

U.S. Equal Employment Opportunity Commission
Component 2 EEO-1 Upload File Validation Process for 2017 and 2018

ISSUED: September 9, 2019, Version 4

Document Overview

The EEOC Component 2 EEO-1 Compensation Data Collection offers employers two options to report their 2017 and 2018 pay and hours worked data through the *Component 2 EEO-1 Online File System* (<https://eeocomp2.norc.org>): (1) manually enter data into an online form (available since July 15, 2019), or (2) upload data via comma separated value (CSV) file (available in mid-August 2019). The *Component 2 EEO-1 Upload File Layout Specifications for 2017 and 2018* that provides specifications for developing Component 2 EEO-1 upload files were made available publicly on July 11, 2019. Employers may use these specifications to prepare Component 2 EEO-1 data upload files in advance of the upload system availability.

This document provides employers with information about how upload files will be validated in the *Component 2 EEO-1 Online File System*. The validation system component, entitled the *Upload Validator*, is integrated into the *Online Filing System* to ensure that all uploaded files conform to specifications. In the event there is an issue with an upload file, employers will receive immediate feedback through the filing system with the specific errors identified during validation. This document has two purposes: (1) it makes the validation process transparent to the employers who are creating upload files to guide their quality assurance review during upload file construction, and (2) it documents the error messages that upload filers may encounter during the filing process.

Overview of Upload Validation Process

At the start of the upload process in the *Component 2 EEO-1 Online File System*, the filer will confirm the upload file's name conforms to the specified upload file naming convention.¹ Once the file name is confirmed, the *Upload Validator* is designed to validate files uploaded in a series of three stages.

- (1) File Format Validation** - The *Upload Validator* will perform a file format validation to confirm the file is in fact a comma separated value (CSV) format file. If it is not, an error message will be generated asking the user to make a correction to the file's format.
- (2) Structural Validations** - Once the file passes the file format validation, the *Upload Validator* will perform a series of structural validations to ensure that the file conforms to the specifications laid out in the [Component 2 EEO-1 Upload File Layout Specifications](#). If it does not, an error message will be generated detailing the type of issue identified.
- (3) Content Validations** - Once the file passes the structure validations, the *Upload Validator* will perform a series of content validations to ensure that requirements provided in the [Component 2 EEO-1 Upload File Layout Specifications](#) and in this document are followed. If it does not, an error message or messages will be generated detailing the type of issue(s) identified.

Table 1 details the specific error messages that might be generated from the *Upload Validator* at any of the three stages.

Once an upload file passes all three stages of validation, the filer will be directed forward in the Online Filing System to review the Upload Summary, provide employer-level data² and certify the file. Figure 1 illustrates the Component 2 EEO-1 file upload submission and validation process.

¹ From the [Component 2 EEO-1 Upload File Layout Specifications for 2017 and 2018](#), the upload file naming convention is UploadFileyyyyCCCCCCC where the file name starts with the text "UploadFile" followed by "yyyy" which is the survey year, 2017 or 2018, and "CCCCCCC" which is the 8-digit User ID provided by NORC.

² Employer-level and certification data captured after successful file upload are detailed in Table 2 of the [Component 2 EEO-1 Upload File Layout Specifications for 2017 and 2018](#).

