Introduction to UCDW Star Schemas and Data Marts Live Demonstration

## Data Infrastructure IRAP Training
4/12/2017

Agenda

- Schemas of Interest
- Facts and Dimensions in UCDW
- Time Dimensions
- Dimensions Keys
- Fact Keys
- Materialized Query Tables (MQT)
  - UCDW MQT Examples
- Views
  - UCDW Regular View Examples
- Difference Between MQT and Views
- Q&A

UCDW Schemas of Interest

- Payroll/Personnel
  - DSS_BI
- Graduate Admissions
  - GAD_STG
  - GAD_BASE
  - GAD_BI
- Contracts and Grants
  - SP_STG
  - SP_BASE
  - SP_BI
- Student & Course Enrollment, Degree and Financial Aid
  - STUD_STG
  - STUD_BASE
  - STUD_BI
- Undergraduate Admissions
  - UAD_STG
  - UAD_BASE
  - UAD_BI
- Data Marts
  - IRAP_BI
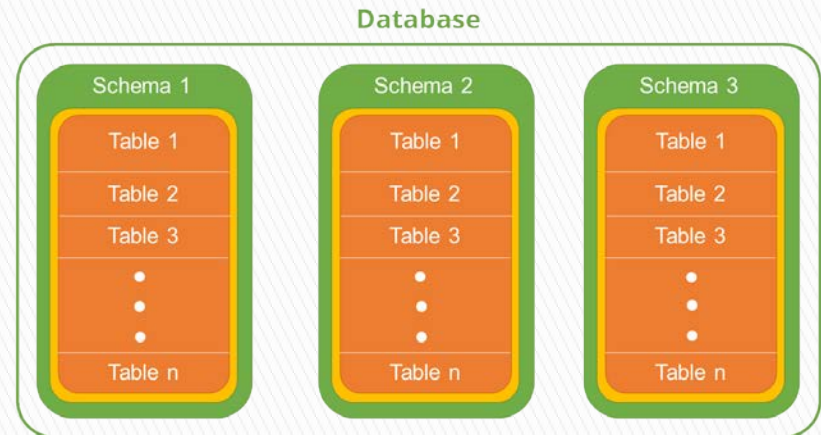
Facts and Dimension Tables
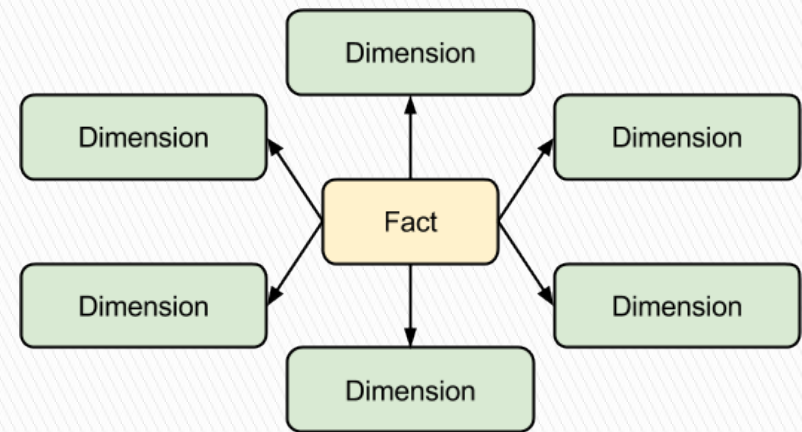
- Student Enrollment Facts
  - ENROLLMENT_F
  - SUMMER_ENROLLMENT_F
- Student Enrollment Dimensions
  - ACADEMIC_SUB_TERM_D
  - AGE_BAND_D
  - CAMPUS_COLLEGE_MAJOR_D
  - CAMPUS_LOCATION_D
  - COUNTRY_STATE_COUNTY_D
  - ENROLLMENT_STATUS_D
  - IPEDS_ETHNIC_CODE_D
  - STUDENT_D
  - STUDENT_ETHNIC_CODE_D
  - STUDENT_LEVEL_D
  - UC_ETHNIC_CODE_LEVEL_1_D
  - UC_ETHNIC_CODE_LEVEL_2_D

