

LEHD Origin-Destination Employment Statistics (LODES) Dataset Structure

Format Version 7.1

Introduction

The LEHD Origin-Destination Employment Statistics (LODES) datasets are released both as part of the OnTheMap application and in raw form as a set of comma separated variable (CSV) text files. This document describes the structure of those raw files and provides basic information for users who want to perform analytical work on the data outside of the OnTheMap application.

File Organization

At the root distribution level, each state (or state-equivalent) has a directory named by its lowercase, 2-letter postal code. Within each state directory are directories of data as well as metadata and geography files for the state. These are described below.

Data Files

The CSV data files are released at the state level and they are organized into three groups within each state. The three groups of files are named as follows, according to their contents:

- OD** – Origin-Destination data, jobs totals are associated with both a home Census Block and a work Census Block
- RAC** – Residence Area Characteristic data, jobs are totaled by home Census Block
- WAC** – Workplace Area Characteristic data, jobs are totaled by work Census Block

Metadata Files and Geography

Within the state directory are two text files. One is a version document, called “version.txt”, and the other is a “checksum” file. Also in the state directory is a geographic relationship file (referred to here as a “geography crosswalk”).

The version file contains information about the state, the data vintage, and the format version. The “data vintage” is an eight character code based on the date the data were prepared for release; it is in the form of YYYYMMDD.¹ New or corrected data may cause newer vintages of data to be released. Only files that have new or changed data will be included in future vintages. The “format version” is a number that refers to file organization and structure of the data release and is fully defined by this document. Restructuring of the existing data into different file formats or a different organization structure would result in a change in the format version.

The checksum file in each state directory is an MD5SUM file, called “otm_[ST].md5sum” where [ST] is the respective state postal code. With the correct software tools, the MD5SUM file can be used to check that the individual data files were downloaded without any file corruption or loss of data. Further information on MD5SUM can be found at <http://en.wikipedia.org/wiki/Md5sum>.

¹ Note: The data vintage is *not* the same as the createdate provided within the data files. The createdate refers to an internal processing point, while the data vintage is the code that refers to a specific release of public OnTheMap data.

Datasets for 2010 and later contain additional Job Types that cover Federal employment as supplied by the Office of Personnel Management (OPM). The data structure is outlined in the File Naming Conventions and File Structure segment of this document.

States and territories which are not in “regular production” do have RAC datasets available, however these only include jobs for workers employed in the 50 OnTheMap states. The partners that are *not* in regular production for the purposes of LODES/OnTheMap are Massachusetts, Puerto Rico, and the U.S. Virgin Islands. Additionally, due to historical unavailability/data sharing limitations/significant QA issues, a handful of state-year combinations of data are *not* part of LODES/OnTheMap. These full list of missing stat-year combinations is given in the table below:

State-Year Combinations without OD/WAC data	
State (or equivalent)	Years
Arizona	2002-2003
Arkansas	2002
District of Columbia	2002-2009
Kansas	2012-2013
Massachusetts	All Years
Mississippi	2002-2003
New Hampshire	2002
Puerto Rico	All Years
U.S. Virgin Islands	All Years

These specific state-year combinations will have only RAC files and not OD or WAC files, similar to those states not in regular production. For a state-year with a full complement of data, in 2009 and earlier there are 8 OD files, 40 RAC files, and 40 WAC files. Starting with 2010 and going forward there are 12 OD files, 60 RAC files, and 60 WAC files. In some cases, individuals files produced for a state-year combination may contain no data because there are no jobs in the specific combination of state-year-jobtype-filetype. When this occurs the individual file still exists, but only contains the file header as specified below.

For more information about Firm Age and Firm Size variables, please see <http://lehd.ces.census.gov/research/>. For more information about Federal employment data in LODES, please see <http://lehd.ces.census.gov/doc/help/onthemap/FederalEmploymentInOnTheMap.pdf>. For more information on the LED partnership, please visit <http://lehd.ces.census.gov/>.

