**Research & Data Integrity Meeting Notes  
November 23th, 2016   
9:00am**

|  |  |  |
| --- | --- | --- |
| 1. | dbo\_Student\_Test table :  My concern is about lots of missing values for students test score on this table. (98,064 Cases), as well as many missing data on the Test\_Date filed. (114,951 Cases)  In addition, I would like to know the role of the snapshot on this table, while we have test date.   * Cristina   + the counts of cases with null dates AND non-null scores of valid Test\_Types (CTEEP, CTERP, MDTP\*P 1-4 and TELD) have dropped from about 8000 cases a snapshot in 2009 to a few hundred in 2016 - is this due to a better data entering practice at the colleges? Or a better way of capturing the data?  Does Colleague allow Test\_Scores to be entered without a Test\_Date? And vice-versa?      * There are 5923 cases of null Test\_Date and Test\_Scores (see query below)     SELECT DISTINCT  ID\_Number,  Snapshot,  Test\_Type    FROM Student\_Test  WHERE School='CC' AND  Test\_Type IN ('TELDP','CTEEP','MDTP1P','MDTP2P','MDTP3P','MDTP4P','CTERP') AND  Test\_Date IS NULL AND  Test\_Score IS NULL     * Couldn't see a specific pattern other than the fact that in more recent semesters (2016), the count of null Test\_Date and Test\_Scores got reduced to a few cases a snapshot.     SELECT DISTINCT  Snapshot, COUNT (\*) AS Expr1, Test\_Type    FROM Student\_Test    WHERE School='CC' AND Test\_Type IN ('TELDP','CTEEP','MDTP1P','MDTP2P','MDTP3P','MDTP4P','CTERP') AND  Test\_Date IS NULL AND  Test\_Score IS NULL  GROUP BY Snapshot, Test\_Type  ORDER BY Snapshot ASC     * It would be great if we could add a test-location identifier like we have for CCTD (Chemistry test) CHEMP - test was taken at SAC and CHEM - test was taken at SCC. * Below is sample query of how to efficiently extract valid tests for English/Math tests for students taking from 2/1/2016 (that's the earliest students can take a test to be used to enroll in Fall 2016) to 11/01/2016.    Need to add STNC.INSTITUTION to STUDENT\_TEST table.    SELECT DISTINCT  School,  ID\_Number,  Test\_Type,  Test\_Date,  Test\_Score    FROM  Student\_Test    WHERE  School = 'CC' AND  Test\_Type IN ('TELDP', 'CTEEP','CTERP','MDTP1P', 'MDTP2P', 'MDTP3P', 'MDTP4P') AND  Test\_Score IS NOT NULL AND  Test\_Date <= CONVERT(DATETIME, '2016-11-01 00:00:00', 102) AND                        Test\_Date >= CONVERT(DATETIME, '2016-02-01 00:00:00', 102)                        ORDER BY ID\_Number, Test\_Type ASC  Stuart to follow up with Paul to determine what is the criteria used to pull test scores and test dates. DW has many with NULL values although Colleague has data stored on TSUM. | Haydeh/ Cristina |
| 2. | Missing SSN Report | Nga |
| 3. | Continuing Education Application Discussion | Stuart |
| 4. | Add “CCC-ID” to the DW  STUDENT\_STATIC add new Column for Alternate ID for only CCID, keep existing Alternate ID column for CPERM and/or NPERM. Are other codes stored here, ITS to research and provide data to team. | Rudy |
| 5. | * Old item from October 13, 2015: add internal course id to the Course\_Catalog and Course\_Sections tables to remove duplication   Add Internal Course ID to Course\_Catalog & Course\_Sections table   * What is the difference between Academic\_Program and Calculated\_Major? Is Calculated\_Major the first position of Academic\_Program? It looks like it is for some students but not all. What are the data element sources for these fields and how are they entered/changed in Colleague?   How is calculated major pulled into DW? Is it position driven or date driven? Stuart to follow up. | Cristina |
| 6. | Daily Snapshots Discussion  Separate hard drives vs one large hard drive: Possible restoration issues if all stored on a single hard drive. Separate drives reduces risk of data loss however you would have to jump databases to match data across drives.  Which tables should we copy daily? (Research team to discuss)  Separate database vs. inside Research DW (Research team to discuss)  Stuart & Paul to estimate costs based on current size of snapshots  Possible piecemeal costs if Blade is required | Nga |
|  | **Brought Forward: Follow up with Paul to provide status**  Unknown Race Discussion  In 2015-16, both CEC and OEC “Unknown Ethnicity” figure experienced a significant jump from about 4% and 10%, respectively in 2014-15 to 25% and 45%, respectively. Additional data explaining what may have caused this sudden jump will be presented tomorrow. This will not be rectified until new CE application is completed  There are approximately 4500 OEC students in 2015-16 with records in Student\_Seat and Student\_Static but no corresponding records in the Student\_EthnicRaces table. This will not be rectified until new CE application is completed  Major differences between custom race table in Colleague for CE vs. Colleague base table. Stuart is working with Phuong and Pat at CEC and OEC to resolve race issues. This is currently affecting all reporting for CE: MIS, IPEDS, Internal and Research. Phuong and Pat to go back and talk with Jim & Jose along with Yuri Betancourt to resolve differences. This will not be rectified until new CE application is completed  Data errors stem from custom drop down menus on RGPE that do not match the College Credit Colleague delivered drop downs. Paul identified that the Data Warehouse is not pulling in the XCED.ETHNIC.BACKGROUND at this time. ITS can pull this data into DW now, however this will be a temporary solution only for the DW as it will not address the MIS, IPEDS, ARCC Scorecard issues. This will not be rectified until new CE application is completed |  |
|  | Home Campus: designating the college campus to which a student belongs (based on application). This data element used to exist in the Student\_Session table, but it has been removed at some point in the past 2-5 years(?).  Need to add this element back into the DW as an element. It does not currently exist there now. |  |
|  | Student Placement Table: Want to add Reading Dept. Test Scores, that Stuart provides on an as needed basis. Stuart can provide Paul the query he uses.  Stuart just sent this to Paul on 9/1/2016, Paul to follow up |  |
|  | Prior Units:  CC: Looked into OpenCCCApply Application to download Prior Units as part of the import process. However, at this time that data is unavailable. Will continue to look into downloading this data along with Foster Youth, Self-Reported Veterans and Parent Guardian Educational Levels, General Financial Assistance, US Military, US Dependent of Military, Adults in Foster Care or have aged out of the system.  CE: Need to see if we can add Highest Degree Completed and Highest Diploma/Degree Earned from RGPE. This will not be rectified until new CE application is completed |  |
|  | Elements to pull from RGPE:  Also pull Labor Force, Annual Household Income, Number in Household, Personal Status, Primary Goal & Secondary Goal for all CE students. This will not be rectified until new CE application is completed |  |
|  | Student Session Semester Units GPA Fields and Cum. Units GPA:  Need to be calculated and populated back to Fall 2009.  Sem\_Units\_GPA & Cum\_Units\_GPA in the Data Warehouse.  Need to see if there is a Sem\_GPA and Cum\_GPA that needs to be historically updated as well.  These fields are being populated now for all current runs. This is coming from RSCCD\_GPU from Albert’s modified transcript. |  |

Attendees: Stuart Davis, Cristina Gheorghe, Janice Love, Haydeh Kaveh, Nga Pham, Aaron Voelcker, Janet Enriquez, Annette Wickman

**Next Meeting:   
January 5, 2016  
Training Room, DO-251**