Q3

rhadoop

2019년 3월 27일

library(tidyverse)

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

## ✔ ggplot2 3.1.0 ✔ purrr 0.3.2   
## ✔ tibble 2.1.1 ✔ dplyr 0.8.0.1  
## ✔ tidyr 0.8.3 ✔ stringr 1.4.0   
## ✔ readr 1.3.1 ✔ forcats 0.4.0

## ─ 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

library(plotly)

##   
## Attaching package: 'plotly'

## The following object is masked from 'package:ggplot2':  
##   
## last\_plot

## The following object is masked from 'package:stats':  
##   
## filter

## The following object is masked from 'package:graphics':  
##   
## layout

# 

# 

# 나이 및 성별 월급 차이 (incomebyagebysex)

load("koweps/welfare.rda")

welfare %>% filter(!is.na(income)) %>%   
 group\_by(age,sex) %>%   
 select(sex, age,income) %>%   
 summarise(mean=mean(income)) -> incomebysexbyage  
  
ggplot(incomebysexbyage,aes(x=age, y=mean, fill=sex)) +   
 geom\_col(position = "dodge") +   
 labs(x="나이",y="평균 월급",fill="성별")

