tự nhiên | log(17) |
| **log10** | Logarit cơ số 10 | log(170) |
| **library** | Truy cập thư viện | library(dequer) |
| **setwd** | Đặt vị trí của thư mục làm việc | setwd(dir)  setwd("C:/Ngôn ngữ lập trình R") |
| **q** | Đóng R | q() |
| **citation** | Cung cấp trích dẫn | citation() |

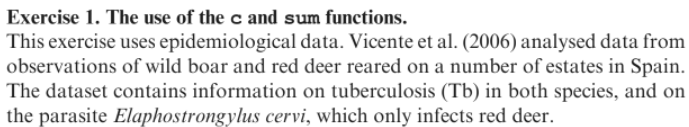
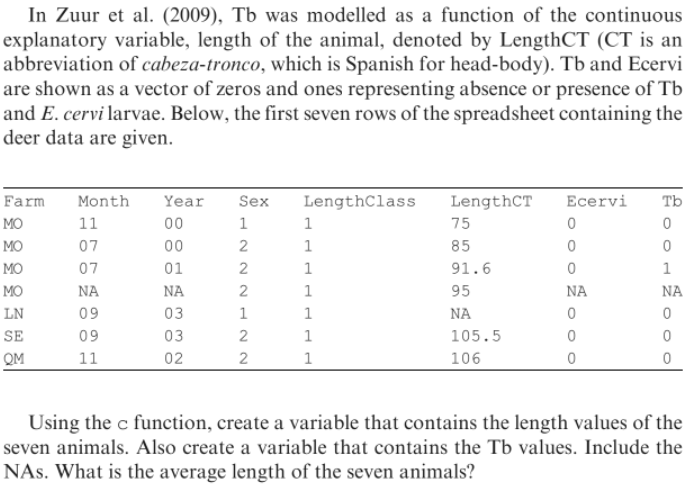
**Chương 2: Getting Data into R**

Ở chương 1 đã trình bày cách nhập dữ liệu từ bảng tính, tiếp theo Trong chương 2, chúng ta sẽ đề cập đến việc nhập dữ liệu vào R và tổ chức nó dưới dạng vô hướng (giá trị đơn), vectơ, ma trận, khung dữ liệu hoặc danh sách. Chúng tôi cũng chứng minh việc lấy dữ liệu từ Excel, tệp ascii, cơ sở dữ liệu và các chương trình thống kê khác.

****

|  |  |  |
| --- | --- | --- |
| **FUNCTION** | **PURPOSE** | **EXAMPLE** |
| **sum** | Trả về giá trị tổng của mảng truyền vào |  |
| **median** | Trả về giá trị trung vị của mảng truyền vào |  |
| **max** | Trả về giá trị lớn nhất |  |
| **min** | Tính giá trị nhỏ nhất |  |
| **c()** | Nhập số liệu trực tiếp, tạo ra một cột dữ liệu |  |
| **cbind** | Lấy một chuỗi các đối số vectơ, ma trận hoặc khung dữ liệu và kết hợp thành các cột tương ứng. |  |
| **rbind** | Lấy một chuỗi vectơ, ma trận hoặc dataframe, kết hợp chúng tương ứng theo hàng. |  |
| **vector** | Trả về một vector với độ dài tùy chọn và với kiểu dữ liệu tùy chọn |  |
| **matrix** | Kết hợp dữ liệu thành một ma trận |  |
| **data.frame** | Kết hợp dữ liệu trong một khung dữ liệu |  |
| **list** | List chứa các phần tử thuộc các kiểu khác nhau như - số, chuỗi, vectơ hoặc một list khác bên trong danh sách.List cũng có thể chứa ma trận hoặc một hàm. |  |
| **rep** | Tạo bản sao của giá trị hoặc biến |  |
| **seq** | Tạo ra 1 dãy số |  |
| **dim** | Cho biết có bao nhiêu cột (hay variable = biến số) và dòng số liệu (0bservations) trong dữ liệu. |  |
| **colnames** | Đổi tên của cột trong khum dữ liệu |  |
| **rownames** | Đổi tên của hàng trong khum dữ liệu |  |
| **setwd** | Đặt vị trí của thư mục làm việc | setwd(dir)  setwd("C:/Ngôn ngữ lập trình R") |
| **read.table** | Đọc dữ liệu từ file ASCII |  |
| **scan** | Đọc dữ liệu từ file ASCII |  |

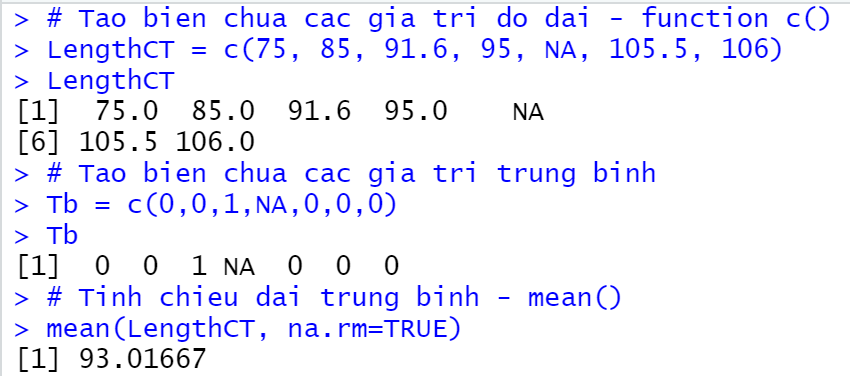
**Exercises**

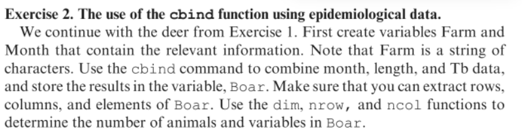
**** ****

**DỊCH:**

      Bài tập này sử dụng dữ liệu dịch tễ học. Vicente và cộng sự. (2006) đã phân tích dữ liệu từ việc phát hiện lợn rừng và hươu đỏ được nuôi trên một số khu đất ở Tây Ban Nha. . (2009), Tb được mô hình hóa dưới dạng hàm của biến số liên tục, chiều dài của động vật, được ký hiệu là LengthCT (CT là từ viết tắt của cabeza-tronco, tiếng Tây Ban Nha có nghĩa là thân đầu). Tb và Ecerviare được hiển thị dưới dạng véc tơ gồm các số không và các véc tơ đại diện cho sự vắng mặt hoặc hiện diện của ấu trùngTb và E. cervi. Dưới đây, bảy hàng đầu tiên của bảng tính chứa dữ liệu hươu được đưa ra

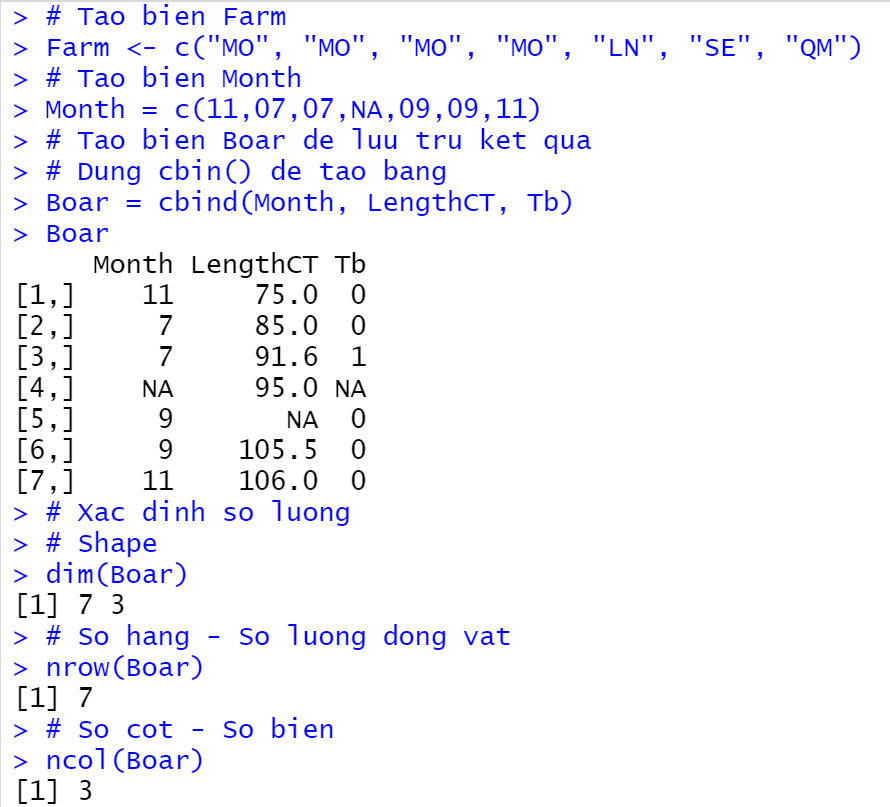
    Sử dụng hàm c, tạo một biến chứa các giá trị độ dài của bảy con vật. Đồng thời tạo một biến chứa các giá trị Tb. Bao gồm cả  NAs. Chiều dài trung bình của bảy con vật là bao nhiêu?

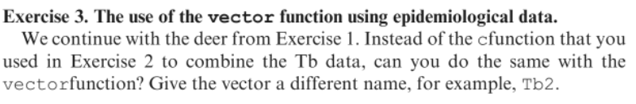




**DỊCH:**

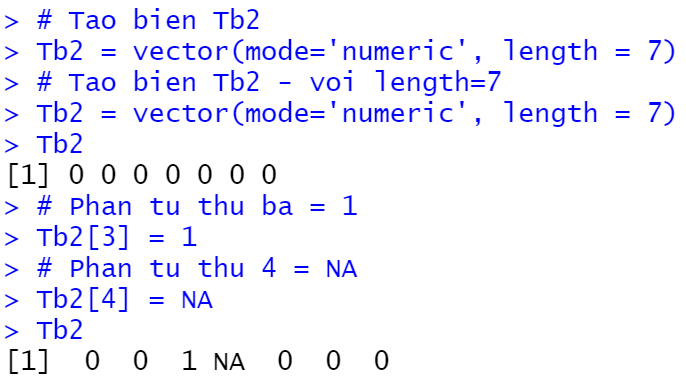
Chúng ta tiếp tục với con hươu từ Exercise 1. Đầu tiên tạo các biến Farm và Tháng có chứa thông tin liên quan. Lưu ý rằng Trang trại là một chuỗi nhân vật. Sử dụng lệnh **cbind** để kết hợp dữ liệu tháng, độ dài và Tb, và lưu trữ kết quả trong biến Boar. Đảm bảo rằng bạn có thể trích xuất các hàng, cột và các phần tử của Boar. Sử dụng các hàm **dim**, **nrow** và **ncol** để xác định số lượng động vật và các biến trong Boar.





**DỊCH:**

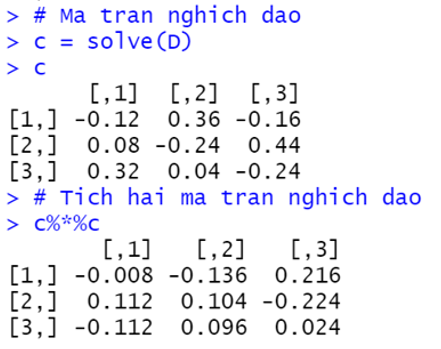
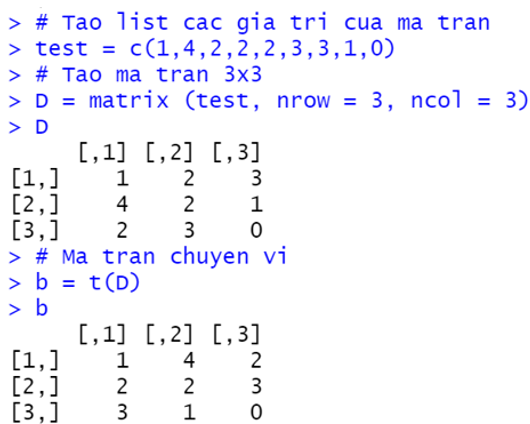
Chúng ta tiếp tục với con nai từ Exercise 1. Thay vì chức năng mà bạn được sử dụng trong Exercise 2 để kết hợp dữ liệu Tb, bạn có thể làm tương tự với hàm **vector**? Đặt tên khác cho vector, chẳng hạn như Tb2.

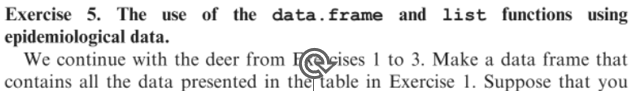
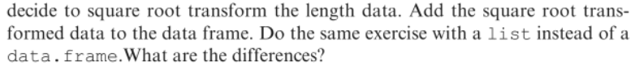




**DỊCH:**

Tạo ma trận sau bằng R và xác định chuyển vị, nghịch đảo của nó và bội D với nghịch đảo của nó (kết quả phải là ma trận nhận dạng).

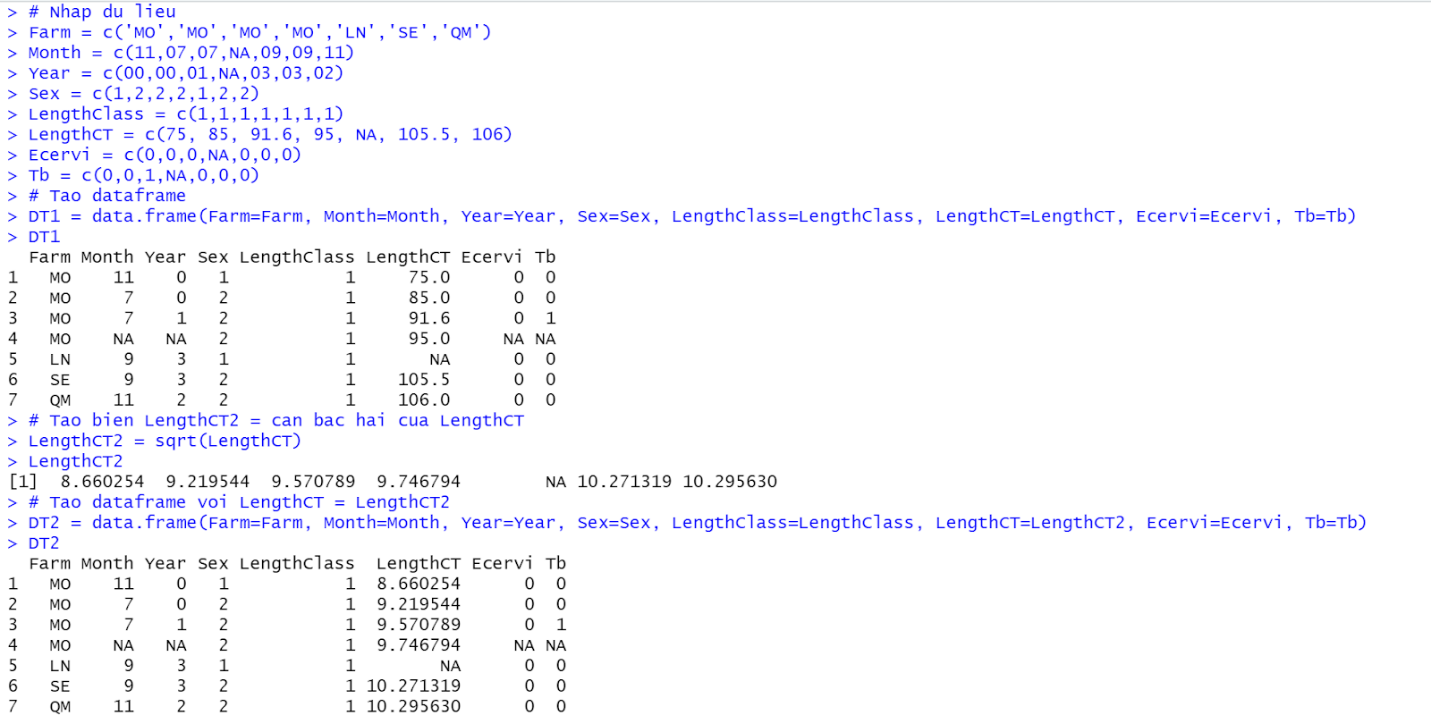
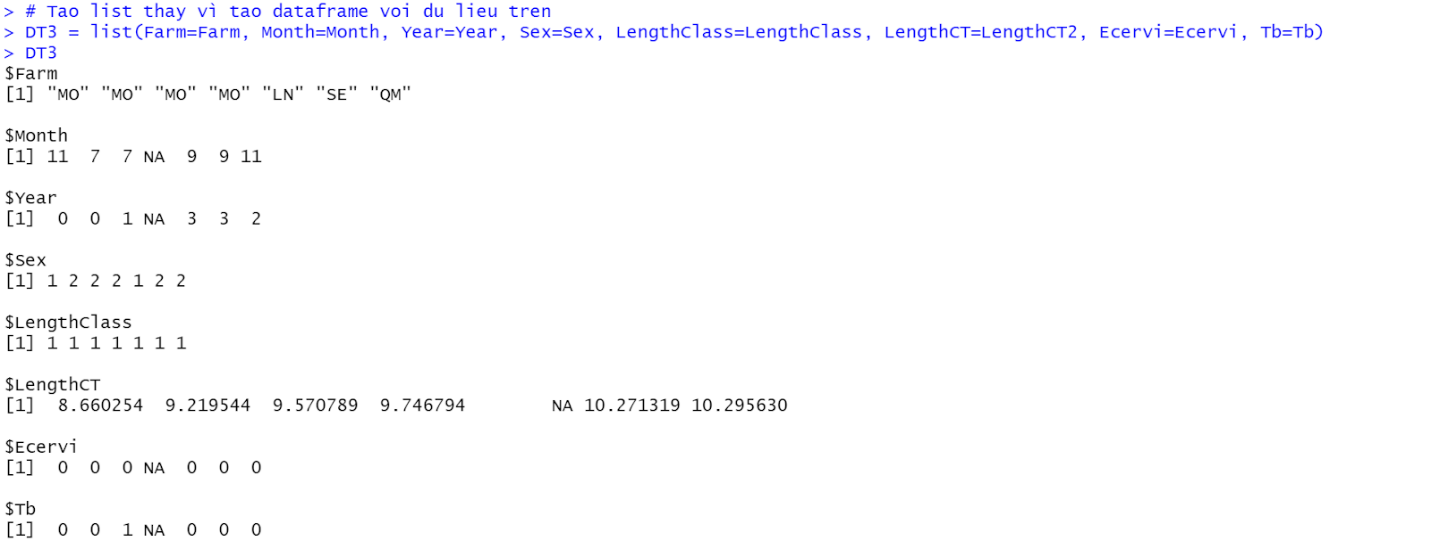


**DỊCH:**

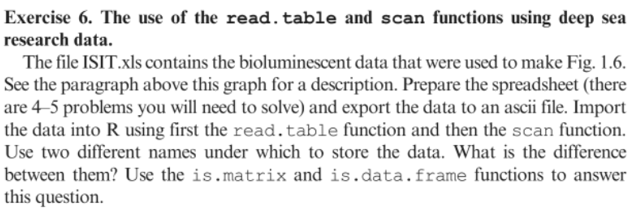
Chúng ta tiếp tục với hươu từ Exercise  1 đến Exercise 3. Tạo khung dữ liệu chứa tất cả dữ liệu được trình bày trong bảng trong Exercise 1. Giả sử rằng bạn quyết định căn bậc hai biến đổi dữ liệu độ dài. Thêm căn bậc hai chuyển đổi thành dữ liệu vào khung dữ liệu. Thực hiện bài tập tương tự với danh sách thay vì data.frame. sự khác biệt là gì?

https://github.com/hoangnuthuphuong/HK2\_freshman/blob/main/NNLT\_R/Chapter2\_Getting\_Data\_intoR/Exercise2.5.R

* **Sự khác biệt:**

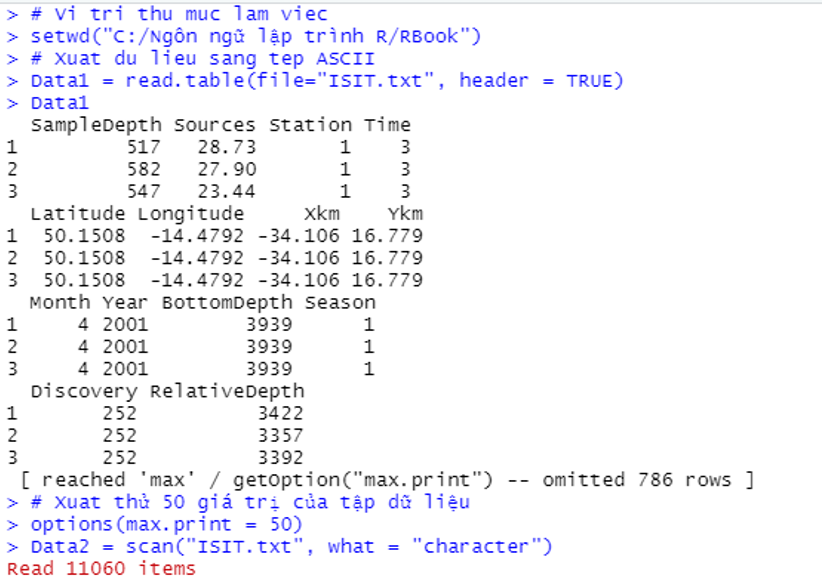
**-** Data frame cho output dạng bảng giống với bảng dữ liệu được cho sẵn. Dễ đọc dễ nhìn hơn so với list.

