THE UNIVERSITY OF NORTH CAROLINA GENERAL ADMINISTRATION

STUDENT DATA MART

User Guide – 5.0.0

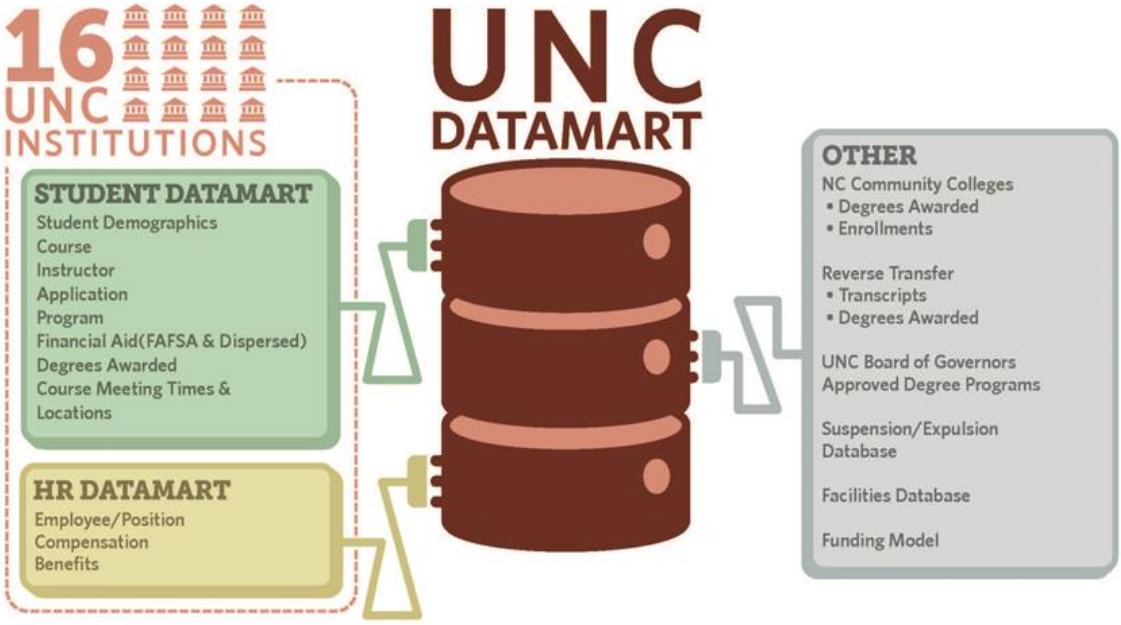
The purpose of this manual is to provide instructions for navigating Student Data Mart.

Contents

1 Introduction

1.1 What is the Student Data Mart?

The Student Data Mart (SDM) is a data warehouse of UNC system data. Effective with fall 2016 reporting, the SDM fully replaced the file-based data system in place since 1972. Data on students, applications, courses, degrees, programs, financial aid and faculty are included.



1.2 Who uses the Student Data Mart?

All 16 UNC system campuses supply data to the SDM via the student information systems, Banner or Peoplesoft. Campus operational offices such as Institutional Research, Registrar, Admissions, Enrollment Management and Financial Aid log in to verify data integrity. Consumers of the SDM data and reports include the same campus users as well as UNC General Administration (GA) offices such as Data & Analytics and Information Technology.

Login access to SDM application is restricted to UNC System approved users. The SDM Dictionary is available to anyone with or without login access. The products of the SDM data are presented in publicly accessible mediums such as the UNC Info Center without a login.

2 Getting Started

2.1 Responsibility for Student Data

Before accessing the SDM, users must understand the responsibility of working with student data. Student data is protected minimally by the Family Educational Rights and Privacy Act (FERPA) and may have additional protections applied by the institution. All users must meet the institution - campus or GA - requirements for secure data access and FERPA qualification.

2.2 Terminology

The Glossary contains more terminology. Helpful terms to know getting started include:

- *Snapshot* - A collection of data feeds. Also called a submission.
- *Data feed* - A collection of data elements. Data feeds are defined by subject area and are the basic elements sent from the campus ERP (Banner or Peoplesoft) to the data mart. Data feed elements are found in the SDM technical documentation, data feeds document.
- *Data set* - A collection of data elements that have been joined from one or many data feeds. Data sets are produced after a successful submission and include the basic element (usually a code) as well as the descriptive value of the data element and derived elements. Data set elements are found in the dictionary.

2.3 Requesting a Login

Each campus and GA has a point of contact charged with approving access to the SDM. New users must request access by contacting their campus official. Individual campuses may implement specific access policies. Minimally by FERPA requirements, the request must include an explanation of why access to student data is relevant to the job.

2.3.1 Production

In addition to making the request to the campus official, the user must register their ID in the SDM by attempting a login to https://uncdm.northcarolina.edu. The attempt will fail, but will register the user ID in SDM so that the campus official can add permissions to the user account.

2.3.2 Development (QA)

Campus users involved in testing new releases or code changes will require access to the development (QA) environment. The QA SDM usually points to a development instance of the campus ERP for the purpose of testing or quality checking before moving to the production instance. Access is granted separately from Production. Users must first attempt the login at https://devuncdm.northcarolina.edu/QA.

3 Logging In

The Student Data Mart is accessed via the Web at https://uncdm.northcarolina.edu in production or https://devuncdm.northcarolina.edu/QA for development (QA).

On the first log in, you will be prompted to select your campus. SDM retains your campus affiliation on subsequent log-ins until your internet browser cache is cleared.

3.1 User ID and Password

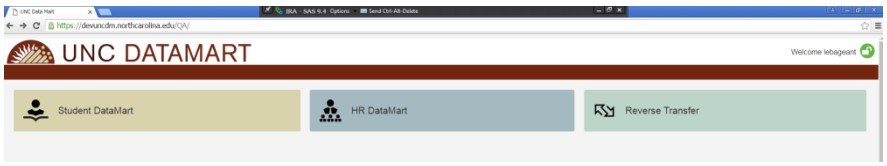
Never save your login information (user ID and/or password). Most internet browsers (e.g. Chrome, Internet Explorer, Safari, Mozilla) prompt, asking if you want to save your login information. Never do this for SDM. Forcing the entry of your user ID and password on login reduces the risk of any unauthorized user accessing SDM from your computer.

The user ID and password are the same as your campus network user ID and password. SDM uses single sign on (SSO) federated identity via Shibboleth.

3.2 Home Page

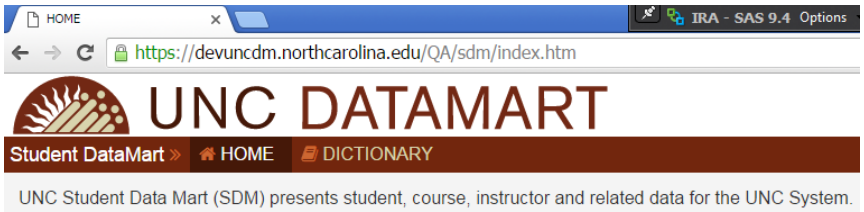
After a successful login, the UNC Data Mart home page appears. The Home Page is shared by all UNC Data Marts. This includes Student, Human Resources and Reverse Transfer. Student Data Mart will advance to the SDM Main Menu.

To confirm a successful login, look to the upper right corner. Instead of “Login”, “Welcome UserID” will appear after successful login.

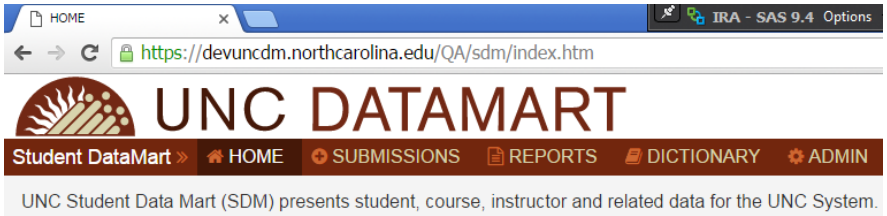


3.3 SDM Main Menu

If you have not yet been granted SDM permissions, then you will see the icons that are publicly available. These include the Home Page and the Dictionary.



If you have been granted SDM permissions, then you will see icons for all provisioned permissions.



3.4 Logging Out

Log out by clicking the lock in the upper right corner. Always log out and close the browser window when done working in SDM.

If leaving the computer for a short time, remember to lock your workstation.

Ctrl+Alt+Delete

Before You Leave Your Seat

Always lock your workstation before walking away from your computer.

4 Submissions (Snapshots)

The submissions page is where you'll spend most of your days, checking out warnings, errors, looking for data outside of the norm. ♪*adapted from WS&DJJJ*

The page displays the list and details of all snapshots. The main page may be used to see the upcoming schedule for planning purposes or it may be used to search for a particular snapshot to examine in detail.

4.1 Submissions - View the Schedule of Snapshots

By default, the list of snapshots is sorted by upcoming due date which provides the schedule of open or future snapshots.

4.1.1 Open Snapshots

Open snapshots are currently running as part of the automatic overnight processes.

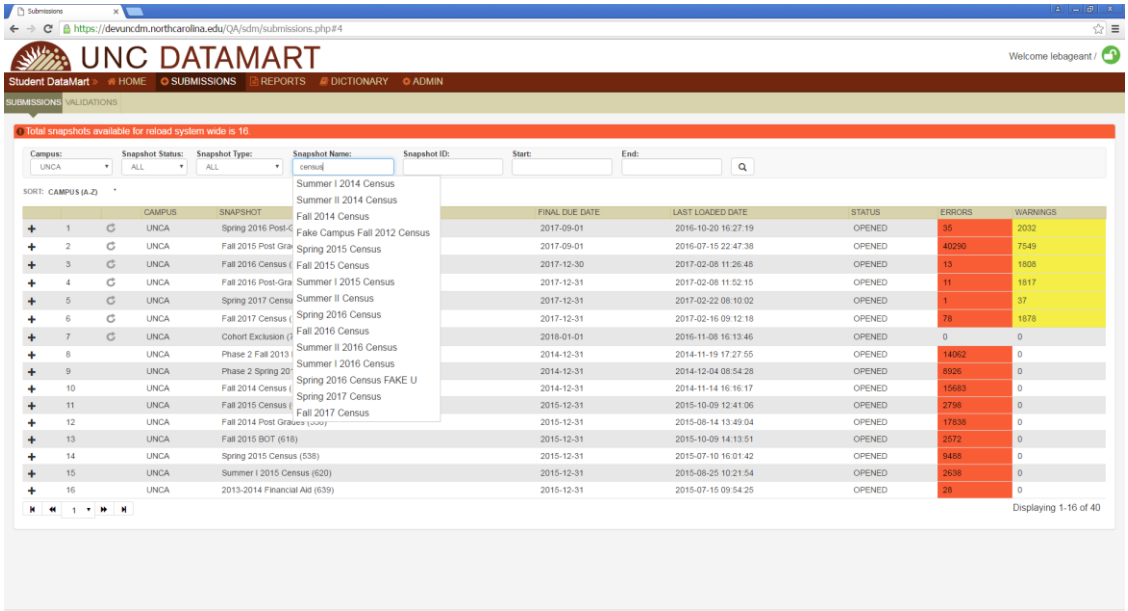
4.1.2 Future Snapshots

Future snapshots have not started running. HINT: Upon request, some future snapshots can be opened earlier than scheduled.

4.2 Submissions - Sort and Search Snapshots

By default, the list of snapshots is sorted by upcoming due date. The list can be sorted in ascending or descending order by any of the labeled columns:

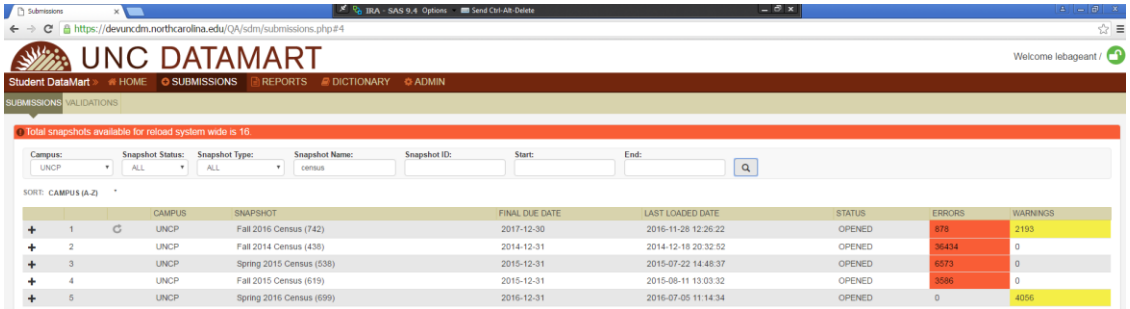
Campus	Campus users will only see their own campus.
Snapshot	The descriptive name of a snapshot. Snapshots are generally named by time frame (e.g. Fall 2016 or 2015-2016) and snapshot type (e.g. census or Financial Aid).
Snapshot ID	The database unique identifier for a snapshot.
Snapshot Type	The snapshot type determines which data feeds and data validations apply.
Initial Load Date	The earliest date that the snapshot was run.
Final Due Date	The date the snapshot will run for the final time
Last Loaded Date	The most recent date and time the snapshot was run.
Status	The snapshot status reflects the snapshot status as a whole as opposed to an individual campus status
Errors	The count of critical errors in the snapshot. Critical errors must be resolved before the final due date.
Warnings	The count of warnings in the snapshot. Warnings must be reviewed prior to the final due date, but may reflect valid data and not require any corrections.



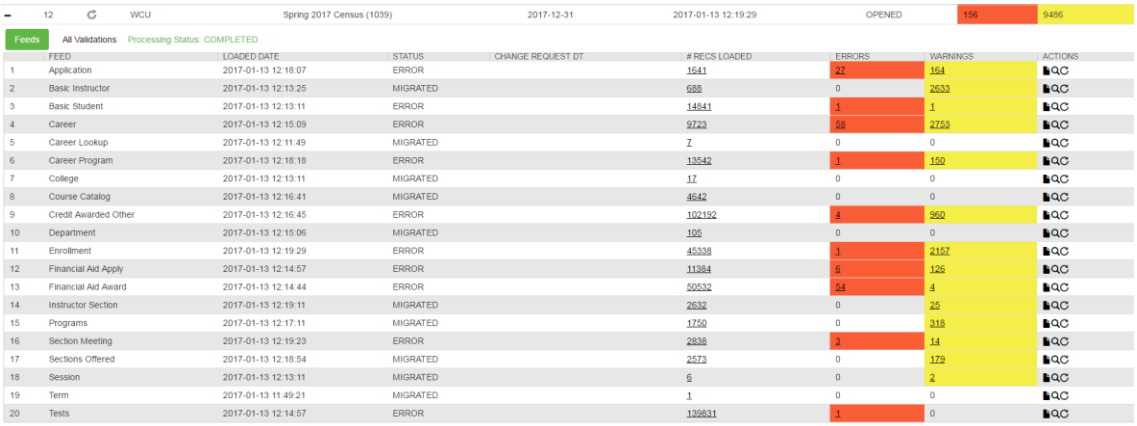
In addition to sorting the snapshot list, snapshots can be searched or filtered by campus, status, name or partial name, ID, start or end dates. Enter criteria and click the magnifying glass to execute the search. The list will be filtered by the search criteria.

4.3 Submissions - View Errors and Warnings by Snapshot

In addition to viewing the list of all snapshots, the Submissions pages are useful for examining the errors and warnings that exist in upcoming snapshots. Errors and warnings must be reviewed before the snapshot due date. Errors will prevent a snapshot from migrating.



The aggregate count of errors and warnings is available from the main list of snapshots. If one or more errors exist, then the Error cell appears in red background. If one or more warnings exist, then the Warning cell appears in yellow background. To view the details of these validations, select the + to expand the snapshot view to display the component data feeds.



