R스터디#2

2023-10-02

# Chapter 4  
# Chapter 4-1) 데이터 수집하기  
ID<-c(1,2,3,4,5)  
ID

## [1] 1 2 3 4 5

SEX<-c('F','M','F','M','F')  
SEX

## [1] "F" "M" "F" "M" "F"

DATA<-data.frame(ID=ID,SEX=SEX)  
View(DATA)  
# View 함수는 반드시 첫글자 대문자로  
# 외부 데이터 가져오기  
ex\_data=read.table("C:/Users/MYCOM/data/data\_ex.txt",encoding='EUC-KR',fileEncoding = 'UTF-8')  
View(ex\_data)  
# header 옵션  
ex\_data1=read.table("C:/Users/MYCOM/data/data\_ex.txt",encoding='EUC-KR',fileEncoding = 'UTF-8',header=TRUE)  
View(ex\_data1)  
# 원시데이터 1행이 변수명인지 아닌지 확인  
# header=TRUE 옵션을 추가하면 원시데이터 1행이 변수명으로 지정됨  
# 변수명으로 사용할 행이 없을 때  
varname<-c('ID','SEX','AGE','AREA')  
ex1\_data=read.table("C:/Users/MYCOM/data/data\_ex.txt",encoding='EUC-KR',fileEncoding = 'UTF-8',col.names=varname)  
View(ex1\_data)  
# skip: 데이터 전체가 아닌 옵션에 지정한 특정 행까지 제외하고 그이후 행부터 가져옴  
ex\_data2=read.table("C:/Users/MYCOM/data/data\_ex.txt",encoding='EUC-KR',fileEncoding = 'UTF-8',header=TRUE,skip=2)  
View(ex\_data2)  
# nrows: 몇개의 행을 불러올지 지정  
ex\_data3=read.table("C:/Users/MYCOM/data/data\_ex.txt",encoding='EUC-KR',fileEncoding = 'UTF-8',header=TRUE,nrows=7)  
  
View(ex\_data3)  
# sep: 구분자 지정하는 옵션  
ex\_data5=read.table("C:/Users/MYCOM/data/data\_ex.txt",encoding='EUC-KR',fileEncoding = 'UTF-8',header=TRUE,sep=',')  
  
View(ex\_data5)  
  
## JSON 파일: 데이터 안에 다시 데이터가 정의된 중첩데이터구조  
#json\_data<-fromJSON("C:/Users/MYCOM/data/Rstudy/data\_ex.json");str(json\_data)  
# Chapter 4-1) 데이터 관측하기  
# 내장 데이터 세트 가져오기  
data('iris');iris #데이터세트 확인하기

