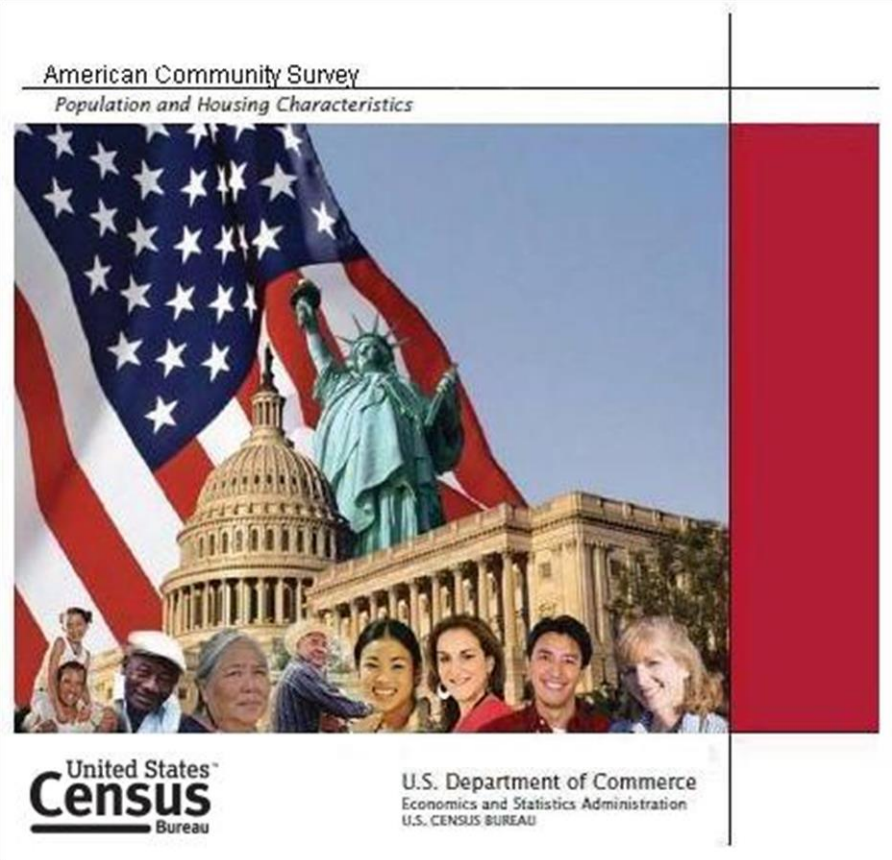
Version 1

EEO 2018 FTP Site Technical Documentation



Social, Economic, and Housing Statistics Division
American Community Survey Office
Version 1 February 5, 2021

United States Census Bureau

Table of Contents

Chapter 1: Introduction

Introduction

The EEO Tabulation is a custom tabulation of data from the American Community Survey (2014-2018 5-Year Estimates) that is designed to measure the effects of and compliance with Equal Employment Opportunity (EEO) laws. It serves as the primary external benchmark for comparing the race, ethnicity, and sex composition of an organization's internal workforce, to the analogous external labor market, within a specified geography and occupation. The tabulation additionally provides data on other demographic and economic characteristics.

The tables include estimates and margins of error of the race, ethnicity and sex composition of the workforce for specified occupations and geographies. Occupation data are provided in Detailed Census Occupations, and in 4 aggregations— EEO Occupational Groups, EEO-1 Job Categories, Federal Sector Job Groups, and State and Local Government Job Groups. Other subject matter characteristics include U.S. citizenship, educational attainment, age, industry, earnings, and unemployment status. The EEO provides information at three geographic types - worksite, residence-to-worksite commuting flows, and residence.

The Equal Employment Opportunity (EEO) Tabulation is a custom product developed by a consortium of four Federal agencies, consisting of the Equal Employment Opportunity Commission (EEOC), the Department of Justice (DOJ), the Department of Labor (DOL), and the Office of Personnel Management (OPM), in conjunction with the U.S. Census Bureau. This tabulation was created according to the specifications of the agencies in the consortium. However, it contains information like comparable tabulations from the 1970, 1980, 1990, and 2000 censuses.

The EEO 2004-2018 (5-year ACS data) provides information at three geographic types- worksite, residence-to-worksite commuting flows, and residence. Tables are tabulated for one or more of the following geographic summary levels:

* U.S. Total
* All states and the District of Columbia, including Puerto Rico
* Core Based Statistical Areas (CBSAs)
* Counties
* County sets (counties with residence populations of 50,000 or more or aggregations of counties that together have a population of 50,000 or more)
* Places

Population thresholds of either 50,000 or 100,000 apply to all geographic areas, depending on specific table content. Population thresholds are based on residence population.

