MA RAQUEL SORIANO
HHA 507

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;

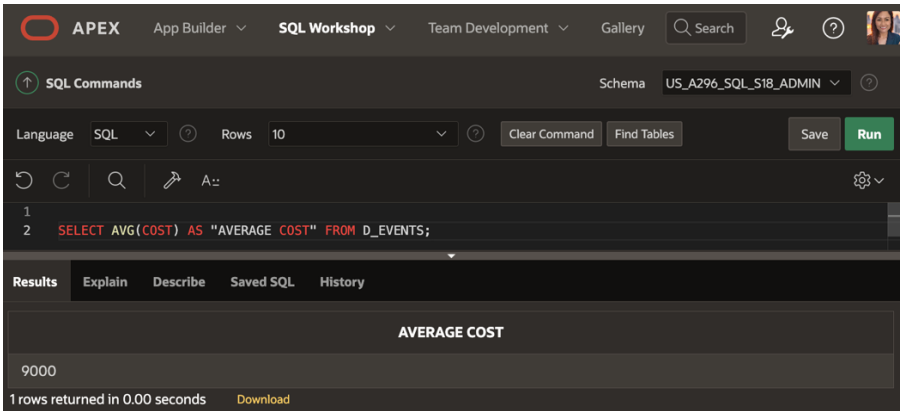
- **MIN**
  - Used with columns that store any data type to return the minimum value
  - Example: SELECT MIN(SALARY) FROM F_STAFFS;

- **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;

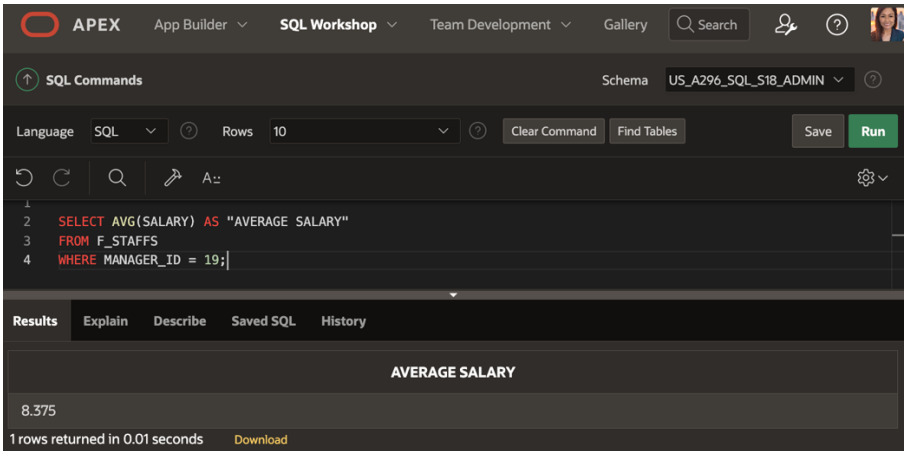
- **SUM**
  - Used with columns that store numeric data to find the total or sum of values
  - Example: SELECT SUM(PRICE) FROM F_FOOD_ITEMS;

- **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;

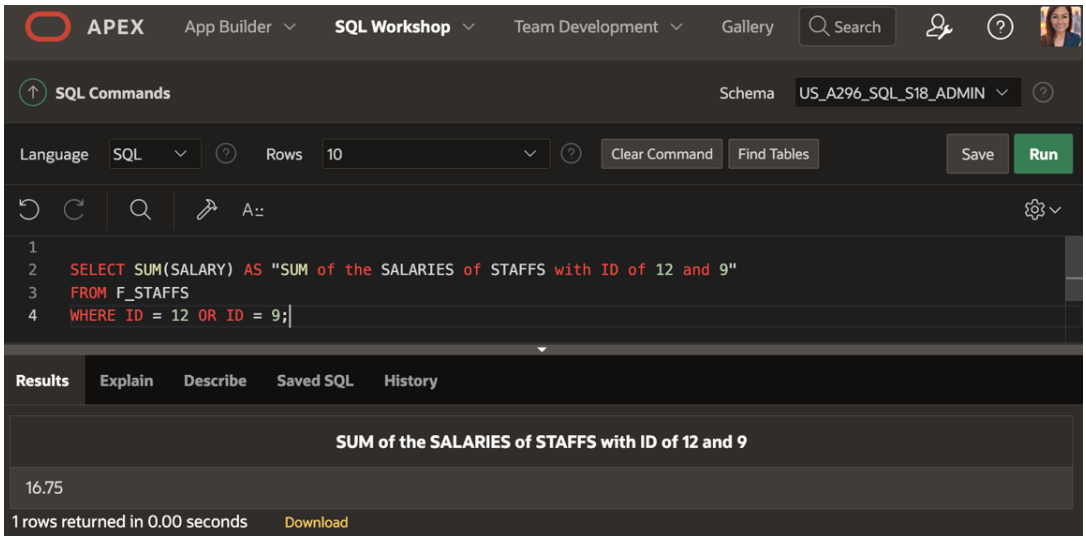
2. Create a query that will show the average cost of the DJs on Demand events. Round to two decimal places.



