

## **GoDaddy Platform (Microbusiness Density)**

### **Data Dictionary**

*Updated February 2023*

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.

*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). HUD crosswalk files are used for each update, making the geographic area defined the most up-to-date as possible.

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

*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.

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

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

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

*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). HUD crosswalk files are used for each update, making the geographic area defined the most up to date as possible.

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

*total\_pop\_2020*: Total population for a given geographic area in 2020. This is not the population figure used to calculate density anywhere.

*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.

*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 county or CBSA.

*md\**: A microbusiness density measure, derived by dividing the 18+ population by the raw number of active microbusinesses in a geographic area, and multiplying by 100. Population figures are on a two-year lag due to the pace of update provided by the U.S. Census Bureau, which provides the underlying population data annually. 2021 density figures are calculated using 2019 population figures, etc.