

```

> print("abc")
[1] "abc"
> plot(iris$Sepal.Length)
> sqrt(11)
[1] 3.316625
> x1 <- 12.234
> # 판별
> is.numeric(x1) # 숫자
[1] TRUE
> is.integer(x1) # 정수
[1] FALSE
> is.double(x1) # 소수점
[1] TRUE
> is.character(x1) # 문자형
[1] FALSE
> is.logical(x1) # 논리형
[1] FALSE
> is.complex(x1) # 복소수
[1] FALSE
> is.null(x1) # null
[1] FALSE
> is.na(x1) # 오류
[1] FALSE
> is.finite(x1) # 유한수
[1] TRUE

```

```

> # 바꾸기
> as.numeric(x1) # 숫자로
[1] 12.234
> as.integer(x1) # 정수로
[1] 12
> as.double(x1) # 소수형태로
[1] 12.234
> as.character(x1) # 문자로
[1] "12.234"
> as.logical(x1) # 논리형으로
[1] TRUE
> as.complex(x1) # 복소수로
[1] 12.234+0i

> abs(x1) # 절대값
[1] 12.234
> sqrt(x1) # 제곱근
[1] 3.497714
> pi #파이
[1] 3.141593
> sign(x1) # 사인
[1] 1
> round(x1) # 반올림
[1] 12
> ceiling(x1) # 무조건 올림
[1] 13
> floor(x1) # 무조건 내림
[1] 12
> exp(x1) # 지수
[1] 205664.2
> log(x1) # 자연로그
[1] 2.504219
> log10(10) # 상용로그
[1] 1
> factorial(x1) # 계승
[1] 866933546
> prod(x1) # 곱
[1] 12.234
> sin(x1) # 사인
[1] -0.3262848
> cos(x1) # 코사인
[1] 0.9452715
> tan(x1) # 탄젠트
[1] -0.3451758

```

```

> c(x1) # vector 생성
[1] 12.234
> x2 <- seq(1:5) # 일정 값만큼 증가하는 벡터생성
> x2
[1] 1 2 3 4 5
> x3 <- c(x2,x2)
> x3
[1] 1 2 3 4 5 1 2 3 4 5
> levels(x2) <- c("a","b","c","d","e") # 벡터에 라벨 생성
> x2
[1] 1 2 3 4 5
attr(,"levels")
[1] "a" "b" "c" "d" "e"
> m1 <- matrix(1:4, nrow=2, ncol=2) # 행렬생성. 행 개수 열 개수 설정
> m1
      [,1] [,2]
[1,]    1    3
[2,]    2    4
> a1 <- array(1:8, dim=c(2, 2, 2)) # 행렬생성. 행렬값 설정
> a1
, , 1

      [,1] [,2]
[1,]    1    3
[2,]    2    4

, , 2

      [,1] [,2]
[1,]    5    7
[2,]    6    8

```

```

> df1 <- data.frame(id=1:3, age=c(20, 30, 40)) # 데이터프레임생성
> df1
  id age
1  1  20
2  2  30
3  3  40
> f1 <- factor(c("A", "O", "A", "AB", "O", "B")) # 팩터생성
> f1
[1] A O A AB O B
Levels: A AB B O
> result <- list(x2,x3) # 리스트값 생성
> result
[[1]]
[1] 1 2 3 4 5
attr(,"levels")
[1] "a" "b" "c" "d" "e"

[[2]]
[1] 1 2 3 4 5 1 2 3 4 5

> length(x2) # 벡터 길이 계산
[1] 5
> dimnames(m1) <- list(rownames <- paste("row", 1:2, sep="."), colnames <- paste("col", 1:2, sep=".")) #행렬 이름 부여
> m1
      col.1 col.2
row.1     1     3
row.2     2     4
> dim(m1) # 행렬 개수 표시
[1] 2 2

> L3 <- letters[1:4] # 알파벳 소문자 표시
> data(iris) # 데이터 호출
> str(iris) # 구조확인
'data.frame':  150 obs. of  5 variables:
 $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
 $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
 $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
 $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
 $ Species      : Factor w/ 3 levels "setosa","versicolor",...: 1 1 1 1 1 1 1 1 1 1 ...
> nrow(iris) # 데이터 행 개수
[1] 150
> ncol(iris) # 데이터 열 개수
[1] 5
> head(iris) # 데이터 위에서 다섯 행 표시
  Sepal.Length Sepal.Width Petal.Length Petal.Width species
1           5.1           3.5           1.4           0.2  setosa
2           4.9           3.0           1.4           0.2  setosa
3           4.7           3.2           1.3           0.2  setosa
4           4.6           3.1           1.5           0.2  setosa
5           5.0           3.6           1.4           0.2  setosa
6           5.4           3.9           1.7           0.4  setosa
> tail(iris) # 데이터 밑에서 다섯 행 표시
   Sepal.Length Sepal.Width Petal.Length Petal.Width species
145           6.7           3.3           5.7           2.5 virginica
146           6.7           3.0           5.2           2.3 virginica
147           6.3           2.5           5.0           1.9 virginica
148           6.5           3.0           5.2           2.0 virginica
149           6.2           3.4           5.4           2.3 virginica
150           5.9           3.0           5.1           1.8 virginica