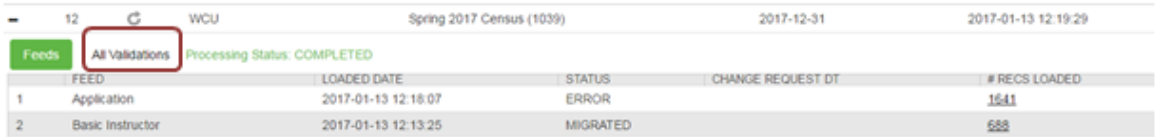
-	12		WCU	Spring 2017 Census (1039)	2017-12-31	2017-01-13 12:19:29	OPENED	156	9486	

Feeds | All Validations | Processing Status: COMPLETED

	FEED	LOADED DATE	STATUS	CHANGE REQUEST DT	# RECS LOADED	ERRORS	WARNINGS	ACTIONS
1	Application	2017-01-13 12:18:07	ERROR		1641	27	164	QC
2	Basic Instructor	2017-01-13 12:13:25	MIGRATED		688	0	2633	QC
3	Basic Student	2017-01-13 12:13:11	ERROR		14641	1	1	QC
4	Career	2017-01-13 12:15:09	ERROR		9723	58	2753	QC
5	Career Lookup	2017-01-13 12:11:49	MIGRATED		7	0	0	QC
6	Career Program	2017-01-13 12:18:18	ERROR		13542	1	150	QC
7	College	2017-01-13 12:13:11	MIGRATED		17	0	0	QC
8	Course Catalog	2017-01-13 12:16:41	MIGRATED		4642	0	0	QC
9	Credit Awarded Other	2017-01-13 12:16:45	ERROR		102192	4	960	QC
10	Department	2017-01-13 12:15:06	MIGRATED		105	0	0	QC
11	Enrollment	2017-01-13 12:19:29	ERROR		45338	1	2157	QC
12	Financial Aid Apply	2017-01-13 12:14:57	ERROR		11384	6	126	QC
13	Financial Aid Award	2017-01-13 12:14:44	ERROR		50532	54	4	QC
14	Instructor Section	2017-01-13 12:19:11	MIGRATED		2632	0	25	QC
15	Programs	2017-01-13 12:17:11	MIGRATED		1750	0	318	QC
16	Section Meeting	2017-01-13 12:19:23	ERROR		2838	3	14	QC
17	Sections Offered	2017-01-13 12:18:54	MIGRATED		2573	0	179	QC
18	Session	2017-01-13 12:13:11	MIGRATED		6	0	2	QC
19	Term	2017-01-13 11:49:21	MIGRATED		1	0	0	QC
20	Tests	2017-01-13 12:14:57	ERROR		139831	1	0	QC

4.3.1 View All Validations by Snapshot

Errors and warnings, collectively called validations, can be viewed for an entire snapshot or by individual feed. Click All Validations to display the list of all validations.



-	12	WCU	Spring 2017 Census (1039)	2017-12-31	2017-01-13 12:19:29

Feeds | All Validations | Processing Status: COMPLETED

	FEED	LOADED DATE	STATUS	CHANGE REQUEST DT	# RECS LOADED
1	Application	2017-01-13 12:18:07	ERROR		1641
2	Basic Instructor	2017-01-13 12:13:25	MIGRATED		688

By default, the list of validations is sorted by Type (Errors, then Warnings) and Edit ID. Alternative sort criteria can be applied by selecting one of the options from the Sort. Any of these can be sorted in ascending or descending order:

Type	The validation type, either error or warning.
ID	The validation ID. This is the unique identifier in the SDM for the validation.
Name	The descriptive name of the validation.
Feed	The data feed to which the validation applies.
Feed Record Key	The unique identifier for the record to which the validation applies.
Description	A description of the validation. In many cases, this parallels the validation name.
Details	An elaborated description of the validation, including details of the affected record.

4.3.2 Exporting All Validations by Snapshot

Exporting validations will be useful if using external tools, e.g. Excel or SAS, to analyze the validations and/or to distribute the validations to operations offices for resolution.

Do not export validations for feeds with student information on any non-secure computer. Validation details omit student names but may still contain identifiable or FERPA-protected information that should never be downloaded to a non-secure computer.

Only the following feeds are free from student or instructor details: Career Lookup, College, Course Catalog, Department, Programs, Section Meeting, Sections Offered, Session, Term.

From the All Validations list view, export options include:

Details - Excel list of all validations and their details.

Summary - An Excel list of all unique validations and the count of each.

Summary CSV - A comma separated (CSV) list of all validations and the count of each. Note that a csv file will download more quickly than Excel and can still be opened in Excel. In addition to download speed, the advantage of csv files are that leading zeroes are retained when imported into most tools other than Excel.

Header/Footer option - With the header/footer option, details about the snapshot or download may be included as a header or footer in the export. To import these lists into external tools, you may want to omit the header/footer information by de-selecting this option.

Previous submission - In the event the current submission is in process or in error status, it may be useful to download validations from the previous submission in order to review or share corrective actions needed without having to wait for the current submission to complete.

4.3.3 Viewing Submission Details and Validations by Snapshot by Feed

Validations can also be reviewed by individual feed. From the expanded view, these details are available by data feed:

Feed Loaded Date	The most recent date and time the feed was loaded.
Feed Status	The status of the individual feed within the snapshot.
Feed Change Request Date	If a change request was made after the snapshot closed, then the request date will display. Normal daily loads of a snapshot are not change requests.
Feed Records Loaded	This is the count of records included in the feed. This number is a hyperlink to a Widget report of all submitted records.

4.3.3.1 View Records Loaded

The # RECS LOADED number is a hyperlink to a Widget report of all submitted records. This hyperlink displays a Widget list report of all loaded records in the feed from the Staging table. Since the list is displayed as a Widget report, it contains all sort, search, edit and export functionality of a Widget report. See the Reporting section for additional details.

Note that this report is run from the staging table, which means it will have some descriptive data elements compared to the raw data feed but will not have all of the derived elements that a full dataset may have.

Reviewing the aggregate count of records loaded is an initial step in data validation. For example, if there are 20 academic departments in the university, then the department feed should have 20 records. Similarly, if the university has 20,000 headcount, then the career feed should have approximately 20,000 records loaded in the career feed in a census snapshot as the semester begins.

4.3.3.2 View Feed Submission Log

The page icon will display details about the stages of the submission. In the event a feed fails to load, it may be useful to view the log for details. The log can be exported to Excel by clicking the Excel icon. Remove the log from the display by clicking the X in the upper right corner of the window.

	FEED	LOADED DATE	STATUS	CHANGE REQUEST DT	# RECS LOADED	ERRORS	WARNINGS	ACTIONS
1	Application	2017-03-21 04:15:12	ERROR		18582	1539	498	
2	Basic Instructor	2017-03-21 04:11:48	MIGRATED		809	0	0	
3	Basic Student	2017-03-21 04:11:53	ERROR		18827	23	2	

4.3.3.3 View Feed Validations

The magnifying glass icon will display the full list and details of validations for the feed. This is the same list displayed in the All Validations screen, but filtered for the specified data feed. The list can be sorted or searched by any of the listed options. This may be useful for quickly checking if a specific error or record has been resolved.

4.3.4 Exporting Validations by Snapshot by Feed

To export the validations for an individual feed, from the expanded view of feeds within the snapshot, view the list of validations for that feed by selecting the magnifying glass. (See 4.3.3.3 View Feed Validations)

From the list of validations for that feed, the following export options are available:

Details	Excel list of all validations and their details.
Summary	Excel list of all unique validations and count of each.
Summary CSV	A comma separated (CSV) list of all validations and the count of each. Note that a csv file will download more quickly than Excel and can still be opened in Excel. In addition to download speed, the advantage of csv files is that leading zeroes are retained when imported into tools other than Excel.
Header/Footer option	The With the header/footer option, details about the snapshot or download may be included as a header or footer in the export. To import these lists into external tools, you may want to omit the header/footer information by de-selecting this option.

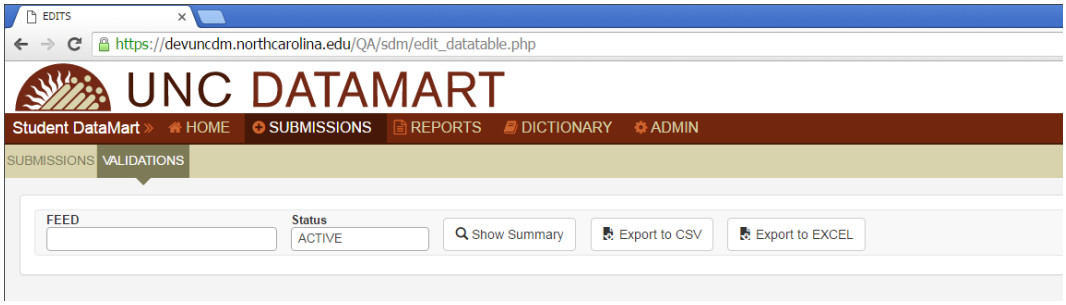
4.3.5 Viewing Records With Errors or Warnings

To view the full records that are incurring errors or warnings, then click the number of errors or number of warnings. This hyperlink will display a Widget list report of the affected records from the Staging table. Since the list is displayed as a Widget report, it contains all sort, search, edit and export functionality of a Widget report. See the Reporting section for additional details.

Note that this report is run from the staging table, which means it will have some descriptive data elements compared to the raw data feed but will not have all of the derived elements that a full dataset may have.

4.3.6 Viewing or Exporting All Possible Validations in SDM

From the Submissions Page, the Validations Tab displays all Errors and Warnings that exist in SDM



View options include selecting one or more feeds, and active or inactive validation status.

Show Summary displays a list of all validations.

The default sort order is by Edit ID. This list can be sorted or searched by the content in any of the labeled columns:

Edit ID	The unique identifier for a validation within the SDM environment.
Data Feed Name	The data feed to which the validation applies.
Edit Type	The type of validation, critical error versus warning.
Edit Name	The descriptive name of the validation.
Status	The status of the validation, active versus inactive.
Initial Effective Date	The date the validation was first in place. If null, the validation was in place since the initial implementation of SDM.
Final Effective Date	The latest date the validation was in effect. If null, then validation is still in place.
Edit Number	The unique identifier for a validation across SDM environments. This number identifies the validation whether in Production, QA or Development.
Edit Description	A description of the validation. In many cases, this parallels the validation name.

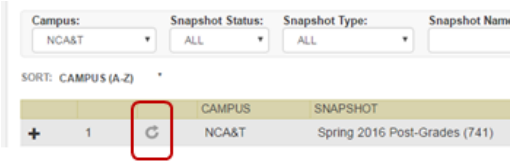
The list can be exported to CSV or Excel. This is a generic list of all validations. No institution or student-specific information is included so the list can be safely exported and/or distributed without concern for sharing personally identifiable information (PII) or other FERPA-protected information.

4.4 Reloading a Snapshot

Snapshots reload automatically each night during the submission window. A snapshot may need to be reloaded manually.

Depending on the snapshot and the quantity of data, reloading an entire snapshot can take multiple hours. Communicate with your team to make sure no one else working on the snapshot will be affected.

Users with the "Upload Data" explicit permission will see the Reload arrow. To reload an entire snapshot, click the Reload arrow that appears to the left of the Campus name.



Examples of occasions when manual reload may occur include:

4.4.1 Census Day

The automatic load of the snapshot will have occurred before or at the beginning of the day on the census date. To capture activity as of the end of the day, the designated campus user, usually someone in Institutional Research, will rerun (reload) the snapshot manually after having verified that operational offices such as Admissions, Registrar, Financial Aid are reasonably certain data errors have been resolved.

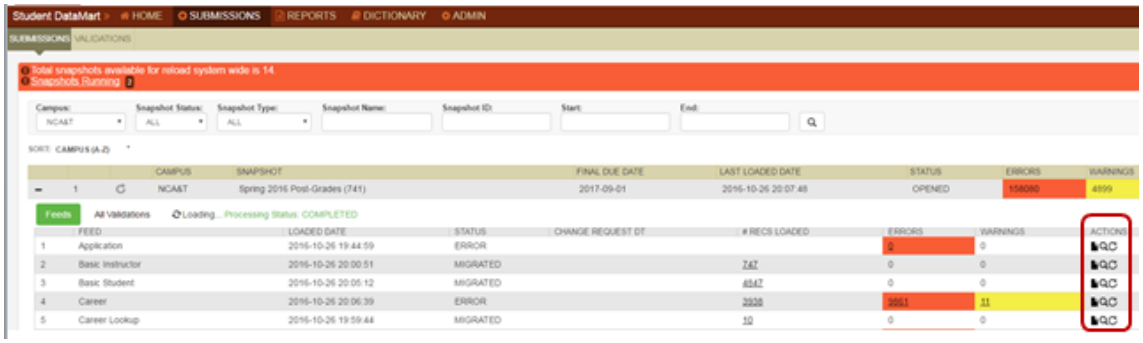
4.4.2 Resolving Errors

To confirm if errors have been resolved without waiting for the scheduled overnight run, the snapshot may be rerun (reloaded) manually.

4.5 Reloading a Data Feed

For campus users, the ability to reload a data feed in isolation is only available in the QA environment, not in production. The reason for this is to ensure the integrity of the population included in the snapshot since the population of students depends on the order of feeds.

Users with the "Can Upload" explicit permission will see the reload arrow. The feed reload arrow appears under the Actions column. The reload arrow is available after the feed has finished running. If the feed status is Running or Staged, then the reload arrow will not be available. If the feed status is Error or Migrated then the reload arrow is available.



Examples of occasions when manual reload of a data feed may be desirable:

4.5.1 Testing Code Changes

When testing code changes - either upgrades to new releases or interim fixes -- reloading only the data feed will test the code change and will require far less time than rerunning the entire snapshot.

4.5.2 Testing Data Fixes

Similar to testing code changes, when testing data fixes running the affected feed in isolation will require far less time than running an entire snapshot.

Running a feed in isolation or running feeds outside the normal order can produce populations and results that are incorrect. The population pulled in some feeds depends on the population pulled in other feeds. See Appendix C - How Things Work: The Role of the Aggregator Table.

4.6 Submissions - Request Modification

The modification request feature, also called snapshot change request, is only available for closed snapshots.

The snapshot status applies to the snapshot as a whole, inclusive of all campuses. A campus finding errors soon after their individual due date may have to wait for all campuses to complete the snapshot before the change request is automatically available. GA staff can manually intervene to close individual campus snapshot feeds if needed.

An approved snapshot change request will re-pull selected data from the source (Banner or Peoplesoft) and will re-run all applicable data validations. With any change request there is a risk of introducing new errors that will have to be resolved before the snapshot can be considered closed again.

4.6.1 When to Request a Snapshot Modification

After determining data errors exist in the closed snapshot, those data must be corrected in the ERP, Banner or Peoplesoft, BEFORE submitting a change request. If code or crosswalks resulted in the errors, then the code and/or crosswalks must be corrected BEFORE submitting a change request. When the GA team approves a snapshot modification, SDM will immediately re-pull the affected data elements and/or records from the source. For this reason, all fixes must be in place in the ERP before requesting the change.

4.6.2 When to Request Multiple Snapshot Modifications

Only one change request per feed per snapshot can be entered at one time. The wrench icon will not be visible until a pending request has closed.

4.6.2.1 Snapshot Modifications – Multiple Feeds

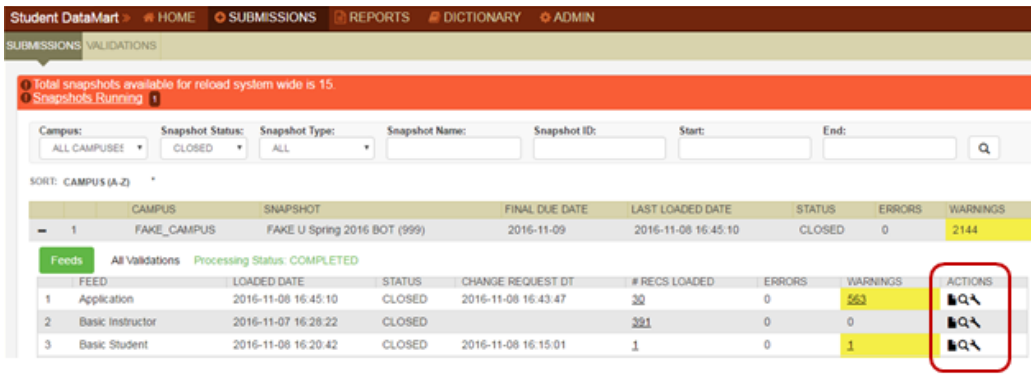
If multiple feeds need to be corrected, then one request for each feed must be entered. Note in the details of the request which feeds are affected. While the order of submitting requests does not matter, the GA team reviewing the request will need to approve the requests in a specific order due to the data validations that check across feeds.

