Week 11 NOV 3 - NOV 9 Lab 8 Race and Ethnic Inequality Lab

Lysa Vanible

November 03, 2021

#**Introduction** In this lab, we will explore the racial and ethnic make up of six states in the US using data from the 2014 American Community Surveys(ACS).

1). First replace “Write your name here” in the YAML header above with your name before you start. MAKE SURE NOT TO MODIFY ANY OTHER PART OF THE HEADER.

2). Again the same general rules used in the previous labs hold for this lab. Do not delete or modify anything unless it was written by you. All R code must be entered in the grey boxes/code chunks.

3). To run R code as you work your way through the assignment, highlight the code and click on the “Run” command in the R studio menu and then on “Run Selected Lines” or click on the green button at the top right of a particular code chunk to run all the code within it.

4). When you get to the end of the assignment you will knit the document to produce a pdf document using the Knit command in the Rstudio menu.

5). Continue reading from Part A.

knitr::opts\_chunk$set(error = TRUE)

PART A CREATE GEOGRAPHIES Similar to what we did in the internal migration lab, we’ll start by creating geographies. What this does is tell R which states we want data for. We first load the acs package which has the routines we want to use to create the geographies and then we create the 6 geographies. We put all the 6 geographies in an object that we call selectedstates.

library(acs)

## Loading required package: stringr

## Loading required package: XML

##   
## Attaching package: 'acs'

## The following object is masked from 'package:base':  
##   
## apply

california <- geo.make(state = "CA")  
florida <- geo.make(state = "FL")  
newyork <- geo.make(state = "NY")  
wyoming <- geo.make(state = "WY")  
westvirginia <- geo.make(state = "WV")  
kentucky <- geo.make(state = "KY")  
selectedstates <- california + florida + newyork + wyoming +  
westvirginia + kentucky

Then we pull the data from the ACS. Enter in the year 2014 for endyear to tell the ACS we want data for the year 2014. We extract the race data from the census bureau table B02001, which is located here <https://censusreporter.org/tables/B02001/> and place them in the variable race2014. We get the Hispanic Latino ethnic data from the census bureau table B03002, which is located here <https://censusreporter.org/tables/B03002/> and place them in the variable HispanicLatino.

#census\_api\_key(key="3d302849c18ce3b9aa077b96cc724d20d200bdb7", overwrite = TRUE)  
race2014 <- acs.fetch(geography = selectedstates, table.number = "B02001", endyear = 2014, col.names = "pretty", key="3d302849c18ce3b9aa077b96cc724d20d200bdb7", overwrite = TRUE)

## Warning in (function (endyear, span = 5, dataset = "acs", keyword, table.name, : temporarily downloading and using archived XML variable lookup files;  
## since this is \*much\* slower, recommend running  
## acs.tables.install()

HispanicLatino <- acs.fetch(geography = selectedstates, table.number = "B03002", endyear = 2014, col.names = "pretty", key="3d302849c18ce3b9aa077b96cc724d20d200bdb7", overwrite = TRUE)

## Warning in (function (endyear, span = 5, dataset = "acs", keyword, table.name, : temporarily downloading and using archived XML variable lookup files;  
## since this is \*much\* slower, recommend running  
## acs.tables.install()

The formatting of the variable names are a bit messy when we first pull them from the ACS, so we clean up those names. You’ll notice we’ve pulled two sets of data, one for race, and one for Hispanic or Latino ethnicity. ACS gives us the data with states in the first column and the variables across the first row. While this data is correct, the table is very wide and hard to read, so we transpose the datasets to have the variables we are interested in are listed down the first column and the states across the top. At the end of this chunk of code, you’ll display the transposed datasets as we have them so far. You’ll also notice when you knit to PDF, that even these transposed tables are very wide and may not fit on the page neatly. That’s ok.

race2014.names <- acs.colnames(race2014)  
new.names <- stringr::str\_sub(race2014.names, start = 7)  
acs.colnames(race2014) <- new.names  
race <- as.data.frame(estimate(race2014))  
HispanicLatino.names <- acs.colnames(HispanicLatino)  
new1.names <- stringr::str\_sub(HispanicLatino.names, start = 29)  
acs.colnames(HispanicLatino) <- new1.names  
HispanicLatino <- as.data.frame(estimate(HispanicLatino))  
racetransposed <- t(race)  
HispanicLatinotransposed <- t(HispanicLatino)  
racetransposed

