**U.S. election 3 primary datasets**

1. governors\_country\_candidate\_2020\_kaggle.csv

|  |  |  |
| --- | --- | --- |
|  | columns | description |
| 1 | State | The name of state |
| 2 | County | The name of county |
| 3 | Candidate | The name of Candidate |
| 4 | Party | The party of the Candidate |
| 5 | Votes | The number of votes |
| 6 | Won | The results of the election for the candidate |

2. president\_country\_candidate\_2020\_kaggle.csv

|  |  |  |
| --- | --- | --- |
|  | columns | description |
| 1 | State (51) | The name of state |
| 2 | County | The name of county |
| 3 | Candidate | The president Candidate |
| 4 | Party | The party of the Candidate |
| 5 | Total\_votes | The number of votes |
| 6 | Won | The results of the election for the candidate |

3. senate\_county\_candidate\_2020\_kaggle.csv

|  |  |  |
| --- | --- | --- |
|  | columns | description |
| 1 | State | The name of state |
| 2 | County | The name of county |
| 3 | Candidate | The name of Candidate |
| 4 | Party | The party of the Candidate |
| 5 | Votes | The number of votes |
| 6 | Won | The results of the election for the candidate |

Final dataframe with description:

|  |  |  |
| --- | --- | --- |
| # | Column | Description |
| 1 | state | The state name |
| 2 | county | The county name |
| 3 | votes\_tru | The number of votes for Trump |
| 4 | Votes\_bid | The number of votes for Bider |
| 5 | Votes\_oth | The number of votes for other candidates |
| 6 | Won\_tru | If trump won, true or false |
| 7 | Won\_bid | If biden won. True or false |
| 8 | TOT\_POP | Total population |
| 9 | TOT\_MALE | Total male population |
| 10 | TOT\_FEMALE | Total female population |
| 11 | WA\_MALE | White alone male population |
| 12 | WA\_FEMALE | White alone female population |
| 13 | BAC\_MALE | Black or African American alone or in combination male population |
| 14 | BAC\_FEMALE | Black or African American alone or in combination female population |
| 15 | IAC\_MALE | American Indian and Alaska Native alone or in combination male population |
| 16 | IAC\_FEMALE | American Indian and Alaska Native alone or in combination male population |
| 17 | AAC\_MALE | Asian alone or in combination male population |
| 18 | AAC\_FEMALE | Asian alone or in combination female population |
| 19 | H\_MALE | Hispanic male population |
| 20 | H\_FEMALE | Hispanic female population |
| 21 | child | 0-17 |
| 22 | yg\_adlt | 18-39 |
| 23 | mid\_age | 40-60 |
| 24 | senior | 60+ |
| 25 | Child\_m | 0-17 male |
| 26 | Yg\_adlt\_m | 18-39 male |
| 27 | Mid\_age\_m | 40-60 male |
| 28 | Senior\_m | 60+ male |
| 29 | Child\_f | 0-17 female |
| 30 | Yg\_adlt\_f | 18-39 female |
| 31 | Mid\_age\_f | 40-60 female |
| 32 | Senior\_f | 60+ female |
| 33 | State\_code | State Abbreviation |
| 34 | Rural\_urban\_code | Rural-urban Continuum Code,2013 |
| 35 | Urban\_infl\_code | Urban influence Code, 2013 |
| 36 | Econ\_type\_code | County economic types, 2015 edition |
|  |  |  |
| 38 | Pop\_chg | Change of the population |
| 39 | R\_birth | Birth rate in period 7/1/2018 to 6/30/2019 |
| 40 | R\_death | Death Rate in period 7/1/2018 to 6/30/2019 |
| 41 | R\_nat\_increase | Natural Increase rate in period 7/1/2018 to 6/30/2019 |
| 42 | R\_intl\_mig | Net International migration rate in period 7/1/2018 to 6/30/2019 |
| 43 | R\_domestic\_mig | Net domestic migration rate in period 7/1/2018 to 6/30/2019 |
| 44 | R\_net\_mig | Net migration rate in period 7/1/2018 to 6/30/2019 |
|  |  |  |

**State-County Coordinates for maps:**

uscities.csv

|  |  |  |
| --- | --- | --- |
|  | columns | description |
| 1 | City | The name of city |
| 2 | City\_ascii | ASCII code of city |
| 3 | State\_id | State abbreviation |
| 4 | State\_name | The name of the state |
| 5 | County\_name | The name of the county |
| 6 | Lat | Latitude |
| 7 | Lng | Longitude |