4.6.2.2 Snapshot Modifications – Multiple Populations

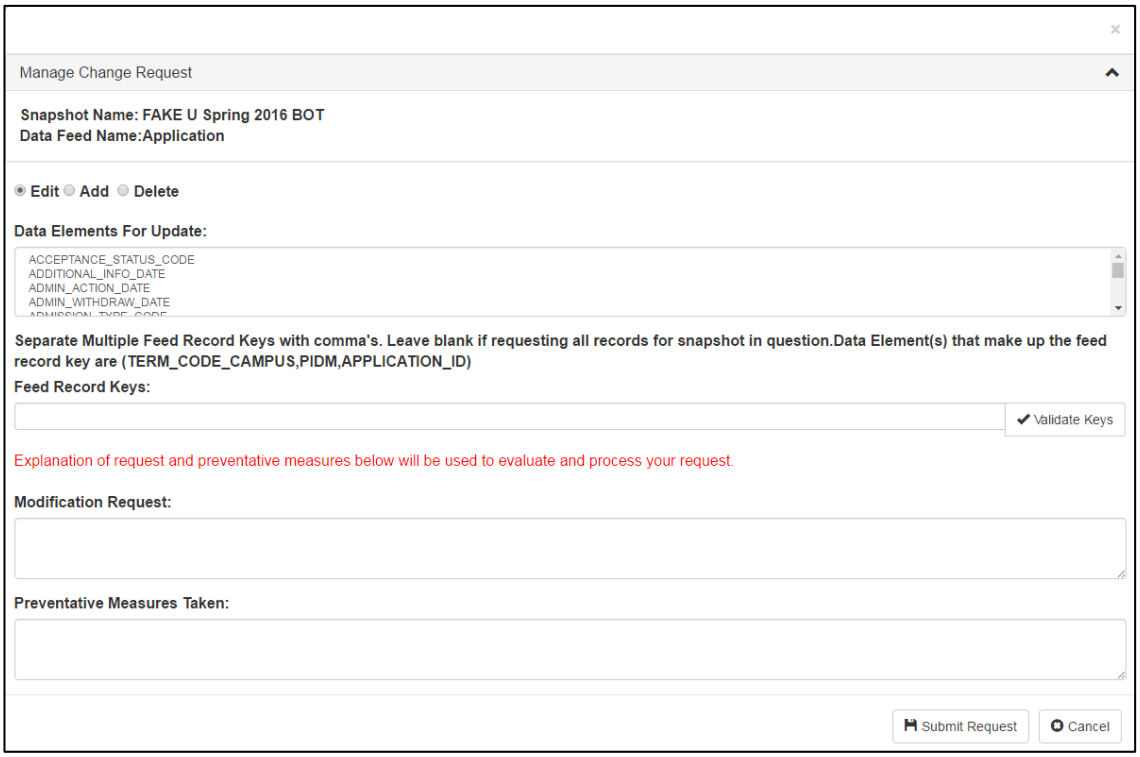
If multiple populations and/or multiple changes are needed within the same feed, then it is advisable to enter separate requests. Each request should specify the distinct population and distinct data elements to be re-pulled. While it is technically possible to combine the requests into one, doing so would mean re-pulling records and elements that do not need to be revised.

4.6.3 How to Request a Snapshot Modification

After determining a data change is required to a closed snapshot, decide which feed or feeds are affected. The Submissions Page - Expanded View of the snapshot will display the wrench icon under the Actions column.



The Manage Change Request form will display.



4.6.3.1 Snapshot Change Type

Determine if the change request is an edit, add or delete and click the appropriate radio button.

4.6.3.2 Data Elements for Update

For Edits only, the list of data elements in the feed will display. Select the element(s) to be updated. Multiple fields can be selected by using ctrl-click (non-consecutive) or shift-click (consecutive). Note: Knowledge of the relationship between data feed elements is needed. For example, if the major_code is selected then the major_desc is the related campus description and must also be selected.

If no elements are selected for update then ALL elements for the record(s) identified will be updated.

4.6.3.3 Feed Record Keys

Feed record keys are the unique identifier for a record in a data feed submission. Entering the feed record key(s) identifies which records will be updated. Data elements that comprise the feed record key differ by data feed and are displayed as help text on the change request form.

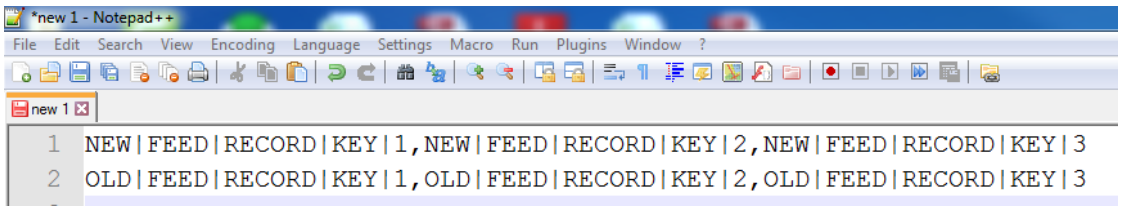
For Add requests, the feed record key must be created based on what it will be when the record is added from the ERP.

For Delete or Edit requests, there are a few ways to obtain existing feed record keys from the SDM.

- From the submissions page, click the hyperlink with the number of Records Loaded for the affected data feed. Export the results to csv and search for the applicable records. Feed record key will be one of the columns in the output.
- From the Widget, create a report using the Staging dataset for the affected data feed. Select any identifier needed, filter by the appropriate snapshot and export the results.

Multiple feed record keys must be separated by commas.

When working with a list of feed record keys, it may be helpful to paste the records into a text editor like Notepad++ in order to confirm that each key is separated by comma, has no spaces and aligns in order with the new and old records.



For Edit requests that involve updating data elements that are components of the feed record key, an additional prompt will appear to distinguish between the old and new feed record keys. Enter the existing feed record key(s) in the Old Feed Record Key field. Enter what will be the new feed record key in the Feed Record Keys field. If there are multiple records to update then the new feed record keys must be entered in precisely the same order as the old feed record keys they are replacing. E.g.

Old_1,Old_2,Old_3

New_1,New_2,New_3

4.6.3.4 Validate Keys

For Edits or Deletes, the Validate Keys feature is available. Clicking Validate Keys will check that the keys entered match records in the data that were submitted. Validate keys presents a list view of all of the feed record keys submitted and a check (√) if the record is found in the previously submitted data or an X if the record is not found in the submitted data.

Only records found in previously submitted data can be validated. If records receive an X, confirm that the key was indeed entered correctly. If not, then correct the key and attempt to Validate again.

4.6.3.5 Modification Request (Details)

In the Modification Request field, explain in detail what needs to be corrected. A detailed description helps the GA team check that all necessary fields have been selected for the request. The information may also let the GA team know to expect additional corrections to other data feeds.

4.6.3.6 Preventive Measures Taken

In the Preventative Measures Taken field, explain what measures have been or will be put into place in order to identify and resolve this type of issue before census (or snapshot close date) in the future.

4.6.3.7 Submit Request

After confirming:

- Campus stakeholders agree that a change to previously submitted data is needed,
- The data and/or code has been corrected in the ERP,
- The fields and records in the change request form are complete and accurate,
- The explanations of the Modification Request and Preventive Measures are complete and comprehensive

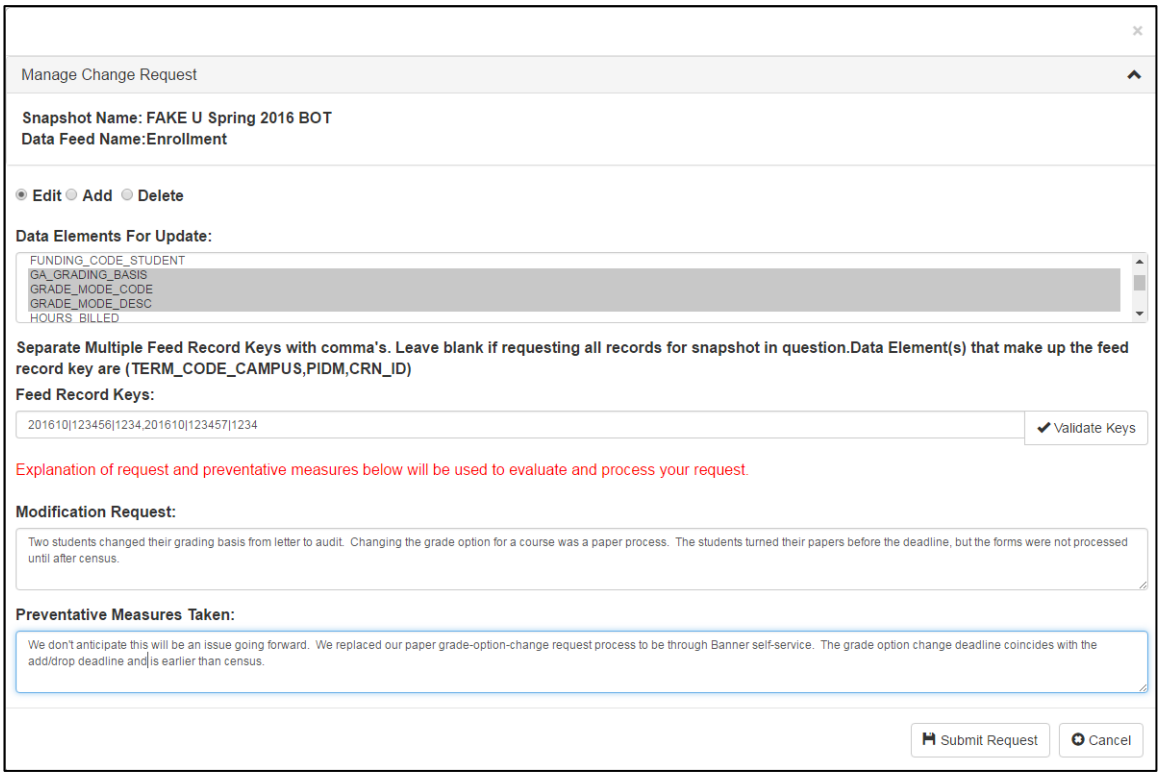
Then click Submit Request. Submit Request will initiate a request to the GA team for review. Results of the review will be communicated via system-generated email.

4.6.3.7.1 Snapshot Change: Edit Existing Records

Edits are the most common change request. An edit will update existing data elements for specified records. For example, an application record already exists for Joe Smith but his HS_CLASS_RANK is corrected in the ERP. There are two basic types of Edit requests depending on whether the feed record key is involved in the change.

4.6.3.7.1.1 Editing Records When the Feed Record Key <u>is Not Involved</u>

Complete the sections of the Manage Change Request form as described below. Only the designated element(s) and record(s) will be updated. If no elements are selected, then all elements will be updated. If no Feed Record Keys are entered, then all records will be updated.

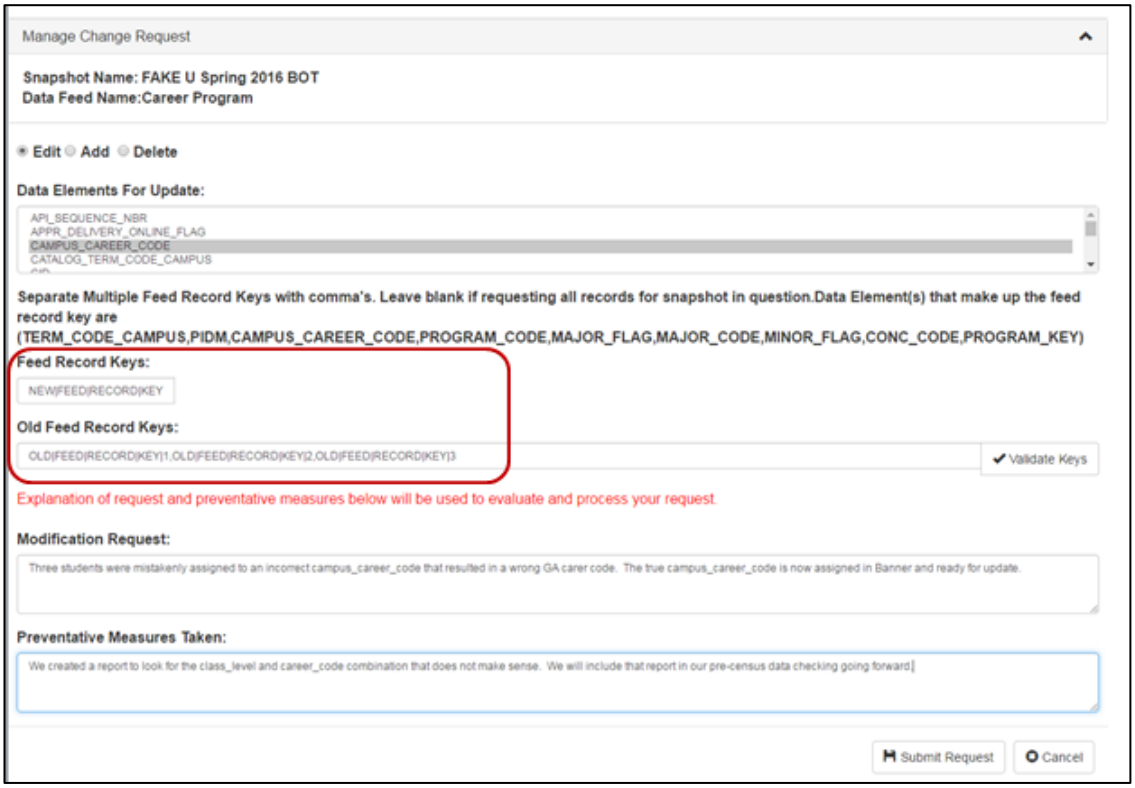


4.6.3.7.1.2 Editing Records When the Feed Record Key <u>is Involved</u>

The Feed Record Key identifies which record to re-pull from the ERP. If components of the Feed Record Key change then change request requires an additional step in order to map the new Feed Record Key to an old Feed Record Key that can be used to update the record correctly.

If any of the components of the FEED_RECORD_KEY are selected for edit, a box for the new FEED_RECORD_KEY and old FEED_RECORD_KEY will display. The new and old

FEED_RECORD_KEYs must be entered in the same order. That is, the first new FEED_RECORD_KEY listed must correspond to the first old FEED_RECORD_KEY listed, and so on. Multiple feed record keys must be separated by commas.



The old feed record keys are obtained in the same manner as if the feed record key were not involved in the update. The new feed record keys must be created with the knowledge of the data elements that changed.

If there are doubts about the correct new feed record key then you may be able to obtain or confirm the details using a currently open future snapshot. For example, if a record in a census snapshot needs to change and you need to confirm the feed record key as it exists in Banner/Peoplesoft right now, then you can usually find the relevant student or record in the End of Term snapshot. If the End of Term snapshot is not yet open, you can request it to be opened for this purpose.

There are some cases where it may be easier to delete a record and then add a new record rather than identifying all items that need to be picked for update. If you think adding and deleting will be easier than changing, then please submit a ticket and describe the change you are trying to make. The GA team will review and advise on the best way to accomplish the update.

4.6.3.7.2 Snapshot Change: Add Records

Adds are rare. Adds will query the ERP system (Banner or PeopleSoft) to insert records that were not included prior to the snapshot's final pull date. Selecting the Add will prompt for the feed record key of the record to be added. You will have to compile this using the data feed specifications for the elements included in the feed record key.

When adding an entirely new student to the snapshot, simply adding them (by PIDM) to the BASIC_STUDENT data feed will ensure they are pulled into all other appropriate data feeds.

4.6.3.7.3 Snapshot Change: Delete Records

Deletes are designed to remove records from the dataset that were not intended to be pulled originally. For example, this could happen when a process did not exclude a record because it did not complete for technical or timing reasons such as an enrollment drop for non-payment at census.

4.6.4 Approved Requests

Approved requests will immediately re-run SDM submission code to update the records and elements from the ERP. After the run, confirm that the intended corrections were made. If so, the process is complete. If not, more root-cause investigation may be required or additional modifications may be needed. See *Appendix E – How Things Work: Approved Snapshot Change Requests*.

4.6.5 Denied Requests

Denied requests will not update any records or elements. An explanation for the denial will be recorded with the request and sent via system email.

4.6.6 Pending Requests

The GA reviewer may request additional information for the change request.

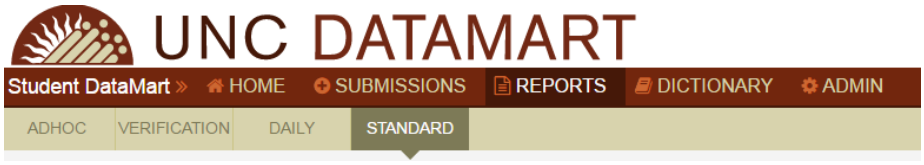
5 Reports

The reporting interface for the SDM is the Report Widget. The Widget allows users to run saved reports, to create reports and to download reports and/or the data behind reports. Whether running saved reports or creating reports, familiarity with the reporting datasets and element definitions is useful.

Reports will only produce results if the datasets that supply the report are error-free. If the underlying data feeds have errors that prevent them from migrating to the data model, then reports cannot pull the data. The Applications – Preliminary Applications, Acceptances report is the only report that runs from Staging Datasets as opposed to fully migrated datasets. As long as the underlying data feeds reach the Staged status then the Preliminary Applications report will produce results. For preliminary reporting, select the Snapshot Type – Beginning of Term. This snapshot type has far fewer error-checks and so will frequently migrate when the Census snapshot type will not.

5.1 Standard Reports

Standard Reports, also called stock reports, are accessed from the Reports->Standard tab. These are reports that have been written and saved for access by permissioned SDM users.



