



UCOP Data Users Group

April 27, 2018

Exercise Goal 1

- ▶ **(UAD: Campus Level Data)**
 - Write a SQL query to find total number of undergraduate applicants applied for Engineering/Computer Sciences and got admitted into Berkeley or Los Angeles.

Exercise 1 – Rules and GroupBy

▶ Rules:

- Academic Year: 2017
- Term Name: Fall
- Campus name: BERKELEY, LOS ANGELES
- Broad Discipline 9 Category: Engineering/Computer Sciences
- Registration Status: Use CASE Statement to set the value to 'Registered' if the value is 'Y' else set to 'Not Registered'

▶ Group By:

- Academic Year
- Term Name
- Campus name
- Broad Discipline 9 Category
- Citizenship Status 2 Category
- Registration Status

CASE Statement Syntax

- ▶ CASE

 - WHEN column name = 'Y'
THEN 'Registered'
ELSE 'Not Registered'

- ▶ END AS REGISTERED_FLAG

Execution Steps

- ▶ Find the required fact and dimension tables
- ▶ Find the joining surrogate key columns
- ▶ Find the fact and dimension non key columns required in the query
- ▶ Build the query by joining the fact table with all the required dimensions using surrogate keys.
- ▶ Add the filter conditions on the dimension columns based on the required values.
- ▶ Finally group the data set based on the mentioned columns and generate the aggregated result set.

Exercise 1 – Technical Requirements

▶ Schema:

- UAD_BI

▶ Table Name:

- CAMPUS_APPLIED_F
- TERM_D
- APPLICANT_D
- CAMPUS_COLLEGE_MAJOR_CODE_D
- CAMPUS_LOCATION_D

▶ Column Names:

- Campus Applied Fact: Applicant ID(**CMP_APPLD_APPL_ID**), Admitted Flag, Registered Flag
- Term dimension: Academic year and term name
- Campus Dimension: campus long description
- Campus College Major Dimension: Broad Discipline 9 Category name
- Applicant Dimension: Citizenship status 2 Category

Exercise 1 – Expected Result Set

YEAR	TERM	CAMPUS	BRD_DISCIPLINE_9CAT	CITIZENSHIP_2CAT	REGISTERED_FLAG	CNT
2017	Fall	BERKELEY	Engineering/Computer Sciences	Domestic	Not Registered	1407
2017	Fall	BERKELEY	Engineering/Computer Sciences	Domestic	Registered	1480
2017	Fall	BERKELEY	Engineering/Computer Sciences	Foreign	Not Registered	124
2017	Fall	BERKELEY	Engineering/Computer Sciences	Foreign	Registered	240
2017	Fall	LOS ANGELES	Engineering/Computer Sciences	Domestic	Not Registered	2106
2017	Fall	LOS ANGELES	Engineering/Computer Sciences	Domestic	Registered	853
2017	Fall	LOS ANGELES	Engineering/Computer Sciences	Foreign	Not Registered	357
2017	Fall	LOS ANGELES	Engineering/Computer Sciences	Foreign	Registered	144

Exercise 1 – SQL Query

```
SELECT
YEAR , TERM, CAMPUS, BRD_DISCIPLINE_9CAT, CITIZENSHIP_2CAT, REGISTERED_FLAG, count(CMP_APPLD_APPL_ID)
FROM (
SELECT t_d.T_ACAD_YR as YEAR, t_d.T_NAM as TERM, ca_f.CMP_APPLD_APPL_ID, ca_f.CMP_APPLD_AL_ID,
cl_d.CMP_LOC_LOC1_LNG_DESC as CAMPUS,
CASE WHEN ca_f.CMP_APPLD_APPL_REGD_FL = 'Y' THEN 'Registered'
ELSE 'Not Registered'
END AS REGISTERED_FLAG, a_d.AL_CTZN_STAT_ADM_2CAT_DESC as CITIZENSHIP_2CAT,
ccm_d.CMP_CG_MAJ_CD_BROAD_DISCIPLN_9CAT_NAM as BRD_DISCIPLINE_9CAT
FROM UAD_BI.CAMPUS_APPLIED_F ca_f
    INNER JOIN UAD_BI.APPLICANT_D a_d    ON ca_f.AL_KEY = a_d.AL_KEY
    INNER JOIN UAD_BI.TERM_D t_d        ON ca_f.T_KEY = t_d.T_KEY
    INNER JOIN UAD_BI.CAMPUS_COLLEGE_MAJOR_CODE_D ccm_d ON ca_f.CMP_CG_MAJ_CD_KEY =
ccm_d.CMP_CG_MAJ_CD_KEY
    INNER JOIN UAD_BI.CAMPUS_LOCATION_D cl_d  ON ca_f.CMP_LOC_KEY = cl_d.CMP_LOC_KEY
WHERE
t_d.T_ACAD_YR IN (2017)
AND t_d.T_NAM = 'Fall'
AND cl_d.CMP_LOC_LOC1_LNG_DESC IN( 'BERKELEY', 'LOS ANGELES')
AND CMP_CG_MAJ_CD_BROAD_DISCIPLN_9CAT_NAM IN ('Engineering/Computer Sciences')
AND ca_f.CMP_APPLD_ADT_CNT = '1'
)
GROUP BY YEAR , TERM, CAMPUS, BRD_DISCIPLINE_9CAT, CITIZENSHIP_2CAT, REGISTERED_FLAG
ORDER BY YEAR , TERM, CAMPUS, BRD_DISCIPLINE_9CAT, CITIZENSHIP_2CAT, REGISTERED_FLAG
```