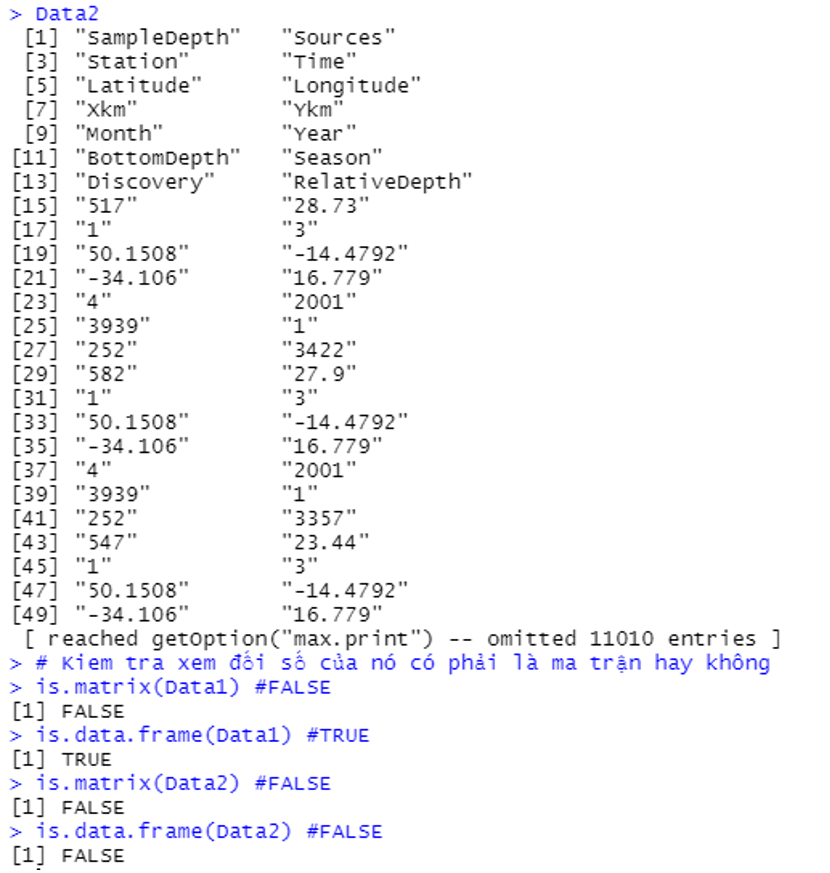
          - List cho output ra từng dữ liệu của từng cột theo dạng liệt kê.   
  
****

**DỊCH:**

Tệp ISIT.xls chứa dữ liệu phát quang sinh học được sử dụng để tạo Hình 1.6. Xem đoạn trên biểu đồ này để biết mô tả. Chuẩn bị bảng tính (ở đó là 4–5 vấn đề bạn sẽ cần giải quyết) và xuất dữ liệu sang tệp ascii. Nhập dữ liệu vào R trước tiên bằng cách sử dụng hàm **read.table** và sau đó là hàm **scan**. Sử dụng hai tên khác nhau để lưu trữ dữ liệu. Sự khác biệt là gì? Sử dụng các hàm **is.matrix** và **is.data.frame** để trả lời câu hỏi này.

<https://github.com/hoangnuthuphuong/HK2_freshman/blob/main/NNLT_R/Chapter2_Getting_Data_intoR/Exercise2.6.R>

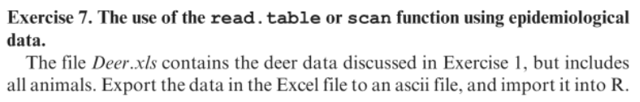




* **Sự khác biệt:**

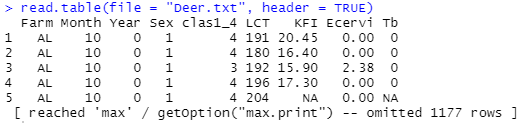
**-** Đối số của cả 2 tập dữ liệu đều không phaỉ là ma trận

          - Chỉ có tập tập dữ liệu Data1 có đối số là dataframe còn Data 2 thì không



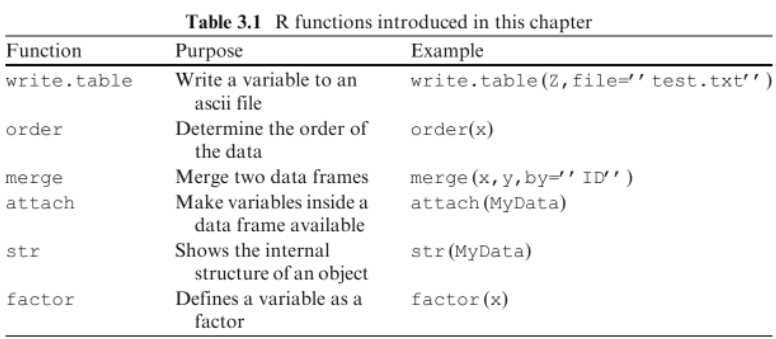
**DỊCH:**

Tệp Deer.xls chứa dữ liệu Deer được thảo luận trong Bài tập 1, nhưng bao gồm tất cả các loài động vật. Xuất dữ liệu trong tệp Excel sang tệp ascii và nhập dữ liệu đó vào R.

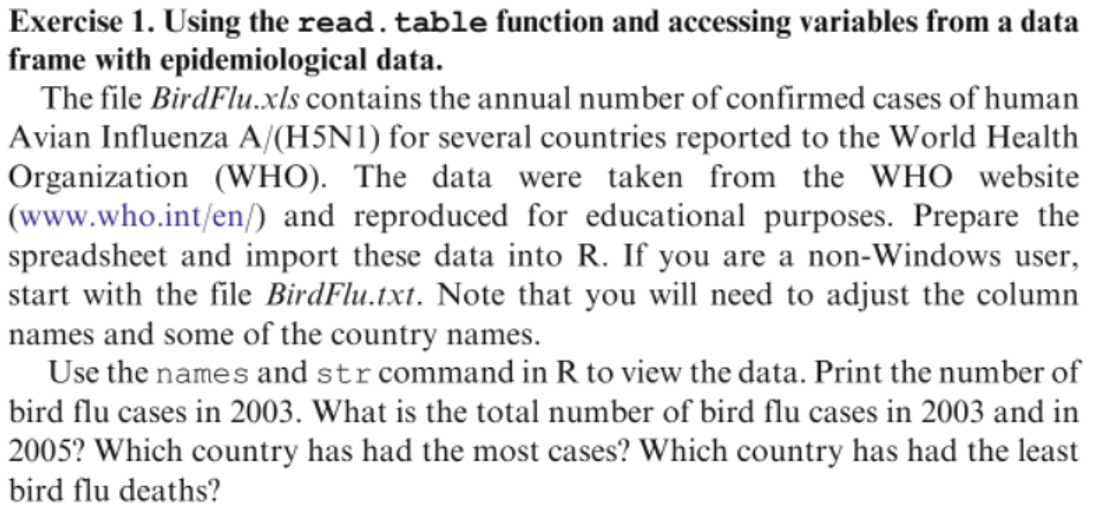


**Chương 3: Accessing Variables and Managing Subsets of Data**

Ở chương 3, chúng ta thảo luận về việc truy cập các tập con của dữ liệu.bằng các hàm như write.table, order, merge, attach…Trình bày cách nhập dữ liệu; đọc dữ liệu từ bảng tính, tệp ascii hoặc cơ sở dữ liệu; và trích xuất các tập dữ liệu con.



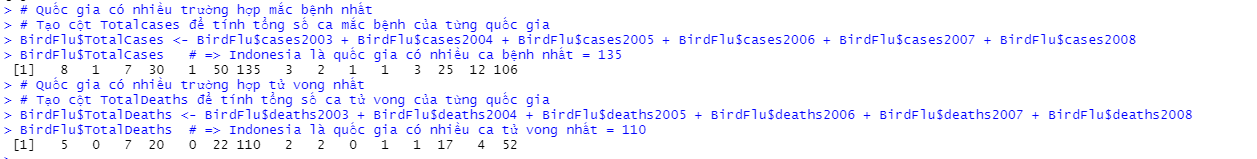
|  |  |  |
| --- | --- | --- |
| **FUNCTION** | **PURPOSE** | **EXAMPLE** |
| **write.table** | Ghi một biến vào tệp ascii | write.table (Z,file=’’test.txt’’) |
| **order** | Xác định thứ tự của dữ liệu | Order(x) |
| **merge** | Hợp nhất hai khung dữ liệu | merge (x,y,by=’’ID’’) |
| **attach** | Làm cho các biến bên trong khung dữ liệu có sẵn | attach (MyData) |
| **str** | Hiển thị cấu trúc bên trong của một đối tượng | str (MyData) |
| **factor** | Định nghĩa một biến như một nhân tố | factor(x) |



**DỊCH:**

Tệp BirdFlu.xls chứa số trường hợp được xác nhận hàng năm của con người Cúm gia cầm A / (H5N1) ở một số quốc gia đã được báo cáo cho Y tế Thế giới Tổ chức (WHO). Dữ liệu được lấy từ trang web của WHO(www.who.int/en/) và được sao chép lại cho các mục đích giáo dục. Chuẩn bị bảng tính và nhập những dữ liệu này vào R. Nếu bạn không phải là người dùng Windows,bắt đầu với tệp BirdFlu.txt. Lưu ý rằng bạn sẽ cần điều chỉnh cột tên và một số tên nước.

Sử dụng **names** và lệnh **str** trong R để xem dữ liệu. In số ca bệnh cúm gia cầm năm 2003. Tổng số ca bệnh cúm gia cầm trong năm 2003 và năm Năm 2005? Quốc gia nào đã có nhiều trường hợp nhất? Quốc gia nào có ít nhất tử vong do cúm gia cầm?





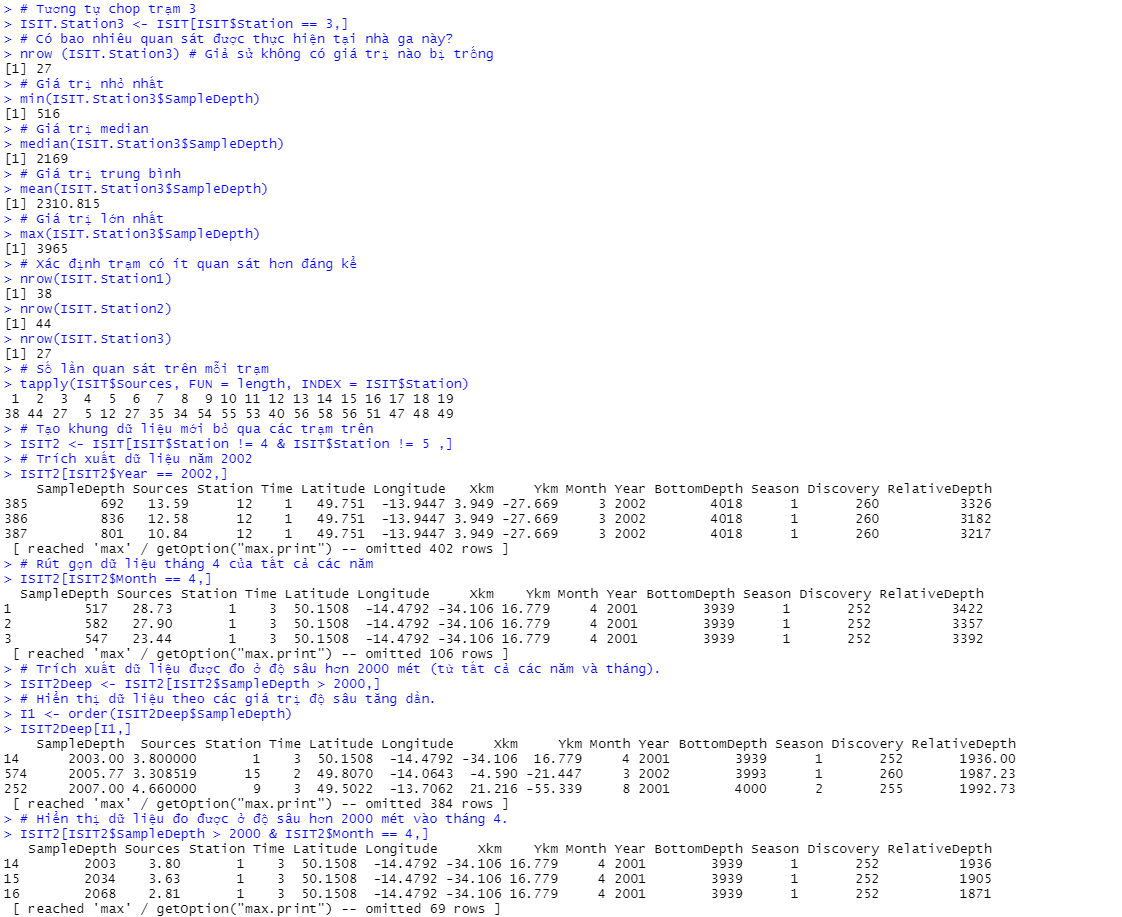
**DỊCH:**

Nếu bạn chưa hoàn thành Bài tập 6 trong Chương 2, hãy làm như vậy và nhập dữ liệu từ tệp ISIT.xls.

Trong R, trích xuất dữ liệu từ trạm 1. Có bao nhiêu quan sát được thực hiện tại ga này? Mẫu tối thiểu, trung bình, trung bình và tối đa là gìđộ sâu tại trạm 1? Giá trị tối thiểu, trung bình, trung bình và tối đa là bao nhiêu độ sâu lấy mẫu tại trạm 2? Ở trạm 3?

Xác định bất kỳ trạm nào có ít quan sát hơn đáng kể. Tạo một cái mới khung dữ liệu bỏ qua các trạm này.

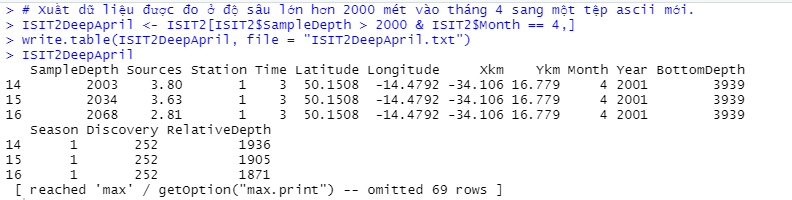
Trích xuất dữ liệu từ năm 2002. Trích xuất dữ liệu từ tháng 4 (của tất cả các năm). Trích xuất dữ liệu được đo ở độ sâu lớn hơn 2000 mét (từ tất cả các năm và tháng). Hiển thị dữ liệu theo giá trị độ sâu tăng dần. Hiển thị dữ liệu được đo ở độ sâu lớn hơn 2000 mét trong Tháng tư. 

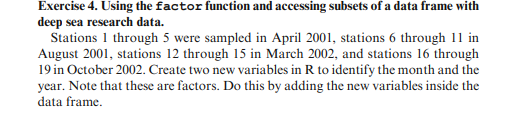




**DỊCH:**

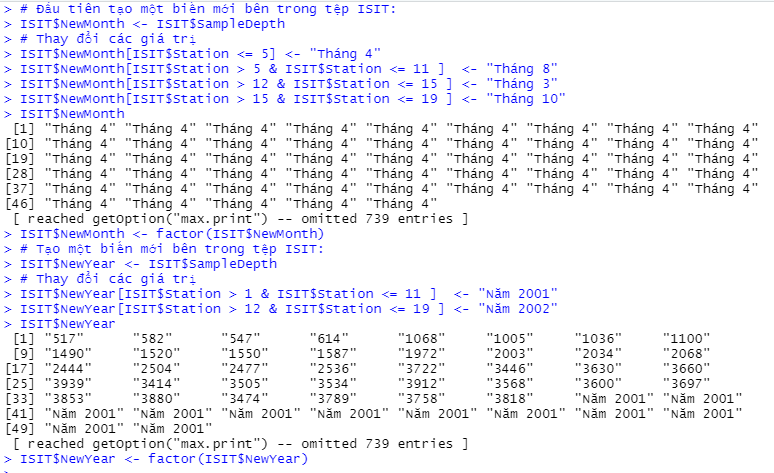
Trong bước cuối cùng của bài tập trước, dữ liệu được đo ở độ sâu lớn hơn 2000 mét vào tháng 4 đã được khai thác. Xuất những dữ liệu này sang một tệp ascii mới.





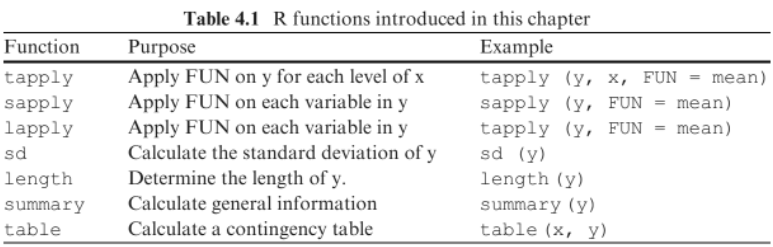
**DỊCH:**

Các trạm từ 1 đến 5 được lấy mẫu vào tháng 4 năm 2001, các trạm từ 6 đến 11 ở Tháng 8 năm 2001, các trạm từ 12 đến 15 vào tháng 3 năm 2002, và các trạm từ 16 đến 19 vào tháng 10 năm 2002. Tạo hai biến mới trong R để xác định tháng và năm. Lưu ý rằng đây là những yếu tố. Làm điều này bằng cách thêm các biến mới bên trong khung dữ liệu.



