

# **US Microbusiness Data**

## **Dictionary**

*Last updated Spring 2025*

Below is a list of the columns contained within the files that contain all the microbusiness count and density information. For ease of use, these files are split into the city center-, CBSA-, county-, and state-level files.

*active\**: A raw count of the number of active microbusinesses in a community, defined by GoDaddy as microbusiness with an active website. This is an integer derived by summing across the zip codes within a geographic unit.

*md*: The number of microbusinesses per 100 people in the geographic unit, during the month described by the row's data. Updated through March 2025.

*cbsa*: A unique numeric identifier (up to 5 digits) for each Core-Based Statistical Area (CBSA). These are the micropolitan and metropolitan areas defined by the U.S Department of Housing and Urban Development (HUD).

*cbsa\_name*: In a string, the legal description of the CBSA denoted by the row, as understood by HUD.

*cfips*: A unique numeric identifier (up to 5 digits) for each county. The first two digits represent the state Federal Information Processing System (FIPS) code, while the last 3 digits are assigned to each of the counties within each state, as defined by the Department of Housing and Urban Development (HUD).

*city*: In a string, the legal description of the municipality denoted by the row, as understood by the U.S. Postal Service (USPS).

*city\_name*: In a string, the legal description of the CBSA denoted by the row, as understood by HUD.

*city\_id*: A unique numeric identifier (up to 5 digits) for each city center/municipality, generated by GoDaddy for ease of data wrangling.

*county*: The legal description in a string of the county denoted by the row, as understood by HUD.

*county\_name*: The legal description in a string of the county denoted by the row, as understood by HUD.

*date*: The month of observation for the associated row of data, one per each geographic entity

*fips*: The first two digits representing the state in the Federal Information Processing System (FIPS) code, as defined by the Department of Housing and Urban Development (HUD).

*MAI\_composite*: The overall composite index score for the associated geography in the associated row-date.

*engagement*: The engagement sub-index score for the associated geography in the associated row-date. This measures how active websites are in the community.

*participation*: The participation sub-index score for the associated geography in the associated row-date. This measures the number of online microbusinesses and operators in the community.

*infrastructure*: The infrastructure sub-index score for the associated geography in the associated row-date. This measures how prepared a community is to use the internet, in terms of necessary physical and intellectual infrastructure.

*micro\_metro*: A flag written in string to denote whether a CBSA is a micropolitan (under 50k population but over 10K) or a metropolitan statistical area (over 50k population).

*pop\_18over\_YYYY*: Adult population 18 years of age and older for the given year YYYY. This is the figure used to calculate density for all geographic units.

*state\_abbrev*: In a string, the legal description of the state denoted by the row, as understood by the USPS.

*state\_name*: Full state name.

*total\_pop\_YYY*: Total population for a given geographic area in the given year YYYY. This is not the population figure used to calculate density anywhere.

*topind\_shoppers*: Character string denoting the industry which had the largest number of active merchants for the month denoted by the row's year and month combined.

*topind\_orders*: Character string denoting the industry which had the largest number of completed transactions/orders for the month denoted by the row's year and month combined.

*topind\_sales*: Character string denoting the industry which had the largest total sales in US dollars for the month denoted by the row's year and month combined.