5.1.1 Standard Reports - Purpose

The primary uses of these reports are:

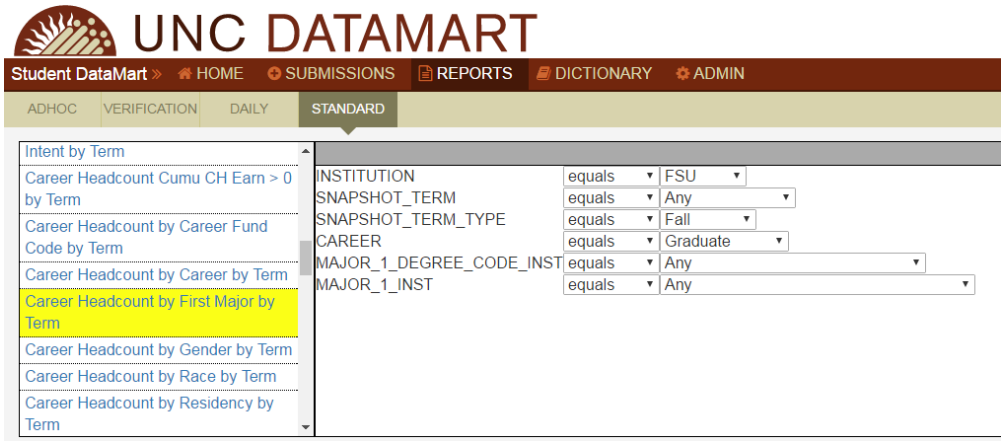
- Data Integrity - Data can pass submission validations and still have incorrect information. Standard reports can show common counts and distributions as an additional tool for validating data leading up to snapshot due dates.
- Frequently Asked Questions - Standard Reports contain quick reports that answer common questions about count of enrolled students, count of applications, etc. The reports may be useful in answering these high level questions.

5.1.2 Standard Reports - Naming Convention

Report names begin with the dataset used to create the report and then contain descriptive information about the report structure.

5.1.3 Standard Reports - Running the Reports

To run a report, click the report name, respond to the prompts and click Run Report. Note that prompts will vary by report.

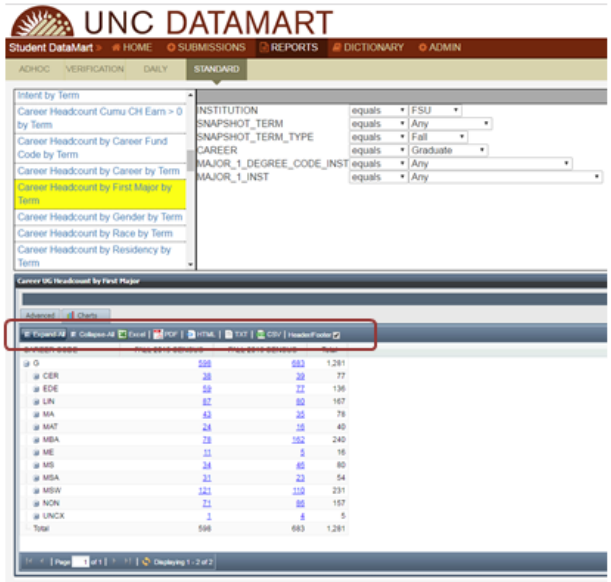


If the user has permissions for the underlying dataset and data exists for the specified parameters, then the report will display in the lower window.

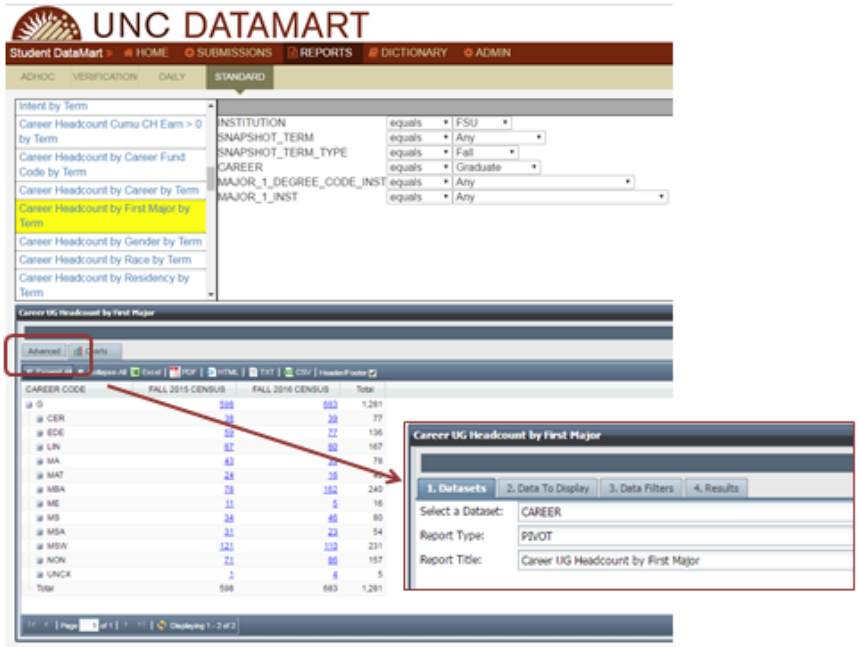
5.1.4 Standard Reports - Viewing, Exporting and Manipulating Results

Results can be exported in any of the listed formats: Excel, PDF, HTML, TXT or CSV.

The Footer will contain the Report ID, Report Name, Underlying Dataset name, URL and the SQL logic for the selection criteria. To suppress this information, uncheck the Header/Footer box.

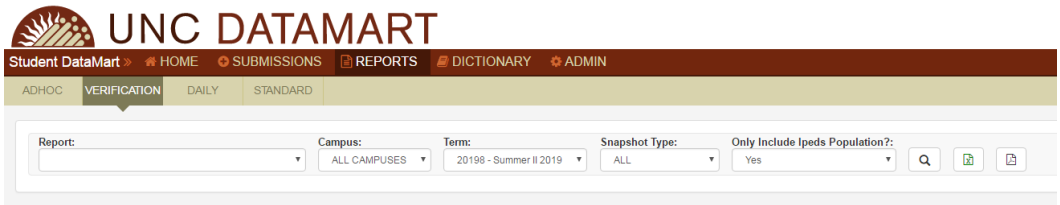


To view or manipulate the selection criteria, click the Advanced Tab. From the Advanced Tab, confirm the dataset used, view or edit the data selected, the sort criteria and/or the filters. Revised criteria can also be Saved and will generate a new Report ID.



5.2 Verification Reports

Verification reports are accessed from the Reports->Verification tab. These are also pre-written reports saved for access by permissioned SDM users. The difference between Verification reports and Standard reports are that Verification reports make use of SAS programs behind the scenes to display more comparisons or derivations, are formatted for export but do not allow modification to report criteria.



5.2.1 Verification Reports - Purpose

The purpose of Verification Reports is very similar to the purpose for Standard Reports.

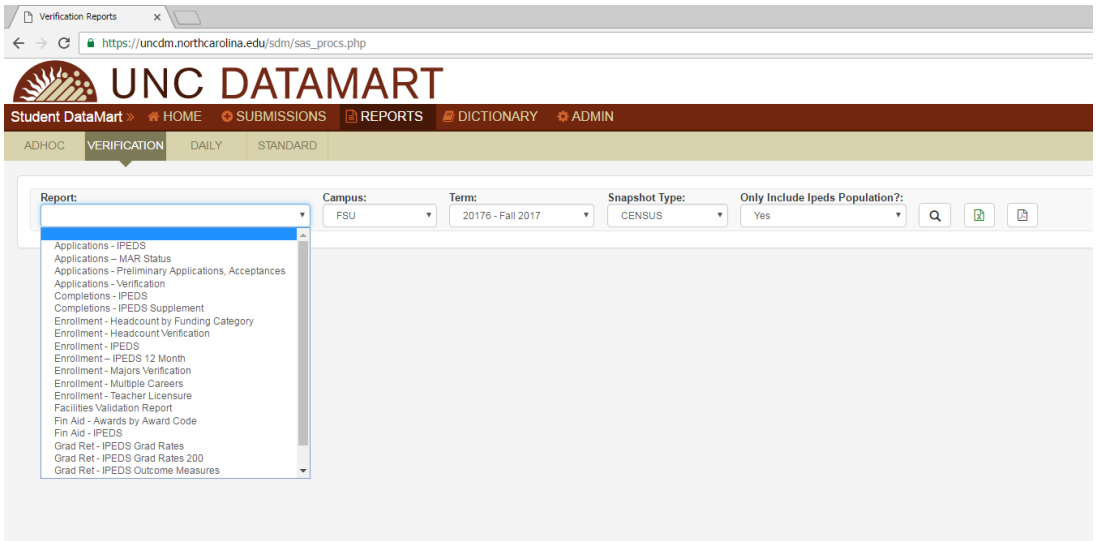
- Data Integrity - Data can pass submission validations and still have incorrect information. Verification reports show common counts and distributions as an additional tool for validating data leading up to snapshot due dates. Some verification

reports are written specifically for IPEDS verification. The reports content and format will mimic IPEDS surveys.

- Frequently Asked Questions - Verification Reports contain common reports that show high level numbers over time. The reports may be useful in answering these common questions from student data stakeholders.

5.2.2 Verification Reports - Naming Convention

Report names begin with the subject area and then some descriptive information about the report content. Note that Verification Reports intended for IPEDS verification will include IPEDS in the name.

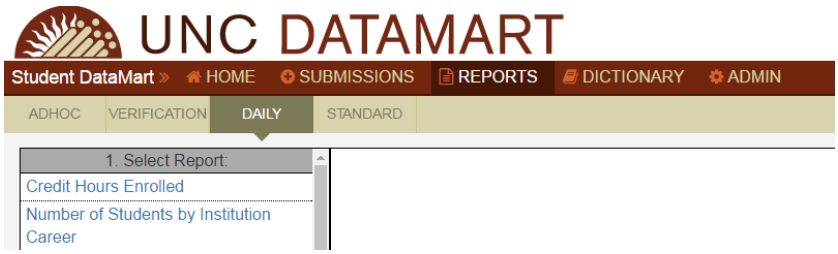


5.2.3 Verification Reports - Running, Viewing and Exporting the Reports

To run a report, select the report name from the Report drop down list, answer the filters and select the Output format: Run displays the output in html. Excel or PDF will create a file for download. The naming convention for the output file is:
Institution_ID||Report_ID||datetime.filetype.

5.3 Daily Reports

Daily reports are pre-defined reports that save a set of aggregate data for census snapshots on a daily basis.



5.3.1 Daily Reports - Purpose

Daily reports offer minimal functionality in the current iteration of SDM. They save aggregate data from each census snapshot with the objective of showing trends over time, leading up to census and comparing to the same point in time to prior years. The daily information is not needed for official reporting but will be useful in managerial (operational) reporting. The primary purpose currently is to save information for use in future tool development that will focus on managerial information.

5.3.2 Daily Reports - Naming Convention

Daily reports are named based on the content of the report.

5.3.3 Daily Reports - Running, Viewing and Exporting

Run the report by selecting the Report Name, answer the filters and select Run Report. The output will display in the window below. To view or manipulate the selection criteria, click the Advanced Tab. From the Advanced Tab, confirm the dataset used, view or edit the data selected, the sort criteria and/or the filters. Revised criteria can also be Saved and will generate a new Report ID.

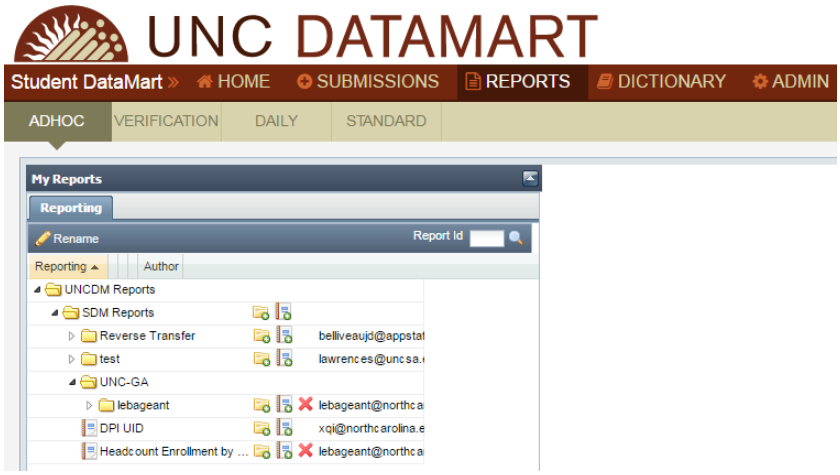
5.4 Ad Hoc Reports - Saved

Another location for saved reports in in the Reports->Ad Hoc->My Reports tab. These are reports that have been created by SDM users and saved in the Ad Hoc report folder structure.

5.4.1 Ad Hoc Reports Folder Structure

The high level folders (UNCDM Reports, SDM Reports, and UNC-GA) are accessible to all of the folders that appear below them.

The last indented list is an example of private reports visible only to a specific user. This personal sub-directory is the location where any personal saved reports will be located.



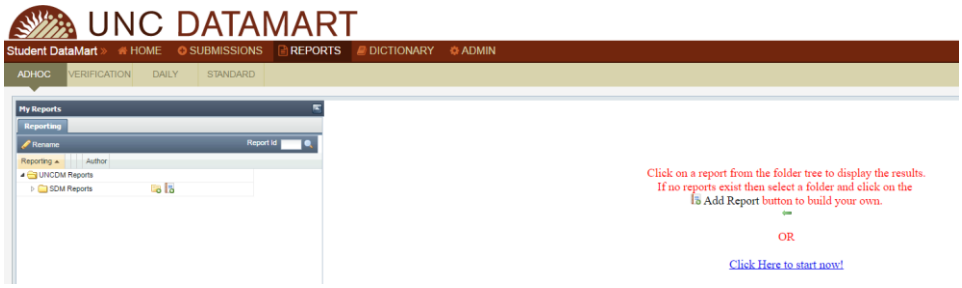
5.4.2 Ad Hoc Reports - Running, Viewing and Exporting

Ad Hoc Reports that are already saved in the My Reports folder structure are run by clicking the Report Name. Results will display and can be downloaded in Excel, HTML, PDF, TXT or CSV. The Footer will contain the Report ID, Report Name, Underlying Dataset name, URL and the SQL logic for the selection criteria. To suppress this information, uncheck the Header/Footer box.

From the Advanced Tab, criteria and/or the filters can be revised and saved and will generate a new Report ID.

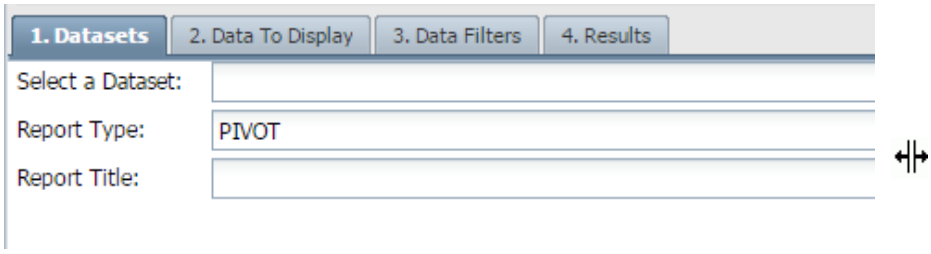
5.5 Ad Hoc Reports - Create Your Own

The Reporting Widget allows you to create and save your own customized reports in addition to viewing or modifying customized reports made by others. Ad Hoc reports are available from the Reports->Ad Hoc tab.



5.5.1 Create Your Own - From New

To start creating reports Click Here to Start Now and then click the Advanced tab. Then walk through the steps to generate a report.



5.5.1.1 Datasets

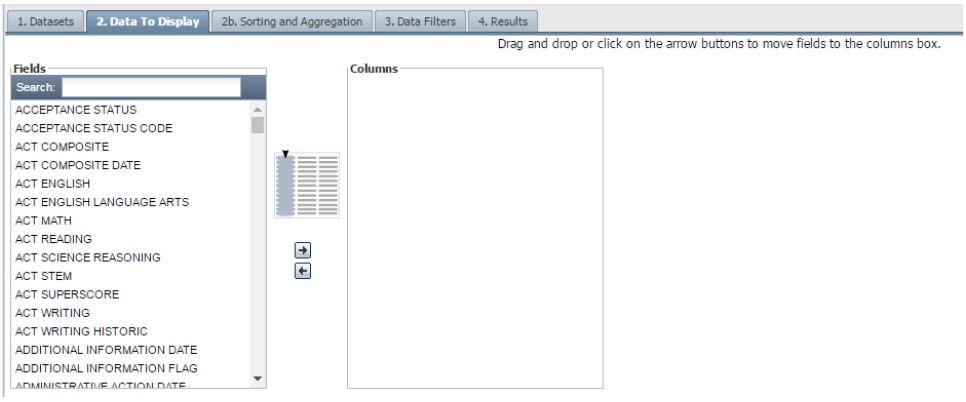
Select a Dataset from the drop down list of all permissioned datasets. Hover over the dataset to see a description and to link to more information about the contents.

5.5.1.2 Report Type

Select Flat or Pivot table. Flat reports display unit record information. In flat reports you can sort, count, total or give the average of data using columns and rows. Pivot tables display aggregated information. Pivot tables allow automatic sort, count, total or give the average of data in one table.

5.5.1.3 Report Title

Optionally, specify a Report Title. This can be anything, though best practice recommends including the dataset and/or descriptive information in the report title such that external users may efficiently understand their contents.



5.5.1.4 Select and Order Data

Advance to the Data to Display tab in order to select the data to include in the report. Available data elements are listed alphabetically in the Fields box. Hovering over the element name displays a link to the dictionary with the element's full definition and description.