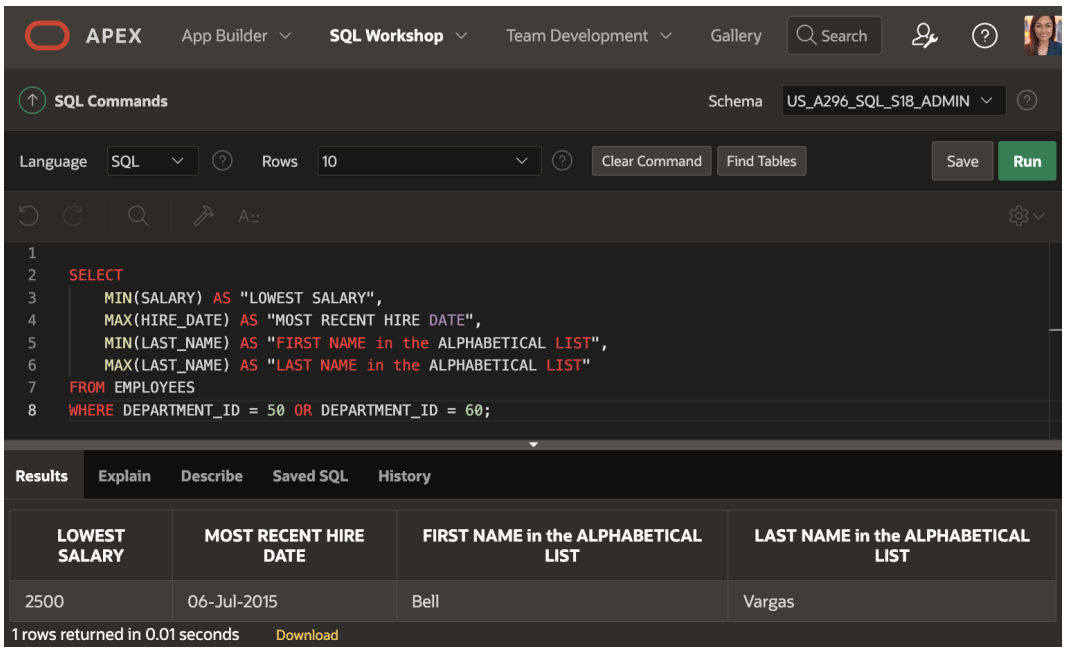
3. Find the average salary for Global Fast Foods staff members whose manager ID is 19.



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



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.



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?

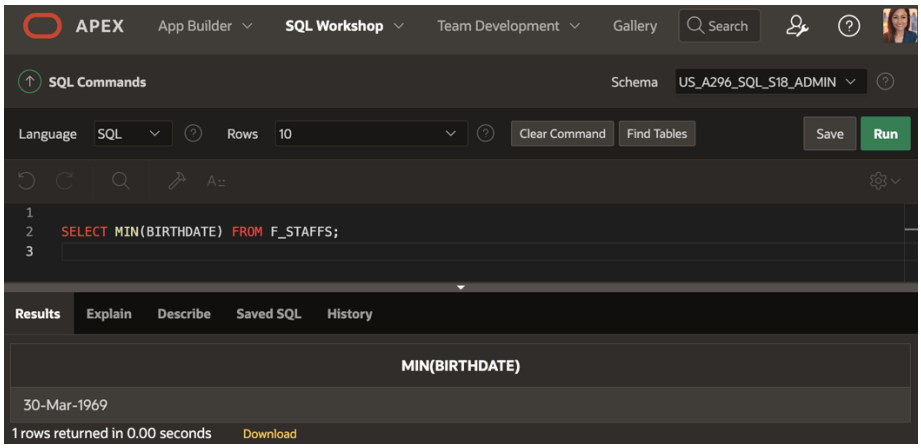
```
SELECT sum(total_sales)
FROM orders;
```

