DCI Components v7.csv

Data Dictionary 08/23/2018

Dimensions: 42 columns / variables, 26125 rows / observations

1. **Zip\_code**: U.S. Postal Service five-digit zip codes in the 50 states that have valid numerical entries in all seven DCI raw components. Each zip code is preceded by the letter “Z” to ensure that the data is treated as character, not numeric, and to ensure that leading zeroes are not removed when transferring data between software platforms. Values range from Z01001 to Z99929.
2. **City**: The name of the town or city containing most or all of the area of the zip code. The first letter of each word in the city name is capitalized and the other letters are lower case.
3. **State**: The two letter USPS abbreviation for the state, with both letters capitalized.
4. **Region**: Region of United States: Midwest, Northeast, South, West
5. **SubRegion**: East North Central, East South Central, Mid-Atlantic, Mountain, New England, Pacific, South Atlantic, West North Central, West South Central
6. **Population**: The estimated number of people living in that zip code from the 2011-2015 American Community Survey. Values range from 15 to 114982.
7. **No\_Diploma**: The estimated percentage of people, aged 25 and over, living in this zip code, who do not have a high school diploma or equivalent (i.e. GED) from the 2011-2015 American Community Survey. Values range from 0.0 to 82.1 and have one decimal place. A lower number is preferable.
8. **Vacancy**: The estimated percentage of homes in this zip code that are permanently vacant, which does not include vacation homes, from the 2011-2015 American Community Survey. Values range from 0.1 to 71.3 and have one decimal place. A lower number is preferable.
9. **No\_Job**: The estimated percentage of people, aged 25 and over, living in this zip code, who do not have a job from the 2011-2015 American Community Survey. Values range from 0.0 to 100.0 and have one decimal place. A lower number is preferable.
10. **Poverty**: The estimated percentage of people living in this zip code, who are living in a home in which the household income is below the poverty line calculated for the number of people in that household from the 2011-2015 American Community Survey. Values range from 0.0 to 92.0 and have one decimal place. A lower number is preferable.
11. **Med\_Inc**: The ratio of the median income estimated for households in this zip code to the median income estimated for households in this state from the 2011-2015 American Community Survey. Values are expressed as a percentage and range from 14.4 to 476.6 and have one decimal place. A higher number is preferable.
12. **Chg\_Job**: The ratio of the number of employees working in this zip code in 2015 to the number in 2011 from the 2015 Business Patterns survey and the 2011 Business Process survey. Values are expressed as a percentage and range from –97.5 to 2180.8 and have one decimal place. A higher number is preferable.
13. **Chg\_Biz**: The ratio of the number of operating businesses in this zip code in 2015 to the number in 2011 from the 2015 Business Patterns survey and the 2011 Business Process survey. Values are expressed as a percentage and range from –78.2 to 275.0 and have one decimal place. A higher number is preferable.
14. **No\_Diploma\_Rank**: A vector of 1 to 26125 assigned after sorting the **No\_Diploma** column from lowest to highest.
15. **Vacancy\_Rank**: A vector of 1 to 26125 assigned after sorting the **Vacancy** column from lowest to highest.
16. **No\_Job\_Rank**: A vector of 1 to 26125 assigned after sorting the **No\_Job** column from lowest to highest.
17. **Poverty\_Rank**: A vector of 1 to 26125 assigned after sorting the **Poverty** column from lowest to highest.
18. **Med\_Inc\_Rank**: A vector of 1 to 26125 assigned after sorting the **Med\_Inc** column from highest to lowest.
19. **Chg\_Job\_Rank**: A vector of 1 to 26125 assigned after sorting the **Chg\_Job** column from highest to lowest.
20. **Chg\_Biz\_Rank**: A vector of 1 to 26125 assigned after sorting the **Chg\_Biz** column from highest to lowest.
21. **Avg\_Rank**: The unweighted, unrounded mean of the seven rank columns. Values range from 717.7 to 24313.3.
22. **DCI\_Rank**: A vector of 1 to 26125 assigned after sorting the **Avg\_Rank** column from lowest to highest.
23. **DCI\_Percentile**: DCI\_Rank scaled to the range 0.0 to 100.0 by dividing DCI\_Rank by 261.25 and rounding to one decimal place.
24. **Rank\_Label**: The EIG label for each quintile (0.0 – 19.9, 20.0 – 39.9, etc.): Prosperous, Comfortable, Mid-tier, At risk, Distressed.
25. **Decile\_Label**: The EIG label applied to each decile (0.0 – 9.9, 10.0 – 19.9, etc.): Prosperous 1, Prosperous 2, Comfortable 3, … Distressed 10.
26. **White**: the percentage (0.0 – 100.0) of the zip code’s population that identifies as “white.”
27. **White\_Class**: the categorical variable that bins the continuous White variable: “0-25%”, “25-50%”, “50-75%”, “75-100%”
28. **Hispanic**: the percentage (0.0 – 100.0) of the zip code’s population that identifies as “hispanic.”
29. **Hispanic\_Class**: the categorical variable that bins the continuous Hispanic variable: “0-25%”, “25-50%”, “50-75%”, “75-100%”
30. **Black**: the percentage (0.0 – 100.0) of the zip code’s population that identifies as “black.”
31. **Black\_Class**: the categorical variable that bins the continuous Black variable: “0-25%”, “25-50%”, “50-75%”, “75-100%”
32. **Asian**: the percentage (0.0 – 100.0) of the zip code’s population that identifies as “Asian.”
33. **Asian\_Class**: the categorical variable that bins the continuous Asian variable: “0-25%”, “25-50%”, “50-75%”, “75-100%”
34. **Mixed\_Other**: the percentage (0.0 – 100.0) of the zip code’s population that identifies as more than one of white, Hispanic, black, and Asian or identifies as other.
35. **Mixed\_Other\_Class**: the categorical variable that bins the continuous Mixed\_Other variable: “0-25%”, “25-50%”, “50-75%”, “75-100%”
36. **Pop\_Density\_Land**: not documented
37. **Pop\_Density\_Land\_Class**: the categorical variable that bins the continuous Pop\_Density\_Land variable, with continuous value being greater than or equal to the lower bound and less than the upper bound: “0-25”, “25-50”, “50-75”, “75-100”, “100-250”, “250-500”, “500-750”, “750-1000”, “1000-2500”, “2500-5000”, “5000-7500”, “7500-10000”
38. **Pop\_Density\_Water**: not documented
39. **Area\_Land**: land area in square miles
40. **Area\_Water**: water area in square miles
41. **Latitude** in decimal degrees with negative numbers being below the equator
42. **Longitude** in decimal degrees with negative numbers being west of Greenwich, England