LEHD Public Use Data Schema V4.1d-draft

Contents

(Printable version)

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 data schema for LEHD files. For each variable, a set of allowable values is defined. Definitions are provided as CSV files, with header variable definitions. Changes relative to the original v4.0 version are listed at the end.

2 File naming

The naming conventions of the data files is documented in lehd_csv_naming.pdf.

3 Replaces

This version replaces v4.0. Any file compliant with LEHD or QWI Schema v4.0 will also be compliant with this schema, but some changes may have been applied to auxiliary data.

4 Basic Schema

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

4.1 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.

4.2 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.

4.2.1 Mapping for Identifiers

(lehd_mapping_identifiers.csv)

Each of the released files has a set of variables uniquely identifying records (*Identifiers*). The table below relates the set of identifier specifications to the released files. The actual CSV files containing the identifiers for each set are listed after this table. Each identifier can take on a specified list of values, documented in the section on Categorical Variables.

identifiers	QWI	NQWI	J2J	J2JR	J2JOD	LODES
lehd_identifiers_qwi	1	1				
lehd_identifiers_j2j			1	1		
lehd_identifiers_j2jod					1	

4.2.2 Identifiers for j2j

(lehd_identifiers_j2j.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
agg_level	Num	Aggregation Level Indicator

4.2.3 Identifiers for j2jod

(lehd_identifiers_j2jod.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 of destination job
geography	Char(8)	Group: Geography code of destination job
ind_level	Char(1)	Group: Industry level of aggregation of destination job
industry	Char(5)	Group: Industry code of destination job
ownercode	Char(3)	Group: Ownership group code of destination job
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
agg_level	Num	Aggregation Level Indicator
geo_level_orig	Char(1)	Group: Geographic level of aggregation of origin job
geography_orig	Char(8)	Group: Geography code of origin job
ind_level_orig	Char(1)	Group: Industry level of aggregation of origin job
industry_orig	Char(5)	Group: Industry code of origin job
ownercode_orig	Char(3)	Group: Ownership group code of origin job
firmage_orig	Char(1)	Group: Firm Age group of origin job
firmsize_orig	Char(1)	Group: Firm Size group of origin job

4.2.4 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

4.3 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.

4.3.1 National QWI and state-level QWI (QWIPU)

(variables_qwi.csv)

Indicator Variable	Alternate name	Status Flag	Indicator Name	Units
Emp	B	sEmp	Beginning-of-Quarter Employment	Count
EmpEnd	E	sEmpEnd	End-of-Quarter Employment	Count
EmpS	F	sEmpS	Full-Quarter Employment (Stable)	Count
EmpSpv	Fpv	sEmpSpv	Full-Quarter Employment in the Previous Quarter	Count
EmpTotal	M	sEmpTotal	Employment - Reference Quarter	Count
HirA	A	sHirA	Hires (All Accessions)	Count
HirN	H	sHirN	New Hires	Count
HirR	R	sHirR	Recall Hires	Count
Sep	S	sSep	Separations (All)	Count
HirAEnd	CA	sHirAEnd	End-of-Quarter Hires	Count
HirAEndR	CAR	sHirAEndR	End-of-Quarter Hiring Rate	Rate
SepBeg	CS	sSepBeg	Beginning-of-Quarter Separations	Count
SepBegR	CSR	sSepBegR	Beginning-of-Quarter Separation Rate	Rate
HirAS	FA	sHirAS	Hires (All Hires into Full-Quarter Employment)	Count
HirNS	FH	sHirNS	New Hires (New Hires into Full-Quarter Employment)	Count
SepS	FS	sSepS	Separations (Flows out of Full-Quarter Employment)	Count
SepSnx	FSnx	sSepSnx	Separations in the Next Quarter (Flows out of Full-Quarter Employment)	Count
TurnOvrS	FT	sTurnOvrS	Turnover (Stable)	Rate
FrmJbGn	JC	sFrmJbGn	Firm Job Gains (Job Creation)	Count
FrmJbLs	JD	sFrmJbLs	Firm Job Loss (Job Destruction)	Count
FrmJbC	JF	sFrmJbC	Firm Job Change (Net Change)	Count
HirAEndRepl	EI	sHirAEndRepl	Replacement Hires	Count
HirAEndReplr	EIR	sHirAEndReplr	Replacement Hiring Rate	Rate
FrmJbGnS	FJC	sFrmJbGnS	Firm Job Gains (Stable)	Count
FrmJbLsS	FJD	sFrmJbLsS	Firm Job Loss (Stable)	Count
FrmJbCS	FJF	sFrmJbCS	Firm Job Change (Stable; Net Change)	Count
EarnS	ZW3	sEarnS	Average Monthly Earnings (Full-Quarter Employment)	Dollars
EarnBeg	ZW1	sEarnBeg	Average Monthly Earnings (Beginning-of-Quarter Employment)	Dollars
EarnHirAS	ZWFA	sEarnHirAS	Average Monthly Earnings (All Hires into Full-Quarter Employment)	Dollars
EarnHirNS	ZWFH	sEarnHirNS	Average Monthly Earnings (New Hires into Full-Quarter Employment)	Dollars
EarnSepS	ZWFS	sEarnSepS	Average Monthly Earnings (Flows out of Full-Quarter Employment)	Dollars
Payroll	W1	sPayroll	Total Quarterly Payroll	Dollars