```

```

> summary(iris) # 데이터 요약
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
Min.      :4.300   Min.      :2.000   Min.      :1.000   Min.      :0.100   setosa      :50
1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300   versicolor:50
Median :5.800   Median :3.000   Median :4.350   Median :1.300   virginica  :50
Mean    :5.843   Mean    :3.057   Mean    :3.758   Mean    :1.199
3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800
Max.    :7.900   Max.    :4.400   Max.    :6.900   Max.    :2.500
> min(iris$Sepal.Length) # iris 데이터 sepal.Length 열 최소값 표시
[1] 4.3
> max(iris$Sepal.Length) # iris 데이터 sepal.Length 열 최대값 표시
[1] 7.9
> range(iris$Sepal.Length) # iris 데이터 sepal.Length 열 범위 표시
[1] 4.3 7.9
> diff(range(iris$Sepal.Length)) # iris 데이터 sepal.Length 최대 최소 차이 표시
[1] 3.6
> median(iris$Sepal.Length) # iris 데이터 sepal.Length 열 중간값 표시
[1] 5.8
> mean(iris$Sepal.Length) # iris 데이터 sepal.Length 열 평균 표시
[1] 5.843333
> var(iris$Sepal.Length) # iris 데이터 sepal.Length 열 분산 표시
[1] 0.6856935
> sum(iris$Sepal.Length) # iris 데이터 sepal.Length 열 합계 표시
[1] 876.5
> quantile(iris$Sepal.Length) # iris 데이터 sepal.Length 열 4분위수 표시
 0%  25%  50%  75% 100%
4.3  5.1  5.8  6.4  7.9
> IQR(iris$Sepal.Length) # iris 데이터 sepal.Length 열 4분위수 범위 표시
[1] 1.3
> fivenum(iris$Sepal.Length) # iris 데이터 sepal.Length 열 다섯수치요약 표시
[1] 4.3 5.1 5.8 6.4 7.9

> stem(iris$Sepal.Length) # iris 데이터 sepal.Length 열 줄기숫자로 출력

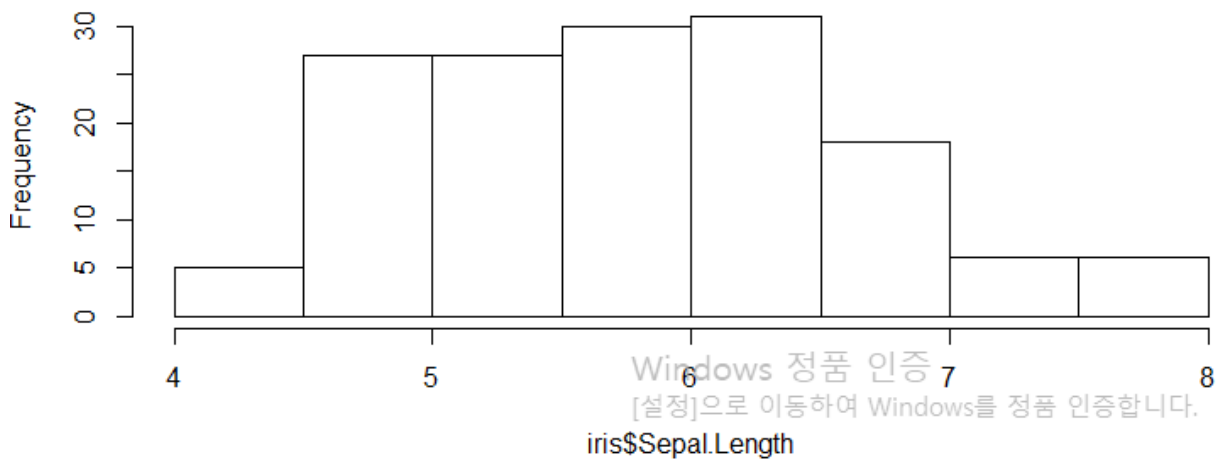
The decimal point is 1 digit(s) to the left of the |

42 | 0
44 | 0000
46 | 000000
48 | 000000000000
50 | 00000000000000000000
52 | 00000
54 | 000000000000000
56 | 0000000000000000
58 | 00000000000
60 | 00000000000000
62 | 000000000000000
64 | 00000000000000
66 | 00000000000
68 | 0000000
70 | 00
72 | 0000
74 | 0
76 | 00000
78 | 0

> hist(iris$Sepal.Length) # iris 데이터 sepal.Length 열 히스토그램 출력
> sd(iris$Sepal.Length) # iris 데이터 sepal.Length 열 표준편차 표시
[1] 0.8280661