FIGURE 1. Component 2 EEO-1 Upload Submission and Validation Process

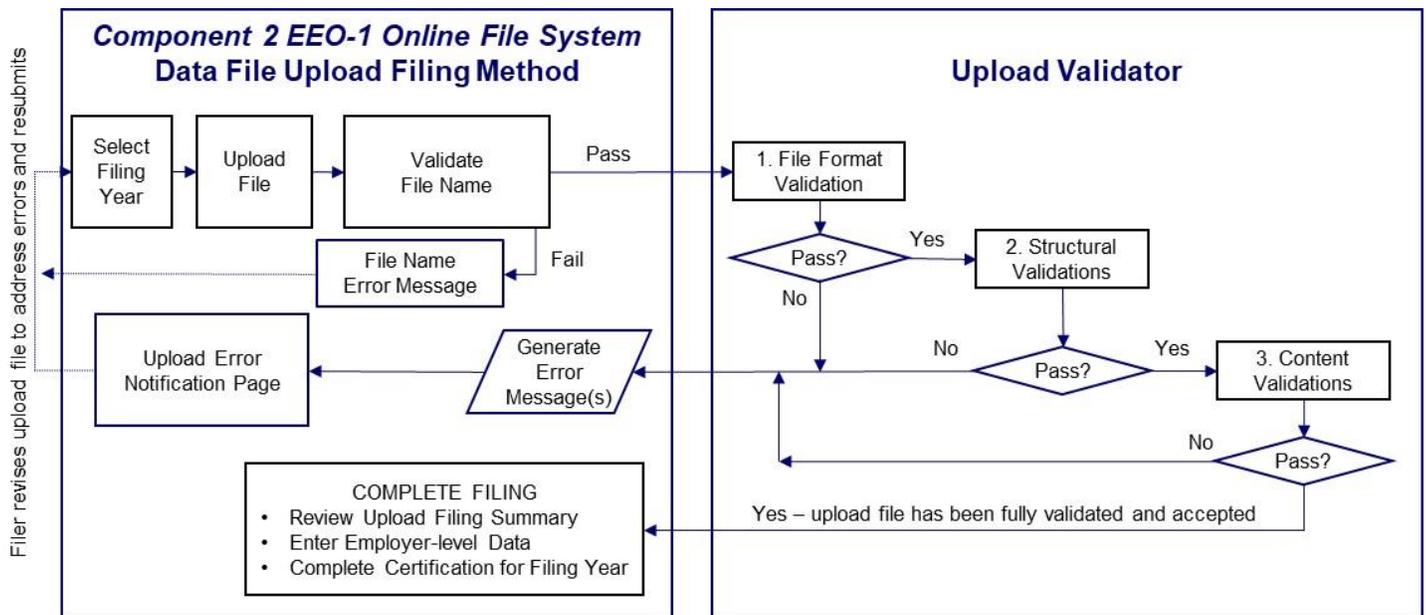


TABLE 1. EEOC Component 2 EEO-1 Upload Validations for 2017 and 2018 Upload Files

ERROR CODE	UPLOAD FILE COLUMN NUMBER	UPLOAD FILE COLUMN NAME (where applicable)	ERROR DESCRIPTION
0.0	NA	FILE VALIDATION	Error Code 0.0 - File Error: The Upload Validator was not able to recognize the uploaded file as a comma separated value file (CSV). Please ensure that the file uploaded is CSV formatted file.
0.1	NA	FILE VALIDATION	Error Code 0.1 - File Error: Empty rows within the file are not allowed and the file had one or more empty rows.
0.2	NA	FILE VALIDATION	Error Code 0.2 - File Error: Column headers in the first row are required and the file is missing column headers.
0.3	NA	FILE VALIDATION	Error Code 0.3 - File Error: The file does not have the expected columns. Please see the Component 2 EEO-1 Upload File Specification and ensure that every column specified is included in your file with field names that are identical to those in the specification.
0.4	NA	FILE VALIDATION	Error Code 0.4 - File Error: The file contains more commas than expected and therefore cannot be read properly by the <i>Upload Validator</i> . Please ensure that columns do not contain commas and that commas are only used as delimiters.
0.5	NA	FILE VALIDATION	Error Code 0.5 - File Error: The file contains an invalid combination of report types. <ol style="list-style-type: none"> When a filer is <u>NOT</u> filing with Status code 6 (i.e., Report Type 6 Small Establishment List Reports), do not include a Status Code 2 (i.e., Report 2 Consolidated Report) in the upload file. Only the following Status Code combinations are permissible and expected: <ul style="list-style-type: none"> 1, and only 1; 3 with at least one 4, 9, OR 8 report, where there can be any possible combination of 4, 9, and 8 reports; and 3 with at least one 6 report AND a single 2 report, where there can be any possible combination of 4, 9, and 6 reports. There cannot be both Status Code 6 and 8 in the records.
1.1	1	USERID	Error Code 1.1 - USERID column: Values are required for every row in this column and at least one is missing.
1.2	1	USERID	Error Code 1.2 - USERID column: USERID column must match the 8-digit User ID provided by NORC.
2.1	1	STATUSCODE	Error Code 2.1 - STATUSCODE column: Values are required for every row in this column and at least one is missing.
2.2	2	STATUSCODE	Error Code 2.2 - STATUSCODE column: An invalid Status Code has been reported. Only Status Codes of 1, 2, 3, 4, 6, 8 or 9 are valid.
4.1	4	UNITNAME	Error Code 4.1 - UNITNAME column: Values are required for every row in this column and at least one is missing.