4.3.2 National QWI and state-level QWI rates (QWIPUR)

Rates are computed from published data, and are provided as a convenience. The column **Base** indicates the denominator used to compute the rate.

(variables_qwir.csv)

Indicator Variable	Alternate name	Status Flag	Indicator Name	Units	Base
HirAR	AR	sHirAR	Hiring Rate (All Accessions)	Rate	(Emp + EmpEnd)/2
HirNR	HR	sHirNR	New Hiring Rate	Rate	(Emp + EmpEnd)/2
HirRR	RR	sHirRR	Recall Rate	Rate	(Emp + EmpEnd)/2
SepR	SR	sSepR	Separation Rate (All Separations)	Rate	(Emp + EmpEnd)/2
HirAEndR	CAR	sHirAEndR	End-of-Quarter Hiring Rate	Rate	(Emp + EmpEnd)/2
SepBegR	CSR	sSepBegR	Beginning-of-Quarter Separation Rate	Rate	(Emp + EmpEnd)/2
HirAsR	FAR	sHirAsR	Hiring Rate (Flows into Full-Quarter Employment)	Rate	(EmpSpv + EmpS)/2
HirNsR	FHR	sHirNsR	New Hiring Rate (New Hires to Full-Quarter Status)	Rate	(EmpSpv + EmpS)/2
SepSR	FSR	sSepSR	Separation Rate (Flows out of Full-Quarter Employment)	Rate	(EmpSpv + EmpS)/2
SepSnxR	FSnxR	sSepSnxR	Separation Rate in the Next Quarter (Flow out of Full-Quarter Employment)	Rate	(EmpSpv + EmpS)/2
TurnOvrSR	FTR	sTurnOvrSR	Turnover Rate (Stable)	Rate	2*EmpS
FrmJbGnR	JCR	sFrmJbGnR	Firm Job Gain Rate (Job Creation Rate)	Rate	(Emp + EmpEnd)/2
FrmJbLsR	JDR	sFrmJbLsR	Firm Job Loss Rate (Job Destruction Rate)	Rate	(Emp + EmpEnd)/2
FrmJbCR	JFR	sFrmJbCR	Firm Job Change Rate (Net Change Rate)	Rate	(Emp + EmpEnd)/2
HirAEndReplR	EIR	sHirAEndReplR	Replacement Hiring Rate	Rate	(Emp + EmpEnd)/2
FrmJbGnSR	FJCR	sFrmJbGnSR	Firm Job Gain Rate (Stable)	Rate	(EmpSpv + EmpS)/2
FrmJbLsSR	FJDR	sFrmJbLsSR	Firm Job Loss Rate (Stable)	Rate	(EmpSpv + EmpS)/2
FrmJbCSR	FJFR	sFrmJbCSR	Firm Job Change Rate (Stable; Net Change Rate)	Rate	(EmpSpv + EmpS)/2