```

### Histogram of iris\$Sepal.Length



```
> ## t.test
> t.test(Bwt ~ Sex, data=cats, var.equal = T) # p-value 검증. 단일 모집단의 모평균에 대한 신뢰구간 추정과 검정
```

Two Sample t-test

```
data: Bwt by Sex
t = -7.3307, df = 142, p-value = 1.59e-11
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.6861584 -0.3946927
sample estimates:
mean in group F mean in group M
 2.359574      2.900000
```

```
> ## 상관계수
> cor(iris[,1],iris[,2])
[1] -0.1175698
> cor.test(iris[,1],iris[,2])
```

Pearson's product-moment correlation

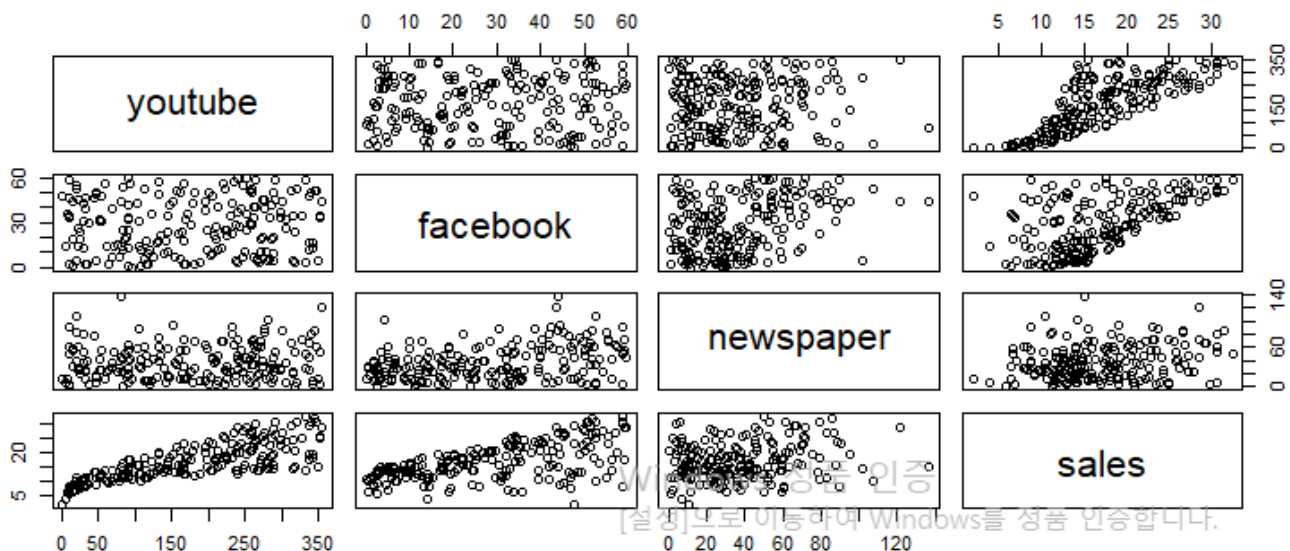
```
data: iris[, 1] and iris[, 2]
t = -1.4403, df = 148, p-value = 0.1519
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.27269325 0.04351158
sample estimates:
cor
-0.1175698
```

```

> ## 다중회귀모형
> if(!require(devtools)) install.packages("devtools")
> library(devtools)
> devtools::install_github("kassambara/datarium") # github에서 "kassambara/datarium" 설치
> install.packages("datarium")
> library(datarium)

> # 광고효과 데이터
> # 광고예산(천달러)과 판매량
> data("marketing", package = "datarium")
> head(marketing, 4)
  youtube facebook newspaper sales
1  276.12    45.36    83.04  26.52
2   53.40    47.16    54.12  12.48
3   20.64    55.08    83.16  11.16
4  181.80    49.56    70.20  22.20
> tail(marketing, 4)
  youtube facebook newspaper sales
197  113.04     5.88     9.72  11.64
198  212.40    11.16     7.68  15.36
199  340.32    50.40    79.44  30.60
200  278.52    10.32    10.44  16.08
> summary(marketing)
  youtube          facebook      newspaper      sales
Min.   : 0.84      Min.   : 0.00      Min.   : 0.36      Min.   : 1.92
1st Qu.: 89.25      1st Qu.:11.97      1st Qu.: 15.30      1st Qu.:12.45
Median :179.70      Median :27.48      Median : 30.90      Median :15.48
Mean   :176.45      Mean   :27.92      Mean   : 36.66      Mean   :16.83
3rd Qu.:262.59      3rd Qu.:43.83      3rd Qu.: 54.12      3rd Qu.:20.88
Max.   :355.68      Max.   :59.52      Max.   :136.80      Max.   :32.40
> #산점도
> plot(marketing)

