Social Determinants of Health Variables

Filename: pubuse.sdoh.census.tract.level.data.csv

File type: Comma separated value (CSV)

Number of records: 73,056 + 1 (header in row 1)

Number of variables: 47

Geography: Census tract level data in 2010 tract geographies

Notes:

- The file contains one record per census tract for every census tract in the U.S.
- Missingness (reasons for missing values):
  - Census tract has zero population
  - Data were suppressed in census tracts with a population of less than 200.
  - Segregation indices were suppressed if the population in a census tract was less than 20 for any of the races used in the calculation.
  - Variable(s) used in creating the index could not be calculated because the universe for one or more of variables have a population of zero.
  - Segregation indices are only created for census tracts located inside of Metropolitan Statistical Areas (MSA)
- SAS code is at the end of this document for reading the file.

Social Determinants of Health Variables

Listing of Variables

1. GeoID: Geographic ID (State/County/Tract)
2. State: State (FIPS)
3. County: County (FIPS)
4. Tract: Census Tract
5. Race_NHisp_White: Race: % White Alone, Non-Hispanic
6. Race_NHisp_Black: Race: % Black or African American Alone, Non-Hispanic
7. Race_NHisp_AIAN: Race: % American Indian and Alaska Native Alone, Non-Hispanic
8. Race_NHisp_Asian: Race: % Asian Alone, Non-Hispanic
9. Race_NHisp_NatHaw: Race: % Native Hawaiian and Other Pacific Islander Alone, Non-Hispanic
10. Race_NHisp_Other: Race: % Some Other Race Alone, Non-Hispanic
11. Race_NHisp_Multi: Race: % Two or More Races, Non-Hispanic
12. Race_Hisp_White: Race: % White Alone, Hispanic
13. Race_Hisp_Black: Race: % Black or African American Alone, Hispanic
14. Race_Hisp_AIAN: Race: % American Indian and Alaska Native Alone, Hispanic
15. Race_Hisp_Asian: Race: % Asian Alone, Hispanic
16. Race_Hisp_NatHaw: Race: % Native Hawaiian and Other Pacific Islander Alone, Hispanic
17. Race_Hisp_Other: Race: % Some Other Race Alone, Hispanic
18. Race_Hisp_Multi: Race: % Two or More Races, Hispanic
19. Educ_Less9th: Education: % less that 9th grade
20. Educ_9th_12th: Education: % 9th-12th grade (no diploma)
21. Educ_HSGrad: Education: % high school graduate
22. Educ_SomeColl: Education: % some college, no degree
23. Educ_AssocDeg: Education: % associate's degree
24. Educ_BachDeg: Education: % bachelor's degree
25. Educ_MastProfDeg: Education: % master's or professional school degree
26. Educ_DoctDeg: Education: % doctorate degree
27. Person_Below_Pov: Poverty: % of person below poverty level
28. RUCA4A: RUCA 4-level categorization 'A'
29. Flag_Tract_Pop_Zero: Flag: Census Tract Pop Zero
30. LQ_White_Alone: Location Quotient (LQ): White Alone
31. LQ_Black_Alone: Location Quotient (LQ): Black or African American Alone
32. LQ_API_Alone: Location Quotient (LQ): Asian or Pacific Islander Alone
33. LQ_Hispanic: Location Quotient (LQ): Hispanic
34. LQ_NH_White_Alone: Location Quotient (LQ): White Alone, Non-Hispanic
35. ICE_Black_Alone_White_Alone: Index of Concentration at Extremes (ICE): Black or African American Alone/White Alone
36. ICE_API_Alone_White_Alone: Index of Concentration at Extremes (ICE): Asian or Pacific Islander Alone/White Alone
37. ICE_Hispanic_NH_White_Alone: Index of Concentration at Extremes (ICE): Hispanic/White Alone, Non-Hispanic
38. Lex_Is_Black_Alone_White_Alone: Local Exposure and Isolation (Lex/IS): Black or African American Alone/White Alone
39. Lex_Is_API_Alone_White_Alone: Local Exposure and Isolation (Lex/IS): Asian or Pacific Islander Alone/White Alone
40. Lex_Is_Hispanic_NH_White_Alone: Local Exposure and Isolation (Lex/IS): Hispanic/White Alone, Non-Hispanic
41. LIS_White_Alone: Local Isolation Score: White Alone
42. LIS_Black_Alone: Local Isolation Score: Black or African American Alone
43. LIS_API_Alone: Local Isolation Score: Asian or Pacific Islander Alone
44. LIS_Hispanic: Local Isolation Score: Hispanic
45. LIS_NH_White_Alone: Local Isolation Score: White Alone, Non-Hispanic
46. Yost_Overall_Quintile: Yost Quintile (Total US)
47. Yost_State_Quintile: Yost Quintile (by state)