For more information about the EEO Tabulation, including details about content and geography, visit the Equal Employment Opportunity Tabulation Main Page on; https://www.census.gov/topics/employment/equal-employment-opportunity-tabulation.html

A custom EEO Data Tool was developed as a demonstration project to make the most used Census EEO data tables more accessible to end users while the Census Bureau creates an EEO retrieval tool on the new data dissemination platform. The EEO Tables Beta tool initial release is limited to the two most-accessed tables (1W and 1R) from the previous tool on American FactFinder. Based on user feedback, additional 2014-2018 and 2006-2010 5-Year EEO Tables will be added to the EEO Data Tool on an ongoing process. **The EEO Data Tool** is available on https://www.census.gov/acs/www/data/data-tables-and-tools/eeo-tables/#DemoProject. Please check this site for additional updates. All geographies from the previous data tool are available on the Census FTP site on https://www2.census.gov/EEO_2006_2010/

The American Community Survey

The American Community Survey (ACS) is a part of the U.S. Census Bureau's Decennial Census Program and is designed to provide more current demographic, social, economic, and housing estimates throughout the decade. The ACS provides information on more than 40 topics, including education, language ability, the foreign-born, marital status, migration and many more subjects. Each year the survey randomly samples around 3.5 million addresses and produces statistics that cover 1-year, 3-year, and 5-year periods for geographic areas in the United States and Puerto Rico. The 5-year estimates are available for many distinct geographies including the nation, all 50 states, DC, Puerto Rico, counties, places, census tracts, and block groups. For more information about the ACS, please visit our home page at: https://www.census.gov/programs-surveys/acs/.

ACS data are based on a sample and are subject to sampling variability. The degree of uncertainty for an estimate arising from sampling variability is represented using a margin of error. The value shown here is the 90 percent margin of error. The margin of error can be interpreted roughly as providing a 90 percent probability that the interval defined by the estimate minus the margin of error and the estimate plus the margin of error (the lower and upper confidence bounds) contains the true value. In addition to sampling variability, the ACS estimates are subject to nonsampling error (for a discussion of nonsampling variability, reference https://www2.census.gov/programs-surveys/acs/tech_docs/accuracy/MultiyearACSAccuracyofData2018.pdf for more information. The effect of nonsampling error is not represented in these tables.

Chapter 2: Geographic Overview of the EEO Tables

The EEO custom tabulation was produced for the following geographies, as residence and worksite geographies, using the following restrictions: Nation (010), State (040), County (050) {with populations greater than 50,000}, County (050) {with populations greater than 100,000}, Place (160) {with populations greater than 50,000}, Place (160) {with populations greater than 100,000}, Core Based Statistical Areas (CBSAs) (310) {with populations greater than 50,000}, Core Based Statistical Areas (CBSAs) (310) {with populations greater than 100,000}, County Sets (User Defined Area) {with populations greater than 50,000}, and County Sets (User Defined Area) {with populations greater than 100,000}. Geographic thresholds of 50,000 and 100,000 are applied to specific tables. Tables are produced for Places of Work (POW) and residence geography for all the geographies listed above. Population thresholds are always based on the residence population, even for tabulations at the worksite geography.

The American Community Survey 2014-2018 5-Year Special EEO File consists of tables for the United States, each of the 50 states and the District of Columbia, most Metropolitan Statistical Areas, counties, county sets, and places. Depending on the amount of detail within a table, information for counties, county sets, and places will only be provided for those with 50,000 or more and, in other instances, those with 100,000 or more. Counties with a population of less than 50,000 were combined with other contiguous counties within a state to create county sets with a population of 50,000 or more.

For the Residence Datasets on the American Community Survey 2014-2018 5-Year Special EEO File, data for individual counties of 50,000 or more is provided wherever possible. For counties of less than 50,000 in the Residence Datasets, the data from these counties were merged with other counties to meet a 50,000 or more threshold.

County Sets are a special EEO User Defined Area and will be defined by the US Census Bureau's Population Division. County Sets are collections of two or more counties that have a population threshold of 50,000 and 100,000 or more people. These areas do not overlap or contain silvers. They are constructed based on existing county definitions. County Sets are a user defined area specifically developed for the EEO Tabulation. They are aggregations of two or more counties, one of which is less than 50,000 population, so that the combined total population of the County Set is 50,000 or more and no county is shown with less than 50,000 population. No county set crosses state lines. All counties within any county set are within the same state. Counties are grouped within CBSAs. These are basically county-based groups that nest within states and must have residence population of 50,000 or more. These tabulations will all use county sets with a population threshold of 50,000 or more for consistency. Please reference the directory "**EEO_FTP_Site_Documentation"** and the sub-directory "**EEO_County_Sets_Reference_Maps**" on FTP site for more details.

