Shared Prosperity Analysis

Data Sources:

1. American Community Survey microdata. Downloaded from [IPUMS USA](https://usa.ipums.org/usa/)
   * Select the following samples:
     + 2018 ACS
     + 2019 ACS
   * Select the variables specified in ACS\_vars.xslx
   * Download as a .dta file
   * Save as acs\_2018\_2019.dta
2. MSA PUMA Crosswalk. Downloaded from IPUMS [here](https://usa.ipums.org/usa-action/variables/met2013#description_section)
   * This data set helps match the cost-of-living adjustments (RPP data) to the ACS microdata.
3. Regional Price Parity by Metropolitan Statistical Area (MARPP\_MSA\_2008\_2019.csv)
   * Download from [BEA Regional Accounts](https://apps.bea.gov/regional/downloadzip.cfm)
     + Under *Personal Income (State and Local)* select *MARPP tables: Real Personal Income by MSA.*
4. Regional Price Parity for Metro/Nonmetropolitan Portion (PARPP\_PORT\_2008\_2019.csv)
   * Download from [BEA Regional Accounts](https://apps.bea.gov/regional/downloadzip.cfm)
     + Under *Personal Income (State and Local)* select *PARPP tables: Real Personal Income by Metro/Nonmetro Portion.*
5. Personal Consumption Expenditure: Chain-type Price Index (Annual Frequency, download from [FRED](https://fred.stlouisfed.org/series/PCEPI))
   * Save as PCE\_annual.csv

File Strucutre and Code:

1. There should be 3 subfolders for containing data
   1. Raw\_data – This contains the 5 datasets above all downloaded directly from the internet
   2. Intermediate\_data – Contains various temporary datasets created in analysis.
   3. Final\_data – Contains two CSV files
      1. income\_by\_puma.csv – Displays mean and median income for 2018 and 2019 for every PUMA in the SEMCOG region.
      2. Income\_categories.csv - Contains the percent of SEMCOG residents classified as lower income, middle income, and higher income for years 2018/2019.
2. Code – All the code is in *semcog\_shared\_prosperity.do*
   1. To run the code you will need to change line 15 to reflect the file path where the code is on your library.
   2. Lines 1-159 are simply setting up the environment and cleaning the regional price/PCE data. The ACS data is read in on line 169 and core analysis follows from there.