

Python Project

Visualization

- Rural / Urban Disparities
 - Industry Composition
 -

Model Building

1. Regression (Median Income / House Value)
2. Clustering
 - a. different counties

Data

1. Education
 - a. High School Graduation Rate
 - i. Population over 25 years Highschool Graduation
 - b. Bachelors or Higher Rate
 - i. Population over 25 years Bachelor or Higher
 - c. Youth School / unemployment
 - i. "S0902_C01_017E","Estimate!!Total!!Population 16 to 19 years!!IDLENESS!!Not enrolled in school and not in the labor force"
 - d. Stem vs non stem degree percentage
2. Employment
 - a. Mean usual work hours
 - i. Work Status: Usual Hours worked
3. Economic
4. Infrastructure Quality

Variables:

- Geo Data
 - Dataset
 - geo_county_2025.csv
 - Data updated as of 24.01.2025.
 - <https://simplemaps.com/data/us-counties>
 - All except in US territories (3144)
 - Variables

- lat
- lng
- population
- county
- county_fips
- state
- Urban Rural Classification (Continuum Codes)
 - Dataset
 - ruralurban_county_2023.csv
 - 2023 Data (Last Updated 22.01.2024)
 - <https://www.ers.usda.gov/data-products/rural-urban-continuum-codes>
 - <https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/documentation>
 - Variables
 - FIPS
 - State
 - County_Name
 - Attribute (filter for RUCC_2023)
 - Value
 - 1,2,3 Metro
 - 4,5,6,7,8,9 Nonmetro
- Unemployment Rate
 - Dataset
 - unemployment_county_2023.csv
 - <https://www.ers.usda.gov/data-products/county-level-data-sets/county-level-data-sets-download-data>
 - Variables
 - FIPS_Code
 - State
 - filter out US and State cumulative
 - Area_Name
 - Attribute (filter for Unemployment_rate_2023)
 - Value
- Poverty Rate
 - Dataset
 - poverty_county_2023.csv

- <https://www.ers.usda.gov/data-products/county-level-data-sets/county-level-data-sets-download-data>
- Variables
 - FIPS_Code
 - Stabr (State)
 - filter out US and State cumulative
 - Area_Name
 - Attribute (filter for PCTPOVALL_2023)
 - PCTPOVALL_2023 = Poverty rate (%) for the total population in 2023
 - Value
- Zillow Home Value Index (ZHVI)
 - Dataset
 - zillowhousevalue_county_2023.csv
 - ZHVI All Homes (SFR, Condo/Co-op) Time Series, Smoothed, Seasonally Adjusted(\$)
 - <https://www.zillow.com/research/data/>
 - It reflects the typical value for homes in the 35th to 65th percentile range.
 - Variables
 - RegionName
 - StateName
 - FIPS Code (merge these 2)
 - StateCodeFIPS
 - MunicipalCodeFIPS
 - 2023-12-31 (house values)
- Median Population Age
 - Dataset
 - age_county_2023.csv
 - 2023: 5 Year Estimates
 - Age and Sex
 - https://data.census.gov/table/ACSST5Y2023.S0101?q=county_age
 - Median Age
 - "S0101_C01_032E","Estimate!!Total!!Total population!!SUMMARY INDICATORS!!Median age (years)"

US Census:

https://www.census.gov/content/dam/Census/library/publications/2020/acs/acs_api_handbook_2020_ch02.pdf
<https://censusreporter.org/topics/table-codes/>

- Internet Access in Households

- Dataset
 - internet_county_2023.csv
 - 2023: 5 Year Estimates
 - Internet Access in Housholds
 - https://data.census.gov/table/ACSDT5Y2023.B28011?q=county_subscription
- Variables
 - Total Households
 - "B28011_001E","Estimate!!Total:"
 - Households with Internet Subscription
 - "B28011_002E","Estimate!!Total!!With an Internet subscription"
- Output Variables
 - Internet Access Percentage
 - Pct_HH_Internet_Subscription
 - =
- Household Size
 - Dataset
 - 2023: 5 Year Estimates
 - Selected Social Characteristics in the United States
 - https://data.census.gov/table/ACSDP5Y2023.DP02?q=county_education&y=2023
 - Total Households
 - "DP02_0001E","Estimate!!HOUSEHOLDS BY TYPE!!Total households"
 - Average Household Size
 - "DP02_0016E","Estimate!!HOUSEHOLDS BY TYPE!!Total households!!Average household size"
- Bachelors or Higher Percentage
 - Dataset
 - education_county_2023.csv
 - 2023: 5 Year Estimates
 - Educational Attainment
 - https://data.census.gov/table/ACSST5Y2023.S1501?q=county_education&y=2023
 - Variables
 - Total Population 25y and over
 - "S1501_C01_006E","Estimate!!Total!!AGE BY EDUCATIONAL ATTAINMENT!!Population 25 years and over"
 - Total Population 25y and over: Highschool or higher