**Chương 4: SIMPLE FUNCTION**

Trong chương này, chúng ta thảo luận về việc áp dụng một số hàm đơn giản cho dữ liệu, chẳng hạn như giá trị trung bình hoặc giá trị trung bình của một tập hợp con dữ liệu đơn lẻ, tóm tắt các yếu tố, tính độ lệch chuẩn, xác định chiều dài, tính toán thống kê, …..

****

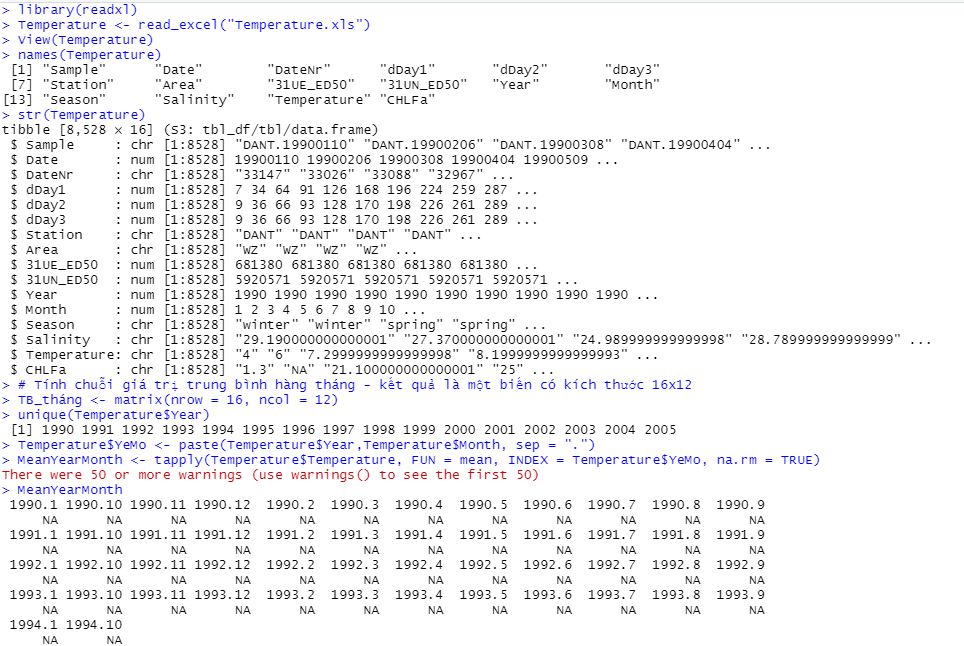
|  |  |  |
| --- | --- | --- |
| **FUNCTION** | **PURPOSE** | **EXAMPLE** |
| **tapply** | Cho phép bạn tạo tóm tắt nhóm dựa trên các mức yếu tố. | set.seed(2)  data\_set <- data.frame(price = round(rnorm(25, sd = 10, mean = 30)),  type = sample(1:4, size = 25, replace = true),  store = sample(paste("store", 1:4),  size = 25, replace = true))  head(data\_set)  price <- data\_set$price  store <- data\_set$store  type <- factor(data\_set$type,  labels = c("toy", "food", "electronics", "drinks"))  mean\_prices <- tapply(price, type, mean)  mean\_prices |
| **sapply** | Cho phép bạn lặp qua một danh sách hay vector không cần vòng lặp. | sapply(1:10, function(i) i ^ 2) |
| **lapply** | Cho phép áp dụng một hàm trên một danh sách hoặc một vector, trả về một danh sách. | #áp dụng với danh sách  a <- list(a = c(8, 9, 7, 5),  b = data.frame(x = 1:5, y = c(5, 1, 0, 2, 3)))  a  lapply(a, sum)  #áp dụng với vector  b <- c(12, 18, 6)  lapply(b, sqrt) |
| **sd** | tính độ lệch chuẩn. | x <- c(10, 25, 12, 18, 5, 16, 14, 20)  sd(x) # 6.21059 |
| **length** | Xác định chiều dài. | x <- c(1:5)  length( c) |
| **summary** | Tính toán các tóm tắt thống kê của dữ liệu và các đối tượng mô hình. | vec <- 1:10  vec  summary(vec) |
| **table** | Tạo nhanh các bảng tần số. | df <- data.frame(player = c(‘aj’, ‘bob’, ‘chad’, ‘dan’, ‘eric’, ‘frank’), position = c(‘a’, ‘b’, ‘b’, ‘b’, ‘b’, ‘a’), points = c(1, 2, 2, 1, 0, 0))  df  table(df$position, df$point) |

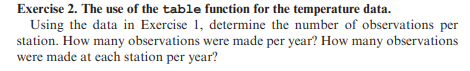
****

**DỊCH:**

Tệp tin Temperature.xls chứa các quan sát nhiệt độ được thực hiện tại 31 địa điểm dọc theo đường bờ biển Hà Lan. Dữ liệu được thu thập và cung cấp bởi Viện Hà Lan RIKZ (thuộc chương trình giám sát MWTL;  Giám sát Waterstaatkundige Toestand des Lands). Việc lấy mẫu bắt đầu vào năm 1990, và lần cuối cùng các phép đo trong bảng tính được thực hiện vào tháng 12 năm 2005, khoảng thời gian 16 năm. Tần suất lấy mẫu là 0–4 lần mỗi tháng, tùy thuộc vào mùa.Tính chuỗi giá trị trung bình hàng tháng một lần bằng cách sử dụng dữ liệu từ tất cả các trạm.

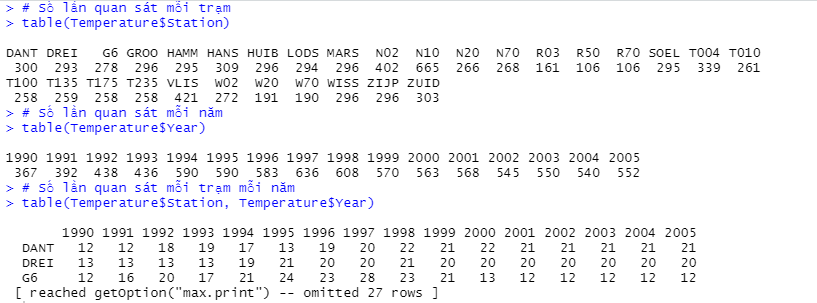
Kết quả cuối cùng phải là một biến có kích thước 16 12. Cũng tính toán độ lệch chuẩn và số lần quan sát mỗi tháng.





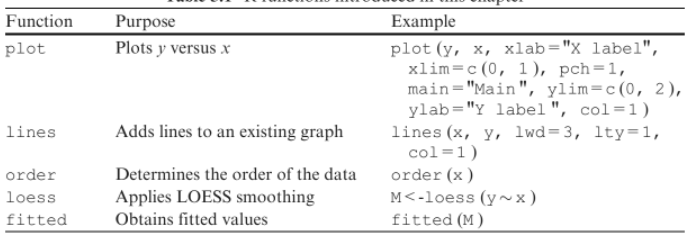
**DỊCH:**

Sử dụng dữ liệu trong Bài tập 1, xác định số lần quan sát mỗi trạm. Có bao nhiêu quan sát được thực hiện mỗi năm? Có bao nhiêu quan sát được thực hiện ở mỗi trạm mỗi năm?

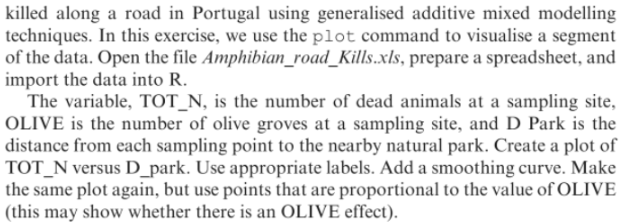


**Chương 5: An Introduction to Basic Plotting Tools**

Trong chương này, chúng tôi giới thiệu các công cụ vẽ đồ thị cơ bản. Nếu bạn chỉ quan tâm đến việc vẽ đồ thị đơn giản, chương này sẽ đủ; tuy nhiên, để xây dựng các đồ thị phức tạp hơn hoặc để thêm các phần tô điểm phức tạp hơn như dấu tích, hoặc phông chữ và kích thước phông chữ chuyên biệt, cho đồ thị cơ bản, bạn sẽ cần các kỹ thuật vẽ đồ thị nâng cao hơn được trình bày trong Chương 7 và 8.

****

|  |  |  |
| --- | --- | --- |
| **FUNCTION** | **PURPOSE** | **EXAMPLE** |
| **plot** | Plots y versus x | plot (y, x, xlab="X label",  xlim=c (0, 1 ), pch=1,  main="Main ", ylim=c (0, 2 ),  ylab="Y label ", col=1 ) |
| **lines** | Thêm các đường vào biểu đồ hiện có | lines (x, y, lwd=3, lty=1,col=1 ) |
| **order** | Xác định thứ tự của dữ liệu | Order(x) |
| **loess** | Áp dụng làm mịn LOESS | M<-loess (y x ) |
| **fitted** | Đạt được các giá trị phù hợp | fitted (M ) |

****

**DỊCH:**

Trong chương16 của Zuur và cộng sự. (2009), một nghiên cứu được trình bày phân tích số lượng các loài lưỡng cư bị giết dọc theo một con đường ở Bồ Đào Nha bằng cách sử dụng mô hình hỗn hợp phụ gia tổng quát kỹ thuật. Trong bài tập này, chúng ta sử dụng lệnh plot để hình dung một đoạn của dữ liệu. Mở tệp Amphibian roadKils.xls, chuẩn bị bảng tính và nhập dữ liệu vào R.

Biến, TOT\_N, là số lượng động vật chết tại một địa điểm lấy mẫu,OLIVE là số lượng rừng ô liu tại một địa điểm lấy mẫu, và D Park là khoảng cách từ mỗi điểm lấy mẫu đến công viên tự nhiên gần đó. Tạo một plot của TOT\_N so với D Park. Sử dụng nhãn thích hợp. Thêm một đường cong làm mịn. Làm Lại cùng một plot, nhưng sử dụng điểm tỷ lệ với giá trị của OLIVE(điều này có thể cho biết liệu có hiệu ứng OLIVE hay không).

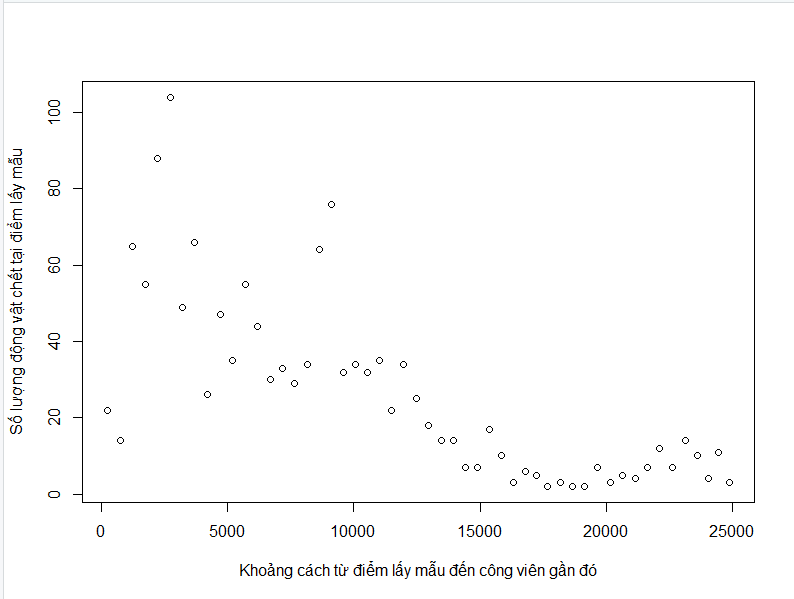


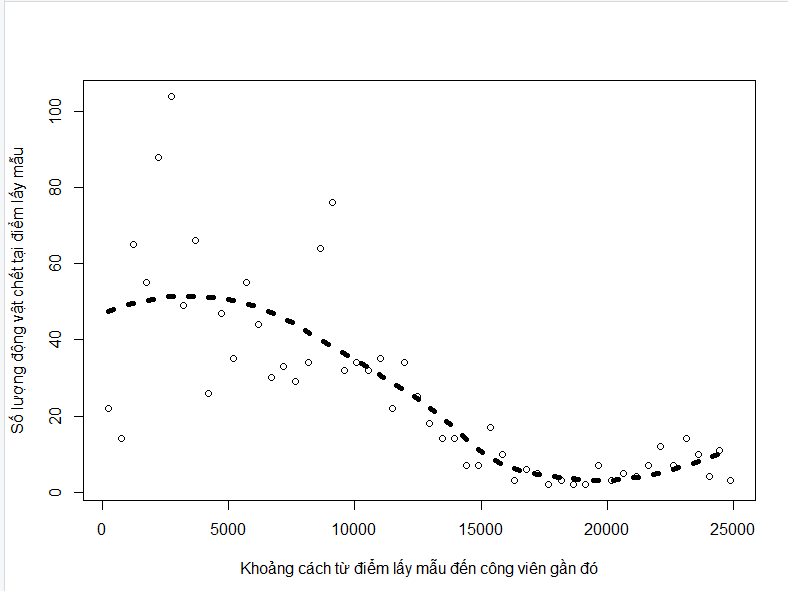
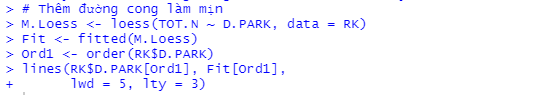
Biến, TOT\_N, là số lượng động vật chết tại từng điểm lấy mẫu,

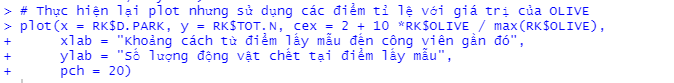
OLIVE là số vườn ô liu tại địa điểm lấy mẫu

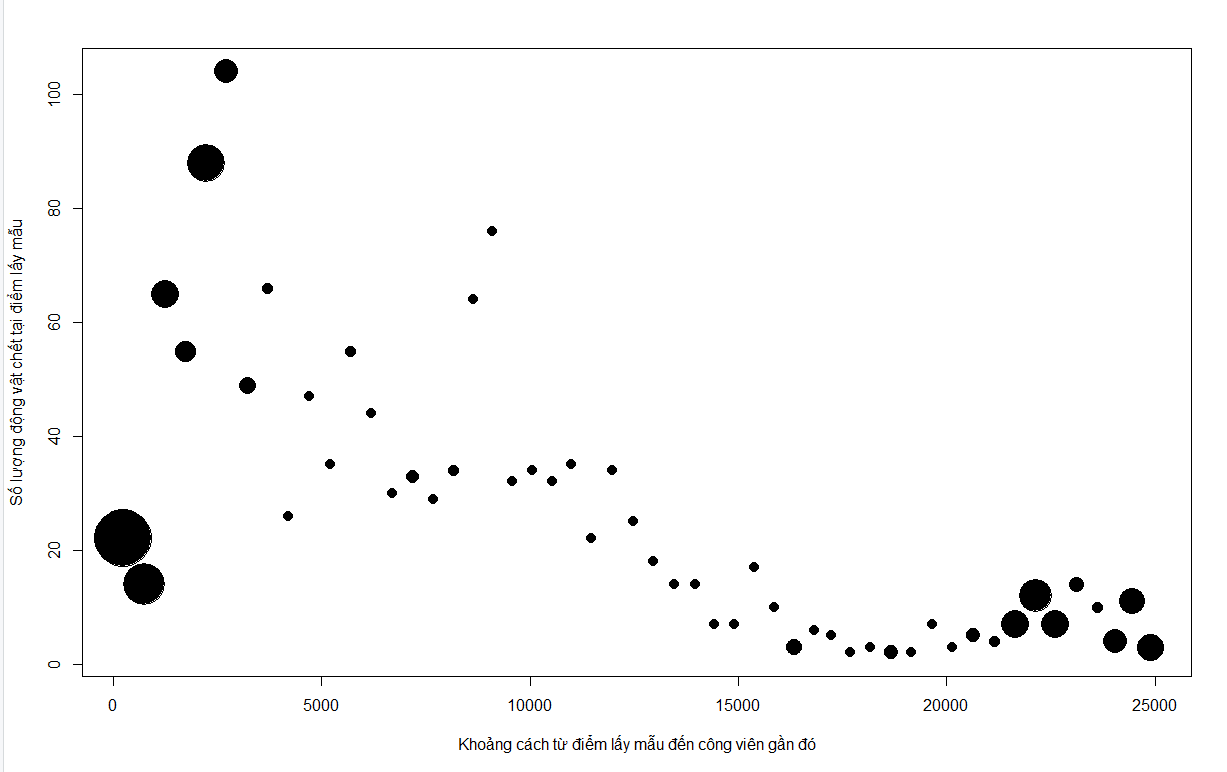
D\_Park là khoảng cách từ điểm lấy mẫu đến công viên gần đó.

