#4/28(목)\_HW1(for dplyr[filter/select/arrange함수])

#mpg 데이터를 이용해 분석 문제를 해결해 보세요.

#• Q1. 자동차 배기량에 따라 고속도로 연비가 다른지 알아보려고 합니다.

#자동차 중 drv 타입(f/r/4)중 어떤 자동차의 hwy(고속도로 연비)가

#평균적으로 더 높은지 알아보세요.

library(ggplot2)

mpg=data.frame(ggplot2::mpg)

HWT1=mpg %>% select(drv,hwy) %>% arrange(drv)

drv4=HWT1 %>% filter(drv==4)

avg4=mean(drv4$hwy)

avg4

> avg4

[1] 19.17476

drvr=HWT1 %>% filter(drv=="r")

avgr=mean(drvr$hwy)

avgr

> avgr

[1] 21

drvf=HWT1 %>% filter(drv=="f")

avgf=mean(drvf$hwy)

avgf

> avgf

[1] 28.16038

#• Q2. 자동차 제조 회사에 연도(year)에 따라 도시 연비가 다른지 알아보려고 합니다.

#"audi"와 "toyota" 중 어느 manufacturer(자동차 제조 회사)의 cty(도시 연비)가

#평균적으로 더 높은지 알아보세요.

#• Q3. "chevrolet", "ford", "honda" 자동차의 고속도로 연비 평균을 알아보려고 합니다.

#이 회사들의 자동차를 추출한 뒤 도시연비 전체 평균을 구해보세요

HWT3=mpg %>% filter(manufacturer%in% c("chevrolet","ford","honda")) %>% select(cty,hwy)

HWT3

mean(HWT3$cty)

mean(HWT3$hwy)

> mean(HWT3$cty)

[1] 16.13208

> mean(HWT3$hwy)

[1] 22.50943

#• Q4. mpg 데이터는 11 개 변수로 구성되어 있습니다.

#이 중 일부만 추출해서 분석에 활용하려고 합니다. mpg

#데이터에서 class(자동차 종류), cty(도시 연비) drv(구동타입) 변수를 추출해 새로운 데이터를 만드세요.

#새로 만든 데이터의 일부를 출력해서 세 변수로만 구성되어 있는지 확인하세요.

#• Q5. 자동차 종류에 따라 도시 연비가 다른지 알아보려고 합니다. 앞에서 추출한

#데이터를 이용해서 class(자동차 종류)가 "suvcompact"인 자동차와 "minivan"인 자동차 중

#어떤 자동차의 hwy(고속도로 연비)가 더 높은지 알아보세요.