Q4

rhadoop

2019년 3월 19일

# 문제4. popasian은 해당 지역의 아시아인 인구를 나타냅니다. ‘전체 인구 대비 아시아인 인구 백분율’ 변수를 추가하고, 하위 10개 지역의 state(주), county(지역명), 아시아인 인구 백분율을 출력하세요.

library(tidyverse)

## ─ Attaching packages ───────────────────────── tidyverse 1.2.1 ─

## ✔ ggplot2 3.1.0 ✔ purrr 0.2.4   
## ✔ tibble 2.1.1 ✔ dplyr 0.8.0.1  
## ✔ tidyr 0.8.3 ✔ stringr 1.2.0   
## ✔ readr 1.1.1 ✔ forcats 0.2.0

## ─ Conflicts ────────────────────────── tidyverse\_conflicts() ─  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

library(sqldf)

## Loading required package: gsubfn

## Loading required package: proto

## Loading required package: RSQLite

# 

### sql 해법

midwest\_sql <- as.data.frame(ggplot2::midwest)

sqldf("select state, county, cast(popasian as float)/cast(poptotal as float) \*100   
 as ratio\_asian from midwest\_sql order by ratio\_asian limit 10")

## state county ratio\_asian  
## 1 WI MENOMINEE 0.00000000  
## 2 IN BENTON 0.01059210  
## 3 IN CARROLL 0.01594981  
## 4 OH VINTON 0.02703190  
## 5 WI IRON 0.03250447  
## 6 IL SCOTT 0.05315379  
## 7 IN CLAY 0.06071645  
## 8 MI OSCODA 0.06375925  
## 9 OH PERRY 0.06654625  
## 10 IL PIATT 0.07074865

# 

# 

# 

### dplyr 해법

midwest <- midwest %>% mutate(ratio\_asian=(popasian/poptotal)\*100)  
midwest %>% select(state,county,ratio\_asian) %>% arrange(ratio\_asian) %>% head(10)

## # A tibble: 10 x 3  
## state county ratio\_asian  
## <chr> <chr> <dbl>  
## 1 WI MENOMINEE 0   
## 2 IN BENTON 0.0106  
## 3 IN CARROLL 0.0159  
## 4 OH VINTON 0.0270  
## 5 WI IRON 0.0325  
## 6 IL SCOTT 0.0532  
## 7 IN CLAY 0.0607  
## 8 MI OSCODA 0.0638  
## 9 OH PERRY 0.0665  
## 10 IL PIATT 0.0707

# 

# 

# 

# 

# 

### r syntax 해법

midwest\_r <- as.data.frame(ggplot2::midwest)

(midwest\_r$popasian/midwest\_r$poptotal)\*100 -> ratio\_asian  
midwest\_r<- data.frame(midwest\_r$state,midwest\_r$county,ratio\_asian)  
midwest\_r[order(midwest\_r$ratio\_asian),] -> midwest\_r  
midwest\_r[1:10,]

## midwest\_r.state midwest\_r.county ratio\_asian  
## 405 WI MENOMINEE 0.00000000  
## 106 IN BENTON 0.01059210  
## 110 IN CARROLL 0.01594981  
## 359 OH VINTON 0.02703190  
## 391 WI IRON 0.03250447  
## 86 IL SCOTT 0.05315379  
## 113 IN CLAY 0.06071645  
## 262 MI OSCODA 0.06375925  
## 341 OH PERRY 0.06654625  
## 74 IL PIATT 0.07074865