## California  
## Total: 38066920  
## White alone 23650913  
## Black or African American alone 2262323  
## American Indian and Alaska Native alone 287360  
## Asian alone 5130536  
## Native Hawaiian and Other Pacific Islander alone 147286  
## Some other race alone 4890329  
## Two or more races: 1698173  
## Two or more races: Two races including Some other race 398497  
## Two or more races: Two races excluding Some other race, and three or more races 1299676  
## Florida  
## Total: 19361792  
## White alone 14747196  
## Black or African American alone 3114841  
## American Indian and Alaska Native alone 59121  
## Asian alone 490833  
## Native Hawaiian and Other Pacific Islander alone 12128  
## Some other race alone 484274  
## Two or more races: 453399  
## Two or more races: Two races including Some other race 86393  
## Two or more races: Two races excluding Some other race, and three or more races 367006  
## New York  
## Total: 19594330  
## White alone 12733518  
## Black or African American alone 3060048  
## American Indian and Alaska Native alone 74572  
## Asian alone 1522549  
## Native Hawaiian and Other Pacific Islander alone 6258  
## Some other race alone 1649123  
## Two or more races: 548262  
## Two or more races: Two races including Some other race 155051  
## Two or more races: Two races excluding Some other race, and three or more races 393211  
## Wyoming  
## Total: 575251  
## White alone 522535  
## Black or African American alone 5958  
## American Indian and Alaska Native alone 13144  
## Asian alone 5377  
## Native Hawaiian and Other Pacific Islander alone 332  
## Some other race alone 12387  
## Two or more races: 15518  
## Two or more races: Two races including Some other race 2242  
## Two or more races: Two races excluding Some other race, and three or more races 13276  
## West Virginia  
## Total: 1853881  
## White alone 1735816  
## Black or African American alone 59129  
## American Indian and Alaska Native alone 3075  
## Asian alone 13144  
## Native Hawaiian and Other Pacific Islander alone 645  
## Some other race alone 5151  
## Two or more races: 36921  
## Two or more races: Two races including Some other race 1561  
## Two or more races: Two races excluding Some other race, and three or more races 35360  
## Kentucky  
## Total: 4383272  
## White alone 3845535  
## Black or African American alone 345035  
## American Indian and Alaska Native alone 9000  
## Asian alone 53448  
## Native Hawaiian and Other Pacific Islander alone 2109  
## Some other race alone 41518  
## Two or more races: 86627  
## Two or more races: Two races including Some other race 7448  
## Two or more races: Two races excluding Some other race, and three or more races 79179

Since there are spaces in the variable names, we’ll need to change those spaces to periods to make them easier to use in the code. We’ll also want to get rid of other odd characters like colons and commas that R has trouble with. The code below cleans up the variables names to make them usable.

new1.names <- str\_replace\_all(str\_trim(new1.names), " ", ".")  
new1.names <- str\_replace\_all(str\_trim(new1.names), ":", "")  
new1.names <- str\_replace\_all(str\_trim(new1.names), ",", "")  
colnames(HispanicLatino) <- new1.names  
new.names <- str\_replace\_all(str\_trim(new.names), " ", ".")  
new.names <- str\_replace\_all(str\_trim(new.names), ":", "")  
new.names <- str\_replace\_all(str\_trim(new.names), ",", "")  
colnames(race) <- new.names

It is easier to work with a single dataset. So, we are taking the two variables we want from the HispanicLatino dataset and adding them to the race dataset. The two variables we are adding are Not.Hispanic.or.Latino and Hispanic.or.Latino.

race$Not.Hispanic.or.Latino <- HispanicLatino$Not.Hispanic.or.Latino   
race$Hispanic.or.Latino <- HispanicLatino$Hispanic.or.Latino

The ACS provides the data in counts (number of people), but it is easier to interpret percents, especially if we are comparing across states. Below, we calculate percentages for the variables we are interested in using. We compute the percentages by taking the count of the race-ethnic group and dividing it by the total population of the state and multiplying by 100.