County Commuter flows. Commuting flows consist of all the ACS Counties that meet the 50,000 County Set population threshold, for all workplace tables, i.e., county worksite total tables, county flows limited to three highest flows. All commuting flows must have at least a 50,000-population threshold and 50 unweighted commuters across all occupations.

This information below provides some additional detail concerning the creation of county sets by the US Census Bureau's Population Division. See URL for more details on EEO County Sets: https://www.eeoc.gov/eeoc/statistics/census/countysets.html.

Guidelines for Creating and Naming Aggregated County Areas for Counties Containing fewer than 50,000 Persons for Residence Datasets for the American Community Survey 2014-2018 5-Year Special EEO File

Residence Datasets on the American Community Survey 2014-2018 5-Year Special EEO File provide data only for counties of 50,000 or more. Counties of less than 50,000 were merged with other counties. These are the guidelines used to create and name these county sets.

General Rules for Creating and Naming County Sets

1. All counties are accounted for within each state and the total American Community Survey 2014-2018 5-Year population count within each state is correct.
2. All counties with population of 50,000 or more standalone unless they are aggregated with a county of less than population of 50,000 or with another county if needed to show a Metropolitan Statistical Area (MSA).
3. Groups of small counties were aggregated to create county sets of 50,000 or more.
4. The counties within each county set are contiguous.
5. No county set crosses state lines. All counties within any county set are within the same state.
6. Counties are grouped within MSAs if possible.
7. The counties and county sets are listed alphabetically by county set name.
8. The words "County Set" are not included in the county set name.
9. The counties and county sets are numbered consecutively starting with the state FIPS code followed by "CS" and the consecutive number of the alphabetized county sets starting at 001.
10. The state name is not included in the area name unless it is preceded by a directional description; in this case, the name is abbreviated. For example: "NW KS" (Northwest Kansas).
11. If the county set is composed of four or fewer counties, the county set name is the same as all the counties themselves in alphabetical order. For example: "Covington+Simpson+Smith."
12. For large groupings of small counties, the county set may be named using a general description that fits the part of the state in which it is found. For example, "NW Panhandle" (of Texas).
13. If a general geographic definition cannot be used, the most centrally located county is picked and "and surrounding cos." is added. If more than one county is centrally located, the county with the largest American Community Survey 2014-2018 5-Year Population Count is used. For example: "Gaines and surrounding cos."

    Note: In some instances, not all the counties included in the county set actually surround the named county. Some of these counties may form a "C." Or, some counties that are next to the named county in the county set name may not be included in the designated

county set. The county named in the county set name is picked based on the number of counties it is adjacent to within the county set designated geography. In some cases, the named county is the smallest, but touches the most other counties within the group. For example: In Iowa, "Audubon and surrounding cos." (19CS002) is made up of five counties and Audubon County is the smallest but is the most centrally located.

14. The Consortium discussed all county sets created by the Federal Agencies that caused disclosure problem with a MSA/PMSA. The Federal Agencies decided which MSA/PMSA to keep and which MSA/PMSA to eliminate. Most MSA/PMSAs are shown in the file; however, there are seven that were eliminated due to disclosure concerns. Less than three changes were made to each county sets for the MSA/PMSA to be shown. (In most cases only one change was required.)
15. Following are the seven MSA/PMSAs should not be shown on the EEO File based on Census 2000 Counts:

    Arizona: Flagstaff AZ-UT MSA - MSA FIPS Code #2620
    Arkansas: Fort Smith AR-OK MSA - MSA FIPS Code #2720
    Maryland: Baltimore MD PMSA - PMSA FIPS Code #0720
    New York: Newburgh NY-PA PMSA - PMSA FIPS Code #5660
    North Dakota: Grand Forks ND-MN MSA - MSA FIPS Code #2985
    Pennsylvania: Pittsburgh PA MSA - MSA FIPS Code #6280
    Wisconsin: La Crosse WI-MN MSA - MSA FIPS Code #3870

Syntax and Abbreviations

- Multiple names are separated by pluses in the name. For example: "Floyd+Hale+Swisher."
- If the county set title requires the state name, it is abbreviated and in all capitals. For example: "Northern AK" (Northern Alaska).
- Directional descriptions are abbreviated if they are nouns and written out in full if they are adjectives. For example: "SW GA" (Southwest Georgia) "Southern ID" (Southern Idaho).
- County/counties are abbreviated as "Co/cos." Except is abbreviated as "exc."