ERROR CODE	UPLOAD FILE COLUMN NUMBER	UPLOAD FILE COLUMN NAME (where applicable)	ERROR DESCRIPTION
4.2	4	UNITNAME	Error Code 4.2 - UNITNAME column: UNITNAME (Establishment name) Establishment name cannot be less than 3 characters long.
4.3	4	UNITNAME	Error Code 4.3 - UNITNAME column: UNITNAME (Establishment name) Establishment name cannot begin with special characters (i.e., @\$%^, etc.)
4.4	4	UNITNAME	Error Code 4.4 - UNITNAME column: UNITNAME (Establishment name) The establishment name cannot be listed as "WORKING FROM HOME", "HOME OFFICE", "CORPORATE", "HEAD QUARTERS", or "HEADQUARTERS."
4.5	4	UNITNAME	Error Code 4.5 - UNITNAME column: UNITNAME (Establishment name) The establishment name cannot be the same as the address for that establishment. Please change the establishment name.
4.6	4	UNITNAME	Error Code 4.6 - UNITNAME and CITYNAME columns: The establishment name cannot be same as city name.
5.1	5	UNITADDRESS	Error Code 5.1 - UNITADDRESS column: Values are required for every row in this column and at least one is missing.
5.2	5	UNITADDRESS	Error Code 5.2 - UNITADDRESS column: UNITADDRESS is invalid. Address cannot be "HOME", "REMOTE", "HOME BASED", "TELECOMMUTE", "HOME ADDRESS", "WORK FROM HOME."
5.3	5	UNITADDRESS	Error Code 5.3 - UNITADDRESS column: UNITADDRESS Address cannot be less than 4 characters long.
7.1	7	CITY	Error Code 7.1 - CITY column: Values are required for every row in this column and at least one is missing.
8.1	8	STATE	Error Code 8.1 - STATE column: Values are required for every row in this column and at least one is missing.
8.2	8	STATE	Error Code 8.2 - STATE column: Invalid state code reported. Please see https://eeoccomp2.norc.org/additional-reference.html for a list of valid values, only the 50 states and DC are allowable values.
9.1	9	ZIPCODE	Error Code 9.1 - ZIPCODE column: Values are required for every row in this column and at least one is missing.
9.2	9	ZIPCODE	Error Code 9.2 - ZIPCODE column: Invalid ZIP code. Invalid ZIP code reported. Please see https://eeoccomp2.norc.org/additional-reference.html for a list of valid values. NOTE: Upload files constructed in Excel should format this field as text to ensure any leading zeros (0) are preserved when the file is converted to CSV format. 5-digit numeric values are expected in this field.
10.1	10	COUNTYNAME	Error Code 10.1 - COUNTYNAME column: Values are required for every row in this column and at least one is missing. Please see https://eeoccomp2.norc.org/additional-reference.html for a list of valid values.
11.1	11	FEIN	Error Code 11.1- FEIN column: Values are required for every row in this column and at least one is missing.
11.2	11	FEIN	Error Code 11.2- FEIN column: A Federal Employer Identification Number (FEIN) reported is invalid. <ul style="list-style-type: none"> • FEINs must be 9 characters long • FEINs must consist only of any combination of the following characters: 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. • FEINs cannot start with the following values combinations: "07", "08", "09", "17", "18", "19", "28", "29", "49", "69", "70", "78", "79", "89", "96", "97" • FEINs cannot be sequenced numbers such as "987654321" or "123456789" • FEINs cannot be the same number repeated nine times such as "00000000", "111111111", "222222222", "333333333", "444444444", "555555555", "666666666", "777777777", "888888888", or "999999999"
12.1	12	NAICSCODE	Error Code 12.1 - NAICSCODE column: Values are required for every row in this column and at least one is missing.
12.2	12	NAICSCODE	Error Code 12.2 - NAICSCODE column: Invalid NAICS code reported. Please see https://eeoccomp2.norc.org/additional-reference.html for a list of valid values or for Type 6 Reports use "999999" .
13.1	13	QUESTIONB2C	Error Code 13.1 - QUESTIONB2C column: Values are required for every row in this column and at least one is missing.
13.2	13	QUESTIONB2C	Error Code 13.2- QUESTIONB2C column: QUESTIONB2C has an invalid value. Question B2C can only be 1 (Yes) or 2 (No).
14.1	14	QUESTIOND2	Error Code 14.1 - QUESTIOND2 column: Values are required for every row in this column and at least one is missing.

