HomeworkAPI

Yue Shang

4/6/2021

library(tidycensus)  
library(tidyverse)

## -- Attaching packages -----------

## v ggplot2 3.3.2 v purrr 0.3.4  
## v tibble 3.0.3 v dplyr 1.0.2  
## v tidyr 1.1.1 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.5.0

## -- Conflicts --------------------  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(dplyr)  
library(ggplot2)

# 1)   
census\_api\_key("242c6a1c9b843410c00133a330a9bb73463b6e80")

## To install your API key for use in future sessions, run this function with `install = TRUE`.

install=TRUE

# 2)   
## a)   
data1 <- load\_variables(2015, "acs5", cache = TRUE)  
view(data1)

data2<- get\_acs(geography = "county",   
 variables = c(mediancome = "B01001A\_011"),   
 state = "California",  
 year = 2015)

## Getting data from the 2011-2015 5-year ACS

data2

## # A tibble: 58 x 5  
## GEOID NAME variable estimate moe  
## <chr> <chr> <chr> <dbl> <dbl>  
## 1 06001 Alameda County, California mediancome 51644 667  
## 2 06003 Alpine County, California mediancome 50 26  
## 3 06005 Amador County, California mediancome 1809 72  
## 4 06007 Butte County, California mediancome 9962 128  
## 5 06009 Calaveras County, California mediancome 1927 74  
## 6 06011 Colusa County, California mediancome 1147 79  
## 7 06013 Contra Costa County, California mediancome 42756 605  
## 8 06015 Del Norte County, California mediancome 1629 90  
## 9 06017 El Dorado County, California mediancome 8609 141  
## 10 06019 Fresno County, California mediancome 34979 714  
## # ... with 48 more rows

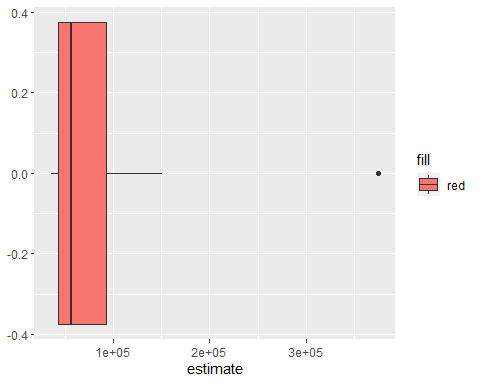
## b)   
data2 %>%   
 filter(estimate>30000) %>%   
 arrange(desc(estimate))->data3  
data3

## # A tibble: 13 x 5  
## GEOID NAME variable estimate moe  
## <chr> <chr> <chr> <dbl> <dbl>  
## 1 06037 Los Angeles County, California mediancome 375435 2332  
## 2 06073 San Diego County, California mediancome 150891 1008  
## 3 06059 Orange County, California mediancome 126819 1152  
## 4 06065 Riverside County, California mediancome 92346 1004  
## 5 06071 San Bernardino County, California mediancome 80925 1160  
## 6 06085 Santa Clara County, California mediancome 63036 879  
## 7 06067 Sacramento County, California mediancome 56066 553  
## 8 06001 Alameda County, California mediancome 51644 667  
## 9 06013 Contra Costa County, California mediancome 42756 605  
## 10 06075 San Francisco County, California mediancome 42307 542  
## 11 06029 Kern County, California mediancome 42121 575  
## 12 06111 Ventura County, California mediancome 41155 557  
## 13 06019 Fresno County, California mediancome 34979 714

## c)   
data3 %>%   
 filter(estimate==51644 , moe==667)->data4  
data4

## # A tibble: 1 x 5  
## GEOID NAME variable estimate moe  
## <chr> <chr> <chr> <dbl> <dbl>  
## 1 06001 Alameda County, California mediancome 51644 667

## d)   
ggplot(data3,aes(x=estimate),horizontal=TRUE)+  
 geom\_boxplot(aes(fill="red"))



## e)   
data3 %>%  
 mutate(NAME = gsub(" County, California", "", NAME)) %>%  
 ggplot(aes(x = estimate, y = reorder(NAME, estimate))) +  
 geom\_errorbarh(aes(xmin = estimate - moe, xmax = estimate + moe)) +  
 geom\_point(color = "blue", size = 2) +  
 labs(title = "Median Income for White Males by County",  
 subtitle = "2014-2018 American Community Survey",  
 y = "",  
 x = "ACS estimate (bars represent margin of error)")