Xây dựng plot với x=D.PARK và y=TOT.N

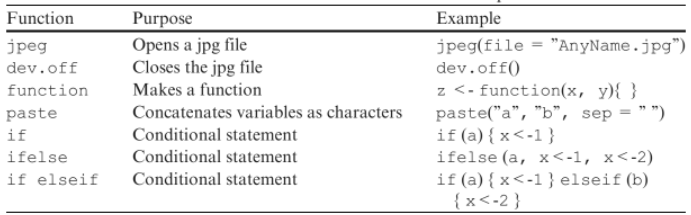




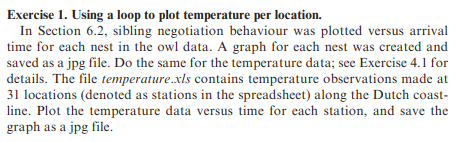


**Chương 6: Loops and Functions**

Chương này sẽ hướng dẫn ta xây dựng các hàm và các vòng lặp, cách đóng mở file jpg,…

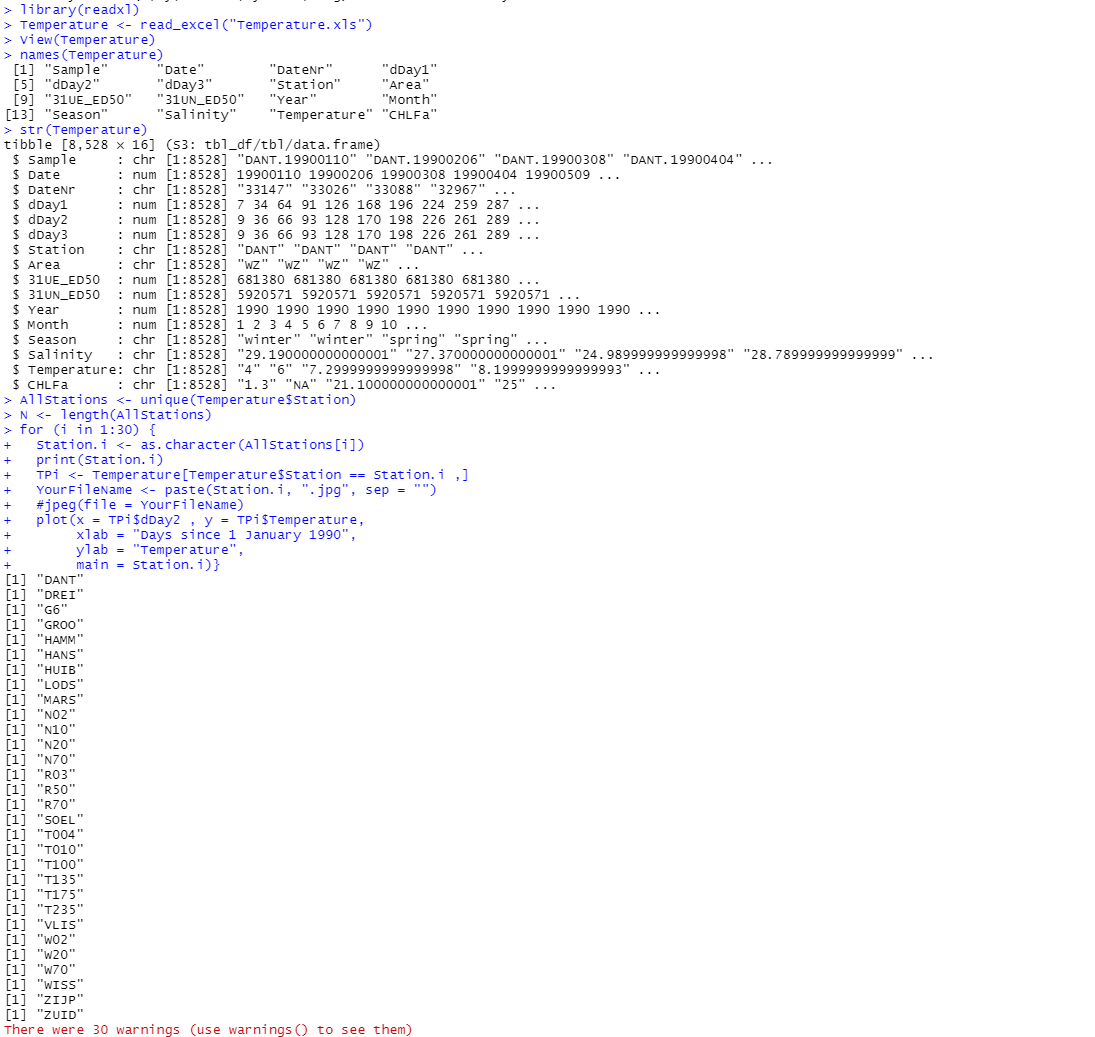
****

|  |  |  |
| --- | --- | --- |
| **FUNCTION** | **PURPOSE** | **EXAMPLE** |
| **jpeg** | Mở một tệp jpg | jpeg(file=’’AnyName.jpg’’) |
| **dev.off** | Đóng tập tin jpg | Dev.off() |
| **function** | Thực hiện một chức năng | Z <- function(x,y){} |
| **paste** | Nối các biến dưới dạng ký tự | paste(’’a’’, ’’b’’, sep = ’’ ’’) |
| **if**  **ifelse**  **if elseif** | Câu lệnh điều kiện | if(a) {x<-1}  ifelse(a, x<-1, x <-2)  if(a) {x<-1} elseif(b) {x<-2} |

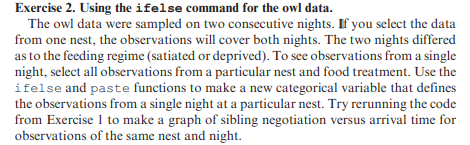
****

**DỊCH:**

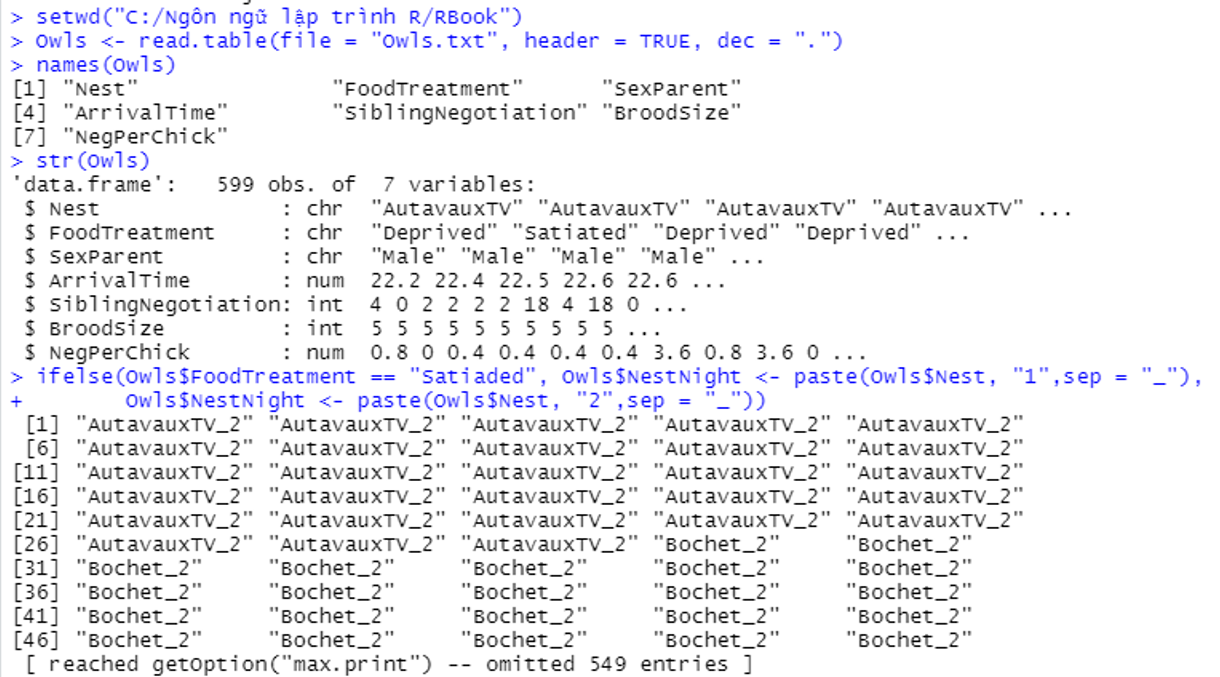
Trong Phần 6.2, hành vi thương lượng anh chị em được lập kế hoạch so với việc đến thời gian cho mỗi tổ trong dữ liệu cú. Một biểu đồ cho mỗi tổ đã được tạo và được lưu dưới dạng tệp jpg. Làm tương tự đối với dữ liệu nhiệt độ; xem Bài tập 4.1 cho thông tin chi tiết. Tập Temperature.xls chứa các quan sát nhiệt độ được thực hiện tại 31 vị trí (được ký hiệu là các trạm trong bảng tính) dọc theo đường bờ biển Hà Lan. Vẽ biểu đồ dữ liệu nhiệt độ so với thời gian cho từng trạm và lưu biểu đồ dưới dạng tệp jpg.

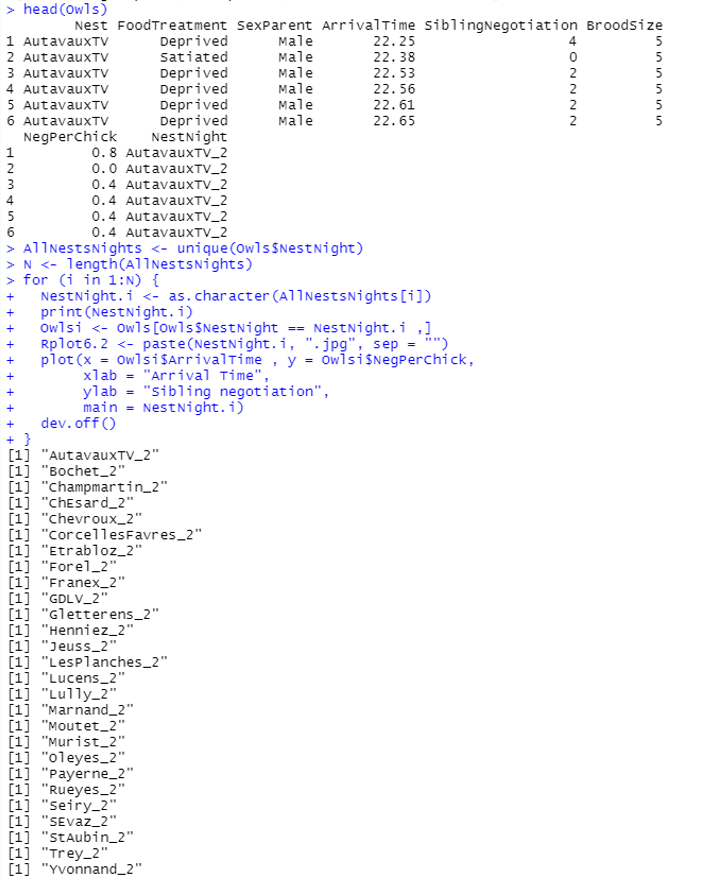
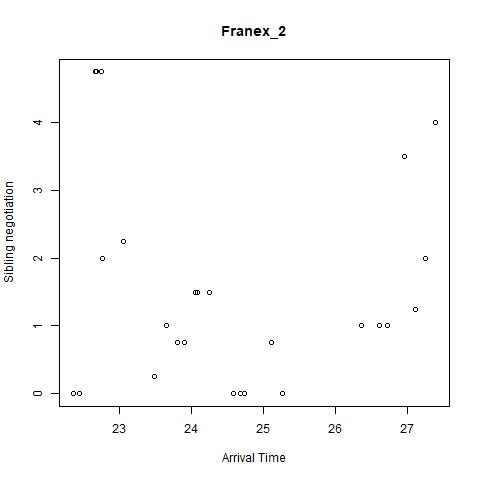




****

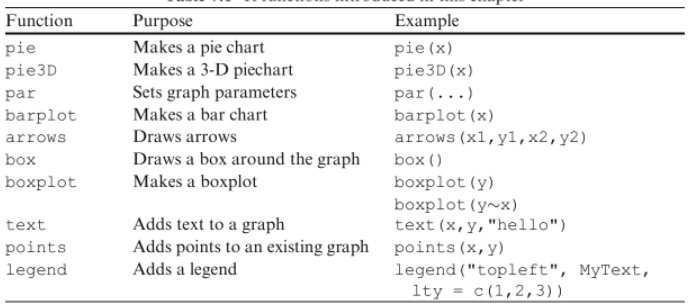
**DỊCH:**

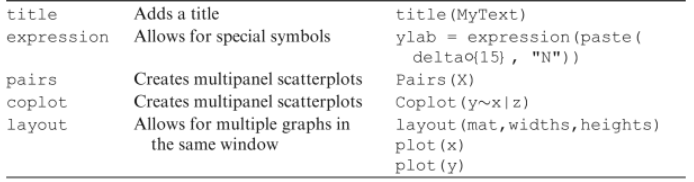
Dữ liệu về cú được lấy mẫu trong hai đêm liên tiếp. Nếu bạn chọn dữ liệu từ một tổ, các quan sát sẽ bao gồm cả hai đêm. Hai đêm khác nhau đối với chế độ cho ăn (ăn no hoặc thiếu thốn). Để xem các quan sát từ một ban đêm, chọn tất cả các quan sát từ một tổ cụ thể và xử lý thức ăn. Sử dụng if else và dán các hàm để tạo một biến phân loại mới xác định những quan sát từ một đêm duy nhất tại một tổ cụ thể. Thử chạy lại mã từ Bài tập 1 để tạo biểu đồ thương lượng giữa anh chị em với thời gian đến quan sát về cùng một tổ và đêm. 

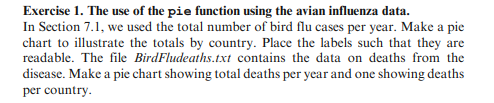
**Chapter 7 : Graphing Tools**

Chương này giới thiệu thêm về các công cụ vẽ đồ thị như biểu đồ tròn, biểu đồ thanh. Trong các phần từ 7.1 đến 7.4 thảo luận về các công cụ để phát hiện ngoại lệ boxplot và cleverland dotplot, 7.5 đến 7.7 thảo luận biểu đồ phân tán, và các công cụ nâng cao để hiển thị nhiều đồ thị trong một cửa sổ duy nhất được trình bày trong phần 7.8

****

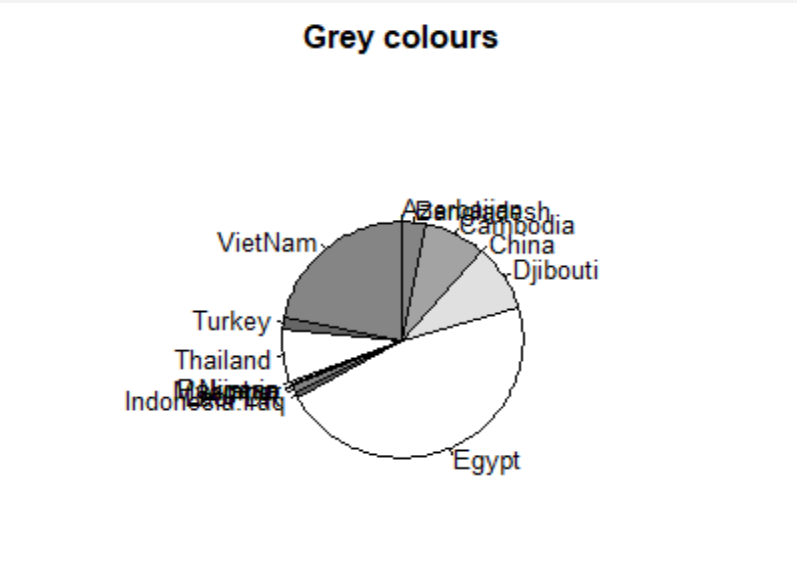
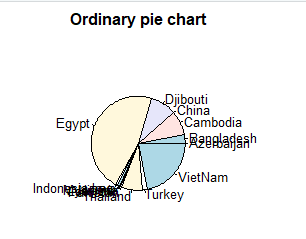
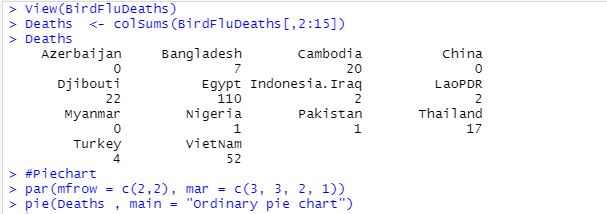
****

|  |  |  |
| --- | --- | --- |
| **FUNCTION** | **PURPOSE** | **EXAMPLE** |
| **pie** | Tạo biểu đồ hình tròn | pie(x) |
| **pie3D** | Tạo biểu đồ hình tròn 3-D | pie3D(x) |
| **par** | Thiết lập các thông số đồ thị | par(…) |
| **barplot** | Tạo một biểu đồ thanh | barplot(x) |
| **arrows** | Vẽ xung quanh | arrows(x1,y1,x2,y2) |
| **box** | Vẽ một hộp xung quanh biểu đồ | box(x) |
| **boxplot** | Vẽ biểu đồ boxplot | boxplot(y)  boxplot(y~x) |
| **text** | Thêm thử nghiệm vào biểu đồ | text(x,y,"hello") |
| **points** | Thêm điểm vào biểu đồ hiện có | points(x,y) |
| **legend** | Thêm một chú giải | legend("topleft", MyText,  lty = c(1,2,3)) |
| **title** | Thêm một tiêu đề | title(MyText) |
| **expression** | Cho phép các ký hiệu đặc biệt | ylab = expression(paste(  delta{15}, "N")) |
| **pairs** | Tạo ra nhiều bảng phân tán | Pairs(X) |
| **coplot** | Tạo ra nhiều bảng phân tán | Coplot(y~x|z) |
| **layout** | Cho phép nhiều đồ thị trong cùng một cửa sổ | layout(mat,widths,heights)  plot(x)  plot(y) |

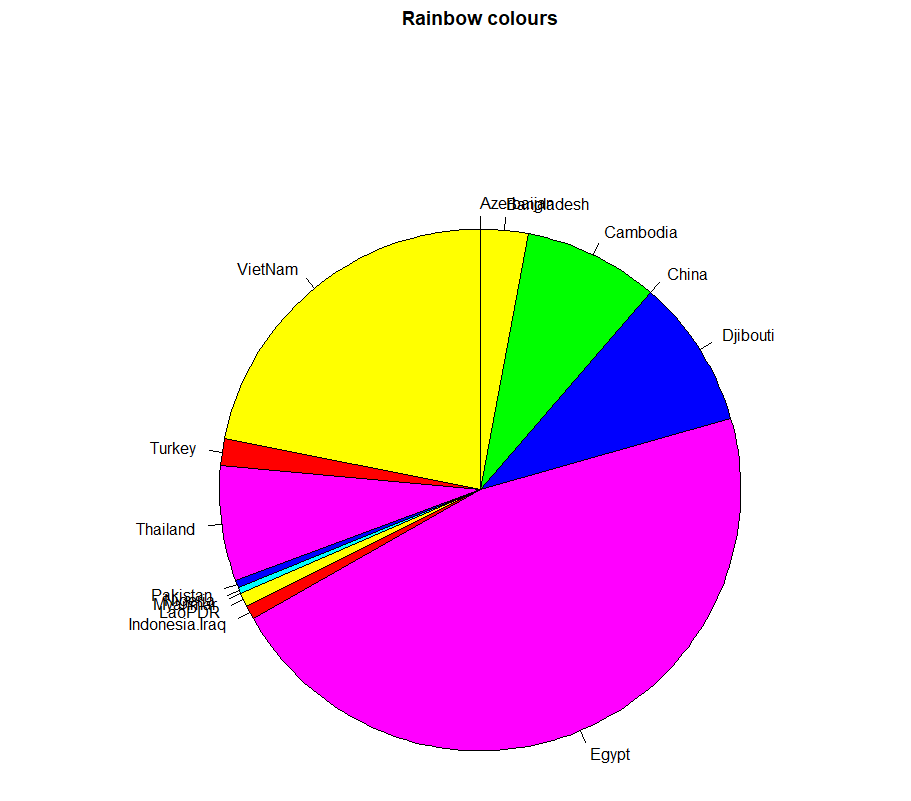
****

**DỊCH:**

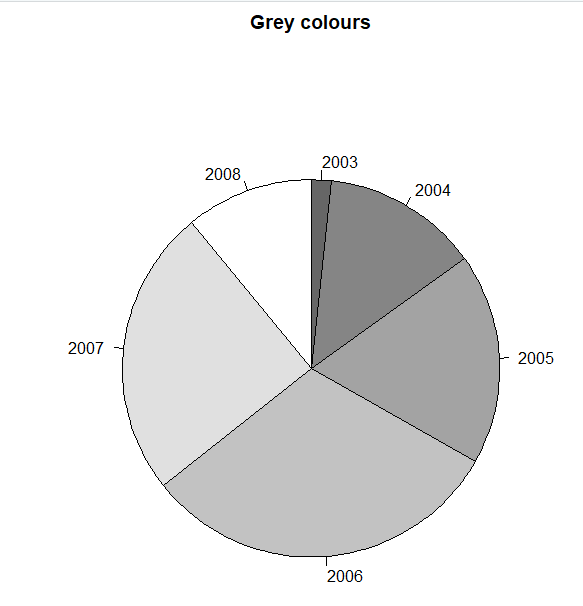
Trong Phần 7.1, chúng tôi đã sử dụng tổng số ca bệnh cúm gia cầm mỗi năm. Làm một chiếc bánh biểu đồ để minh họa tổng số theo quốc gia. Đặt các nhãn sao cho chúng có thể đọc được. Tệp BirdFludeaths.txt chứa dữ liệu về các trường hợp tử vong do dịch bệnh. Lập biểu đồ hình tròn thể hiện tổng số ca tử vong mỗi năm và một biểu đồ thể hiện số ca tử vong mỗi quốc gia.