```

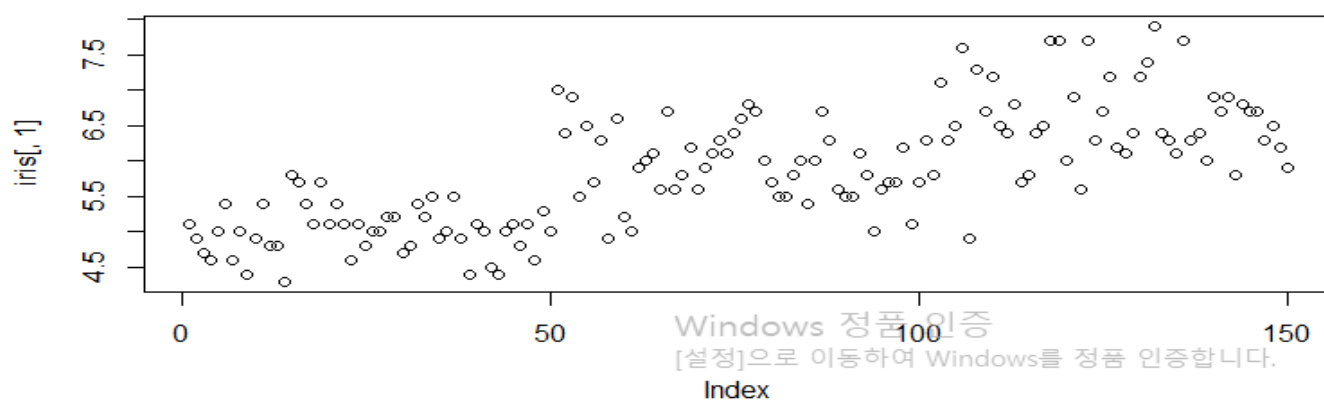


[실정] 윈도우 인증 Windows를 정품 인증합니다.

```

> dt2<-cbind(x1, x2) # cbind : 열 결합
> cov(iris[,1],iris[,2]) # 공분산 계산
[1] -0.042434
> iris_2 <- na.omit(iris[,1]) #결측치 제거
> iris_2
[1] 5.1 4.9 4.7 4.6 5.0 5.4 4.6 5.0 4.4 4.9 5.4 4.8 4.8 4.3 5.8 5.7 5.4 5.1 5.7 5.1 5.4 5.1 4.6 5.1 4.8 5.0 5.0 5.2 5.2 4.7 4.8
[32] 5.4 5.2 5.5 4.9 5.0 5.5 4.9 4.4 5.1 5.0 4.5 4.4 5.0 5.1 4.8 5.1 4.6 5.3 5.0 7.0 6.4 6.9 5.5 6.5 5.7 6.3 4.9 6.6 5.2 5.0 5.9
[63] 6.0 6.1 5.6 6.7 5.6 5.8 6.2 5.6 5.9 6.1 6.3 6.1 6.4 6.6 6.8 6.7 6.0 5.7 5.5 5.5 5.8 6.0 5.4 6.0 6.7 6.3 5.6 5.5 5.5 6.1 5.8
[94] 5.0 5.6 5.7 5.7 6.2 5.1 5.7 6.3 5.8 7.1 6.3 6.5 7.6 4.9 7.3 6.7 7.2 6.5 6.4 6.8 5.7 5.8 6.4 6.5 7.7 7.7 6.0 6.9 5.6 7.7 6.3
[125] 6.7 7.2 6.2 6.1 6.4 7.2 7.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.0 6.9 6.7 6.9 5.8 6.8 6.7 6.7 6.3 6.5 6.2 5.9
> plot(iris[,1]) #산점도

```



```
> set.seed(20150123) # 데이터 시트 생성
> id <- paste("comp", seq(1:20)) # paste : 벡터 결합
> area <- sample(c("gumi","daegu"), 20, replace = T) # sample : 무작위 생성
