

## SQL Database Programming: Section 8-1: Group Functions

### Vocabulary

**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 FUNCTION** – 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

**COUNT** - To gather into a sum or whole

### Try It / Solve It

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

- **AVG**
  - Used with columns that store numeric data to compute the average
  - Example: `SELECT AVG(SALARY) FROM EMPLOYEES;`
- **COUNT**
  - Returns the number of rows
  - Example: `SELECT COUNT(*) FROM F_STAFFS;`
- **MAX**
  - Used with columns that store any data type to return the maximum value
  - Example: `SELECT MAX(SALARY) FROM F_STAFFS;`

- **MIN**
  - Used with columns that store any data type to return the minimum value
  - Example: `SELECT MIN(SALARY) FROM F_STAFFS;`
- **STDDEV**
  - Standard deviation measures the spread of data. For two sets of data with approximately the same mean, the greater the spread, the greater the standard deviation.
  - Example: `SELECT STDDEV(SALARY) FROM F_STAFFS;`
- **SUM**
  - Used with columns that store numeric data to find the total or sum of values
  - Example: `SELECT SUM(PRICE) FROM F_FOOD_ITEMS;`
- **VARIANCE**
  - Used with columns that store numeric data to calculate the spread of data around the mean
  - Example: `SELECT VARIANCE(PRICE) FROM F_FOOD_ITEMS;`

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

The screenshot shows the APEX SQL Workshop interface. The SQL Commands panel contains the following query:

```
1
2 SELECT AVG(COST) AS "AVERAGE COST" FROM D_EVENTS;
```

The Results panel shows the output of the query:

AVERAGE COST
9000

1 rows returned in 0.00 seconds

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

The screenshot shows the APEX SQL Workshop interface. The SQL Commands panel contains the following query:

```
1
2 SELECT AVG(SALARY) AS "AVERAGE SALARY"
3 FROM F_STAFFS
4 WHERE MANAGER_ID = 19;
```

The Results panel shows the output of the query:

AVERAGE SALARY
8.375

1 rows returned in 0.01 seconds

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

The screenshot shows the APEX SQL Workshop interface. The SQL command entered is:

```

1
2 SELECT SUM(SALARY) AS "SUM of the SALARIES of STAFFS with ID of 12 and 9"
3 FROM F_STAFFS
4 WHERE ID = 12 OR ID = 9;

```

The results tab shows the following output:

SUM of the SALARIES of STAFFS with ID of 12 and 9
16.75

1 rows returned in 0.00 seconds

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.

The screenshot shows the APEX SQL Workshop interface. The SQL command entered is:

```

1
2 SELECT
3     MIN(SALARY) AS "LOWEST SALARY",
4     MAX(HIRE_DATE) AS "MOST RECENT HIRE DATE",
5     MIN(LAST_NAME) AS "FIRST NAME in the ALPHABETICAL LIST",
6     MAX(LAST_NAME) AS "LAST NAME in the ALPHABETICAL LIST"
7 FROM EMPLOYEES
8 WHERE DEPARTMENT_ID = 50 OR DEPARTMENT_ID = 60;

```

The results tab shows the following output:

LOWEST SALARY	MOST RECENT HIRE DATE	FIRST NAME in the ALPHABETICAL LIST	LAST NAME in the ALPHABETICAL LIST
2500	06-Jul-2015	Bell	Vargas

1 rows returned in 0.01 seconds

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?

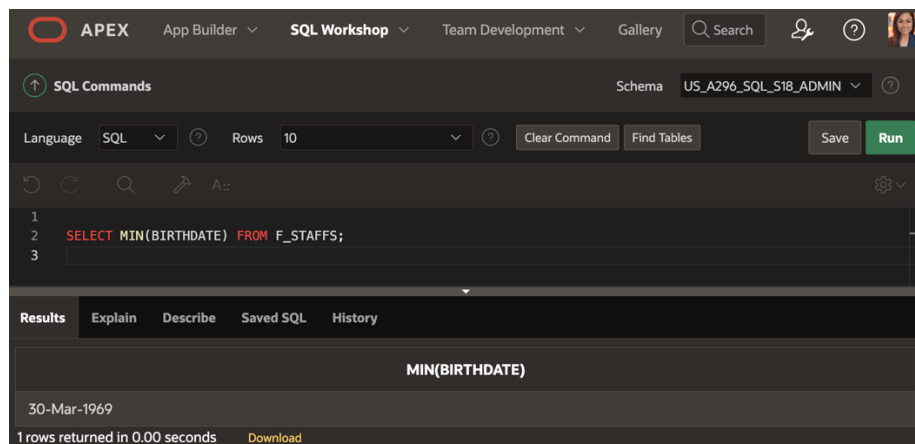
```
SELECT sum(total_sales)
FROM orders;
```

→ It will return ONE row. SUM combines values from multiple rows into a single value which will produce a single row.