4.3.3 Job-to-job flow counts (J2J)

(variables_j2j.csv)

Indicator Variable	Alternate name	Status Flag	Indicator Name
MHire	all_doma2	sMHire	Main Job Accessions - Counts
MSep	all_doms2	sMSep	Main Job Separations - Counts
MJobStart	all_dest	sMJobStart	Main Job Starts - Counts (hires + main source of earnings changes)
MJobEnd	all_orig	sMJobEnd	Main Job Ends - Counts (separations + main source of earnings changes)
EEHire	ee_doma2	sEEHire	EE Accessions - Counts
EESep	ee_doms2	sEESep	EE Separations - Counts
AQHire	aq_doma2	sAQHire	AQ Accessions - Short Duration Nonemployment Spell - Counts
AQSep	aq_doms2	sAQSep	AQ Separations - Short Duration Nonemployment Spell - Counts
J2JHire	eeall_doma2	sJ2JHire	Job-to-Job Accessions - Counts
J2JSep	eeall_doms2	sJ2JSep	Job-to-Job Separations - Counts
NEHire	ne_doma2	sNEHire	Main Job Accessions - Any Nonemployment Spell - Counts
ENSep	en_doms2	sENSep	Main Job Separations - Any Nonemployment Spell - Counts
NEPersist	ne2_doma2	sNEPersist	Main Job Accessions - Persistent Nonemployment Spell - Counts
ENPersist	en2_doms2	sENPersist	Main Job Separations - Persistent Nonemployment Spell - Counts
NEFullQ	ne2p_doma2	sNEFullQ	Main Job Accessions - Full-qtr Nonemployment Spell - Counts
ENFullQ	en2p_doms2	sENFullQ	Main Job Separations - Full-qtr Nonemployment Spell - Counts
MainB	domB	sMainB	Main Job - Beginning of Quarter - Counts
MainE	domE	sMainE	Main Job - End of Quarter - Counts

4.3.4 Job-to-job flow rates (J2JR)

(variables_j2jr.csv)

Rates are computed from published data, and are provided as a convenience. The column **Base** indicates the denominator used to compute the rate.

Indicator Variable	Alternate name	Status Flag	Indicator Name	Base
MHireR	all_doma2_rate	sMHireR	Main Job Accessions - Rate	(MainB+MainE)/2
MSepR	all_doms2_rate	sMSepR	Main Job Separations - Rate	(MainB+MainE)/2
MJobStartR	all_dest_rate	sMJobStartRate	Main Job Starts - Rate	(MainB+MainE)/2
MJobEndR	all_orgin_rate	sMJobEndRate	Main Job Ends - Rate	(MainB+MainE)/2
EEHireR	eea_rate	sEEHireRate	EE Accessions - Rate	(MainB+MainE)/2
EESepR	ees_rate	sEESepR	EE Separations - Rate	(MainB+MainE)/2
AQHireR	aq_doma2_rate	sAQHireR	AQ Accessions - Short Duration Nonemployment Spell - Rate	(MainB+MainE)/2
AQSepR	aq_doms2_rate	sAQSepR	AQ Separations - Short Duration Nonemployment Spell - Rate	(MainB+MainE)/2
J2JHireR	eeall_doma2_rate	sJ2JHireR	Job-to-Job Accessions - Rate	(MainB+MainE)/2
J2JSepR	eeall_doms2_rate	sJ2JSepR	Job-to-Job Separations - Rate	(MainB+MainE)/2
NEHireR	ne_doma2_rate	sNEHireR	Main Job Accessions - Any Nonemployment Spell - Rate	(MainB+MainE)/2
ENSepR	en_doms2_rate	sENSepR	Main Job Separations - Any Nonemployment Spell - Rate	(MainB+MainE)/2
NEPersistR	ne2_doma2_rate	sNEPersistR	Main Job Accessions - Persistent Nonemployment Spell - Rate	(MainB+MainE)/2
ENPersistR	en2_doms2_rate	sENPersistR	Main Job Separations - Persistent Nonemployment Spell - Rate	(MainB+MainE)/2
NEFullQR	ne2p_doma2_rate	sNEFullQR	Main Job Accessions - Full-qtr Nonemployment Spell - Rate	(MainB+MainE)/2
ENFullQR	en2p_doms2_rate	sENFullQR	Main Job Separations - Full-qtr Nonemployment Spell - Rate	(MainB+MainE)/2