> df <- data.frame(id,area)
> df
```

```
      id  area
1  comp 1 daegu
2  comp 2 daegu
3  comp 3 daegu
4  comp 4 daegu
5  comp 5  gumi
6  comp 6  gumi
7  comp 7  gumi
8  comp 8  gumi
9  comp 9 daegu
10 comp 10  gumi
11 comp 11 daegu
12 comp 12 daegu
13 comp 13 daegu
14 comp 14  gumi
15 comp 15  gumi
16 comp 16  gumi
17 comp 17 daegu
18 comp 18 daegu
19 comp 19  gumi
20 comp 20  gumi
```

```
> write.csv(df, "df.csv") # write : 파일 생성
```

```
> # lm()
> x <- sample(1:10)
> y <- sample(2:11)
> df<-data.frame(x,y)
> df_lm <- lm(y ~x, data = df) # 선형 모델 생성
> df_lm
```

```
call:
lm(formula = y ~ x, data = df)
```

```
Coefficients:
```

```
(Intercept)          x
      8.7333      -0.4061
```

```
> anova(df_lm) # 분산분석 : 모형의 적합도 #####
```

```
Analysis of Variance Table
```

```
Response: y
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
x	1	13.603	13.6030	1.5795	0.2443
Residuals	8	68.897	8.6121		

```
> #신뢰구간 계산
```

```
> confint(df_lm)
```

	2.5 %	97.5 %
(Intercept)	4.110390	13.3562763
x	-1.151115	0.3389941