Social Determinants of Health Variables

Race/Ethnicity:

Table: B03002 Hispanic or Latino Origin by Race **Universe: Total population** **Data source: 2008-12 ACS5-Year Estimates**	
B03002e01	Total:
B03002e02	Not Hispanic or Latino:
B03002e03	White alone
B03002e04	Black or African American alone
B03002e05	American Indian and Alaska Native alone
B03002e06	Asian alone
B03002e07	Native Hawaiian and Other Pacific Islander alone
B03002e08	Some other race alone
B03002e09	Two or more races:
B03002e10	Two races including Some other race
B03002e11	Two races excluding Some other race, and three or more races
B03002e12	Hispanic or Latino
B03002e13	White alone
B03002e14	Black or African American alone
B03002e15	American Indian and Alaska Native alone
B03002e16	Asian alone
B03002e17	Native Hawaiian and Other Pacific Islander alone
B03002e18	Some other race alone
B03002e19	Two or more races:
B03002e20	Two races including Some other race
B03002e21	Two races excluding Some other race, and three or more races

```
If B03002e01^=0 Then Do;
  Race_NHisp_White=(B03002e03/B03002e01)*100;      /* % White, non-Hispanic  */
  Race_NHisp_Black=(B03002e04/B03002e01)*100;      /* % Black or African American, non-Hispanic  */
  Race_NHisp_AIAN=(B03002e05/B03002e01)*100;       /* % Amer. Indian or AK Nat, non-Hispanic  */
  Race_NHisp_Asian=(B03002e06/B03002e01)*100;      /* % Asian, non-Hispanic  */
  Race_NHisp_NatHaw=(B03002e07/B03002e01)*100;     /* % Native HI or other Pac Isl., non-Hispanic  */
  Race_NHisp_Other=(B03002e08/B03002e01)*100;      /* % Other, non-Hispanic  */
  Race_NHisp_Multi=(B03002e09/B03002e01)*100;      /* % Two or more races, non-Hispanic  */

  Race_Hisp_White=(B03002e13/B03002e01)*100;       /* % White, Hispanic  */
  Race_Hisp_Black=(B03002e14/B03002e01)*100;       /* % Black or African American, Hispanic  */
  Race_Hisp_AIAN=(B03002e15/B03002e01)*100;        /* % American Indian or Alaska Nat, Hispanic  */
  Race_Hisp_Asian=(B03002e16/B03002e01)*100;       /* % Asian, Hispanic  */
  Race_Hisp_NatHaw=(B03002e17/B03002e01)*100;      /* % Native HI or other Pacific Isl., Hispanic  */
  Race_Hisp_Other=(B03002e18/B03002e01)*100;       /* % Other, Hispanic  */
  Race_Hisp_Multi=(B03002e19/B03002e01)*100;       /* % Two or more races, Hispanic  */
  End;
```

Social Determinants of Health Variables

Education:

Table: B15002 Sex by Educational Attainment for the Population 25 Years and Over Universe: Population 25 Years and Over Data source: 2008-12 ACS5-Year Estimates	
B15002e01	Total:
B15002e02	Male:
B15002e03	No schooling completed
B15002e04	Nursery to 4th grade
B15002e05	5th and 6th grade
B15002e06	7th and 8th grade
B15002e07	9th grade
B15002e08	10th grade
B15002e09	11th grade
B15002e10	12th grade, no diploma
B15002e11	High school graduate (includes equivalency)
B15002e12	Some college, less than 1 year
B15002e13	Some college, 1 or more years, no degree
B15002e14	Associate's degree
B15002e15	Bachelor's degree
B15002e16	Master's degree
B15002e17	Professional degree
B15002e18	Doctorate degree
B15002e19	Female
B15002e20	No schooling completed
B15002e21	Nursery to 4th grade
B15002e22	5th and 6th grade
B15002e23	7th and 8th grade
B15002e24	9th grade
B15002e25	10th grade
B15002e26	11th grade
B15002e27	12th grade, no diploma
B15002e28	High school graduate (includes equivalency)
B15002e29	Some college, less than 1 year
B15002e30	Some college, 1 or more years, no degree
B15002e31	Associate's degree
B15002e32	Bachelor's degree
B15002e33	Master's degree
B15002e34	Professional degree
B15002e35	Doctorate degree

If B15002e01^=0 Then Do;
  Educ_Less9th= ((Sum(Of B15002e03-B15002e06)+Sum(Of B15002e20-B15002e23))/B15002e01)*100;
  Educ_9th_12th=((Sum(Of B15002e07-B15002e10)+Sum(Of B15002e24-B15002e27))/B15002e01)*100;
  Educ_HSGrad=((B15002e11+B15002e28)/B15002e01)*100;
  Educ_SomeColl=((B15002e12+B15002e13+B15002e29+B15002e30)/B15002e01)*100;
  Educ_AssocDeg=((B15002e14+B15002e31)/B15002e01)*100;
  Educ_BachDeg=((B15002e15+B15002e32)/B15002e01)*100;
  Educ_MastProfDeg=((B15002e16+B15002e17+B15002e33+B15002e34)/B15002e01)*100;
  Educ_DoctDeg=((B15002e18+B15002e35)/B15002e01)*100;
  End;

Social Determinants of Health Variables

Poverty:

Table: C17002 Ratio of Income to Poverty Level in the Past 12 Months **Universe: Population for Whom Poverty Status is Determined** **Data source: 2008-12 ACS5-Year Estimates**	
C17002e01	Total:
C17002e02	Under .5
C17002e03	.50 to .99
C17002e04	1.00 to 1.24
C17002e05	1.25 to 1.49
C17002e06	1.50 to 1.84
C17002e07	1.85 to 1.99
C17002e08	2.00 and over

If C17002e01^=0 Then Person_Below_Pov=((C17002e02+C17002e03)/C17002e01)*100;
Label Person_Below_Pov = "Poverty: % of person below poverty level";

Urban/Rural:

RUCA census tract codes for 2010 from: https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx

If Secondary_RUCA_Code In (1,1.1,2.0,2.1,3,4.1,5.1,7.1,8.1,10.1) Then RUCA4A=1;
Else If Secondary_RUCA_Code In (4,4.2,5,5.2,6,6.1) Then RUCA4A=2;
Else If Secondary_RUCA_Code In (7,7.2,7.3,7.4,8,8.2,8.3,8.4,9,9.1,9.2) Then RUCA4A=3;
Else If Secondary_RUCA_Code In (10,10.2,10.3,10.4,10.5,10.6) Then RUCA4A=4;
Else If Secondary_RUCA_Code=99 Then RUCA4A=.;

Label RUCA4A = "RUCA 4-level categorization 'A'";

1	Urban focused
2	Large rural city/town
3	Small rural town focused
4	Isolated small rural town focused
.	Not coded: Census tract has zero population and no rural-urban identifier information

Social Determinants of Health Variables

Indicator variable for population=0:

	Table: B01001 Sex by Age Universe: Total Population Data source: 2008-12 ACS5-Year Estimates	
B01001e01	Total:	
B01001e02	Male:	
B01001e03	Under 5 years	
B01001e04	5 to 9 years	
	. . . .	
B01001e24	80 to 84 years	
B01001e25	85 years and over	
B01001e26	Female:	
B01001e27	Under 5 years	
B01001e28	5 to 9 years	
	. . . .	
B01001e48	80 to 84 years	
B01001e49	85 years and over	

