This dataset and codebook was produced by the EMERGENS research team based out of the University of California San Diego

emergens-project.com

# 

EMERGENS County-Level Dataset Codebook

Variables

[FIPS 2](#_Toc109815305)

[Year 2](#_Toc109815306)

[State\_FIPS 2](#_Toc109815307)

[County 2](#_Toc109815308)

[Unemployment\_Rate 2](#_Toc109815309)

[Median\_Household\_Income 2](#_Toc109815310)

[Mean\_Household\_Income 3](#_Toc109815311)

[Proportion\_Poverty 3](#_Toc109815312)

[Proportion\_Uninsured 3](#_Toc109815313)

[Proportion\_Home\_No\_Vehicle 3](#_Toc109815314)

[Proportion\_Homeowners\_35Perc\_Income\_on\_Home 3](#_Toc109815315)

[Proportion\_Renters\_35Perc\_Income\_on\_Rent 3](#_Toc109815316)

[Proportion\_White 4](#_Toc109815317)

[Proportion\_Black 4](#_Toc109815318)

[Proportion\_American\_Indian\_Alaska\_Native 4](#_Toc109815319)

[Proportion\_Asian 4](#_Toc109815320)

[Proportion\_Native\_Hawaiian\_Pacific\_Islander 4](#_Toc109815321)

[Proportion\_Male 4](#_Toc109815322)

[Proportion\_High\_School\_Greater 4](#_Toc109815323)

[Proportion\_Bachelors\_Degree\_or\_Greater 5](#_Toc109815324)

[Employee\_Capacity 5](#_Toc109815325)

[Total\_Employees\_March\_Snapshot 5](#_Toc109815326)

[Total\_Annual\_Payroll 5](#_Toc109815327)

[Employee\_Capacity\_Change 5](#_Toc109815328)

[Total\_Employees\_March\_Snapshot\_Change 5](#_Toc109815329)

[Total\_Annual\_Payroll\_Change 6](#_Toc109815330)

[Urbanicity 6](#_Toc109815331)

[ORx\_per\_100 6](#_Toc109815332)

[Fentanyl\_Total 6](#_Toc109815333)

[Deaths 6](#_Toc109815334)

[Population 6](#_Toc109815335)

[overdose\_death\_rate 7](#_Toc109815336)

[overdose\_gravity\_add 7](#_Toc109815337)

## FIPS

Variable Type: ID (Numeric)

Source: US Census Bureau

Description: Each county in the United States has a 5-digit identification number. The first two digits represent the state, and the final three digits represent the county within the state. Every county within the same state will, thus, start with the same two digits.

## Year

Variable Type: Numeric

Source: n/a

Description: Each observation in the dataset is a county-year. This variable indicates which year the observation corresponds to.

## State\_FIPS

Variable Type: Numeric

Source: Derived from FIPS, US Census

Description: The FIPS code for a county is made up of 5 digits. The first 2 digits identify the state of the county. This variable is the state FIPS. This variable is useful for organizing/clustering observations by state.

## County

Variable Type: Text

Source: US Census

Description: Contains the actual name of the county, for example: “Geauga County, OH”

## Unemployment\_Rate

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: The percentage of people, aged 16 and over, in the labor force who are currently unemployed. Values are between 0 and 100. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Median\_Household\_Income

Variable Type: Numeric (in US $)

Source: American Community Survey

Description: Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Mean\_Household\_Income

Variable Type: Numeric (in US $)

Source: American Community Survey

Description: Estimate of county’s mean household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Proportion\_Poverty

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: Percentage of families and people in the county whose income in past 12 months is below the poverty level. Values range from 0 – 100. Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Proportion\_Uninsured

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: Percentage of civilian, non-institutionalized population that have no health insurance coverage. Values range from 0 – 100. Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Proportion\_Home\_No\_Vehicle

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: Percentage of occupied housing units in the county that have no vehicle available for use. Values range from 0 – 100. Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Proportion\_Homeowners\_35Perc\_Income\_on\_Home

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: The percentage of homeowners who spend 35% or more of their monthly income on their mortgage. Values range from 0 – 100. Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Proportion\_Renters\_35Perc\_Income\_on\_Rent

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: The percentage of housing renters who spend 35% or more of their monthly income on their rent. Values range from 0 – 100. Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Proportion\_White

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: Percentage of the county population that identifies as white. Values range from 0 – 100. Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Proportion\_Black

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: Percentage of county population that identifies as Black. Values range from 0 – 100. Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Proportion\_American\_Indian\_Alaska\_Native

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: Percentage of county population that identifies as American Indian and/or Alaskan Native – this does not include individuals who identify as Native Hawaiian. Values range from 0 – 100. Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Proportion\_Asian

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: Percentage of county population that identifies as Asian. Values range from 0 – 100. Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Proportion\_Native\_Hawaiian\_Pacific\_Islander

