Homework9key

James Dickens

4/10/2021

library(tidycensus)  
library(tidyverse)

## -- Attaching packages ---------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.2 v purrr 0.3.4  
## v tibble 3.0.3 v dplyr 1.0.5  
## v tidyr 1.1.1 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.5.0

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

library(dplyr)  
library(ggplot2)  
census\_api\_key("2153faf1e1d25707ef71c8c464cb4c2c08be4e76")

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

install = TRUE  
  
v15 <- load\_variables(2015, "acs5", cache = TRUE)  
v15

## # A tibble: 22,768 x 3  
## name label concept   
## <chr> <chr> <chr>   
## 1 B00001\_001 Estimate!!Total UNWEIGHTED SAMPLE COUNT OF THE PO~  
## 2 B00002\_001 Estimate!!Total UNWEIGHTED SAMPLE HOUSING UNITS   
## 3 B01001\_001 Estimate!!Total SEX BY AGE   
## 4 B01001\_002 Estimate!!Total!!Male SEX BY AGE   
## 5 B01001\_003 Estimate!!Total!!Male!!Under 5~ SEX BY AGE   
## 6 B01001\_004 Estimate!!Total!!Male!!5 to 9 ~ SEX BY AGE   
## 7 B01001\_005 Estimate!!Total!!Male!!10 to 1~ SEX BY AGE   
## 8 B01001\_006 Estimate!!Total!!Male!!15 to 1~ SEX BY AGE   
## 9 B01001\_007 Estimate!!Total!!Male!!18 and ~ SEX BY AGE   
## 10 B01001\_008 Estimate!!Total!!Male!!20 years SEX BY AGE   
## # ... with 22,758 more rows

CA <- get\_acs(geography = "county",   
 variables = c(medincome = "B01001A\_011"),   
 state = "CA",   
 year = 2015)

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

CA

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

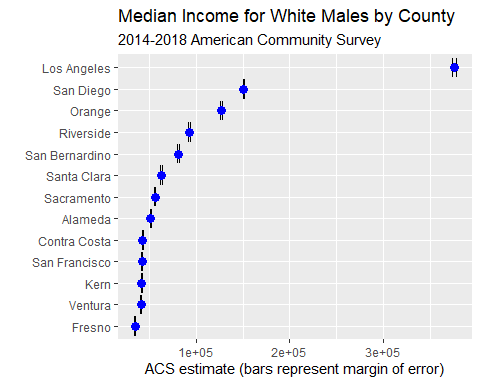
CA%>%  
 filter(estimate > 30000)%>%  
 arrange(desc(estimate)) -> CA1  
CA1

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

CA1%>%  
 filter(estimate == 51644 , moe == 667) -> CA2  
CA2

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

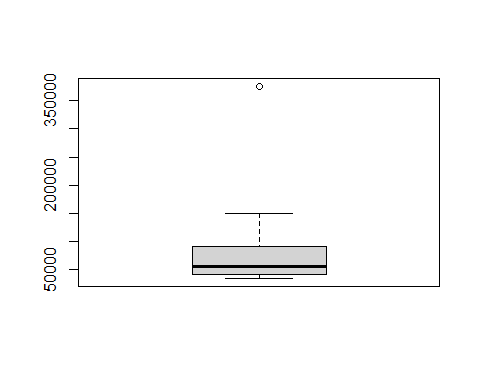
CA1 %>%  
 mutate(NAME = gsub(" County, California", "", NAME)) %>%  
 ggplot(aes(x = estimate, y = reorder(NAME, estimate))) +  
 geom\_errorbarh(aes(xmin = estimate - moe, xmax = estimate + moe)) +  
 geom\_point(color = "blue", size = 3) +  
 labs(title = "Median Income for White Males by County",  
 subtitle = "2014-2018 American Community Survey",  
 y = "",  
 x = "ACS estimate (bars represent margin of error)")



CA1

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

boxplot(CA1$estimate)



ggplot(data = CA1) +  
 geom\_boxplot(mapping = aes(y = estimate), fill = "red") +  
 coord\_flip()