Geography Vintage

LODES Version 7.1 is enumerated with 2010 census blocks. Prior to LODES Version 6.0, data releases used 2000 census blocks as the base vintage. Basic information on 2010 census blocks can be found at http://www.census.gov/geo/maps-data/data/tiger-line.html#tab_2010. General information on the relationships between 2000 census blocks and 2010 census blocks can be found at <http://www.census.gov/geo/maps-data/data/relationship.html>. The methods used to

translate historical data into 2010 census blocks can be found at <http://lehd.ces.census.gov/doc/help/onthemap/OnTheMap2010Geography.pdf>.

Format Version

This document – LODES Format Version 7.1 – describes the structure for the data released along with the following OnTheMap Application versions:

OnTheMap Version 6.4.1 – Released August 12, 2015

OnTheMap Version 6.4 – Released July 7, 2015

LODES Format Version 7.0 describes the structure of the data released with the earlier OnTheMap Application Versions.⁵

OnTheMap Version 6.3 – Released December 3, 2014

OnTheMap Version 6.2 – Released December 3, 2013

OnTheMap Version 6.1.2 – Released May 15, 2013

LODES Format Version 6.1 describes the structure of the data released with the earlier OnTheMap Application Versions.⁶

OnTheMap Version 6.1.1 – Released August 6, 2012

OnTheMap Version 6.1 – Released May 31, 2012

LODES Format Version 6.0 describes the structure of data released with the earlier OnTheMap Application versions.^{7,8}

Application Version 6.0 – Released January 26, 2012

Format version 5.0⁹ describes the structure of data released along with the earlier OnTheMap Application versions:

Application Version 5.0 – Released December 13, 2010

Application Version 5.1 – Released February 15, 2011

Format version 4.1¹⁰ describes the structure of data released along with the earlier OnTheMap Application versions:

Application Version 3.0 – Released September 4, 2008

Application Version 3.1 – Released September 11, 2008

Application Version 3.2 – Released December 15, 2008

Application Version 4.0 – Released December 14, 2009

Application Version 4.1 – Released May 27, 2010

⁵ See <http://lehd.ces.census.gov/onthemap/LODES7/LODESTechDoc7.0.pdf> for more information.

⁶ See <http://lehd.ces.census.gov/onthemap/LODES6/LODESTechDoc6.1.pdf> for more information.

⁷ See <http://lehd.ces.census.gov/onthemap/LODES6/LODESTechDoc6.0.pdf> for more information.

⁸ The only difference in structure between Version 6.0 and Version 6.1 is the addition of two new job types to cover Federal employment.

⁹ See <http://lehd.ces.census.gov/onthemap/LODES5/OnTheMapDataTechDoc5.0.pdf> for more information.

¹⁰ See <http://lehd.ces.census.gov/onthemap/LODES5/OnTheMapDataTechDoc4.1.pdf> for more information.

For further information on the background and release history of OnTheMap data, please visit http://lehd.did.census.gov/applications/help/onthemap.html#!what_is_onthemap.

File Naming Conventions and File Structure

The structure of the three types of files are described on the following pages:

OD

Filename of the OD datasets are described by the following templates

[ST]_od_[PART]_[TYPE]_[YEAR].csv.gz where

[ST] = lowercase, 2-letter postal code for a chosen state

[PART] = Part of the state file, can have a value of either “main” or “aux”. Complimentary parts of the state file, the main part includes jobs with both workplace and residence in the state and the aux part includes jobs with the workplace in the state and the residence outside of the state.

[TYPE] = Job Type, can have a value of “JT00” for All Jobs, “JT01” for Primary Jobs, “JT02” for All Private Jobs, “JT03” for Private Primary Jobs, “JT04” for All Federal Jobs, or “JT05” for Federal Primary Jobs.

[YEAR] = Year of job data. Can have the value of 2002-2013 for most states.

