

LEHD Public Use Data Schema v4.0.1

| REVISION HISTORY | | | |
|------------------|------------------|-------------|------|
| NUMBER | DATE | DESCRIPTION | NAME |
| | 24 February 2015 | | LV |

Contents

| | | |
|----------|--|-----------|
| 1 | Basic Schema | 1 |
| 1.1 | Generic structure | 1 |
| 1.2 | Identifiers | 2 |
| 1.2.1 | Identifiers for qwi | 2 |
| 1.3 | Indicators | 3 |
| 2 | Categorical Variables | 4 |
| 2.1 | agegrp | 4 |
| 2.2 | education | 4 |
| 2.3 | ethnicity | 4 |
| 2.4 | firmage | 4 |
| 2.5 | firmsize | 5 |
| 2.6 | ownercode | 5 |
| 2.7 | periodicity | 5 |
| 2.8 | race | 5 |
| 2.9 | seasonadj | 5 |
| 2.10 | sex | 6 |
| 2.11 | Industry | 7 |
| 2.11.1 | Industry levels | 7 |
| 2.11.2 | Industry | 7 |
| 2.12 | Geography | 8 |
| 2.12.1 | Geographic levels | 8 |
| 2.12.2 | State-level values | 8 |
| 2.12.3 | Detailed state and substate level values | 8 |
| 3 | Status flags | 10 |
| 4 | Changes | 11 |
| 4.1 | Version 4.0.1 | 11 |

([Printable version](#))



Important

This document is not an official Census Bureau publication. It is compiled from publicly accessible information by Lars Vilhuber ([Labor Dynamics Institute, Cornell University](#)). Feedback is welcome. Please write us at lars.vilhuber@cornell.edu.

The public-use Quarterly Workforce Indicators (QWI) data from the Longitudinal Employer-Household Dynamics Program are available for download with the following data schema. These data are available as Comma-Separated Value (CSV) files through the LEHD website's Data page at <http://lehd.ces.census.gov/data/> and at an (occasional) mirror site at <http://download.vrdc.cornell.edu/-qwipu/>.

This document describes the data schema for QWI files. For each variable, a set of allowable values is defined. Definitions are provided as CSV files, with header variable definitions. The naming conventions of the data files is documented in [lehd_csv_naming.html](#). Changes relative to the original v4.0 version are listed [at the end](#).

1 Basic Schema

Each file is structured as a CSV file. The first columns contain [\[identifiers\]](#), subsequent columns contain [\[indicators\]](#), followed by [status flags](#).

1.1 Generic structure

| Column name |
|-------------------|
| [Identifier1] |
| [Identifier2] |
| [Identifier3] |
| [...] |
| [Indicator 1] |
| [Indicator 2] |
| [Indicator 3] |
| [...] |
| [Status Flag 1] |
| [Status Flag 2] |
| [Status Flag 3] |
| [...] |

Note: A full list of indicators for each type of file are shown below in the [Indicators](#) section. While all indicators are included in the CSV files, only the requested indicators will be included in data outputs from the LED Extraction Tool.

1.2 Identifiers

Records, unless otherwise noted, are parts of time-series data. Unique record identifiers are noted below, by file type. Identifiers without the year and quarter component can be considered a series identifier.

1.2.1 Identifiers for qwi

([lehd_identifiers_qwi.csv](#))

| Variable | Type | label |
|-------------|---------|--|
| periodicity | Char(1) | Periodicity of report |
| seasonadj | Char(1) | Seasonal Adjustment Indicator |
| geo_level | Char(1) | Group: Geographic level of aggregation |
| geography | Char(8) | Group: Geography code |
| ind_level | Char(1) | Group: Industry level of aggregation |
| industry | Char(5) | Group: Industry code |
| ownercode | Char(3) | Group: Ownership group code |
| sex | Char(1) | Group: Gender code |
| agegrp | Char(3) | Group: Age group code (WIA) |
| race | Char(2) | Group: race |
| ethnicity | Char(2) | Group: ethnicity |
| education | Char(2) | Group: education |
| firmage | Char(1) | Group: Firm Age group |
| firmsize | Char(1) | Group: Firm Size group |
| year | Num | Time: Year |
| quarter | Num | Time: Quarter |

1.3 Indicators

The following tables and associated mapping files list the indicators available on each file. The '*Indicator Variable*' is the short name of the variable on the CSV files, suitable for machine processing in a wide variety of statistical applications. When given, the '*Alternate name*' may appear in related documentation and articles. The '*Status Flag*' is used to indicate publication or data quality status (see [Status Flags](#)). The '*Indicator Name*' is a more verbose description of the indicator.