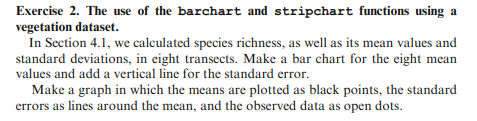








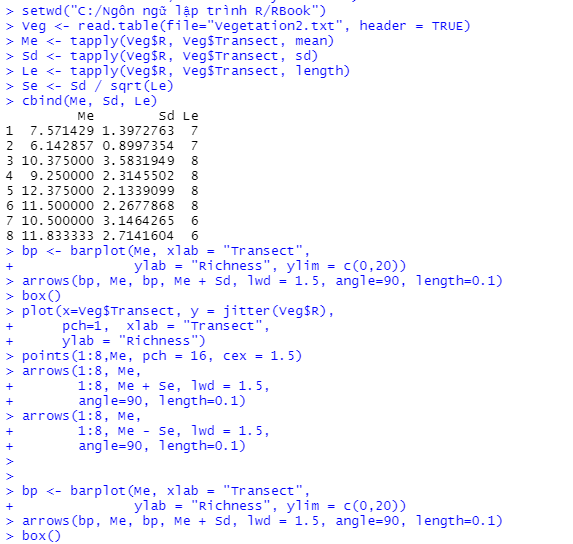
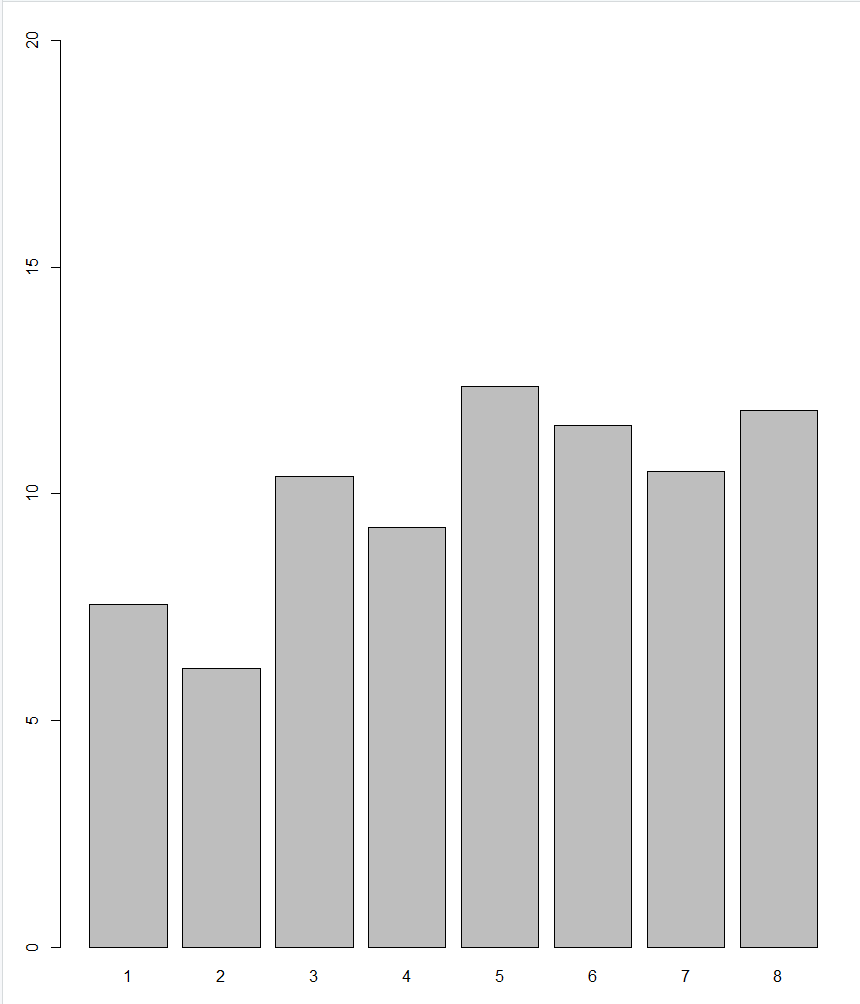
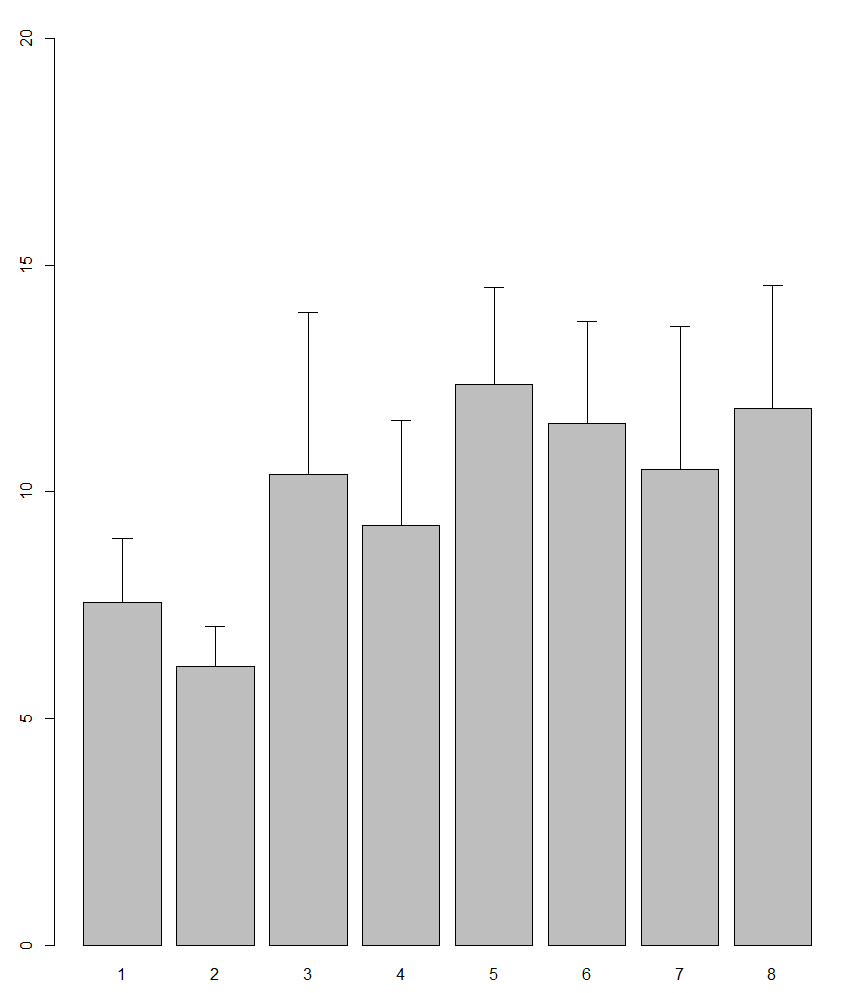
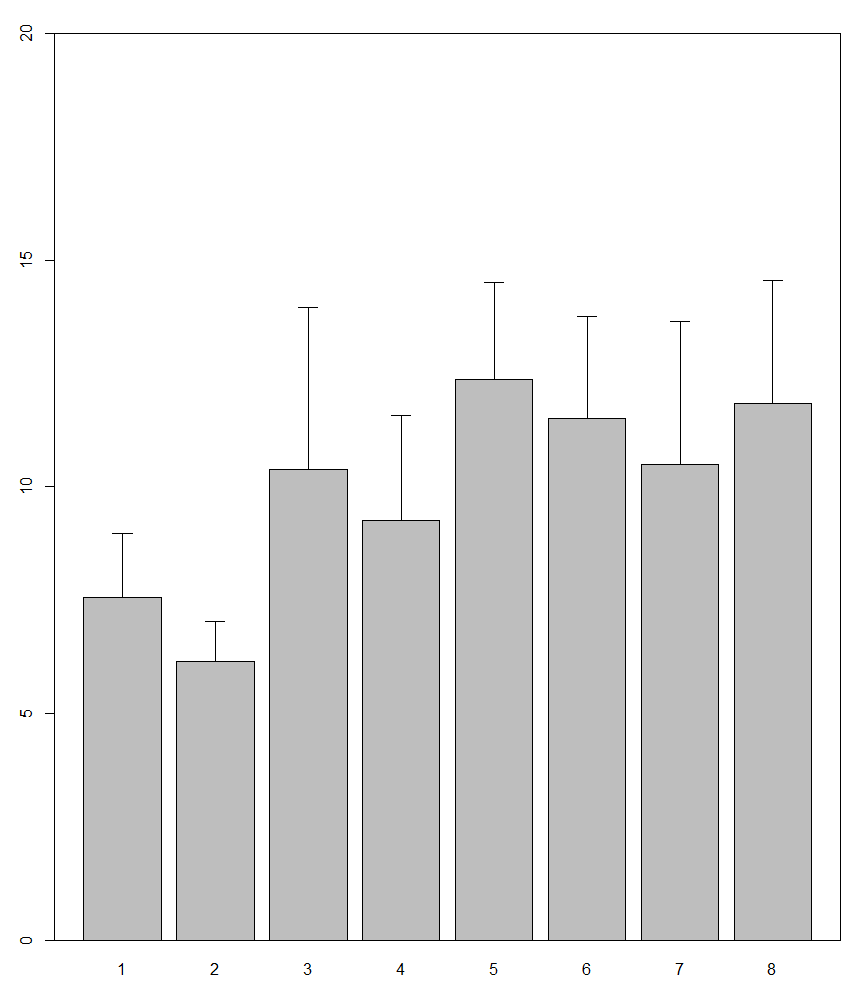


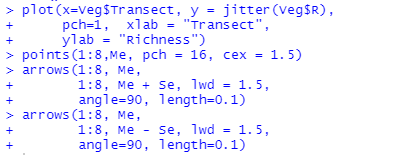
****

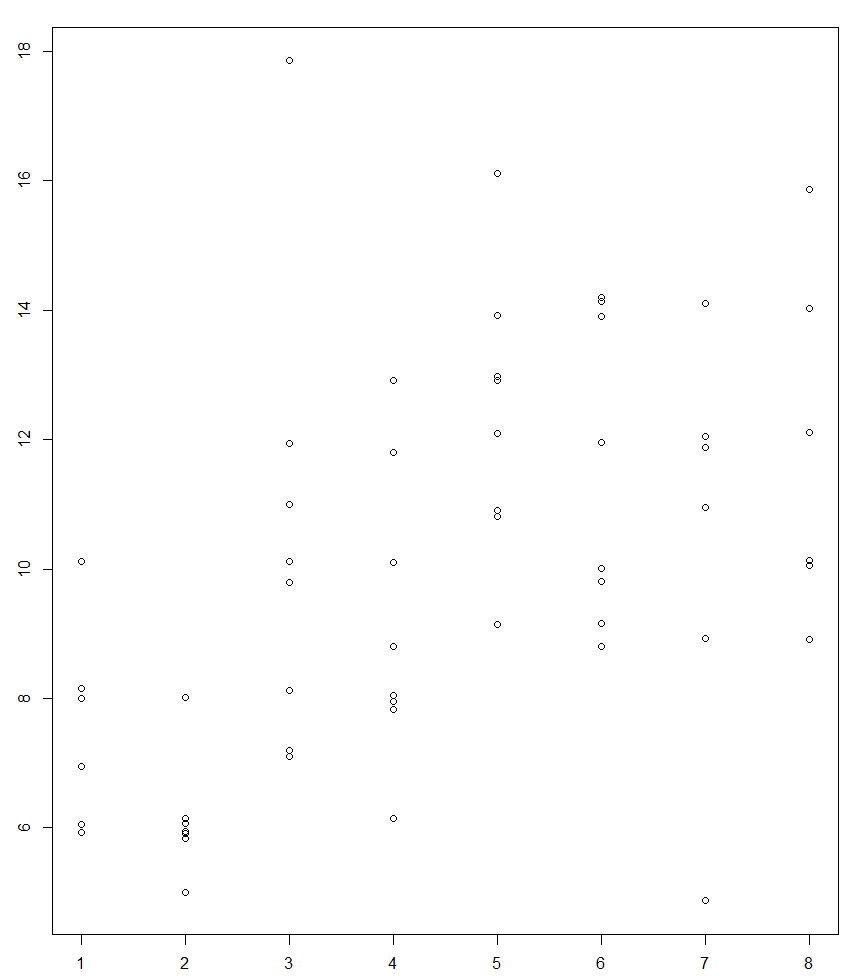
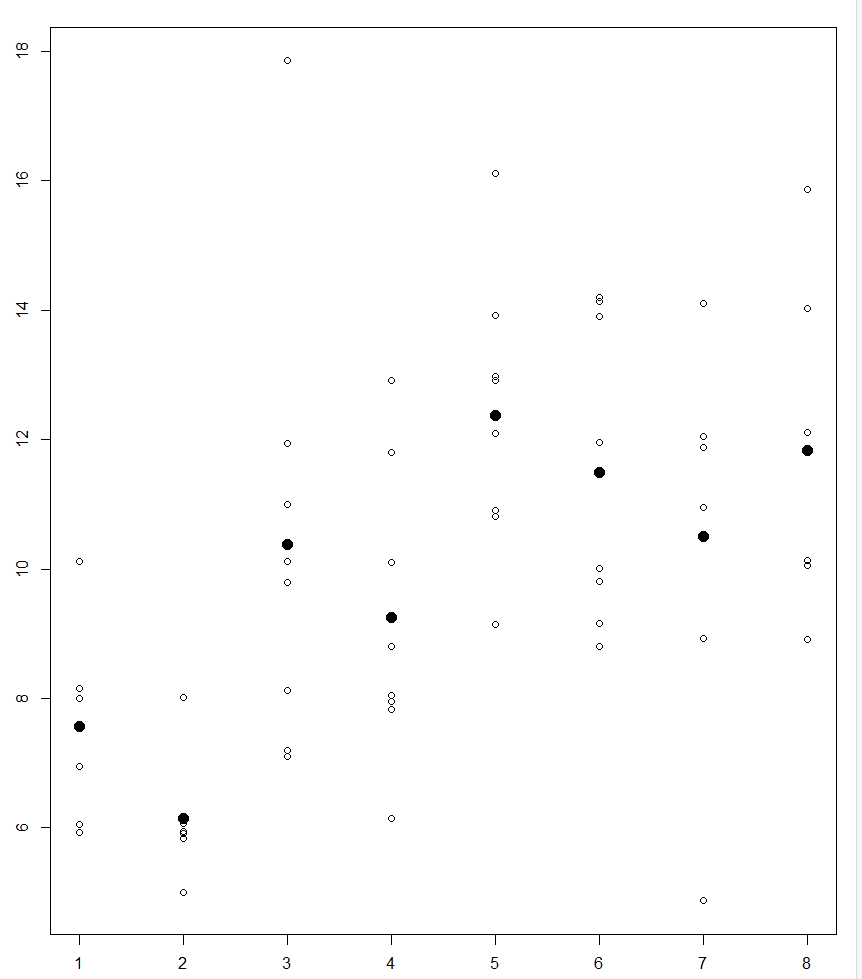
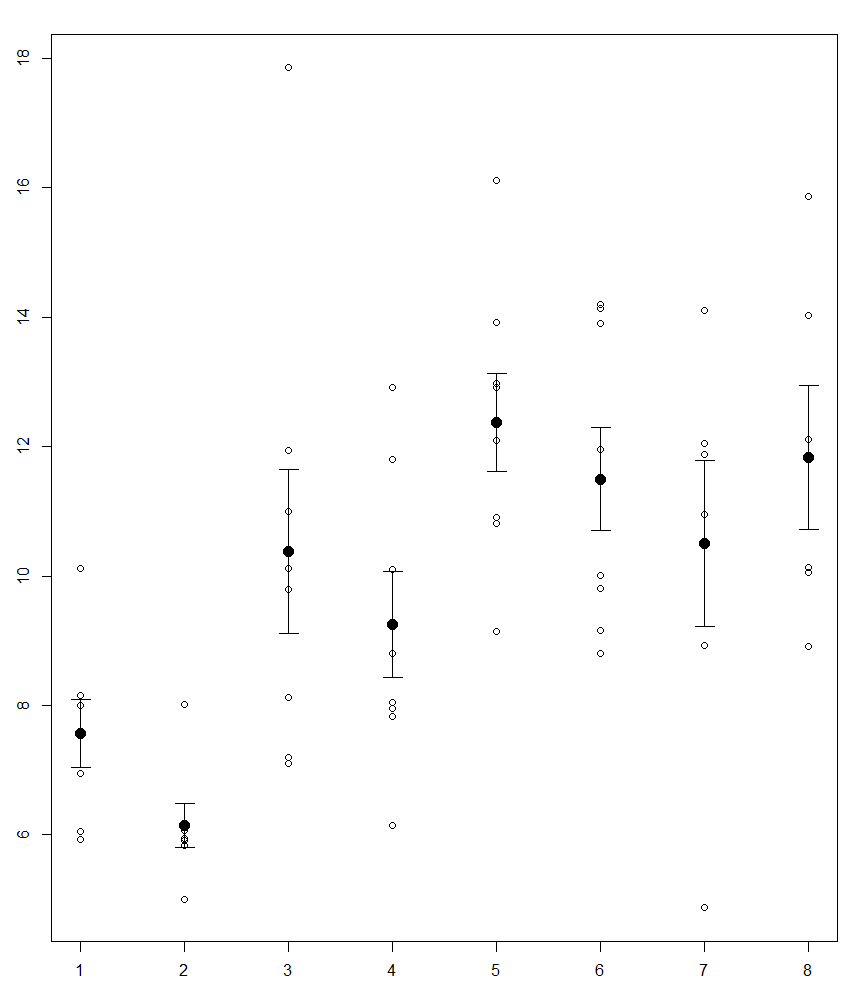
**DỊCH:**

Trong Phần 4.1, chúng tôi đã tính toán mức độ phong phú của các loài, cũng như giá trị trung bình của nó và độ lệch chuẩn, trong tám đường cắt ngang. Lập biểu đồ thanh cho tám giá trị trung bình và thêm một đường thẳng đứng cho lỗi tiêu chuẩn.

Lập biểu đồ trong đó các phương tiện được vẽ dưới dạng điểm đen, tiêu chuẩn lỗi dưới dạng các đường xung quanh giá trị trung bình và dữ liệu được quan sát dưới dạng các dấu chấm mở.

