

LEHD Public Use CSV Naming Schema V4.2.0 - CSV Naming Convention

Contents

1 Purpose	1
2 Schema for Data File Contents	1
3 Extends	1
4 Supersedes	1
5 Basic Schema	1
5.1 QWIPU from the LED Extraction Tool	2
5.2 Metadata for QWI data files (version.txt)	2
5.3 Metadata for J2J data files (version.txt)	2
5.4 Metadata for J2JOD files	2
6 Description of Filename Components	2
6.1 Types	2
6.2 geohi	3
6.3 demo	3
6.4 fas	5
6.5 geocat	6
6.6 indcat	7
6.7 owncat	8
6.8 sa	9
6.9 ext	10
7 Changes	11
7.1 This version (revisions)	11
7.2 Version 4.2.0 from 4.1.3	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.

1 Purpose

The public-use data from the Longitudinal Employer-Household Dynamics Program, including the Quarterly Workforce Indicators (QWI) and Job-to-Job Flows (J2J), 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 through LED Extraction Tool at <http://ledextract.ces.census.gov/>.

This document describes the file naming schema for LEHD-provided CSV files.

2 Schema for Data File Contents

The contents (schema) for LEHD data files are described in [lehd_public_use_schema.pdf](#). The contents (schema) for shapefiles are described in [lehd_shapefiles.pdf](#).

3 Extends

This version reimplements some features from V4.0. Many files compliant with LEHD or QWI Schema v4.0 will also be compliant with this schema, but compatibility is not guaranteed.

4 Supersedes

This version supersedes V4.1, for files released as of R2018Q1

5 Basic Schema

All files are preceded by a file type definition, followed by additional information on the aggregation level of the file, or some other identifier.

```
[TYPE]_[DETAILS].[EXT]
```

([naming_convention.csv](#))

type	Used By	Description
[type]_[id]	qwix	CSV files queried from the LED Extraction Tool at http://ledextract.ces.census.gov/ .
[type]_[geohi]_[time]_[fwi]_[geo]_[j2j]_[j2j]_[geo]_[sa]	lehd_public_use_data	Comma-separated value files for currently provided files.
lehd_shp_[geohi]	shapefiles	ZIP files containing data for shapefiles
version_[demo]_[fwi]	text	Machine-readable text files with metadata on geographic and temporal coverage for QWI
version\[type]	metadata	Machine-readable text files with metadata on geographic and temporal coverage for J2J
j2jod_[geography]	data.csv	Complex metadata (used only in J2JOD) - see below

5.1 QWIPU from the LED Extraction Tool

Files downloaded through the LED Extraction Tool at <http://ledextract.ces.census.gov/> follow the following naming convention:

```
[type]_[id].[EXT]
```

where [id] is the Request ID (a unique string of characters) generated every time Submit Request” is clicked. The ID references each query submission made to the database.

5.2 Metadata for QWI data files (version.txt)

Metadata accompanies the data files, identifying provenance, geographic and temporal coverage. These files follow the following naming convention:

```
version_[demo]_[fas].txt
```

where each name component is described in more detail [below](#).

5.3 Metadata for J2J data files (version.txt)

Metadata accompanies the data files, identifying provenance, geographic and temporal coverage. These files follow the following naming convention:

```
version_[type].txt
```

where each name component is described in more detail [below](#).

5.4 Metadata for J2JOD files

Because the origin-destination (J2JOD) data link two regions, we provide an auxiliary file with the time range for which cells containing data for each geographic pairing may appear in a data release. The reference region will always be either the origin or the destination. These files follow the following naming convention:

```
j2jod_[geography]_avail.csv
```

where each name component is described in more detail [below](#).