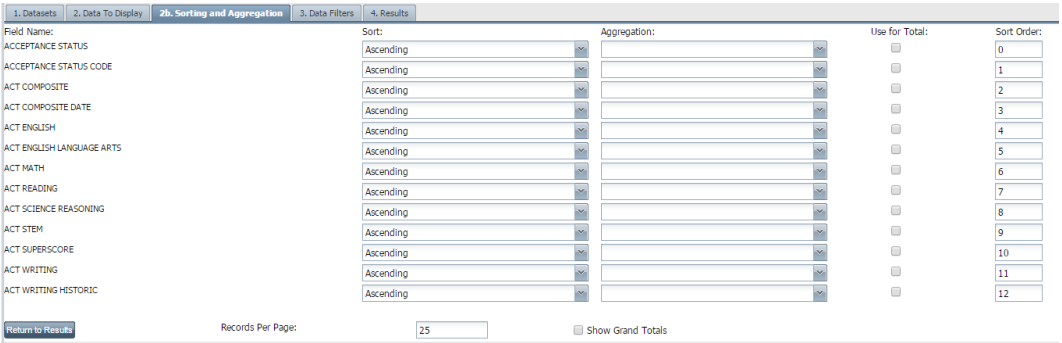
In a pivot report, specify which data should be in columns and which should be rows. At least one row and one column must be selected. A maximum of three (3) columns can be displayed in Results. Note: When a fourth (4th) column is selected, the Results tab disappears.

The data elements will appear in your report in the order they appear in the columns or rows boxes. Click and drag the data elements within the boxes to change the order of appearance.

5.5.1.5 Sorting and Aggregating

In Pivot tables, aggregation operators and elements are selected when Rows and Columns are defined. The default is the count of records.

In Flat tables, an additional Tab - 2b Sorting and Aggregation -- appears. Selections on this tab are optional. If nothing is defined here, the results will produce a list of unique records with no aggregation.

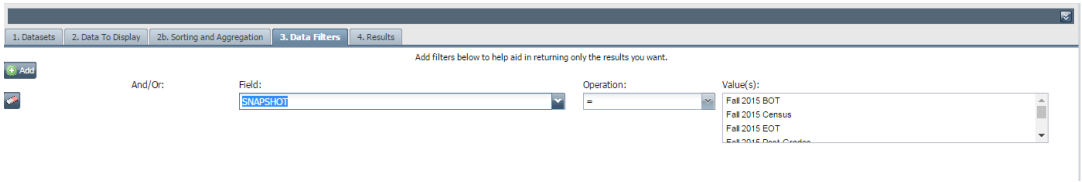


Sort	Ascending (A->Z) or Descending (Z->A)
Aggregation	Options for character fields are Count or Count Distinct. Options for numeric fields are SUM, Average, Count, Count Distinct, Min or Max.
Use for Total	If checked, the chosen aggregation will be added to any other checked field to create a row total.
Sort Order	A numeric value that shows the order that the fields will appear. The order is based on selections made on the Data to Display tab.
Records Per Page	Controls the size of the pagination within the web browser. Defaults to 25 and can be extended but will slow down processing.
Show Grand Totals	Will display a total row.

5.5.1.6 Data Filters

Advance to the Data Filters tab in order to select appropriate filters for the report.

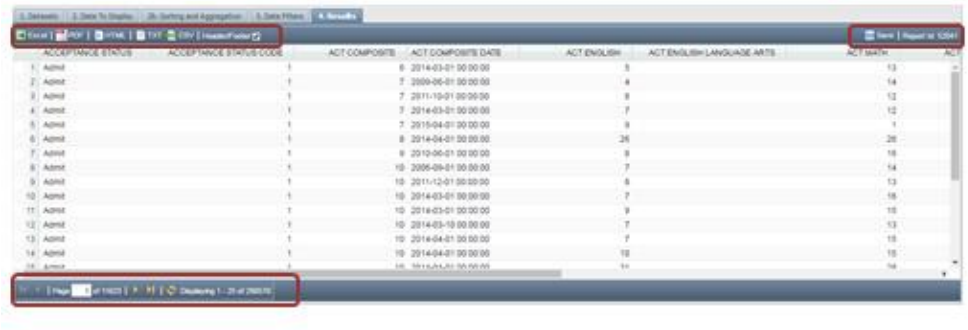
Filters are also technically optional. Practically, though, filters are desirable to control the size of results and processing time.



Field	Fields available for filter are listed alphabetically in the drop down box. Hovering over the element name displays a link to the dictionary entry with full description.
Operation	Depending on the element data type - character, numeric, dimension value or free-form text -- options can be selected from the drop down.
Operation (=)	The field selected will equal the value selected.
Operation (!=)	The field selected will not be equal to the value selected.
Operation (IN)	The field selected will be included in the value(s) selected.
Operation (NOT IN)	The field selected will be excluded from the value(s) you select.
Value(s)	Data elements that are supplied by a finite list of values (dimension tables) will display the possible values from the dimension table. Otherwise, values can be entered. Ctrl+select will select multiple, non-consecutive fields. Shift+select will select multiple consecutive fields. To enter multiple values, use a pipe delimiter. That is use the | to separate values. For example, multiple student pidms would be entered as 123456|789456|123123 with no spaces, commas or delimiters other than the pipe (|) between the values. For a long list of values, pasting values into a text editor and then into the Widget filter may be more efficient.
Add Filter(s)	Multiple filters can be applied by clicking the Add (+) button and defining additional filter(s).
Remove Filter(s)	Filters can be removed by selecting the Eraser icon.
Advanced Nesting	Filters within filters can be defined. This feature is useful when OR conditions are needed.

5.5.1.7 Viewing and Exporting Results

Advance to the Results tab to view the report.



The number of results displayed on the first screen depends on the data and on the Records per Page defined previously. The lower ribbon displays Page # of Total and Record 1 - # of Total. Left and right arrows will navigate through the pages.

From the upper ribbon, results can be downloaded in Excel, HTML, PDF, TXT or CSV. The Footer will contain the Report ID, Report Name, Underlying Dataset name, URL and the SQL logic for the selection criteria. To suppress this information, uncheck the Header/Footer box.

Widget reports frequently contain student information that should only be exported to a secure environment.

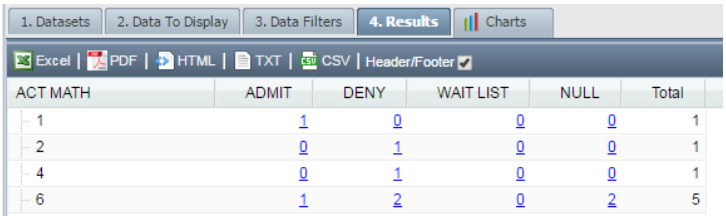
5.5.1.8 Saving Results

On the Results tab a **Report ID** is automatically generated for every report. Note this number in order to recall the report or to share the report with others.

Selecting **Save** is optional. If selected, a prompt to name the report appears and this Report Name will be saved to the user folder for recall.

5.5.1.9 Charts

If the selected data and results allow, then the Charts tab will also display. Bar and Pie Chart views are available.



ACT MATH	ADMIT	DENY	WAIT LIST	NULL	Total
1	1	0	0	0	1
2	0	1	0	0	1
4	0	1	0	0	1
6	1	2	0	2	5

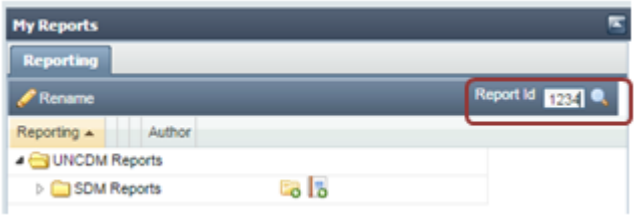
Only the first level of aggregation, rows or columns, will be displayed and used for the legend.

5.5.2 Create Your Own - From Existing

The Widget offers multiple ways to manipulate existing saved reports.

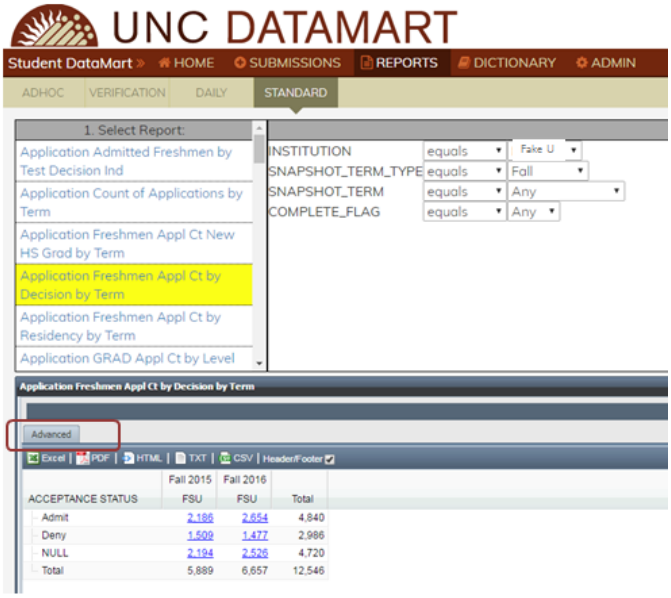
5.5.2.1 Report ID

If the Report ID is known, it can be entered into the Report ID search box. Click the magnifying glass to search. The results of the Report will be displayed. To view or manipulate the report criteria, click the Advanced tab and make revisions to the selected Title, data elements, filters, sorts or aggregate information.



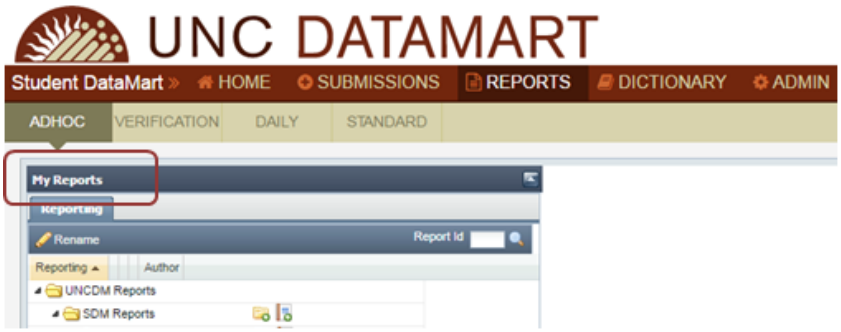
5.5.2.2 Report Name

From any reports page, other than Verification, widget reports can be executed and then modified. If run, then an existing report will display Results. These reports can be modified by clicking the Advanced tab and making revisions to the selected Title, data elements, filters, sorts or aggregate information.



5.5.3 Managing Saved Reports

Named reports are saved in the My Reports folder structure.



Arrows on the left of each folder can be clicked to expand the root folder and the subsequent personal folder.

Add Folder 	Creates a new folder within the displayed folder.
Add Report 	Adds a new report within the selected folder. Report criteria must then be defined.
Delete Folder or Report 	Removes the selected folder or report. Reports or folders can only be deleted by the user who created the item.

5.5.4 Sharing Reports

The Report ID can be shared to any SDM user. Any user can enter the Report ID and generate the report as long as s/he has permissions to the underlying dataset.

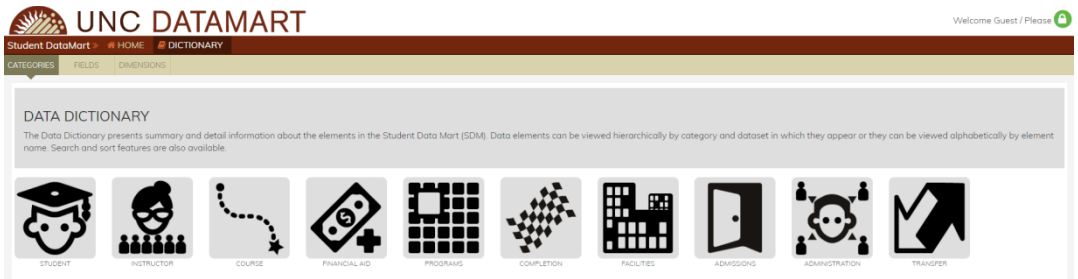
In addition to sharing the Report ID, a saved report can be moved up in the folder structure to allow others to see and run. A report in a specific user's folder is only visible to that user. If the report is created in or clicked and dragged up to the Campus folder then any permissioned user for that Campus can see and run the report.

6 Dictionary and Dataset Elements

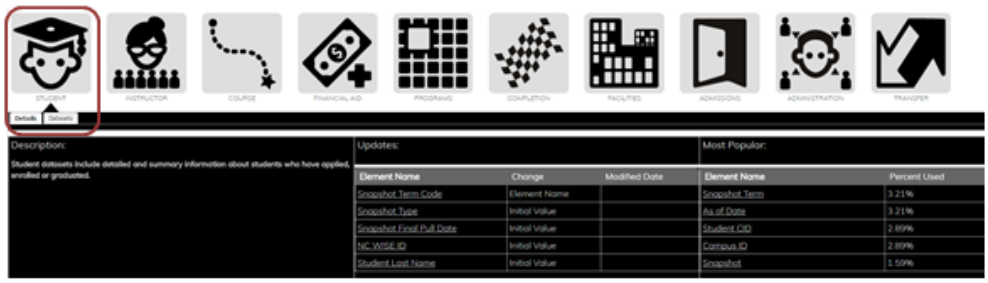
Access to the dictionary and dataset metadata is available without an SDM login. Knowledge of the datasets and element definitions is useful in report writing.

6.1 View Data Elements - By Category and Dataset

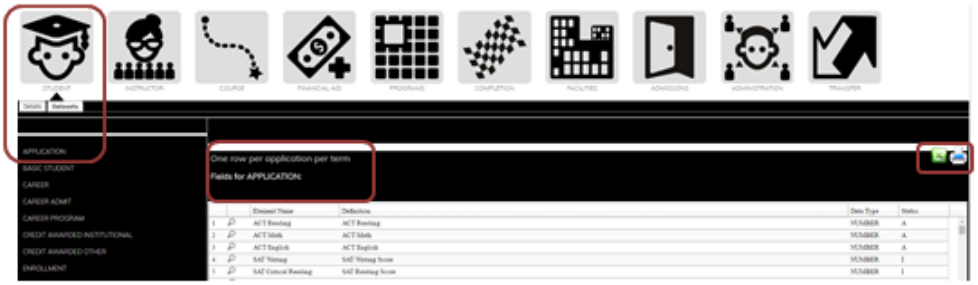
From the Dictionary page, the default view is by Category.



Selecting a Category will display a description of the Category.

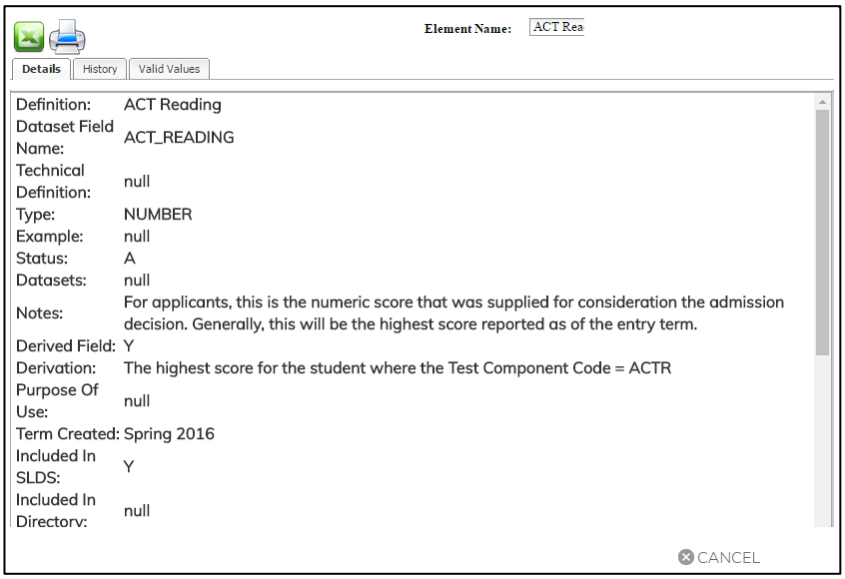


Selecting Datasets will display available Datasets within the Category. Selecting the Dataset will display a description of the dataset and a list of all elements in that dataset.



The list of elements can be printed or exported to Excel.

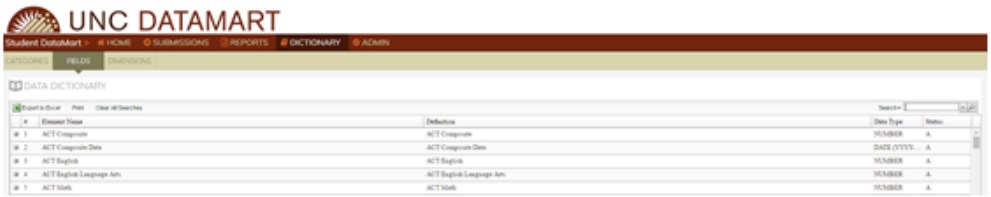
Selecting the magnifying glass icon will display additional details about the data element.