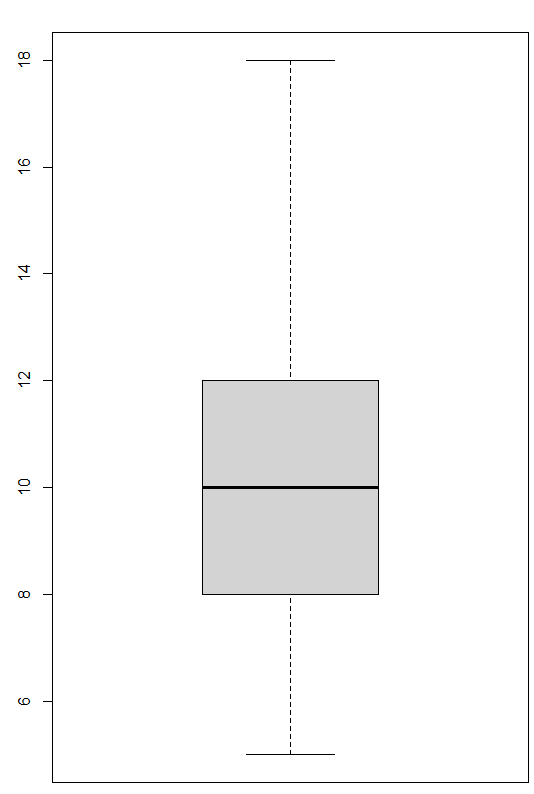
   



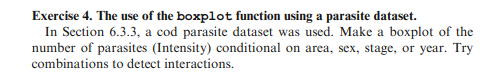
  

****

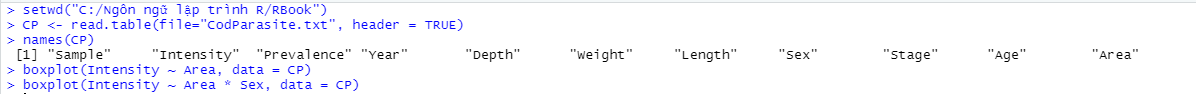
**DỊCH:**

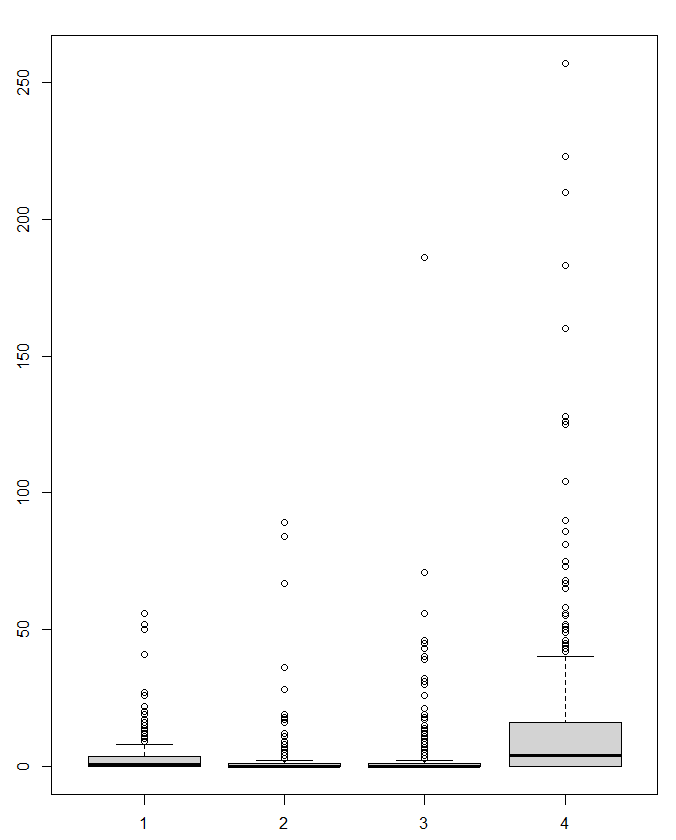
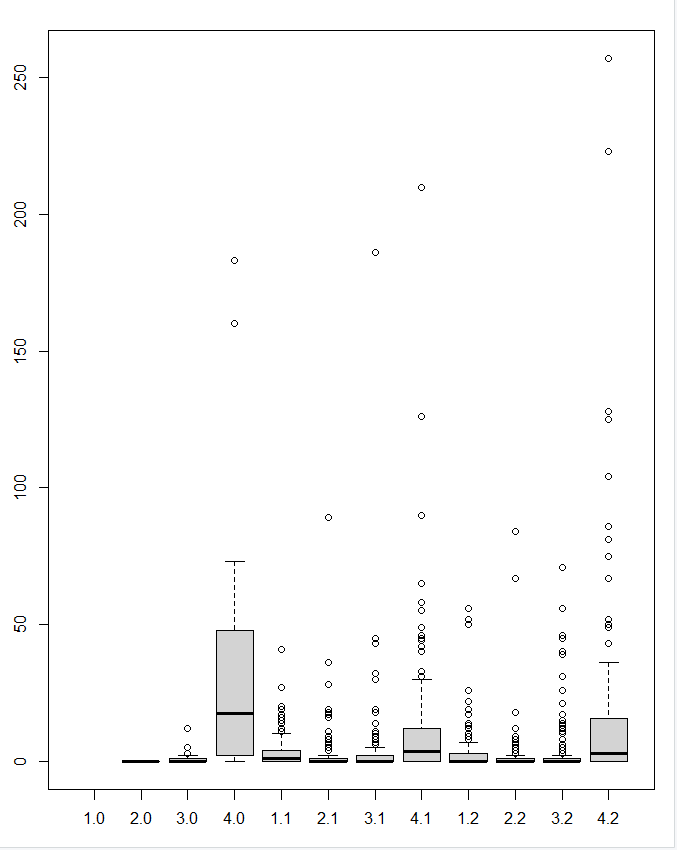
Sử dụng số liệu về thảm thực vật trong Bài tập 2, hãy lập biểu đồ hình hộp thể hiện sự phong phú các giá trị.

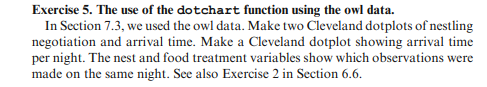


****

**DỊCH:**

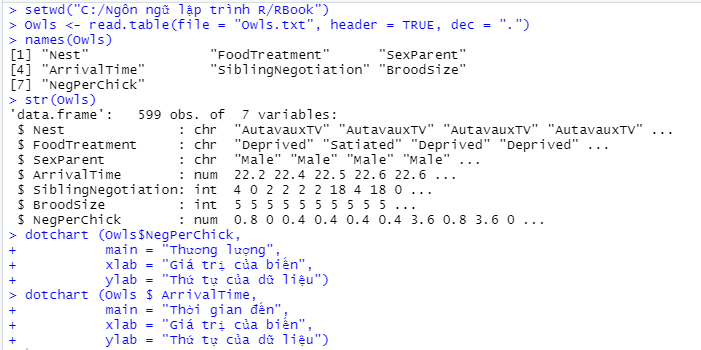
Trong Phần 6.3.3, một tập dữ liệu về ký sinh trùng cá tuyết đã được sử dụng. Tạo một ô vuông của số lượng ký sinh trùng (Cường độ) tùy theo khu vực, giới tính, giai đoạn hoặc năm. Thử Các kết hợp để phát hiện các tương tác. 

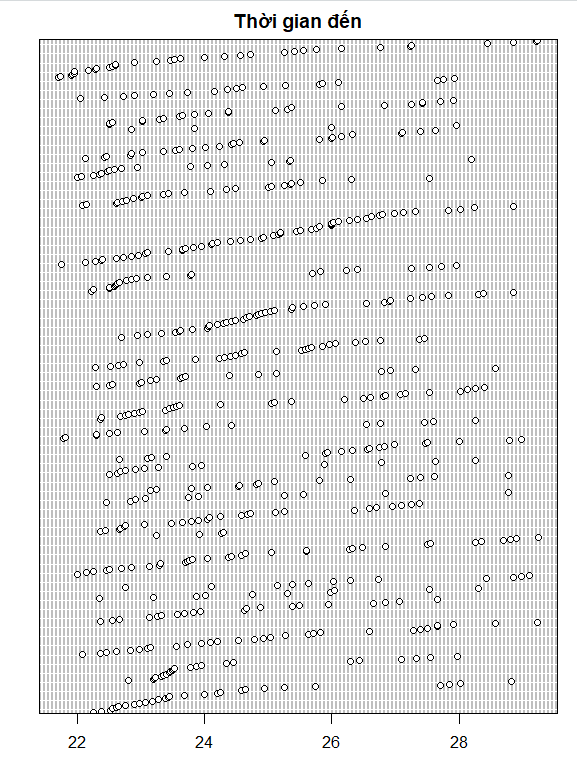
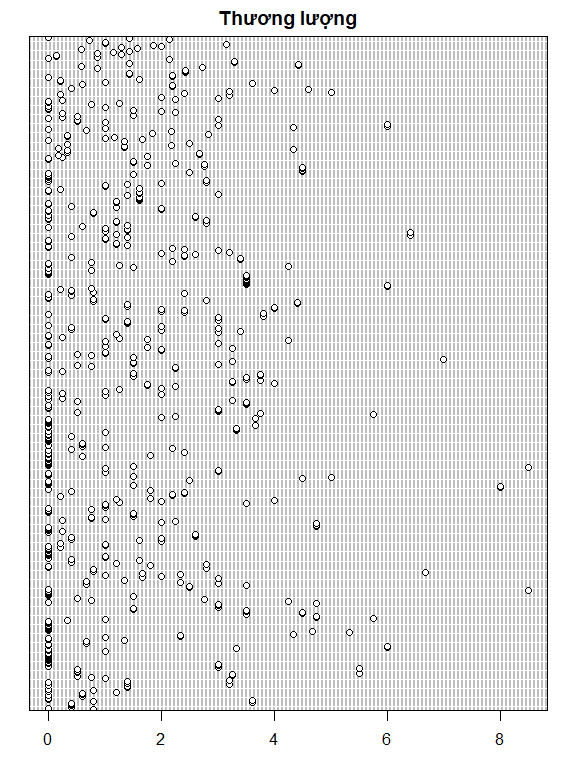
 

****

**DỊCH:**

Trong Phần 7.3, chúng tôi đã sử dụng dữ liệu về con cú. Tạo hai khe chấm Cleveland lồng vào nhau thương lượng và thời gian đến. Thực hiện một dấu chấm Cleveland hiển thị thời gian đến mỗi đêm. Các biến số xử lý tổ và thức ăn cho thấy những quan sát nào đã được thực hiện trong cùng một đêm. Xem thêm Bài tập 2 trong Mục 6.6.

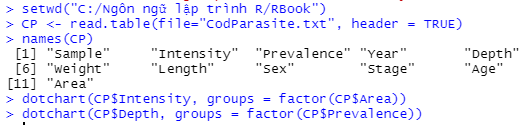


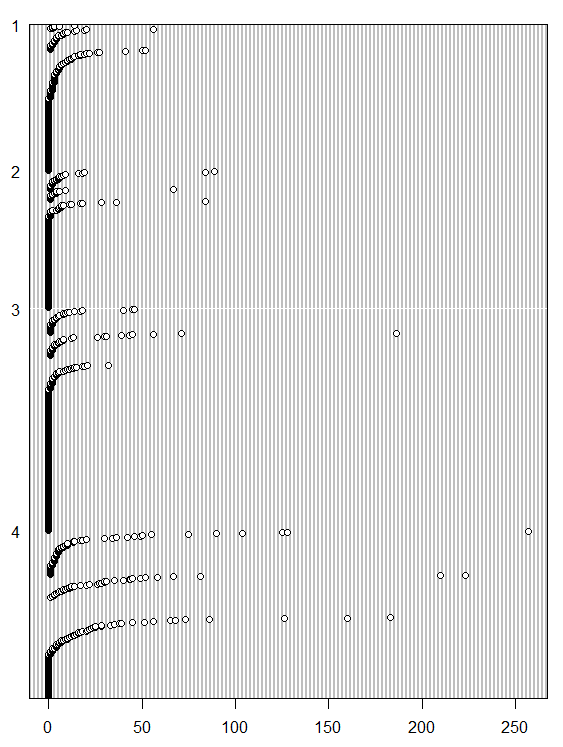
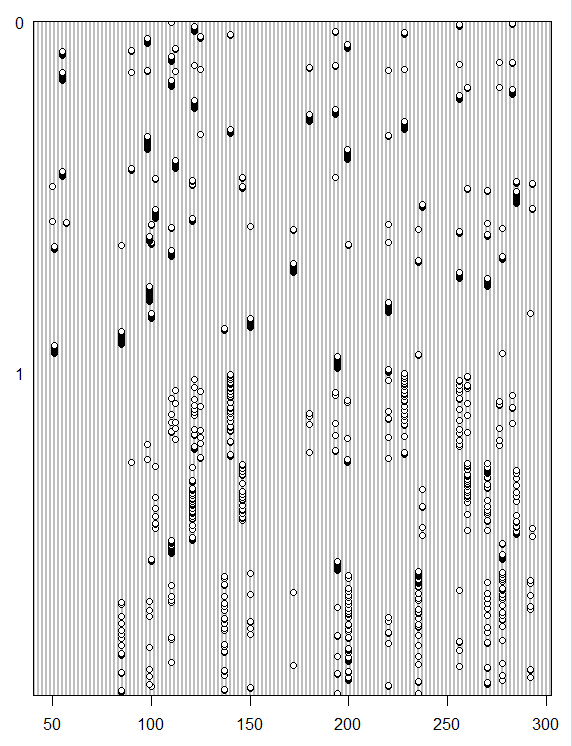
 

** **

**DỊCH:**

Tạo một dotplot Cleveland cho dữ liệu về ký sinh trùng được sử dụng trong Bài tập 4. Sử dụng số lượng ký sinh trùng (Cường độ) và nhóm các quan sát theo khu vực, giới tính, giai đoạn hoặc theo năm. Tạo một dotplot Cleveland hiển thị độ sâu và nhóm các quan sát theo tỷ lệ hiện mắc.

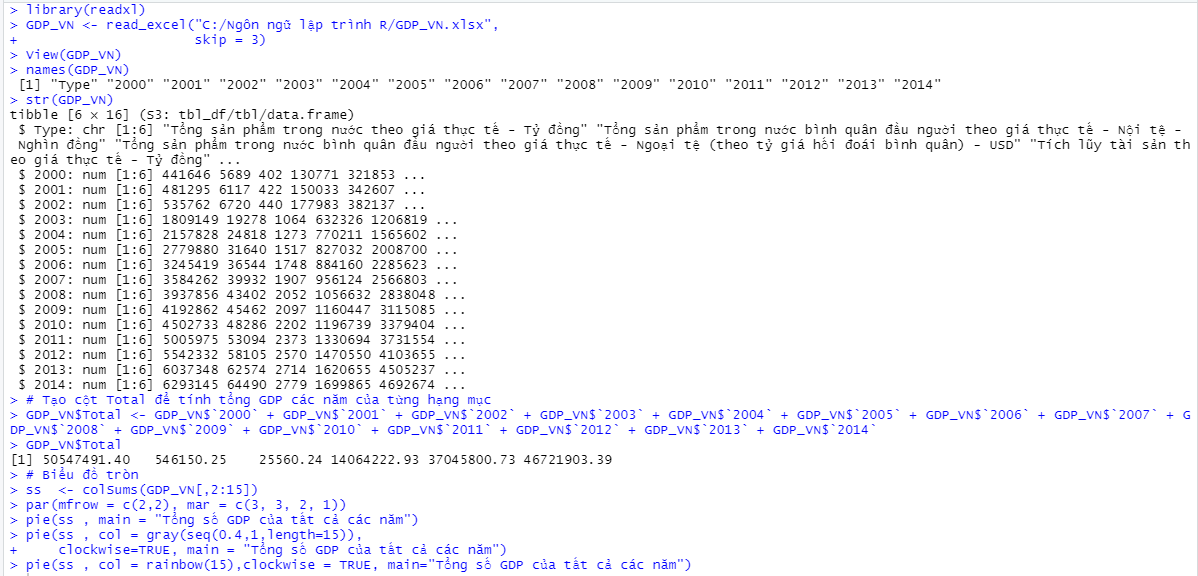
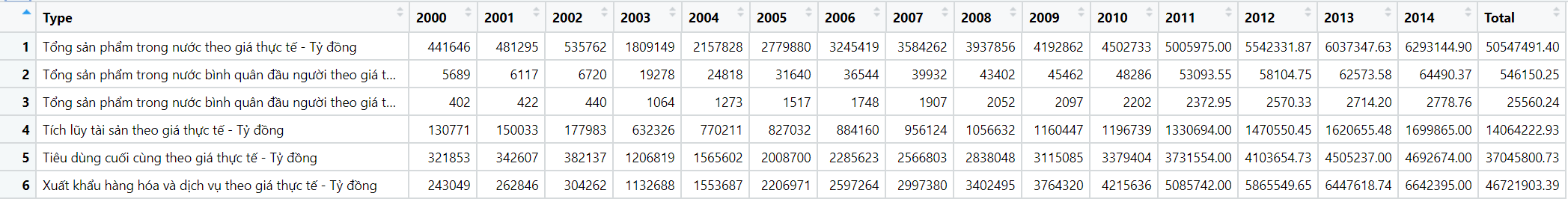
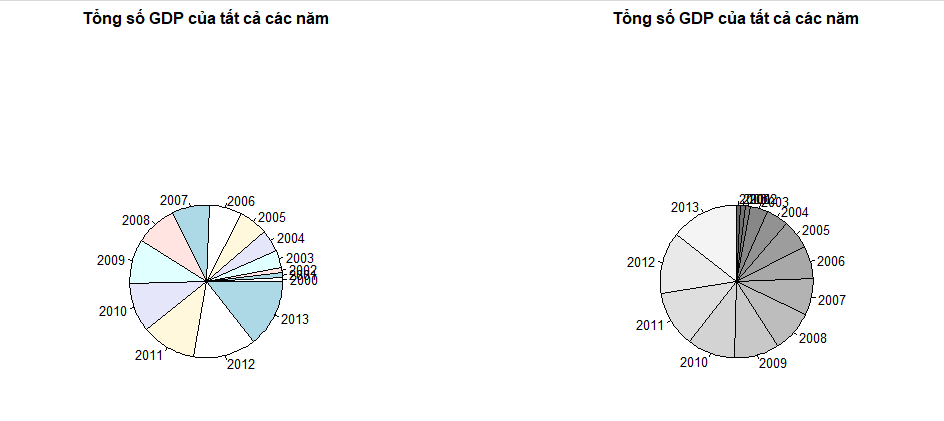


