

LEHD Public Use Schema V4.3-rc1 - File and Directory Naming Convention

Contents

1	Purpose	1
2	Schema for Data File Contents	1
3	Extends	1
4	Supersedes	1
5	Basic Filename Schema	1
5.1	QWIPU from the LED Extraction Tool	2
5.2	Metadata for QWI and J2J data files (version.txt)	2
5.3	Metadata for J2JOD files	2
6	Directory Paths	2
7	Description of Naming Components	2
7.1	Release	2
7.2	Types and Products	2
7.3	geohi	3
7.4	demo	4
7.5	fas	5
7.6	geocat	6
7.7	indcat	7
7.8	owncat	8
7.9	sa	9
7.10	ext	10
8	Changes	11
8.1	This version (revisions)	11
8.2	Version 4.3.0 from 4.2.0	11



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.

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

type	Used By	Description
[type]_[id]	qwx	CSV files queried from the LED Extraction Tool at http://ledextract.ces.census.gov/ .
[type]_[geohi]_[id]_[fasi]_[geoi]_[id]_[jod]_[id]_[id]	qwx	Current convention for currently provided files.
lehd_shp_[geoshp]	shpfiles	ZIP files containing data for shapefiles
version\[type]_[id]	metadata	Machine-readable text files with metadata on geographic and temporal coverage
j2jod_[geography]_[id].csv	metadata	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 and 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](#).

5.3 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 Directory Paths

Downloadable data files are organized on the download server (at the time of this writing: <https://lehd.ces.census.gov/data/>) in directories organized as follows:

```
SERVER/data/[PRODUCT]/[RELEASE]/[GEOHI] [/[TYPE] ]/{files}
```

where [\[PRODUCT\]](#), [\[RELEASE\]](#), [\[GEOHI\]](#), and [\[TYPE\]](#) are each described below. [\[TYPE\]](#) is optional.

7 Description of Naming Components

7.1 Release

A **release** is defined by the calendar year quarter when data production occurs. It is thus generically constructed as

```
R[YEAR]Q[QUARTER]
```

where [YEAR] is the 4-digit year and [QUARTER] the single-digit calendar year quarter (1-4).

7.2 Types and Products

([naming_type.csv](#))

type	product	Description	Explanation	url
j2j	j2j	Job-to-Job Flows Counts (J2J)	Counts of hires and separations, to and from employment and non-employment, summed across all establishments	http://lehd.ces.census.gov/-data/#j2j
j2jod	j2j	Job-to-Job Flows Origin-Destination	Origin and destination flows and earnings, tabulated by the destination establishment	http://lehd.ces.census.gov/-data/#j2j
j2jr	j2j	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
lodes	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	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	qwi	Quarterly Workforce Indicators Public Use Rates	Rates for flow variables in the Quarterly Workforce Indicators (QWI)	tbd
qwirv	qwi	Quarterly Workforce Indicators Variability Measures for Rates	Variability measures for rates in the Quarterly Workforce Indicators (QWI)	tbd
qwiv	qwi	Quarterly Workforce Indicators Variability Measures	Variability measures for Quarterly Workforce Indicators (QWI)	tbd
lehd_shp	all	LEHD Shapefiles	Shapefiles are used to provide mapping functionality in QWI Explorer and Job-to-Job Explorer	lehd_shapefiles.html

7.3 geohi

([naming_geohi.csv](#))

The geohi component in **filenames** 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.

For **directories**, files at the CBSA level are collected under a single **metro** directory.

type	Description
all	Collection of all available states
us	National data (50 states + DC)
metro	Indicates collection of CBSA-level files (directory names only)
[st]	Any legal 2-character lower-case state postal code
[NNNNN]	Any valid CBSA-derived code listed in label_geography_metro.csv

7.4 demo

([naming_demo.csv](#))

type	Description
d	No demographic detail
rh	Race by Ethnicity tabulations
sa	Sex by Age tabulations
se	Sex by Education tabulations
sarhe	Collection of sa-rh-se tabulations

7.5 fas

([naming_fas.csv](#))

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

7.6 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

7.7 indcat

([naming_indcat.csv](#))

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

7.8 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)

7.9 sa

([naming_sa.csv](#))

type	Description
s	Seasonally adjusted
u	Not seasonally adjusted

7.10 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)

8 Changes

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

8.1 This version (revisions)

- 2018-06-25: Initial release

8.2 Version 4.3.0 from 4.2.0

- Changed the format of the version.txt file. This does not affect data files, but it does affect the metadata, and thus requires an increase in the minor version.
- Changed the description of the download locations of the QWI files, as part of the "naming" schema.

**Important**

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

This revision: Mon Jun 25 15:14:36 EDT 2018