4.3.5 Job-to-job flow Origin-Destination (J2JOD)

(variables_j2jod.csv)

Indicator Variable	Alternate name	Status Flag	Indicator Name
EE	ee	sEE	Direct Job Flows - Counts
AQHire	aq_doma2	sAQHire	Main Job Accessions - Short Nonemployment Spell - Counts
EEFullQ	fee	sEEFullQ	Stable Job to Stable Job Direct Flows - Counts
AQFullQHire	faq_doma2	sAQFullQHire	Stable Job Accessions From Stable Jobs - Short Nonemployment Spell Between Jobs - Counts

4.4 Variability measures

The following tables and associated mapping files list the variability measures available on each file. The *'Variability Measure*' 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 *'Variable Name*' is a more verbose description of the variability measure.

Six variability measures are published:

- Total variability, prefixed by vt_
- Standard error, prefixed by st_, and computed as the square root of Total Variability
- Between-implicate variability, prefixed by vb_
- Average within-implicate variability, prefixed by vw_
- Degrees of freedom, prefixed by df_
- Missingness ratio, prefixed by mr_

A missing variability measure indicates a structural zero in the corresponding indicator. This is currently not associated with a flag.

4.4.1 Generic structure

Column name
[Identifier1]
[Identifier2]
[Identifier3]
[. . .]
[Standard error for Indicator 1]
[Standard error for Indicator 2]
[Standard error for Indicator 3]
[. . .]
[Total variation for Indicator 1]
[Total variation for Indicator 2]
[Total variation for Indicator 3]
[. . .]
[Between-implicate variability for Indicator 1]
[Between-implicate variability for Indicator 2]
[Between-implicate variability for Indicator 3]
[. . .]
[Average within-implicate variability for Indicator 1]
[Average within-implicate variability for Indicator 2]
[Average within-implicate variability for Indicator 3]
[. . .]
[Degrees of freedom for Indicator 1]
[Degrees of freedom for Indicator 2]

Column name
[Degrees of freedom for Indicator 3]
[. . .]
[Missingness ratio for Indicator 1]
[Missingness ratio for Indicator 2]
[Missingness ratio for Indicator 3]
[. . .]

Note: A full list of indicators for each type of file are shown in the Indicators section. In the tables below, only a sample of variability measures are printed, but the complete list is available in the linked CSV schema files.

4.4.2 National QWI and state-level QWI

(variables_qwiv.csv)

Variability measure	Alternate name	Variable name	Units
st_Emp	st_B	Standard error of Beginning-of-Quarter Employment	Count
st_EmpEnd	st_E	Standard error of End-of-Quarter Employment	Count
st_EmpS	st_F	Standard error of Full-Quarter Employment (Stable)	Count
...			
vt_Emp	vt_B	Total variation of Beginning-of-Quarter Employment	Count
vt_EmpEnd	vt_E	Total variation of End-of-Quarter Employment	Count
vt_EmpS	vt_F	Total variation of Full-Quarter Employment (Stable)	Count
...			
vb_Emp	vb_B	Between-implicate variability for Beginning-of-Quarter Employment	Count
vb_EmpEnd	vb_E	Between-implicate variability for End-of-Quarter Employment	Count
vb_EmpS	vb_F	Between-implicate variability for Full-Quarter Employment (Stable)	Count
...			
...			
df_Emp	df_B	Degrees of freedom for VT of Beginning-of-Quarter Employment	Count
df_EmpEnd	df_E	Degrees of freedom for VT of End-of-Quarter Employment	Count
df_EmpS	df_F	Degrees of freedom for VT of Full-Quarter Employment (Stable)	Count
...			
mr_Emp	mr_B	Missingness ratio for Beginning-of-Quarter Employment	Count
mr_EmpEnd	mr_E	Missingness ratio for End-of-Quarter Employment	Count
mr_EmpS	mr_F	Missingness ratio for Full-Quarter Employment (Stable)	Count

