



Your Economy Time Series (YE Time Series) Database Description 1998 - 2022

Assembled exclusively by the Business Dynamics Research Consortium (BDRC), YE Time Series (Your Economy Time Series) is an establishment level time series database used by academic researchers, policy makers, economic development analysts, and for profit businesses seeking to evaluate companies at their unique locations across the U.S.

YE Time Series data is the resource behind the popular website youreconomy.org, and is available for aggregate modeling, research, and direct licensing.

Comprehensive Data on U.S. Businesses

YE Time Series tracks all establishments¹ (and their jobs and sales) at their unique location, including for-profit (both privately-owned and publicly-traded), non-profits, agriculture, and government establishments.

Dynamic “In-business” Establishments

YE Time Series focuses on establishments that are “in-business”, meaning they are or intend to be conducting commercial activities. By contrast, businesses that are created for the purpose of housing financial, real estate, and tax reporting entities, or are suspected of never actually conducting commercial activities, are not included in YE Time Series.

Jobs in YE Time Series vs. FTE

Jobs in the YE Time Series are defined as all persons currently working at an establishment including full-time, part-time, and temporary workers. This includes proprietorships and sole-proprietorships (self-employed), and no distinction is made between employee or non-employee designations. Reference the BDRC exclusive chart “Comparing Economic Data Sources”² for more information government data sources and their definitions.

¹ Establishment - business or organization at a unique location

² Link - [Comparing Economic Data Sources](#)

Current Data

Annual establishment data through the last calendar year and quarterly establishment data through the last calendar quarter.

YE Time Series Assembly

Each quarter, Data Axle³ delivers over 300 data variables per record to BDRC for all establishments in their database (approximately 91 million records) including verified, suspect, pre-verified, and closed:

- Verified - establishment verified "in-business" at location (>97% accuracy)
- Pre-verified - new establishment has not met all verified rules
- Suspect - establishment removed from verified status for further review
- Closed - establishment verified out of business at location

From 1998 thru 2019, the YE Time Series was assembled from Data Axle Business Data verified annual historical files. Starting in March of 2020 BDRC began receiving data on a quarterly basis (March, June, September, and December) which is now used to assemble the YE Time Series.

All records are assembled using the Data Axle unique ABI (American Business Identifier or Infogroup_ID) to link all historical and quarterly files together. Each year 168 time-series variables, including BDRC uniquely developed variables (eg. CIK, congressional districts, census tract FIPS), are assembled from the original Data Axle data variables and several other external data sources.

Only Verified data is used when assembling the YE Time Series databases, with the Pre-verified, Suspect, and Closed data used to examine and improve the Verified data. All ATM machines, video kiosks, and other automated kiosks are further removed from the Verified data.

Annual Assembly

After the establishment records are assembled and added to the YE Time Series, some (1-2%) have gaps between years of data (eg. active in 2018, missing data in

³ Link - [Data Axle](#) - BDRC business data provider

2019 and 2020, active again in 2021). These gaps are repaired by rolling values into the missing year as needed to complete the record. No averaging is used in this process. The latest year (last year) of information indicating an establishment's name, address, and location is also assembled and included. YE Time Series remains transparent. In certain circumstances, major data outliers (primarily in employment, industry, and parent relationship variables) are identified, verified, and repaired (imputed) algorithmically. This typically accounts for an average of 1.5% of all YE Time Series data. Any data that is in question will be examined, verified with large internal establishment databases and external sources, and corrected as necessary. In the rare cases where this outlier process is not able to verify data at an acceptable level, the individual establishment record is temporarily deactivated for further review.

Jobs - Approximately 39% (as of second quarter 2022) of the Data Axle Verified businesses have their location employment size verified as actual through telephone interview. When an employment number cannot be verified through the telephone interview process, a model is built to estimate the employment size. About 51% of businesses have their location employment size modeled by Data Axle. The model uses a multi-step approach, with over 7 million telephone verified employment figures as the cornerstone, to create the most accurate estimated employment information possible.

Sales - YE Time Series contains over 5,600 actual sales for corporate (HQ) establishments obtained from SEC filings from 2003-2021.

Establishment Sales - BDRC has developed a separate establishment sales model that replaces all sales values taken from Data Axle modeled sales by reviewing modeled establishment and actual firm (corporate) sales, 6-digit NAICS, and verified actual jobs data. To keep in step with GDP (Gross Domestic Product), each new establishment sales value is multiplied by an annual inflation value using Gross Output (sales and revenue) factors from the BEA (Bureau of Economic Analysis).

Quarterly Model Assembly

On average, Data Axle contacts every company in the US at least once per year in order to up-date the establishment records for start-ups, establishment deaths, and going concerns. At a quarterly level this means that approximately one-fourth of records have been verified current by Data Axle.

BDRC imputes values of the files not updated by Data Axle in each quarter (approximately three-fourths of records). Imputation estimates the status (expansion, contraction, or death) and job levels of the non-updated records by distilling information from the updated records and applying this back to the non-updated

records. The relevant variables used in the model include industry, employment size, and geographic location (based on FIPS). The model employs a combination of multiple imputation and linear regression techniques. When the records are updated by Data Axle, BDRC exchanges any old imputation information with this new 'actual' information so that each record is as current and correct as possible.

