



Office of the President

**Institutional Research and
Academic Planning (IRAP)**

UCDW Student Enrollment User Guide

THIS PAGE HAS BEEN INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

1.	INTRODUCTION.....	7
1.1	Overview of Student Enrollment	7
1.2	Student Privacy	7
1.3	How Student Enrollment Data in UCDW Is Stored?.....	7
1.4	Important points to know when using UCDW Student Enrollment.....	9
2.	Accessing UCDW Student Enrollment.....	10
3.	UCDW Student Enrollment Tables Guide	10
3.1	UCDW Student Enrollment Star	10
3.2	Summer Enrollment Star.....	11
3.3	Details on Fact and Dimension Tables in UCDW Student Enrollment	12
3.3.1	ENROLLMENT_F	12
3.3.2	SUMMER_ENROLLMENT_F.....	13
3.3.3	STUDENT_D.....	13
3.3.4	CAMPUS_LOCATION_D.....	13
3.3.5	STUDENT_LEVEL_D	14
3.3.6	ENROLLMENT_STATUS_D	14
3.3.7	CAMPUS_COLLEGE_MAJOR_CODE_D	14
3.3.8	AGE_BAND_D.....	14
3.3.9	ACADEMIC_SUB_TERM_D	15
3.3.10	ACADEMIC_DEGREE_D	15
3.4	DataMart in UCDW Student Enrollment.....	15
3.5	MQTS in UCDW Student Enrollment.....	16
3.5.1	ENROLLMENT_HEAD_COUNT_M.....	16
3.5.2	SUMMER_ENROLLMENT_HEAD_COUNT_M	18
4.	Sample Queries	18
5.	Terms and Abbreviations	18

TABLE OF TABLES

Table 1 : Enrollment files from UC campuses 8
Table 2: Glossary 18

TABLE OF FIGURES

Figure 1: Three-Tier UCDW architecture 8
Figure 2: Current UCDW Infrastructure 9
Figure 3 : Fact and dimension tables in UCDW Student Enrollment Star 11
Figure 4: Fact and dimension tables in UCDW Summer Enrollment star 12
Figure 5: Enrollment Fact Row Counts..... 17
Figure 6: Enrollment Head Count in MQT..... 17

THIS PAGE HAS BEEN INTENTIONALLY LEFT BLANK

1. INTRODUCTION

1.1 OVERVIEW OF STUDENT ENROLLMENT

UCDW-Student Enrollment data warehouse stores data related to 3rd week and end of term enrollments for (fall, winter and spring terms), summer enrollments (both 3rd week and EOT), and medical resident enrollments. UC's Student Enrollment data is consolidated into one enterprise-wide system, and will help improve accuracy, timeliness, flexibility, and reporting of Student enrollment data. UCDW-Student Enrollment provides system-wide reporting data for student enrollments across all 10 UC campuses. The data in UCDW-Student Enrollment is dependent on the submission of campus 3rd week, end of term, medical resident and summer enrollment files.

1.2 STUDENT PRIVACY

Personally identifiable information for education records includes direct identifiers, such as a student's name or identification number, indirect identifiers, such as a student's date of birth, or other information which can be used to distinguish or trace an individual's identity either directly or indirectly through linkages with other student information.

Please remember that student data is covered under UC's privacy policy and FERPA. UCDW users will be required to adhere to all Federal, State and local statutes, regulations, and other requirements pertaining to the security, confidentiality, and privacy of data provided by UCOP including, but not limited to, the Family Educational Rights and Privacy Act (FERPA), the California Information Practices Act, and the Privacy Act of 1974, as amended.

Proper security rules will be enforced and access will be restricted only to qualified users and permissions will be appropriately assigned based on user-access groups.

1.3 HOW STUDENT ENROLLMENT DATA IN UCDW IS STORED?

All 10 UC campuses submit student enrollment data to UCOP according to the file specifications and the file submission deadline. Berkeley and Merced are on the semester system whereas all other campuses are on the quarter system. Berkeley and Merced do not submit files for the winter quarter cycle.

