

GoDaddy Platform (Industry and Commerce)

Data Dictionary

Updated August 2022

Below is a list of the columns contained within the files that contain all the Industry and Commerce data. For ease of use, these files are split into a CBSA-level, a county-level, a state-level, and a city center-level file.

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.

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). 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.

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).

state: Abbreviated state name.

stateName: Full state name.

city_id: A unique numeric identifier for each city center, generated by GoDaddy.

city_name: In a string, the legal description of the city as provided by the U.S. Postal Service.

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

total_pop_20: The total population per the U.S. Census Bureau in 2020.

orders_rank_[month][year]: Monthly rank in terms of orders per 100 people for sites GoDaddy can see transactions in the respective geographic area

- *Methodology*: Monthly orders are summed across each zip code. Those zip codes are then aggregated up to the larger geography (ex: sum of all the zip codes in the city of Atlanta). Ranks are assigned based on population cuts in each file:

Cities: Group 1 are the top 100 cities by population (305,000 and higher), Group 2 are the cities with populations between 50,000 – 305,000, and Group 3 are cities with populations less than 50,000

Counties: Group 1 are the top 300 counties by population (225,000 and higher), Group 2 are the counties with populations between 30,000 – 225,000, and Group 3 are counties with populations less than 30,000

CBSAs: Group 1 are the top 100 CBSAs by population (555,000 and higher), Group 2 are the CBSAs with populations between 50,000 – 555,000, and Group 3 are CBSAs with populations less than 50,000

States: no groupings are made, states are all ranked comparatively to one another.

- *Dates:* Aug 2019 – Jun 2022

merchants_rank_[month][year]: Monthly rank in terms of active merchants (online stores) per 100 people for sites GoDaddy can see transactions in the respective geographic area

- *Methodology:* Monthly merchant counts are summed across each zip code. Those zip codes are then aggregated up to the larger geography (ex: sum of all the zip codes in the city of Atlanta). Ranks are assigned based on population cuts in each file:

Cities: Group 1 are the top 100 cities by population (305,000 and higher), Group 2 are the cities with populations between 50,000 – 305,000, and Group 3 are cities with populations less than 50,000

Counties: Group 1 are the top 300 counties by population (225,000 and higher), Group 2 are the counties with populations between 30,000 – 225,000, and Group 3 are counties with populations less than 30,000

CBSAs: Group 1 are the top 100 CBSAs by population (555,000 and higher), Group 2 are the CBSAs with populations between 50,000 – 555,000, and Group 3 are CBSAs with populations less than 50,000

States: no groupings are made, states are all ranked comparatively to one another.

- *Dates:* Aug 2019 – Jun 2022

gmv_rank_[month][year]: Monthly rank in terms of transaction dollars per 100 people for sites GoDaddy can see transactions in the respective geographic area

- *Methodology:* Monthly dollars transacted are summed across each zip code. Those zip codes are then aggregated up to the larger geography (ex: sum of all the zip codes in the city of Atlanta). Ranks are assigned based on population cuts in each file:

Cities: Group 1 are the top 100 cities by population (305,000 and higher), Group 2 are the cities with populations between 50,000 – 305,000, and Group 3 are cities with populations less than 50,000

Counties: Group 1 are the top 300 counties by population (225,000 and higher), Group 2 are the counties with populations between 30,000 – 225,000, and Group 3 are counties with populations less than 30,000

CBSAs: Group 1 are the top 100 CBSAs by population (555,000 and higher), Group 2 are the CBSAs with populations between 50,000 – 555,000, and Group 3 are CBSAs with populations less than 50,000

States: no groupings are made, states are all ranked comparatively to one another.

- *Dates:* Aug 2019 – Jun 2022

*top3industries_[month][year]:_**: The top three industries by transaction dollars for the given month in the respective geographic area.

*avg_traffic_[month][year]:_**: Average monthly traffic to the sites of GoDaddy's website + marketing customers in the respective geographic area

- *Methodology:* Monthly traffic is calculated for each individual website. Then, those traffic numbers are averaged across each zip code. Next, a weighted average of traffic between all the zip codes within the larger geography (ex: average of all the zip codes in the city of Atlanta) is calculated, using the number of websites in each zip code as a weight.
- *Dates:* Aug 2019 – Jun 2022
- *Notes:* Any geo (city, county, etc.) containing less than 5 total websites in a month will receive an NA traffic value that month. Any cities with less than 3 zip codes OR < 10,000 population (as of 2020) are excluded from the data.

avg_lifespan_mths: Average lifespan of existing websites, in months, for GoDaddy website + marketing customers, in the respective geographic area.

- *Methodology:* Lifespan in months (time between create date and current date) is calculated for each existing website in a zip code. Then an average lifespan is calculated across each zip code. The final average lifespan comes from a weighted average of each of the zip code lifespans in the larger geographic area, with the weight being the # of sites in each zip code.
- *Dates:* Websites exist as of Aug 2022
- *Notes:* Any geo (city, county, etc.) containing less than 5 total websites in a month will receive an NA average lifespan value. Any cities with less than 3 zip codes OR < 10,000 population (as of 2020) are excluded from the data.

Note: more information on the Industry and Commerce data can be found at <https://www.godaddy.com/ventureforward/explore-the-data/>