## 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  
## 7 4.6 3.4 1.4 0.3 setosa  
## 8 5.0 3.4 1.5 0.2 setosa  
## 9 4.4 2.9 1.4 0.2 setosa  
## 10 4.9 3.1 1.5 0.1 setosa  
## 11 5.4 3.7 1.5 0.2 setosa  
## 12 4.8 3.4 1.6 0.2 setosa  
## 13 4.8 3.0 1.4 0.1 setosa  
## 14 4.3 3.0 1.1 0.1 setosa  
## 15 5.8 4.0 1.2 0.2 setosa  
## 16 5.7 4.4 1.5 0.4 setosa  
## 17 5.4 3.9 1.3 0.4 setosa  
## 18 5.1 3.5 1.4 0.3 setosa  
## 19 5.7 3.8 1.7 0.3 setosa  
## 20 5.1 3.8 1.5 0.3 setosa  
## 21 5.4 3.4 1.7 0.2 setosa  
## 22 5.1 3.7 1.5 0.4 setosa  
## 23 4.6 3.6 1.0 0.2 setosa  
## 24 5.1 3.3 1.7 0.5 setosa  
## 25 4.8 3.4 1.9 0.2 setosa  
## 26 5.0 3.0 1.6 0.2 setosa  
## 27 5.0 3.4 1.6 0.4 setosa  
## 28 5.2 3.5 1.5 0.2 setosa  
## 29 5.2 3.4 1.4 0.2 setosa  
## 30 4.7 3.2 1.6 0.2 setosa  
## 31 4.8 3.1 1.6 0.2 setosa  
## 32 5.4 3.4 1.5 0.4 setosa  
## 33 5.2 4.1 1.5 0.1 setosa  
## 34 5.5 4.2 1.4 0.2 setosa  
## 35 4.9 3.1 1.5 0.2 setosa  
## 36 5.0 3.2 1.2 0.2 setosa  
## 37 5.5 3.5 1.3 0.2 setosa  
## 38 4.9 3.6 1.4 0.1 setosa  
## 39 4.4 3.0 1.3 0.2 setosa  
## 40 5.1 3.4 1.5 0.2 setosa  
## 41 5.0 3.5 1.3 0.3 setosa  
## 42 4.5 2.3 1.3 0.3 setosa  
## 43 4.4 3.2 1.3 0.2 setosa  
## 44 5.0 3.5 1.6 0.6 setosa  
## 45 5.1 3.8 1.9 0.4 setosa  
## 46 4.8 3.0 1.4 0.3 setosa  
## 47 5.1 3.8 1.6 0.2 setosa  
## 48 4.6 3.2 1.4 0.2 setosa  
## 49 5.3 3.7 1.5 0.2 setosa  
## 50 5.0 3.3 1.4 0.2 setosa  
## 51 7.0 3.2 4.7 1.4 versicolor  
## 52 6.4 3.2 4.5 1.5 versicolor  
## 53 6.9 3.1 4.9 1.5 versicolor  
## 54 5.5 2.3 4.0 1.3 versicolor  
## 55 6.5 2.8 4.6 1.5 versicolor  
## 56 5.7 2.8 4.5 1.3 versicolor  
## 57 6.3 3.3 4.7 1.6 versicolor  
## 58 4.9 2.4 3.3 1.0 versicolor  
## 59 6.6 2.9 4.6 1.3 versicolor  
## 60 5.2 2.7 3.9 1.4 versicolor  
## 61 5.0 2.0 3.5 1.0 versicolor  
## 62 5.9 3.0 4.2 1.5 versicolor  
## 63 6.0 2.2 4.0 1.0 versicolor  
## 64 6.1 2.9 4.7 1.4 versicolor  
## 65 5.6 2.9 3.6 1.3 versicolor  
## 66 6.7 3.1 4.4 1.4 versicolor  
## 67 5.6 3.0 4.5 1.5 versicolor  
## 68 5.8 2.7 4.1 1.0 versicolor  
## 69 6.2 2.2 4.5 1.5 versicolor  
## 70 5.6 2.5 3.9 1.1 versicolor  
## 71 5.9 3.2 4.8 1.8 versicolor  
## 72 6.1 2.8 4.0 1.3 versicolor  
## 73 6.3 2.5 4.9 1.5 versicolor  
## 74 6.1 2.8 4.7 1.2 versicolor  
## 75 6.4 2.9 4.3 1.3 versicolor  
## 76 6.6 3.0 4.4 1.4 versicolor  
## 77 6.8 2.8 4.8 1.4 versicolor  
## 78 6.7 3.0 5.0 1.7 versicolor  
## 79 6.0 2.9 4.5 1.5 versicolor  
## 80 5.7 2.6 3.5 1.0 versicolor  
## 81 5.5 2.4 3.8 1.1 versicolor  
## 82 5.5 2.4 3.7 1.0 versicolor  
## 83 5.8 2.7 3.9 1.2 versicolor  
## 84 6.0 2.7 5.1 1.6 versicolor  
## 85 5.4 3.0 4.5 1.5 versicolor  
## 86 6.0 3.4 4.5 1.6 versicolor  
## 87 6.7 3.1 4.7 1.5 versicolor  
## 88 6.3 2.3 4.4 1.3 versicolor  
## 89 5.6 3.0 4.1 1.3 versicolor  
## 90 5.5 2.5 4.0 1.3 versicolor  
## 91 5.5 2.6 4.4 1.2 versicolor  
## 92 6.1 3.0 4.6 1.4 versicolor  
## 93 5.8 2.6 4.0 1.2 versicolor  
## 94 5.0 2.3 3.3 1.0 versicolor  
## 95 5.6 2.7 4.2 1.3 versicolor  
## 96 5.7 3.0 4.2 1.2 versicolor  
## 97 5.7 2.9 4.2 1.3 versicolor  
## 98 6.2 2.9 4.3 1.3 versicolor  
## 99 5.1 2.5 3.0 1.1 versicolor  
## 100 5.7 2.8 4.1 1.3 versicolor  
## 101 6.3 3.3 6.0 2.5 virginica  
## 102 5.8 2.7 5.1 1.9 virginica  
## 103 7.1 3.0 5.9 2.1 virginica  
## 104 6.3 2.9 5.6 1.8 virginica  
## 105 6.5 3.0 5.8 2.2 virginica  
## 106 7.6 3.0 6.6 2.1 virginica  
## 107 4.9 2.5 4.5 1.7 virginica  
## 108 7.3 2.9 6.3 1.8 virginica  
## 109 6.7 2.5 5.8 1.8 virginica  
## 110 7.2 3.6 6.1 2.5 virginica  
## 111 6.5 3.2 5.1 2.0 virginica  
## 112 6.4 2.7 5.3 1.9 virginica  
## 113 6.8 3.0 5.5 2.1 virginica  
## 114 5.7 2.5 5.0 2.0 virginica  
## 115 5.8 2.8 5.1 2.4 virginica  
## 116 6.4 3.2 5.3 2.3 virginica  
## 117 6.5 3.0 5.5 1.8 virginica  
## 118 7.7 3.8 6.7 2.2 virginica  
## 119 7.7 2.6 6.9 2.3 virginica  
## 120 6.0 2.2 5.0 1.5 virginica  
## 121 6.9 3.2 5.7 2.3 virginica  
## 122 5.6 2.8 4.9 2.0 virginica  
## 123 7.7 2.8 6.7 2.0 virginica  
## 124 6.3 2.7 4.9 1.8 virginica  
## 125 6.7 3.3 5.7 2.1 virginica  
## 126 7.2 3.2 6.0 1.8 virginica  
## 127 6.2 2.8 4.8 1.8 virginica  
## 128 6.1 3.0 4.9 1.8 virginica  
## 129 6.4 2.8 5.6 2.1 virginica  
## 130 7.2 3.0 5.8 1.6 virginica  
## 131 7.4 2.8 6.1 1.9 virginica  
## 132 7.9 3.8 6.4 2.0 virginica  
## 133 6.4 2.8 5.6 2.2 virginica  
## 134 6.3 2.8 5.1 1.5 virginica  
## 135 6.1 2.6 5.6 1.4 virginica  
## 136 7.7 3.0 6.1 2.3 virginica  
## 137 6.3 3.4 5.6 2.4 virginica  
## 138 6.4 3.1 5.5 1.8 virginica  
## 139 6.0 3.0 4.8 1.8 virginica  
## 140 6.9 3.1 5.4 2.1 virginica  
## 141 6.7 3.1 5.6 2.4 virginica  
## 142 6.9 3.1 5.1 2.3 virginica  
## 143 5.8 2.7 5.1 1.9 virginica  
## 144 6.8 3.2 5.9 2.3 virginica  
## 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