	Fall		Winter		Spring		Summer	
	3WK	EOT	3WK	EOT	3WK	EOT	3WK	EOT
Berkeley	Y	Y	N	N	Y	Y	Y	Y
San Francisco	Y	N	Y	N	Y	N	N	N
Davis	Y	Y	Y	Y	Y	Y	Y	Y
Los Angeles	Y	Y	Y	Y	Y	Y	Y	Y
Riverside	Y	Y	Y	Y	Y	Y	Y	Y
San Diego	Y	Y	Y	Y	Y	Y	Y	Y
Santa Cruz	Y	Y	Y	Y	Y	Y	Y	Y
Santa Barbara	Y	Y	Y	Y	Y	Y	Y	Y

	Fall		Winter		Spring		Summer	
	3WK	EOT	3WK	EOT	3WK	EOT	3WK	EOT
Irvine	Y	Y	Y	Y	Y	Y	Y	Y
Merced	Y	Y	N	N	Y	Y	Y	Y

Table 1 : Enrollment files from UC campuses

File specifications and the file submission deadline for enrollment data are available in this [link](#).

UCOP-UCDW infrastructure stores information in three layers-stage, base and business intelligence (BI).

All 10 campuses’ student enrollment related data is first stored in staging layer which is a parking lot for all the campus enrollment data. Staging data is transformed if necessary and stored in base layer. Staging and base data is validated and certified by data stewards in IRAP and then stored in BI layer. Canned reports for business use are generated from BI layer data.

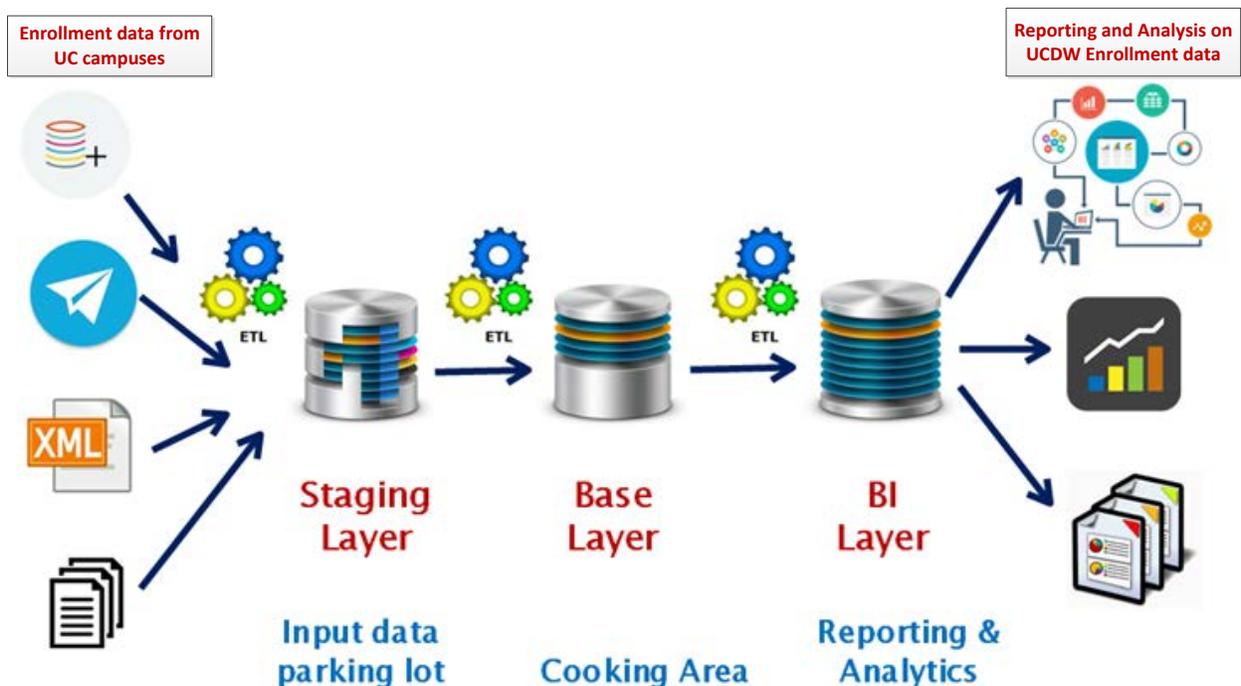


Figure 1: Three-Tier UCDW architecture

For analyses, analytical and reporting purposes, business users use data in BI layer. BI layer data is stored as star schema, one of the styles of dimensional data modelling. Star schema provides important benefits to the user: improved understandability of the data, performance and enhanced reporting capability. Star schema consists of one or more fact tables referencing any number of dimension tables.