**Cartographic Boundary Files (.shp files) can be found at:**

<https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html>

States : cb\_2018\_us\_county\_20m.shp

Counties : cb\_2018\_us\_state\_20m.shp

2013 Urban Influence Codes

<https://www.ers.usda.gov/data-products/urban-influence-codes/documentation.aspx#.UYKQ2kpZRvY>

Metropolitan Counties\*

Code

1 In large metro area of 1+ million residents

2 In small metro area of less than 1 million residents

Nonmetropolitan Counties

3 Micropolitan area adjacent to large metro area

4 Noncore adjacent to large metro area

5 Micropolitan area adjacent to small metro area

6 Noncore adjacent to small metro area and contains a town of at least 2,500 residents

7 Noncore adjacent to small metro area and does not contain a town of at least 2,500 residents

8 Micropolitan area not adjacent to a metro area

9 Noncore adjacent to micro area and contains a town of at least 2,500 residents

10 Noncore adjacent to micro area and does not contain a town of at least 2,500 residents

11 Noncore not adjacent to metro or micro area and contains a town of at least 2,500 residents

12 Noncore not adjacent to metro or micro area and does not contain a town of at least 2,500 residents

County Typology Data:

See also

https://www.ers.usda.gov/data-products/county-typology-codes/documentation/

Metro-nonmetro status 2013 Classification of counties by metro or nonmetro definition, where 1=metro county (Urban); 0=nonmetro county (Rural); metro areas include all counties containing one or more urbanized areas: high-density urban areas containing 50,000 people or more; metro areas also include outlying counties that are economically tied to the central counties, as measured by the share of workers commuting on a daily basis to the central counties. Nonmetro counties are outside the boundaries of metro areas and have no cities with 50,000 residents or more.

Econ Type Code Non-overlapping economic-dependence county indicator. 0=Nonspecialized 1=Farm-dependent 2=Mining-dependent 3=Manufacturing-dependent 4=Federal/State government-dependent 5=Recreation

Overlapping Economic Indicators:

Farming Farm-dependent county indicator. 0=no 1=yes. Faming accounted for at 25% or more of the county's earnings or 16% or more of the employment averaged over 2010-2012.

Mining Mining-dependent county indicator. 0=no 1=yes. Mining accounted for 13% or more of the county's earnings or 8% of the employment averaged over 2010-12.

Manufacturing Manufacturing-dependent county indicator. 0=no 1=yes. Manufacturing accounted for 23% or more of the county's earnings or 16% of the employment averaged over 2010-12.

Government Federal/State government-dependent county indicator. 0=no 1=yes. Federal and State government accounted for 14% or more of the county's earnings or 9% or more of the employment averaged over 2010-2012.

Recreation Recreation county indicator 0=no 1=yes. See documentation page: http://www.ers.usda.gov/data-products/county-typology-codes.aspx

Nonspecialized Nonspecialized indicator 0=no 1=yes. The county was not a farming, mining, manufacturing, government-dependent, or recreation county.

Low\_Education Low education county indicator. 0=no 1=yes. At least 20% or more of the residents age 25 to 64 did not have a high school diploma or equivalent between 2008-12.

Low\_Employment Low-employment county indicator. 0=no 1=yes. Less than 65% of residents age 25-64 were employed in 2008-12

Pop\_Loss Population loss county indicator. 0=no 1=yes. Number of resident declined between the 1990 and 2000 censuses and also between the 2000 and 2010 censuses.

Retirement\_Dest Retirement destination county indicator. 0=no 1=yes. Number of resident 60 and older grew by 15 percent or more between 2000 and 2010.

Persistent\_Poverty Classification of counties by level of poverty over three decades, where 1=persistent poverty county; 0=all other counties; a county was classified as persistent poverty if 20 percent or more of its residents were poor as measured by the 1980, 1990, and 2000 decennial censuses and the American Community Survey 5-year estimates for 2007-11.

Persistent\_Related\_Child\_Poverty Classification of counties by level of poverty over three decades, where 1=persistent related child poverty county; 0=all other counties; a county was classified as persistent related child poverty if 20 percent or more of related children under 18 years old were poor as measured by the 1980, 1990, and 2000 decennial censuses and the American Community Survey 5-year estimates for 2007-11. See the Census Bureau website.