

## ID1 and ID2 in ERC data

There are two identifiers for matching:

ID1 - Only appears with TEA Data. The ID1 is a unique replacement for the data posted in the PEIMS ID (PID) field.  
ID2 - Appears in all data sets TEA and Non-TEA. This variable is a unique substitute for what is posted in the SSN field.

## PK-12 students

The basis for some of this ID reasoning is state law. Texas educates all children in PK-12 even if they are not citizens and do not have a SSN. That said ID1 is basically a state assigned id for PK-12 students (for the PEIMS data).

TEA, THECB, and TWC all check and crosswalk the original ids and the algorithm of the original ids.

ID2 is an algorithm of the SSN. You have to have a SSN for higher education and workforce data. In the past, people did not need to apply for a SSN until they wanted to work or go to college. Now, they encourage new parents to obtain a SSN at birth, but this is voluntary and even if they obtain one, it does not need to be shared at the PK-12 level.

For TEA, the crosswalk will contain the best unique {pid/id1} student record **by district**. Crosswalk variables include id2, id1, district, state\_assigned\_flag, invalid\_id1\_flag, yy, age, age\_mmdyyy.

The **yy** variable is the actual year where the best available crosswalk record was extracted.

The **age** variable is the student age on September 1 of the year the **yy** variable was extracted.

The **age\_mmdyyy** variable is the student age based on the crosswalk year.

The PEIMS {pid/id1} record is used to track a student through the twelfth grade. Each record includes the {ssn\_altid/id2} for student tracking when matching with non-PEIMS data and the {pid/id1} which is used to track a student within the TEA PEIMS data. Some issues have been raised over the consistency of identifiers for a single student as they are posted through the many school years. For instance, the same student may have several {state\_assigned\_ssn and valid SSNs/id2s} over a twelve year school history.

This Crosswalk provides the best hierarchical {ssn\_altid/id2} and {pid/id1}, **by district**, for all PEIMS records beginning 1992 through the current year. The crosswalk can then be matched by **ID1 and district** with all PEIMS data to assure the best consistent identifiers for each student.

### Here is why the TEA PID/ID1s are listed by District:

Matching the crosswalk with any individual PEIMS file by **{pid/id1} and district**, will assure the matched data is for the same student.

If the same {pid/id1} appears in more than one district, the student is probably the same, but there is a slightly greater chance that the student may not be the same. For these records the crosswalk will list a separate record for each {pid/id1} / district combination.

The crosswalk provides the best unique student/district level record available based on the following hierarchy:

- 1) If one or more year's student id1/district record includes a **valid** {ssn\_altid/id2} where {state\_assigned\_flag='0' and Invalid id1\_flag='0'}, then the most recent valid {ssn\_altid/id2} will be included on the crosswalk list.
- 2) Else, if there is not a single year where a student's id1/district records meet criteria #1 above, then the student record on the list will be the most recent year's record where the {ssn\_altid/id2} reflects a state assigned number {state\_assigned\_flag = '1' and invalid\_id1\_flag = '0'}.
- 3) Else, if there is not a single year where a student's id1/district records meet criteria #1 or #2 above, then if the {state\_assigned\_flag='0' and the invalid\_id1\_flag = '1'} these will be posted next in the hierarchy.
- 4) Else, if there is not a single year where a student's id1/district records meet criteria #1,#2 or #3 above, if the **{state\_assigned\_flag='1' and the invalid\_id1\_flag = '1'}** these will be posted last.

## Higher Education and Adults in TEA files

THECB follows the protocol below.

Identify Valid **ID2s** in **NEW** data set (recently released file to THECB for the ERC warehouse).

**TEA – Valid ID2 criteria (PEIMS/TAKS/TAAS)** - State\_Assigned\_Flag='0', Invalid\_ID1\_Flag='0', ID1 field is not blank,

**TWC** – ID2 is always valid unless the ID2 field is blank

**All Other Sources (THECB SBEC etc.) – Valid ID2 Criteria** - ID2 field is not blank, and Invalid SSN Flag = '0';

**Step 2 – Match valid ID2s with target data.** No criteria are needed for the target data. If the ID2 is valid in the **NEW** file, the ID2 that matches in the target data will be also be a valid ID2.

Generally, all students in TEA data sets have either a valid SSN or a State Assigned (state assigned flag='1'). Either way, there is a unique ID2 posted for both SSNs and State Assigned SSNs. Technically, the ID1s with a State Assigned Flag = 1 cannot be matched with non-TEA data. Since matching within TEA data uses the ID1, those ID2s where there is a State Assigned Flag='1' are problematic.

Each year there can be a unique record for each ID1/ID2 combination. Where there may just be two records where the ID1 matches more than one ID2 on one file, each of those unique ID1/ID2 combinations may actually show up over the years, possibly even once for each year. So each record of two ID1/ID2 combinations could potentially have many mismatches over time.

### Additional Matching TEA data with THECB Data:

Beginning in FY2009 the THECB Enrollment Report CBM001 includes the field TEA\_ID2. When this field is populated, it means the corresponding TEA data for this student included a State Assigned SSN/ID2 (state\_assigned\_flag = '1'). In these cases, the THECB SSN/ID2 variable will not match the TEA's State Assigned SSN/ID2. However, by matching the populated CBM001 TEA\_ID2 field to the TEA's ID2 field you will create a match. Although the TEA\_ID2 field only appears on the THECB CBM001, you can carry the TEA\_ID2 variable to other THECB data sets by matching on the ID2.

**Teacher IDs**

1. There is not an invalid\_ID1 field on the SBEC certification data files because TEA only keeps the certification records that match TEAs crosswalk. So there are no invalid ID1s.
2. There are no invalid SSNs or PID numbers in P.employ, and there are no state-assigned numbers given to staff.