As an example the main OD file of Primary Jobs in 2007 for California would be the file:

ca_od_main_JT01_2007.csv.gz

The structure of the OD files is as follows:

Origin-Destination (OD) File Structure			
Pos	Variable	Type	Explanation
1	w_geocode	Char15	Workplace Census Block Code
2	h_geocode	Char15	Residence Census Block Code
3	S000	Num	Total number of jobs
4	SA01	Num	Number of jobs of workers age 29 or younger ¹¹
5	SA02	Num	Number of jobs for workers age 30 to 54 ¹¹
6	SA03	Num	Number of jobs for workers age 55 or older ¹¹
7	SE01	Num	Number of jobs with earnings \$1250/month or less
8	SE02	Num	Number of jobs with earnings \$1251/month to \$3333/month
9	SE03	Num	Number of jobs with earnings greater than \$3333/month
10	SI01	Num	Number of jobs in Goods Producing industry sectors
11	SI02	Num	Number of jobs in Trade, Transportation, and Utilities industry sectors
12	SI03	Num	Number of jobs in All Other Services industry sectors
13	createdate	Char	Date on which data was created, formatted as YYYYMMDD

¹¹ For the 2012 data year and later, the job frame has been constrained to include only workers age 14 to 99, so that the youngest age group becomes 14 to 29 and the oldest age group becomes 55 to 99.

RAC

Filenames of the RAC datasets are described by the following template:

[ST]_rac_[SEG]_[TYPE]_[YEAR]_1.csv.gz where

[ST] = lowercase, 2-letter postal code for a chosen state

[SEG] = Segment of the workforce, can have the values of “S000”, “SA01”, “SA02”, “SA03”, “SE01”, “SE02”, “SE03”, “SI01”, “SI02”, or “SI03”. These correspond to the same segments of the workforce as are listed in the OD file structure above.

[TYPE] = Job Type, can have a value of “JT00” for All Jobs, “JT01” for Primary Jobs, “JT02” for All Private Jobs, “JT03” for Private Primary Jobs, “JT04” for All Federal Jobs, or “JT05” for Federal Primary Jobs.

[YEAR] = Year of job data. Can have the value of 2002-2013 for most states.

As an example the RAC file of All Jobs for workers age 29 or younger in 2004 for New Jersey would be the file: nj_rac_SA01_JT00_2004.csv.gz

Residence Area Characteristics (RAC) File Structure			
Pos	Variable	Type	Explanation
1	h_geocode	Char15	Residence Census Block Code
2	C000	Num	Total number of jobs
3	CA01	Num	Number of jobs for workers age 29 or younger ¹¹
4	CA02	Num	Number of jobs for workers age 30 to 54 ¹¹
5	CA03	Num	Number of jobs for workers age 55 or older ¹¹
6	CE01	Num	Number of jobs with earnings \$1250/month or less
7	CE02	Num	Number of jobs with earnings \$1251/month to \$3333/month
8	CE03	Num	Number of jobs with earnings greater than \$3333/month
9	CNS01	Num	Number of jobs in NAICS sector 11 (Agriculture, Forestry, Fishing and Hunting)
10	CNS02	Num	Number of jobs in NAICS sector 21 (Mining, Quarrying, and Oil and Gas Extraction)
11	CNS03	Num	Number of jobs in NAICS sector 22 (Utilities)
12	CNS04	Num	Number of jobs in NAICS sector 23 (Construction)
13	CNS05	Num	Number of jobs in NAICS sector 31-33 (Manufacturing)
14	CNS06	Num	Number of jobs in NAICS sector 42 (Wholesale Trade)
15	CNS07	Num	Number of jobs in NAICS sector 44-45 (Retail Trade)
16	CNS08	Num	Number of jobs in NAICS sector 48-49 (Transportation and Warehousing)
17	CNS09	Num	Number of jobs in NAICS sector 51 (Information)
18	CNS10	Num	Number of jobs in NAICS sector 52 (Finance and Insurance)
19	CNS11	Num	Number of jobs in NAICS sector 53 (Real Estate and Rental and Leasing)
20	CNS12	Num	Number of jobs in NAICS sector 54 (Professional, Scientific, and Technical Services)
21	CNS13	Num	Number of jobs in NAICS sector 55 (Management of Companies and Enterprises)
22	CNS14	Num	Number of jobs in NAICS sector 56 (Administrative and Support and Waste Management and Remediation Services)
23	CNS15	Num	Number of jobs in NAICS sector 61 (Educational Services)
24	CNS16	Num	Number of jobs in NAICS sector 62 (Health Care and Social Assistance)
25	CNS17	Num	Number of jobs in NAICS sector 71 (Arts, Entertainment, and Recreation)
26	CNS18	Num	Number of jobs in NAICS sector 72 (Accommodation and Food Services)
27	CNS19	Num	Number of jobs in NAICS sector 81 (Other Services [except Public Administration])
28	CNS20	Num	Number of jobs in NAICS sector 92 (Public Administration)
29	CR01	Num	Number of jobs for workers with Race: White, Alone ¹²

¹² Race, Ethnicity, Educational Attainment, and Sex variables are only available for data years 2009 and later and are made available through a Beta release.