Tables with _F are **facts**

Tables with _D are **dimensions**

Dimension

Dimension

Dimension

Fact

Dimension

Dimension

Dimension

TEXT   Shared

TEXT   Conformed

Time Dimensions

- Payroll/Personnel
  - CALENDAR_D
  - CALENDAR_YEAR_D
- Financial Aid
  - FILE_CYCLE_D
- Undergraduate & Graduate Admissions
  - TERM_D
- Student & Course Enrollment
  - ACADEMIC_SUB_TERM_D
- Degree
  - ACADEMIC_TERM_D

Dimension Key

- One primary Key
- System-Generated – Surrogate Keys – No Business Context
  - Examples of Surrogate Keys:
    - STUDENT_D.STUD_KEY
    - COURSE_D.CRSE_KEY
    - AGE_BAND.AGE_BAND_KEY
    - ENROLLMENT_STATUS_D.ENRL_STAT_KEY
- Natural Key – Have Business Context
  - Examples of Natural Keys
    - CMP_CG_MAJ_CD_CIP_CD
    - CMP_LOC_LOC1_CD
    - CRSE_SECT_ID
    - STUD_ID

Columns with _KEY are surrogate keys

Columns with _CD or _ID are natural keys

Fact Keys

▸ Primary key is a combination of all connected dimension keys as composite key

▸ No fact specific surrogate key
  ◦ Example STUD_BI.ENROLLMENT_F has the following keys:
    · ACAD_SUB_T_KEY *link to* ACADEMIC_SUB_TERM_D
    · STUD_KEY *link to* STUDENT_D
    · STUD_LVL_KEY *link to* STUDENT_LEVEL_D
    · ENRL_STAT_KEY *link to* ENROLLMENT_STATUS_D
    · CMP_CG_MAJ_CD_KEY *link to* CAMPUS_COLLEGE_MAJOR_CODE_D
    · ACAD_DGR_KEY *link to* ACADEMIC_DEGREE_D
    · AGE_BAND_KEY *link to* AGE_BAND_D

Materialized Query Tables (MQT)

