

# Статистик програмчлалын R хэл

## Өгөгдөл оруулах ба гаргах, өгөгдлийн сантай холбох

Г.Махгал

© 2016 Г.Махгал

᠑ 2016/10/2



## Удиртгал

1. R хэлэн дэх өгөгдөлд холбогдох объектийн төрлүүдийн заримтай танилцана.
2. Өгөгдлийг файлаас оруулахад ашиглагддаг `scan()`, `read.table()` функцүүд мөн өгөгдлийг файлд хадгалахад ашиглагддаг `write.table()` функцийг товч авч үзнэ.
3. Өгөгдлийн сантай ажиллах суурь ойлголтыг MySQL өгөгдлийн санд тулгуурлан өгнө.



# Өгөгдөлд холбогдох объектийн төрлүүд

- ▶ scalar (тоо, текст, логик)
- ▶ **vector**
- ▶ **matrix**
- ▶ array
- ▶ **factor**
- ▶ list
- ▶ **data frame**
- ▶ ts



# Вектор

## Векторыг үүсгэх буюу гараас оруулах

- ▶ Векторыг гараас оруулах байдал

```
| prime.numbers = c(2, 3, 5, 7)
```

```
| club.members = c("Ankhaa", "Undraa", "Gerelt")
```

```
| logical.values = c(TRUE, FALSE)
```

- ▶ Тоон дараалал буюу тоон векторыг үүсгэх зарим арга

- ▶  $1, 2, \dots, 10$

```
| numbers = 1:10
```

⚠  $10:1 \equiv 1, 2, \dots, 10$

- ▶  $2, 4, \dots, 10$

```
| even.numbers = seq(from = 2, to = 10, by = 2)
```

- ▶  $0, 0, 0$

```
| zeros = rep(x = 0, times = 3)
```



# Вектор

## Векторыг файлаас оруулах

1.4,2.6,3.2

4.4,5.1,6.9

өгөгдлийг CSV форматаар "data\_matrix.csv" файлд хадгалсан гэе.

```
a = scan(file = "data_matrix.csv", sep = ",", dec = ".",  
        skip = 0, nlines = 2, na.strings = "NA")
```

`file` файлын нэр

`sep` баганыг зааглагч

`dec` аравтын бутархайн  
"цэг"

`skip` алгасах мөрийн тоо,  
файлын эхнээс

`nlines` унших мөрийн тоо

`na.strings` орхигдсон утгыг<sup>1</sup>  
илэрхийлэх текст

Файлыг ажлын хавтаснаас дуудна.

Хэрэгцээт зарим функцүүд: `getwd()`, `setwd("path")`

<sup>1</sup>missing value



# Матриц

Матрицийг үүсгэх буюу гараас оруулах

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$$

```
A = matrix(data = 1:6, nrow = 2, ncol = 3, byrow = FALSE)
print(A)
```

```
      [,1] [,2] [,3]
[1,]     1     3     5
[2,]     2     4     6
```



# Матриц

## Матрицийг файлаас оруулах

```
A = matrix(  
  data = scan(file = "data_matrix.csv", sep = ","),  
  ncol = 3, byrow = TRUE  
)  
print(A)
```

```
      [,1] [,2] [,3]  
[1,]  1.4  2.6  3.2  
[2,]  4.4  5.1  6.9
```

Баганын тоо тодорхойгүй үед `read.table()` ба `as.matrix()` зэрэг функцүүдийн тусламжтайгаар матрицийг импортлон авах боломж бий. Тодруулбал

```
A = as.matrix(read.table(file = "data_matrix.csv", sep = ","))
```



# Фактор ба Дата фрэйм

Фактор ба Дата фрэймийг гараас оруулах

ID	Нэр	Хүйс <sup>2</sup>	Нас
2	Эрдэнэ	эр	21
3	Халиун	эм	19
1	Од	эр	20

```
X = data.frame(  
  name = c("Erdene", "Khaliun", "Od"),  
  sex = factor(x = c(1, 2, 1), levels = c(1, 2), labels =  
    c("male", "female"), ordered = TRUE),  
  age = c(21, 19, 20),  
  row.names = c(3, 2, 1),  
  stringsAsFactors = FALSE  
)  
print(X)
```

```
   name    sex age  
3 Erdene  male  21  
2 Khaliun female 19  
1      Od   male  20
```

<sup>2</sup>чанарын өгөгдөл, фактор байдлаар оруулбал зохино





# Дата фрэйм

## Дата фрэймийг файлаас оруулах

```
"id","name","sex","age"  
3,"Erdene","male",21  
2,"Khalium","female",19  
1,"Od","male",20
```