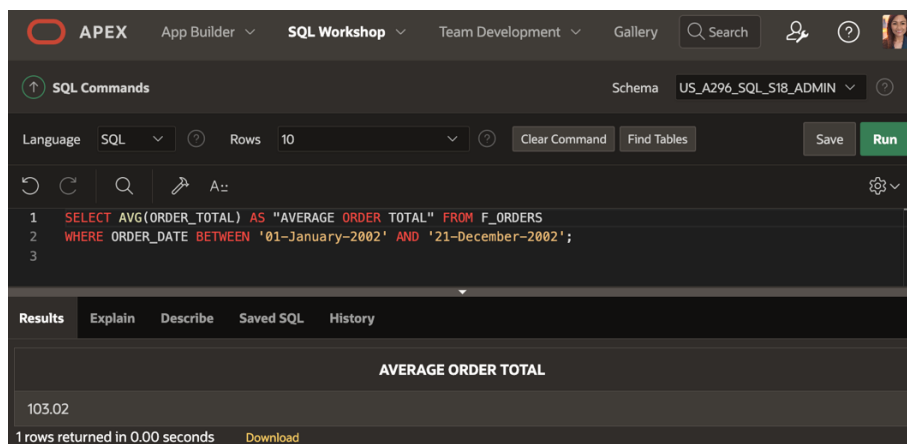
7. You were asked to create a report of the average salaries for all employees in each division of the company. Some employees in your company are paid hourly instead of by salary. When you ran the report, it seemed as though the averages were not what you expected—they were much higher than you thought!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

→ The culprit could be due to employees that are paid hourly instead of by salary. Calculating the average salary and the average hourly rate separately can possibly solve this issue.

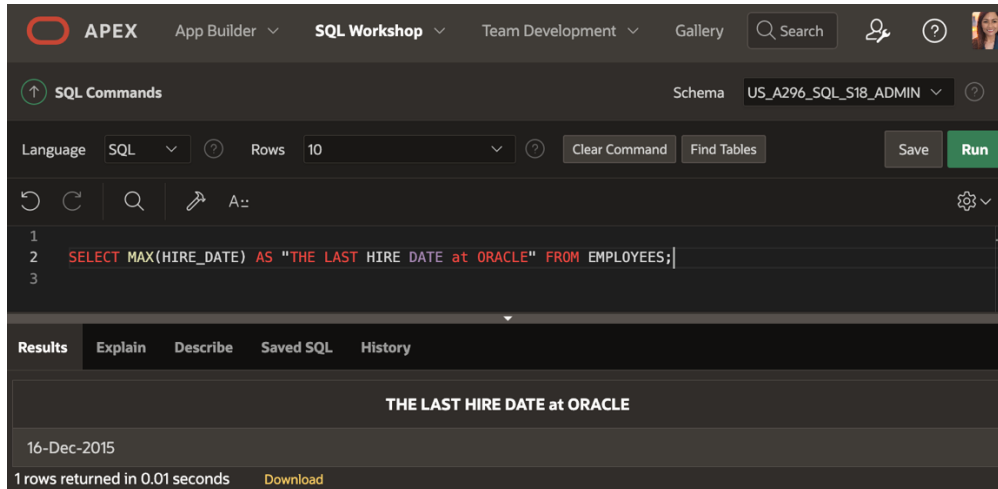
8. Employees of Global Fast Foods have birth dates of July 1, 1980, March 19, 1979, and March 30, 1969. If you select MIN(birthdate), which date will be returned? -- **March 19, 1979**



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.



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



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

→ It's not possible to say exactly which will be larger without knowing the values in the 'operating cost' column. Although if the values are lots of small amounts or small costs, then the total sum would come out larger/higher. However, if there are few large costs combined with small costs, the average might come out bigger/larger than the sum.