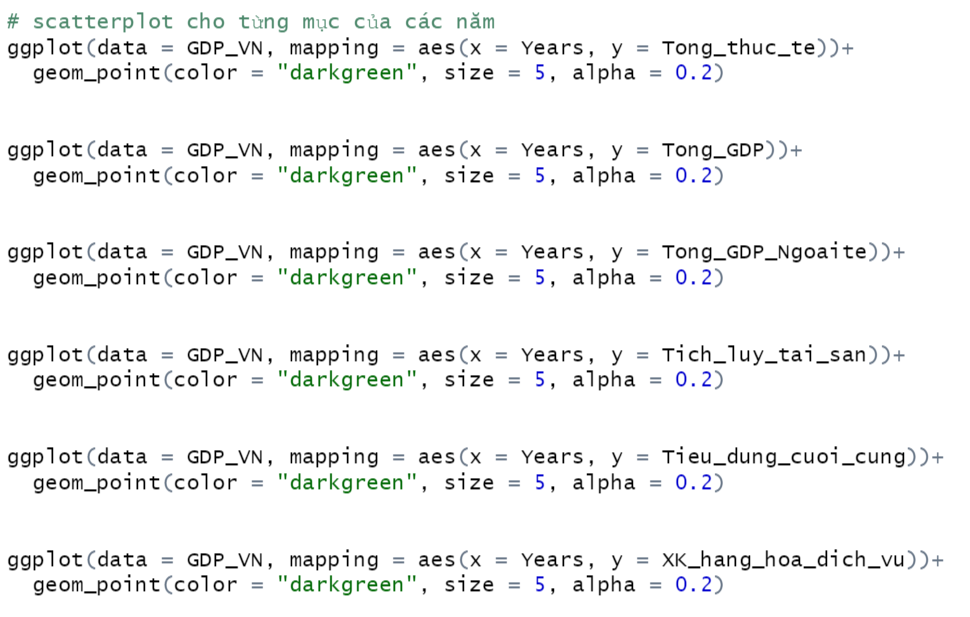
 

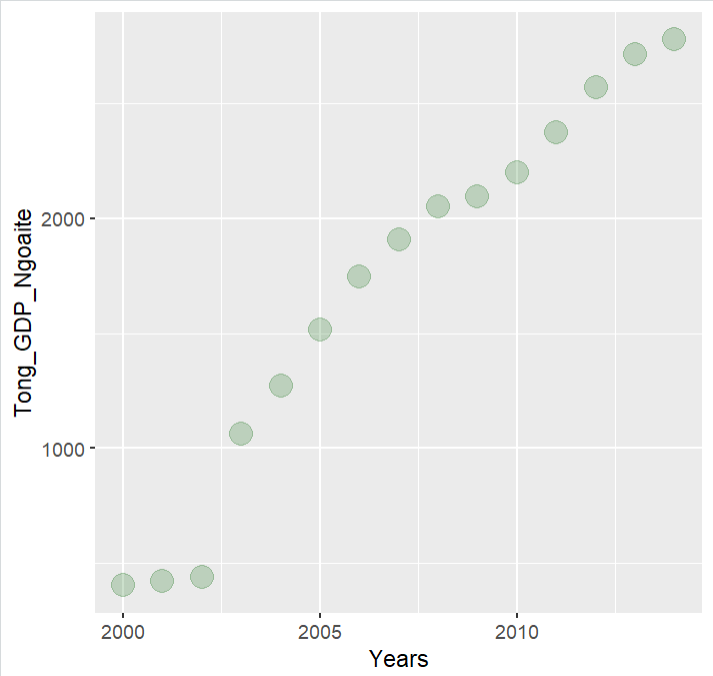
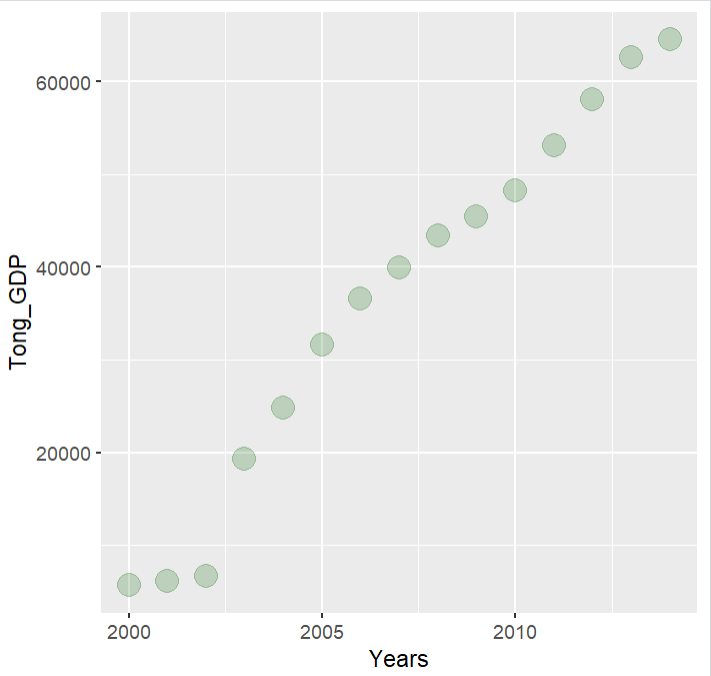
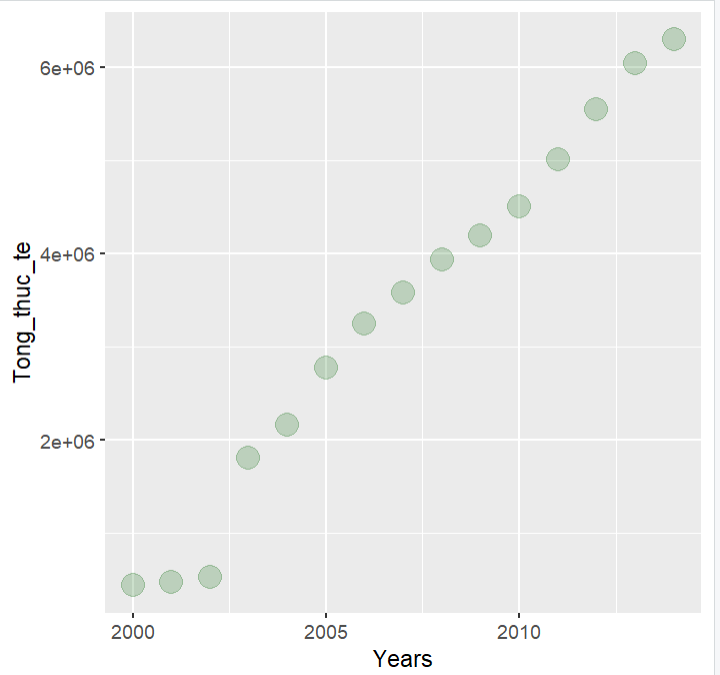
**BIỂU ĐỒ**

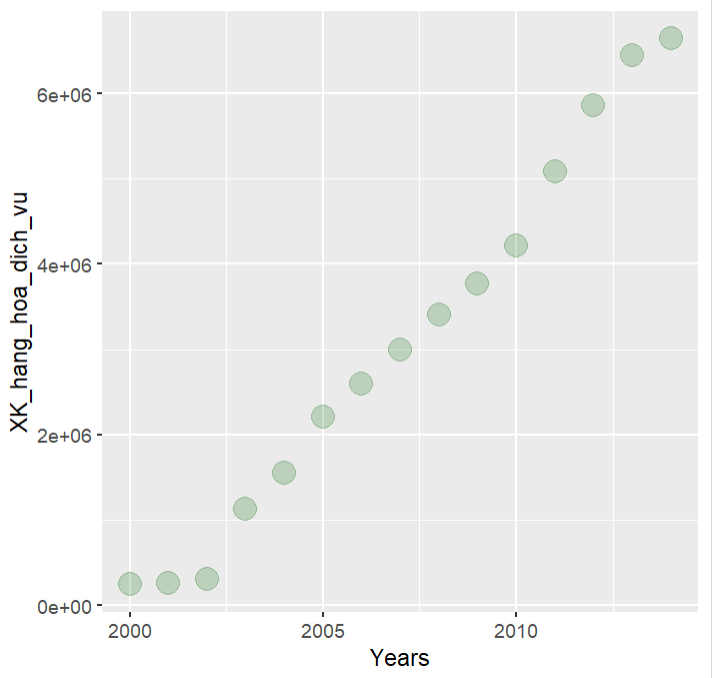
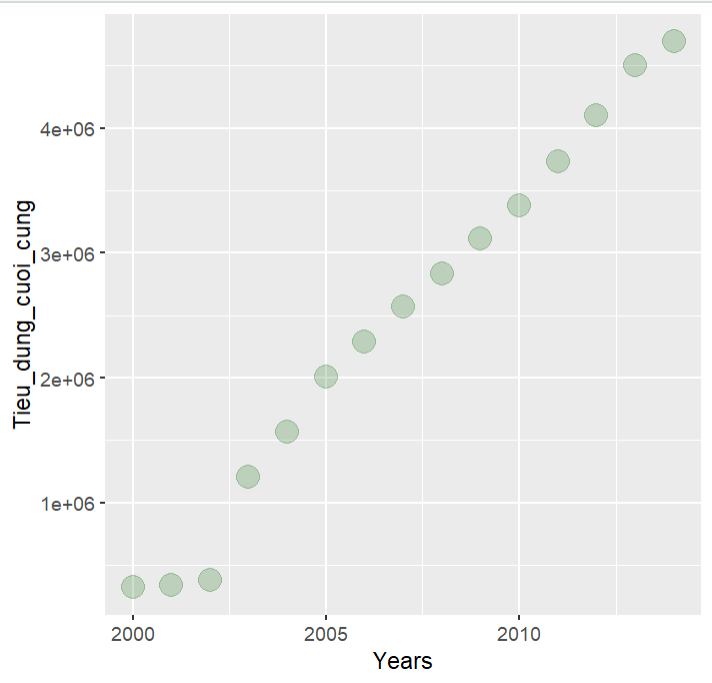
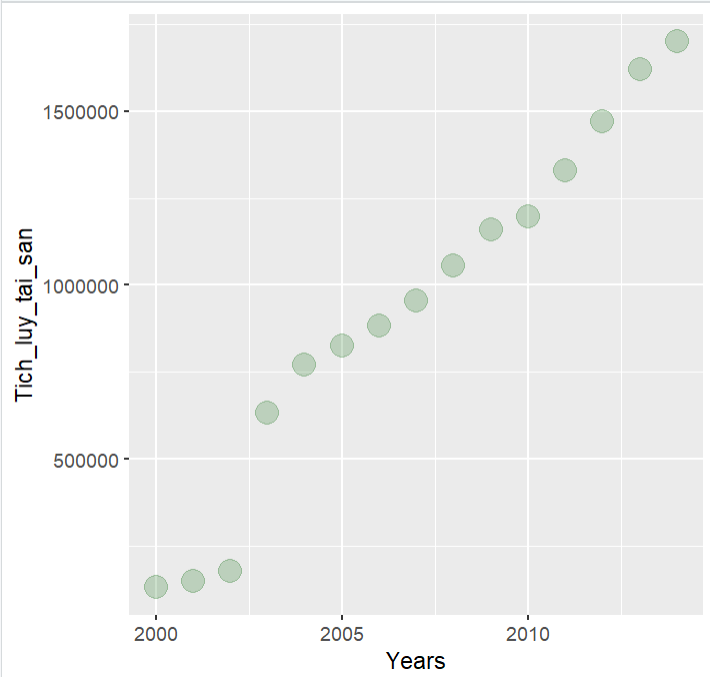
Đầu tiên ta cần xử lí và làm sạch dữ liệu

* Xóa bỏ 3 hàng đầu tiên
* Xóa bỏ các giá trị rỗng

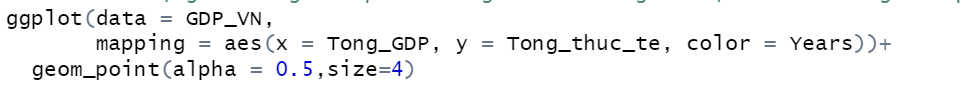
****

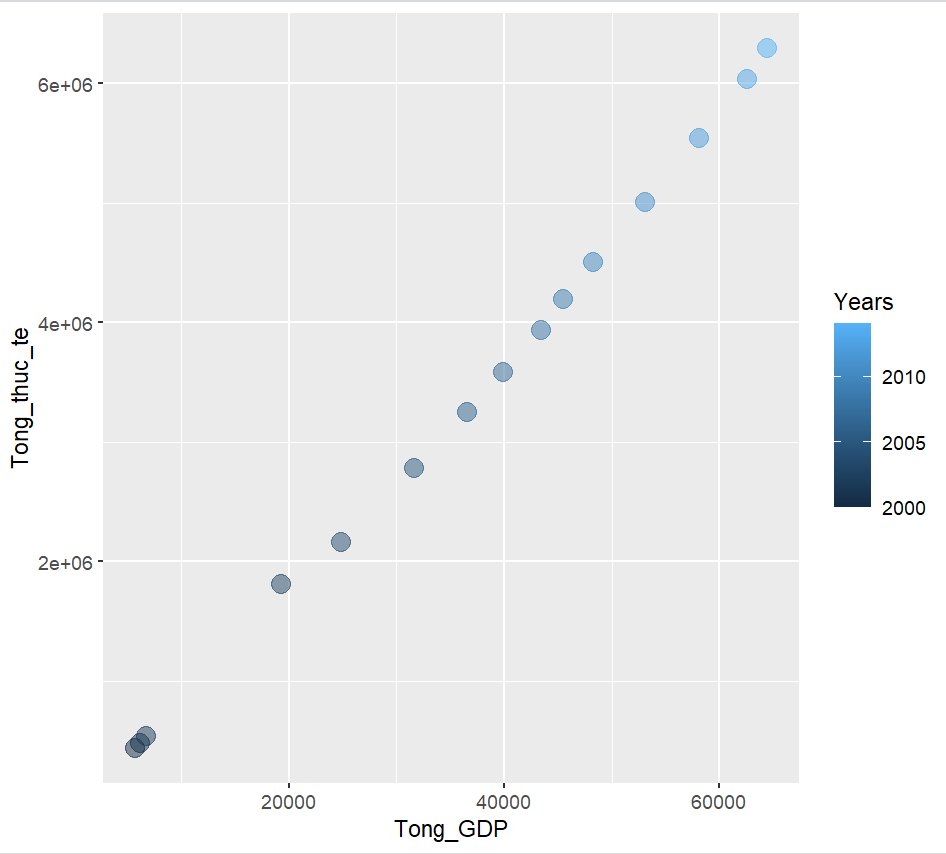
****

****

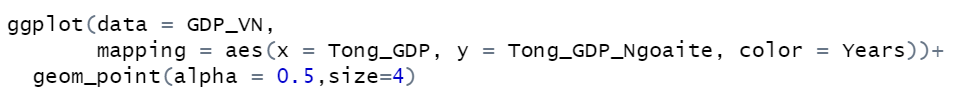
****

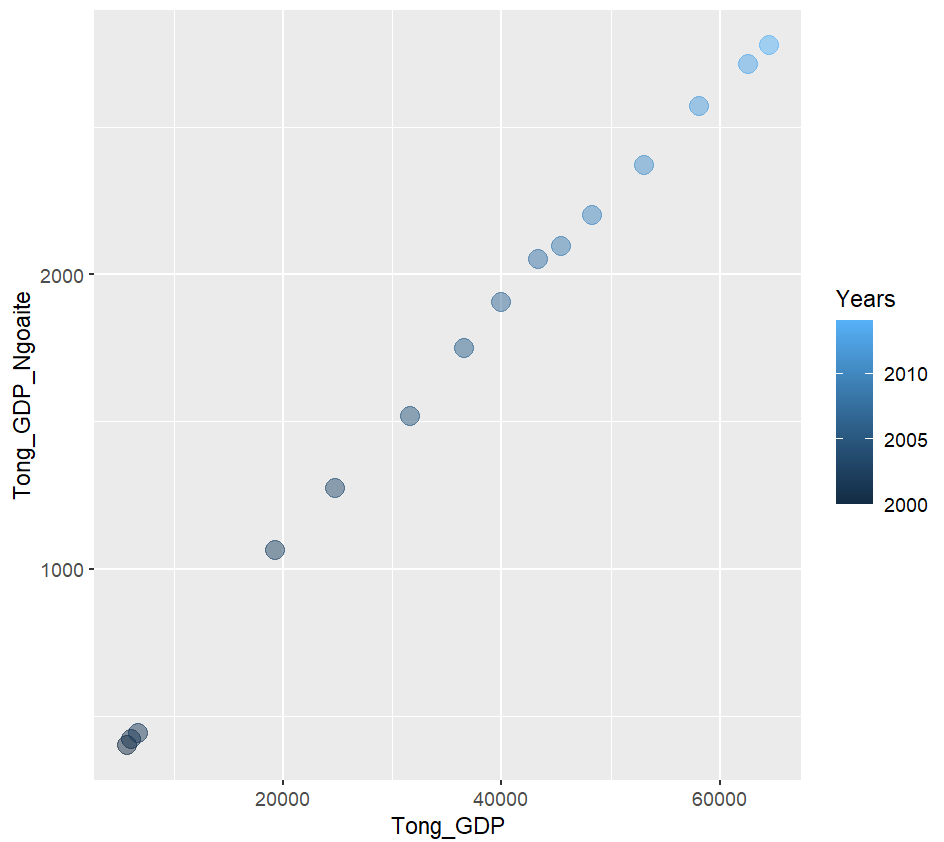
**Mối liên hệ giữa tổng sản phẩm trong nước theo giá thực tế với tổng sản phẩm trong nước bình quân đầu người hàng năm**



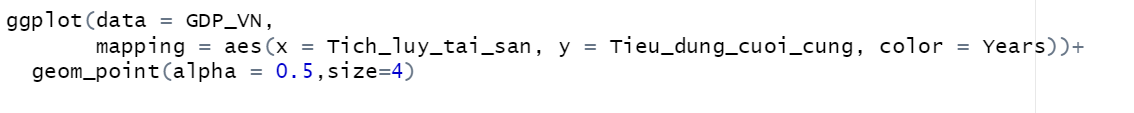
****

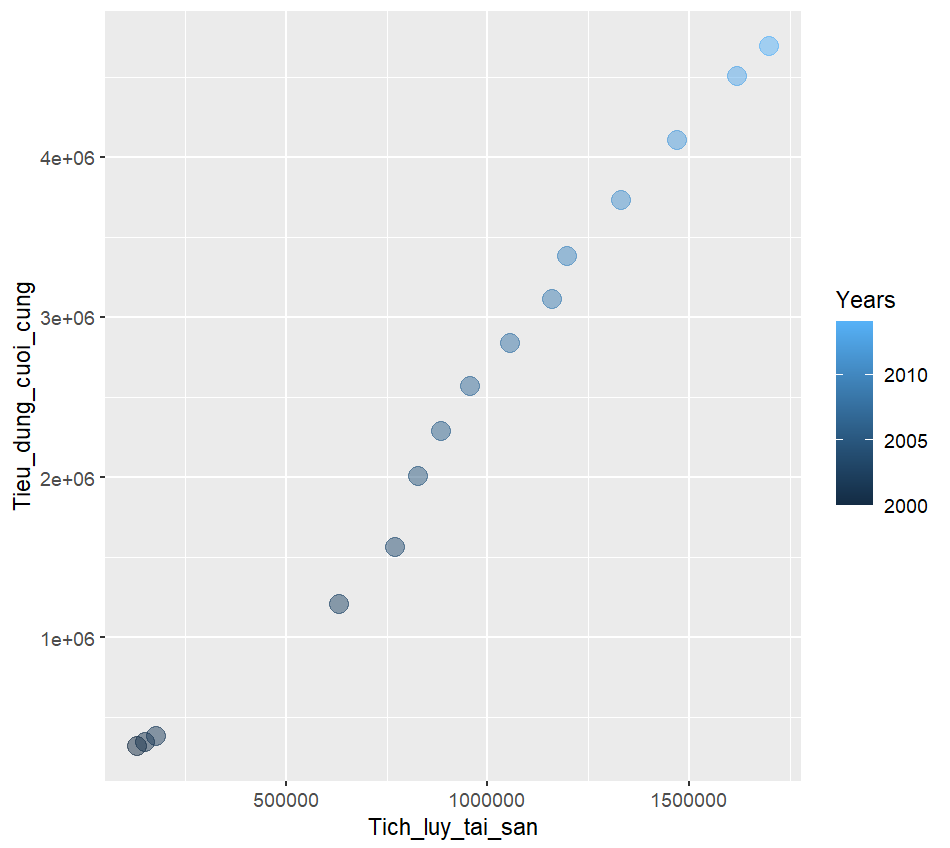
**Mối liên hệ giữa tổng sản phẩm trong nước bình quân đầu người (trong nước) với tổng sản phẩm trong nước bình quân đầu người (Ngoại tệ) hàng năm**

