

Assignment Exemplar

Libraries

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.5.3
```

```
## -- Attaching packages ----- tidyverse 1.2.1 --
```

```
## v ggplot2 3.2.1    v purrr  0.3.2
## v tibble  2.1.3    v dplyr  0.8.1
## v tidyr   0.8.3    v stringr 1.4.0
## v readr   1.3.1    v forcats 0.4.0
```

```
## Warning: package 'ggplot2' was built under R version 3.5.3
```

```
## Warning: package 'tibble' was built under R version 3.5.3
```

```
## Warning: package 'tidyr' was built under R version 3.5.3
```

```
## Warning: package 'purrr' was built under R version 3.5.3
```

```
## Warning: package 'dplyr' was built under R version 3.5.3
```

```
## Warning: package 'stringr' was built under R version 3.5.3
```

```
## Warning: package 'forcats' was built under R version 3.5.3
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
```

```
## x dplyr::lag()     masks stats::lag()
```

```
library(rvest)
```

```
## Warning: package 'rvest' was built under R version 3.5.3
```

```
## Loading required package: xml2
```

```
## Warning: package 'xml2' was built under R version 3.5.3
```

```
##
## Attaching package: 'rvest'

## The following object is masked from 'package:purrr':
##
##   pluck

## The following object is masked from 'package:readr':
##
##   guess_encoding
```

```
library(acs)
```

```
## Warning: package 'acs' was built under R version 3.5.3

## Loading required package: XML

## Warning: package 'XML' was built under R version 3.5.3

##
## Attaching package: 'XML'

## The following object is masked from 'package:rvest':
##
##   xml

##
## Attaching package: 'acs'

## The following object is masked from 'package:dplyr':
##
##   combine

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

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':
##
##   date
```

```
library(noncensus)
library(tigris)
```

```
## Warning: package 'tigris' was built under R version 3.5.3

## To enable
## caching of data, set `options(tigris_use_cache = TRUE)` in your R script or .Rprofile.

##
## Attaching package: 'tigris'

## The following object is masked from 'package:graphics':
##
##      plot
```

1. Download data for all of the zip codes in Los Angeles county on education levels.

```
## Look up fips code for county
lookup_code("CA","Los Angeles")
```

```
## [1] "The code for California is '06' and the code for Los Angeles County is '037'."
```

```
state_fips<-"06"
county_stub<-"037"
```

Next, we'll combine the state and county fips into a single object

```
county_fips<-paste0(state_fips,county_stub)
```

```
# Get dataset that matches all zip codes to cities, counties and states.
county_to_zip<-read_csv("http://www2.census.gov/geo/docs/maps-data/data/rel/zcta_county_rel_10.txt")
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   ZCTA5 = col_character(),
##   COUNTY = col_character()
## )
```

```
## See spec(...) for full column specifications.
```

```
save(county_to_zip,file="county_to_zip.Rdata")

#easier names to work with
names(county_to_zip)<-tolower(names(county_to_zip))

#Just zip codes in selected county
county_to_zip<-county_to_zip%>%
  filter(state==as.numeric(state_fips),county==county_stub)%>%
```

```

select(zcta5,state,county)

#list of zip codes
ziplist<-county_to_zip$zcta5

#City names
data(zip_codes)

city_zip<-zip_codes%>%filter(zip%in%ziplist)%>%select(zip,city)

#Arrange in order
city_zip<-city_zip%>%arrange(as.numeric(zip))

```

2. Compute the proportion of the population that has a bachelor's degree or above by zip code.

```

# Get your own key and save as my_acs_key.txt
#my_acs_key<-readLines("my_acs_key.txt",warn = FALSE)
#acs_key<-my_acs_key

acs_key<-"b27a265fe0dc7c49bd9281d6bc778637f10685e3"

#List of tables: https://www.census.gov/programs-surveys/acs/technical-documentation/summary-file-documentation
# b15002: education of pop over 25, by sex
# b19001: household income over last 12 months

api.key.install(acs_key, file = "key.rda")

select_zip<-geo.make(zip.code=ziplist)

county_educ=acs.fetch(geography=select_zip,
                      endyear=2016,
                      table.number="B15002",
                      col.names="pretty",verbose=T) # <---- this may take a long time :(

## Warning in (function (endyear, span = 5, dataset = "acs", keyword, table.name, : acs.lookup for endyear
## (See ?acs.lookup for details)

## Warning in (function (endyear, span = 5, dataset = "acs", keyword, table.name, : temporarily downloading
## since this is *much* slower, recommend running
## acs.tables.install()

#save(county_educ,file="county_educ_la.Rdata")
acs.colnames(county_educ)

## [1] "Sex by Educational Attainment for the Population 25 Years and over: Total:"
## [2] "Sex by Educational Attainment for the Population 25 Years and over: Male:"
## [3] "Sex by Educational Attainment for the Population 25 Years and over: Male: No schooling completed"
## [4] "Sex by Educational Attainment for the Population 25 Years and over: Male: Nursery to 4th grade"

```

