**Description of Variable Creation**

**1. Population**

* **Variable**: Total\_Population  
  **Description**: Represents the total population for each ZCTA and year. Extracted from the column POPULATION\_Estimate Total and renamed to Total\_Population.

**2. Poverty**

* **Variables**: Poverty\_Count, Poverty\_Population, Poverty\_Rate, Child\_Poverty\_Count, Child\_Poverty\_Rate  
  **Description**:
  + Poverty\_Count: Total number of individuals with income below the poverty level.
  + Poverty\_Population: Total population considered in the poverty estimate.
  + Child\_Poverty\_Count: Total count of children living below the poverty level, calculated by summing various child age groups (under 5 to 17 years) for both genders.
  + Poverty\_Rate: Calculated as the ratio of Poverty\_Count to Poverty\_Population.
  + Child\_Poverty\_Rate: Calculated as the ratio of Child\_Poverty\_Count to Poverty\_Population.

**3. Median Income**

* **Variable**: Median\_Income  
  **Description**: Represents the median household income in inflation-adjusted dollars. Extracted from the column INCOME\_Estimate Median household income in the past 12 months (in 2011 inflation-adjusted dollars).

**4. Race and Ethnicity**

* **Variables**: Hispanic, White\_non\_Hispanic, Black\_non\_Hispanic, Asian\_non\_Hispanic, Other\_non\_Hispanic, Race\_Population  
  **Description**:
  + Hispanic: Population identified as Hispanic or Latino.
  + White\_non\_Hispanic: Population identified as White but not Hispanic or Latino.
  + Black\_non\_Hispanic: Population identified as Black or African American but not Hispanic or Latino.
  + Asian\_non\_Hispanic: Population identified as Asian but not Hispanic or Latino.
  + Other\_non\_Hispanic: Sum of populations from other race categories (Native American, Pacific Islander, etc.).
  + Race\_Population: Total population used for race/ethnicity estimates.

**5. Sex**

* **Variables**: Male, Female, Sex\_Population  
  **Description**:
  + Male: Total number of males.
  + Female: Total number of females.
  + Sex\_Population: Total population used for sex estimates, extracted from the column SEX\_Estimate Total.

**6. Age Groups**

* **Variables**: Age\_0\_4, Age\_5\_9, Age\_10\_17, Age\_18\_24, Age\_25\_34, Age\_35\_44, Age\_45\_54, Age\_55\_64, Age\_65\_Plus, Age\_Population  
  **Description**:  
  Age group variables are created by summing the male and female population estimates for various age ranges (e.g., 0-4 years, 5-9 years, etc.). The total population across all age groups is captured in Age\_Population.

**7. Education**

* **Variables**: No\_HS, HS\_Graduate, Some\_College, College\_Graduate, Education\_Population  
  **Description**:
  + No\_HS: Population with less than a high school education.
  + HS\_Graduate: Population that graduated from high school, including GED recipients.
  + Some\_College: Population with some college experience or an associate’s degree.
  + College\_Graduate: Population with a bachelor’s degree or higher.
  + Education\_Population: Total population used for education estimates.

**8. Marital Status**

* **Variables**: Never\_Married, Married\_Spouse\_Present, Married\_Spouse\_Absent, Widowed, Divorced, Marital\_Status\_Population  
  **Description**:
  + Never\_Married: Individuals who have never been married.
  + Married\_Spouse\_Present: Individuals who are married and have a spouse present.
  + Married\_Spouse\_Absent: Individuals who are married but have an absent spouse.
  + Widowed: Individuals who are widowed.
  + Divorced: Individuals who are divorced.
  + Marital\_Status\_Population: Total population used for marital status estimates.

**9. Employment and Industry**

* **Variables**: Employment\_Population, Employed, Unemployed, Armed\_Forces, Not\_in\_Labor\_Force, Agriculture, Construction, Manufacturing, Wholesale\_Trade, Retail\_Trade, Transportation\_Utilities, Information, Finance, Professional\_Services, Education\_Health, Arts\_Entertainment, Other\_Services, Public\_Administration  
  **Description**:  
  Employment and industry variables capture labor force participation and employment across different industries for populations aged 16 years and older.
  + Employment\_Population: Total population aged 16 years and older in the labor force estimate.
  + Employed: Civilian labor force employed.
  + Unemployed: Civilian labor force unemployed.
  + Armed\_Forces: Individuals in the armed forces.
  + Not\_in\_Labor\_Force: Individuals not participating in the labor force.
  + **Industry variables**: Capture employment across various sectors like agriculture, construction, retail trade, etc.

**10. Rates**

**Rate Variables Created**:  
For each category, rates are calculated by dividing the relevant subpopulation by the total population for that category. For example:

* Poverty\_Rate: Calculated as Poverty\_Count / Poverty\_Population.
* Employment\_Rate: Calculated as Employed / Employment\_Population.
* Race Rates: Calculated as the population of each race (e.g., Hispanic, White\_non\_Hispanic) divided by Race\_Population.
* **Industry Rates**: The percentage of the population employed in each industry is calculated by dividing the number employed in that industry by Employment\_Population.

**11. Weighted Averages**

**Weighted Variables Created**:  
Weighted averages are calculated for each category to provide an aggregate measure across years or ZCTAs. The weights used are typically the total population for each variable. For example:

* Weighted\_Poverty\_Rate: Calculated as the weighted mean of Poverty\_Rate using Poverty\_Population as the weight.
* Weighted\_Median\_Income: Calculated as the weighted mean of Median\_Income using Total\_Population as the weight.
* Weighted\_Race Rates: Each race rate (e.g., Hispanic\_Rate, White\_non\_Hispanic\_Rate) is averaged across ZCTAs using Race\_Population as the weight.