When thinking about data and the way it is analyzed, the fact tables provides measurements and the dimension tables provide context. For example, in a data report the dimension tables supply the values for the row and column headings as well as the constraint filters and grouping capabilities. The fact table supplies the measures, or what is being counted.

Refer to sections 3.1, 3.2 and 3.3 for details on the fact and dimension tables in UCDW-Student Enrollment.

Enrollment Data Mart is a subset of the enrollment data in UCDW. Usually data marts are oriented to a specific business need. Data marts are easily accessible by business users, because they don't require queries that are too complicated or resource-consuming.

Refer to section 3.4 for details on the data mart in UCDW-Student Enrollment.

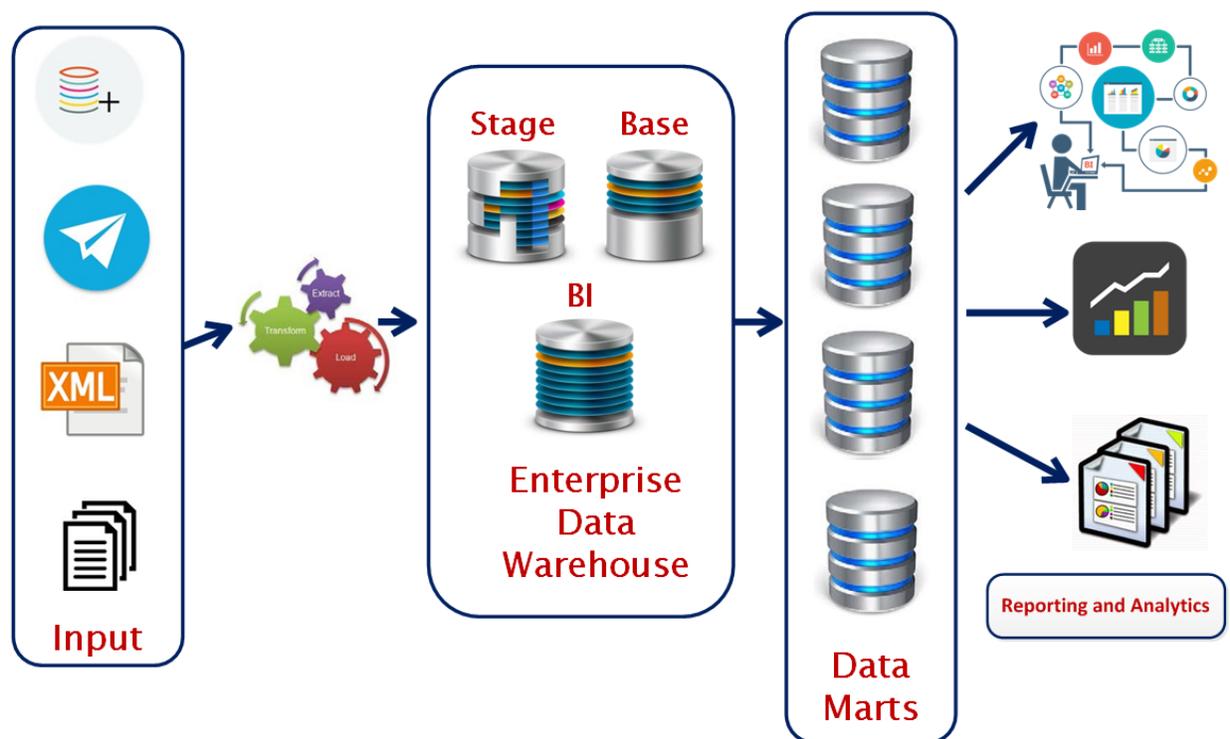


Figure 2: Current UCDW Infrastructure

1.4 IMPORTANT POINTS TO KNOW WHEN USING UCDW STUDENT ENROLLMENT

- i. 3rd week and EOT data for fall, winter, spring and summer terms from applicable UC campuses are stored in Student Enrollment star.
- ii. Medical resident enrollment data from UC campuses (Berkeley, Davis, Irvine, Los Angeles, San Diego, Riverside and San Francisco) are stored in Student Enrollment Star.
- iii. Summer de-duplicated data i.e. consolidation of summer data with student's home campus information is only available in Summer Enrollment Fact.
Home Campus – Campus where a student is enrolled for fall or winter or spring term

- Host Campus – Campus where a student is enrolled for summer term that is not their home campus.
- iv. To report on fall, winter and spring term, use Student Enrollment star.
 - v. To report on summer term with integrity to student’s home campus information, use Summer Enrollment Star.
 - vi. To report on summer term with respect to student’s host campus information, use Student Enrollment star.
 - vii. To report on all terms in an academic year, use Student Enrollment star for fall, winter, spring terms and use Summer Enrollment star for summer term.
 - viii. Do’s and Don’ts on querying UCDW is available via this [link](#).

