LEHD Origin-Destination Employment Statistics (LODES) Dataset Structure
Format Version 6.0

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
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

At the root distribution level, each state has a directory named by its lowercase, 2-letter postal code. Within each state directory are three directories, one for each of the files groups described above. Also within the state directory are two text files. One is a version document, called "version.txt", and the other is a "checksum" file.

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 was prepared for release; it is 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.

A typical directory tree looks as follows:
```
\
|-- ak
|   |-- od
|   |   `-- (individual origin-destination data files)
|   |-- rac
|   |   `-- (individual residence area characteristics data files)
|   `-- wac
```

---

[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.

```
|   |      `-- (individual workplace area characteristics data files)
|   |-- otm_ak.md5sum
|   `-- version.txt
|-- az
|
(more states)
```

The CSV files are compressed using a GZip algorithm, which appends a ".gz" to the end of the filename. Current compression/extraction software packages should be able to access these files.

Data Coverage and Availability
Data is currently available for most states in most years. The LED Partnership consists of 53 State and Territorial Partners, of which 49 are producing data for OnTheMap. With the exceptions listed below all 49 states have OD, RAC, and WAC data for all 9 years of data (2002-2010).

Datasets for 2009 and 2010 contain additional variables (Race, Ethnicity, Education, and Sex) on the RAC and WAC files that are not available in other years of data.[2] The data structure for all years of data has been expanded to included space for these new variables, but in years before 2009 these variables contain only zeroes. See the File Naming Conventions and File Structure segment of this document for more information.

States and territories which are not in "regular production" do have RAC datasets available, however these only include jobs for workers employed in the 49 OnTheMap states. The partners that are *not* in regular production for the purposes of LODES/OnTheMap are Massachusetts, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands.

Due to historical unavailability a handful of state-year combinations of data are *not* part of OnTheMap. These are Arizona (2002 and 2003), Arkansas (2002), Mississippi (2002 and 2003), and New Hampshire (2002). 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, there are 8 OD files, 40 RAC files, and 40 WAC files.

For more information on the LED partnership, please visit http://lehd.did.census.gov/.

Geography Vintage
LODES Version 6.0 is the first release of data to be enumerated with 2010 census blocks. Prior releases of LODES used 2000 census blocks as the base vintage. Basic information on 2010 census blocks can be found at http://www.census.gov/geo/www/tiger/tgrshp2010/tgrshp2010.html and at http://www.census.gov/geo/www/2010census/. General information on the relationships between 2000 census blocks and 2010 census blocks can be found at http://www.census.gov/geo/www/2010census/rel_blk.html. The methods used to translate

---

[2] Race, Ethnicity, Educational Attainment, and Sex variables are made available through a Beta release.

historical data into 2010 census blocks can be found at
http://lehd.ces.census.gov/led/datatools/doc/OnTheMap2010Geography.pdf.

Further questions about changes of the geographic vintage within the LODES dataset should be
sent to CES.OnTheMap.Feedback@census.gov

Format Version
With the release of data accompanying Version 6 of the OnTheMap application, the format
version of the data release is being re-baselined. This document (format version 6.0) describes
the structure for the data released along with the following OnTheMap Application versions:

        Application Version 6.0 – Released January 26, 2012

Format version 5.0[3] 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 4.0[4] 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

For further information on the background and release history of OnTheMap data, please visit
http://lehd.did.census.gov/led/datatools/onthemap.html

---

[3] See http://lehd.did.census.gov/led/onthemap/LODES5/OnTheMapDataTechDoc5.0.pdf for more information.
[4] See http://lehd.did.census.gov/led/onthemap/LODES5/OnTheMapDataTechDoc4.1.pdf for more information.