4.4.3 National QWI and state-level QWI rates

(variables_qwirv.csv)

Variability measure	Alternate name	Variable name	Units
st_HirAR	st_AR	Standard error of Hiring Rate (All Accessions)	Rate
st_HirNR	st_HR	Standard error of New Hiring Rate	Rate
st_HirRR	st_RR	Standard error of Recall Rate	Rate
...			
vt_HirAR	vt_AR	Total variation of Hiring Rate (All Accessions)	Rate
vt_HirNR	vt_HR	Total variation of New Hiring Rate	Rate
vt_HirRR	vt_RR	Total variation of Recall Rate	Rate
...			
vb_HirAR	vb_AR	Between-implicate variability for Hiring Rate (All Accessions)	Rate
vb_HirNR	vb_HR	Between-implicate variability for New Hiring Rate	Rate
vb_HirRR	vb_RR	Between-implicate variability for Recall Rate	Rate
...			
...			
df_HirAR	df_AR	Degrees of freedom for VT of Hiring Rate (All Accessions)	Rate
df_HirNR	df_HR	Degrees of freedom for VT of New Hiring Rate	Rate
df_HirRR	df_RR	Degrees of freedom for VT of Recall Rate	Rate
...			
mr_HirAR	mr_AR	Missingness ratio for Hiring Rate (All Accessions)	Rate
mr_HirNR	mr_HR	Missingness ratio for New Hiring Rate	Rate
mr_HirRR	mr_RR	Missingness ratio for Recall Rate	Rate

4.4.4 Job-to-job flow counts (J2J)

Soon.

4.4.5 Job-to-job flow rates (J2JR)

Soon.

4.4.6 Job-to-job flow Origin-Destination (J2JOD)

Soon.

5 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.

5.1 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

5.2 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)

5.3 ethnicity

(label_ethnicity.csv)

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

5.4 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
N	Firm Age Not Available

5.5 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
N	Firm Size Not Available

5.6 ownercode

(label_ownercode.csv)

ownercode	label
A00	All (1-5)
A01	Federal government
A05	All Private (5)

5.7 periodicity

(label_periodicity.csv)

periodicity	label
A	Annual data
Q	Quarterly data

5.8 quarter

(label_quarter.csv)

quarter	label
1	1st Quarter of the Year (January-March)
2	2nd Quarter of the Year (April-June)
3	3rd Quarter of the Year (July-September)
4	4th Quarter of the Year (October-December)

5.9 race

(label_race.csv)

race	label
A0	All Races

race	label
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

5.10 seasonadj

(label_seasonadj.csv)

seasonadj	label
S	Seasonally adjusted
U	Not seasonally adjusted

5.11 sex

(label_sex.csv)

sex	label
0	All Sexes
1	Male
2	Female

5.12 stusps

(label_stusps.csv)

geography	stusps
01	AL
02	AK
04	AZ
05	AR
06	CA
08	CO
09	CT
10	DE
11	DC
12	FL
13	GA
15	HI
16	ID
17	IL
18	IN
19	IA
20	KS
21	KY
22	LA
23	ME
24	MD
25	MA
26	MI

geography	stusps
27	MN
28	MS
29	MO
30	MT
31	NE
32	NV
33	NH
34	NJ
35	NM
36	NY
37	NC
38	ND
39	OH
40	OK
41	OR
42	PA
44	RI
45	SC
46	SD
47	TN
48	TX
49	UT
50	VT
51	VA
53	WA
54	WV
55	WI
56	WY
72	PR
78	VI

5.13 Industry

5.13.1 Industry levels

(label_ind_level.csv)

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

5.13.2 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. QWI releases prior to R2015Q3 used the 2007 NAICS classification (see Schema v4.0.1). For a full listing of all valid 2012 NAICS codes, see http://www.census.gov/cgi-bin/sssd/naics/naicsrch?chart=2012.