2. ACCESSING UCDW STUDENT ENROLLMENT

The following are the different ways to access UCDW Student Enrollment data:

- i. UCDW Student Enrollment data can be accessed using the SQL case tool, Db Visualizer. Instructions on how to connect to UCDW through the tool are available via this [link](#).
- ii. Canned and validation reports generated from UCDW Student Enrollment are available in Cognos. Instructions on how to connect to Cognos and access reports, are available under **‘UCDW and Cognos’** section on ‘How-To’s’ in ‘Support & Training’ tab of Data Operations Hub. It can be accessed via this [link](#).
- iii. UCDW Student Enrollment data can be accessed via Tableau. Instructions on how to connect to UCDW through Tableau are available under **‘UCDW and Tableau’** section on ‘How-To’s’ in ‘Support & Training’ tab of Data Operations Hub. It can be accessed via this [link](#).
- iv. UCDW Student Enrollment data can also be accessed using other data analysis and mining tools such as SAS.

3. UCDW STUDENT ENROLLMENT TABLES GUIDE

3.1 UCDW STUDENT ENROLLMENT STAR

Student Enrollment star has Enrollment Fact table and many dimension tables. These tables are stored in STUD_BI schema of uc_dss database in the production server.

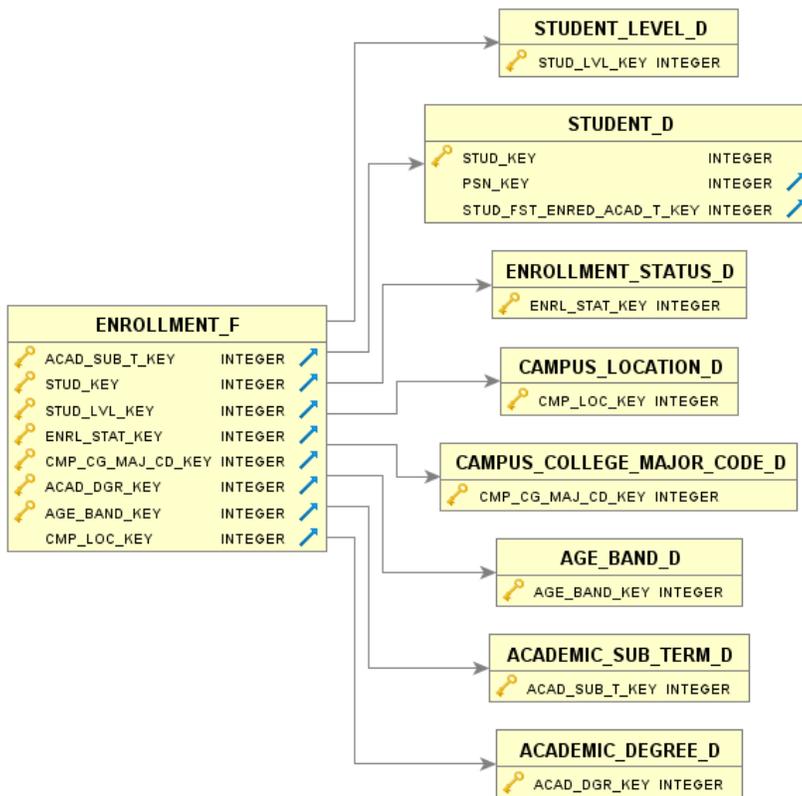


Figure 3 : Fact and dimension tables in UCDW Student Enrollment Star

Note the fact tables are suffixed with _F and dimension tables are suffixed with _D.

3.2 SUMMER ENROLLMENT STAR

Summer Enrollment star has Summer Enrollment Fact table and many dimension tables. These tables are stored in STUD_BI schema in uc_dss database in dwp2 server.

