



## Student Transcript and Academic Record Repository (STARR) Data Quality Technical Document

The purpose of this document is to provide the end user with specific business rules used in the creation of potential data quality reports that will be run by CEPI's data quality team during the current STARR collection. Please note that this document is not an exhaustive listing of all data quality checks. The listing below highlights the checks CEPI will be focusing on for the current STARR collection. Please refer to the CEPI [\(IHE\) Web page](#) for information and support documents regarding the STARR Collection.

**NOTE: CEPI may periodically update this document.**

Report Number	Report Title	Validation Rules	What to Look for?	Why?
ST1405	Session Designator after July of Year Being Reported	{AcademicSession}.[SessionDesignator] > 201X-07	Records where the {AcademicSession}.[SessionDesignator] is after July of the current year being reported.	A Session Designator cannot be a date greater than July of the current year being reported.
ST1406	Session Designator Reported with a Date more than 50 years ago	The year from {AcademicSession}.[SessionDesignator] + 50 years > current year	The sum of the year from {AcademicSession}.[SessionDesignator] and 50 years is after the current year being reported.	CEPI does not expect to see a lot of older records submitted in the STARR collection.
ST1419	IPEDS Code does not Align with the College Navigator**	{FileUpload}.[IPEDS] NOT = College Navigators IPEDS Code	{FileUpload}.[IPEDS] does not match the IPEDS Code Found on the College Navigator website.	You must submit the correct IPEDS code in STARR to accurately identify your college or university.
ST1420	Student's Entry Date changed from a previous STARR Collection	Prior{StudentDemographic}.[EntryDate] NOT = Current {StudentDemographic }.[EntryDate]	Records where the {StudentDemographic}. [EntryDate] differs from the previous year within your Institution.	Once a student has officially entered the IHE, the first entry date should never change. The entry date should remain constant for the student at that IHE.
ST1422*	No Students Reported as Dual Enrolled, Other School	{AcademicSession}.[HighSchoolStudentId] is missing or not reported	Records where no students were reported as Dual Enrolled, Other	It is important to report this field for MI School Data reporting

Report Number	Report Title	Validation Rules	What to Look for?	Why?
	Student, Concurrent Enrollment or Early Middle College		Enrollment, or Early Middle College using the {AcademicSession}.[HighschoolStudentId] field.	purposes as well as assisting colleges in submitting portions of the Section 209 and 245 Budget Transparency reporting when the IHE has students who participate in these programs.
<b>ST1429</b>	Academic Session Data not aligned with other Academic Session Elements	Year in {AcademicSession}.[SessionDesignator] is NOT contained in {AcademicSession}.[AcademicYearDesignator]	Records where the year in the {AcademicSession}.[SessionDesignator] is not one of the two years contained in {AcademicSession}.[AcademicYearDesignator].	Academic Sessions should reflect the academic year in which they occurred. The Academic Session data identifies the specific timeframe that the student attended a college.
<b>ST1502*</b>	Duplicate Course Records	Records where the distinct count of the Unique course is not equal to the count of unique course per UIC. Unique Course is defined as {AcademicSession}.[SessionDesignator], {Course}.[Number], {Course}.[CIPCode], {Course}.[Title], {Course}.[SubjectAbbreviation] and {Course}.[CreditAttempted]	The same unique course reported multiple times in one {AcademicSession}.[SessionDesignator].	A student course record should only be reported once in the same session. When all of the course information is the same in the session it suggests that the course is a duplicate.
<b>ST1503*</b>	Same Academic Sessions Reported Multiple Times, No Primary	Count of {AcademicSession}.[SessionDesignator] NOT = Count distinct of {AcademicSession}.[SessionDesignator] AND duplicated {AcademicSession}.[SessionDesignator] does NOT contain	{AcademicSession}.[SessionDesignator] is reported multiple times for the same student and it is not indicated which {AcademicSession} is primary in {AcademicSession}.[PrimaryAcademicLevelId].	When the same Academic Session is reported multiple times for the same student, you must also indicate which Academic Session was primary for MI School Data reporting purposes.