If B01001e01=0 Then Flag_Tract_Pop_Zero=1;
Else If B01001e01>0 Then Flag_Tract_Pop_Zero=0;
Label Flag_Tract_Pop_Zero = "Flag: Census Tract Pop Zero";

Social Determinants of Health Variables

Yost SES index used by SEER:

- The index is constructed using a factor analysis of seven variables of SES. This index is described in "Using a composite index of socioeconomic status to investigate health disparities while protecting the confidentiality of cancer registry data" by Mandi Yu, Zaria Tatalovich, Kathleen A. Cronin and James T. Gibson**.** https://pubmed.ncbi.nlm.nih.gov/24178398/
- The file contains two population weighted quintile variables. One generated for the total U.S. (across states) and one by state. Census tract population estimates obtained from Woods & Poole Economics, Inc. are used in weighting due to the large sampling errors in the American Community Survey data.

Segregation Indices:

- Race combinations – Indices incorporate two racial/ethnic groups. The indexes are generated for the following race combinations: Asian or Pacific Islander/white, black/white, Hispanic-any race/non-Hispanic white.
- The four segregation indices to be measured include the following.
  - Location Quotient (LQ)
  - Local Exposure and Isolation (LEx/Is)
  - Local Isolation (LIs)
  - Index of Concentration of Extremes (ICE)
- This article describes all segregation indices and provides equations: Amin Bemanian and Kirsten M.M. Beyer (CEBP, April 2017) *Measures Matter: The Local Exposure/Isolation (Lex/Is) Metrics and Relationships between Local Level Segregation and Breast Cancer Survival;* https://www.ncbi.nlm.nih.gov/pubmed/28325737
- Segregation indices are only created for census tracts located inside of Metropolitan Statistical Areas (MSA), thus there will be many patient records that do not have segregation measures available. In addition, data were suppressed if the population in a census tract was less than 20 for any of the races used in the calculation. Residential racial segregation is created and sustained by labor and housing markets that operate at the MSA level, thus segregation isn't conceptually relevant outside of MSAs. Investigators interested in examining racial/ethnic density in non-metropolitan areas can use the percent of a population of a particular race/ethnicity. There are decades of published work in the social sciences that provide the foundation to better understand the conceptual meaning, mechanistic pathways, and empirical measurement of segregation.

Social Determinants of Health Variables

SAS Code to Read Data File

```
Data SDoH;
   Infile "pubuse.sdoh.census.tract.level.data.csv" dlm="," FirstObs=2;
   Input GeoID :$11.
         State :$2.
         County :$3.
         Tract :$6.
         Race_NHisp_White
         Race_NHisp_Black
         Race_NHisp_AIAN
         Race_NHisp_Asian
         Race_NHisp_NatHaw
         Race_NHisp_Other
         Race_NHisp_Multi
         Race_Hisp_White
         Race_Hisp_Black
         Race_Hisp_AIAN
         Race_Hisp_Asian
         Race_Hisp_NatHaw
         Race_Hisp_Other
         Race_Hisp_Multi
         Educ_Less9th
         Educ_9th_12th
         Educ_HSGrad
         Educ_SomeColl
         Educ_AssocDeg
         Educ_BachDeg
         Educ_MastProfDeg
         Educ_DoctDeg
         Person_Below_Pov
         RUCA4A
         Flag_Tract_Pop_Zero
         LQ_White_Alone
         LQ_Black_Alone
         LQ_API_Alone
         LQ_Hispanic
         LQ_NH_White_Alone
         ICE_Black_Alone_White_Alone
         ICE_API_Alone_White_Alone
         ICE_Hispanic_NH_White_Alone
         Lex_Is_Black_Alone_White_Alone
         Lex_Is_API_Alone_White_Alone
         Lex_Is_Hispanic_NH_White_Alone
         LIS_White_Alone
         LIS_Black_Alone
         LIS_API_Alone
         LIS_Hispanic
         LIS_NH_White_Alone
         Yost_Overall_Quintile
         Yost_State_Quintile;
```