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
...	

5.14 Geography

5.14.1 Geographic levels

Geography labels for data files are provided in separate files, by scope. Each file *label_geograpy_SCOPE.csv* may contain one or more types of records as flagged by geo_level. For convenience, a composite file containing all geocodes is available as label_geography.csv. The 2015 vintage of Census TIGER/Line geography is used for all tabulations as of the R2015Q4 release.

Shapefiles are described in a separate document.

(label_geo_level.csv)

geo_level	label	description	sourceurl
B	Metropolitan/Micropolitan (complete)	Identifies 5-digit CBSA code provided by the Census Bureau's Geography Division	http://www.census.gov/-population/metro/
C	Counties	Identifies 5-digit FIPS code	https://www.census.gov/geo/-reference/codes/cou.html
M	Metropolitan/Micropolitan (state part)	Identifies 7-digit code constructed from the 2-digit state FIPS code and the 5-digit CBSA code provided by the Census Bureau's Geography Division	
N	National (50 States + DC)	Custom code using *00* to denote national scope	
S	States	Identifies 2-digit FIPS code (also called "ANSI" codes)	https://www.census.gov/geo/-reference/ansi_statetables.html
W	Workforce Investment Areas	2-digit state FIPS code and the 6-digit WIA identifier provided by LED State Partners	

5.14.2 National and state-level values

(label_fipsnum.csv)

The file label_fipsnum.csv contains values and labels for all entities of geo_level *N* or *S*, and is a summary of separately available files.

geography	label	geo_level
00	National (50 States + DC)	N
01	Alabama	S
02	Alaska	S
04	Arizona	S
05	Arkansas	S
06	California	S
08	Colorado	S
...		
45	South Carolina	S
46	South Dakota	S
47	Tennessee	S
48	Texas	S
49	Utah	S
50	Vermont	S
51	Virginia	S
53	Washington	S

5.14.3 Detailed state and substate level values

Note: cross-state CBSA, in records of type geo_level = M, are present on files of type *label_geography_XX.csv*. A particular cross-state CBSA will appear on multiple files.

Scope	Format file	
US	label_geography_us.csv	
CBSA	label_geography_cbsa.csv	
States		
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	
MA	label_geography_ma.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	

Scope	Format file	
WI	label_geography_wi.csv	
WV	label_geography_wv.csv	
WY	label_geography_wy.csv	

5.15 Aggregation level

(label_agg_level.csv)

Measures within the J2J and QWI data products are tabulated on many different dimensions, including demographic characteristics, geography, industry, and other firm characteristics. For Origin-Destination (O-D) tables, characteristics of the origin and destination firm can be tabulated separately. Every tabulation level is assigned a unique aggregation index, represented by the agg_level variable. This index starts from 1, representing a national level grand total (all industries, workers, etc.), and progresses through different combinations of characteristics. There are gaps in the progression to leave space for aggregation levels that may be included in future data releases.

agg_level is currently reported only for J2J data products.

The following variables are included in the label_agg_level.csv file:

Variable	Description
agg_level	index representing level of aggregation reported on a given record
worker_char	demographic (worker) characteristics reported on record
firm_char	firm characteristics reported on record. These will be the characteristics of the destination firm in O-D tabulations
firm_orig_char	characteristics of origin firm reported on record (O-D tabulations only)
j2j	Flag: Aggregation level available on J2J counts tables
j2jr	Flag: Aggregation level available on J2J rates tables
j2jod	Flag: Aggregation level available on J2J O-D tables
qwi	Flag: Aggregation level available on QWI

The characteristics available on an aggregation level are repeated using a series of flags following the standard schema:

- geo_level - geographic level of table
- ind_level - industry level of table
- by_ variables - flags indicating other dimensions reported, including ownership, demographics, firm age and size.

