

# 5주차 복습 문제

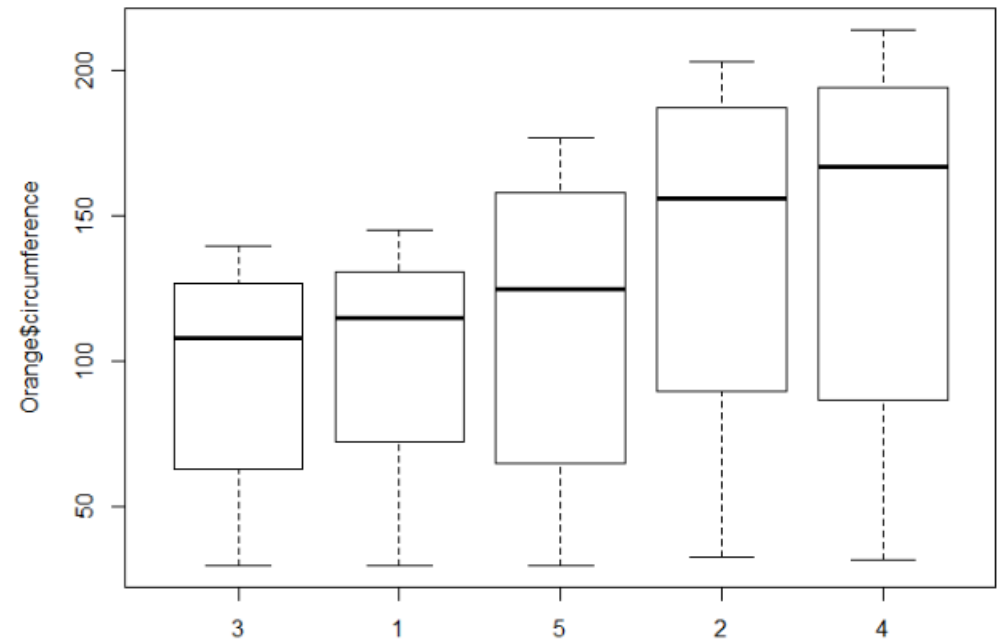
- 기초 통계량과 시각화

Made by 2조

# 1번.

datasets 패키지를 설치한 다음 Orange 라는 데이터를 상위 10개 까지 출력하고 변수 Tree별 circumference의 상자 그림을 그리시오. 최솟값과 사분위수도 함께 구하시오.

