Doit R 4,5,pdf Assignment

LEE SANGIN

2017년 9월 24일

#### 04-2 혼자서해보기

fruits <- data.frame("제품" = c("사과","딸기","수박"),  
 "가격" = c(1800,1500,3000),  
 "판매량" = c(24,38,13))  
fruits

## 제품 가격 판매량  
## 1 사과 1800 24  
## 2 딸기 1500 38  
## 3 수박 3000 13

#### 05-2 혼자서해보기

Q1. ggplot2 패키지의 mpg 데이터를 사용할 수 있도록 불러온 후 복사본 만들기

library(ggplot2)  
data(mpg)  
  
mpg\_copy <- mpg  
head(mpg\_copy)

## # A tibble: 6 x 11  
## manufacturer model displ year cyl trans drv cty hwy fl  
## <chr> <chr> <dbl> <int> <int> <chr> <chr> <int> <int> <chr>  
## 1 audi a4 1.8 1999 4 auto(l5) f 18 29 p  
## 2 audi a4 1.8 1999 4 manual(m5) f 21 29 p  
## 3 audi a4 2.0 2008 4 manual(m6) f 20 31 p  
## 4 audi a4 2.0 2008 4 auto(av) f 21 30 p  
## 5 audi a4 2.8 1999 6 auto(l5) f 16 26 p  
## 6 audi a4 2.8 1999 6 manual(m5) f 18 26 p  
## # ... with 1 more variables: class <chr>

Q2. 복사본 데이터를 이용해 cty는 city로, hwy는 highway로 수정하기

library(dplyr)  
mpg\_copy <- rename(mpg\_copy, city = cty, highway = hwy)

Q3. 변수명 바뀌었는지 확인

head(mpg\_copy)

## # A tibble: 6 x 11  
## manufacturer model displ year cyl trans drv city highway  
## <chr> <chr> <dbl> <int> <int> <chr> <chr> <int> <int>  
## 1 audi a4 1.8 1999 4 auto(l5) f 18 29  
## 2 audi a4 1.8 1999 4 manual(m5) f 21 29  
## 3 audi a4 2.0 2008 4 manual(m6) f 20 31  
## 4 audi a4 2.0 2008 4 auto(av) f 21 30  
## 5 audi a4 2.8 1999 6 auto(l5) f 16 26  
## 6 audi a4 2.8 1999 6 manual(m5) f 18 26  
## # ... with 2 more variables: fl <chr>, class <chr>

#### Doit\_part06 pdf

##### 06-2 혼자서해보기

Q1. 자동차 배기량에 따라 고속도로 연비가 다른지 알아보려고 합니다. displ(배기량)이 4 이하인 자동차와 5 이상인 자동차 중 어떤 자동차의 hwy(고속도로 연비)가 평균적으로 더 높은지 알아보세요.

displ4 <- as.data.frame(mpg %>% filter(displ <= 4))  
displ5 <- as.data.frame(mpg %>% filter(displ <= 5))  
  
mean\_hwy <- data.frame(data = c("displ4","displ5"),  
 Mean = c(mean(displ4$hwy),mean(displ5$hwy)))  
mean\_hwy

## data Mean  
## 1 displ4 25.96319  
## 2 displ5 24.40404

* displ이 4 이하인 자동차가 hwy가 평균적으로 더 높다.

Q2. "audi"와 "toyota" 중 어느 manufacturer의 cty가 평균적으로 더 높은지 알아보기

mpg\_audi <- as.data.frame(mpg %>% filter(manufacturer =="audi"))  
mpg\_toyo <- as.data.frame(mpg %>% filter(manufacturer == "toyota"))  
  
mean\_cty <- data.frame(data = c("mpg\_audi","mpg\_toyo"),  
 Mean = c(mean(mpg\_audi$cty),mean(mpg\_toyo$cty)))  
mean\_cty

## data Mean  
## 1 mpg\_audi 17.61111  
## 2 mpg\_toyo 18.52941

* toyota의 cty가 평균적으로 더 높다.