30	CR02	Num	Number of jobs for workers with Race: Black or African American Alone ¹²
31	CR03	Num	Number of jobs for workers with Race: American Indian or Alaska Native Alone ¹²
32	CR04	Num	Number of jobs for workers with Race: Asian Alone ¹²
33	CR05	Num	Number of jobs for workers with Race: Native Hawaiian or Other Pacific Islander Alone ¹²
34	CR07	Num	Number of jobs for workers with Race: Two or More Race Groups ¹²
35	CT01	Num	Number of jobs for workers with Ethnicity: Not Hispanic or Latino ¹²
36	CT02	Num	Number of jobs for workers with Ethnicity: Hispanic or Latino ¹²
37	CD01	Num	Number of jobs for workers with Educational Attainment: Less than high school ^{12,13}
38	CD02	Num	Number of jobs for workers with Educational Attainment: High school or equivalent, no college ^{12,13}
39	CD03	Num	Number of jobs for workers with Educational Attainment: Some college or Associate degree ^{12,13}
40	CD04	Num	Number of jobs for workers with Educational Attainment: Bachelor's degree or advanced degree ^{12,13}
41	CS01	Num	Number of jobs for workers with Sex: Male ¹²
42	CS02	Num	Number of jobs for workers with Sex: Female ¹²
43	createdate	Char	Date on which data was created, formatted as YYYYMMDD

¹³ Educational Attainment is only available for workers age 30 and older.

WAC

Filenames of the WAC datasets are described by the following template:

[ST]_wac_[SEG]_[TYPE]_[YEAR].csv.gz where

[ST] = lowercase, 2-letter postal code for a chosen state

[SEG] = Segment of the workforce, can have the values of “S000”, “SA01”, “SA02”, “SA03”, “SE01”, “SE02”, “SE03”, “SI01”, “SI02”, or “SI03”. These correspond to the same segments of the workforce as are listed in the OD file structure above.

[TYPE] = Job Type, can have a value of “JT00” for All Jobs, “JT01” for Primary Jobs, “JT02” for All Private Jobs, “JT03” for Private Primary Jobs, “JT04” for All Federal Jobs, or “JT05” for Federal Primary Jobs.

[YEAR] = Year of job data. Can have the value of 2002-2013 for most states.

As an example the WAC file of Private Primary Jobs for workers earning more than \$3333/mo in 2008 for Iowa would be the file: ia_wac_SE03_JT03_2008.csv.gz

