UNIVERSITY OF CALIFORNIA

Office of the President
Institutional Research and Academic Planning (IRAP)
UCDW Course Enrollment User Guide
# Table of Contents

## 1. Introduction

1.1 Overview of Course Enrollment ................................................................. 7
1.2 Student Privacy ....................................................................................... 7
1.3 How Course Enrollment Data in UCDW Is Stored? ................................. 7
1.4 Important points to know when using UCDW Course Enrollment ............. 9

## 2. Accessing UCDW Course Enrollment .......................................................... 9

## 3. UCDW Course Enrollment Tables Guide ...................................................... 10

3.1 UCDW Course Enrollment Star ................................................................. 10
3.2 Details on Fact and Dimension Tables in UCDW Course Enrollment ........ 10

3.2.1 COURSE_ENROLLMENT_F .......................................................... 10
3.2.2 STUDENT_D ................................................................................. 11
3.2.3 CAMPUS_LOCATION_D ............................................................... 11
3.2.4 COURSE_ENROLLMENT_STATUS_D ........................................... 11
3.2.5 AGE_BAND_D ............................................................................. 12
3.2.6 ACADEMIC_SUB_TERM_D ......................................................... 12
3.2.7 COURSE_D .................................................................................. 12

3.3 DataMart in UCDW Student Enrollment ................................................... 13
3.4 MQTS in UCDW Student Enrollment ....................................................... 13

## 4. Sample Queries ....................................................................................... 13

## 5. Terms and Abbreviations .......................................................................... 14
TABLE OF TABLES

Table 1: Course Enrollment files from UC campuses ................................................................. 7
Table 2: Glossary .......................................................................................................................... 14
## 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 .............. 10
1. INTRODUCTION

1.1 OVERVIEW OF COURSE ENROLLMENT

UCDW-Course Enrollment data warehouse stores data on all courses the students engage in during the academic year – both 3rd week and end of term in fall, winter, spring and summer terms. UC's Course Enrollment data is consolidated into one enterprise-wide system which will help improve accuracy, timeliness, flexibility, reporting of course enrollment data and also provides system-wide reporting data for course enrollments across all 10 UC campuses. The data in UCDW-Course Enrollment is dependent on the submission of campus 3rd week, end of term, and summer course 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 COURSE ENROLLMENT DATA IN UCDW IS STORED?

All 10 UC campuses submit course 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.

<table>
<thead>
<tr>
<th></th>
<th>Fall</th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3WK</td>
<td>EOT</td>
<td>3WK</td>
<td>EOT</td>
</tr>
<tr>
<td>Berkeley</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>San Francisco</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Davis</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Riverside</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>San Diego</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Santa Cruz</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Santa Barbara</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Irvine</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Merced</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

Table 1: Course Enrollment files from UC campuses
File specifications and the file submission deadline for course enrollment data are available in this link.

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

All 10 campuses’ course enrollment related data is first stored in staging layer which is a parking lot for all the campus course 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.

For analyses, analytical and reporting purposes, business users use data in BI layer. BI layer data is stored in a star schema relational database design, 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 and 3.2 for details on the fact and dimension tables in UCDW- Course Enrollment.
1.4 IMPORTANT POINTS TO KNOW WHEN USING UCDW COURSE ENROLLMENT

i. 3rd week and EOT course enrollment data for fall, winter, spring and summer terms from applicable UC campuses are stored in course enrollment star.

ii. For a given term (summer, fall, winter, spring) and sub-term (3WK, EOT), enrollment data should be loaded into enrollment star before course enrollment star can be populated.

iii. Do’s and Don’ts on querying UCDW is available via this link.

2. ACCESSING UCDW COURSE ENROLLMENT

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

i. UCDW Course 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 Course 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 Course 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 Course Enrollment data can also be accessed using other data analysis and mining tools such as SAS.

3. UCDW Course Enrollment Tables Guide

3.1 UCDW Course Enrollment Star

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

![Diagram of UCDW Course Enrollment Star]

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 Details on Fact and Dimension Tables in UCDW Course Enrollment

3.2.1 COURSE_ENROLLMENT_F

Description: Course Enrollment fact table tracks course 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 course and the course enrollment status (P, N, C, R, V or W).

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, CRSE_KEY, and CRSE_ENRL_STAT_KEY is primary key to COURSE_ENROLLMENT_F table.

3.2.2 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.

3.2.3 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

3.2.4 COURSE_ENROLLMENT_STATUS_D

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

Course Enrollment Status dimension is used in course enrollment 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. CRSE_ENRL_STAT_KEY is the primary key for COURSE_ENROLLMENT_STATUS_D table.

Natural Key: N/A

3.2.5 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

3.2.6 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

3.2.7 COURSE_D

Description: Course dimension contains information about all types of courses offered by UC campuses.

Course dimension is used in course enrollment subject area. 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. CRSE_KEY is the primary key for COURSE_D table.

Natural Key: N/A
3.3 DATAMART IN UCDW STUDENT ENROLLMENT

Data marts are easily accessible by business users, because they don’t require queries that are too complicated or resource-consuming. Currently, there is no data mart for course enrollment in UCDW.

3.4 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 a 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. Currently, UCDW Course Enrollment does not have MQTs.

4. SAMPLE QUERIES

Below are few sample queries to use course enrollment star.

Example:

Query to see count of students enrolled in MATH courses in fall 3rd week of academic year 2015.

Sample SQL:

```sql
select
    t.ACAD_SUB_T_ACAD_YR,
    t.ACAD_SUB_T_NAM,
    t.ACAD_SUB_T_REC_TY,
    l.CMP_LOC_LOC1_SHRT_DESC,
    c.CRSE_DEPT_CD,
    COUNT(*) AS COUNT
from STUD_BI.COURSE_ENROLLMENT_F f
inner join STUD_BI.ACADEMIC_SUB_TERM_D t
    on f.ACAD_SUB_T_KEY = t.ACAD_SUB_T_KEY
inner join STUD_BI.CAMPUS_LOCATION_D l
    on f.CMP_LOC_KEY = l.CMP_LOC_KEY
inner join STUD_BI.COURSE_D c
    on f.CRSE_KEY = c.CRSE_KEY
inner join STUD_BI.COURSE_ENROLLMENT_STATUS_D ces
    on f.CRSE_ENRL_STAT_KEY = ces.CRSE_ENRL_STAT_KEY
inner join STUD_BI.AGE_BAND_D a
    on f.AGE_BAND_KEY = a.AGE_BAND_KEY
where t.ACAD_SUB_T_ACAD_YR = 2015
    and t.ACAD_SUB_T_NAM = 'Fall'
    and t.ACAD_SUB_T_REC_TY = '3WK'
    and UPPER(c.CRSE_DEPT_CD) = 'MATH'
    and ces.CRSE_ENRL_STAT_DESC = 'Enrolled'
group by
    t.ACAD_SUB_T_ACAD_YR,
```
More sample queries on UCDW Course 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%20priority%20DESC%20updated%20DESC

5. TERMS AND ABBREVIATIONS

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

Table 2: Glossary