成為初級資料分析師 I R 程式設計與資料科學應用

向量操作

郭耀仁

大綱

- 什麼是向量
- 創建
- 索引與切割
- 更新與刪除
- 向量特性

什麼是向量

截至目前為止，我們都在面對長度為 1 的向量

為什麼在 Console 印資料前面都會顯示 `[1]`？

In [ ]:
```
print("Hello world")
```

對於 R 語言來說，最小的資料單位是長度為 1 的向量，而不是純量

創建

使用 c() 函數來創建長度大於 1 的向量

In [ ]:
```
avengers <- c("The Avengers", "Avengers: Age of Ultron", "Avengers: Infinity War",
"Avengers: Endgame")
avengers_ratings <- c(8.1, 7.3, 8.5, 8.7)
avengers
avengers_ratings
```

快速創建數值向量的函數

- `seq()`
- `rep()`

In [ ]:

```
seq(from = 11, to = 21) # same as 11:21
seq(from = 11, to = 21, by = 2)
seq(from = 11, to = 21, length.out = 6)
```

In [ ]:

```
rep(5, times = 6)
rep("5566", times = 5)
rep(TRUE, times = 4)
```

索引與切割

使用 length() 函數得知向量長度

In [ ]:
```
avengers <- c("The Avengers", "Avengers: Age of Ultron", "Avengers: Infinity War",
"Avengers: Endgame")
length(avengers)
```

使用 [INDEX] 进行索引

In [ ]:
```
avengers <- c("The Avengers", "Avengers: Age of Ultron", "Avengers: Infinity War",
"Avengers: Endgame")
avengers[1]
avengers[2]
avengers[3]
avengers[length(avengers)] # in case we have a long vector
```

使用 [c(INDICES)] 進行切割

In [ ]:
```
avengers <- c("The Avengers", "Avengers: Age of Ultron", "Avengers: Infinity War",
"Avengers: Endgame")
avengers
avengers[c(1, 3, 4)]
```

隨堂練習：使用負的索引值有什麼效果?

In [ ]:
```
# Try it yourself
```

更新與刪除

修正向量資料

In [ ]:
```
avengers <- c("the avengers", "Avengers: Age of Ultron", "Avengers: Infinity War",
"Avengers: Endgame")
avengers
avengers[1] <- "The Avengers"
avengers
```

新增向量資料

In [ ]:
```
avengers <- c("The Avengers", "Avengers: Age of Ultron", "Avengers: Infinity War")
avengers
avengers <- c(avengers, "Avengers: Endgame")
avengers
```

刪除向量資料

In [ ]:
```
avengers <- c("The Avengers", "Avengers: Age of Ultron", "Avengers: Infinity War",
"Avengers: Endgame")
avengers
avengers <- avengers[-2]
avengers
```

向量特性

向量有這幾個值得注意的特性

- 元素級別（element-wise）的運算
- 同樣的 `class`
- 支援 `logical` 的資料篩選

元素級別（element-wise）的運算

In [ ]:
```
avengers_ratings <- c(8.1, 7.3, 8.5, 8.7)
avengers_ratings > 8
```

同樣的 class

In [ ]:
```
endgame <- c(8.7, "Avengers: Endgame")
class(endgame)
```

支援 logical 的資料篩選

In [ ]:
```
avengers <- c("The Avengers", "Avengers: Age of Ultron", "Avengers: Infinity War",
"Avengers: Endgame")
avengers_ratings <- c(8.1, 7.3, 8.5, 8.7)
avengers_ratings > 8
avengers[avengers_ratings > 8]
```

隨堂練習：計算三位球員的 BMI

In [1]:
```
players <- c("Jeremy Lin", "Michael Jordan", "Shaquille O'Neal")
player_heights <- c(191, 198, 216)
player_weights <- c(91, 98, 148)
# player_bmis
```

In [3]:
```
player_bmis
```

24.9444916531893 24.9974492398735 31.721536351166

隨堂練習：哪個球員的 BMI 超過 30

In [4]:
```
players <- c("Jeremy Lin", "Michael Jordan", "Shaquille O'Neal")
player_heights <- c(191, 198, 216)
player_weights <- c(91, 98, 148)
# overweight_player
```

In [6]:
```
overweight_player
```

'Shaquille O\'Neal'

隨堂練習：從 random_numbers 中找出奇數

In [ ]:
```r
set.seed(87)
random_numbers <- sample(1:1000, size = 100, replace = FALSE)
# odds
```

In [ ]:
```r
odds
```