These flags will be expanded to include origin characteristics in a later release. A shortened representation of the file is provided below, the complete file is available in the link above.

agg_level	worker_char	firm_char	firm_orig_char	j2j	j2jr	j2jod	qwi	geo_level
1				1	1	1	0	N
2	Sex			1	1	1	0	N
3	Age			1	1	1	0	N
4	Sex * Age			1	1	1	0	N
5	Race			1	1	1	0	N
9	Ethnicity			1	1	1	0	N
13	Race * Ethnicity			1	1	1	0	N
...								
129		Firm Size		1	1	1	0	N
257		NAICS Sector		1	1	1	0	N
513		NAICS Subsector		0	0	0	0	N

agg_level	worker_char	firm_char	firm_orig_char	j2j	j2jr	j2jod	qwi	geo_level
...								
3329		NAICS Sector * CBSA-state part		1	1	1	0	B
12289			Origin [Firm Age]	0	0	1	0	N
12353		Destination [Firm Age]	Origin [Firm Age]	0	0	0	0	N
...								

6 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.

Important
Note: Currently, the J2J tables only contain status flags *-1* and *1*. Status flags with values 10 or above only appear in online applications, not in CSV files.

flag	label
-2	no data available in this category for this quarter
-1	data not available to compute this estimate
1	OK
5	Value suppressed because it does not meet US Census Bureau publication standards.
6	Value calculated from other released measures - no significant distortion
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

7 Changes

7.1 This version from previous releases of V4.1 draft schema documents

- 2015-02-25: corrected flag values
- 2015-02-25: documents are now identified by date, not revision
- 2015-03-10: Correction of the TIGER vintage that is used for geographic references
- 2015-03-11: Added URL for J2J
- 2015-03-11: Correction of typo in type naming convention, rename of naming_fipsalpha.csv to naming_geohi.csv to be consistent.
- 2015-03-17: Changing of naming convention for to-be-released files based on federal government from fg → of. At this time, no such files have been released.
- 2015-04-24: Changes to alternate name of SepSnx and EmpSpv, tentative rate names
- 2015-04-26: Changes to name of variable schema files (qwipu → qwi), addition of variability variable schema files.
- 2015-04-28: Fixed small typos in QWI variable short names
- 2015-05-18: Updated agg_level description, replaced agg_level.csv file
- 2015-05-22: Fixed minor rendering bug for QWI rate variability names. No change to actual metadata.
- 2015-06-09: Fixed a minor coding error in label_fipsnum.csv, added a concatenation of geography files as label_geography_all.csv.
- 2015-08-07: Minor text change for agg_level, modified agg_level file.
- 2015-08-12: Removed the last 4 rows of variables_j2jod.csv, since they are not on the current beta J2JOD files.
- 2015-08-25: Added a extension component [ext] to the file naming convention to reflect availability of Excel files (and PDF files)
- 2016-03-16: Removed extraneous empty lines
- 2016-04-12: Fixed typo in variables_qwi.csv (FrmJbLsS, EarnHirNS and status variables)
- 2016-04-12: Incorporated all state-level geography from 4.0.5
- 2016-04-12: Added additional geo_level for CBSA (complete)
- 2016-04-12: Added additional agg_level values
- 2016-04-12: Clarified labels of j2jod identifiers
- 2016-04-12: Renamed label_geography_all as label_geography
- 2016-04-12: Added a column geo_level to all label_geography_* files

7.2 Version 4.1d-draft from 4.0.1

- added J2J, National QWI spec
- added geohi category of ALL, US
- added definitions of variability measures
- added definitions of rates on separate file
- added naming convention for additional files

- added agg_level variable
- added additional geo_level for CBSA (complete)
- added SHP files and description thereof

This revision: Mon Aug 8 18:16:05 EDT 2016