Name: Danny Chen
Assignment: Database Programming with SQL: Section 7, 8
Date: 10/17/24

Table of Content:

7.1 - Oracle Equijoin and Cartesian Product
Vocab

Word	Definition
Cartesian Product	Results from an invalid or omitted join condition; all combinations of rows are displayed
EQUIJOIN	Values in a column in one table are equal to a value in another table; also called an inner join or simple join
Proprietary Join	Connection command exclusive to a specific company
Table Aliases	Gives a table another name to simplify queries and improve performance
Join Conditions	Display data from two or more related tables

1. Create a Cartesian product that displays the columns in the d_play_list_items and the d_track_listings in the DJs on Demand database.

ANS:

```
SELECT *
FROM d_play_list_items, d_track_listings;
```

EVENT_ID	SONG_ID	COMMENTS	SONG_ID	CD_NUMBER	TRACK
100	45	Play late	45	92	1
100	46	-	45	92	1
100	47	Play early	45	92	1
105	48	Play after cake cutting	45	92	1
105	49	Play first	45	92	1
105	47	Play for the father	45	92	1
100	45	Play late	46	93	1
100	46	-	46	93	1
100	47	Play early	46	93	1
105	48	Play after cake cutting	46	93	1

More than 10 rows available. Increase rows selector to view more rows.

10 rows returned in 0.02 seconds Download

2. Correct the Cartesian product produced in question 1 by creating an equijoin using a common column

ANS:

```
SELECT *
FROM d_play_list_items p, d_track_listings t
WHERE p.song_id = t.song_id;
```

EVENT_ID	SONG_ID	COMMENTS	SONG_ID	CD_NUMBER	TRACK
100	45	Play late	45	92	1
100	46	-	46	93	1
105	47	Play for the father	47	91	2
100	47	Play early	47	91	2
105	48	Play after cake cutting	48	95	5
105	49	Play first	49	91	3

3. Write a query to display the title, type, description, and artist from the DJs on Demand database.

ANS:

```
SELECT s.title, s.type_code, t.description, s.artist
FROM d_songs s, d_types t
WHERE s.type_code = t.code;
```

TITLE	TYPE_CODE	DESCRIPTION	ARTIST
Meet Me At the Altar	1	Jazz	Bobby West
Its Finally Over	12	Pop	The Hobbits
Im Going to Miss My Teacher	12	Pop	Jane Pop
Hurrah for Today	77	New Age	The Jubilant Trio
Lets Celebrate	77	New Age	The Celebrants
All These Years	88	Country	Diana Crooner

4. Rewrite the query in question 3 to select only those titles with an ID of 47 or 48.

ANS:

```
SELECT s.id, s.title, s.type_code, t.description, s.artist
FROM d_songs s, d_types t
WHERE s.type_code = t.code
   AND s.id IN (47, 48);
```

ID	TITLE	TYPE_CODE	DESCRIPTION	ARTIST
47	Hurrah for Today	77	New Age	The Jubilant Trio
48	Meet Me At the Altar	1	Jazz	Bobby West

5. Write a query that extracts information from three tables in the DJs on Demand database, the d_clients table, the d_events table, and the d_job_assignments table.

ANS:

```
SELECT
   c.client_number,
   c.first_name,
   c.last_name,
```

```
    c.phone,
    c.email,
    e.name,
    e.event_date,
    e.description,
    j.job_date,
    j.status
FROM d_clients c, d_events e, d_job_assignments j
WHERE c.client_number = e.client_number
    AND e.id = j.event_id;
```

CLIENT_NUMBER	FIRST_NAME	LAST_NAME	PHONE	EMAIL	NAME	EVENT_DATE	DESCRIPTION	JOB_DATE	STATUS
6133	Lauren	Vigil	4072220090	lbv@lbv.net	Vigil wedding	28-Apr-2004	Black tie at Four Season hotel	02-Feb-2004	Visited

6. Create and execute an equijoin between DJs on Demand tables d_track_listings and d_cds. Return the song_id and the title only.