# 데이터 구조 확인하기  
str(iris)

## '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 ...

# 데이터 세트 컬럼 및 관측치 확인학  
ncol(iris)

## [1] 5

nrow(iris)

## [1] 150

dim(iris)

## [1] 150 5

length(iris) # iris 데이터 세트 열의 개수 확인

## [1] 5

length(iris$Species) # iris 데이터 세트 Species 열의 데이터 개수 확인

## [1] 150

# 데이터 앞부분 확인   
head(iris)

## 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

# 데이터 뒷부분 확인   
tail(iris,n=5)

## Sepal.Length Sepal.Width Petal.Length Petal.Width Species  
## 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

# 기술통계량 확인하기   
mean(iris$Sepal.Length)

## [1] 5.843333

median(iris$Sepal.Length)

## [1] 5.8

# 최솟값, 최댓값과 범위   
min(iris$Sepal.Length)

## [1] 4.3

max(iris$Sepal.Length)

## [1] 7.9

range(iris$Sepal.Length)

## [1] 4.3 7.9

# 사분위값 구하기  
quantile(iris$Sepal.Length)

## 0% 25% 50% 75% 100%   
## 4.3 5.1 5.8 6.4 7.9

quantile(iris$Sepal.Length,probs=0.25)

## 25%   
## 5.1

quantile(iris$Sepal.Length,probs=0.5)

## 50%   
## 5.8

quantile(iris$Sepal.Length,probs=0.75)

## 75%   
## 6.4

quantile(iris$Sepal.Length,probs=0.8)

## 80%   
## 6.52

# 분산과 표준편차 구하기   
var(iris$Sepal.Length)

## [1] 0.6856935

sd(iris$Sepal.Length)

## [1] 0.8280661

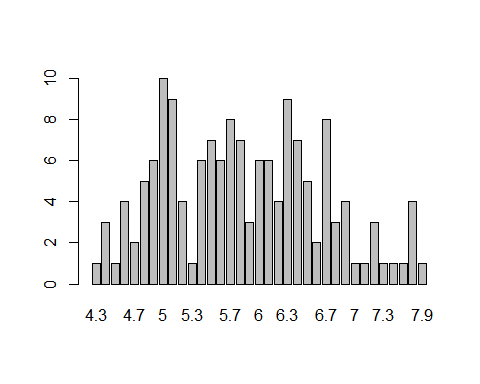
# 첨도와 왜도 구하기  
library(psych)  
  
kurtosi(iris$Sepal.Length)

## [1] -0.6058125

skew(iris$Sepal.Length)

## [1] 0.3086407

# 데이터 빈도분석 하기   
library(descr)  
freq\_test=freq(iris$Sepal.Length,plot=T)



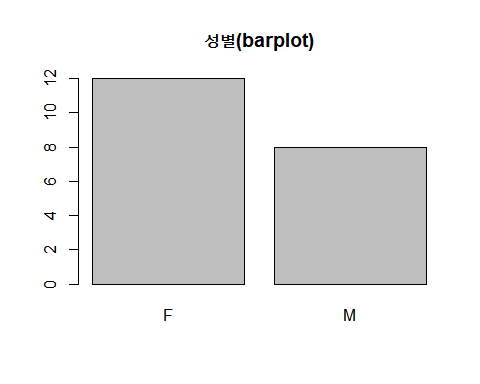
# plot=F 옵션을 통해 막대 그래프 출력을 제외할 수 있음  
freq\_test

