

데이터베이스 기초

13장, 14장 과제

<13장_연습4>

- d1, d2 가 다음과 같을 때 질문에 답하시오

d1 <- 1:50

d2 <- 51:100

- d1, d2 의 값을 보이시오

d1 <- 1:50

d2 <- 51:100

d1

d2

- d1+d2, d2-d1, d1*d2, d2/d1의 결과를 각각 보이시오

d1+d2

d2-d1

d1*d2

d2/d1

- d1, d2 의 값들의 합(sum)을 각각 보이시오

sum(d1)

sum(d2)

- d1, d2 에 있는 모든 값들의 합(sum)을 보이시오

sum(c(d1,d2))

- d2 에서 가장 큰 값과 가장 작은 값을 보이시오

max(d2)

min(d2)

- d2 와 d1 의 값들의 평균값을 각각 구하고 두 평균의 차이를 보이시오

```
mean(d1)
mean(d2)
abs(mean(d2)-mean(d1))
```

- d1 의 값들을 큰 수에서 작은 수 순으로 정렬하여 보이시오
`sort(d1, decreasing = TRUE)`

- d1 과 d2 에서 큰 수 순으로 각각 10개씩을 추출하여 d3 에 저장하시오 (결과적으로 d3 에는 20개의 수가 저장)

```
d3 <- c(sort(d1, decreasing=T)[1:10], sort(d2, decreasing=T)[1:10])
```

<13장_연습5>

- 다음과 같이 벡터 v1 을 생성하시오

```
v1 <- 51:90
```

- 1) v1 에서 60 보다 작은 수 들을 보이시오

```
v1 <- 51:90
```

```
v1[v1<60]
```

- 2) v1 에서 70 보다 작은 수 들은 몇 개인지 보이시오

```
sum(v1<70)
```

- 3) v1 에서 65 보다 큰 수들의 합을 보이시오

```
sum(v1[v1>65])
```

- 4) v1 에서 60보다 크고 73 보다 작은 수들을 보이시오

```
v1[v1>60 & v1<73]
```

- 5) v1 에서 65 보다 작거나 80 보다 큰 수들을 보이시오

```
v1[v1<65 | v1>80]
```

6) v1 에서 7로 나눈 나머지가 3 인 숫자들만 보이시오

```
v1[v1%%7==3]
```

7) v1 에서 짝수들의 합계를 보이시오

```
sum(v1[v1%%2==0])
```

8) v1 에서 홀수이거나 80 보다 큰 수를 보이시오

```
sum(v1[v1%%2==1 | v1>80])
```

9) v1 에서 3과 5의 공배수를 보이시오

```
v1[v1%%3==0 & v1%%5==0]
```

<14장_연습2>

1. 다음과 같은 내용의 matrix 를 생성하시오 (이름은 score)

```
m <- c(10, 40, 60, 20)
```

```
f <- c(21, 60, 70, 30)
```

```
score <- cbind(m, f)
```

2. 컬럼의 이름을 각각 male, female 로 바꾸시오

```
colnames(score) <- c("male", "female")
```

3. 2행에 있는 모든 값을 보이시오

```
score[2, ]
```

4. female 의 모든 값을 보이시오

```
score[ , "female"]
```

5. 3행 2열의 값을 보이시오

```
score[3,2]
```

<14장_연습3>

- R 에서 제공하는 state.x77 데이터셋을 이용하여 실습하시오

1. state.x77 를 st 에 data.frame 으로 저장하시오

```
st <- data.frame(state.x77) # matrix 를 data.frame 으로 변환
```

2. st 의 내용을 보이시오

```
st
```

3. st 의 열 이름을 보이시오

```
colnames(st)
```

4. st 의 행 이름을 보이시오

```
rownames(st)
```

5. st 의 행의 개수와 열의 개수를 보이시오

```
nrow(st)
```

```
ncol(st)
```

6. st 의 요약정보를 보이시오

```
str(st)
```

7. st 의 행별 합계와 평균을 보이시오

```
rowSums(st)
```

```
rowMeans(st)
```

8. st 의 열별 합계와 평균을 보이시오

```
colSums(st)
```

```
colMeans(st)
```

9. Florida 주의 모든 정보를 보이시오

```
st["Florida", ]
```

10. 50개 주의 Income 정보만 보이시오

```
st[ , "Income"]
```

11. texas 주의 면적(Area) 을 보이시오

```
st["Texas", "Area"]
```


12. ohio 주의 인구(population) 와 수입(income)을 보이시오

```
st["Ohio", c("Population", "Income")]
```

13. 인구가 5000 이상인 주의 데이터만 보이시오

```
subset(st, Population >= 5000)
```

14. 수입이 4500 이상인 주의 인구, 수입, 면적을 보이시오

```
subset(st, Income >= 4500, select=c("Population", "Income", "Area"))
```

또는

```
subset(st[c("Population", "Income", "Area")], Income >= 4500)
```

15. 수입이 4500 이상인 주는 몇 개인지 보이시오

```
sum(st[, "Income"] >= 4500)
```

또는

```
sum(st$Income >= 4500)
```

16. 전체면적(area)이 100000 이상이고 결빙일수(frost) 가 120 이상인 주의 정보를 보이시오

```
subset(st, Area >= 100000 & Frost >= 120)
```

17. 전체면적(area)이 100000 이상이고 결빙일수(frost) 가 120 이상인 주의 정보를 보이시오

```
subset(st, Area >= 100000 & Frost >= 120)
```

18. 문맹률(illiteracy)이 2.0 이상인 주의 평균 수입은 얼마인가

```
mean(subset(st[, "Income"], "Illiteracy" >= 2.0))
```

19. 문맹률(illiteracy)이 2.0 미만인 주와 2.0 이상인 주의 평균 수입의 차이를 보이시오

```
mean(abs(subset(st[, "Income"], "Illiteracy" < 2.0) - mean(subset(st[, "Income"], "Illiteracy" >= 2.0))))
```

20. 기대수명(life.exp)이 가장 높은 주는 어디인가

```
rownames(subset(st, st$Life.Exp==max(st$Life.Exp)))
```

21. Pennsylvania 보다 수입이 높은 주들을 보이시오

```
rownames(subset(st, Income > st["Pennsylvania", "Income"]))
```

<14장_연습4>

1. R 에서 제공하는 state.x77 데이터셋에서 수입이 5000 이상인 주의 데이 터만 추출하여 rich_state.csv 에 저장하시오

```
mynew <- subset(st, Income >= 5000)
mynew
write.csv(mynew, "rich_state.csv", row.names=F, quote=F)
```

2. rich_state.csv 파일을 읽어서 ds 변수에 저장후 ds 의 내용을 보이시오

```
setwd("C:/works")
mydata <- read.csv("rich_state.csv", header = TRUE)
mydata
```