- Physical/Materialized View
- Created for Simplicity
- Created for Join Avoidance
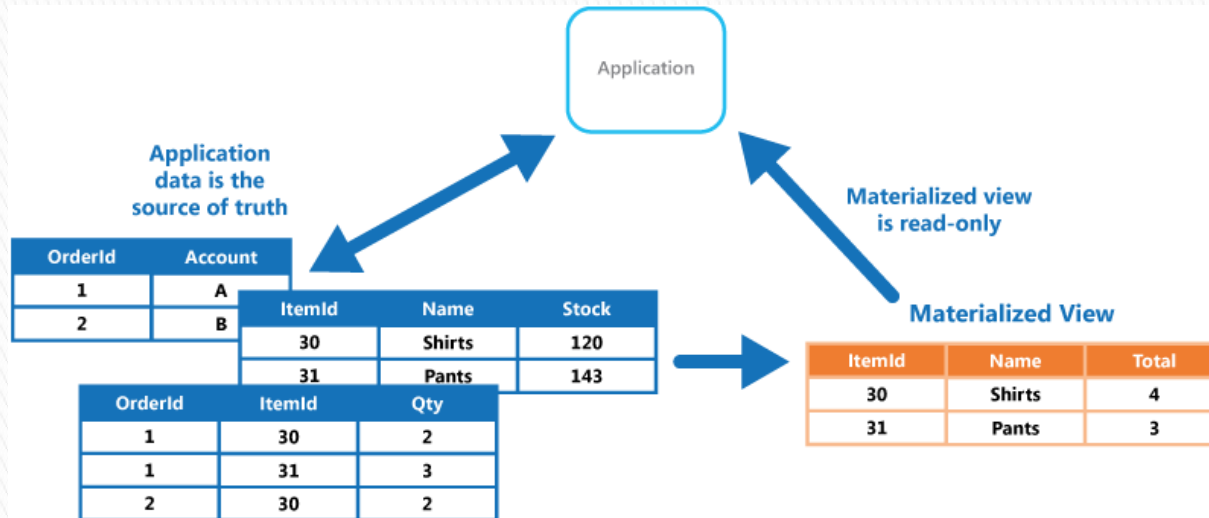- Identical to a Query
- Stores Query Results as Data
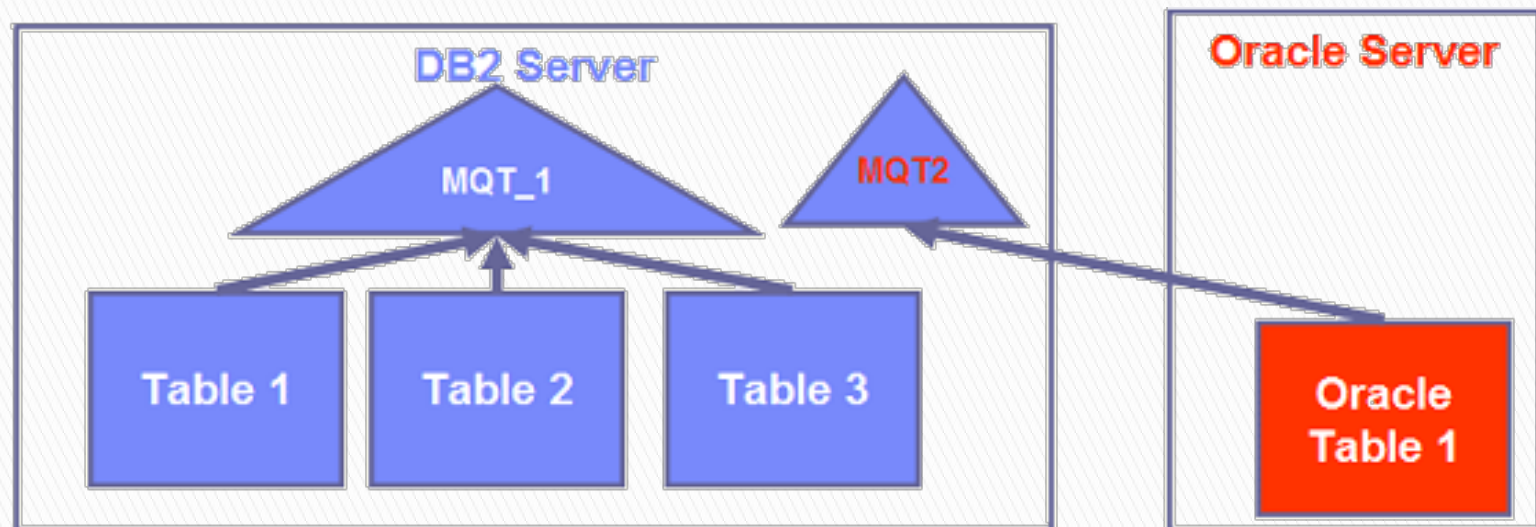
UCDW MQT Examples

▸ STUD_BI Schema
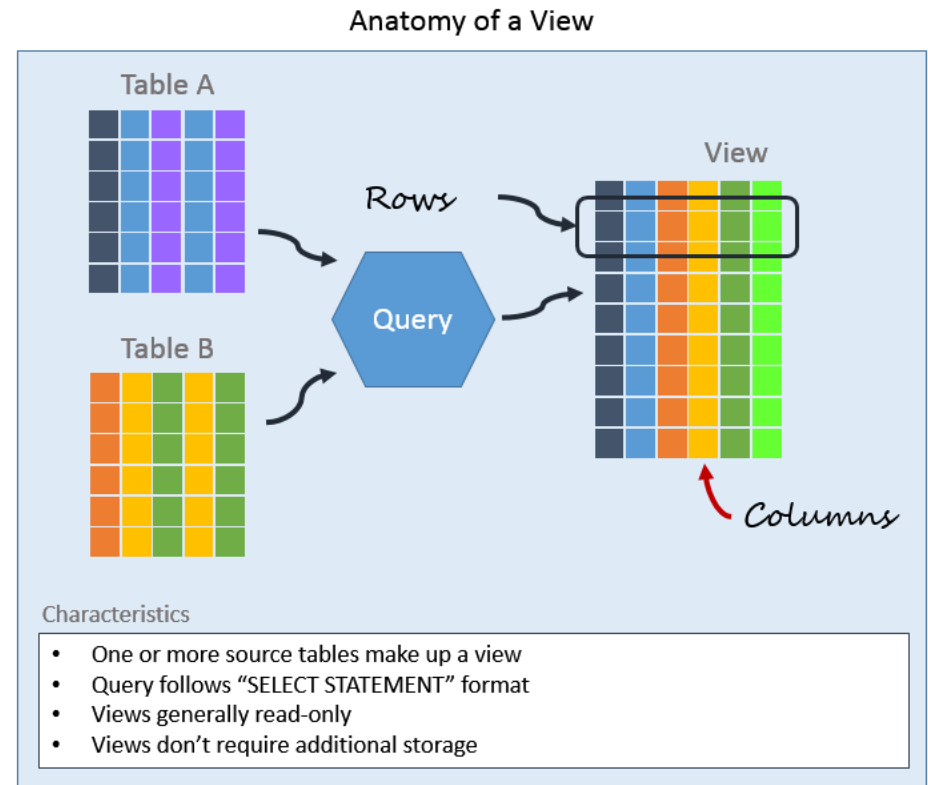- DEGREE_HEAD_COUNT_**M**
- ENROLLMENT_HEAD_COUNT_**M**
- STUDENT_IPEDS_ETHNIC_CODE_**M**
- STUDENT_UC_ETHNIC_CODE_LEVEL_1_**M**
- STUDENT_UC_ETHNIC_CODE_LEVEL_2_**M**
- SUMMER_ENROLLMENT_HEAD_COUNT_**M**