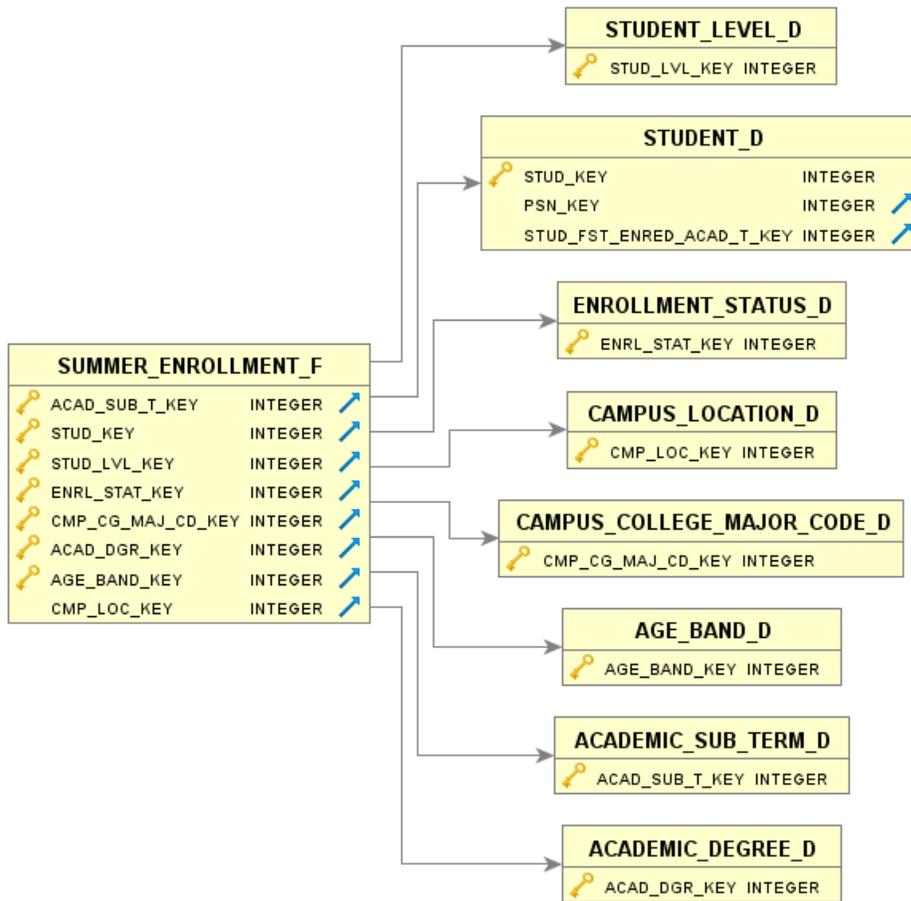


Figure 4: Fact and dimension tables in UCDW Summer Enrollment star

Note the fact tables are suffixed with _F and dimension tables are suffixed with _D.

3.3 DETAILS ON FACT AND DIMENSION TABLES IN UCDW STUDENT ENROLLMENT

3.3.1 ENROLLMENT_F

Description: Enrollment fact table tracks enrollments reported by campuses in the 3rd week and end of term (EOT).

Grain (lowest level at which data is captured): One row per student, per academic sub term (3WK or EOT) in a given year for each combination of enrollment status (P, N, C, R, V or W), degree objective code and major code.

Primary Key: Primary key of the fact table is a composite key made up of foreign keys to its dimension tables. Combination of ACAD_SUB_T_KEY, STUD_KEY, STUD_LVL_KEY, ENRL_STAT_KEY, CMP_CG_MAJ_CD_KEY, ACAD_DGR_KEY and AGE_BAND_KEY is primary key to ENROLLMENT_F table.

3.3.2 SUMMER_ENROLLMENT_F

Description: Summer Enrollment fact table tracks summer enrollments reported by campuses in the 3rd week and end of term (EOT). For students taking courses at one or more host campuses the measures are merged and represented under their home campus.

Grain (lowest level at which data is captured): Students who have enrolled in summer will have one row per student, per academic sub term (3WK or EOT) in a given year for each combination of enrollment status (P, N, C, R, V or W), degree objective code and major code.