YE Time Series 2022 Standard Field Variable List

- ✓ One-time variable (last period in the database)
 - ➔ Time series variables from 1998-2021 (eg. FIPS2021)
 - ★ YE Time Series exclusive variable
-
- ✓ **ABI** - also known as Infogroup_id, this provides a unique identifier for each establishment in the Data Axle business database.
 - ✓ **Company** - name of establishment (or industry description)
 - ✓ **City** - last year city
 - ✓ **State** - last year state
 - ✓ **ZipCode** - last year Zip-code
 - ✓ **Gender** - gender of the contact or executive. Female, Male, Undetermined
 - ✓ **Latitude** - parcel level assigned via point geocoding. Half of a pair of coordinates (the other being longitude) provided in a decimal degree format, with a negative sign for longitude. Not available in Puerto Rico or Virgin Island
 - ✓ **Longitude** - see Latitude
 - ✓ **EIN** - unique number assigned to an establishment for easy IRS identification for tax reporting purposes. Possibly containing more than one number
 - ✓ **ForeignParent** - 1 indicates foreign affiliation
 - ✓ **Ethnicity** - A South Asian, B African American, C Central & Southwest Asian, E Eastern European, F Far Eastern, H Hispanic, I Middle Eastern, J Jewish, M Mediterranean, N Native American, P Pacific Islander, S Scandinavian, W Western European, X Other, Z Uncoded
 - ✓ **FortuneRanking** - a list of 500 (1-500) of the largest businesses in the United States compiled by Fortune magazine every year. Businesses are ranked by their annual revenues and includes both public and private companies
 - ✓ **Creditscore** - Developed From The Demographic Information In The Data Axle Database Including Employees And Years In Business As Well As Historical Performance Data And Other Information. Data Axle Developed Predictive Statistical Models Based On Multivariate Regression Analysis To Determine Creditworthiness Of Businesses. Each Credit Score Is Assigned A Number Based Upon The Alpha Score
 - ✓ **StockSymbol** - a series of letters assigned to a security for trading purposes

- ✓ **SICDescription** - the SIC code description for for the last year
- ✓ **NAICSDescription** - the description for the NAICS code for the last year
- ➔ **HQABI** - the parent number identifies the corporate parent of the business and also serves as the ABI number for the headquarters site of the parent. This field provides corporate ownership linkage information. This information is not collected or maintained for the types of organization for which ownership is ambiguous. Churches and schools, in particular, are not linked for this reason
- ➔ **FIPS** - the Federal Information Processing Standard Publication is a five-digit code which uniquely identified counties and county equivalents. The first two digits identify the state while the last 3 digits identify the county
- ➔ **CBSA** - a core-based statistical area (CBSA) is a U.S. geographic area defined by the Office of Management and Budget (OMB) that consists of one or more counties (or equivalents) anchored by an urban center of at least 10,000 people plus adjacent counties that are socioeconomically tied to the urban center by commuting. These standards are used to replace the definitions of metropolitan areas that were defined in 1990
- ➔ **Zip** - each year Zip-code
- ➔ **Jobs** - all persons currently working at an establishment including full-time, part-time, and temporary workers
- ➔ **JobsEstimate** - Actual (A) or Estimated (blank) Jobs beginning in 2020
- ➔ **CTBlock** - a combination of census blocks within a census tract (values 1-9). From 2019
- ➔ **PlaceType** - indicates if record is HQ (1), subsidiary (3), branch (2), or independent or individual (9): (one job where individual name=company name)
- ➔ **SIC** - the Standard Industrial Classification System is a numerical scheme used to classify businesses according to industry type - 6 digits
- ➔ **NAICS** - the 2017 definition of the North American Industry Classification System is the standard used by Federal statistical agencies in classifying business establishments for the purpose of collecting, analyzing, and publishing statistical data.
- ★ **FirstYear** - first year establishment enters the database
- ★ **LastYear** - last year establishment is in the database
- ★ **HQFIPS** - parent establishment's geographic location by county
- ★ **Seg** - market segments grouping by 6-digit NAICS: ET = external trade, LT = local trade, NTH = non trade/ healthcare
- ★ **Type** - 3 establishment types totaling 100%: R= resident (either standalone or HQ in the state examined) NR = nonresident (HQ not in the state examined) NC = noncommercial (government, nonprofit, determined by NAICS)
- ★ **Sales** - BDRC modeled sales volume at that location (in thousands)

- ★ **SalesCorporate** - Corporate (Firm) total sales taken from SEC (Public Companies) filings
- ★ **StockExchange** - the abbreviation for the name of the stock exchange
- ★ **CIK** - the Central Index Key (CIK) is used on the SEC's computer systems to identify corporations and individual people who have filed disclosure with the SEC
- ★ **CDIST** - congressional district identification (State FIPS+district) from the 117th Congress
- ★ **CTFIPS** - an 11-digit statistical subdivision of a county format: FIPS+census-tract from 2019