## iris$Sepal.Length   
## Frequency Percent  
## 4.3 1 0.6667  
## 4.4 3 2.0000  
## 4.5 1 0.6667  
## 4.6 4 2.6667  
## 4.7 2 1.3333  
## 4.8 5 3.3333  
## 4.9 6 4.0000  
## 5 10 6.6667  
## 5.1 9 6.0000  
## 5.2 4 2.6667  
## 5.3 1 0.6667  
## 5.4 6 4.0000  
## 5.5 7 4.6667  
## 5.6 6 4.0000  
## 5.7 8 5.3333  
## 5.8 7 4.6667  
## 5.9 3 2.0000  
## 6 6 4.0000  
## 6.1 6 4.0000  
## 6.2 4 2.6667  
## 6.3 9 6.0000  
## 6.4 7 4.6667  
## 6.5 5 3.3333  
## 6.6 2 1.3333  
## 6.7 8 5.3333  
## 6.8 3 2.0000  
## 6.9 4 2.6667  
## 7 1 0.6667  
## 7.1 1 0.6667  
## 7.2 3 2.0000  
## 7.3 1 0.6667  
## 7.4 1 0.6667  
## 7.6 1 0.6667  
## 7.7 4 2.6667  
## 7.9 1 0.6667  
## Total 150 100.0000

# Chapter 4-3) 데이터 탐색하기  
library(descr)  
library(readxl)  
exdata1<-read\_excel("C:/Users/MYCOM/data/Rstudy/Sample1.xlsx")  
head(exdata1)

## # A tibble: 6 × 13  
## ID SEX AGE AREA CAR\_YN Y21\_AMT Y21\_CNT Y21F\_AMT Y21O\_CNT Y20\_AMT  
## <dbl> <chr> <dbl> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 1 F 50 서울 1 1300000 50 170000 25 1000000  
## 2 2 M 40 경기 1 450000 25 50000 10 700000  
## 3 3 F 28 제주 0 275000 10 7500 3 500000  
## 4 4 M 50 서울 0 2300000 8 50000 3 2500000  
## 5 5 M 27 서울 1 845000 30 130000 11 760000  
## 6 6 F 23 서울 0 42900 1 0 1 300000  
## # ℹ 3 more variables: Y20\_CNT <dbl>, Y20F\_AMT <dbl>, Y20O\_CNT <dbl>

freq(exdata1$SEX,plot=T,main='성별(barplot)')

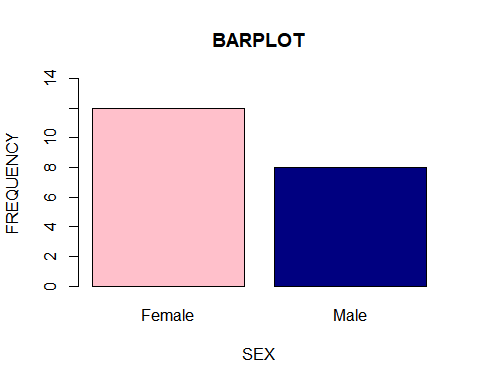


## exdata1$SEX   
## Frequency Percent  
## F 12 60  
## M 8 40  
## Total 20 100

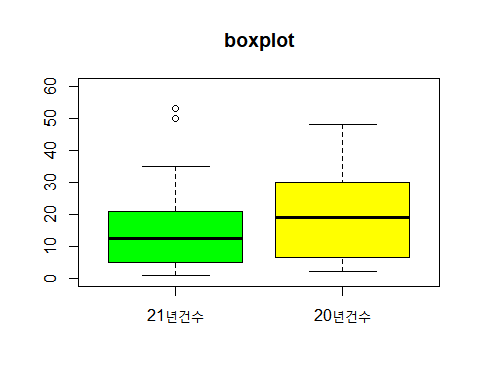
# 빈도분포를 구하고 막대 그래프 그리기   
dist\_sex<-table(exdata1$SEX)  
dist\_sex

##   
## F M   
## 12 8

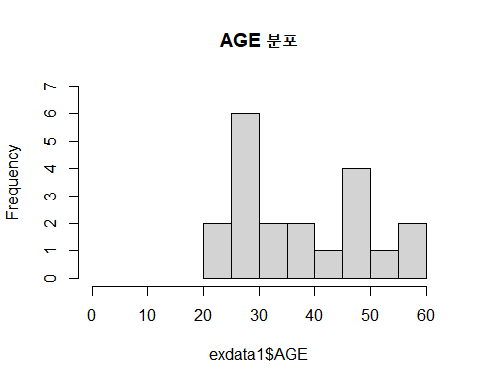
barplot(dist\_sex,ylim=c(0,14),main='BARPLOT',xlab='SEX',  
 ylab='FREQUENCY',names=c('Female','Male'),  
 col=c('pink','navy'))