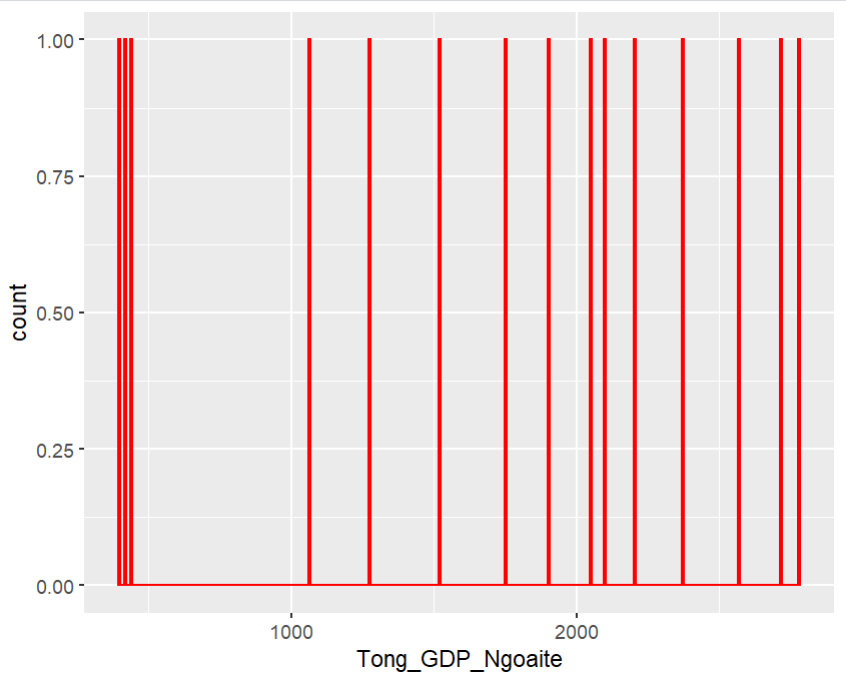
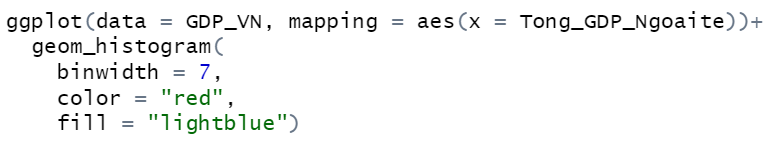
****

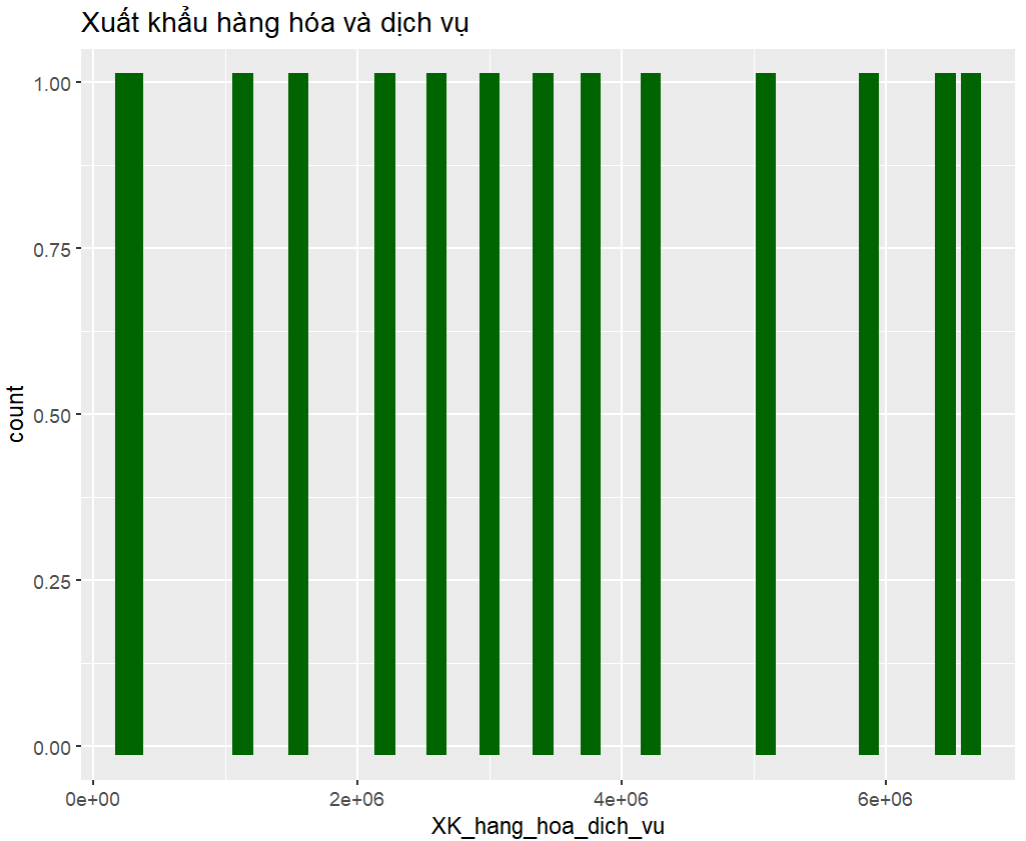
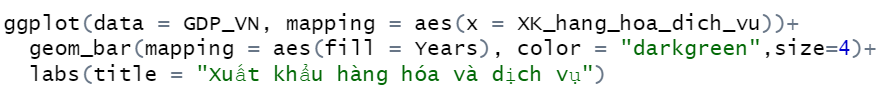
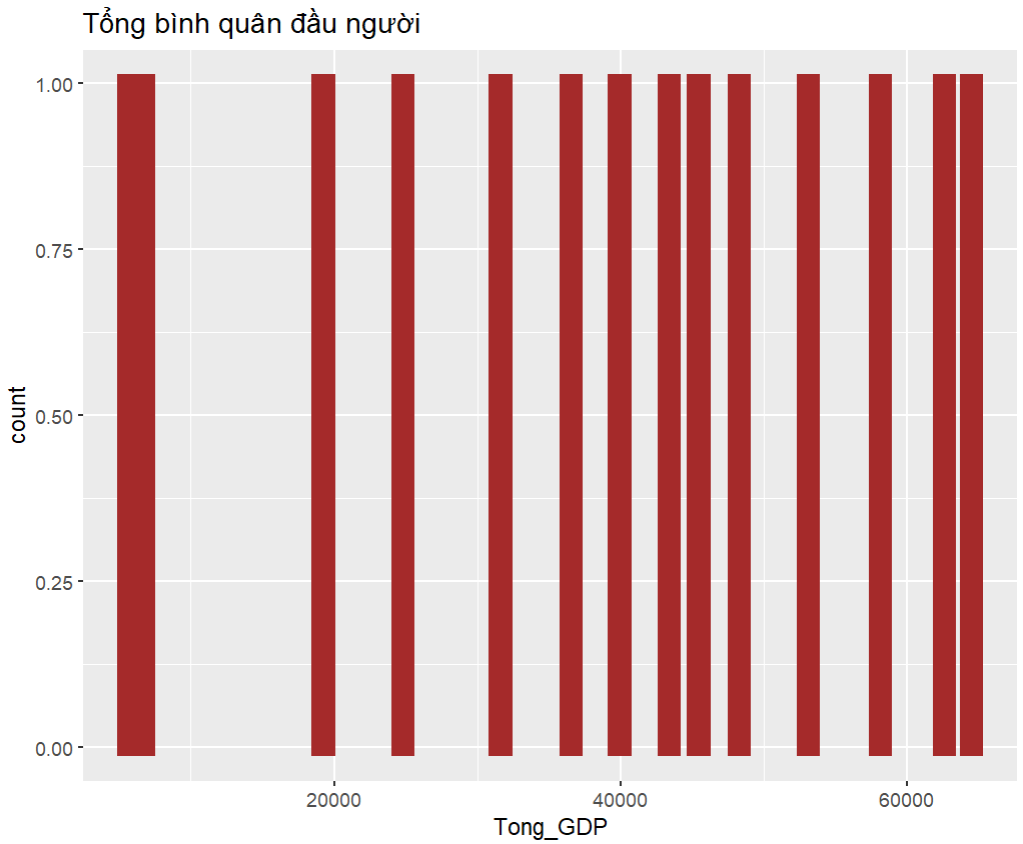
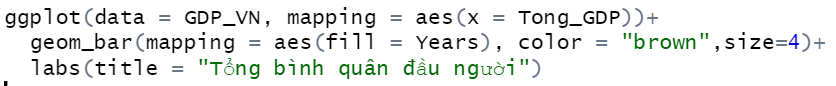
****

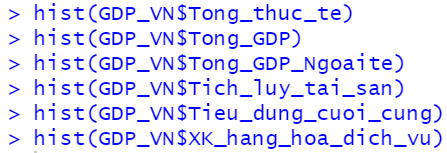
**Mối liên hệ giữa tích lũy tài sản và tiêu dùng theo giá thực tế hàng năm**

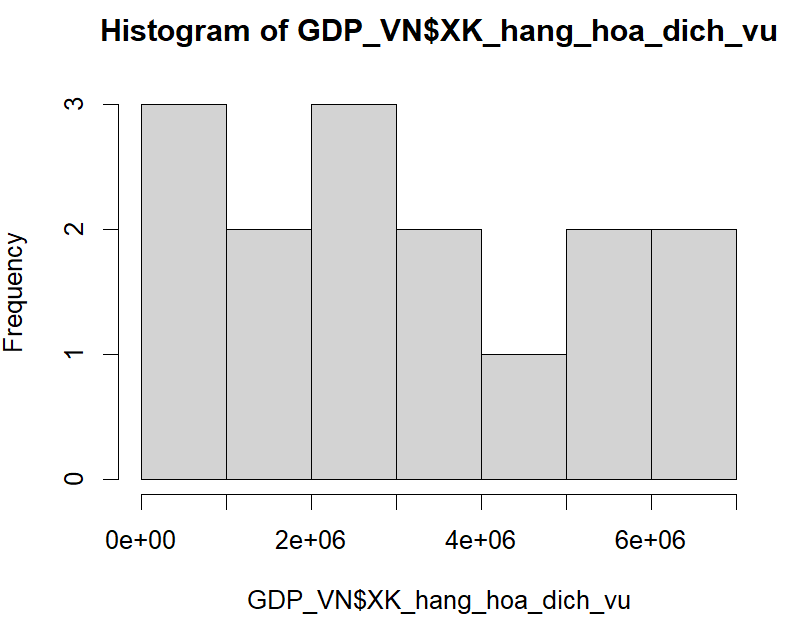
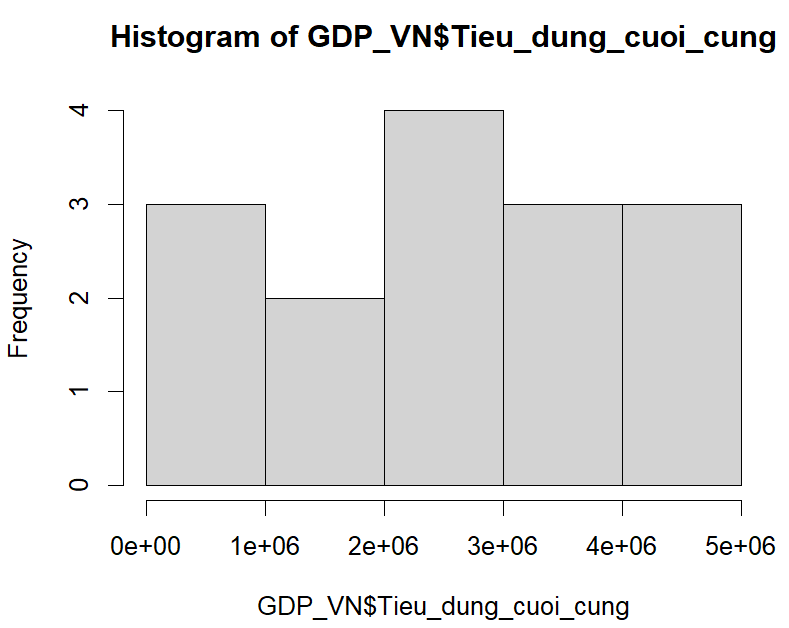
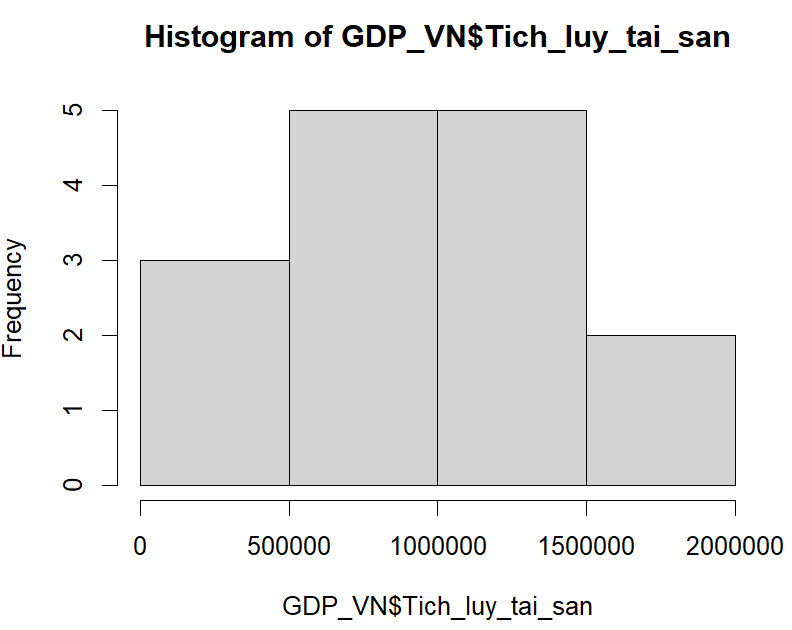
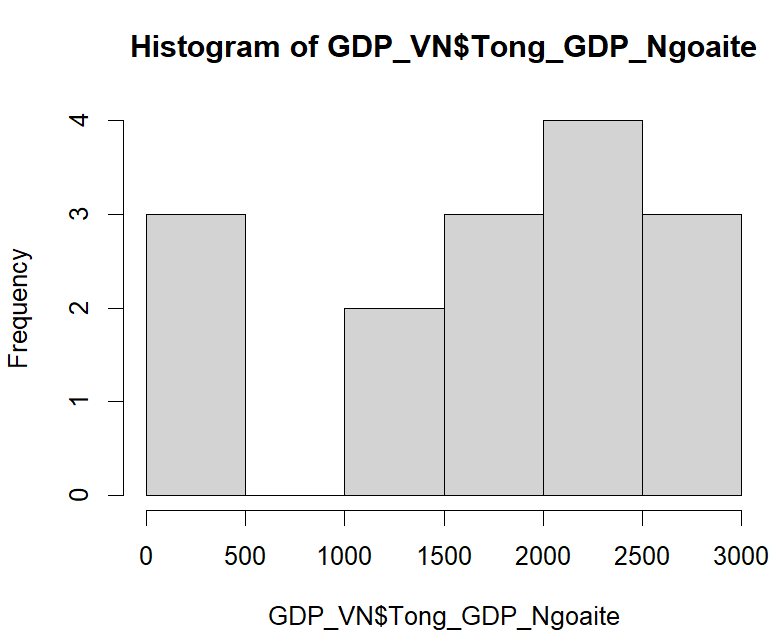
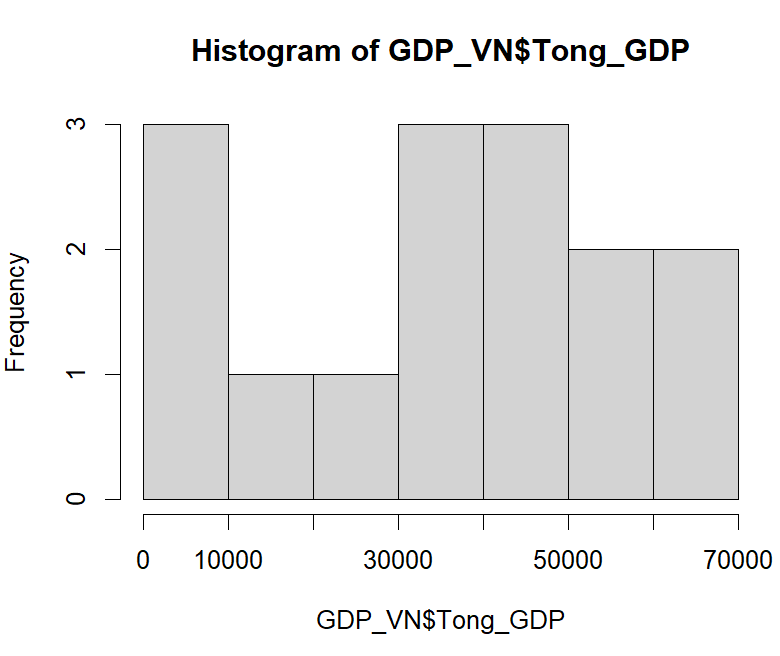
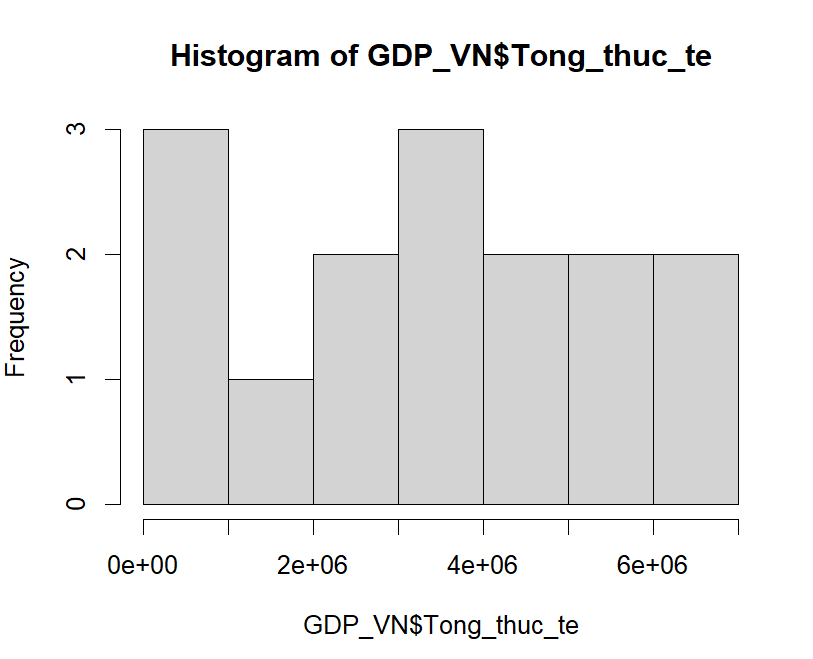
****

****

****

****

****

****