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)` → this is because it calculates the COST column in the "D\_EVENTS" table.
- \_\_\_\_\_ c. `SELECT SUM(event_date)`
- \_\_\_\_\_ d. `SELECT AVG(cost) AS "Expense"` → this is because it calculates the 'average cost', and 'Expense' is the alias.
- \_\_\_\_\_ e. WHERE MIN(id) = 100
- \_\_\_\_\_ f. `SELECT MAX(AVG(cost))`
- \_\_\_\_\_ g. `SELECT MIN(event_date)` → this is because 'MIN' shows the earliest event date from the 'EVENT\_DATE' column.

---

## SQL Database Programming: Section 8-2: Count, Distinct, NVL

### Vocabulary

**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.

### Try It / Solve It

1. How many songs are listed in the DJs on Demand D\_SONGS table?

The screenshot shows the APEX SQL Workshop interface. The top navigation bar includes 'APEX', 'App Builder', 'SQL Workshop', 'Team Development', and 'Gallery'. The 'SQL Commands' tab is active, showing a query in the editor:

```
1  
2 SELECT COUNT(TITLE) AS "Total List of Songs"  
3 FROM D_SONGS;
```

The 'Schema' dropdown is set to 'US\_A296\_SQL\_S18\_ADMIN'. The 'Language' is 'SQL' and 'Rows' is set to 10. The 'Run' button is highlighted. Below the editor, the 'Results' tab is selected, displaying the query results:

Total List of Songs
6

At the bottom, it indicates '1 rows returned in 0.00 seconds' and provides a 'Download' link.

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

The screenshot shows the APEX SQL Workshop interface. The top navigation bar includes 'APEX', 'App Builder', 'SQL Workshop', 'Team Development', and 'Gallery'. The 'SQL Commands' tab is active, showing a query in the editor:

```
1  
2 SELECT COUNT(LOC_TYPE) AS "Total Different Location Types"  
3 FROM D_VENUES;
```

The 'Schema' dropdown is set to 'US\_A296\_SQL\_S18\_ADMIN'. The 'Language' is 'SQL' and 'Rows' is set to 10. The 'Run' button is highlighted. Below the editor, the 'Results' tab is selected, displaying the query results:

Total Different Location Types
6

At the bottom, it indicates '1 rows returned in 0.01 seconds' and provides a 'Download' link.

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?

The screenshot shows the APEX SQL Workshop interface. The SQL Commands tab is active, displaying a query to count songs and distinct CD numbers from the d\_track\_listings table. The Results tab shows a table with two columns: Total Songs and Total Different CD Numbers. The results are 5 and 4 respectively. The query is: `SELECT COUNT(SONG_ID) AS "Total Songs", COUNT(DISTINCT CD_NUMBER) AS "Total Different CD Numbers" FROM D_TRACK_LISTINGS;`

Total Songs	Total Different CD Numbers
5	4

1 rows returned in 0.00 seconds [Download](#)

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

The screenshot shows the APEX SQL Workshop interface. The SQL Commands tab is active, displaying a query to count customers with email addresses from the d\_clients table. The Results tab shows a table with one column: CUSTOMERS WITH EMAIL ADDRESSES. The result is 3. The query is: `SELECT COUNT(EMAIL) AS "CUSTOMERS WITH EMAIL ADDRESSES" FROM D_CLIENTS;`

CUSTOMERS WITH EMAIL ADDRESSES
3

1 rows returned in 0.02 seconds [Download](#)

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?

The screenshot shows the APEX SQL Workshop interface. The SQL Commands tab is active, displaying the following query:

```

1
2 SELECT COUNT(AUTH_EXPENSE_AMT) AS "Partners with No Authorized Expense Amounts"
3 FROM D_PARTNERS;

```

The Results tab shows the query execution results:

Partners with No Authorized Expense Amounts
1

1 rows returned in 0.01 seconds

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;

- COUNT(shoe\_color) – 4 [brown, tan x2, black]
- COUNT(DISTINCT shoe\_color) – 3 [brown, tan, black]

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.

The screenshot shows the APEX SQL Workshop interface. The SQL Commands tab is active, displaying the following query:

```

1
2 SELECT ROUND(AVG(NVL(AUTH_EXPENSE_AMT, 100000)), 2)
3 FROM D_PARTNERS;

```

The Results tab shows the query execution results:

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

1 rows returned in 0.00 seconds



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 datatype 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.
- ☐ 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.