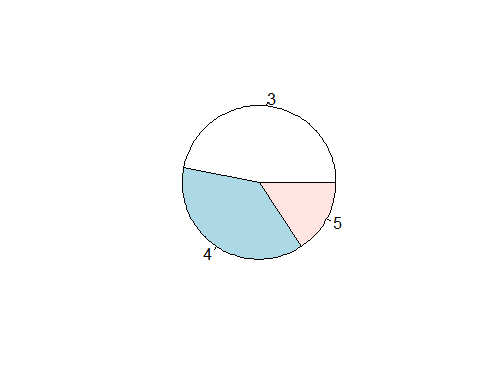
# 상자 그림 그리기  
boxplot(exdata1$Y21\_CNT,exdata1$Y20\_CNT,ylim=c(0,60),  
 main='boxplot',names=c('21년건수','20년건수'),  
 col=c('green','yellow'))



# 히스토그램 그리기   
hist(exdata1$AGE,xlim=c(0,60),ylim=c(0,7),main='AGE 분포')



# 파이차트 그리기  
data(mtcars)  
x<-table(mtcars$gear)  
pie(x)



# 줄기잎그림 그리기  
x<-c(1,2,3,4,7,8,8,5,9,6,9)  
stem(x)

##   
## The decimal point is at the |  
##   
## 0 | 0  
## 2 | 00  
## 4 | 00  
## 6 | 00  
## 8 | 0000

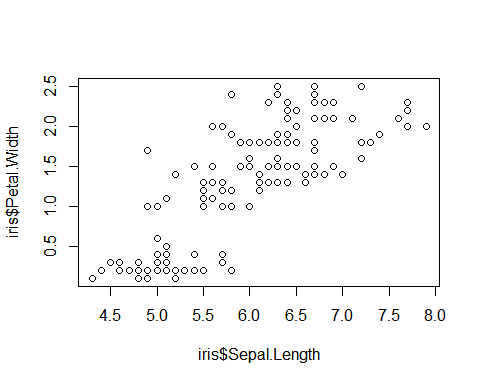
stem(x,scale=2)

##   
## The decimal point is at the |  
##   
## 1 | 0  
## 2 | 0  
## 3 | 0  
## 4 | 0  
## 5 | 0  
## 6 | 0  
## 7 | 0  
## 8 | 00  
## 9 | 00

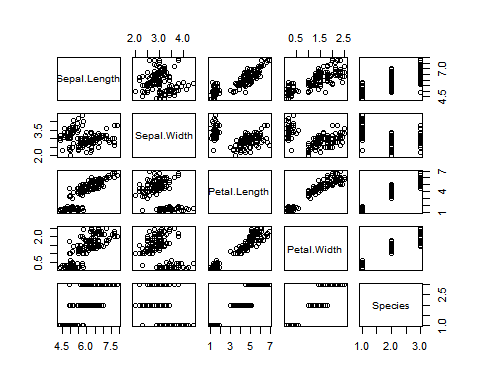
stem(x,scale=0.5)

##   
## The decimal point is 1 digit(s) to the right of the |  
##   
## 0 | 1234  
## 0 | 5678899

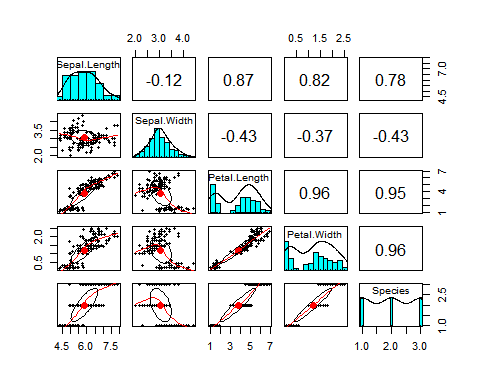
# 산점도 그리기  
data(iris)  
plot(x=iris$Sepal.Length,y=iris$Petal.Width)



# 산점도 행렬  
data(iris)  
pairs(iris)



# psych 패키지로 산점도 행렬 그리기   
library(psych)  
data(iris)  
pairs.panels(iris)



# Chapter 6  
# Chapter 6-1) 그래프 그리기  
library(ggplot2)

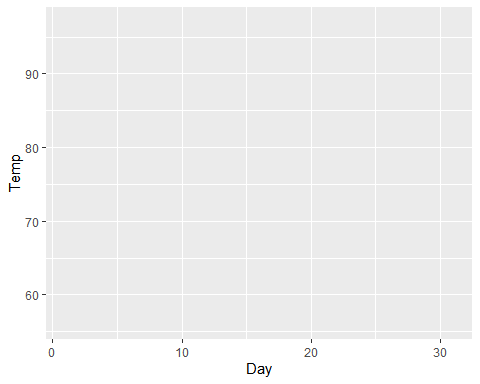
##   
## 다음의 패키지를 부착합니다: 'ggplot2'