race$percent\_total <- round(100\*race$Total/race$Total)  
race$percent\_White.alone <- round(100\*race$White.alone/race$Total)  
race$percent\_Black.alone <-round(100\*race$Black.or.African.American.alone/race$Total)  
race$percent\_American.Indian.alone <-round(100\*race$American.Indian.and.Alaska.Native.alone/  
race$Total)  
race$percent\_Asian.alone <-round(100\*race$Asian.alone/race$Total)  
race$percent\_Pacific.Islander.alone <-round(100\*race$Native.Hawaiian.and.Other.Pacific.Islander.alone/  
race$Total)  
race$percent\_Some.other.race.alone <-round(100\*race$Some.other.race.alone/race$Total)  
race$percent\_Two.or.more.races <-round(100\*race$Two.or.more.races/race$Total)  
race$percent\_Not.Hispanic.or.Latino <-round(100\*race$Not.Hispanic.or.Latino/race$Total)  
race$percent\_Hispanic.or.Latino <-round(100\*race$Hispanic.or.Latino/race$Total)

Now, to separate out the new variables we created with percents, we create a new dataset that only includes the new variables measured in percents. Then, since the datasets are easier to read if we flip the columns and rows, we transpose the new dataset so states are in the columns and race-ethnic groups are in the rows. The last line of code will display your completed table.

percentrace <- dplyr::select(race, percent\_total,percent\_White.alone, percent\_Black.alone,  
percent\_American.Indian.alone, percent\_Asian.alone, percent\_Pacific.Islander.alone,  
percent\_Some.other.race.alone, percent\_Two.or.more.races, percent\_Not.Hispanic.or.Latino,  
percent\_Hispanic.or.Latino)  
percentracetransposed <- t(percentrace)  
percentracetransposed

## California Florida New York Wyoming  
## percent\_total 100 100 100 100  
## percent\_White.alone 62 76 65 91  
## percent\_Black.alone 6 16 16 1  
## percent\_American.Indian.alone 1 0 0 2  
## percent\_Asian.alone 13 3 8 1  
## percent\_Pacific.Islander.alone 0 0 0 0  
## percent\_Some.other.race.alone 13 3 8 2  
## percent\_Two.or.more.races 4 2 3 3  
## percent\_Not.Hispanic.or.Latino 62 77 82 91  
## percent\_Hispanic.or.Latino 38 23 18 9  
## West Virginia Kentucky  
## percent\_total 100 100  
## percent\_White.alone 94 88  
## percent\_Black.alone 3 8  
## percent\_American.Indian.alone 0 0  
## percent\_Asian.alone 1 1  
## percent\_Pacific.Islander.alone 0 0  
## percent\_Some.other.race.alone 0 1  
## percent\_Two.or.more.races 2 2  
## percent\_Not.Hispanic.or.Latino 99 97  
## percent\_Hispanic.or.Latino 1 3

PART F CONCLUSION

Great work!!!! You are almost at the end of your eighth lab. You will now prepare to knit the lab. Click on the “Knit” command at the top of this window to turn this document with your code and all the results into a pdf document. When the knit is complete, Rstudio will open the pdf document. Inspect it very carefully to make sure that you have not missed any part of the assignment. If you have missed any tasks please go back and fix them and knit the document again. Once the pdf document is knitted and you are sure you have completed all the assignments in the lab prepare to submit it to the week 11 folder of SOC339’s blackboard site. Before submitting a few more things to do.

1). First save your rmarkdown file (the file you have been working on in Rstudio) by clicking on the FILE command in the Rstudio menu and then on SAVE. You will see the rmarkdown document in the FILES folder of the window to the bottom right of your screen. It will be named Lab 8 Race and Ethnicity Lab.Rmd.

2). To save the pdf file, locate it in the FILES folder of the window to the bottom right of your screen. It will be named Lab 8 Race and Ethnicity Lab.pdf. Click the little box to the left of it. Then click on the MORE tab near the top of the FILES window. Then on EXPORT and then when prompted by RSTUDIO click on DOWNLOAD. RSTUDIO will save the lab 8 Race and Ethnicity Lab.pdf to the DOWNLOAD folder of your computer.

3). Go to the WEEK 11 folder in blackboard and follow the instructions inside that folder to upload the Lab 8 Race and Ethnicity Lab.pdf and answer the remaining questions in this lab.