ANS:

```
SELECT t.song_id, c.title
FROM d_track_listings t, d_cds c
WHERE t.cd_number = c.cd_number;
```

SONG_ID	TITLE
45	Back to the Shire
48	Here Comes the Bride
47	Party Music for All Occasions
49	Party Music for All Occasions
46	Songs from My Childhood

7. Mark T for the statements that are true and F for the statements that are false.
    a. A join is a type of query that gets data from more than one table based on columns with the same name.
        i. F
    b. To join tables using an equijoin, there must be a common column in both tables and that column is usually a primary key in one of the tables.
        i. T
    c. A Cartesian product occurs because the query does not specify a WHERE clause.
        i. T

d. Table aliases are required to create a join condition.
   i. **F**
e. If a table alias is used for a table name in the FROM clause, it must be substituted for the table name throughout the SELECT statement.
   i. **T**
f. Table alias must be only one character in length.
   i. **F**
g. A simple join or inner join is the same as an equijoin.
   i. **T**

8. What advantage does being able to combine data from multiple tables have for a business?
   a. **ANS:** Combining data from multiple tables allows businesses to view desired data in one go rather than needing to manually go through each individual table and ensuring that the desired data from the rows match up correctly. This helps save a significant amount of time, which can be spent in other critical business operations.

7.2 - Oracle Nonequijoins and Outer Joins

1. Create a join based on the cost of the event between the DJs on Demand tables D_EVENTS and D_PACKAGES. Show the name of the event and the code for each event.

ANS:

```
SELECT name, code
FROM d_events, d_packages
WHERE cost BETWEEN low_range AND high_range;
```

NAME	CODE
Vigil wedding	112
Peters Graduation	112

2. Using the Oracle database, create a query that returns the employee last name, salary, and job-grade level based on the salary. Select the salary between the lowest and highest salaries.

ANS:

```
SELECT last_name, salary, grade_level
FROM employees, job_grades
```

WHERE salary BETWEEN lowest_sal AND highest_sal;

LAST_NAME	SALARY	GRADE_LEVEL
Vargas	2500	A
Matos	2600	A
Heiden	2600	A
Davies	3100	B
Rajs	3500	B
Bell	3500	B
Stocks	3700	B
Fay	3900	B
TAYLOR	4000	B
Ricci	4100	B

More than 10 rows available. Increase rows selector to view more rows.

10 rows returned in 0.02 seconds Download

3. What condition requires the creation of a nonequijoin?
    a. **ANS:** The tables being used must not have a column with the same names

4. Rewrite the following nonequijoin statement using the logical condition operators (AND, OR, NOT): WHERE a.ranking BETWEEN g.lowest_rank AND g.highest_rank

ANS:
WHERE ranking BETWEEN lowest_rank AND highest_rank

5. How do you know when to use a table alias and when not to use a table alias?
    a. **ANS:** A table alias is needed when the tables being joined have one or more columns with the same name. When they do not have a same-named column, an alias is not needed. A tables alias is also very helpful when a table name is long, which can be tedious to type out; it helps to simplify queries and improve performance too.

6. What kind of join would you use if you wanted to find data between a range of numbers?
    a. ANS: nonequijoin

7. You need to produce a report for Global Fast Foods showing customers and orders. A customer must be included on the report even if the customer has had no orders.

ANS:

```
SELECT c.first_name, c.last_name, o.order_number, o.order_total
FROM f_customers c, f_orders o
WHERE c.id = o.cust_id(+);
```

FIRST_NAME	LAST_NAME	ORDER_NUMBER	ORDER_TOTAL
Cole	Bee	5678	103.02
Zoe	Twee	-	-

8. Create a query of the Oracle database that shows employee last names, department IDs, and department names. Include all employees even if they are not assigned to a department.