## The following objects are masked from 'package:psych':  
##   
## %+%, alpha

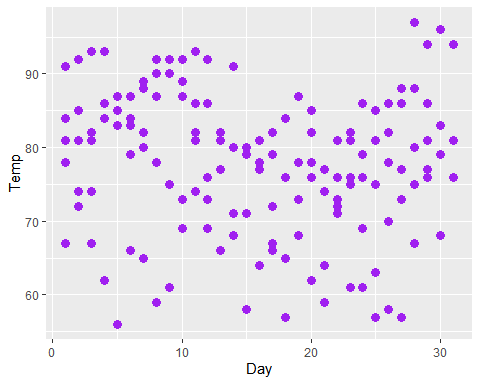
str(airquality)

## 'data.frame': 153 obs. of 6 variables:  
## $ Ozone : int 41 36 12 18 NA 28 23 19 8 NA ...  
## $ Solar.R: int 190 118 149 313 NA NA 299 99 19 194 ...  
## $ Wind : num 7.4 8 12.6 11.5 14.3 14.9 8.6 13.8 20.1 8.6 ...  
## $ Temp : int 67 72 74 62 56 66 65 59 61 69 ...  
## $ Month : int 5 5 5 5 5 5 5 5 5 5 ...  
## $ Day : int 1 2 3 4 5 6 7 8 9 10 ...

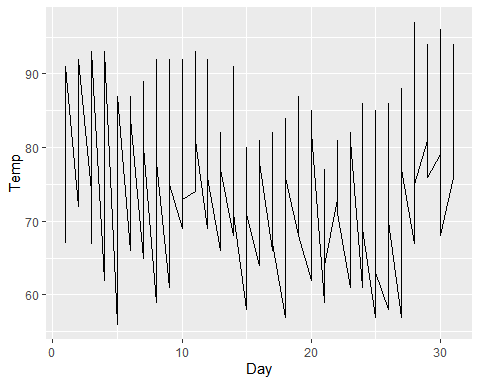
# 그래프 기본 틀 생성하기  
ggplot(airquality,aes(x=Day,y=Temp))



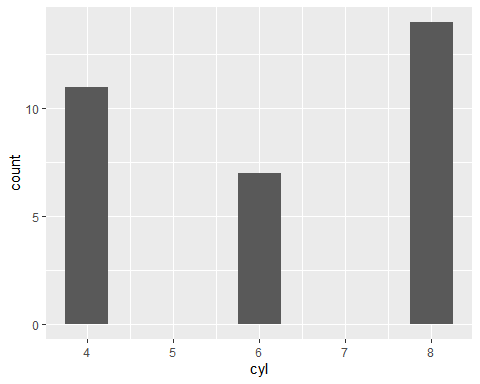
# 산점도 그리기   
ggplot(airquality,aes(x=Day,y=Temp))+geom\_point(size=3,color='purple')



# 선그래프 그리기  
ggplot(airquality,aes(x=Day,y=Temp))+geom\_line()



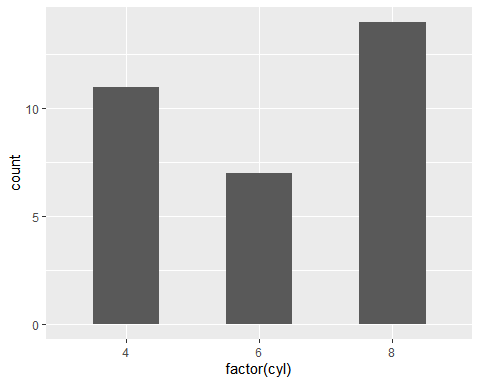
# 막대 그래프 그리기  
ggplot(mtcars,aes(x=cyl))+geom\_bar(width=0.5)



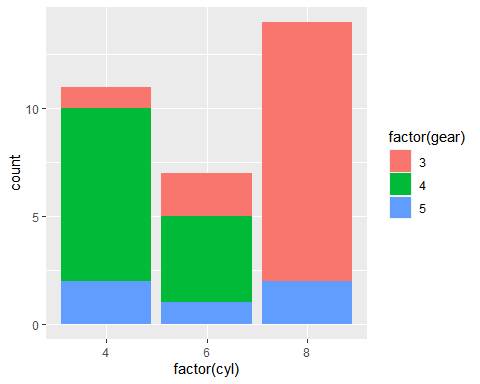
# cyl 변수의 속성 확인해보기  
str(mtcars$cyl)

## num [1:32] 6 6 4 6 8 6 8 4 4 6 ...

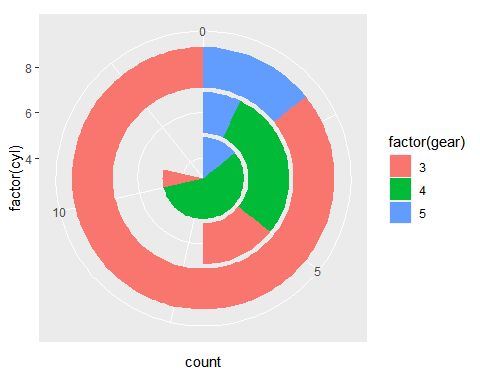
# 빈범주를 제외하고 싶다면 수치형인 데이터를 범주화할 수 있는 factor 함수 사용하기  
ggplot(mtcars,aes(x=factor(cyl)))+geom\_bar(width=0.5)



