

----- SQL PROJECT: LILIANA BARRERA -----

--#1: Joining
--#2: Aggregating
--#3: Subqueries
--#4: Functions
--#5: Date format manipulation

--Describe how you made sure the data was clean and accurate. Include at least the following concepts:

--Deduplication
--Accounting for null values (if null values are expected, explain why)
--Ensuring no data that was needed was lost in joins or filtering
--Ensuring no unreasonable values appear in the data

-----#1 JOINS-----

INNER JOINS- JOINING SIMILAR TABLES TOGETHER IN THE ACADEMIC SCHEMA

-- JOINED THREE TABLES: STUDENT COURSE DETAILS, STUDENT DETAILS, AND PARENT INFORMATION

```
select *
from AD.AD_STUDENT_DETAILS
inner join
AD.AD_PARENT_INFORMATION on AD.AD_STUDENT_DETAILS.PARENT_ID =
AD.AD_PARENT_INFORMATION.PARENT_ID
inner join
AD.AD_STUDENT_COURSE_DETAILS on AD.AD_STUDENT_DETAILS.STUDENT_ID =
AD.AD_STUDENT_COURSE_DETAILS.STUDENT_ID
ORDER BY COURSE_ID;
```

```
select* --JOINED TWO TABLES: FACULTY DETAILS AND JOBS
from AD.AD_FACULTY_DETAILS
inner join
AD.AD_JOBS on AD.AD_JOBS.JOB_ID = AD.AD_FACULTY_DETAILS.JOB_ID;
```

```
select*
from AD.AD_STUDENT_DETAILS
left join
AD.AD_EXAM_TYPE on AD.AD_STUDENT_DETAILS.STUDENT_ID =
AD.AD_EXAM_TYPE.EXAM_TYPE;
```

-----LEFT JOIN-----

--NEW TABLE CREATED TO CONCAT FIRST AND LAST NAME IN FACULTY DETAILS TABLE, THEN USED NEW TABLE USED LEFT INNER JOINED IT ON DEPARTMENTS TABLE

```
create table NEW_FACULTY_DETAILS as
(select FIRST_NAME || ' ' || LAST_NAME as FULL_NAME
from AD.AD_FACULTY_DETAILS)

select*
from NEW_FACULTY_DETAILS

select*
from AD.AD_DEPARTMENTS
left join
NEW_FACULTY_DETAILS on NEW_FACULTY_DETAILS.FULL_NAME =
AD.AD_DEPARTMENTS.HOD
```

-----#2: FUNCTIONS: MONTHS_BETWEEN FUNCTION-----

--BUSINESS QUESTION_1. HOW LONG HAVE THE FACULTY BEEN EMPLOYED AT THIS SCHOOL? WHAT IS THE LENGTH OF TIME OF THEIR EMPLOYMENT? HOW MANY MONTHS HAVE PASSED BETWEEN TODAY'S DATE AND THE HIREDATE?

```
select FACULTY_ID, FIRST_NAME, LAST_NAME, JOB_ID, SALARY, sysdate,
HIRE_DATE, round(months_between(sysdate, HIRE_DATE)/12) as
YEARS_EMPLOYED
from AD.AD_FACULTY_DETAILS
order by YEARS_EMPLOYED desc;h
```

-----#3: DATE FORMAT-----

```
select FACULTY_ID, FIRST_NAME, LAST_NAME, JOB_ID, SALARY,
to_char(HIRE_DATE, 'MON-DD-YYYY') as EMPLOYMENT_START_DATE --FOR
FACULTY DETAILS TABLE
from AD.AD_FACULTY_DETAILS
order by HIRE_DATE desc;

select STUDENT_ID, FIRST_NAME, to_char(STUDENT_REG_YEAR, 'MON-DD-YYYY')
as STUDENT_YEAR --FOR STUDENT DETAILS TABLE
from AD.AD_STUDENT_DETAILS
order by STUDENT_REG_YEAR desc;
```