ANS:

```
SELECT e.last_name, e.department_id, d.department_name
FROM employees e, departments_pl d
WHERE e.department_id = d.department_id(+);
```

LAST_NAME	DEPARTMENT_ID	DEPARTMENT_NAME
Whalen	10	Administration
Hernandez	10	Administration
Ricci	10	Administration
Saikawa	10	Administration
Hartstein	20	Marketing
Reinhard	110	Accounting
Duric	110	Accounting
Loermans	110	Accounting
Grant	-	-

9. Modify the query in problem 8 to return all the department IDs even if no employees are assigned to them.

ANS:
SELECT e.last_name, e.department_id, d.department_name
FROM employees e, departments_pl d
WHERE e.department_id(+) = d.department_id;

LAST_NAME	DEPARTMENT_ID	DEPARTMENT_NAME
Ricci	10	Administration
Saikawa	10	Administration
Duric	110	Accounting
Loermans	110	Accounting
Reinhard	110	Accounting
-	-	Contracting

10. There are one or more errors in each of the following statements. Describe the errors and correct them.
    a. WHERE e.department_id(+) = d.department_id (+);
        i. **ANS:** Oracle syntax has no direct equivalent to a FULL OUTER JOIN.

           Corrected:
           employees e FULL OUTER JOIN departments d
           ON (e.department_id = d.department_id)

    b. SELECT e.employee id, e. last name, d. location id
       FROM employees, departments
       WHERE e.department_id = d.department_id(+);
        i. **ANS:**
            1. SELECT → The column names are incorrect because they are missing underscores to complete their names. There should also be no whitespace between each table alias and the column name that comes after.
            2. FROM → The table names are missing aliases that the SELECT statements refer to them as, which are “e” and “d.”

Corrected:

```
SELECT e.employee_id, e.last_name, d.location_id
FROM employees e, departments_pl d
WHERE e.department_id = d.department_id(+);
```

11. Create a query that will show all CD titles and song IDs in the DJs on Demand database even if there is no CD number in the track-listings table.

ANS:

```
SELECT c.title, t.song_id
FROM d_cds c, d_track_listings t
WHERE c.cd_number = t.cd_number(+);
```

TITLE	SONG_ID
Back to the Shire	45
Songs from My Childhood	46
Party Music for All Occasions	47
Here Comes the Bride	48
Party Music for All Occasions	49
Whirled Peas	-
The Celebrants Live in Concert	-
Graduation Songbook	-
Carpe Diem	-

12. How many times has someone asked you: “What do you want to be when you grow up?” For most of us, the first thing that comes to mind is something like business manager, engineer, teacher, game designer, doctor, scientist, computer programmer, or accountant -- all pretty much traditional career choices. Have you ever thought about working in an odd job or nontraditional career? There are people who are professional shoppers for busy executives, directors of zoos, recipe designers, insecticide chemists, golf-course designers, and turf managers. Picture yourself in a dream job or nontraditional career doing something that you think would be interesting, life fulfilling, and profitable.

    Use Internet resources to explore your idea. Write a brief description of the job to share with the class.

ANS:

I think being a voice actor (VA) could be fun. This job involves performing voice overs for new shows and also dubbing over existing shows too. The VA would need to articulate well, understand the character they're voicing in order to ensure their portrayal is accurate, and be adaptable since the voice director may want different takes of the same line.

8.1 - Group Functions
Vocab

Word	Definition
AVG	Calculates average value excluding nulls
COUNT	Returns the number of rows with non-null values for the expression
STDDEV	For two sets of data with approximately the same mean, the greater the spread, the greater the standard deviation
GROUP Functions	Operate on sets of rows to give one result per group
MIN	Returns minimum value ignoring nulls
VARIANCE	Used with columns that store numeric data to calculate the spread of data around the mean
SUM	Calculates the sum ignoring null values
MAX	Returns the maximum value ignoring nulls
SUM	To gather into a sum or whole