ERROR CODE	UPLOAD FILE COLUMN NUMBER	UPLOAD FILE COLUMN NAME (where applicable)	ERROR DESCRIPTION
14.2	14	QUESTIOND2	Error Code 14.2 - QUESTIOND2 column: QUESTIOND2 has an invalid value. Question B2C can only be 1 (Yes) or 2 (No).
15.1	15	JOBCATEGORY	Error Code 15.1 - JOBCATEGORY column: Values are required for every row in this column and at least one is missing.
15.2	15	JOBCATEGORY	Error Code 15.2 - JOBCATEGORY column: JOBCATEGORY has an invalid value. Job categories can only be 1-10 <u>except for status type 6</u> which reports "99" in this column.
16.1	16	RACEETHNICITYGENDER	Error Code 16.1- RACEETHNICITYGENDER column: Values are required for every row in this column and at least one is missing.
16.2	16	RACEETHNICITYGENDER	Error Code 16.2 - RACEETHNICITYGENDER column: RACEETHNICITYGENDER has an invalid value. Race/Ethnicity/Gender categories can only be A-N <u>except for status type 6</u> which reports "Z" in this column.
17.1	17	ANNUALSALARY	Error Code 17.1 - ANNUALSALARY column: Values are required for every row in this column and at least one is missing.
17.2	17	ANNUALSALARY	Error Code 17.2 - ANNUALSALARY column: ANNUALSALARY has an invalid value. Annual Salary (Pay Band) categories can only be 1-12 <u>except for status type 6</u> which reports "99" in this column.
18.1	18	TOTALEMPLOYEES	Error Code 18.1 - TOTALEMPLOYEES column: Values are required for every row in this column and at least one is missing.
18.2	18	TOTALEMPLOYEES	Error Code 18.2 - TOTALEMPLOYEES column: TOTALEMPLOYEES There has is an invalid value in the column (<=0 or a character).
18.3	18	TOTALEMPLOYEES	Error Code 18.3 - TOTALEMPLOYEES column: Total number of reported employees is less than 50 and not status code type 8 or 6.
18.4	18	TOTALEMPLOYEES	Error Code 18.4 - TOTALEMPLOYEES column: Total employees reported for status code type 6 or 8 cannot be greater than 49.
18.5	18	TOTALEMPLOYEES	Error Code 18.5 - TOTALEMPLOYEES column: When there is a Status Code 6, the total summed number of employees from all Status Code 2 records must equal the total summed number of employees from all Status Code 3, 4, 9, and 6 records.
19.1	19	TOTALHOURS	Error Code 19.1 - TOTALHOURS column: Values are required for every row in this column and at least one is missing.
19.2	19	TOTALHOURS	Error Code 19.2 - TOTALHOURS column: The total number of hours cannot be negative or a non-number <u>except for status type 6, which reports "-3"</u> in this column.
19.3	19	TOTALHOURS	Error Code 19.3 - TOTALHOURS column: Total hours column cannot contain non-numbers.
99.99	NA	ALL COLUMNS	Error Code 99.99 - File Error. An unexpected error has occurred. Please check the upload file and try again.

