



UCOP Data Users Group

October 18, 2018

Exercise Goal

- ▶ Degree and Financial Aid Star:
 - Write a SQL query to find the Academic Masters degrees awarded in academic year 2016 for UC Davis graduate students, that are in top five for average student debt at the time of graduation.

Execution steps to join Degree and FAI

- ▶ Build query from degree star using degree awarded fact and associated dimensions (Query1)
- ▶ Build query from FAI star using financial aid fact and associated dimensions (Query2)
- ▶ Join Query1 to Query2 using cross schema join(INNER JOIN) using campus_cd+student_id(CMP_STUD_ID)
- ▶ Group the data set based on the required elements

Syntax to join Degree and FAI

```
SELECT ACAD_DGR_LNG_NAM, AVG_DEBT
FROM
(
SELECT
DEGREE_STAR.ACAD_DGR_LNG_NAM, AVG(FAI_STAR.TOTAL_DEBT_LOAN)
as AVG_DEBT
FROM
(QUERY1) as DEGREE_STAR
INNER JOIN
(QUERY2) as FAI_STAR
ON DEGREE_STAR.CMP_STUD_ID = FAI_STAR.CMP_STUD_ID
GROUP BY DEGREE_STAR.ACAD_DGR_LNG_NAM
)
ORDER BY AVG_DEBT DESC
FETCH FIRST 5 ROWS ONLY;
```

- *FAI_STAR.TOTAL_DEBT_LOAN is a calculated field from FAI star. It is the sum of debt amounts from all the years up until academic year 2016, for every student.*
- *Loans awarded prior to 2016 for a student, could also include the students undergraduate debts.*

Execution steps to write a query

- ▶ 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.
- ▶ Group the data set based on the mentioned columns and generate the aggregated result set.
- ▶ Rank the result set based on the required attributes.

Query1 – Technical Requirements

▶ Schema:

- STUD_BI

▶ Table Name:

- DEGREE_AWARDED_F
- STUDENT_D
- CAMPUS_LOCATION_D
- STUDENT_LEVEL_D
- ACADEMIC_DEGREE_D
- ACADEMIC_TERM_D

▶ Column Names:

- Degree Awarded fact: Degree awarded major component number
- Student dimension: Student ID
- Campus Location dimension: Campus code
- Student Level dimension: Student level code and description
- Academic Degree dimension: Academic degree long name
- Academic Term dimension: Academic Year

Query1: Degree star Query

▶ Rules:

- ▶ File Cycle Academic Year = 2016
- ▶ Student Level Code = '5'
- ▶ Campus Code = '03'
- ▶ Degree Awarded Major Component Number = 1
- ▶ Create Campus Student ID:
 - `CMP_STUD_ID` = campus code||student id (student_d).
- ▶ For students who have earned multiple degrees in different years, select only the degrees earned in the latest year.

Query2 – Technical Requirements

▶ Schema:

- STUD_BI

▶ Table Name:

- STUDENT_FINANCIAL_AID_F
- STUDENT_D
- CAMPUS_LOCATION_D
- FILE_CYCLE_D
- AWARD_D
- AWARD_REFERENCE_D

▶ Column Names:

- Financial Aid fact: Paid to date amount
- Student dimension: Student ID
- Campus Location dimension: Campus Code
- File cycle dimension: Academic year, and Final file flag
- Award dimension: Award type code
- Award reference dimension: Award acceptance code

Query2: FAI star Query

▶ Rules:

- ▶ File Cycle Academic Year \leq 2016
- ▶ Award Type Code = '3'(Loan)
- ▶ Award reference acceptance code = 'A'
- ▶ File cycle Final File = 'Y'
- ▶ Campus Code = '03'
- ▶ Create Campus Student ID:
 - `CMP_STUD_ID = campus code||student id (student_d)`.

Cumulative columns

- ▶ Calculate the SUM of all loan amounts awarded for every student until the academic year 2016. Debt amounts could also include loans awarded prior to masters degree such as undergraduate degree.
- ▶ Calculate AVG of the students total debt amount grouped by academic degree awarded in 2016.
- ▶ Order the result set by calculated average amounts and use FETCH FIRST 5 ROWS ONLY to get the top 5 records.