Workplace Area Characteristics (WAC) File Structure			
Pos	Variable	Type	Explanation
1	w_geocode	Char15	Workplace Census Block Code
2	C000	Num	Total number of jobs
3	CA01	Num	Number of jobs for workers age 29 or younger ¹¹
4	CA02	Num	Number of jobs for workers age 30 to 54 ¹¹
5	CA03	Num	Number of jobs for workers age 55 or older ¹¹
6	CE01	Num	Number of jobs with earnings \$1250/month or less
7	CE02	Num	Number of jobs with earnings \$1251/month to \$3333/month
8	CE03	Num	Number of jobs with earnings greater than \$3333/month
9	CNS01	Num	Number of jobs in NAICS sector 11 (Agriculture, Forestry, Fishing and Hunting)
10	CNS02	Num	Number of jobs in NAICS sector 21 (Mining, Quarrying, and Oil and Gas Extraction)
11	CNS03	Num	Number of jobs in NAICS sector 22 (Utilities)
12	CNS04	Num	Number of jobs in NAICS sector 23 (Construction)
13	CNS05	Num	Number of jobs in NAICS sector 31-33 (Manufacturing)
14	CNS06	Num	Number of jobs in NAICS sector 42 (Wholesale Trade)
15	CNS07	Num	Number of jobs in NAICS sector 44-45 (Retail Trade)
16	CNS08	Num	Number of jobs in NAICS sector 48-49 (Transportation and Warehousing)
17	CNS09	Num	Number of jobs in NAICS sector 51 (Information)
18	CNS10	Num	Number of jobs in NAICS sector 52 (Finance and Insurance)
19	CNS11	Num	Number of jobs in NAICS sector 53 (Real Estate and Rental and Leasing)
20	CNS12	Num	Number of jobs in NAICS sector 54 (Professional, Scientific, and Technical Services)
21	CNS13	Num	Number of jobs in NAICS sector 55 (Management of Companies and Enterprises)
22	CNS14	Num	Number of jobs in NAICS sector 56 (Administrative and Support and Waste Management and Remediation Services)
23	CNS15	Num	Number of jobs in NAICS sector 61 (Educational Services)
24	CNS16	Num	Number of jobs in NAICS sector 62 (Health Care and Social Assistance)
25	CNS17	Num	Number of jobs in NAICS sector 71 (Arts, Entertainment, and Recreation)
26	CNS18	Num	Number of jobs in NAICS sector 72 (Accommodation and Food Services)
27	CNS19	Num	Number of jobs in NAICS sector 81 (Other Services [except Public Administration])
28	CNS20	Num	Number of jobs in NAICS sector 92 (Public Administration)
29	CR01	Num	Number of jobs for workers with Race: White, Alone ¹²
30	CR02	Num	Number of jobs for workers with Race: Black or African American Alone ¹²
31	CR03	Num	Number of jobs for workers with Race: American Indian or Alaska Native Alone ¹²
32	CR04	Num	Number of jobs for workers with Race: Asian Alone ¹²

33	CR05	Num	Number of jobs for workers with Race: Native Hawaiian or Other Pacific Islander Alone ¹²
34	CR07	Num	Number of jobs for workers with Race: Two or More Race Groups ¹²
35	CT01	Num	Number of jobs for workers with Ethnicity: Not Hispanic or Latino ¹²
36	CT02	Num	Number of jobs for workers with Ethnicity: Hispanic or Latino ¹²
37	CD01	Num	Number of jobs for workers with Educational Attainment: Less than high school ^{12,13}
38	CD02	Num	Number of jobs for workers with Educational Attainment: High school or equivalent, no college ^{12,13}
39	CD03	Num	Number of jobs for workers with Educational Attainment: Some college or Associate degree ^{12,13}
40	CD04	Num	Number of jobs for workers with Educational Attainment: Bachelor's degree or advanced degree ^{12,13}
41	CS01	Num	Number of jobs for workers with Sex: Male ¹²
42	CS02	Num	Number of jobs for workers with Sex: Female ¹²
43	CFA01	Num	Number of jobs for workers at firms with Firm Age: 0-1 Years ¹⁴
44	CFA02	Num	Number of jobs for workers at firms with Firm Age: 2-3 Years ¹⁴
45	CFA03	Num	Number of jobs for workers at firms with Firm Age: 4-5 Years ¹⁴
46	CFA04	Num	Number of jobs for workers at firms with Firm Age: 6-10 Years ¹⁴
47	CFA05	Num	Number of jobs for workers at firms with Firm Age: 11+ Years ¹⁴
48	CFS01	Num	Number of jobs for workers at firms with Firm Size: 0-19 Employees ^{14,15}
49	CFS02	Num	Number of jobs for workers at firms with Firm Size: 20-49 Employees ^{14,15}
50	CFS03	Num	Number of jobs for workers at firms with Firm Size: 50-249 Employees ^{14,15}
51	CFS04	Num	Number of jobs for workers at firms with Firm Size: 250-499 Employees ^{14,15}
52	CFS05	Num	Number of jobs for workers at firms with Firm Size: 500+ Employees ^{14,15}
53	createdate	Char	Date on which data was created, formatted as YYYYMMDD

¹⁴ Firm Age and Firm Size variables are only available for data years 2011 and later, for All Private Jobs (JT02), and are made available through a Beta release.

¹⁵ For firms 1 year and older, firm size is the national size of the firm in March of the *previous* year.