Primary Key: Primary key of the fact table is a composite key made up of foreign keys to its dimension tables. Combination of ACAD_SUB_T_KEY, STUD_KEY, STUD_LVL_KEY, ENRL_STAT_KEY, CMP_CG_MAJ_CD_KEY, ACAD_DGR_KEY and AGE_BAND_KEY is primary key to SUMMER_ENROLLMENT_F table.

3.3.3 STUDENT_D

Description: Student dimension contains a record for each student that has enrolled or has been awarded a degree.

Student dimension is a conformed dimension that is shared within Student subject area. Also, this dimension is Slowly Changing Dimension (SCD) type 2 i.e., it contains complete history of changes in different attributes stored for a student.

Primary Key: Primary key of the dimension table is a system generated key to uniquely identify a row in a dimension table. STUD_KEY is the primary key for STUDENT_D table.

Natural Key: This is a type of unique key which is formed of attributes within real data. Combination of Student location campus code (STUD_LOC_CMP_CD) and Student id (STUD_ID) is the natural key in STUDENT_D table.

3.3.4 CAMPUS_LOCATION_D

Description: Campus Location dimension identifies the 10 UC campuses.

Campus Location dimension is a conformed dimension that is shared across multiple subject areas. Also, this dimension is SCD type 2 i.e., it contains complete history of changes in different attributes stored for a campus.

Primary Key: Primary key of the dimension table is a system generated key to uniquely identify a row in a dimension table. CMP_LOC_KEY is the primary key for CAMPUS_LOCATION_D table.

Natural Key: N/A

3.3.5 STUDENT_LEVEL_D

Description: Student level dimension contains various categorizations of academic levels such as Graduate, Undergraduate, and Credential and Certificate.

Student level dimension is a conformed dimension that is shared within Student subject area. Also, this dimension is SCD type 1 i.e., the new information overwrites the original information. No history is kept.

Primary Key: Primary key of the dimension table is a system generated key to uniquely identify a row in a dimension table. STUD_LVL_KEY is the primary key for STUDENT_LEVEL_D table.

Natural Key: N/A

3.3.6 ENROLLMENT_STATUS_D

Description: This dimension holds various miscellaneous codes related to registration status of the students.

Enrollment Status dimension is a conformed dimension that is shared within Student subject area. Also, this dimension is SCD type 1 i.e., the new information overwrites the original information. No history is kept.

Primary Key: Primary key of the dimension table is a system generated key to uniquely identify a row in a dimension table. ENRL_STAT_KEY is the primary key for ENROLLMENT_STATUS_D table.

Natural Key: N/A

3.3.7 CAMPUS_COLLEGE_MAJOR_CODE_D

Description: This dimension contains information about majors offered by various colleges in each campus within the UC system.

Campus College Major dimension is a conformed dimension that is shared across multiple subject areas. Also, this dimension is SCD type 2 i.e., it contains complete history of changes in different attributes stored in this dimension.

Primary Key: Primary key of the dimension table is a system generated key to uniquely identify a row in a dimension table. CMP_CG_MAJ_CD_KEY is the primary key for CAMPUS_COLLEGE_MAJOR_CODE_D table.

Natural Key: N/A

3.3.8 AGE_BAND_D

Description: This dimension contains the age ranges required for banding students by age.

Age Band dimension is a conformed dimension that is shared across multiple subject areas. Also, this dimension is SCD type 1 i.e., the new information overwrites the original information. No history is kept.

Primary Key: Primary key of the dimension table is a system generated key to uniquely identify a row in a dimension table. AGE_BAND_KEY is the primary key for AGE_BAND_D table.

Natural Key: N/A

3.3.9 ACADEMIC_SUB_TERM_D

Description: Academic Sub-Term dimension is the master table for all academic term-related conforming dimensions. It identifies the 3rd week and end-of-term of an academic term in each calendar and academic year. The grain is one record per sub-term (3rd week or EOT)

Academic Sub Term dimension is a conformed dimension that is shared within Student subject area. Also, this dimension is SCD type 1 i.e., the new information overwrites the original information. No history is kept.

Primary Key: Primary key of the dimension table is a system generated key to uniquely identify a row in a dimension table. ACAD_SUB_T_KEY is the primary key for ACADEMIC_SUB_TERM_D table.

Natural Key: N/A

3.3.10 ACADEMIC_DEGREE_D

Description: Academic degree dimension contains all types of degrees conferred by UC campuses.