Query – Part 1

```

SELECT ACAD_DGR_LNG_NAM, AVG_DEBT
FROM
(
  SELECT
  DEGREE_STAR.ACAD_DGR_LNG_NAM, AVG(FAI_STAR.TOTAL_DEBT_LOAN) as AVG_DEBT
  FROM
  (
    SELECT ACAD_T_ACAD_YR, CMP_STUD_ID, ACAD_DGR_LNG_NAM, STUD_LVL_DESC
    FROM (
      SELECT
      s_d.STUD_LOC_CMP_CD||s_d.STUD_ID as CMP_STUD_ID, s_d.STUD_LOC_CMP_CD, s_d.STUD_ID,
      ad_d.ACAD_DGR_SHRT_NAM, ad_d.ACAD_DGR_LNG_NAM, ad_d.ACAD_DGR_CATG_NAM, ad_d.ACAD_DGR_TY_NAM,
      at_d.ACAD_T_ACAD_YR, at_d.ACAD_T_NAM, sl_d.STUD_LVL_DESC,
      DENSE_RANK() OVER(PARTITION BY s_d.STUD_LOC_CMP_CD, s_d.STUD_ID order by at_d.
      ACAD_T_CYCLE_BEG_DT desc) as ROWNUM1 /*Don't use row_number because a student might have
      multiple degrees awarded same year. We need to consider all the degrees earned*/
      FROM STUD_BI.DEGREE_AWARDED_F da_f
      INNER JOIN STUD_BI.STUDENT_D s_d ON da_f.STUD_KEY = s_d.STUD_KEY
      INNER JOIN STUD_BI.STUDENT_LEVEL_D sl_d ON da_f.STUD_LVL_KEY = sl_d.STUD_LVL_KEY
      INNER JOIN STUD_BI.ACADEMIC_DEGREE_D ad_d ON da_f.ACAD_DGR_KEY = ad_d.ACAD_DGR_KEY
      INNER JOIN STUD_BI.ACADEMIC_TERM_D at_d ON da_f.ACAD_T_KEY = at_d.ACAD_T_KEY
      WHERE
      at_d.ACAD_T_ACAD_YR = 2016
      AND sl_d.STUD_LVL_CD IN ('5')
      AND da_f.DGR_AWRDED_MAJ_CMPNT_NUM = 1
      AND s_d.STUD_LOC_CMP_CD = '03'
      ORDER BY s_d.STUD_LOC_CMP_CD
    )
    WHERE ROWNUM1 = 1
    ORDER BY ACAD_T_ACAD_YR, CMP_STUD_ID, ACAD_DGR_LNG_NAM, STUD_LVL_DESC
  ) as DEGREE_STAR
  )
  )
  )

```

Query – Part2

```

LEFT OUTER JOIN
(
SELECT * FROM (
SELECT CMP_STUD_ID,  SUM(STUD_FINL_AID_AWRD_PD_TO_DT_AMT) AS TOTAL_DEBT_LOAN
FROM
(
SELECT cl_d.CMP_LOC_LOC1_CD||s_d.STUD_ID as "CMP_STUD_ID", fc_d.FILE_CYCLE_MTH_SHRT_NAM, fc_d.
FILE_CYCLE_DPNDT_ON_REC_TY, fc_d.FILE_CYCLE_ACAD_YR, aw_d.AWRD_CD, fai_f.
STUD_FINL_AID_AWRD_PD_TO_DT_AMT
FROM STUD_BI.STUDENT_FINANCIAL_AID_F fai_f
INNER JOIN STUD_BI.STUDENT_D s_d      ON fai_f.STUD_KEY = s_d.STUD_KEY
INNER JOIN STUD_BI.CAMPUS_LOCATION_D cl_d  ON fai_f.CMP_LOC_KEY = cl_d.CMP_LOC_KEY
INNER JOIN STUD_BI.FILE_CYCLE_D fc_d      ON fai_f.FILE_CYCLE_KEY = fc_d.FILE_CYCLE_KEY
INNER JOIN STUD_BI.AWARD_D aw_d          ON fai_f.AWRD_KEY = aw_d.AWRD_KEY
INNER JOIN STUD_BI.AWARD_REFERENCE_D awr_d ON fai_f.AWRD_REF_KEY = awr_d.AWRD_REF_KEY
WHERE
aw_d.AWRD_TY_CD = '3'
AND fc_d.FILE_CYCLE_ACAD_YR <= 2016
AND awr_d.AWRD_REF_AWRD_ACPTC_CD = 'A'
AND fc_d.FILE_CYCLE_FN_FILE_FL = 'Y'
and  cl_d.CMP_LOC_LOC1_CD = '03'
)
)
GROUP BY CMP_STUD_ID)
) as FAI_STAR
ON DEGREE_STAR.CMP_STUD_ID = FAI_STAR.CMP_STUD_ID
WHERE
FAI_STAR.TOTAL_DEBT_LOAN IS NOT NULL
GROUP BY DEGREE_STAR.ACAD_DGR_LNG_NAM
)
ORDER BY AVG_DEBT DESC
FETCH FIRST 5 ROWS ONLY;

```

Expected Result Set

ACAD_DGR_LNG_NAM	AVG_DEBT
MASTER OF HEALTH SERVICES	133856.61
MASTER OF PREVENTATIVE VETERINARY MEDICINE	110210.5
JURIS DOCTOR	105935.95
MASTER OF BUSINESS ADMINISTRATION	79864.67
MASTER OF LAWS	59415