```

> #데이터 핸들링
> iris.dt <- iris
> #열추가
> iris.dt$test1 <- iris.dt[,2]
> head(iris.dt)
  Sepal.Length Sepal.width Petal.Length Petal.width Species test1
1           5.1          3.5         1.4         0.2  setosa  3.5
2           4.9          3.0         1.4         0.2  setosa  3.0
3           4.7          3.2         1.3         0.2  setosa  3.2
4           4.6          3.1         1.5         0.2  setosa  3.1
5           5.0          3.6         1.4         0.2  setosa  3.6
6           5.4          3.9         1.7         0.4  setosa  3.9
> iris.dt <- cbind(iris.dt,iris.dt[,4])
> head(iris.dt)
  Sepal.Length Sepal.width Petal.Length Petal.width Species test1 iris.dt[, 4]
1           5.1          3.5         1.4         0.2  setosa  3.5         0.2
2           4.9          3.0         1.4         0.2  setosa  3.0         0.2
3           4.7          3.2         1.3         0.2  setosa  3.2         0.2
4           4.6          3.1         1.5         0.2  setosa  3.1         0.2
5           5.0          3.6         1.4         0.2  setosa  3.6         0.2
6           5.4          3.9         1.7         0.4  setosa  3.9         0.4
> colnames(iris.dt)[6]<- "test.p.width"
> head(iris.dt)
  Sepal.Length Sepal.width Petal.Length Petal.width Species test.p.width iris.dt[, 4]
1           5.1          3.5         1.4         0.2  setosa          3.5         0.2
2           4.9          3.0         1.4         0.2  setosa          3.0         0.2
3           4.7          3.2         1.3         0.2  setosa          3.2         0.2
4           4.6          3.1         1.5         0.2  setosa          3.1         0.2
5           5.0          3.6         1.4         0.2  setosa          3.6         0.2
6           5.4          3.9         1.7         0.4  setosa          3.9         0.4
> #행추가
> tail(iris.dt)
  Sepal.Length Sepal.width Petal.Length Petal.width test.p.width iris.dt[, 4]
146           6.7          3.0         5.2         2.3          3.0         2.3
147           6.3          2.5         5.0         1.9          2.5         1.9
148           6.5          3.0         5.2         2.0          3.0         2.0
149           6.2          3.4         5.4         2.3          3.4         2.3
150           5.9          3.0         5.1         1.8          3.0         1.8
310           4.7          3.2         1.3         0.2          3.2         0.2
> iris.dt <- rbind(iris.dt, iris.dt[3,])
> tail(iris.dt)
  Sepal.Length Sepal.width Petal.Length Petal.width test.p.width iris.dt[, 4]
147           6.3          2.5         5.0         1.9          2.5         1.9
148           6.5          3.0         5.2         2.0          3.0         2.0
149           6.2          3.4         5.4         2.3          3.4         2.3
150           5.9          3.0         5.1         1.8          3.0         1.8
310           4.7          3.2         1.3         0.2          3.2         0.2
410           4.6          3.1         1.5         0.2          3.1         0.2
> #열제거
> iris.dt <- iris.dt[,-5]
> head(iris.dt)
  Sepal.Length Sepal.width Petal.Length Petal.width iris.dt[, 4]
1           5.1          3.5         1.4         0.2         0.2
3           4.7          3.2         1.3         0.2         0.2
4           4.6          3.1         1.5         0.2         0.2
5           5.0          3.6         1.4         0.2         0.2
6           5.4          3.9         1.7         0.4         0.4
7           4.6          3.4         1.4         0.3         0.3

```

```

> #행제거
> iris.dt <- iris.dt[-2,]
> head(iris.dt)
  Sepal.Length Sepal.width Petal.Length Petal.width iris.dt[, 4]
1           5.1           3.5           1.4           0.2           0.2
4           4.6           3.1           1.5           0.2           0.2
5           5.0           3.6           1.4           0.2           0.2
6           5.4           3.9           1.7           0.4           0.4
7           4.6           3.4           1.4           0.3           0.3
8           5.0           3.4           1.5           0.2           0.2