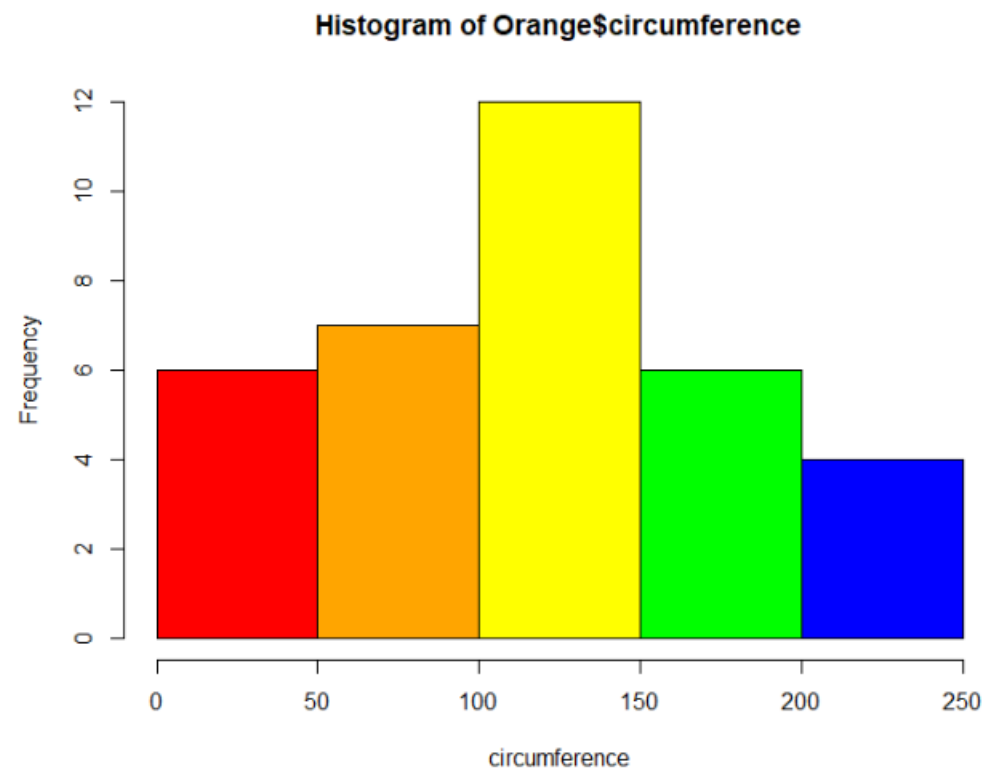
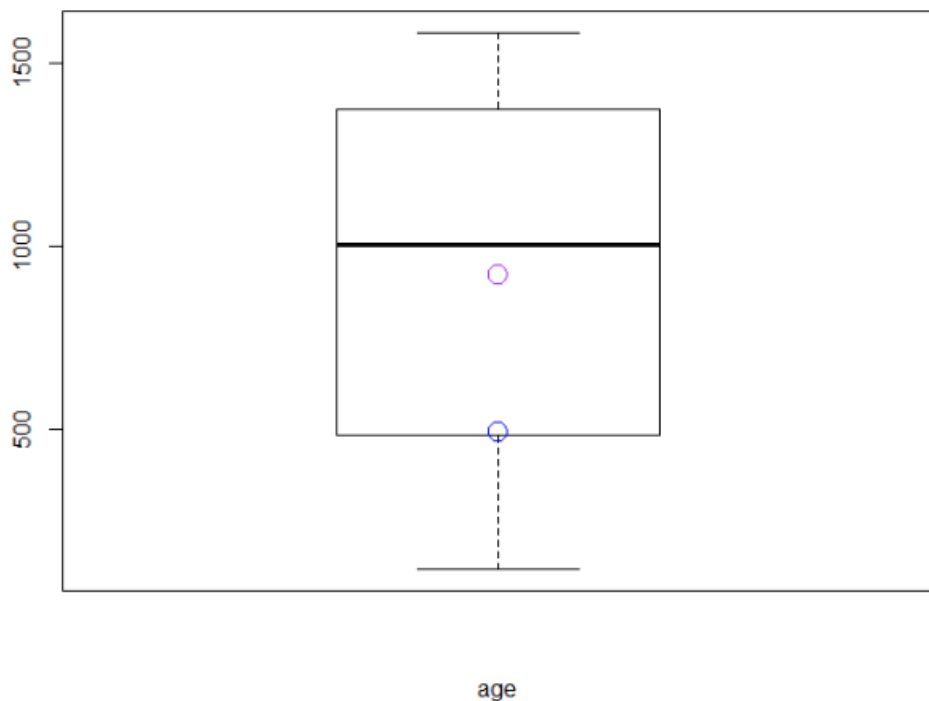
	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	30	30.0	30	33.0	32
[2,]	63	72.5	65	90.0	87
[3,]	108	115.0	125	156.0	167
[4,]	127	131.0	158	187.5	194
[5,]	140	145.0	177	203.0	214



## 2번.

IQR 대비 배수를 1.2로 하여 변수 age에 대한 상자그림을 그리고 age의 표준편차와 평균을 각각 다른 색으로 한 번에 나타내시오. (x축 이름을 age로 지정하시오.)

또한, circumference에 대한 5칸(막대)짜리 히스토그램을 만들고 x축 이름을 'circumference'로 지정한 후 각 막대를 빨주노초파 색으로 채우시오.



3번.

주어진 데이터를 대문자로 변환한 후 N(콧물)을 C(감기)로 간주하고 합계가 포함된 표를 만들어라  
J=joint swelling(관절의 부어오름), F=fatigue(피로), B=back pain(등 통증), M=muscle weakness(근육의 허약함), C=cold(감기),  
N=nose running(콧물), O=other(기타))

##( "o", "o", "n", "j", "c", "f", "b", "b", "f", "o", "j", "o", "o", "m", "o", "f", "f", "o", "o", "n", "o", "n", "j", "f", "j", "b", "o", "c", "j", "o", "j", "j",  
"f", "n", "o", "b", "m", "o", "j", "m", "o", "b", "o", "f", "j", "o", "o", "b", "n", "c", "o", "o", "o", "m", "b", "f", "j", "o", "f", "n").

```
complaint <-  
  c( "o", "o", "n", "j", "c", "f", "b", "b", "f", "o", "j", "o", "o",  
    "m", "o", "f", "f",  
      "o", "o", "n", "o", "n", "j", "f", "j", "b", "o", "c", "j", "o",  
    "j", "j", "f", "n",  
      "o", "b", "m", "o", "j", "m", "o", "b", "o", "f", "j", "o", "o",  
    "b", "n", "c", "o",  
      "o", "o", "m", "b", "f", "j", "o", "f", "n")  
complaint<-toupper(complaint)  
complaint[complaint=="N"]<-"C"  
complaint_table<-addmargins(table(complaint))  
complaint_table
```