Other Notes

- In Alaska: Borough = County
- In Louisiana: Parish = County

Worksite geography presents data according to where people worked at the time of survey. These tables provide the number of people who were employed “at work,” that is, those who did any work at all during the reference week as paid employees, worked in their own business or profession, worked on their own farm, or worked 15 hours or more as unpaid workers on a family farm or in a family business in a given county or place.

Residence geography presents data according to where people lived, regardless of where they worked. These tables include people who were employed at work; employed but not at work, because they were temporarily absent due to illness, bad weather, industrial dispute, vacation, or other personal reasons; and the unemployed, who were actively looking for work in the last four weeks and available to start a job, whose last job was not a military-specific job.

The Following Geographic Areas are needed to tabulate the EEOC Custom Tabulation.

1. Nation (010) (already defined)
2. State (040) (already defined)
3. County (050) {with populations greater than 50,000}
4. County (050) {with populations greater than 100,000}
5. Place (160) {with populations greater than 50,000}
6. Place (160) {with populations greater than 100,000}
7. Core Based Statistical Areas (CBSAs) (310) {with populations greater than 50,000}
8. Core Based Statistical Areas (CBSAs) (310) {with populations greater than 100,000}
9. County Sets (902) – Definitions provided by Population Division
10. County Flows – Based on individual Counties in County Sets (902)

The Census Bureau refers to geographic areas as summary levels, which specifies the content and the hierarchical relationships of the geographic elements that are required to tabulate and summarize data. For example, summary level code 040 represents states, Washington D.C. and Puerto Rico, while summary level code 050 represents counties and county equivalents within states.

Chapter 3: EEO 2014-2018 Table Overview and List

The EEO Custom Tabulation was produced for all the table sets and geographic areas listed in the table below. Twelve sets of ACS subject tables are in .CSV files on the FTP site.

Table ID	Geography Type	Nation	States	Counties	County Sets	Places	CBSAs
EEO Table Set 1. Detailed Census Occupation Categories (236/237) by Race/Ethnicity (7), Sex (2)							
EEO-ALL01W	**Worksite**	X	X	X (50K+)		X (50K+)	X (50K+)
EEO-ALL01W	**Worksite Flow**			X(50K+)			
EEO-ALL01R	**Residence**	X	X		X (50K+)	X (50K+)	X (50K+)
EEO Table Set 2. Detailed Census Occupation Categories (236/237) by Race/Ethnicity (7), Sex (2), Citizenship (1)							
EEO-ALL02W, EEO-CIT02W	**Worksite**	X	X	X (100K+)		X (100K+)	X (100K+)

Table ID	Geography Type	Nation	States	Counties	County Sets	Places	CBSAs
EEO-ALL02W, EEO-CIT02W	**Worksite Flow**			X(50K+)			
EEO-ALL02R, EEO-CIT02R	**Residence**	X	X	X (100K+)		X (100K+)	X (100K+)
EEO Table Set 3. EEO Occupational Groups (14) by Race/Ethnicity (7), Sex (2), Citizenship (1)							
EEO-ALL03W, EEO-CIT03W	**Worksite**	X	X	X (50K+)		X (50K+)	X(50K+)
EEO-ALL03W, EEO-CIT03W	**Worksite Flow**			X(50K+)			
EEO-ALL03R, EEO-CIT03R	**Residence**	X	X	X(50K+)		X (50K+)	X(50K+)
EEO Table Set 4. EEO-1 Job Categories (9) by Race/Ethnicity (7), Sex (2)							
EEO-ALL04W	**Worksite**	X	X	X(50K+)		X (50K+)	X(50K+)
EEO-ALL04W	**Worksite Flow**			X(50K+)			
EEO-ALL04R	**Residence**	X	X	X(50K+)		X (50K+)	X(50K+)
EEO Table Set 5. Federal Sector Job Groups (9) by Race/Ethnicity (7), Sex (2), Citizenship (1)							
EEO-ALL05W, EEO-CIT05W	**Worksite**	X	X	X(50K+)		X (50K+)	X(50K+)
EEO-ALL05W, EEO-CIT05W	**Worksite Flow**			X(50K+)			
EEO-ALL05R, EEO-CIT05R	**Residence**	X	X	X(50K+)		X (50K+)	X(50K+)
EEO Table Set 6. State and Local Government Job Groups (8) by Race/Ethnicity (7), Sex (2), Citizenship (1)							
EEO-ALL06W, EEO-CIT06W	**Worksite**	X	X	X(50K+)		X (50K+)	X(50K+)
EEO-ALL06W, EEO-CIT06W	**Worksite Flow**			X(50K+)			
EEO-ALL06R, EEO-CIT06R	**Residence**	X	X	X(50K+)		X (50K+)	X(50K+)
EEO Table Set 7. Educational Attainment (5/6) by Younger Age Groups (6), Race/Ethnicity (7), Sex (2), Citizenship (1)							
EEO-ALL07R, EEO-CIT07R	**Worksite**	X	X	X(100K+)		X (50K+)	X(100K+)
EEO-ALL07W, EEO-CIT07W	**Worksite Flow**	X	X	X(50K+)		X (50K+)	X (100K+)