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

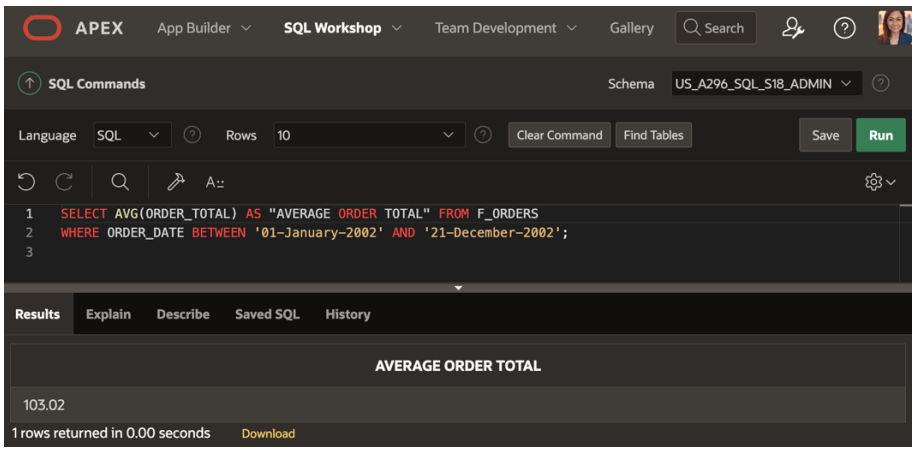
7. You were asked to create a report of the average salaries for all employees in each division of the company. Some employees in your company are paid hourly instead of by salary. When you ran the report, it seemed as though the averages were not what you expected—they were much higher than you thought!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

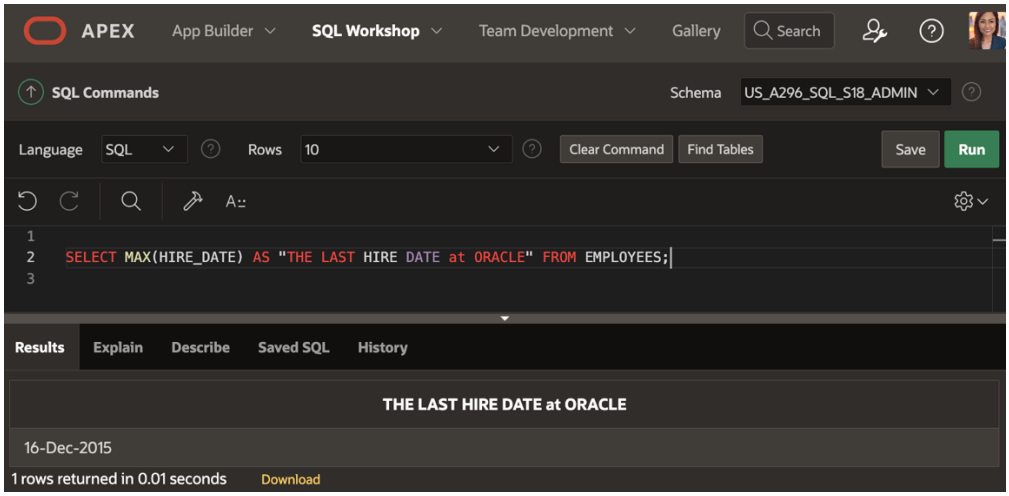
8. Employees of Global Fast Foods have birth dates of July 1, 1980, March 19, 1979, and March 30, 1969. If you select MIN(birthdate), which date will be returned? -- March 19, 1979



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.