6 Description of Filename Components

6.1 Types

([naming_type.csv](#))

type	Description	Explanation	url
j2j	Job-to-Job Flows Counts (J2J)	Counts of hires and separations, to and from employment and non-employment, summed across all establishments	tbd
j2jod	Job-to-Job Flows Origin-Destination	Origin and destination flows and earnings, tabulated by the destination establishment	tbd
j2jr	Job-to-Job Flows Rates	Hiring and separation rates, to and from employment and nonemployment, summed across all establishments	http://lehd.ces.census.gov/data/j2j_beta.html

type	Description	Explanation	url
lodes	LEHD Origin-Destination Employment Statistics (LODES)	LEHD Origin-Destination Employment Statistics (LODES) used by OnTheMap. Version 7 and higher of LODES was enumerated by 2010 census blocks. Version 6 and lower of LODES were enumerated with 2000 census blocks.	http://lehd.ces.census.gov/data/#lodes
qwi	Quarterly Workforce Indicators Public Use Counts (QWIPU)	The Quarterly Workforce Indicators (QWI) are a set of economic indicators including employment, job creation, earnings, and other measures of employment flows. The QWI are reported using detailed firm characteristics (geography, industry, age, size) and worker demographics information (sex, age, education, race, ethnicity).	http://lehd.ces.census.gov/data/#qwi
qwir	Quarterly Workforce Indicators Public Use Rates	Rates for flow variables in the Quarterly Workforce Indicators (QWI)	tbd
qwirv	Quarterly Workforce Indicators Variability Measures for Rates	Variability measures for rates in the Quarterly Workforce Indicators (QWI)	tbd
qwiv	Quarterly Workforce Indicators Variability Measures	Variability measures for Quarterly Workforce Indicators (QWI)	tbd
lehd_shp	LEHD Shapefiles	Shapefiles are used to provide mapping functionality in QWI Explorer and Job-to-Job Explorer	lehd_shapefiles.html

6.2 geohi

([naming_geohi.csv](#))

The geohi component is based on one of two possible code sets:

- the lower-case alphabetic FIPS or postal state code based on [FIPS PUB 5-2](#) (see also [label_stusps.csv](#) for upper-case variants).
- the numeric CBSA code corresponding to the metropolitan areas (see [label_geography_metro.csv](#))
- It is expanded to encompass additional codes denoting national coverage, or a collection of states.

type	Description
all	Collection of all available states
us	National data (50 states + DC)
st	Any legal 2-character lower-case state postal code
NNNN	Any valid CBSA-derived code listed in label_geography_metro.csv

6.3 demo

([naming_demo.csv](#))

type	Description
d	No demographic detail

type	Description
rh	Race by Ethnicity tabulations
sa	Sex by Age tabulations
se	Sex by Education tabulations
sarhe	Collection of sa-rh-se tabulations

6.4 fas

([naming_fas.csv](#))

type	Description
f	No firm size or age detail
fa	Tabulations by firm age
fs	Tabulations by firm size

6.5 geocat

([naming_geocat.csv](#))

type	Description
gb	Metropolitan (complete)
gc	Counties
gm	Metropolitan/Micropolitan (state parts)
gn	National (50 States + DC)
gs	States
gw	Workforce Investment Areas

6.6 indcat

([naming_indcat.csv](#))

type	Description
n	all industries
ns	NAICS sectors
n3	NAICS subsectors
n4	NAICS industry groups

6.7 owncat

([naming_owncat.csv](#))

type	Description
of	Federal government (QWI Code A01)
op	All Private (QWI Code A05)
oslp	State, local, and private ownership categories (QWI Code A00)

6.8 sa

([naming_sa.csv](#))

type	Description
s	Seasonally adjusted
u	Not seasonally adjusted

6.9 ext

([naming_ext.csv](#))

type	Description
csv	Comma-Separated file
pdf	PDF file (printable)
xls	Excel file (only select output)
xlsx	Excel file (not available for larger files)

7 Changes

For a description of how schema files are versioned, see [main directory](#).

7.1 This version (revisions)

- 2017-12-15: Initial release
- 2018-03-23: Fix EOL issues

7.2 Version 4.2.0 from 4.1.3

- Updated industry classification from NAICS 2012 to NAICS 2017
- Added a column `ind_level` to `label_industry.csv` similar to the `geo_level`
- Added additional columns to the variable metadata schema for greater clarity
 - Description,
 - Concept,
 - Base
- Added a (draft) taxonomy of concepts used in the LEHD data world ([label_concept_draft.csv](#))
- Fixed the labeling of ownership code A00 to correctly reflect scope
- Added files describing the number of quarters of data availability required relative to start and end quarters ([lags_qwi.csv](#) and [lags_j2j.csv](#)), and its metadata ([variables_lags.csv](#))



Important

Some of the data products noted above do not exist yet.

This revision: Tue Apr 17 17:54:57 EDT 2018