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, or "JT03" for Private Primary Jobs.
> [YEAR] = Year of job data. Can have the value of 2002-2009 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:

Pos	Variable	Type	Length	Explanation
1	w_geocode	Char	15	Workplace Census Block Code
2	h_geocode	Char	15	Residence Census Block Code
3	S000	Num	8	Total number of jobs
4	SA01	Num	8	Number of jobs of workers age 29 or younger
5	SA02	Num	8	Number of jobs for workers age 30 to 54
6	SA03	Num	8	Number of jobs for workers age 55 or older
7	SE01	Num	8	Number of jobs with earnings $1250/month or less
8	SE02	Num	8	Number of jobs with earnings $1251/month to $3333/month
9	SE03	Num	8	Number of jobs with earnings greater than $3333/month
10	SI01	Num	8	Number of jobs in Goods Producing industry sectors
11	SI02	Num	8	Number of jobs in Trade, Transportation, and Utilities industry sectors
12	SI03	Num	8	Number of jobs in All Other Services industry sectors
13	createdate	Char	8	Date on which data was created, formatted as YYYYMMDD

Table caption: Origin-Destination (OD) File Structure

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, or "JT03" for Private Primary Jobs.
> [YEAR] = Year of job data. Can have the value of 2002-2009 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

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

29	CR01	Num	8	Number of jobs for workers with Race: White, Alone[5]
30	CR02	Num	8	Number of jobs for workers with Race: Black or African American Alone[5]
31	CR03	Num	8	Number of jobs for workers with Race: American Indian or Alaska Native Alone[5]
32	CR04	Num	8	Number of jobs for workers with Race: Asian Alone[5]
33	CR05	Num	8	Number of jobs for workers with Race: Native Hawaiian or Other Pacific Islander Alone[5]
34	CR07	Num	8	Number of jobs for workers with Race: Two or More Race Groups[5]
35	CT01	Num	8	Number of jobs for workers with Ethnicity: Not Hispanic or Latino[5]
36	CT02	Num	8	Number of jobs for workers with Ethnicity: Hispanic or Latino[5]
37	CD01	Num	8	Number of jobs for workers with Educational Attainment: Less than high school[5,6]
38	CD02	Num	8	Number of jobs for workers with Educational Attainment: High school or equivalent, no college[5,6]
39	CD03	Num	8	Number of jobs for workers with Educational Attainment: Some college or Associate degree[5,6]
40	CD04	Num	8	Number of jobs for workers with Educational Attainment: Bachelor's degree or advanced degree[5,6]
41	CS01	Num	8	Number of jobs for workers with Sex: Male[5]
42	CS02	Num	8	Number of jobs for workers with Sex: Female[5]
43	createdate	Char	8	Date on which data was created, formatted as YYYYMMDD

---

[5] Race, Ethnicity, Educational Attainment, and Sex variables are only available for 2009 and 2010 data and are made available through a Beta release.

[6] 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, or "JT03" for Private Primary Jobs.
> [YEAR] = Year of job data. Can have the value of 2002-2008 for most states.

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

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

31	CR03	Num	8	Number of jobs for workers with Race: American Indian or Alaska Native Alone[5]
32	CR04	Num	8	Number of jobs for workers with Race: Asian Alone[5]
33	CR05	Num	8	Number of jobs for workers with Race: Native Hawaiian or Other Pacific Islander Alone[5]
34	CR07	Num	8	Number of jobs for workers with Race: Two or More Race Groups[5]
35	CT01	Num	8	Number of jobs for workers with Ethnicity: Not Hispanic or Latino[5]
36	CT02	Num	8	Number of jobs for workers with Ethnicity: Hispanic or Latino[5]
37	CD01	Num	8	Number of jobs for workers with Educational Attainment: Less than high school[5,6]
38	CD02	Num	8	Number of jobs for workers with Educational Attainment: High school or equivalent, no college[5,6]
39	CD03	Num	8	Number of jobs for workers with Educational Attainment: Some college or Associate degree[5,6]
40	CD04	Num	8	Number of jobs for workers with Educational Attainment: Bachelor's degree or advanced degree[5,6]
41	CS01	Num	8	Number of jobs for workers with Sex: Male[5]
42	CS02	Num	8	Number of jobs for workers with Sex: Female[5]
43	createdate	Char	8	Date on which data was created, formatted as YYYYMMDD