- "S1501_C01_014E","Estimate!!Total!!AGE BY EDUCATIONAL ATTAINMENT!!Population 25 years and over!!High school graduate or higher"
 - Total Population 25y and over: Bachelor or higher
 - "S1501_C01_015E","Estimate!!Total!!AGE BY EDUCATIONAL ATTAINMENT!!Population 25 years and over!!Bachelor's degree or higher"
- ~~Mean household income~~
 - Dataset
 - 2023: 5 Year Estimates
 - Income in the Past 12 months (in 2023 Inflation Adjusted Dollars)
 - https://data.census.gov/table/ACSST5Y2023.S1901?q=county_income&y=2023
 - Variables
 - Total Households
 - "S1901_C01_001E","Estimate!!Households!!Total"
 - Mean Household Income
 - "S1901_C01_013E","Estimate!!Households!!Mean income (dollars)"
- Limited English Speaking Households Percentage
 - Dataset
 - 2023: 5 Year Estimates
 - Limited English Speaking Households
 - https://data.census.gov/table/ACSST5Y2023.S1602?q=county_speaking&y=2023
 - The Census Bureau defines a "Limited English Speaking Household" as one where no one aged 14 or older speaks English "very well."
 - Variables
 - Total Households
 - "S1602_C01_001E","Estimate!!Total!!All households"
 - Limited English Speaking Households: Spanish
 - "S1602_C01_002E","Estimate!!Total!!All households!!Households speaking --!!Spanish"
 - Limited English Speaking Housholds: European
 - "S1602_C01_003E","Estimate!!Total!!All households!!Households speaking --!!Other Indo-European languages"
 - Limited English Speaking Housholds: Asian
 - "S1602_C01_004E","Estimate!!Total!!All households!!Households speaking --!!Asian and Pacific Island languages"
 - Limited English Speaking Housholds: Other
 - "S1602_C01_005E","Estimate!!Total!!All households!!Households speaking --!!Other languages"
- General Fertility Rate (GFR)

- Dataset
 - 2023: 5 Year Estimates
 - Fertility
 - https://data.census.gov/table?q=county_fertility&y=2023
 - The general fertility rate (GFR) measures births per 1,000 women of childbearing age (usually 15-44 or 15-50).
 - $GFR = (\text{Number of women with births in the past year} / \text{Total number of women aged 15-50}) \times 1000$
- Variables
 - Women ages 15 - 50
 - "S1301_C01_001E","Estimate!!Total!!Women 15 to 50 years"
 - Women aged 15-50 with births
 - "S1301_C02_001E","Estimate!!Number!!Women with births in the past 12 months!!Women 15 to 50 years"
- Industry Composition: Employed over 16 Occupation
 - Dataset
 - 2023: 5 Year Estimates
 - Industry by Occupation for the Civilian Employed Population 16 Years and Over
 - https://data.census.gov/table/ACSST5Y2023.S2405?q=county_occupation&y=2023
 - Variables
 - 16+ Employed Total
 - "S2405_C01_001E","Estimate!!Total!!Civilian employed population 16 years and over"
 - 16+ Employed: Agriculture, forestry, fishing and hunting, and mining
 - "S2405_C01_002E","Estimate!!Total!!Civilian employed population 16 years and over!!Agriculture, forestry, fishing and hunting, and mining"
 - 16+ Employed: Construction
 - "S2405_C01_003E","Estimate!!Total!!Civilian employed population 16 years and over!!Construction"
 - 16+ Employed: Manufacturing
 - "S2405_C01_004E","Estimate!!Total!!Civilian employed population 16 years and over!!Manufacturing"
 - 16+ Employed: Wholesale Trade
 - "S2405_C01_005E","Estimate!!Total!!Civilian employed population 16 years and over!!Wholesale trade"
 - 16+ Employed: Retail Trade
 - "S2405_C01_006E","Estimate!!Total!!Civilian employed population 16 years and over!!Retail trade"

- 16+ Employed: Transportation and warehousing, and utilities
 - "S2405_C01_007E","Estimate!!Total!!Civilian employed population 16 years and over!!Transportation and warehousing, and utilities"
- 16+ Employed: Information
 - "S2405_C01_008E","Estimate!!Total!!Civilian employed population 16 years and over!!Information"
- 16+ Employed: Finance and insurance, and real estate and rental and leasing
 - "S2405_C01_009E","Estimate!!Total!!Civilian employed population 16 years and over!!Finance and insurance, and real estate and rental and leasing"
- 16+ Employed: Professional, scientific, and management, and administrative and waste management services
 - "S2405_C01_010E","Estimate!!Total!!Civilian employed population 16 years and over!!Professional, scientific, and management, and administrative and waste management services"
- "16+ Employed: Educational services, and health care and social assistance
 - S2405_C01_011E","Estimate!!Total!!Civilian employed population 16 years and over!!Educational services, and health care and social assistance"
- "16+ Employed: Arts, entertainment, and recreation, and accommodation and food services
 - "S2405_C01_012E","Estimate!!Total!!Civilian employed population 16 years and over!!Arts, entertainment, and recreation, and accommodation and food services"
- 16+ Employed: Other services, except public administration
 - "S2405_C01_013E","Estimate!!Total!!Civilian employed population 16 years and over!!Other services, except public administration"
- 16+ Employed: Public administration
 - "S2405_C01_014E","Estimate!!Total!!Civilian employed population 16 years and over!!Public administration"