NA = Not applicable.

APPENDIX – Guidance to Support Successful Filing for Component 2 EEO-1 Upload File Validation Process for 2017 and 2018

This appendix provides employers with additional guidance to support the construction of Component 2 EEO-1 upload files and prevent errors when filing in the *Component 2 EEO-1 Online Filing System*.

COMMON THEMES IN FILE VALIDATION ERRORS

Four out of every five upload file errors are the result of these six themes listed below.

ERROR TYPE	COMMON REASON FOR ERROR
File Construction	<ul style="list-style-type: none"> ▪ The upload file must be constructed exactly as designated by the validation specification document and upload specification document with 19 columns. ▪ Errors in file construction occur primarily due to missing column headers, extra spaces, or extra commas in the CSV (causing it to appear there are extra columns in the file).

DO carefully read the specification and upload documents if preparing file by hand, and include column headers exactly as specified.

DO NOT include extra spaces in the column headers or extra commas in the CSV.

ERROR TYPE	COMMON REASON FOR ERROR
Missing Values	<ul style="list-style-type: none"> ▪ Every column and row (excluding address 2) must have a value in it. Almost 2 out of every 5 errors are the result of missing a value in a row. ▪ There are many columns that have missing values, but over 1/3 of these errors are a result of missing hours in the column. ▪ If an employee was employed during the snapshot period, but worked no hours in the entire reporting year, 0 must be reported, not a blank.

DO ensure that every row has a value reported in the file.

DO NOT leave empty cells in rows where data should be reported.

ERROR TYPE	COMMON REASON FOR ERROR
Reporting Decimals and Not Whole Numbers (Integers)	<ul style="list-style-type: none"> ▪ For items that are numeric, such as total hours, decimals cannot be reported. For example, 10 is acceptable, but not 10.3.

DO ensure that numbers such as total hours and total employees are reported as integers, i.e., whole numbers.

DO NOT include any decimals in reported numbers.

ERROR TYPE	COMMON REASON FOR ERROR
ZIP Codes Truncated/ Not Five Digits	<ul style="list-style-type: none"> ▪ Some ZIP codes have leading zeros, and the leading zero is being dropped in file construction. ▪ If the upload file is being constructed in Excel, format the ZIP code field as text.

DO ensure all ZIP codes are five digits including leading zeros. If the ZIP code has leading zeros, ensure this column is not truncating these zeroes. See the [Additional Reference Documentation](#) page of the Component 2 EEO-1 website for list of valid ZIP codes. Do format this field as text when constructing the file in Excel to ensure any leading zeros are preserved when the file is converted to CSV format.

DO NOT format the ZIP code column as a number, or save the file after opening with the ZIP code column as a number.

ERROR TYPE	COMMON REASON FOR ERROR
Reporting Invalid Values	<ul style="list-style-type: none"> ▪ Some columns only allow certain values per the file layout specifications. Columns that only allow certain valid values include: STATUSCODE, STATE, NAICSCODE, QUESTIONB2C, QUESTIOND2, JOBCATEGORY, RACEETHNICITYGENDER, or ANNUALSALARY.

DO ensure columns have valid values according to the specifications. Use “1” and “2”, not “Yes” and “No” or “Y” and “N” in columns QUESTIONB2C and QUESTIOND2.