([variables_qwipu.csv](#))

| Indicator Variable | Alternate name | Status Flag | Indicator Name |
|--------------------|----------------|---------------|--|
| Emp | B | sEmp | Beginning-of-Quarter Employment: Counts |
| EmpEnd | E | sEmpEnd | End-of-Quarter Employment: Counts |
| EmpS | F | sEmpS | Full-Quarter Employment (Stable): Counts |
| EmpSpv | Fprev | sEmpSpv | Full-Quarter Employment in the Previous Quarter: Counts |
| EmpTotal | M | sEmpTotal | Employment - Reference Quarter: Counts |
| HirA | A | sHirA | Hires All: Counts (Accessions) |
| HirN | H1 | sHirN | Hires New: Counts |
| HirR | R1 | sHirR | Hires Recalls: Counts |
| Sep | S | sSep | Separations: Counts |
| HirAEnd | A2 | sHirAEnd | End-of-Quarter Hires |
| HirAEndR | A2R | sHirAEndR | End-of-Quarter Hiring Rate |
| SepBeg | S2 | sSepBeg | Beginning-of-Quarter Separations |
| SepBegR | S2R | sSepBegR | Beginning-of-Quarter Separation Rate |
| HirAs | A3 | sHirAs | Hires All (Stable): Counts (Flows into Full-Quarter Employment) |
| HirNs | H3 | sHirNs | Hires New (Stable): Counts (New Hires to Full-Quarter Status) |
| SepS | S3 | sSepS | Separations (Stable): Counts (Flow out of Full-Quarter Employment) |
| SepSnx | S3R | sSepSnx | Separations (Stable): Next Quarter: Counts (Flow out of Full-Quarter Employment) |
| TurnOvrS | FT | sTurnOvrS | Turnover (Stable) |
| FrmJbGn | JC | sFrmJbGn | Firm Job Gains: Counts (Job Creation) |
| FrmJobLs | JD | sFrmJobLs | Firm Job Loss: Counts (Job Destruction) |
| FrmJbC | JF | sFrmJbC | Firm Job Change: Net Change |
| HirAEndRepl | EI | sHirAEndRepl | Replacement Hires |
| HirAEndReplr | EIR | sHirAEndReplr | Replacement Hiring Rate |
| FrmJbGnS | FJC | sFrmJbGnS | Firm Job Gains (Stable): Counts |
| FrmJbLsS | FJD | sFrmJbLsS | Firm Job Loss (Stable): Counts |
| FrmJbCS | FJF | sFrmJbCS | Job Change (Stable): Net Change |
| EarnS | ZW3 | sEarnS | Full Quarter Employment (Stable): Average Monthly Earnings |
| EarnBeg | ZW1 | sEarnBeg | Beginning-of-Quarter Employment: Average Monthly Earnings |
| EarnHirAS | ZWFA | sEarnHirAS | Hires All (Stable): Average Monthly Earnings |
| EarnHireNS | ZWFH | sEarnHireNS | Hires New (Stable): Average Monthly Earnings |
| EarnSepS | ZWFS | sEarnSepS | Separations (Stable): Average Monthly Earnings |
| Payroll | W1 | sPayroll | Total Quarterly Payroll: Sum |

2 Categorical Variables

Categorical variable descriptions are displayed above each table, with the variable name shown in parentheses. Unless otherwise stated, every possible value/label combination for each categorical variable is listed. Please note that not all values will be available in every table.