The details, history and, if applicable, valid values can also be printed or exported to Excel.

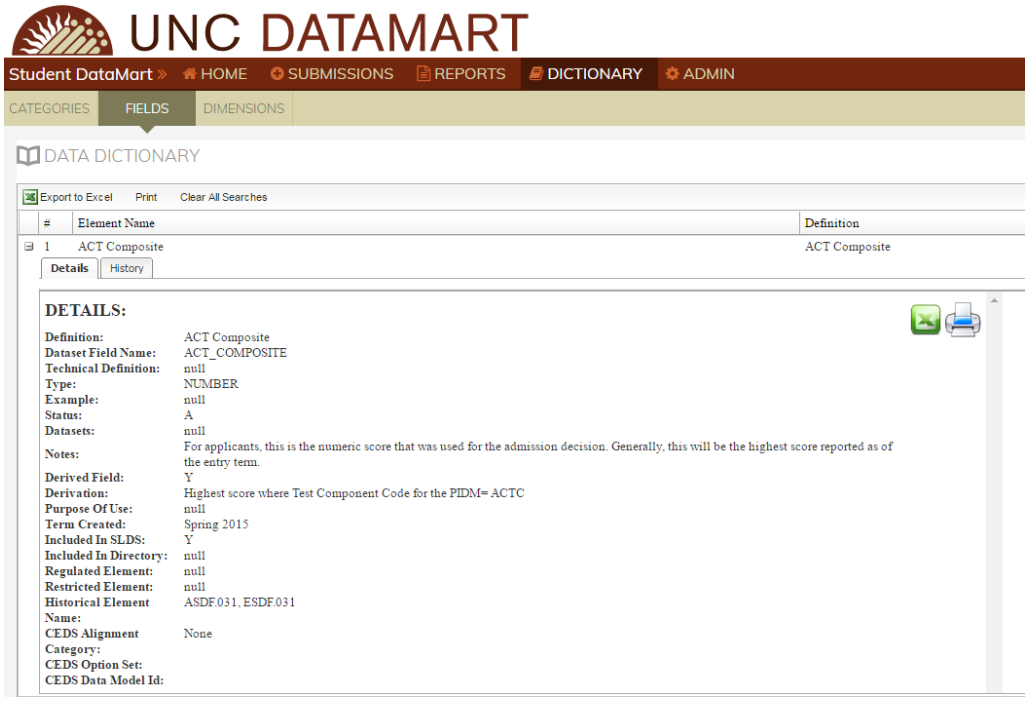
6.2 View Data Elements - All

From the Dictionary page, the Fields tab will display all elements.



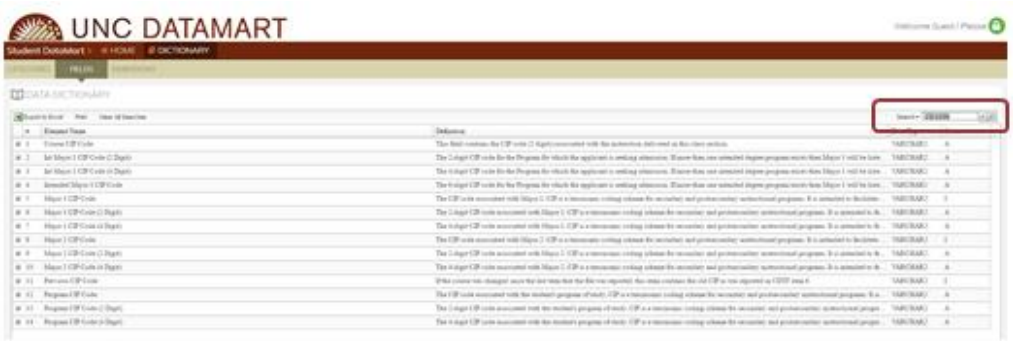
The list of fields can be printed or exported to Excel.

Expanding the data element displays the details, history and, if applicable, valid values.



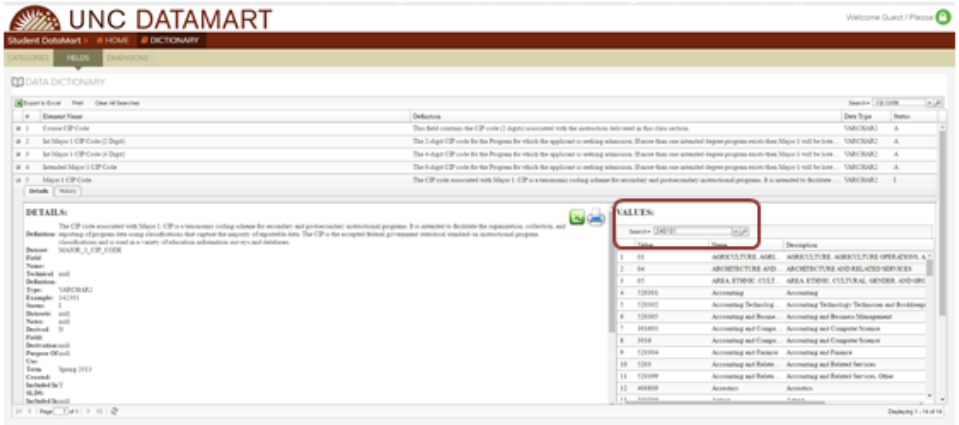
6.3 Search Data Elements

From the Dictionary->Fields screen, the search box allows full or partial word searches. Searches are not case-sensitive.



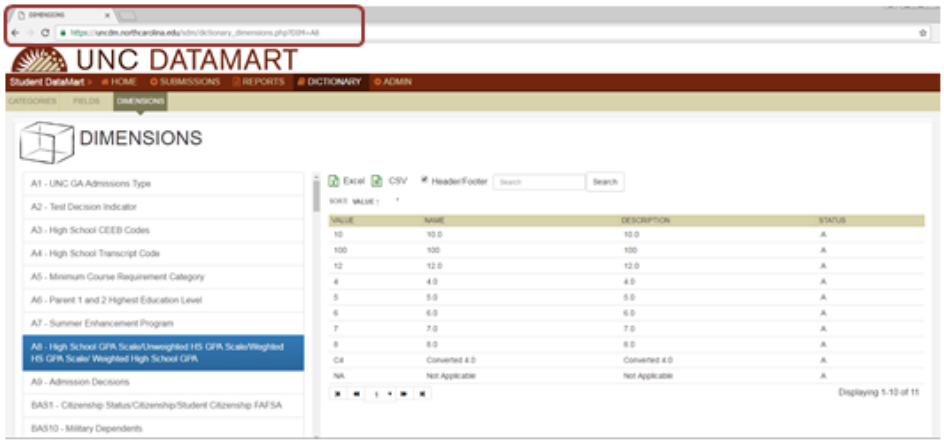
6.3.1 Search Valid Values

Within the Data Element view, a second search box is available to search valid values.



6.4 View Dimensions (Valid Values)

In addition to viewing valid values within the applicable data element, the universe of Dimension tables can be viewed from the Dimensions tab by selecting the Dimension.



6.5 Share Dictionary Elements

The SDM Dictionary is available to users without a login. Options to share the full dictionary with another user:

- Share the full URL: https://uncdm.northcarolina.edu/sdm/dictionary.php
- Export the list of elements to Excel and forward the downloaded file.
- Print the list of elements and share the paper copy. (Not recommended.)

6.6 Share Dimension (Valid Values)

Options to share dimension lists (valid values) with another user:

- Share the full URL for all dimensions: https://uncdm.northcarolina.edu/sdm/dictionary_dimensions.php

- Share the URL for a specific dimension where the dimension table alias occurs after the equal sign (=):
  https://uncdm.northcarolina.edu/sdm/dictionary_dimensions.php?DIM=P2
- Export the list of valid values to Excel and forward the downloaded file.
- Print the list of valid values and share the paper copy. (Not recommended.)

7 Dataset Downloads

There are multiple ways to download datasets.

Always make sure you are in a secure environment to download datasets. Depending on the criteria selected, FERPA-protected data will exist in the contents.

7.1 Dataset Download - Without Administrator Role

From Reports->AdHoc:

1. Click Here to Start Now.
2. Click the Advanced tab.
3. Select the Dataset to download.
4. Select Report Type = Flat.
5. Advance to tab 2, Data to Display, and select all fields of interest.
6. Optionally, advance to tab 2b for Sorting.
7. Advance to tab 3, Data Filters, and select the appropriate filters. While filters are technically optional, recommended filters include the Snapshot, Snapshot Type and Exclude From IPEDS or Exclude from RPT variables.
8. Advance to tab 4, Results.
    a. Optionally, unclick the Header/Footer box to remove from the output.
    b. Optionally, Save the report with a name in order to access from Report folder later.
    c. Download to Excel, TXT or CSV.

7.2 Dataset Download - With Administrator Role

There are a few options to download datasets more directly than using the Report Widget. These options require administrator permissions.

From Admin->Dataset Exports:

1. Campus administrators will have access only to their campus.
2. Optionally, filter by Snapshot Type.
3. Select one or many Snapshots. The list of available snapshots is filtered by the Snapshot Type selection.
4. Select one or many Datasets. The list of available datasets is filtered by the Snapshot selection.
5. Optionally, label your export. This is a free-form text label viewable only in the user interface.
6. Select the download format.

7.2.1 Dataset Download - CSV

If Comma Separated Values (CSV) is selected then a zip file is created containing the selected dataset(s) in csv format.

Naming convention for the zip file is _INSTITUTION_ID_zip_YYYYMMDDHHMMSS.zip

E.g. _2_zip_2017031641808.zip

Naming convention for the csv files within the zip file is INSTITUTION_ID-TERM(s)-SNAPSHOT_TYPE(s)-DATASETNAME-YYYYMMDDHHMMSS.csv

E.g. 2-20161-POST_GRADES-CAREER_PROGRAM-2017031641808.csv

7.2.2 Dataset Download - SAS

If SAS is selected, then a zip file is created containing the selected dataset(s) in SAS format. A transport file is also created for each dataset in case the campus SAS environment requires this.

Naming convention for the zip file is _INSTITUTION_ID_zip_YYYYMMDDHHMMSS.zip

E.g. _2_zip_2017031616189.zip

Naming convention for the SAS file within the zip file is

_INSTITUTION_ID_DATASET_ID_YYYYMMDD_HHMM.sas

7.2.3 Dataset Download - Service

Dataset download as a service is only functional if campus IT has established the environment to receive the datasets. See Appendix F – How Things Work: Dataset Download as a Service for a diagram of the service. Submit a help ticket to initiate the setup on campus.

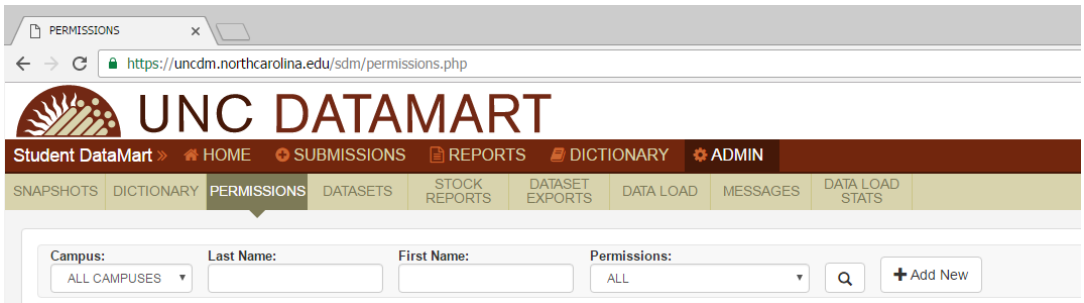
8 Administrator - Manage User Permissions

Campus Administrators (Admins) and UNC GA Admins have the ability to manage other users' access to the data mart and to provision datasets. GA Admins also manage snapshots and uploads of data.

The Admin functions are visible only for permissioned users.

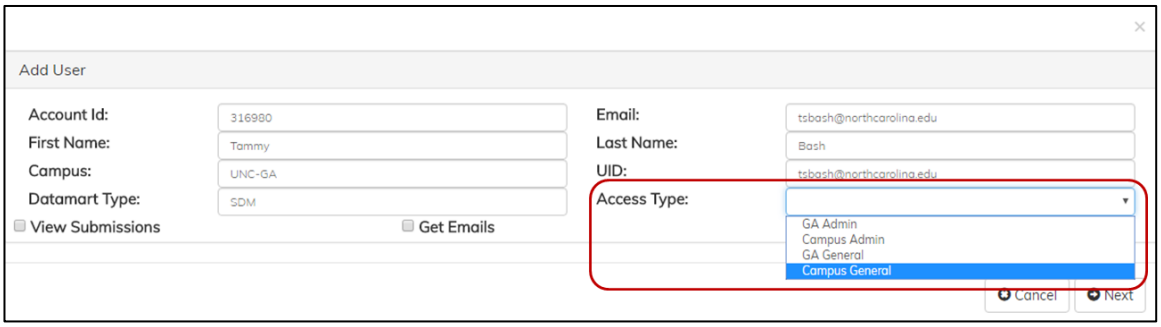
8.1 View Users and Permissions

From the Admin->Permissions tab, permissioned users are displayed in alphabetical order. Campus users will only see designated campus users. GA users will see all users.



8.1.1 Access Type and Implicit Permissions

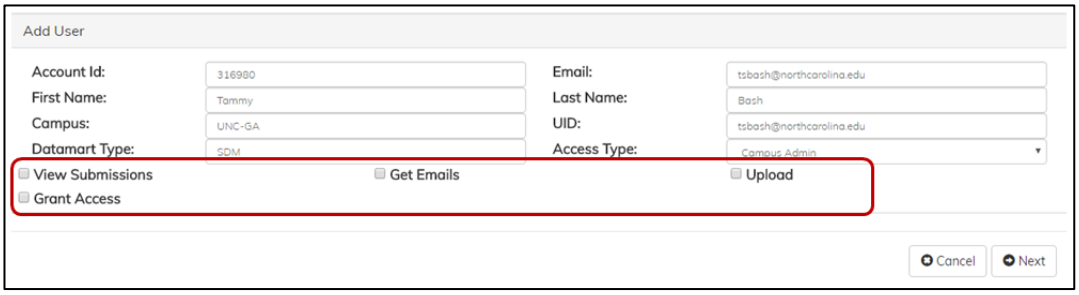
Access Type identifies the Campus Admin or Campus General role and implicit permissions:



Campus Admin	Has unit level access to all datasets, can provision and grant access to SDM.
Campus General	Has aggregate access to datasets released by GA Admin
GA Admin	Has unit level access to all datasets, can provision and grant access to datasets, and can view and edit closed snapshots.
GA General	Has aggregate access to datasets released by GA Admin.

8.1.2 Explicit Permissions

In addition to the Implicit Permissions assigned by role or access type, Explicit Permissions also identify which users on campus have specific permissions:



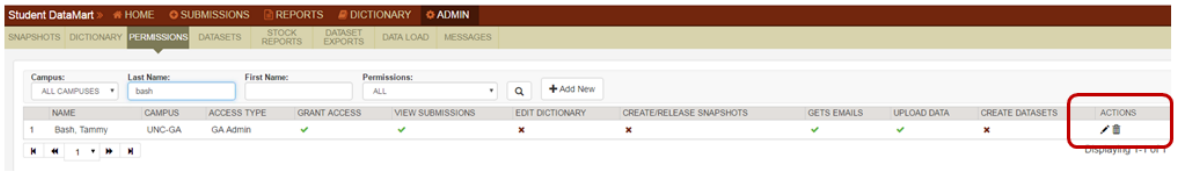
Grant Access	Can grant or edit access to SDM for other campus users.
View Submissions	Can view the submissions tab and Verification Reports.
Edit Dictionary	(GA Admin Only) Can edit the data dictionary.
Create Release Snapshots	(GA Admin Only) Can create snapshots and release them for access.
Get E-mail	Receives e-mail notifications about submissions.
Upload Data	Has access to the reload button in submissions tab in order to re-submit a snapshot.
Create Datasets	(GA Admin Only) Can create new datasets by limiting columns of existing datasets.

8.2 Manage User Permissions - Existing User

Campus Admins with the explicit permission for Grant Access can delete or modify existing users.

8.2.1 Modify Permissions

To modify permissions of an existing user, find the user from Admin->Permissions and select the pencil (edit) icon under Actions column.



See the description of Implicit and Explicit Permissions above for the effect of modifying permissions.

Note: If the user is logged in when the permissions are modified, s/he must log out and then back in in order for the changes to take effect.

