



비즈니스프로그래밍

- Python과 R 기초 -



행이름 다루기



▶ 행이름이 있는 데이터 불러오기

```
데이터프레임이름 <- read.csv("저장된경로/저장된파일이름",  
                             row.names = 1)
```

▶ 행이름 관련 함수들

tibble 함수	형태	의미
has_rownames()	has_rownames(데이터프레임명)	행이름(rowname)을 가지는지 여부 - TRUE: 행이름 있음 - FALSE: 행이름 없음
remove_rownames()	remove_rownames(데이터프레임명)	데이터프레임에서 행이름 제거
rownames_to_column()	rownames_to_column(데이터프레임명, var = '칼럼이름')	행이름을 새로운 변수(칼럼)으로 변환
columns_to_rownames()	columns_to_rownames(데이터프레임명, var = '칼럼이름')	지정한 변수(칼럼)을 행이름으로 변환
rowid_to_column()	rowid_to_column(데이터프레임명)	행번호를 새로운 변수(칼럼)으로 생성



7. 그래프



▶ 7.1 파이차트

- ▶ 하나의 원 안에 데이터의 각 값이 차지하는 비율을 넓이로 나타내는 그래프

pie(데이터이름)

파라미터	설명
labels	데이터의 각 값에 대응하는 이름을 표시한다.
main	파이차트의 전체 이름을 표시한다.
col	각 데이터의 영역을 지정한 색으로 구분한다.
clockwise	각 데이터의 영역을 어느 방향으로 표시할지 결정한다. 값이 TRUE이면 시계방향, FALSE이면 반시계방향으로 표시한다. 기본값은 FALSE이다.
init.angle	데이터의 표시를 시작하는 각도를 나타낸다. 기본값은 0으로 3시 방향을 시작점으로 하여 데이터를 표시한다.



그래프 저장



- ▶ Plots 창에 그려진 그래프를 파일로 저장하여 사용할 수 있다.
- ▶ Plots 창의 [Export]를 클릭
 - ▶ Save as Image: 그래프를 PNG, JPEG, EPS 등과 같은 이미지 파일로 저장
 - ▶ Save as PDF: 그래프를 PDF파일로 저장
 - ▶ Copy to Clipboard: 그래프를 클립보드에 일시 저장해서 다른 프로그램으로 붙여넣기



7. 그래프



▶ ggplot2 패키지

ggplot2 명령어

ggplot(데이터, 좌표축설정) + 그래프함수 + 추가설정

그래프 함수	그래프 형태
geom_col()	막대그래프
geom_bar()	빈도 막대그래프
geom_histogram()	히스토그램
geom_line()	선그래프
geom_point()	산점도
geom_boxplot()	상자그림



7. 그래프



- ▶ 7.2 막대그래프 → `geom_col()`
 - ▶ 데이터의 그룹별 값을 막대의 길이로 나타낸 그래프
 - ▶ 크기순 정렬 → `reorder()`
- ▶ 7.3 빈도막대그래프 → `geom_bar()`
 - ▶ 데이터에 포함되어 있는 각 값들의 개수를 나타내는 그래프



7. 그래프



- ▶ 7.3 히스토그램 ➔ `geom_histogram()`
 - ▶ 데이터가 그룹으로 지정되어 있지 않은 경우에 데이터를 일정 범주로 나누어 그 범주에 속한 데이터의 개수를 막대그래프로 나타낸다.
 - ▶ 축 범위 설정 ➔ `xlim()`, `ylim()`



7. 그래프



- ▶ 7.4 선그래프 → `geom_line()`

- ▶ 데이터들을 직선으로 연결하여 전체 데이터를 선으로 나타낸 그래프



7. 그래프



▶ 7.5 산점도 → `geom_point()`

- ▶ 두 개의 데이터를 각각 x축과 y축의 값으로 하는 점들을 좌표 평면에 나타내어 만드는 그래프



7. 그래프



▶ 7.6 상자그림 → `geom_boxplot()`

- ▶ 집단 상자그림: 데이터를 어떤 특성에 따라 집단으로 분류하고 각 집단의 데이터를 상자그림으로 그려 비교하면 데이터의 특징을 파악하는데 효과적일 수 있다.