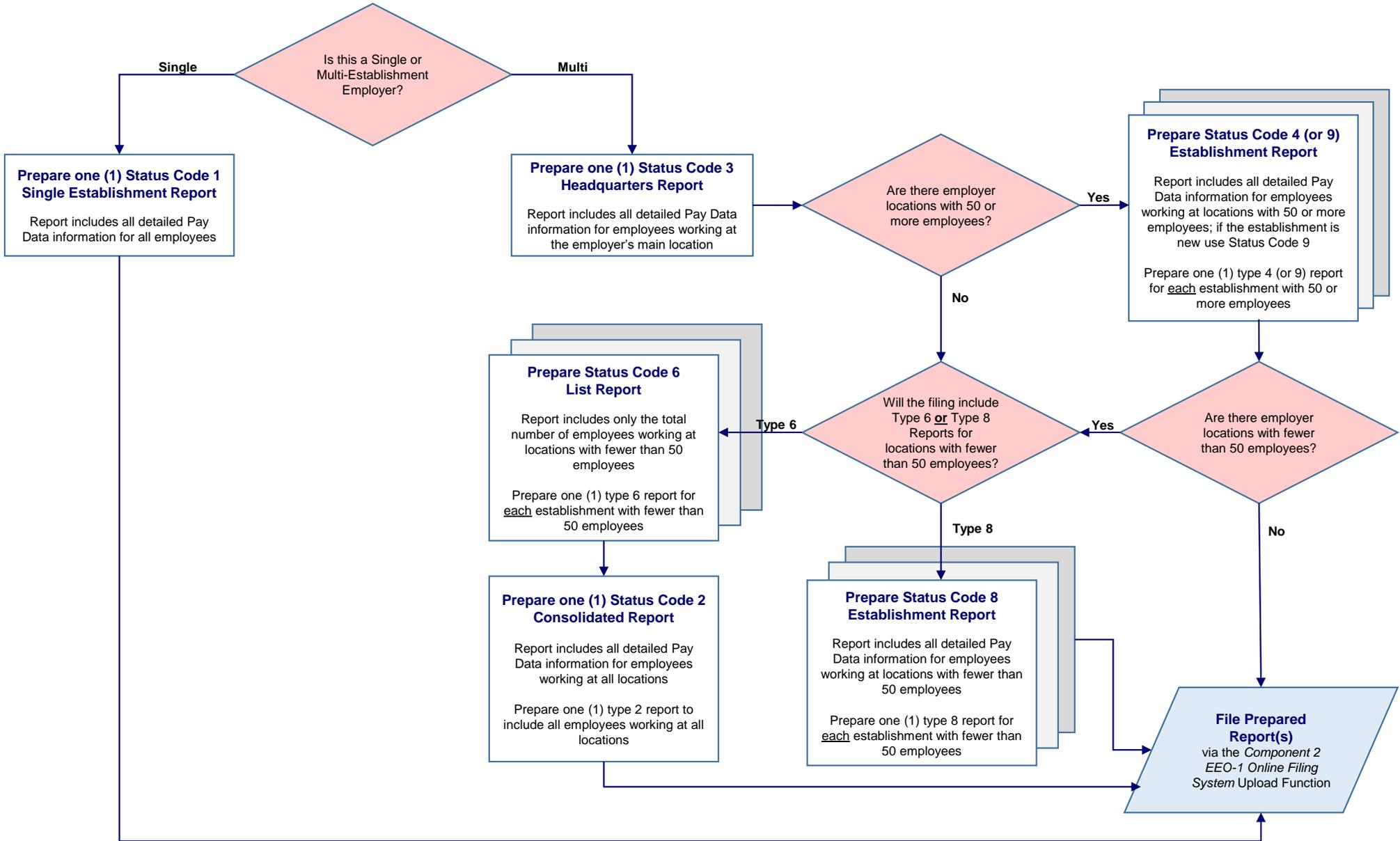
DO NOT report values that are not allowed in the file layout specifications.