2.1 agegrp

([label_agegrp.csv](#))

| agegrp | label |
|--------|------------------|
| A00 | All Ages (14-99) |
| A01 | 14-18 |
| A02 | 19-21 |
| A03 | 22-24 |
| A04 | 25-34 |
| A05 | 35-44 |
| A06 | 45-54 |
| A07 | 55-64 |
| A08 | 65-99 |

2.2 education

([label_education.csv](#))

| education | label |
|-----------|---|
| E0 | All Education Categories |
| E1 | Less than high school |
| E2 | High school or equivalent, no college |
| E3 | Some college or Associate degree |
| E4 | Bachelor's degree or advanced degree |
| E5 | Educational attainment not available (workers aged 24 or younger) |

2.3 ethnicity

([label_ethnicity.csv](#))

| ethnicity | label |
|-----------|------------------------|
| A0 | All Ethnicities |
| A1 | Not Hispanic or Latino |
| A2 | Hispanic or Latino |

2.4 firmage

([label_firmage.csv](#))

| firmage | label |
|---------|---------------|
| 0 | All Firm Ages |
| 1 | 0-1 Years |
| 2 | 2-3 Years |
| 3 | 4-5 Years |

| firmage | label |
|----------------|--------------|
| 4 | 6-10 Years |
| 5 | 11+ Years |

2.5 firmsize

([label_firmsize.csv](#))

| firmsize | label |
|-----------------|-------------------|
| 0 | All Firm Sizes |
| 1 | 0-19 Employees |
| 2 | 20-49 Employees |
| 3 | 50-249 Employees |
| 4 | 250-499 Employees |
| 5 | 500+ Employees |

2.6 ownercode

([label_ownercode.csv](#))

| ownercode | label |
|------------------|-----------------|
| A00 | All (1-5) |
| A05 | All Private (5) |

2.7 periodicity

([label_periodicity.csv](#))

| periodicity | label |
|--------------------|----------------|
| A | Annual data |
| Q | Quarterly data |

2.8 race

([label_race.csv](#))

| race | label |
|-------------|---|
| A0 | All Races |
| A1 | White Alone |
| A2 | Black or African American Alone |
| A3 | American Indian or Alaska Native Alone |
| A4 | Asian Alone |
| A5 | Native Hawaiian or Other Pacific Islander Alone |
| A6 | Some Other Race Alone (Not Used) |
| A7 | Two or More Race Groups |

2.9 seasonadj

([label_seasonadj.csv](#))

| seasonadj | label |
|------------------|---------------------|
| S | Seasonally adjusted |

| seasonadj | label |
|-----------|-------------------------|
| U | Not seasonally adjusted |

2.10 sex

([label_sex.csv](#))

| sex | label |
|-----|-----------|
| 0 | All Sexes |
| 1 | Male |
| 2 | Female |

2.11 Industry

2.11.1 Industry levels

([label_ind_level.csv](#))

| ind_level | label |
|-----------|-----------------------|
| 3 | NAICS Subsectors |
| 4 | NAICS Industry Groups |
| A | All Industries |
| S | NAICS Sectors |

2.11.2 Industry

([label_industry.csv](#))

Only a small subset of available values shown. The 2012 NAICS (North American Industry Classification System) is used for all years. For a full listing of all valid NAICS codes, see <http://www.census.gov/eos/www/naics/>.

| industry | label |
|----------|---|
| 00 | All NAICS Sectors |
| 000 | All NAICS Subsectors |
| 0000 | All NAICS Industry Groups |
| 11 | Agriculture, Forestry, Fishing and Hunting |
| 111 | Crop Production |
| 1111 | Oilseed and Grain Farming |
| 1112 | Vegetable and Melon Farming |
| ... | |
| 2382 | Building Equipment Contractors |
| 2383 | Building Finishing Contractors |
| 2389 | Other Specialty Trade Contractors |
| 31-33 | Manufacturing |
| 311 | Food Manufacturing |
| 3111 | Animal Food Manufacturing |
| 3112 | Grain and Oilseed Milling |
| 3113 | Sugar and Confectionery Product Manufacturing |
| ... | |

2.12 Geography

2.12.1 Geographic levels

([label_geo_level.csv](#))

| geo_level | label |
|-----------|----------------------------|
| C | Counties |
| M | Metropolitan/Micropolitan |
| S | States |
| W | Workforce Investment Areas |

Geography labels are provided in separate files by state. Note that cross-state CBSA will have state-specific parts, and thus will appear in multiple files. A separate [label_fipsnum.csv](#) contains values and labels for all entities of geo_level *n* or *s*, and is a summary of separately available files.