Report Number	Report Title	Validation Rules	What to Look for?	Why?
		{AcademicSession}.[PrimaryAcademicLevelId]		
<b>ST1505</b>	Counts of Military Status by IHE	COUNTS of {StudentDemographics}.[MilitaryStudentid] / Total UICs NOT in range of percentages as other IHEs	The percentage of students reported with {StudentDemographics}.[MilitaryStudentid] falls outside of a reasonable range.	Military Status is a new data element created to indicate the student's military status at the time of admission. Because this is a new data element CEPI is examining how colleges and universities are submitting it.
<b>ST1507</b>	Remedial Course Count Comparison	COUNT Distinct of {CourseCreditBasis}.[CourseCreditBasisid] by {CreditBasis}.[CreditBasisDesc] is within the same range as the previous STARR submission	The counts of remedial courses by category should be within the same range as the previous STARR submission.	This may indicate under or over-reporting of remedial courses in the STARR collection.
<b>ST1516*</b>	First time College Students Reported, Found in Previous Collection	Count of FirstTime > 1 per UIC AND {AcademicSession}.[PostSecondaryEnrollmentTypeId] in prior STARR NOT = {AcademicSession}.[PostSecondaryEnrollmentTypeId] in Current STARR	Records where a student is reported as FirstTime in {AcademicSession}.[PostSecondaryEnrollmentTypeId] and data are found from the previous academic sessions within the same collection for that UIC labeled as first time. This analysis is also executed longitudinally between the current and previous collection.	Student can only be reported as FirstTime once within an IHE.
<b>ST1517</b>	Degree CIP Codes do not align with US Ed's Degree CIP Code List	{DegreeCIPCode}.[DegreeCIPCode] NOT = NCES allowed CIP Codes^	{DegreeCIPCode}.[DegreeCIPCode] does not match NCES allowed CIP Codes. This check is performed at	It is important to report this field for MI School Data reporting purposes as we start to report more

Report Number	Report Title	Validation Rules	What to Look for?	Why?
			the IHE level.	on award data.
<b>ST1518</b>	Identifying Course CIP Codes submitted by Colleges and Universities	<i>Check 1:</i> {Course}.[CIPCode] NOT = NCES allowed CIP Codes <sup>^</sup> <i>Check 2:</i> Count of {Course}.[CIPCode] = 0	<i>Check 1:</i> Evaluate {Course}.[CIPCode] submitted by IHEs and compare them to NCES CIP Codes. Identify invalid CIP Codes. <i>Check 2:</i> Flag IHEs that have not submitted any {Course}.[CIPCode] This check is performed at the IHE level.	In the future, CEPI will start identifying Course Subjects by the Course CIP Codes. This is an initial check to identify a pattern of IHEs submitting Course CIP Codes.
<b>ST1519</b>	Academic Award Date in Future	{AcademicAward}.[AcademicAwardDate] > '01JUN' of Current STARR Submission	The {AcademicAward}.[AcademicAwardDate] is after June of the current STARR submission. This check is performed at the UIC level.	Award records should only reflect the prior years' or the current STARR submission.

Notes:

There are three windows in which STARR data quality will be run: pre-collection, mid-collection and post-collection.

- Pre-collection data quality analyzes prior year's data.
  - For example: The pre-collection for STARR 2015 will analyze data from STARR 2014.
  - If data are being compared to a previous year during the pre-collection window, the comparison will be the prior STARR collection to the STARR collection that occurred before it.
- Mid-collection data quality analyzes the current data that are being submitted up to the snapshot date.
  - The snapshot date will be provided in the applicable data quality email(s).
  - If data are being compared to a previous year during the mid-collection window, the comparison will be the current STARR collection to the prior STARR collection.
- Post-collection data quality analyzes the current data after the STARR collection has been closed.
  - The comparisons utilize the current STARR Collection and/or prior STARR collection, similar to the mid-collection timeframe.

\* These issues are NOT run during the pre-collection window.

\*\* College Navigator's IPEDs code is located at: <https://nces.ed.gov/collegenavigator/>

<sup>^</sup> NCES CIP Codes are located at: <http://nces.ed.gov/ipeds/cipcode/browse.aspx?y=55>