Q3. "chevrolet", "ford", "honda" 자동차의 고속도로 연비 평균을 알아보려고 합니다. 이 회사들의 자동차를 추출한 뒤 hwy 전체 평균을 구하라

mpg3cars <- as.data.frame(mpg %>% filter(manufacturer %in% c("chevrolet", "ford", "honda")))  
  
mean(mpg3cars$hwy)

## [1] 22.50943

#### 06-3 혼자서해보기

Q1. 이 중 일부만 추출해서 분석에 활용하려고 합니다. mpg 데이터에서 class(자동차 종류), cty(도시 연비) 변수를 추출해 새로운 데이터를 만들고 출력하라.

mpgcc <- as.data.frame(mpg %>% select("class","cty"))  
head(mpgcc)

## class cty  
## 1 compact 18  
## 2 compact 21  
## 3 compact 20  
## 4 compact 21  
## 5 compact 16  
## 6 compact 18

Q2. 앞에서 추출한 데이터를 이용해서 class(자동차 종류)가 "suv"인 자동차와 "compact"인 자동차 중 어떤 자동차의 cty(도시 연비)가 더 높은지 확인하라

mpg\_suv <- as.data.frame(mpg %>% filter(class == "suv"))  
mpg\_compact <- as.data.frame(mpg %>% filter(class == "compact"))  
  
mean\_cty2 <- data.frame(data = c("mpg\_suv","mpg\_compact"),  
 Mean =c(mean(mpg\_suv$cty),mean(mpg\_compact$cty)))  
mean\_cty2

## data Mean  
## 1 mpg\_suv 13.50000  
## 2 mpg\_compact 20.12766

* 자동차의 class가 compact인 자동차가 cty가 더 높다.

#### 06-4 혼자서해보기

mpg 데이터를 이용해서 "audi"에서 생산한 자동차 중 hwy가 1~5위에 해당하는 자동차의 데이터를 출력하라.

mpg %>% as.data.frame() %>%   
 filter(manufacturer =="audi") %>%  
 arrange(desc(hwy)) %>%  
 head(5)

## manufacturer model displ year cyl trans drv cty hwy fl class  
## 1 audi a4 2.0 2008 4 manual(m6) f 20 31 p compact  
## 2 audi a4 2.0 2008 4 auto(av) f 21 30 p compact  
## 3 audi a4 1.8 1999 4 auto(l5) f 18 29 p compact  
## 4 audi a4 1.8 1999 4 manual(m5) f 21 29 p compact  
## 5 audi a4 quattro 2.0 2008 4 manual(m6) 4 20 28 p compact

#### 06-5 혼자서해보기

Q1. mpg 데이터 복사본을 만들고, cty와 hwy를 더한 '합산 연비 변수'를 추가하라

mpg\_copy2 <- mpg %>% as.data.frame() %>%  
 mutate(total = cty + hwy)

Q2. 앞에서 만든 '합산 연비 변수'를 2 로 나눠 '평균 연비 변수'를 추가하라

mpg\_copy2 <- mpg\_copy2 %>% mutate(Mean = total/2)

Q3. '평균 연비 변수'가 가장 높은 자동차 3 종의 데이터를 출력하라

mpg\_copy2 %>% arrange(desc(Mean)) %>%  
 head(3)

## manufacturer model displ year cyl trans drv cty hwy fl  
## 1 volkswagen new beetle 1.9 1999 4 manual(m5) f 35 44 d  
## 2 volkswagen jetta 1.9 1999 4 manual(m5) f 33 44 d  
## 3 volkswagen new beetle 1.9 1999 4 auto(l4) f 29 41 d  
## class total Mean  
## 1 subcompact 79 39.5  
## 2 compact 77 38.5  
## 3 subcompact 70 35.0

Q4. 1~3 번 문제를 해결할 수 있는 하나로 연결된 dplyr 구문을 만들어 출하라. 데이터는 복사본 대신 mpg 원본을 이용하라

mpg %>% mutate(total = cty + hwy, Mean = total/2) %>%   
 arrange(desc(Mean)) %>%   
 head(3)

## # A tibble: 3 x 13  
## manufacturer model displ year cyl trans drv cty hwy  
## <chr> <chr> <dbl> <int> <int> <chr> <chr> <int> <int>  
## 1 volkswagen new beetle 1.9 1999 4 manual(m5) f 35 44  
## 2 volkswagen jetta 1.9 1999 4 manual(m5) f 33 44  
## 3 volkswagen new beetle 1.9 1999 4 auto(l4) f 29 41  
## # ... with 4 more variables: fl <chr>, class <chr>, total <int>,  
## # Mean <dbl>