Geography Crosswalk

Filenames of the geography crosswalk datasets are described by the following template:

[ST]_xwalk.csv.gz where

[ST] = lowercase, 2-letter postal code for a chosen state

As an example the geography crosswalk file for Michigan would be the file: mi_xwalk.csv.gz

Code Vintages

The primary key/identifier in the Geography Crosswalk is the 2010 Census Tabulation Block Code (tabblk2010). This code does not change between Decennial Censuses and can be used to link to the geocodes in the OD/RAC/WAC files. All other geographic codes in the crosswalk are “current” and represent the most recent definitions of the legal, statistical, or administrative areas as integrated into LODES and OnTheMap.

Geography Crosswalk File Structure			
Pos	Variable	Type	Explanation¹⁶
1	tabblk2010	Char15	2010 Census Tabulation Block Code
2	st	Char2	FIPS State Code
3	stusps	Char2	USPS State Code
4	stname	Char100	State Name
5	cty	Char5	FIPS County Code
6	ctyname	Char100	County or County Equivalent Name
7	trct	Char11	Census Tract Code
8	trctname	Char100	Tract Name, formatted with County and State
9	bgrp	Char12	Census Blockgroup Code
10	bgrpname	Char100	Census Blockgroup Name, formatted with Tract, County, and State
11	cbsa	Char5	CBSA (Metropolitan/Micropolitan Area) Code
12	cbsaname	Char100	CBSA (Metropolitan/Micropolitan Area) Name
13	zcta	Char5	ZIP Code Tabulation Area (ZCTA) Code
14	zctaname	Char100	ZCTA Name
15	stplc	Char7	Nationally Unique Place Code, (FIPS State + FIPS Place)
16	stplcname	Char100	Place Name
17	ctycsub	Char10	Nationally Unique County Subdivision Code, (FIPS State + FIPS County + FIPS County Subdivision)
18	ctycsubname	Char100	County Subdivision Name
19	std114	Char4	Nationally Unique 114 th Congressional District Code, (FIPS State + 2-digit District Number) ¹⁷
20	std114name	Char100	114 th Congressional District Name
21	stslldl	Char5	Nationally Unique State Legislative District, Lower Chamber, (FIPS State + 3-digit District Number)
22	stslldlname	Char100	State Legislative District Chamber, Lower Chamber

¹⁶ All codes are constructed to be nationally unique. For areas that do not have a FIPS or other identifier that is nationally unique, the State FIPS or both State and County FIPS codes are prepended to create uniqueness. Source information for each geography can be found at http://lehd.ces.census.gov/applications/help/onthemap.html#!geographic_data.

¹⁷ States with At Large districts have a district code of “00”. State equivalents with Resident Commissioners or Delegates have a district code of “98”.

23	stsldu	Char5	Nationally Unique State Legislative District, Upper Chamber, (FIPS State + 3-digit District Number)
24	stslname	Char100	State Legislative District Chamber, Upper Chamber Chamber
25	stschool	Char7	Nationally Unique Unified/Elementary School District Code, (FIPS State + 5-digit Local Education Agency Code)
26	stschoolname	Char100	Unified/Elementary School District Name
27	stsecon	Char7	Nationally Unique Secondary School District Code, (FIPS State + 5-digit Local Education Agency Code)
28	stseconname	Char100	Secondary School District Name
29	trib	Char5	American Indian /Alaska Native/Native Hawaiian Area Census Code
30	tribname	Char100	American Indian /Alaska Native/Native Hawaiian Area Name
31	tsub	Char7	American Indian Tribal Subdivision Code
32	tsubname	Char100	American Indian Tribal Subdivision Name
33	stanrc	Char7	Nationally Unique Alaska Native Regional Corporation (ANRC) Code (FIPS State + FIPS ANRC)
34	stanrcname	Char100	Alaska Native Regional Corporation Name
35	mil	Char22	Military Installation Landmark Code
36	milname	Char100	Military Installation Name
37	stwib	Char8	Nationally Unique Workforce Innovation Board (WIB) Area Code (FIPS State + state-provided 6-digit WIB Area Code)
38	stwibname	Char100	Workforce Innovation Board Area Name
39	createdate	Char8	Date on which data was created, formatted as YYYYMMDD