Q2

rhadoop

2019년 3월 27일

### 나이와 월급의 관계(incomebyage) - 몇 살때 월급을 가장 많이 받을까?

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()

#### 1.데이터LOAD

load("koweps/welfare\_2016.rda")

#### 2.나이변수 전처리

table(is.na(welfarebirth),“birth”] %>% head(10) welfare[order(-welfare$birth),“birth”] %>% head(10)

#### 3.변수추가 및 나이계산

welfarebirth+1

#### 4.나이별 평균 월급요약표

welfare %>% filter(!is.na(income)) %>%   
 group\_by(age) %>%   
 summarise(mean=mean(income)) ->incomebyage

#### 5.라인그래프

ggplot(incomebyage,aes(age,mean))+geom\_line()