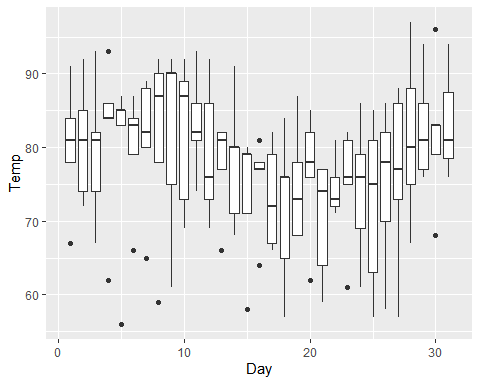
# 누적 막대 그래프 그리기  
ggplot(mtcars,aes(x=factor(cyl)))+geom\_bar(aes(fill=factor(gear)))



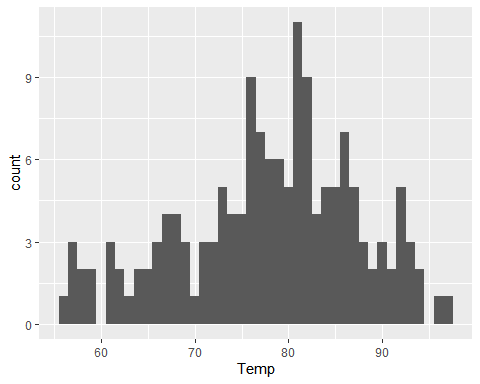
# 누적 막대 그래프로 그리는 선버스트 차트  
# theta='y' 옵션을 통해 그래프 가운데가 도넛처럼 비어있는 선버스트 차트 그리기  
ggplot(mtcars,aes(x=factor(cyl)))+geom\_bar(aes(fill=factor(gear)))+coord\_polar(theta='y')



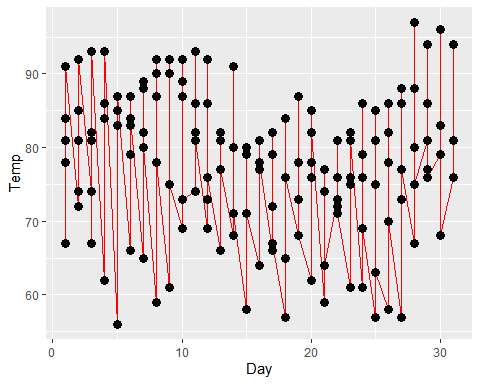
# 상자그림 그리기  
ggplot(airquality,aes(x=Day,y=Temp,group=Day))+geom\_boxplot()



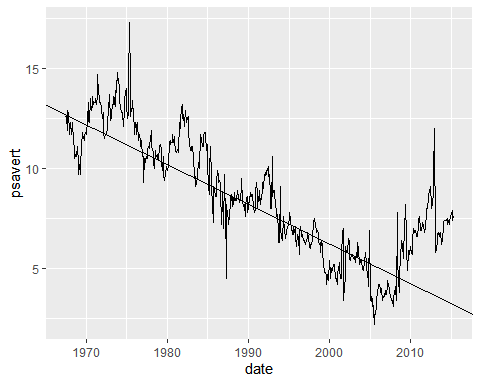
# 히스토그램 그리기  
# binwidth를 통해 막대 그래프 구간 직접 조종  
ggplot(airquality,aes(Temp))+geom\_histogram(binwidth = 1)



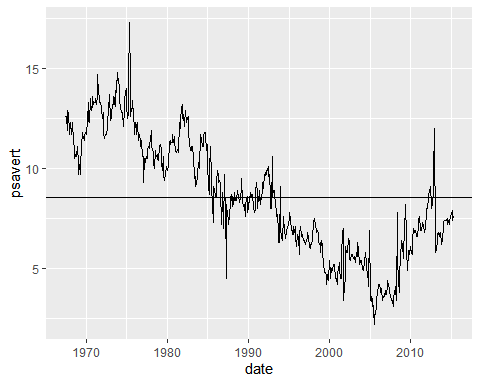
# 선그래프와 산점도 함께 그리기  
ggplot(airquality,aes(x=Day,y=Temp))+geom\_line(color='red')+geom\_point(size=3)



# Chapter 6-2) 그래프에 객체 추가하기  
# 그래프에 사선 그리기  
ggplot(economics,aes(x=date,y=psavert))+  
 geom\_line()+geom\_abline(intercept=12.18671,slope=-0.00054444)