Table ID	Geography Type	Nation	States	Counties	County Sets	Places	CBSAs
EEO-ALL07W, EEO-CIT07W	**Residence**			X(100K+)			
EEO Table 8w. Detailed Census Occupation (236) by Educational Attainment (6), Race/Ethnicity (7), Sex (2), Citizenship (1)							
EEO-ALL08W, EEO-CIT08W	**Worksite**	X	X	X (100K+)		X (100K+)	X (100K+)
EEO Table 9w. Detailed Census Occupation (236) by Industry (20), Race/Ethnicity (7), Sex (2)							
EEO-ALL9W	**Worksite**	X	X	X (100K+)		X (100K+)	X (100K+)
EEO Table 10w. Detailed Census Occupation (236) by Older Age Groups (7), Race/Ethnicity (7), Sex (2)							
EEO-ALL10W	**Worksite**	X	X	X (100K+)		X (100K+)	X (100K+)
EEO Table 11w. Employment, Median and Mean Earnings for Detailed Census Occupation (236) by Race/Ethnicity (7), Sex (2) for Full-Time Year-Round Workers							
EEO-ALL11W	**Worksite**	X	X	X (100K+)		X (100K+)	X (100K+)
EEO Table 12w. Employment, Median and Mean Earnings for EEO-1 Job Categories (9) by Race/Ethnicity (7), Sex (2) for Full-Time Year-Round Workers							
EEO-ALL12W	**Worksite**	X	X	X (100K+)		X (100K+)	X (100K+)

Note: Each worksite table with commuter flows has a residence population threshold of 50,000 that both worksite and residence counties must meet. Each table will show a central worksite county that meets this population threshold with flows from individual counties that contribute commuters. Each table will show a central worksite that meets this population threshold with flows from up to three individual geographies that contribute commuters. One of these flows may be from the worksite itself, but each additional flow must meet the population threshold and contribute at least 50 unweighted commuters across all occupations.

Chapter 4: How to Use These Files

4.1 Locating the Custom EEO Tabulation on the FTP site

The EEO Tabulation is located on the Census FTP site at http://www2.census.gov/EEO_2004_2018/.

4.2 EEO Tabulation File Organization

The EEO Tabulation 2014-2018 (5-year ACS data) on the FTP site consists of separate .CSV files. The contents of these files, how to use the datasets, and a description of the variables and geography contained in the datasets are included in this documentation. The Custom EEO Tabulation on the FTP site are organized as shown below.

EEO_Tables_By_All_Areas – ***Coming soon***, for advanced users who need to download each table for all the EEO geographies for the purpose of pre-packaging the tables or loading them into a database, Commercial Off The Shelf (COTS) software, or a custom system. Files will be in directory below this one, no need for subdirectories. (Files will be Zipped)

EEO_Tables_By_Geographic Area – For data users who need one table for one set of geographies for all states. (Files will be in CSV format)

Files will be on subdirectories will be as follows;

- Nation
- State
- State_County
- State_Place
- Core_Based_Statistical_Areas
- County Sets
- County-Flows – ***Coming soon***

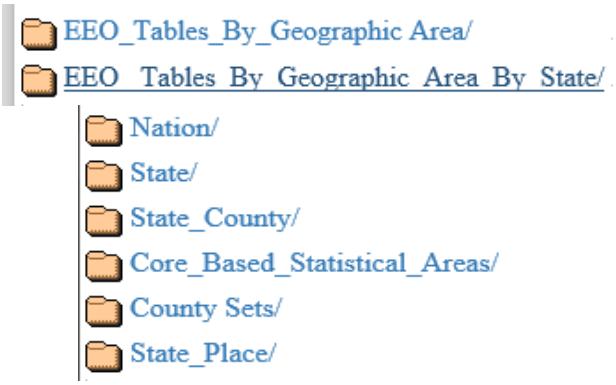
EEO_ Tables_By_Geographic_Area_By_State – For data users, States Governments, or Census State Data Centers, who only need specific EEO Tables for only a set of areas for a specific state. (Files will be in CSV format)

Files will be on subdirectories will be as follows (Each directory is divided by state).