-----#4: AGGREGATE FUNCTION: AVERAGE OF SALARY-----

--BUSINESS QUESTION_2: WHAT IS THE AVERAGE SALARY BY JOB ID IN
DESCENDING ORDER AND CONDITION WHERE JOB ID INCLUDES HOD = HEAD OF
DEPARTMENT....

```
select JOB_ID, FACULTY_ID, FIRST_NAME, LAST_NAME, avg(SALARY) as  
AVG_SALARY from AD.AD_FACULTY_DETAILS  
where JOB_ID like '%HOD'  
group by JOB_ID, FACULTY_ID, FIRST_NAME, LAST_NAME  
order by AVG(SALARY) desc;
```

--AVERAGE SALARY BY JOB TITLE

```
select JOB_TITLE, FIRST_NAME, LAST_NAME, avg(SALARY) as AVG_SALARY  
from AD.AD_FACULTY_DETAILS  
inner join  
AD.AD_JOBS on AD.AD_JOBS.JOB_ID = AD.AD_FACULTY_DETAILS.JOB_ID  
group by JOB_TITLE, FIRST_NAME, LAST_NAME  
order by AVG_SALARY desc;
```

--BUSINESS QUESTION_3: HOW MANY EMPLOYEES HAVE THE SAME JOB TITLE?

--(USED COUNT FUNCTION TO FIND NUMBER OF EMPLOYEES WITH THE SAME JOB
TITLE)

```
select JOB_TITLE, FIRST_NAME, LAST_NAME, MIN_SALARY, MAX_SALARY,  
count(*) as NUMBER_OF_EMPLOYEES  
from AD.AD_FACULTY_DETAILS  
inner join  
AD.AD_JOBS on AD.AD_JOBS.JOB_ID = AD.AD_FACULTY_DETAILS.JOB_ID  
group by JOB_TITLE, FIRST_NAME, LAST_NAME, MIN_SALARY, MAX_SALARY  
having MAX_SALARY > 16000 and MIN_SALARY < 30000;
```

-----#5: SUBQUERIES-----

```
select FACULTY_ID, JOB_ID, FIRST_NAME, LAST_NAME, SALARY, (select
avg(SALARY) from AD.AD_FACULTY_DETAILS) as AVG_SALARY
from AD.AD_FACULTY_DETAILS;
```

-----BUSINESS QUESTION_4: WHAT IS THE COUNT OF UNIQUE COURSES
TAKEN BY STUDENTS? -----

```
select count(distinct COURSE_ID) as NUMBER_OF_COURSES
from
(select*
from AD.AD_STUDENT_DETAILS
inner join
AD.AD_STUDENT_COURSE_DETAILS on AD.AD_STUDENT_DETAILS.STUDENT_ID =
AD.AD_STUDENT_COURSE_DETAILS.STUDENT_ID
ORDER BY COURSE ID);
```

-----BUSINESS QUESTION_5: WHICH JOB TITLES HAVE THE LOWEST AND HIGHEST SALARY range?-----

```
select*
from AD.AD_JOBS
order by MAX SALARY desc, MIN SALARY;
```

[illegible]

----DETECTING ANOMALIES IN THE DATA: REMOVING DUPLICATE ROWS AND NULL
VALUES IN DATA?-----

---TRYING TO TROUBLESHOOT HOW TO REMOVE DUPLICATE ROWS AND NULL VALUES
FROM MY TABLES IN MY SCHEMA

```
select STUDENT_ID, FIRST_NAME, PARENT_ID, STUDENT_REG_YEAR, COURSE_ID
from
(select*
from AD.AD_STUDENT_DETAILS
inner join
AD.AD_STUDENT_COURSE_DETAILS on AD.AD_STUDENT_DETAILS.STUDENT_ID =
AD.AD_STUDENT_COURSE_DETAILS.STUDENT_ID
ORDER BY COURSE_ID)
group by STUDENT_ID, FIRST_NAME, PARENT_ID, STUDENT_REG_YEAR,
COURSE_ID;
```

```
select STUDENT_ID, COURSE_ID, count(*)
from AD.AD_STUDENT_COURSE_DETAILS
group by STUDENT_ID, COURSE_ID
having count(*) > 1;
```

```
-
select*
from AD.AD_STUDENT_DETAILS;
delete from AD.AD_STUDENT_DETAILS
where EMAIL_ADDR = NULL;
```