1. Define and give an example of the seven group functions: AVG, COUNT, MAX, MIN, STDDEV, SUM, and VARIANCE.

   ANS:
   a. AVG: calculates the average of a numerical data type column, excluding the null values
      i. **E.g.** SELECT AVG(cost) FROM d_events
   b. COUNT: counts the total number of rows without a null, assuming a column name is specified. If an asterisk (*) is used as a parameter, the total number of rows will be counted with and without the nulls.
      i. **E.g.** SELECT COUNT(name) FROM d_events
   c. MAX: returns the highest value in the column, ignoring nulls

i. Numbers → returns highest number
    1. **E.g.** SELECT MAX(cost) FROM d_events
ii. Date → returns the most recent date
    1. **E.g.** SELECT MAX(event_date) FROM d_events
iii. String → returns the string whose first letter is furthest in the alphabet
    1. **E.g.** SELECT MAX(cost) FROM d_events

d. MIN: returns the lowest value in the column, ignoring nulls. Numbers, date, and string returned will be opposite of MAX
    i. **E.g.** SELECT MIN(event_date) FROM d_events
e. STDDEV: measures data spread; the greater the spread, the higher the standard deviation
    i. **E.g.** SELECT STDDEV(cost) FROM d_events
f. SUM: adds up all the non-null values in a column together
    i. **E.g.** SELECT SUM(cost) FROM d_events
g. VARIANCE: calculates data spread around mean. Variance is less the more values are closer to the column's mean.
    i. **E.g.** SELECT VARIANCE(cost) FROM d_events

2. Create a query that will show the average cost of the DJs on Demand events. Round to two decimal places.

ANS:

```
SELECT AVG(cost) "Average Cost of Events"
FROM d_events;
```

Average Cost of Events
9000

3. Find the average salary for Global Fast Foods staff members whose manager ID is 19.

ANS:

```
SELECT AVG(salary) "Avg Salary for Staff Members"
FROM f_staffs
WHERE manager_id = 19;
```

Avg Salary for Staff Members
8.375

4. Find the sum of the salaries for Global Fast Foods staff members whose IDs are 12 and 9.

ANS:

```
SELECT SUM(salary) "Salary Sum for Staff with ID 12 or 9"
FROM f_staffs
WHERE id IN (12, 9);
```

Salary Sum for Staff with ID 12 or 9
16.75

5. Using the Oracle database, select the lowest salary, the most recent hire date, the last name of the person who is at the top of an alphabetical list of employees, and the last name of the person who is at the bottom of an alphabetical list of employees. Select only employees who are in departments 50 or 60.

ANS:

```
SELECT MIN(salary), MAX(hire_date), MIN(last_name), MAX(last_name)
FROM employees
WHERE department_id IN (50, 60);
```

MIN(SALARY)	MAX(HIRE_DATE)	MIN(LAST_NAME)	MAX(LAST_NAME)
2500	06-Jul-2015	Bell	Vargas

6. Your new Internet business has had a good year financially. You have had 1,289 orders this year. Your customer order table has a column named total_sales. If you submit the following query, how many rows will be returned?

   ```
   SELECT sum(total_sales)
   FROM orders;
   ```

ANS: Just one row because GROUP functions only return one result

7. You were asked to create a report of the average salaries for all employees in each division of the company. Some employees in your company are paid hourly instead of by salary. When you ran the report, it seemed as though the averages were not what you expected—they were much higher than you thought! What could have been the cause?

ANS: The salary of an employee who is paid hourly is likely null in the table. AVG is a GROUP function, so these null values were ignored, resulting in a higher average.

8. Employees of Global Fast Foods have birth dates of July 1, 1980, March 19, 1979, and March 30, 1969. If you select MIN(birthdate), which date will be returned?

ANS: March 30, 1969 because it is the oldest birth date. Using MIN() in a date data type column would result in the oldest/earliest date returned. MAX() would return the most recent date.

9. Create a query that will return the average order total for all Global Fast Foods orders from January 1, 2002, to December 21, 2002.

