# SQL Data Staging Area - Degree Audit Data

The SQL database that stores production Degree Audit data was designed by SBCTC-ITD staff. This database contains a combination of data created specifically for the Degree Audit application and data that is imported from the HP-UX Eloquence environment.

Meta data has been created to provide descriptions on the tables and fields in the Degree Audit database. Refer to the spreadsheet ***DSAs_DegreeAudit_MetaData*** for this information.

## Degree Audit ETL Process

The data that exists in the Degree Audit system that is populated from data found in the Eloquence database on the legacy HP-UX environment is loaded using an Extract, Transform and Load (ETL) process. Here is the schedule of that process:

- Course and catalog data is updated daily; the process is completed by 7:30 a.m. for all schools.
- Student data is loaded twice each day; the process is completed by 8:00 a.m. and 2:00 p.m. for all schools.

### Characteristics of Degree Audit SQL Data that is loaded by the ETL Process

There are some common characteristics you will find with the data in the tables that are populated from ETL process:

- These tables use the same name as the table name is in the Eloquence environment:
  - The SQL environment uses an underscore (_) not a dash (-) as the separator in the names.
  - The suffix '_M' and '_D' at the end of the SQL table name represents the characteristic of the corresponding table in the Eloquence database. *_M* represents that the table was a Master dataset and *_D* represents that the table was a Detail dataset in Eloquence.

- Concatenated fields are separated into their sub-fields

- Although the schemas for the tables contain every field that is found in the legacy table, these tables in the SQL environment have a limited number of fields populated. Expect to find many columns in each table that are not populated. This is because the ETL process only includes the exact data items that are necessary for the Degree Audit process to function.
  - It is recommended to select only the fields that are populated in your data extracts

- Many character fields are defined using the data type *varchar*. Since the Eloquence data uses a fixed length field to store the data, the data that is provided by the ETL process fills the entire field length. You can expect to find any non-used spaces padded with blank spaces.
  - The additional blank spaces can be removed by using the SQL "right trim" function when loading the extracted data at your campus

- There are fields in the tables that are populated by the ETL process that are not part of the HP-UX legacy data but are populated in the SQL environment:

- **IMAGE_RECNBR** – This is used to create a primary key in the SQL environment and represents the order that a row was added to the table. Note that the IMAGE_RECNBR value from table to table does not match up so these fields should not be used to link data from one table to another table

- **UPDATE_TIMESTAMP** – Indicates the most recent date and time that the row of data was updated. This can be useful to determine what data has changed and eliminate the need to extract all of the data in a table for each download process

- **cs_*table_name*** – This is used for an internal operation in the ETL process and has no value for your data downloads