2.12.2 State-level values

([label_fipsnum.csv](#))

| geography | label |
|-----------|---------------|
| 02 | Alaska |
| 01 | Alabama |
| 05 | Arkansas |
| 04 | Arizona |
| 06 | California |
| 08 | Colorado |
| 09 | Connecticut |
| ... | |
| 47 | Tennessee |
| 48 | Texas |
| 49 | Utah |
| 51 | Virginia |
| 50 | Vermont |
| 53 | Washington |
| 55 | Wisconsin |
| 54 | West Virginia |

2.12.3 Detailed state and substate level values

For a full listing of all valid geography codes (except for WIA codes), see <http://www.census.gov/geo/maps-data/data/tiger.html>. Note about geography codes: Four types of geography codes are represented with this field. Each geography has its own code structure.

- State is the 2-digit **FIPS** code.
- County is the 5-digit FIPS code.
- Metropolitan/Micropolitan codes are constructed from the 2-digit state FIPS code and the 5-digit **CBSA** code provided by the Census Bureau's Geography Division.
 - In the QWI, the metropolitan/micropolitan areas are the state parts of the full CBSA areas.
- The WIA code is constructed from the 2-digit state FIPS code and the 6-digit WIA identifier provided by LED State Partners.

The 2014 vintage of Census TIGER geography is used for all tabulations as of the 2014Q3 release.

| State | Format file |
|-------|--|
| AK | label_geography_ak.csv |
| AL | label_geography_al.csv |
| AR | label_geography_ar.csv |
| AZ | label_geography_az.csv |
| CA | label_geography_ca.csv |
| CO | label_geography_co.csv |
| CT | label_geography_ct.csv |
| DC | label_geography_dc.csv |
| DE | label_geography_de.csv |
| FL | label_geography_fl.csv |
| GA | label_geography_ga.csv |
| HI | label_geography_hi.csv |
| IA | label_geography_ia.csv |
| ID | label_geography_id.csv |
| IL | label_geography_il.csv |
| IN | label_geography_in.csv |
| KS | label_geography_ks.csv |
| KY | label_geography_ky.csv |
| LA | label_geography_la.csv |
| MD | label_geography_md.csv |
| ME | label_geography_me.csv |
| MI | label_geography_mi.csv |
| MN | label_geography_mn.csv |
| MO | label_geography_mo.csv |
| MS | label_geography_ms.csv |
| MT | label_geography_mt.csv |
| NC | label_geography_nc.csv |
| ND | label_geography_nd.csv |
| NE | label_geography_ne.csv |
| NH | label_geography_nh.csv |
| NJ | label_geography_nj.csv |
| NM | label_geography_nm.csv |
| NV | label_geography_nv.csv |
| NY | label_geography_ny.csv |
| OH | label_geography_oh.csv |
| OK | label_geography_ok.csv |
| OR | label_geography_or.csv |
| PA | label_geography_pa.csv |
| RI | label_geography_ri.csv |
| SC | label_geography_sc.csv |
| SD | label_geography_sd.csv |
| TN | label_geography_tn.csv |
| TX | label_geography_tx.csv |
| UT | label_geography_ut.csv |
| VA | label_geography_va.csv |
| VT | label_geography_vt.csv |
| WA | label_geography_wa.csv |
| WI | label_geography_wi.csv |
| WV | label_geography_wv.csv |
| WY | label_geography_wy.csv |

3 Status flags

([label_flags.csv](#))

Each status flag in the tables above contains one of the following valid values. The values and their interpretation are listed in the table below.

| flag | label |
|------|--|
| -2 | no data available in this category for this quarter |
| -1 | data not available to compute this estimate |
| 0 | zero employment estimated or zero estimated denominator in a ratio - zero released |
| 1 | OK |
| 5 | Value suppressed because it does not meet US Census Bureau publication standards. |
| 6 | Value calculated from other released measures |
| 7 | Value calculated from other released measures - some of which have significantly distorted data |
| 9 | Data significantly distorted - fuzzed value released |
| 10 | Aggregate of cells - no significant distortion |
| 11 | Aggregate of cells not released because component cells do not meet U.S. Census Bureau publication standards |
| 12 | Aggregate of cells - some of which have significantly distorted data |

4 Changes

4.1 Version 4.0.1

- switched NAICS coding from 2007 to 2012

| |
|------------------------|
| This version: : 607 \$ |
|------------------------|