ANS:

```
SELECT AVG(order_total)
FROM f_orders
WHERE order_date BETWEEN '01-Jan-2002' AND '21-Dec-2022';
```

AVG(ORDER_TOTAL)
103.02

10. What was the hire date of the last Oracle employee hired?

ANS:

```
SELECT MAX(hire_date) "Last employee hire date"
FROM employees;
```

Last employee hire date
16-Dec-2015

11. In the following SELECT clause, which value returned by the SELECT statement will be larger? SELECT SUM(operating_cost), AVG(operating_cost)

ANS: The sum will be larger because calculating the sum does not involve dividing by anything unlike the average

12. Refer to the DJs on Demand database D_EVENTS table:

    Which code is valid as part of an SQL query?

    a. FROM event_date
    b. SELECT SUM(cost)
    c. SELECT SUM(event_date)
    d. SELECT AVG(cost) AS "Expense"

e. WHERE MIN(id) = 100
f. SELECT MAX(AVG(cost))
g. SELECT MIN(event_date)

8.2 - COUNT, DISTINCT, NVL
Vocab

Word	Definition
COUNT (expression)	Returns the number of non-null values in the expression column
DISTINCT	The keyword used to return only non-duplicate values or combinations of non-duplicate values in a query.
COUNT (DISTINCT expression)	Returns the number of unique non-null values in the expression column.

1. How many songs are listed in the DJs on Demand D_SONGS table?

ANS:
SELECT COUNT(title) total_songs
FROM d_songs;

TOTAL_SONGS
6

2. In how many different location types has DJs on Demand had venues?

ANS:
SELECT COUNT(DISTINCT loc_type) unique_location_types
FROM d_venues;

UNIQUE_LOCATION_TYPES
4

3. The d_track_listings table in the DJs on Demand database has a song_id column and a cd_number column. How many song IDs are in the table and how many different CD numbers are in the table?

ANS:

```
SELECT COUNT(song_id) total_song_ids, COUNT(DISTINCT cd_number)
unique_cd_numbers
FROM d_track_listings;
```

TOTAL_SONG_IDS	UNIQUE_CD_NUMBERS
5	4

4. How many of the DJs on Demand customers have email addresses?

ANS:

```
SELECT COUNT(email)
FROM d_clients;
```

COUNT(EMAIL)
3

5. Some of the partners in DJs on Demand do not have authorized expense amounts (auth_expense_amt). How many partners do have this privilege?

ANS:

```
SELECT COUNT(auth_expense_amt)
FROM d_partners;
```

COUNT(AUTH_EXPENSE_AMT)
1

6. What values will be returned when the statement below is issued?

ID	type	shoe_color
456	oxford	brown
463	sandal	tan
262	heel	black
433	slipper	tan

```
SELECT COUNT(shoe_color), COUNT(DISTINCT shoe_color)
FROM shoes;
```

ANS: COUNT(shoe_color) would return the total non-null shoe colors in the column, including duplicates. COUNT(DISTINCT shoe_color) does the same, except returning duplicates.

7. Create a query that will convert any null values in the auth_expense_amt column on the DJs on Demand D_PARTNERS table to 100000 and find the average of the values in this column. Round the result to two decimal places.

ANS:

```
SELECT ROUND(AVG(NVL(auth_expense_amt, 100000)), 2)
FROM d_partners;
```

ROUND(AVG(NVL(AUTH_EXPENSE_AMT,100000)),2)
166666.67

8. Which statement(s) is/are True about the following SQL statement:
   SELECT AVG(NVL(selling_bonus, 0.10)) FROM bonuses;
   a. The datatypes of the values in the NVL clause can be any data type except date data.
   b. If the selling_bonus column has a null value, 0.10 will be substituted.
   c. There will be no null values in the selling_bonus column when the average is calculated.
      i. NOTE: True in the sense that during the calculation process, any nulls in the selling_bonus column are replaced with 0.10. However, it does not permanently change the null to 0.10 after the calculation is complete.
   d. This statement will cause an error. There cannot be two functions in the SELECT statement.

9. Which of the following statements is/are TRUE about the following query?
   SELECT DISTINCT colors, sizes FROM items;
   a. Each color will appear only once in the result set.
   b. Each size will appear only once in the result set.

c. Unique combinations of color and size will appear only once in the result set.
d. Each color and size combination will appear more than once in the result set.