

Quy ước mã hóa lệnh trong R (R coding convention)

Trình bày: Lê Ngọc Thạch

Nội dung

1

- Phần giới thiệu file

2

- Định dạng mã lệnh

3

- Trình bày logic

Giới thiệu file mã nguồn

1

- Nội dung phần đầu của file

Cho biết file lưu mã R cho việc gì?

Người tạo là ai?

Tạo ngày nào?

Phiên bản

Lịch sử thay đổi

Ví dụ Header file

```
# -----  
# Description: Mã R này phân tích dữ liệu của bệnh nhân được  
# lưu trữ tại  
# https://github.com/phantichdulieu/phantichdulieukhoahoc/blob/master/huong-  
# dan-thuc-hanh/DataFiles/test%20excel%20file%201.xls  
# Creator: Le Ngoc Thach  
# Created: 22-Dec-2015  
# Version: 1.0  
# Record of changes  
# -----  
# -22-Dec-2015: Phiên bản đầu tiên. Tính số bệnh nhân; thống kê bệnh nhận theo  
# lần thăm clinic.  
# -----
```

Định dạng mã lệnh – tên đối tượng

➤ Đặt tên cho đối tượng

- Tên không quá dài, cũng không nên quá vắn tắt mà không đoán được ý nghĩa.
- Khi viết tắt các thuật ngữ thông dụng thì viết tắt cả là chữ thường (nếu bắt đầu tên đối tượng) hoặc **TẤT CẢ LÀ CHỮ HOA** (nếu là chữ thứ hai trở đi) trong tên đối tượng

Vd:

- `bmi = 0.26;` # Minh họa tên đối tượng viết tắt
- `bmiAVG = mean(bmi);` # Minh họa viết thường, viết HOA

Định dạng mã lệnh – khoảng trắng

➤ Dùng các khoảng trắng (space) và dấu ngoặc hợp lý để mã lệnh dễ đọc

– Dùng space trước và sau các ký hiệu:

- Gán (assign): =
- Toán tử tính toán, so sánh: + - * / > < >= <= &

Vd:

```
data$osteo[tscore1 <= -2.50] = "Yes"
```



4 chỗ này cần có space

Định dạng mã lệnh – khoảng trắng

➤ Dùng các khoảng trắng (space) và dấu ngoặc hợp lý để mã lệnh dễ đọc

– Dùng space sau dấu phẩy ngăn cách 2 thông số

Vd:

```
dat = read.csv(t, header = T)
```



Chỗ này cần có space

Định dạng mã lệnh – dòng trắng

- Giữa các đoạn mã có mục đích xử lý khác nhau nên có dòng trắng (empty line)
- Trước mỗi dòng chú thích nên có dòng trắng

Vd:

```
pbcs$AGE = age/365  
pbcs$AGE [age/365 < 40]='tre'  
pbcs$AGE [age/365 >= 40 & age/365 < 60]='trungnien'  
pbcs$AGE [age/365 >= 60]='gia'
```

```
t=file.choose()  
data = read.csv (t, header=T)
```

```
# Xem tên các cột dữ liệu  
names (data)  
attach (data)
```

Nên có dòng
trắng ở đây

Trình bày logic

- Không nên viết 1 dòng thực việc quá nhiều lệnh
 - Nên viết một dòng làm 1 việc càng đơn giản càng tốt

➤ Vd:

Thay vì viết 1 lệnh vẽ biểu đồ thanh với màu sắc và legend như sau:

```
barplot(table(obesity, gender), col = c("Red", "Blue", "Yellow"),  
        legend.text=c('Normal', 'Overweight', 'Obesity'))
```

Có thể viết thành 5 lệnh:

```
legends = c('Normal', 'Overweight', 'Obesity');  
colors = c("Red", "Blue", "Yellow");  
tabDat = table(obesity, gender);  
barplot(tabDate, col = colors, legend.text=legends)
```

Trình bày logic

- Trình bày ký hiệu \pm để diễn đạt giá trị mean (trung bình) và sd (độ lệch chuẩn) trong tài liệu
 - Không nên viết: 50 ± 20
 - Nên viết $50 (20)$
 - Có ghi chú số trong ngoặc là sd