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: Percentage of county population that identifies as Native Hawaiian or Pacific Islander. Values range from 0 – 100. Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Proportion\_Male

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: Percentage of county population that identifies as male. Values range from 0 – 100. Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Proportion\_High\_School\_Greater

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: The percentage of adults, aged 25 and over, who have graduated high school or received equivalent certification. Values range from 0 – 100. Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Proportion\_Bachelors\_Degree\_or\_Greater

Variable Type: Numeric (proportion)

Source: American Community Survey

Description: The percentage of adults, aged 25 and over, who have graduated with a 4-year college degree or higher level of education. Values range from 0 – 100. Estimate of county’s median household income, in US$. This is a 5-year “period” estimate – given that every county across the US is not surveyed each year, using data collected over the prior 5-year window increases the stability of estimates.

## Employee\_Capacity

Variable Type: Numeric

Source: Census County Business Patterns

Description: County Business Patterns data provides a number of businesses of various size ranges. For example, it provides information on the number of businesses with 1-5 employees, 6-9 employees, 10-19 employees, all the way up to 5,000+. This variable was calculated by taking each of these size ranges, counting the number of businesses within each size category, and then summing up the mean value of each range. This provides an approximation for the number of employees businesses in the county are able to employ.

## Total\_Employees\_March\_Snapshot

Variable Type: Numeric

Source: County Business Patterns

Description: This is a cross-sectional snapshot of the number of people employed by businesses in the county in March of that year.

## Total\_Annual\_Payroll

Variable Type: Numeric

Source: County Business Patterns

Description: The sum of annual payrolls for businesses within the county.

## Employee\_Capacity\_Change

Variable Type: Numeric

Source: Calculated from Employee\_Capacity

Description: This measure was calculated by measuring the difference between Employee\_Capacity for the current year and the prior year.

## Total\_Employees\_March\_Snapshot\_Change

Variable Type: Numeric

Source: Calculated from Total\_Employees\_March

Description: This measure was calculated by measuring the difference between Total\_Employees\_March for the current year and the prior year.

## Total\_Annual\_Payroll\_Change

Variable Type: Numeric

Source: Calculated from Total\_Annual\_Payroll

Description: This measure was calculated by measuring the difference between Total\_Annual\_Payroll for the current year and the prior year.

## Urbanicity

Variable Type: Categorical

Source: NCHS Urban-Rural Classification Scheme

Description: Counties are classified into one of 6 urbanicity rankings: 1 = large central metro counties; 2 = large fringe metro counties; 3 = medium metro counties; 4 = small metro counties; 5 = micropolitan counties; 6 = non-core counties.

## ORx\_per\_100

Variable Type: Numeric (rate)

Source: IQVIA Xponent via CDC

Description: Rate of retail opioid ***prescriptions*** dispensed per 100 persons – these rates do not indicate the dosage of prescriptions, just the quantity of prescriptions. Data was collected by IQVIA and is published by the CDC [here](https://www.cdc.gov/drugoverdose/rxrate-maps/index.html).

## Fentanyl\_Total

Variable Type: Numeric (count)

Source: National Forensic Laboratory Information System

Description: State-level variable (i.e., all counties within same state-year assigned the same value) of the number of laboratory tests that identified fentanyl. When drugs are seized by law enforcement, they are then tested in state labs – this counts the number of samples tested that were found to contain fentanyl. This does not capture the amount of fentanyl (i.e., dosage), just the number of tests run that identified it. Higher totals means larger numbers of drug seizures containing fentanyl – a proxy for higher levels of fentanyl in the drug market.

## Deaths

Variable Type: Numeric (count)

Source: CDC Mortality Records

Description: Crude number of overdose deaths reported for that county. If using the unrestricted data, the CDC “suppresses” all overdose death counts less than 10 as a human subjects protection. For these counties, deaths were imputed by identifying the number of “county-less” deaths (i.e., by subtracting all known county-level death counts from the state total) and evenly distributing them across all suppressed counties by population size.

## Population

Variable Type: Numeric (count)

Source: CDC

Description: Estimated population size for the county in that year.

## overdose\_death\_rate

Variable Type: Numeric (rate)

Source: Calculated

Description: Calculated from the number of deaths and population – this is the overdose death rate per 100,000 people.

## overdose\_gravity\_add

Variable Type: Numeric

Source: Calculated

Description: We generated an overdose gravity variable to measure the burden of overdose in neighboring counties – we found this variable to be the most informative in our prior work. To do so, for a given year and county, we identified all neighboring counties within 200 miles of said county (using county centroid approach), identified both the overdose death rate (OD) and distance (d) from said county. Then for each neighboring county we divided OD by the distance squared (i.e., d2) and then took the sum of this value to get the overdose gravity value for said county.