Application

Application data is the source of truth

Materialized view is read-only

OrderId	Account
1	A
2	B

ItemId	Name	Stock
30	Shirts	120
31	Pants	143

OrderId	ItemId	Qty
1	30	2
1	31	3
2	30	2

Materialized View

| ItemId	Name	Total
30	Shirts	4
31	Pants	3

Views

- ▸ Created Dynamically
- ▸ Created for Simplicity
- ▸ Created for Join Avoidance
- ▸ Definition Based on a Query
- ▸ Query Runs When View is Referenced

Anatomy of a View

Table A

Rows

Query

View

Table B

Columns

Characteristics

- • One or more source tables make up a view
- • Query follows "SELECT STATEMENT" format
- • Views generally read-only
- • Views don't require additional storage

UCDW View Examples

▸ STUD_BI Schema
  ◦ ETHNIC_CODE_V
  ◦ FIRST_ENROLLED_V
  ◦ IPEDS_ETHNIC_CODE_V
  ◦ UC_ETHNIC_CODE_LEVEL_1_V
  ◦ UC_ETHNIC_CODE_LEVEL_2_V

Tables with _V are views

View is created from Table 1 (Column2, Column 3) and Table 2 (Column1)
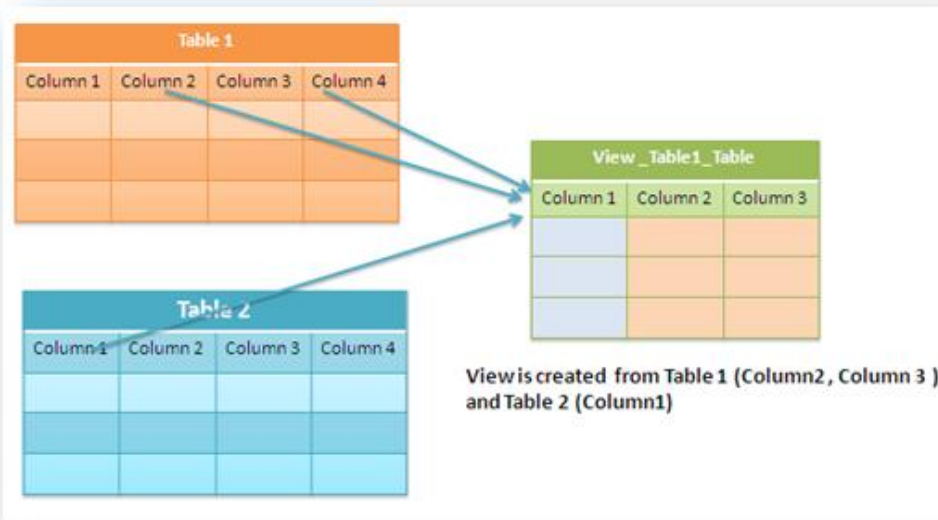
MQTs or View - Differences

- Physical
- Stored as physical object in database
- Indexes can be created on MQT
- Faster performance
- Needs to be updated or refreshed

- Dynamic
- Stored as query against base table(s)
- Indexes on base tables must be used
- No real performance improvement
- No update required

| Materialized Query Table | Regular View |