өгөгдлийг CSV форматаар "data\_dataframe.csv" файлд хадгалсан гээ.

```
X = read.table(file = "data_dataframe.csv",  
  header = TRUE, sep = ",", dec = ".", na.strings = "NA",  
  row.names = "ID",  
  col.names = c("ID", "Name", "Sex", "Age"),  
  colClasses = c("numeric", "character", "factor",  
    "numeric")  
)
```

**header** хүснэгт толгойтой эсэх

**row.names** мөрийн нэрс эсвэл нэрсийг агуулсан баганын нэр

**col.names** баганууд буюу хувьсагчдын нэрс

**colClasses** хувьсагчдын төрөл



# Өгөгдлийг файлд хадгалах

## Өгөгдлийг CSV форматаар файлд хадгалах

```
write.table(x,  
  file = "", append = FALSE,  
  sep = " ", na = "NA", dec = ".",  
  row.names = TRUE, col.names = TRUE  
)
```

`x` экспортлох объект

`append` хуучин файлд нэмж бичих эсэх

`row.names` мөрийн нэрийг экспортлох эсэх

`col.names` баганын нэрийг экспортлох эсэх

Объектийг файлд хадгалах ба буцаан уншуулахад `save()` ба `load()` функцүүд тохиромжтой байдаг.



# Өгөгдлийн сантай ажиллах

## Бэлтгэл ажил

### Урьдчилан суулгаж бэлдсэн байх зүйлс

- ▶ MySQL Server програм
- ▶ R програмын RMySQL багц<sup>3</sup>

Нэмэлт програм суулгах шаардлага гарч болзошгүй. Тухайлбал Ubuntu Linux тархацын хувьд "libmysqlclient". Суулгах тушаал нь Ubuntu 16.04 хувилбарт

```
| sudo apt-get install libmariadb-client-lgpl-dev
```

### Өгөгдлийн сангийн системд холбогдох<sup>4</sup>

```
| library(RMySQL)  
connection = dbConnect(MySQL(),  
  user = "root", password = "", host = "localhost")
```

localhost-оор холбогдох үед алдаа гарвал 127.0.0.1 хаягийг ашиглана уу.

<sup>3</sup>суулгах тушаал нь `install.packages("RMySQL")`

<sup>4</sup>холболтыг салгах `dbDisconnect(connection)`



# Өгөгдлийн сантай ажиллах

## Өгөгдлийн сан ба хүснэгт үүсгэх

Өгөгдлийн сан үүсгэх<sup>5</sup>, нэрийг нь "r" гэе.

```
| dbSendQuery(connection, "CREATE DATABASE r")
```

Өгөгдлийн санд хандах холболтыг үүсгэх

```
| dbSendQuery(connection, "USE r")  
connection <- dbConnect(MySQL(),  
  user = "root", password = "", host = "localhost",  
  dbname = "r")
```

Хүснэгт үүсгэх

```
| dbSendQuery(connection, "  
  CREATE TABLE students (  
    id INT AUTO_INCREMENT PRIMARY KEY,  
    name VARCHAR(50), age INT UNSIGNED  
  ) ENGINE=InnoDB DEFAULT CHARSET=utf8;")
```

<sup>5</sup> устгах тушаал нь `dbSendQuery(connection, "DROP DATABASE r")`



# Өгөгдлийн сантай ажиллах

Өгөгдлийн сангийн хүснэгтэнд мэдээлэл оруулах болон хүснэгтээс мэдээлэл унших

Хүснэгтэнд мэдээлэл оруулах

```
dbSendQuery(connection, "  
  INSERT INTO students  
    (name, age)  
  VALUES  
    ('Od',20), ('Erdene',19);")
```

Хүснэгтээс мэдээлэл унших

```
student = fetch(dbSendQuery(connection, "  
  SELECT  
    id, name, age  
  FROM  
    students  
  WHERE  
    age = 20;"))
```



# Өгөгдлийн сантай ажиллах

Өгөгдлийн сангийн хүснэгтэн дэх мэдээллийг засч өөрчлөх болон устгах

Хүснэгтэн дэх мэдээллийг засч өөрчлөх

```
dbSendQuery(connection, "  
  UPDATE students  
    SET  
      age = 21  
  WHERE  
    name = 'Erdene';")
```

Хүснэгтээс мэдээллийг устгах

```
dbSendQuery(connection, "  
  DELETE FROM students  
  WHERE  
    age >= 21;")
```





© 2016 Г.Махгал