# 평행선 그리기  
ggplot(economics,aes(x=date,y=psavert))+geom\_line()+geom\_hline(yintercept=mean(economics$psavert))



# 그래프에 수직선 그리기  
library(dplyr)

##   
## 다음의 패키지를 부착합니다: 'dplyr'

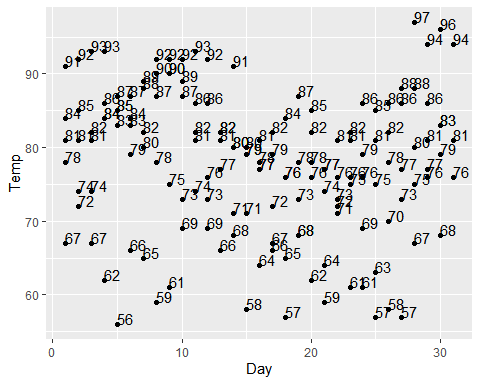
## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

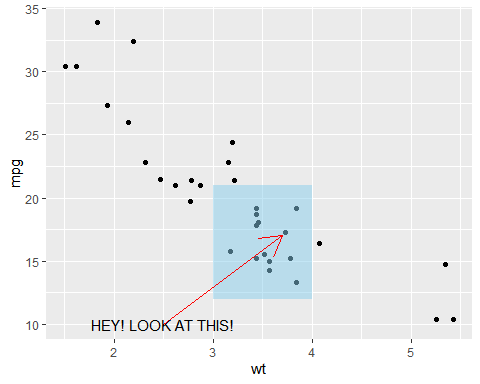
x\_inter<-filter(economics,psavert==min(economics$psavert))$date  
# 수직선의 위치를 파악하지 않고 직접 원하는 날짜에 수직선을 그릴 수 있음  
#ggplot(economics,aes(x=date,y=psavert))+geom\_vline(xintercept=x\_inter)  
# 날짜를 그대로 입력하면 글자로 인식되어 원하는 위치에 수직선이 그려지지 않으므로 as.Date 함수를 이용해 날짜 형식으로 변환해야함  
geom\_vline(xintercept=as.Date('2005-07-01'))

## mapping: xintercept = ~xintercept   
## geom\_vline: na.rm = FALSE  
## stat\_identity: na.rm = FALSE  
## position\_identity

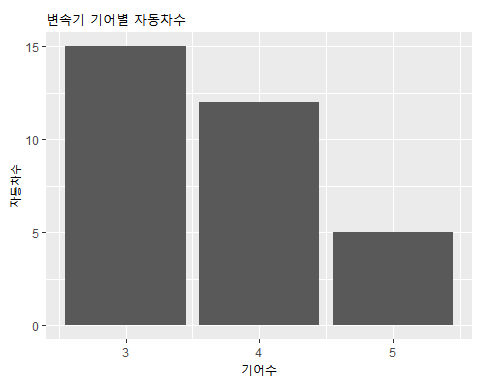
# 레이블 입력하기  
ggplot(airquality,aes(x=Day,y=Temp))+geom\_point()+  
 geom\_text(aes(label=Temp,vjust=0,hjust=0))



# 도형 및 화살표 넣기   
# 사각형+화살표+ 텍스트 넣기  
ggplot(mtcars,aes(x=wt,y=mpg))+geom\_point()+  
 annotate('rect',xmin=3,xmax=4,ymin=12,ymax=21,alpha=0.5,fill='skyblue')+  
 annotate('segment',x=2.5,xend=3.7,y=10,yend=17,  
 color='red',arrow=arrow())+  
 annotate('text',x=2.5,y=10,label='HEY! LOOK AT THIS!')



# 그래프와 축에 제목 추가하고 디자인 테마 적용하기   
ggplot(mtcars,aes(x=gear))+geom\_bar()+labs(x='기어수',y='자동차수',title='변속기 기어별 자동차수')



# 상관분석하여 두 변수간의 상관관계 확인하기   
cor.test(exdata1$Y20\_CNT,exdata1$Y21\_CNT)

##   
## Pearson's product-moment correlation  
##   
## data: exdata1$Y20\_CNT and exdata1$Y21\_CNT  
## t = 4.9343, df = 18, p-value = 0.000107  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.4751688 0.8990895  
## sample estimates:  
## cor   
## 0.7582507

# 절편과 기울기 구하기: 회귀분석  
reg\_result<-lm(Y21\_CNT~Y20\_CNT,data=exdata1)  
reg\_result

##   
## Call:  
## lm(formula = Y21\_CNT ~ Y20\_CNT, data = exdata1)  
##   
## Coefficients:  
## (Intercept) Y20\_CNT   
## 0.7104 0.7864

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

summary(cars)

## speed dist   
## Min. : 4.0 Min. : 2.00   
## 1st Qu.:12.0 1st Qu.: 26.00   
## Median :15.0 Median : 36.00   
## Mean :15.4 Mean : 42.98   
## 3rd Qu.:19.0 3rd Qu.: 56.00   
## Max. :25.0 Max. :120.00

## Including Plots

You can also embed plots, for example:



Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.