ERROR TYPE	COMMON REASON FOR ERROR
Invalid Status Code Combinations Reported	<ul style="list-style-type: none"> ▪ Any number of invalid status codes combinations are reported. ▪ The most common issue is including the Status Code 2 (Consolidated Report) when it is not allowed (i.e., is not needed). ▪ The Consolidated Report is <u>only</u> included in the upload file for multi-establishment employers who are using Status Code 6 (List Reports) to report data for establishments with fewer than 50 employees.

DO compare the status codes reported to the required combinations in the specification document. **FIGURE 1A on APPENDIX Page 7 provides a decision tree to determine which types of Status Codes (i.e., report types) should be included in upload files.**

DO NOT report Status Code 2 when it is not required.

FIGURE 1A. Decision Tree to Determine Types of Status Codes to Include in Component 2 EEO-1 Upload Files



TOP 10 MOST COMMON ERRORS

Three out of every four upload file errors come from the top ten errors listed below.

ERROR CODE & DESCRIPTION	COMMON REASON FOR ERROR
0.2 or 0.3 File does not have the expected columns or column headers	<ul style="list-style-type: none"> ▪ One or more column headers is missing, or in a different order than specified. ▪ Extra spaces in column headings. ▪ Misspellings or added apostrophes. ▪ Commas in the CSV, causing it to appear that there more than the expected 19 columns in the file
0.5 or 2.2 Invalid Status Code Combinations	<ul style="list-style-type: none"> ▪ File includes invalid status code combinations. ▪ See FIGURE 1A on APPENDIX Page 3.
9.2 ZIP Code Invalid Values	<ul style="list-style-type: none"> ▪ File includes ZIP codes that are not five digits. Five digit numeric values are expected in this field. ▪ Leading zeros are not truncated.
18.3 Total Employees Less than 50 and Not 8, 6, or 1 reported	<ul style="list-style-type: none"> ▪ Reporting invalid status code for employee count. ▪ See FIGURE 1A on APPENDIX Page 3.
19.1 Total Hours Missing Values	<ul style="list-style-type: none"> ▪ Hours field is left blank. Values are required for every row in this column.
19.2 or 19.3 Total Hours Has Invalid Values (Non-Integers)	<ul style="list-style-type: none"> ▪ Numbers are written as decimals or fractions, or this field includes characters. Example: "10.3" reported instead of "10".
0.2 Column Headers Row Missing	<ul style="list-style-type: none"> ▪ Column headers not included in row 1 of file. Row 1 should not include any employer data, only the expected headers.
11.2 FEIN Invalid Values	<ul style="list-style-type: none"> ▪ FEIN value reported is not a valid FEIN or is not 9-digits.
12.2 NAICS Code Invalid Values	<ul style="list-style-type: none"> ▪ NAICS code must be a valid code from list on Additional Reference Documentation page of Component 2 EEO-1 website.
4.2 Unit Name Too Short	<ul style="list-style-type: none"> ▪ Unit name (establishment name) is less than 3 characters.

Document Version History

VERSION #	DATE	DESCRIPTION OF VERSION OR REVISION
1	7/30/2019	Specification was created and released to the public.
2	8/07/2019	Specification was revised with the following changes: <ul style="list-style-type: none">- Addition of Error Code 0.5 to specify valid combinations of report types.- Clarification of Error Code 9.2 to specify ZIP codes must be 5 characters long and valid.- Addition of Error Code 18.5 to specify expected sum of total employees in the case of Status Code 6 reports.
3	8/25/2019	Specification was revised with the following changes: <ul style="list-style-type: none">- User guidance was added in the Error Description column for Error Code 0.5 to describe the combination of Status Codes (i.e., report types) allowable in an upload file.- User guidance was added in the Error Description column for Error Code 9.2 (ZIPCODE) for users constructing upload files in Excel.- Removal of Error Code 10.2.
4	09/09/2019	Specification was revised with the following changes: <ul style="list-style-type: none">- Removal of Error Code 5.4 to allow addresses which start with alpha characters.- Addition of an appendix providing guidance on how to prevent or address common validation errors.