Academic Degree dimension is a conformed dimension that is shared across multiple subject areas. Also, this dimension is SCD Type 2 i.e., it contains complete history of changes in different attributes stored in this dimension.

Primary Key: Primary key of the dimension table is a system generated key to uniquely identify a row in a dimension table. ACAD_DGR_KEY is the primary key for ACADEMIC_DEGREE_D table.

Natural Key: N/A

3.4 DATAMART IN UCDW STUDENT ENROLLMENT

Student Enrollment Data Mart is a subset of data generated from Student Enrollment Star Schema. ENROLLMENT_DM is in IRAP_BI schema in uc_dss database in dwp2 server.

Grain (lowest level at which data is captured): One record per student (primary major regardless of their multiple majors) in fall 3wk sub-term of given academic year.

Primary Key: Primary key of the data mart is used to uniquely identify a row in a data mart table. Combination of ACADEMIC_YR, CALENDAR_YR, TERM_CD, TERM_NAME, CAMPUS_CD and STUD_ID is the primary key for ENROLLMENT_DM table.

3.5 MQTS IN UCDW STUDENT ENROLLMENT

A Materialized query table (MQT) is a table whose definition is based upon the result of a query. The data that is contained in an MQT is derived from one or more tables on which the materialized query table definition is based.

MQTs are materialized views. Both views and MQTs are defined on the basis of a query. The query on which a view is based is run whenever the view is referenced; however, an MQT actually stores the query results as data, and you can work with the data that is in the MQT instead of the data that is in the underlying tables. UCDW Student Enrollment currently has two MQTs: ENROLLMENT_HEAD_COUNT_M and SUMMER_ENROLLMENT_HEAD_COUNT_M.

3.5.1 ENROLLMENT_HEAD_COUNT_M

ENROLLMENT_HEAD_COUNT_M stores only one record per student (primary major) regardless of their multiple majors. This MQT also has ENRL_MULTI_MAJOR_FL column that shows whether the student has multiple majors.

Look at the below example to better understand the difference between ENROLLMENT_F and ENROLLMENT_HEAD_COUNT_M tables.

Example:

Join the ACADEMIC_SUB_TERM_D dimension to the ENROLLMENT_F fact table to see a count of rows in the Enrollment star for academic year 2013, broken out by term.

Sample SQL:

```
select
    t.ACAD_SUB_T_ACAD_YR,
    t.ACAD_SUB_T_CD,
    t.ACAD_SUB_T_REC_TY,
    t.ACAD_SUB_T_NAM,
    t.ACAD_SUB_T_CYCLE_BEG_DT,
    COUNT (*) AS FACT_ROW_COUNT
from STUD_BI.ENROLLMENT_F f
inner join STUD_BI.ACADEMIC_SUB_TERM_D t
    on f.ACAD_SUB_T_KEY = t.ACAD_SUB_T_KEY
where ACAD_SUB_T_ACAD_YR = 2013
group by
    t.ACAD_SUB_T_ACAD_YR,
    t.ACAD_SUB_T_CD,
    t.ACAD_SUB_T_REC_TY,
    t.ACAD_SUB_T_NAM,
    t.ACAD_SUB_T_CYCLE_BEG_DT
order by
    t.ACAD_SUB_T_ACAD_YR,
    t.ACAD_SUB_T_CD,
    t.ACAD_SUB_T_REC_TY,
    t.ACAD_SUB_T_NAM,
    t.ACAD_SUB_T_CYCLE_BEG_DT
```

for read only with ur;

*	ACAD_SUB_T_ACAD_YR	ACAD_SUB_T_CD	ACAD_SUB_T_REC_TY	ACAD_SUB_T_NAM	ACAD_SUB_T_CYCLE_BEG_DT	FACT_ROW_COUNT
1	2013	1	3WK	Summer	2013-09-15	92877
2	2013	1	EOT	Summer	2013-11-01	93376
3	2013	2	3WK	Fall	2013-11-08	251918
4	2013	2	EOT	Fall	2014-01-15	245352
5	2013	3	3WK	Winter	2014-02-15	197656
6	2013	3	EOT	Winter	2014-05-02	190769
7	2013	4	3WK	Spring	2014-05-06	237954
8	2013	4	EOT	Spring	2014-06-30	231531