> #파일 저장
> #현재 디렉토리
> getwd()
[1] "C:/Users/PC/Desktop"
> #작업 디렉토리 지정
> setwd("C:\\Users\\PC\\Desktop")
> getwd()
[1] "C:/Users/PC/Desktop"
> #기본 디렉토리 지정
> setwd("C:\\Users\\PC\\Desktop")
> #파일 읽어오기
> #통계청 소음도현황_도로변지역__낮시간대
> dt <- read.csv(file="C:\\Users\\PC\\Desktop\\소음도현황_도로변지역__낮시간대.csv", head = T)
> str(dt)
'data.frame': 46 obs. of 8 variables:
 $ 시도.1. : Factor w/ 17 levels "강원","경기",...: 10 10 9 8 6 12 5 7 11 2 ...
 $ 시도.2. : Factor w/ 38 levels "강릉","고양",...: 18 18 15 15 15 15 15 15 16 ...
 $ x2017.2.2 : Factor w/ 15 levels "49","58","59",...: 15 14 12 11 10 10 6 3 4 9 ...
 $ x2017.2.2.1: Factor w/ 18 levels "-","54","56",...: 18 17 13 15 13 12 11 6 9 11 ...
 $ x2017.2.2.2: Factor w/ 16 levels "-","56","59",...: 16 15 1 11 11 14 11 10 11 10 ...
 $ x2018.1.2 : Factor w/ 15 levels "48","58","59",...: 15 14 11 10 11 10 6 4 4 8 ...
 $ x2018.1.2.1: Factor w/ 16 levels "-","54","56",...: 16 15 11 12 9 10 8 7 6 8 ...
 $ x2018.1.2.2: Factor w/ 15 levels "-","57","62",...: 15 14 1 11 9 10 8 4 9 7 ...
> head(dt)
  시도.1. 시도.2. x2017.2.2 x2017.2.2.1 x2017.2.2.2 x2018.1.2 x2018.1.2.1 x2018.1.2.2
1 시도(1) 시도(2) 용도지역 용도지역 용도지역 용도지역 용도지역 용도지역
2 시도(1) 시도(2) 가 및 나지역 다지역 라지역 가 및 나지역 다지역 라지역
3 서울 소계 68 69 - 67 69 -
4 부산 소계 67 71 68 66 70 70
5 대구 소계 66 69 68 67 67 68
6 인천 소계 66 68 71 66 68 69
> colnames(dt) <- c("area1", "area2", "area_aa", "area_bb", "area_cc", "area_a", "area_b", "area_c") # 열 이름 부여

```

> dt

	area1	area2	area_aa			area_bb			area_cc			area_a	area_b	area_c
1	시도(1)	시도(2)	가	및	응도지역	응도지역	응도지역	응도지역	가	및	응도지역	응도지역	응도지역	
2	시도(1)	시도(2)	가	및	나지역	다지역	라지역	가	및	나지역	다지역	라지역		
3	서울	소계			68	69	-				67	69	-	
4	부산	소계			67	71	68				66	70	70	
5	대구	소계			66	69	68				67	67	68	
6	인천	소계			66	68	71				66	68	69	
7	광주	소계			62	67	68				62	66	67	
8	대전	소계			59	61	67				60	65	63	
9	울산	소계			60	65	68				60	64	68	
10	경기	수원			65	67	67				64	66	66	
11	경기	성남			62	69	70				62	69	70	
12	경기	안양			61	65	67				61	66	67	
13	경기	부천			65	65	71				64	64	69	
14	경기	용인			65	68	-				64	67	-	
15	경기	안산			64	65	68				62	62	69	
16	경기	고양			66	66	-				66	67	-	
17	경기	남양주			63	61	-				63	68	-	
18	경기	의정부			66	56	-				66	56	-	
19	경기	평택			64	63	64				63	62	65	
20	경기	시흥			68	66	-				66	64	-	
21	경기	화성			63	61	-				63	60	-	
22	경기	광명			65	67	-				64	66	-	
23	경기	파주			64	69	66				64	68	66	
24	경기	군포			60	70	70				59	71	69	
25	경기	광주			58	59	60				58	60	62	
26	경기	김포			58	58	59				60	60	57	
27	경기	이천			64	68	-				65	68	-	
28	경기	구리			69	77	-				67	77	-	
29	강원	춘천			69	67	70				68	64	69	
30	강원	원주			69	68	65				69	69	66	
31	강원	강릉			66	65	67				68	65	69	
32	충북	청주			68	66	69				66	66	65	
33	충북	충주			66	66	-				65	66	-	
34	충북	제천			65	65	-				64	64	-	
35	충남	천안			61	66	62				61	66	64	
36	전북	전주			66	68	65				66	68	66	
37	전남	목포			58	64	64				58	65	68	
38	전남	여수			63	67	67				63	68	65	
39	전남	순천			49	54	56				48	54	57	
40	전남	나주			58	61	63				60	62	65	
41	전남	광양			63	65	68				62	65	69	
42	경북	포항			63	65	67				62	66	72	
43	경북	구미			63	-	-				62	-	-	
44	경남	진주			62	66	67				62	67	68	
45	경남	창원			64	66	71				65	66	71	
46	제주	소계			66	63	71				64	66	66	