8.3 Manage User Permissions - Add a New User

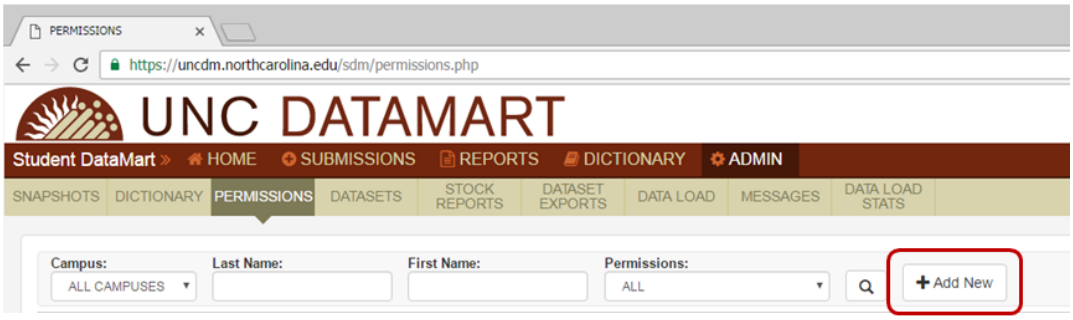
Campus Admins with the explicit permission for Grant Access can add new users to the SDM.

1. Ensure the user is qualified for access to student data per the University's FERPA Guidelines and Data Security policies.

2. Have the user attempt a login to Production and/or QA depending on the SDM environment to which s/he will be added. The SDM uses a federated identity protocol, Shibboleth. New users must first attempt a login to https://uncdm.northcarolina.edu for Production or to https://devuncdm.northcarolina.edu/QA for QA. This identifies the user to SDM and permits the Campus Admin to add permissions to the account.

3. From the Admin->Permissions tab, permissioned users are displayed in alphabetical order. Campus users will only see designated campus users. GA users will see all users.

4. Select +Add New



5. Search for the user by entering part or all of the Last and/or First names.

If there are no results for the search or if the person you are searching for is not listed, then try a broader name search. For example, if you tried searching Last Name = Smith and First Name = John. Try searching just last name Smith in case the person's formal first name is Jon or something other than John. If the person is still not found in the results, then confirm that s/he has indeed attempted the login to the correct URL (uncdm.northcarolina.edu or devuncdm.northcarolina.edu/QA). If everything appears to have been done correctly, submit a ticket and explain that the user has attempted a login but is not found in the list of available users.

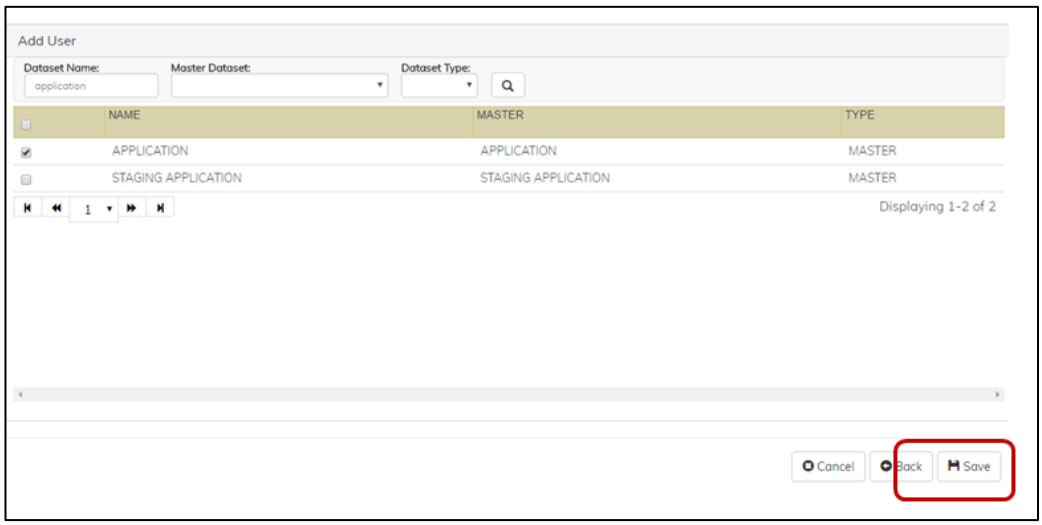
If the user is listed more than one time, then pick the proper affiliation in this order:

i. Staff
ii. Member
iii. Employee
iv. Affiliate

6. Select the plus key + to add permissions to the user.
    a. Select Access Type from the drop down. See above for the description of Implicit Permissions.
    b. Select Explicit Permissions by checking the applicable boxes - See above for the description of Explicit Permissions.

7. Specify datasets the user can access. Dataset access can be modified after initial setup using the Admin->Datasets functions described later.
    a. Admin Access Type: If the access type of Campus Admin or GA Admin is selected, then a warning displays indicating the user will have access to all master datasets. Master datasets are unit level and contain all possible data elements throughout all history that is available in SDM.
    b. General Access Type: Search for the desired dataset(s). Selecting the magnifying glass with no filters will display all datasets or filters can be added to restrict the list. Select a dataset(s) to add by checking the box to the left of the dataset you wish to add. To add all of the datasets in the list, click the checkbox at the top on the left.

8. Select Save to complete adding the user.

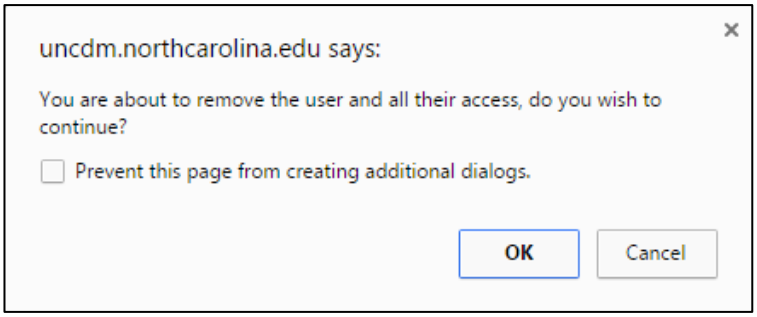


8.4 Manage User Permissions - Remove User Access

Campus Admins with the explicit permission for Grant Access can add new users to the SDM.

From the Admin->Permissions tab, search for the user to be removed. Click the trash can (remove) icon under the Actions column.

When you remove a user, the following warning appears. You cannot undo the action to remove someone once you click "OK." However, you could re-add the user by going back through the add user process.



8.5 Manage User Permissions - Dataset Access

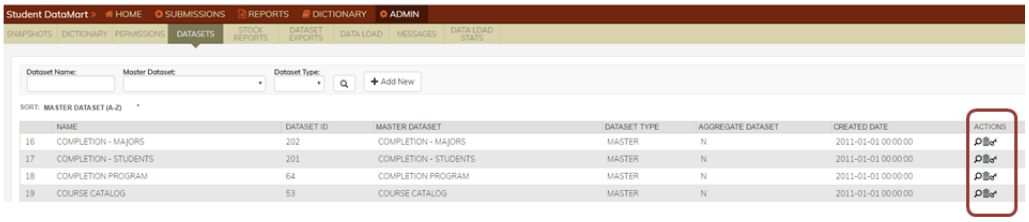
Campus Admins with the explicit permission for Grant Access can add or remove user access to Datasets.

Admin Users have access to all Master Datasets.

General Users must have access granted to specific datasets. This allows SDM users unit level or aggregate access only to information specific to job-related functions.

Granting a user access to datasets can happen during initial user setup or at a later time. See Manage User Permissions - Add New User for steps to add dataset access during initial user setup. To add or remove dataset access for existing users:

1. From the Admin->Datasets tab, find the Dataset to which permissions will be added or removed. The full list will display by default or the search filters can be used to restrict the list displayed.
2. Optionally, View information about the Dataset, including current user access, by clicking the magnifying glass. Close the View window.
3. Select the Key (Permissions) icon under the Actions column. The full list of users will appear in alphabetical order. The list can be searched or filtered using the Magnifying Glass.



4. To grant permissions to a user or multiple users, click the check box on the far left.
5. To remove permissions from a user or multiple users, unclick the check box on the far left.
6. Select Save.

SDM has three dataset types:

Master Dataset	Includes all fields for all time available in SDM. This dataset is not limited in anyway and is the basis for all Core and some Filtered datasets.
Core Dataset	Includes a restricted set of fields and contains data for all time available in SDM. This dataset type can be thought of as restricting by columns or field. For example, you might want to create a dataset that has removed names and IDs and other identifying information to try to protect the identity of individual students. This would be a Core dataset.
Filtered Dataset	Includes either all (when based on a Master dataset) or a restricted (when based on a Core dataset) set of fields based on selected filters. For example, a user may only be interested in looking at students majoring in programs offered by the business department and the dataset can be filtered by department.

Staging datasets are master datasets that are based on the most recent submission. These allow users to view data submitted to SDM prior to migration into the full reporting datasets.

Glossary

Aggregate- (as opposed to "unit level") is a summarized look at data. Aggregate information is often displayed in counts or percentages and does not show data at the individual or unit level.

Auto-Populate- (automatically populated) refers to fields or forms being filled for the user (not by the user) from information previously provided.

Campus Admin- is an access type that has the implicit permissions of unit level access to all datasets, can provision and grant access to datasets.

Campus General- is an access type that has the implicit permissions of aggregate access to datasets released by GA Admin.

C**an Create and Release Snapshots**- is an explicit permission that is only available to the GA Admin access type and allows the user to create snapshots and release them for access.

Can Create Datasets- is an explicit permission that is only available to the GA Admin access type and allows the user to create datasets (limit by columns).

Can Edit Dictionary- is an explicit permission that is only available to the GA Admin access type and allows the user to edit the data dictionary.

Can Grant Access- is an explicit permission that allows the user to grant access and give or take explicit permissions to others.

Can Upload- is an explicit permission that allows the user to have access to the reload button in submissions tab.

Can View Submissions- is an explicit permission that allows the user to view the submissions tab.

Collection- is a grouping a data feeds. The Collections defines the data feeds that will be pulled. Each snapshot type has a corresponding collection. For example, the Census snapshot type has a corresponding collection of census. If you wanted to pull all of the feeds that you typically would for census on an ad hoc basis, you would choose the snapshot type of Ad hoc but choose the collection of census.

Core Dataset- includes a restricted set of fields and contains data for all time available in SDM. This dataset type can be thought of as restricting by columns or field. For example, you might want to create a dataset that has removed names and IDs and other identifying information to try to protect the identity of individual students. This would be a Core dataset.

Daily Reports- are created from aggregate (summary) data, such as counts/ frequencies, saved off of the nightly reload from each snapshot. These will be greatly expanded for production, but there are currently a few examples. The purpose of these reports is to show how counts change from the open of a snapshot to the close. These reports can also be used for year to year comparisons to answer questions such as: "How does student enrollment this year compare to enrollment at the same time last year?" These reports are found under the 'Daily' tab on the reports page.

Data Feed (Feed)- is a grouping of data transferred to and from a Campus, UNC GA, or a Data Broker. Data from the data feeds are combined in various ways to create datasets.

Dataset- (called a "view" in Oracle) is a grouping of data, raw or derived, from underlying database tables that would be used to support views, stock reports, custom reports, ad-hoc reports, dashboards, heat maps, etc. Datasets are groups of data that are used for reporting. Data from the data feeds are combined in various ways to create datasets.

Dataset Creation- making a core or filtered dataset from another dataset (master, core, or filtered).

Dataset Provisioning- involves granting permissions or access to a dataset.

Drop-down- is a menu that is viewable by pressing an arrow and contains a pre-defined set of choices from which users can select. For example:

Explicit Permission- is a kind of access that must be expressly given. Each access type can have different explicit permissions. For example, Get E-mail is an explicit permission that users of any access type can be given.

Field (in a data feed or dataset)- is an individual data element (e.g. Last Name).

Filter- limits the information provided by a specified set of criteria. For example, a dataset filtered by the gender of female would contain only females in the results.

Filtered Dataset- includes either all (when based on a Master dataset) or a restricted (when based on a Core dataset) set of fields based on selected filters. For example, someone might only be interested in looking at students who majored in programs offered by the business department, so you would filter your dataset by department.

Final Pull Date- is the date that the snapshot will be frozen and must be free from errors. For example, the census snapshot will have a final pull day of census day (generally defined as the 10th class day).

Flat reports- display unit (e.g. student, instructor, course, etc.) level information. You must have been given permissions that enable you to see information about individuals or items within a dataset.

GA Admin- is an access type that has the implicit permissions of unit level access to all datasets, can provision and grant access to datasets, and can view and edit closed snapshots.

GA General- is an access type that has the implicit permissions of aggregate access to datasets released by GA Admin.

Get E-mail- is an explicit permission that allows the user to receive E-mail notifications about submissions, etc.

Implicit Permission- is a kind of access that is contained within an access type. It is a permission shared by all individuals with e certain access type and is inherent to the access type. For example, an implicit permission for the Campus Admin access type is unit level access to all master datasets.

Initial Pull Date- is the date where the first test pull for that snapshot is transmitted. The snapshot will refresh (re-pull) every night until the final snapshot pull date. The initial pull date may be days or months before the final pull date.

Master Dataset- includes all fields for all time available in SDM. This dataset is not limited in anyway and is the basis for all Core and some Filtered datasets.

Parallel Reports - are used to look at data being run during the parallel process for comparison of data between the traditional flat file reporting and the data being reported using the Student Data Mart process.

Permissions- are related to security. Permissions refer to the type or level of access that a user has within SDM.

Pivot tables- display only aggregated (summarized) information. Pivot tables allow you to sort, count, total or give the average of data in one table or spreadsheet.

Production tables- are where data is stored once they are free from errors. Once the data are free from validation errors, they can be moved into the production tables. Data can be moved into production when there are warnings present but not when there are errors. However, all warnings should have been addressed to ensure that they are not the result of incorrect data. The production tables are frozen at the end of the final pull date and that is what is used for official reporting.

Reporting Widget- is a tool that is used for ad hoc reporting in SDM and contains a variety of datasets from which to report. The widget contains master, core, and filtered datasets.

Shibboleth- is an open-source project that provides Single Sign-On capabilities and allows sites to make informed authorization decisions for individual access of protected online resources in a privacy-preserving manner.

Snapshot- is a look at data from a particular point in time. Snapshots are frozen at the final pull date and are never updated to a more recent time.

Snapshot Status - There are four different statuses at the snapshot level:

- **Opened** - The campus snapshot is within the initial pull and final pull dates and is being refreshed nightly.
- **Closed**- The final pull date has come and no new data is being accepted. The snapshot is no longer being refreshed nightly. Once closed, only Campus Admin and GA Admin will be able to see the data from the snapshot. The snapshot must be released before other access types can view the data.
- **Released**- The snapshot has been validated and released for campus and UNC GA reporting. UNC GA Institutional Research releases the snapshot.
- **Future**- The campus snapshot has been scheduled but is not yet within the initial pull and final pull date period.

Snapshot Type- conveys the purpose of the snapshot. Snapshot Type designates whether the snapshot is being used for official reporting purposes (Census, Beginning of Term, End of Term, Post Grades); whether the snapshot is being pulled on an informal, impromptu bases (ad hoc); or whether the snapshot is being pulled for the purposes of generating transcript-like data.

Staging tables- are a temporary holding place for data that contain errors. When the data initially enter the SDM, they are first put into staging tables. The data can reside in the staging tables even when the submission has validation errors. . However, there are some critical errors (any SQL error) that would result in the data not even being able to enter the staging tables. Some examples of such errors include:

- Incorrect data type (e.g. a letter in a numeric field)
- Account locks
- Subqueries returning more than one distinct row
- Values exceeding the designated max lengths

Validations are run off of the data in the staging tables, so if the data cannot be loaded into staging then validations cannot be run. In that case, you would see zero validation errors/ warnings but a red status.

Submission- is a transmission of data for the purpose of creating a snapshot. Multiple submissions will be made but only the submission on the final pull date will be saved as a snapshot.
Submission Window - The period of time the snapshot is open - between the initial load date/time and the final due date/time.

Unit Level- (as opposed to "aggregate") is the most base/ foundational level of viewing data. Viewing data at the unit level involves being able to see down to the level of each individual unit of observation or element in the data. For example, a unit level dataset of course sections offered by a university would contain one row per course section.

Validation Reports- provide summary information about your data in order to detect potential issues. Just because a submission has no errors or warnings on the submissions page does not mean that the data being sent is accurate. The purpose of these reports is to compare counts and other summary values against other university

reporting to confirm its accuracy. Any deviation from what is expected should be investigated before the close of the snapshot.

Web Services- is a method of communication and transmission of data between two secured entities. The service is a software function provided at a network address over the web to query data from one entities resources and then package that data into a JSON array and transmit it back to the initiating entity to process.

Appendix A – SDM Data Feeds by Snapshot Type

	Census	Beginning of Term (BOT)	End of Term (EOT)	Post-Grades (Completion)	Financial Aid	Reverse Transfer*	Exclusions Only	Program Only	Crosswalk Only	Term Only
Application	x	x	x	x						
Basic Instructor	x	x	x	x						
Basic Student	x	x	x	x	x	x				
Career	x	x	x	x		x				
Career Look-up	x	x	x	x						
Career Program	x	x	x	x						
Cohort Exclusion							x			
College	x	x	x	x				x		
Completion				x						
Completion Program				x						
Course Catalog	x	x	x	x						
Credit Awarded - Institutional				x		x				
Credit Awarded - Other	x	x	x	x		x				
Crosswalk									x	
Department	x	x	x	x				x		
Enrollment	x	x	x	x						
Financial Aid Apply	x				x					
Financial Aid Award	x				x					
Instructor Section	x	x	x	x						
Programs	x	x	x	x				x		
Section Meeting	x	x	x	x						
Sections Offered	x	x	x	x						
Session	x	x	x	x						
Term	x	x	x	x	x					x
Tests	x	x	x	x						

* Reverse Transfer snapshot type requires the Post-Grades snapshot for the same term to be closed successfully.

