



GEORGIA DEPARTMENT OF EDUCATION (GADOE)

---

Office of Technology Services - Technology Management

# FY2019 Guidelines for Extracting Names

## GaDOE Guidelines for Extracting Names

When extracting names for any Georgia Department of Education data collection, use the following business rules:

### Legal Last Name

There should be some data cleansing on the Legal Last Name field, as well as parsing out of any standard suffixes.

- Parsed out suffixes should be placed in the Legal Name Suffix field if that field exists in the data collection. The following business rules should be applied: Check for the standard suffixes like *Jr, II, III, IV, V* and *VI*. These may be in upper case, lower case or mixed case. They must be preceded by a delimiter. The delimiter can be any special character or a space. The suffix may or may not be the last few characters in the field.
- Remove parentheses including the text between the parentheses.
- Remove all special characters other than hyphen, period, and single quote. Special characters to remove include commas, asterisks, double quote, underscore, and exclamation points.
- Do not remove a single quote (') used in a name or at the end of the name as an accent.
- Do not remove a hyphen or space. Do not remove whatever follows a hyphen or space unless one of the above rules also applies.
- Do not remove a period used in a name.

Using the above rules:

1. If the last name field contained *Smith, III (Trey)*, then the extract program would parse this into Legal Last Name of *Smith* and Legal Name Suffix of *III*. The text (Trey) is not parsed to any field. For the GTID application, the Suffix should be placed in the suffix field in the file layout. In other GaDOE data collections, such as FTE, SR, and Pre-ID, which have no suffix field in the layout, please remove the suffix from the last name and do not send the suffix in the file extract, drop it from the extract. This will allow the last names to match when GaDOE does error checking comparing GTID and FTE/SR/Pre-ID.
2. If the name field contained *Jones \*twin*, then the extract program would parse this into Legal Last Name of *Jones* and drop the *\*twin*.
3. Do not remove a single quote (') used in a name or at the end of the name as an accent, such as *La'Mont* or *O'Brien*.
4. Do not remove a hyphen - used in a name such as *Sanchez-Garcia* or *Smith – Jones*. Do not remove a space or whatever follows a space, such as *Sanchez Garcia*. Do remove other special characters that follow. For example, if the Last Name is *Smith Jones \*double last name*, remove *\*double last name* and extract *Smith Jones*.
5. Do not remove a period, such as *Jill St. John*.

## Legal First and Middle Name

- Remove parentheses including the text between the parentheses.
- Remove all special characters other than hyphen, period, and single quote. Special characters include commas, asterisks, and exclamation points.
- Do not remove a single quote (') used in a name or at the end of the name as an accent.
- Do not remove a hyphen or space. Do not remove whatever follows a hyphen or space unless one of the above rules also applies.
- Do not remove a period used in a name.

Using the above rules:

1. If the first name field contained *Elizabeth (Beth)*, then the extract program would parse this into Legal First Name of *Elizabeth* and drop the parentheses and contents of the parentheses.
2. If the first name field contained *Jacob \*custody issue*, then the extract program would parse this into Legal First Name of *Jacob* and drop the *\* custody issue*.
3. Do not remove a single quote (') used in a name or at the end of the name as an accent, such as *La'Shauna* or *Andre'*.
4. Do not remove a hyphen - used in a name such as *Rose-Marie*. Do not remove a space or whatever follows a space, such as *Mary Kate*. Do remove other special characters that follow. For example, if the first name is *Mary Jane \*MJ*, remove *\*MJ* and extract *Mary Jane*.
5. Do not remove a period, such as *Jill St. John Bell*.

## Student Name Suffix

The Student Name Suffix is used to distinguish a person who shares the same name within a family, such as Jr., II, III, IV. If the last name field contains *Smith, III (Trey)*, then the extract program would parse this into STUDENT LAST NAME of *Smith* and STUDENT NAME SUFFIX of *III*. The nickname (*Trey*) should not be exported for any data collection. Leave blank if blank on the birth certificate.

For suffixes that contain Roman numerals like 'III', the suffix must be the letter 'i' capitalized. A lower case 'L' (which may look like 'l') is not acceptable. The number one (which may look like '1' or '1') is not acceptable.

Only the following suffixes are acceptable:

JR	Jr.	Jr	JR.	II	III	IV	V	VI
VII	VIII	IX	X	XI	XII	XIII	XIV	XV
2nd	3rd	4th	5th	8th				
2ND	3RD	4TH	5TH	8TH				

The suffixes highlighted in yellow are acceptable in an upload but not found in online dropdown selections. The highlighted suffixes, if uploaded, will be transformed to compare more precisely with data in GaDOE databases.

If you have a student with the name suffix of 'VIII' or the '13<sup>th</sup>', contact the GUIDE Administrator to have the suffix added to the record.

The most important name processing guideline is – **the business rules used to process names in one data extract should apply to all data extracts.**