- State (1560 files)
- State_County (1496 files)
- State_Place (1440 files)
- County Sets (52 files)
- County-Flows (614 files) – ***Coming soon***

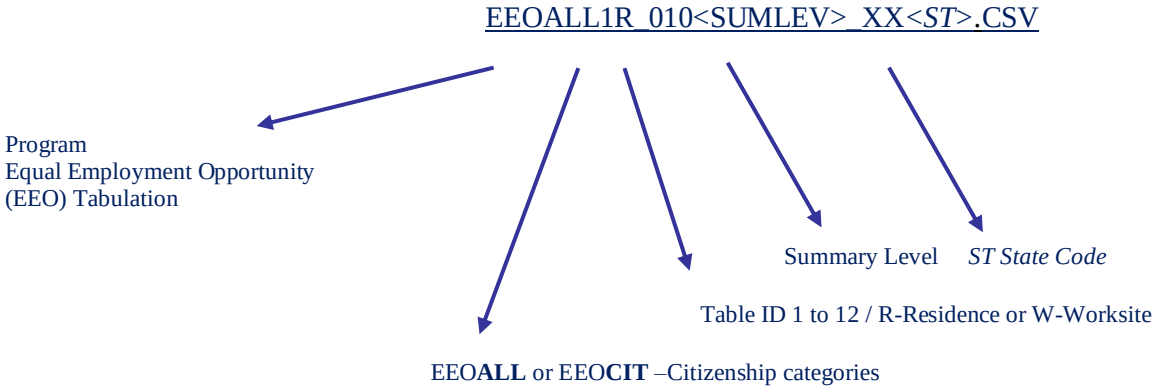
Please refence the **EEO_2018_FTP_Table_Lookup.xls** for more details concerning directory structure and file locations.

EEO_FTP_Site_Documentation – This directory will contain documentation for data users to reference.



4.3 Data Files

The naming convention used for the CSV files in this directory is the following:



File Name: EEOALL1R_010.CSV (Tables_By_Geographic_Area) or EEOALL1R_040_AK (Tables_By_Geographic_Area_By_State)		
Example	**Name**	**Range or Type**
EEO	Program	Equal Employment Opportunity (EEO) Tabulation.
ALL, CIT	Citizenship categories	Contains tables for categories of ALL and CIT for Citizenship categories
1	TBLID	1 to 12
R/W	Geography Type	R-Residence or W-Worksite
010, 040, 050, 310, 902 C21, P01, P06, P07, P10	SUMLEVEL	See description below of Summary Level
ST	State Abbreviation	Only CSV files in **EEO_ Tables_By_Geographic_ Area_By_State** will have a state abbreviation next to the SUMLEV, for example EEOALL1R_040_AK

4.4 Data Format and Access

Each data file contains a single EEO table in a comma-delimited (,) (CSV) ASCII file format. Each file contains the following variables or Columns; **OCC_CODE, OCC_DESCRIPTION,**

GEOID, GEONAME, TBLID, PROFLN, TITLE, ESTIMATE_1, MG_ERROR_1 .... ESTIMATE_N, MG_ERROR_N.

An Explanation of Variables in ACS Custom Tabulation Table CSV Files

Column Header or Variable Name	Description
OCC_CODE	Occupation Code. Described later in this document.
OCC_DESCRIPTION	Occupation Code Description. Described later in this document.
GEOID	This is the geographic ID number used within the Census Bureau. Described below in section **4.5 Geographic Identifiers.**
GEONAME	This is the geography name. Described below in section **4.5 Geographic Identifiers.**
TBLID	The table is assigned a table ID number. The tables you received were designed specifically for this tabulation and are unique See **Chapter 3: EEO 2014-2018 Table Overview and List** for more details
PROFLN	This identifies the line number within the table. The lines will be repeated for each geography.
TITLE	This identifies the line of data, usually as a total or a subtotal.
ESTIMATE_N	This is the estimate as a character value
MG_ERROR_N	This is the margin of error as a character value.

Example EEO CSV File Layout:

OCC_CODE	OCC_DESCRIPTION	GEOID	GEONAME	TBLID	PROFLN	TITLE	ESTIMATE_1	MG_ERROR_1
				EEOALL1R	**0.4**		**Total, race and ethnicity**	**Total, race and ethnicity**
0	**Total, all occupations**	**01000US**	**United States**	**EEOALL1R**	**0.9**	**Total, both sexes**		
0	**Total, all occupations**	**01000US**	**United States**	**EEOALL1R**	**1**	**Number**	**XXX,XXX**	**+/-XXX,XXX**
0	**Total, all occupations**	**01000US**	**United States**	**EEOALL1R**	**2**	**Percent**	**100**	**+/-0.1**

4.5 Geographic Identifiers

Every record (column) in each table .CSV file contains a unique geographic identifier (GEOID) and associated geographic area name GEONAME. All geographic codes for a geographic area

are embedded within the GEOID. The GEOID is made up of a three-digit summary level, a two-digit component (always "00" for EEO), a constant "US," and the unique geographic area code within the summary level. The summary level field (SUMLEVEL) is the critical element in identifying the geographic area type for each record. The GEONAME (Area Name) contains the name of each geographic area.