Figure 5: Enrollment Fact Row Counts

Modify the previous SQL to join to ENROLLMENT_HEAD_COUNT_M fact MQT instead of the ENROLLMENT_F fact table to see a count of rows in the Enrollment star for academic year 2013, broken out by term. Use SUM (ENRL_HEAD_CNT) in your SQL instead of COUNT (*):

```
select
    t.ACAD_SUB_T_ACAD_YR,
    t.ACAD_SUB_T_CD,
    t.ACAD_SUB_T_REC_TY,
    t.ACAD_SUB_T_NAM,
    t.ACAD_SUB_T_CYCLE_BEG_DT,
    SUM (ENRL_HEAD_CNT) AS ENRL_HEAD_CNT
from STUD_BI.ENROLLMENT_HEAD_COUNT_M f
inner join STUD_BI.ACADEMIC_SUB_TERM_D t
    on f.ACAD_SUB_T_KEY = t.ACAD_SUB_T_KEY
where ACAD_SUB_T_ACAD_YR = 2013
group by
    t.ACAD_SUB_T_ACAD_YR,
    t.ACAD_SUB_T_CD,
    t.ACAD_SUB_T_REC_TY,
    t.ACAD_SUB_T_NAM,
    t.ACAD_SUB_T_CYCLE_BEG_DT
order by
    t.ACAD_SUB_T_ACAD_YR,
    t.ACAD_SUB_T_CD,
    t.ACAD_SUB_T_REC_TY,
    t.ACAD_SUB_T_NAM,
    t.ACAD_SUB_T_CYCLE_BEG_DT
for read only with ur;
```

You should see the counts below:

*	ACAD_SUB_T_ACAD_YR	ACAD_SUB_T_CD	ACAD_SUB_T_REC_TY	ACAD_SUB_T_NAM	ACAD_SUB_T_CYCLE_BEG_DT	ENRL_HEAD_CNT
1	2013	1	3WK	Summer	2013-09-15	89139
2	2013	1	EOT	Summer	2013-11-01	89324
3	2013	2	3WK	Fall	2013-11-08	244126
4	2013	2	EOT	Fall	2014-01-15	236617
5	2013	3	3WK	Winter	2014-02-15	191483
6	2013	3	EOT	Winter	2014-05-02	183819
7	2013	4	3WK	Spring	2014-05-06	228626
8	2013	4	EOT	Spring	2014-06-30	221234

Figure 6: Enrollment Head Count in MQT



Did you notice that the counts are lower when you use ENROLLMENT_HEAD_COUNT_M fact MQT instead of ENROLLMENT_F fact table?



Double major students have two rows in the fact table and triple major students have 3 rows in the fact table. The counts on Enrollment fact table are slightly higher because double major students are counted twice and triple major students are counted three times.

Always use the ENROLLMENT_HEAD_COUNT_M instead of ENROLLMENT_F if you want to count a student only once regardless of their multiple majors.

3.5.2 SUMMER_ENROLLMENT_HEAD_COUNT_M

SUMMER_ENROLLMENT_HEAD_COUNT_M stores only one record per student (primary major) regardless of their multiple majors. This MQT also has SUMR_ENRL_MULTI_MAJOR_FL column that shows whether the student is a multiple major student. For students taking courses at one or more host campuses the measures are merged and represented under their home campus.

Always use the SUMMER_ENROLLMENT_HEAD_COUNT_M instead of SUMMER_ENROLLMENT_F if you want to count a student only one regardless of their multiple majors.

4. SAMPLE QUERIES

Sample queries for UCDW Student Enrollment are available in JIRA-IRAP Code Library Project. Use this link to access JIRA-IRAP Code Library:

<https://ucicdwrequirements.atlassian.net/issues/?jql=project%20%3D%20ICL%20AND%20resolution%20%3D%20Unresolved%20ORDER%20BY%20key%20ASC%2C%20priority%20DESC%2C%20updated%20DESC>

5. TERMS AND ABBREVIATIONS

#	ACRONYM	MEANING
1	UCDW	University of California Data Warehouse
2	MQT	Materialized Query Table
3	FERPA	Family Educational Rights and Privacy Act
4	UC	University of California
5	UCOP	University of California Office of the President
6	MQT	Materialized Query Table
7	BI	Business Intelligence
8	EOT	End of Term
9	3WK	Third Week
10	SAS	Statistical Analysis System
11	SCD	Slowly Changing Dimension
12	SQL	Structured Query Language

Table 2: Glossary