```
## [5] "Sex by Educational Attainment for the Population 25 Years and over: Male: 5th and 6th grade"
## [6] "Sex by Educational Attainment for the Population 25 Years and over: Male: 7th and 8th grade"
## [7] "Sex by Educational Attainment for the Population 25 Years and over: Male: 9th grade"
## [8] "Sex by Educational Attainment for the Population 25 Years and over: Male: 10th grade"
## [9] "Sex by Educational Attainment for the Population 25 Years and over: Male: 11th grade"
## [10] "Sex by Educational Attainment for the Population 25 Years and over: Male: 12th grade, no diploma"
## [11] "Sex by Educational Attainment for the Population 25 Years and over: Male: High school graduate"
## [12] "Sex by Educational Attainment for the Population 25 Years and over: Male: Some college, less than 1 year"
## [13] "Sex by Educational Attainment for the Population 25 Years and over: Male: Some college, 1 or more years"
## [14] "Sex by Educational Attainment for the Population 25 Years and over: Male: Associate's degree"
## [15] "Sex by Educational Attainment for the Population 25 Years and over: Male: Bachelor's degree"
## [16] "Sex by Educational Attainment for the Population 25 Years and over: Male: Master's degree"
## [17] "Sex by Educational Attainment for the Population 25 Years and over: Male: Professional school graduate"
## [18] "Sex by Educational Attainment for the Population 25 Years and over: Male: Doctorate degree"
## [19] "Sex by Educational Attainment for the Population 25 Years and over: Female:"
## [20] "Sex by Educational Attainment for the Population 25 Years and over: Female: No schooling completed"
## [21] "Sex by Educational Attainment for the Population 25 Years and over: Female: Nursery to 4th grade"
## [22] "Sex by Educational Attainment for the Population 25 Years and over: Female: 5th and 6th grade"
## [23] "Sex by Educational Attainment for the Population 25 Years and over: Female: 7th and 8th grade"
## [24] "Sex by Educational Attainment for the Population 25 Years and over: Female: 9th grade"
## [25] "Sex by Educational Attainment for the Population 25 Years and over: Female: 10th grade"
## [26] "Sex by Educational Attainment for the Population 25 Years and over: Female: 11th grade"
## [27] "Sex by Educational Attainment for the Population 25 Years and over: Female: 12th grade, no diploma"
## [28] "Sex by Educational Attainment for the Population 25 Years and over: Female: High school graduate"
## [29] "Sex by Educational Attainment for the Population 25 Years and over: Female: Some college, less than 1 year"
## [30] "Sex by Educational Attainment for the Population 25 Years and over: Female: Some college, 1 or more years"
## [31] "Sex by Educational Attainment for the Population 25 Years and over: Female: Associate's degree"
## [32] "Sex by Educational Attainment for the Population 25 Years and over: Female: Bachelor's degree"
## [33] "Sex by Educational Attainment for the Population 25 Years and over: Female: Master's degree"
## [34] "Sex by Educational Attainment for the Population 25 Years and over: Female: Professional school graduate"
## [35] "Sex by Educational Attainment for the Population 25 Years and over: Female: Doctorate degree"
```

```
## Proportion of individuals at college or above=
## number with college degree/
## total number
prop_coll_above<-divide.acs(numerator=(county_educ[,15]+
                                county_educ[,16]+
                                county_educ[,17]+
                                county_educ[,18]+
                                county_educ[,32]+
                                county_educ[,33]+
                                county_educ[,34]+
                                county_educ[,35]),
                                denominator=county_educ[,1]
)### Double check -- are these the correct columns??!??
```

```
#prop with educational attainment=number with level of ed attain/total pop
```

3. Download data for all of the zip codes in LA county on family income by zip code.

Family Income Data

```
# 19001-- family income
county_income<-acs.fetch(geography=select_zip,
                        endyear = 2016,
                        table.number="B19001",
                        col.names="pretty")

## Warning in (function (endyear, span = 5, dataset = "acs", keyword, table.name, : acs.lookup for endyear
##   (See ?acs.lookup for details)

## Warning in (function (endyear, span = 5, dataset = "acs", keyword, table.name, : temporarily downloading
##   since this is *much* slower, recommend running
##   acs.tables.install()
```

```
acs.colnames(county_income)
```

```
## [1] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): Total:"
## [2] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): Less than
## [3] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $10,000 to
## [4] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $15,000 to
## [5] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $20,000 to
## [6] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $25,000 to
## [7] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $30,000 to
## [8] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $35,000 to
## [9] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $40,000 to
## [10] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $45,000 to
## [11] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $50,000 to
## [12] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $60,000 to
## [13] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $75,000 to
## [14] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $100,000 to
## [15] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $125,000 to
## [16] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $150,000 to
## [17] "B19001. Household Income in the Past 12 Months (in 2015 Inflation-Adjusted Dollars): $200,000 to
```

4. Compute the proportion of the population that has family income above 75,000.

```
#Proportion above 75k--
prop_above_75<-divide.acs(numerator=(county_income[,13]+
                                county_income[,14]+
                                county_income[,15]+
                                county_income[,16]+
                                county_income[,17]),
                          denominator=county_income[,1]
                          )

# Convert to tibble
county_df<-tibble(substr(geography(county_educ)[[1]],7,11),
                  as.numeric(estimate(prop_coll_above)),
```

```

        as.numeric(estimate(prop_above_75))
    )

    # Give it easy to use names
    names(county_df)<-c("zip","college_educ","income_75")
    save(county_df,file="dav.RData")

    head(county_df)

```

```

## # A tibble: 6 x 3
##   zip    college_educ income_75
##   <chr>         <dbl>     <dbl>
## 1 90001         0.0443     0.160
## 2 90002         0.0528     0.157
## 3 90003         0.0527     0.140
## 4 90004         0.347      0.263
## 5 90005         0.278      0.146
## 6 90006         0.163      0.116

```

5. Plot the proportion of residents with incomes above 75,000 as a function of income.

```

gg<-ggplot(county_df,aes(x=college_educ,y=income_75))
gg<-gg+geom_point()
gg<-gg+xlab("Prop. College Education")+ylab("Prop. Income Above 75k")+ggtitle("Education & Income By Co
gg

```

```

## Warning: Removed 13 rows containing missing values (geom_point).

```

Education & Income By County