Appendix B – SDM Datasets by Snapshot Type

ID	NAME	OBJECT_NAME	Census	Beginning of Term (BOT)	End of Term (EOT)	Post-Grades (Completion)	Financial Aid	Reverse Transfer*	Exclusions Only	Program Only	Crosswalk Only	Term Only
66	APPLICATION	SDM_APPLICATION_V and SDM_APPLICATION_MV	x	x	x	x						
79	BASIC INSTRUCTOR	SDM_BASIC_INSTRUCTOR_V	x	x	x	x						
78	BASIC STUDENT	SDM_BASIC_STUDENT_V	x	x	x	x	x	x				
323	CAREER ADMIT	SDM_CAREER_ADMIT_V and SDM_CAREER_ADMIT_MV	x	x	x	x						
58	CAREER PROGRAM	SDM_CAREER_PROGRAM_V	x	x	x	x						
57	CAREER	SDM_CAREER_V and SDM_CAREER_MV	x	x	x	x		x				
405	COHORT EXCLUSIONS	SDM_COHORT_EXCLUSION_V							x			
51	COLLEGE	SDM_COLLEGE_V	x	x	x	x						
202	COMPLETION - MAJORS	SDM_COMPLETION_MAJOR_V				x						
201	COMPLETION - STUDENTS	SDM_COMPLETION_V				x						
53	COURSE CATALOG	SDM_COURSE_CATALOG_V	x	x	x	x						
60	CREDIT AWARDED INSTITUTIONAL	SDM_CREDIT_AWARDED_INST_V				x		x				
61	CREDIT AWARDED OTHER	SDM_CREDIT_AWARDED_OTHER_V	x	x	x	x		x				
204	CROSSWALK	SDM_CROSSWALK_V									x	
52	DEPARTMENT	SDM_DEPARTMENT_V	x	x	x	x						
55	ENROLLMENT	SDM_ENROLLMENT_V	x	x	x	x						
70	FINANCIAL AID APPLY	SDM_FINAID_APPLY_V	x			x	x					
205	FINAID YEAR AWARD DETAIL FULL	SDM_FINAID_YR_AWARD_DTL_FULL_V	x			x	x					
206	FINAID YEAR AWARD DETAIL SHORT	SDM_FINAID_YR_AWARD_DTL_SHORT_V	x			x	x					
56	INSTRUCTOR BY SECTION	SDM_INSTRUCTOR_SECTION_V	x	x	x	x						
59	PROGRAM	SDM_PROGRAMS_V	*	*	*	*				x		
67	SECTION MEETING	SDM_SECTION_MEETING_V	x	x	x	x						
54	SECTIONS OFFERED	SDM_SECTIONS_OFFERED_V	x	x	x	x						
81	SESSION	SDM_SESSION_V	*	*	*	*						
80	TERM	SDM_TERM_V	*	*	*	*	x					x
62	TESTS	SDM_TESTS_V	x	x	x	x						

* These Datasets exist in the snapshot type, but will only contain the most recent or changed information.

Appendix C – How Things Work: The Role of the Aggregator Table

The aggregator table's sole purpose is to provide a temporary storage facility for sets of keys that would be used to identify a population of data when joined with an Oracle View created for a data feed. There are currently three aggregator types identified called "Student", "Instructor" and "TERM_CODE_CAMPUS. The "Student" aggregator type will have an aggregator_key of a student PIDM. The "Instructor" aggregator type will have an aggregator_key of an instructor PIDM. The "TERM_CODE_CAMPUS" aggregator type will have an aggregator_key of campus term code for one snapshot run.

Below is a table that describes the web services that will rely on the aggregator table as well as the web services that will populate the aggregator table. The order in which web services are initiated and processed is key in this whole process and will be managed by the code UNC-GA will run when requesting a snapshot of data. The snapshot of data can consist of web service calls to pull data for one or more data feeds. The first column designates the Oracle View that will need to join with the aggregator table from the web service code. Whether the transmission type set is implicit or explicit, the PIDMS will always be populated in this aggregator table and so these Oracle views should always join with the aggregator table in the web services code based on the data element in the last column of this table below. The web services that rely on the aggregator will be processed last in the process of requesting data to allow for the web services specified in the third column below to populate the aggregator table with aggregator keys that will eventually identify the population they need to pull.

View Joining to Aggregator	Aggregator Type	Web service calls that will populate the aggregator if called by snapshot	Data Feed Element to insert into Aggregator
UNCSDM_BASIC_STUDENT	Student	For Reverse Transfer Snapshot: clearAggregator For All Other Snapshots: GetEnrollment, GetCreditAwardInst, GetCompletion,GetCompletionProgram, GetApplication,GetFinaidAward, GetCohortExclusion	PIDM
UNCSDM_BASIC_INSTRUCTOR	Instructor	GetInstructorSection	PIDM
UNCSDM_TESTS	Student	GetEnrollment, GetCreditAwardInst, GetCompletion,GetCompletionProgram, GetApplication,GetfinaidAward	PIDM
UNCSDM_CREDIT_AWARDED_OTHER	Student	For Reverse Transfer Snapshot: clearAggregator For All Other Snapshots: GetEnrollment, GetCreditAwardInst, GetCompletion,GetCompletionProgram, GetApplication,GetFinaidAward	PIDM
UNCSDM_FINAID_APPY	Student	GetEnrollment, GetCreditAwardInst, GetCompletion,GetCompletionProgram, GetApplication,GetFinaidAward	PIDM

Feeds that are currently using the campus term code from aggregator table are listed below:

View Using Term Code from Aggregator	Aggregator Type	Web service calls that will populate the aggregator if called by snapshot	Data Feed Element to insert into Aggregator
UNCSDM_BASIC_STUDENT	TERM_CODE_CAMPUS	GetCampusTerms	TERM_CODE_CAMPUS
UNCSDM_BASIC_INSTRUCTOR	TERM_CODE_CAMPUS	GetCampusTerms	TERM_CODE_CAMPUS
UNCSDM_CAREER	TERM_CODE_CAMPUS	GetCampusTerms	TERM_CODE_CAMPUS
UNCSDM_COMPLETION	TERM_CODE_CAMPUS	GetCampusTerms	TERM_CODE_CAMPUS
UNCSDM_COURSE_CATALOG	TERM_CODE_CAMPUS	GetCampusTerms	TERM_CODE_CAMPUS

Appendix D – How Things Work: Data Flow Through Staging and Production Tables

When a snapshot is loaded, web services are initiated and data are pulled from the campus ERP. The high level process is:

1. Data enter shadow tables and staging tables
2. Validations (warnings and errors) are run against the staging tables
3. If error free, then the data move into Production tables.

Definitions of Shadow, Staging and Production tables for SDM:

Shadow tables - Shadow tables are replicas of the data exactly as pulled from the ERP. Specifically, the field names and valid values from dimension tables will match the Datafeeds and Dimensions specifications. Shadow tables are the basis for drill down reports available on the Submissions page.

Staging tables - Staging tables are similar to Shadow tables in that they are a pre-migrated copy of the submitted data. The difference from the Shadow tables is that the Staging tables use database IDs as opposed to literal codes. Translating shadow tables to staging tables speeds the migration process into production tables as opposed to moving directly from shadow to production.

Data can reside in the staging tables even when the submission has validation errors. However, there are some critical errors (any SQL error) that would result in the data not even being able to enter the staging tables. Some examples of such errors include:

- Incorrect data type (e.g. a letter in a numeric field)
- Account locks
- Subqueries returning more than one distinct row
- Values exceeding the designated max lengths

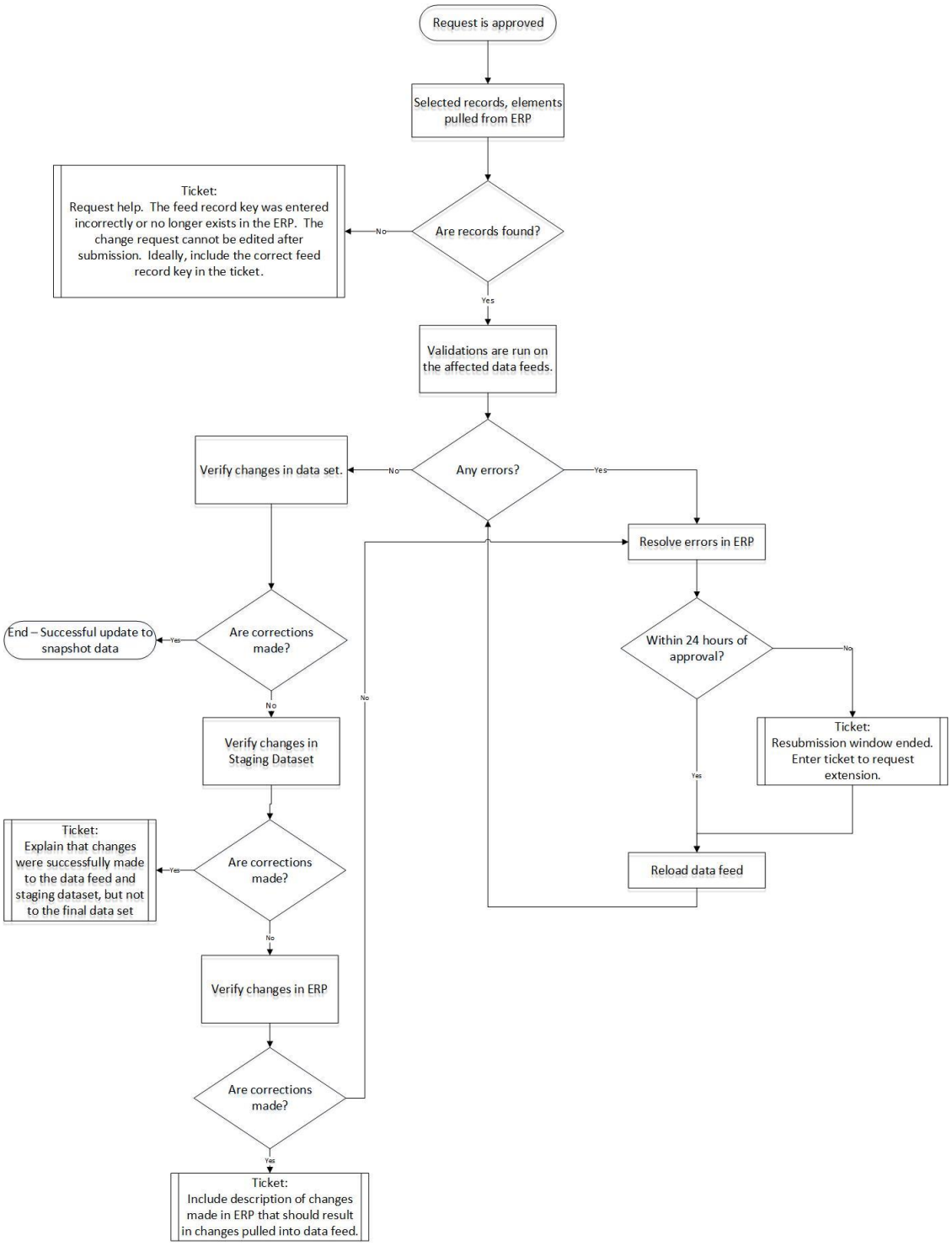
Validations are run off of the data in the staging tables, so if the data cannot be loaded into staging then validations cannot be run. In that case, you would see zero validation errors/ warnings but a red status for the feed.

Each time the snapshot runs, data in the staging tables is deleted and repopulated. Only some aggregate data is saved on a nightly basis.

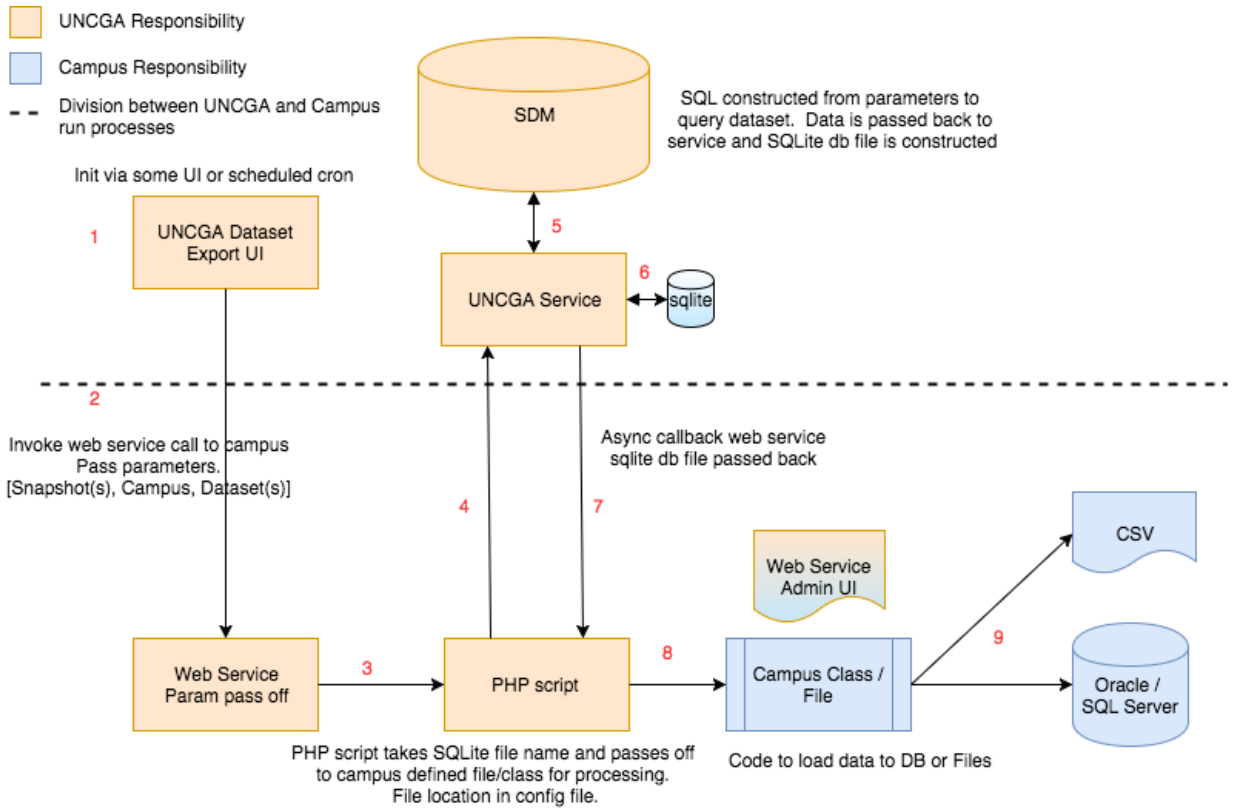
Production tables - Once the data are free from validation errors, they move into the production tables. While the snapshot is open, production tables are the basis for preliminary reports and preliminary dataset downloads used in data verification and in preliminary reporting. When the snapshot closes, the production tables are frozen and become the basis for System reporting.

Appendix E – How Things Work: Approved Snapshot Change Requests

Snapshot Change – Approved Request



Appendix F – How Things Work: Dataset Download as a Service



1. Campus initiates request via SDM UI
2. Web service invoked to pass parameters to campus
3. PHP script invoked with params to call web service to get the data
4. Async web service is called from campus to GA to get sqlite db file of data
5. Service pulls data from SDM using parameters and creates sqlite db file
6. Sqlite db is created and saved to server
7. Sqlite db is passed back to campus
8. Sqlite db is handed off to custom class file
9. Custom class file queries data from sqlite and creates files or loads data directly to another DB

NOTES:
An override class file defined in the config file would be used to pass off the data for the campus to then take over the processing. UNCGA will create a template custom class file that will query the sqlite db file per dataset and move the data into an oracle database table per dataset.

Appendix G – Additional Resources and Helpful Links

UNC Student Data Mart Project and System Documentation

Project documentation, release guides and related files are maintained on the SDM site. Anyone with UNC System credentials can access the site.
https://collab.northcarolina.edu/public/index.php?path_info=projects/student-data-mart/files

UNC System Institutional Research (IR) Reporting Calendar

The academic year reporting calendar, including applicable SDM or HRDM snapshots:
https://collab.northcarolina.edu/public/index.php?path_info=projects/student-data-mart/files/files/66671

UNC SDM Snapshot and Code Release Activities

Snapshot schedule and SDM Version release calendar.
https://collab.northcarolina.edu/public/index.php?path_info=projects/student-data-mart/files/files/53157

UNC Information Center

Dashboards and reports that display UNC System data. http://northcarolina.edu/infocenter