Instructions for accessing 2019 PUMS data in R

If you already have R/RStudio installed, please skip to step 3.

1. Download R
   - Go to www.r-project.org and download R.
2. Download R Studio
   - Go to www.rstudio.com and download the R studio IDE.
3. You will need to get an API key from the Census.
   - Go to https://api.census.gov/data/key_signup.html
   - Fill out form
     - Organization Name: **School Name**
     - Email address: **School email address (or personal)**
     - Agree to terms of services
     - Click "Submit Key request"
4. You will receive an email telling you to activate your API key.
   - Click the link to activate your key.
5. The email will also give you your API Key.
   - Copy the key value
   - Go to RStudio
   - Enter command `census_api_key("KEYVALUE", FALSE, TRUE)`
   - Refresh your R environment with command `readRenviron("~/.Renviron")`
   - Run a "get" command to confirm that you set your census key and it worked with the command `Sys.getenv("CENSUS_API_KEY")`
6. You will need to install a few packages. Please run the following commands:
   - `install.packages(c("survey", "srvyr"))`
   - `install.packages(c("tidyverse", "tidycensus"))`
7. Load and attach the following packages:
   - `library(srvyr, warn.conflicts = FALSE)`
   - `library(tidyverse)`
   - `library(tidycensus)`
8. Next, we will load the variables for the PUMS data.
   - `pums_var_2019 <- pums_variables %>% filter(year == 2019, survey = "acs1")`
9. To see these variables run the following command:
   - `pums_var_2019 <- %>% distinct(var_code, var_label, data_type, level)`
   - Take a look at all of variables available to you and their descriptions
10. It is important to note, some of these variables are at a housing unit level and some are at a person level. To see person level run the following:
    - `pums_var_2019 %>% distinct(var_code, var_label, data_type, level) %>% filter(level=="person")`

11. To access the PUMS data, you will need to run a command similar to below. Please note, this command is only getting a few variables, you will need to dictate in your command which variables you would like to get.
    - `tx_pums <- get_pums(variables = c("PUMA", "SEX", "AGEP", "SCHL"), state = TX, survey = "acs1", year = 2019)`
    - That data is now stored inside of tx_pums and ready for analysis. Remember you can change which variables and which state.

These instructions are inspired and modified from [www.walker-data.com/tidycensus/articles/pums-data.html.](https://www.walker-data.com/tidycensus/articles/pums-data.html) Please go to their website for more details and further instructions if desired.