The GEOID is used to uniquely identifies each geographic area (and each record) in the file. The GEOID is formed by concatenating the following components: summary level code + geo-component code + US + applicable geocodes. Any blanks in the resulting string are replaced with zeros.

The SUMLEVEL represents a three-digit code, which represent the following geographic areas:

- Nation (010)
- Place of Work - Nation (C21)
- State (040)
- Place of Work - State (P06)
- State-County (050)
- Place of Work - State-County (P01)
- State-Place (160)
- Place of Work - State-Place (P07)
- Core Based Statistical Areas (CBSA) (310)
- Place of Work - Core Based Statistical Areas (P10)
- County Sets (902)
- County Flows (000)

Here are a few examples of GEOID's and their associated GEONAME's (Area Name):

01000US - United States
04000US01 - Alabama
05000US02020 - Anchorage Municipality, Alaska
16000US0103076 - Auburn city, Alabama
31000US11020 - Altoona, PA Metro Area
90200US05CS001- Arkansas (county)+Monroe+Phillips+Prairie

Files containing flow geographies will contain a GEOID followed by a FLOW ID (i.e. 0500000US01001, 0100100000), shown separately with a GEOID, then an accompanying FLOWID. These files will be separate from all the other geographic areas.

4.5 Occupation Codes

There is a total of 236 occupation categories that were produced for the place of work EEO Tables sets and 237 for the residence EEO Tables. Each detailed occupation that is shown for each subject table must have at least 10,000 national weighted population.

Example of EEO Occupation Codes:

2014-2018 Census EEO Tabulation Occupation Code	2018 SOC Code	EEO 2014-2018 Tabulation Occupation Code Descriptions	2018 Census PUMS Occupation Code
0010-3550	**11-0000 - 29-0000**	**Management, Business, Science, and Arts Occupations:**	**0010-3550**
0010-0930	***11-0000 - 13-0000***	***Management, Business, and Financial Occupations:***	***0010-0960***
0010-0440	11-0000	Management Occupations:	0010-0440

See EEO reference spreadsheet on the **EEO_FTP_Site_Documentation** directory for more details and a full list of EEO Occupation Codes.

Chapter 5: User Notes

Supplemental Documentation

Supplemental documentation concerning the American Community Survey, to assist users using this technical document, is located on the ACS Website at: www.census.gov/acs/www/data_documentation/documentation_main/.

Geographic Terms and Concepts

The most updated ACS geographic terms and concepts can be found at: https://www.census.gov/programs-surveys/acs/geography-acs.html
Reference the following document for more information; https://www.census.gov/programs-surveys/acs/geography-acs.html

Data Collection and Processing Procedures

The American Community Survey operations involve a complex set of data collection and processing procedures that are too extensive to discuss as a chapter in the EEO tabulation documentation. EEO data users interested in a technical discussion of ACS operational processes and survey design should reviewed on: https://www.census.gov/programs-surveys/acs/methodology.html

Maps and Geographic Reference Materials

Detailed information about the ACS maps can be found in the reference maps section of the ACS website located at: http://www.census.gov/acs/www/data_documentation/reference_maps/.

Table & Geography Changes

Learn more about changes to tables and geography for each American Community Survey tables and geography referenced on https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes.2018.html

Code Lists

The ACS provides code lists to identify all potential response categories for the variables included in the EEO Tabulation

Code Lists, Definitions, and Accuracy are available for reference on: https://www.census.gov/programs-surveys/acs/technical-documentation/code-lists.2018.html

Please note the 2018 Industry and Occupation Changes referenced on: https://www.census.gov/programs-surveys/acs/technical-documentation/user-notes/2019-03.html

Jam Values

Some data values represent unique situations where either the information to be conveyed is an explanation for the absence of data represented by a symbol in the data display, such as "(X).

The following list shows the special data values which can appear in any EEO Table on the FTP Site.

Special Data Values	FTP Site Display Value	Description
-999999999	N	Indicates that an estimate or its margin of error cannot be provided because the number of sample cases is too small for the given geographic area.
-888888888	(X)	Indicates that the estimate is not applicable or not available.
-666666666	-	Indicates that no sample observations were available to compute an estimate, or a ratio of medians cannot be calculated because one or both of the median estimates falls in the lowest interval or upper interval of an open-ended distribution.
-222222222	**	An '**' entry in the margin of error column indicates that either no sample observations or too few sample observations were available to compute a standard error and thus the margin of error. A statistical test is not appropriate.