complaint

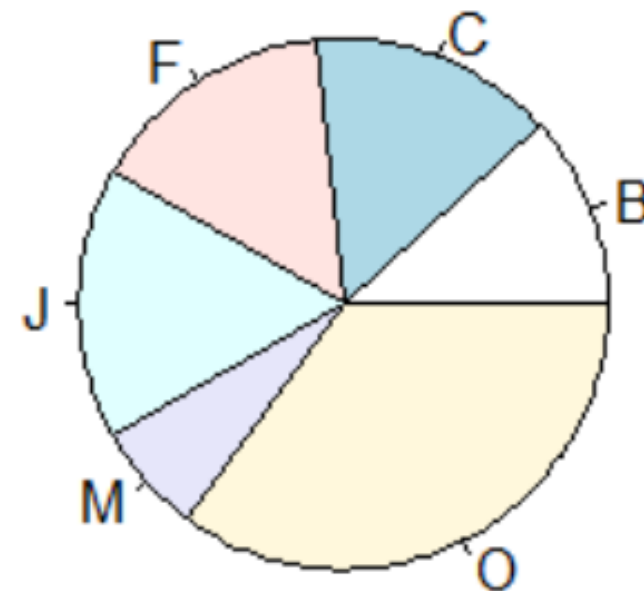
B	C	F	J	M	O	Sum
7	9	9	10	4	21	60

4번.

위에서 만든 complaint를 활용한 표를 가지고 파이차트를 만들어 시각화해보자.(색상 상관 없음. 범주명과 제목 포함).

```
pie(x=table(complaint),labels =  
c("B","C","F","J","M","O"),lty=1,main="건강에 대한 불편함의 형태")
```

건강에 대한 불편함의 형태



# 5번

원소의 개수가 10개이고, N번째 원소의 값이  $2^N$ 인 배열을 만들고, 이 배열 원소들의 평균과 표준편차를 출력하라

(단, 배열을 초기화하는데 while문을 이용하라)

평균 : 204.6

표준편차 : 329.897 이 출력되어야함

# 6번

5번의 응용문제

5번의 배열에서 원소의 값이 500미만인 것들의 평균과 표준편차를 출력하시오

평균 : 63.75

표준편차 : 88.53369 가 출력되어야함

# 7번

(a) R내장 데이터인 `airquality`에 대하여 Month별 Temp를 아래와 같은 table로 만들어라.

```
> addmargins(table(airquality[,c(5,4)]))
```

	Temp	56	57	58	59	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82
Month	5	1	3	2	2	3	2	0	1	1	3	2	2	2	0	0	1	1	2	0	1	0	0	1	0	1	0
	6	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	1	2	1	1	4	3	2	2	2	0	2
	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	5	4
	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	2	2	3	1	3	2
	9	0	0	0	0	0	0	1	1	0	0	1	2	1	1	3	0	1	0	2	3	2	2	0	1	2	1
Sum		1	3	2	2	3	2	1	2	2	3	4	4	3	1	3	3	5	4	4	9	7	6	6	5	11	9

	Temp	83	84	85	86	87	88	89	90	91	92	93	94	96	97	Sum
Month	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	31
	6	1	1	1	0	2	0	0	1	0	1	1	0	0	0	30
	7	3	3	3	3	1	2	1	0	1	2	0	0	0	0	31
	8	0	0	1	4	1	1	1	2	0	1	0	2	1	1	31
	9	0	1	0	0	1	0	0	0	1	1	2	0	0	0	30
Sum		4	5	5	7	5	3	2	3	2	5	3	2	1	1	153

(b) R내장 데이터인 `airquality`를 다시 한 번 더 활용하자.

6월부터 8월동안 기온이 90 이상인 횟수를 나타내는 막대차트를 만들어라

이때, 아래의 형태와 동일하게 만들어라 (참고: 막대 색깔은 `lightpink`)

(단, 막대차트를 만듦에 있어서 반드시 각 월별 기온이 세세하게 구분되지 않도록 만들어라)

(Hint: 'dplyr' 패키지를 설치하고 하기)