Chapter 6: Disclosure Avoidance Review of Statistical Products

Title 13, Section 9 of the United States Code (U.S.C.) requires the Census Bureau to keep confidential the information collected from the public under the authority of Title 13. Section 214 of Title 13, U.S.C., and Sections 3551, 3559 and 3571 of Title 18, U.S.C., provide for the imposition of penalties of up to five years in prison and/or up to $250,000.00 in fines for wrongful disclosure of confidential census information.

Disclosure avoidance is the process for protecting the confidentiality of data, as required under Title 13 U.S.C. A disclosure of data occurs when someone can use published statistical information to identify an individual who has provided confidential information. For data tabulations, the Census Bureau uses disclosure avoidance procedures to modify or remove the characteristics that put confidential information at risk for disclosure. Although a published table may appear to show information about a specific individual, the Census Bureau has taken steps to disguise or suppress the original data, while making sure the results are still useful. The techniques used by the Census Bureau to protect confidentiality in tabulations vary, depending on the type of data.

Noise injection is the Census Bureau's preferred disclosure avoidance technique. By policy, noise infusion is applied to all data products that are reported with geographies smaller than a state. Noise injection may be required for microdata releases, depending on the characteristics of the microdata and the specific variables that are to be released. Data that cannot be publicly released may still be analyzed within the Federal Statistical Research Data Centers (FSRDCs) by individuals who have Special Sworn Status; the results of such analyses must still go through a disclosure avoidance process prior to being publicly released.

The parties understand that Title 13 confidentiality protection and disclosure avoidance techniques apply to all work described in this agreement. The disclosure avoidance methods are defined by the Census Bureau who has the responsibility of carrying out that work. Accordingly, upon completion of the tabulation, the data produced must be reviewed by the Census Bureau to ensure that no identifiable Title 13 data are or may be disclosed. Should the Census Bureau's Disclosure Review Board (DRB) determine that the requested statistical product does or reasonably could result in such disclosure, then the data product will be modified prior to approval for release to the party(ies) of this agreement. The DRB must approve before a research product can be released to an individual who does not have Special Sworn Status and a need to know or moved to a computer not approved for controlled data according to Census' existing policies and procedures.

The Census Bureau will not provide data made confidential by Title 13 in the 2014-2018 5-Year ACS EEO File. This tabulation was approved, on 3/26/2018, Decision #: CBDRB-FY18-225, as per the U.S. Census Bureau's Disclosure Review Board (DRB) on, shown in Attachment D-1 Spanos DRB Cover Sheet ST109 EEO- BC-3034 v3 0-DECISION. Pursuant to this DRB decision, the U.S. Census Bureau may only apply noise infusion and/or rounding methods that are not related to a method known as "differential privacy" to the confidential data under Title 13

in order to generate the special tabulations. Noise injection was not required to be applied to the 2014-2018 5-Year ACS EEO File.

Standard DRB rounding rules were applied to this tabulation.

1. All cells in any American Community Survey special tabulation must be rounded. The rounding schematic for all tables is:

- 0 remains 0
- 1-7 rounds to 4
- 8 or greater rounds to nearest multiple of 5 (i.e., 864 rounds to 865, 982 rounds to 980)
- Any number that already ends in 5 or 0 stays as is.

Any totals or subtotals needed should be constructed before rounding.

2. Percent's will be calculated after rounding, based on rounded estimates.

3. Medians are calculated as an interpolation from a frequency distribution of unrounded data (these are not subject to additional rounding).

APPENDIX A: ADDITIONAL REFERENCES

Equal Employment Opportunity Tabulation FAQ's
https://www.census.gov/topics/employment/equal-employment-opportunity-tabulation/about/faq.html

Guidelines for Creating and Naming Aggregated County Areas for Counties Containing fewer than 50,000 Persons for Residence Datasets for the Census 2000 Special EEO File
https://www.eeoc.gov/eeoc/statistics/census/naming.html

County Set Areas for Residence Datasets for the Census 2000 Special EEO File
https://www.eeoc.gov/eeoc/statistics/census/countysets.html

Information on ACS Sample Size and Data Quality can be found on:
https://www.census.gov/acs/www/methodology/sample-size-and-data-quality/.

The document "Understanding and Using American Community Survey Data: What All Data Users Need to Know" should provide you with more details concerning the ACS:
https://www.census.gov/programs-surveys/acs/guidance/handbooks/general.html.