



databricks

Academy

DAWD 02-6 – Lab – Basic SQL

The two fields below are used to customize queries used in this course. Enter your schema (database) name and username, and press "Enter" to populate necessary information in the queries on this page.

Schema Name:

Username:



Lesson Objective

At the end of this lesson, you will be able to:

- Write basic SQL queries to subset tables using Databricks SQL
 - Join multiple tables together to create a new table
 - Aggregate data columns using SQL functions to answer defined business questions
-



Retrieve Data

The statement we are using in this part of the lab implements `SELECT`, `SELECT ... AS`, `GROUP BY`, and `ORDER BY`. Note that `FROM`, `GROUP BY`, and `ORDER BY` need to occur in a specific order, or an error will be thrown.

Complete the following:

1. Make the required changes to the query below
2. Run the query in Databricks SQL
3. Check your work by entering your answer to the question
4. After pressing ENTER/RETURN, green indicates a correct answer, and red indicates incorrect

```
USE hive_metastore.class_013_odg7_da_dawd;  
SELECT loyalty_segment, count(loyalty_segment) AS Count  
  FILL_IN customers  
  FILL_IN BY loyalty_segment  
  FILL_IN BY loyalty_segment;
```

[Show Answer](#)[Copy](#)

How many customers are in loyalty_segment O?

Use Column Expressions

The `customers` table contains the column `customer_name`, which has the same problem as the `city` column had in the last lesson. The customer names are all in lower-case. Run a `SELECT` query, using the `initcap()` function to examine the results of using this function on the `customer_name`.

Complete the following:

1. Make the required changes to the query below
2. Run the query in Databricks SQL
3. Check your work by entering your answer to the question
4. After pressing ENTER/RETURN, green indicates a correct answer, and red indicates incorrect

```
USE hive_metastore.class_013_odg7_da_dawd;  
SELECT FILL_IN(FILL_IN) AS Customer_Name  
  FROM customers  
  ORDER BY customer_name DESC;
```

[Show Answer](#)[Copy](#)

What is the last name of the last customer (alphabetically) in the table? Ensure answer is in lower case

Update Data

Let's go ahead and implement the changes we examined in the last section. Use an `UPDATE` statement to change the data in the `customers` table.

Complete the following:

1. Make the required changes to the query below
2. Run the query in Databricks SQL
3. Check your work by entering your answer to the question
4. After pressing `ENTER/RETURN`, green indicates a correct answer, and red indicates incorrect

```
USE hive_metastore.class_013_odg7_da_dawd;  
FILL_IN customers FILL_IN customer_name = initcap(customer_name);  
SELECT * FROM customers;
```

[Show Answer](#)[Copy](#)

In which city is Bittner Engineering, Inc. located?

Insert Data

We can implement a fifth loyalty segment in our `loyalty_segments` table. We will use a `unit_threshold` of 130 units.

Complete the following:

1. Make the required changes to the query below
2. Run the query in Databricks SQL
3. Check your work by entering your answer to the question
4. After pressing `ENTER/RETURN`, green indicates a correct answer, and red indicates incorrect

```
USE hive_metastore.class_013_odg7_da_dawd;  
FILL_IN INTO loyalty_segments  
  (loyalty_segment_id, loyalty_segment_description, unit_threshold, valid_from, valid_to)  
  FILL_IN  
  (5, 'level_5', 130, current_date(), Null);  
SELECT * FROM loyalty_segments;
```

[Show Answer](#)[Copy](#)

How many rows are in the table now?



CREATE OR REPLACE VIEW AS statement

In this part of the lab, we are going to create a view that contains only those customers in California (CA) who have a `loyalty_segment` of 3.

Complete the following:

1. Make the required changes to the query below
2. Run the query in Databricks SQL
3. Check your work by entering your answer to the question
4. After pressing ENTER/RETURN, green indicates a correct answer, and red indicates incorrect

```
USE hive_metastore.class_013_odg7_da_dawd;  
CREATE OR REPLACE FILL_IN high_value_CA_customers AS  
  SELECT *  
    FROM customers  
   WHERE state = 'FILL_IN'  
   AND loyalty_segment = FILL_IN;  
SELECT * FROM high_value_CA_customers;
```

[Show Answer](#)[Copy](#)

How many rows are in the view?



Joins

Let's use a `INNER JOIN`, `GROUP BY`, `ORDER BY`, and a function to calculate the total dollar amount of sales to various states in the sales table. Note that `INNER JOIN` is the default join type, so we can just type `JOIN`.

Complete the following:

1. Make the required changes to the query below
2. Run the query in Databricks SQL
3. Check your work by entering your answer to the question
4. After pressing `ENTER/RETURN`, green indicates a correct answer, and red indicates incorrect

```
USE hive_metastore.class_013_odg7_da_dawd;  
FILL_IN customers.state, sum(total_price) AS Total FROM customers  
  FILL_IN sales  
  FILL_IN customers.customer_id = sales.customer_id  
  GROUP BY customers.state  
  ORDER BY Total DESC;
```

[Show Answer](#)[Copy](#)

Which state has the highest sales?



Use Aggregations

We have already used a few aggregate functions in the lab. Let's finish by running a simple aggregation to find the best price our suppliers have been able to provide on a specific piece of vintage electronics, a 5-disk CD changer. The best price means the lowest price, and we can use the `min()` function to make this calculation.

Complete the following:

1. Make the required changes to the query below
2. Run the query in Databricks SQL

3. Check your work by entering your answer to the question
4. After pressing ENTER/RETURN, green indicates a correct answer, and red indicates incorrect

```
USE hive_metastore.class_013_odg7_da_dawd;  
SELECT FILL_IN(total_price)  
FROM sales;
```

[Show Answer](#)[Copy](#)

What is the lowest price? (numbers only)

© 2023 Databricks, Inc. All rights reserved.

Apache, Apache Spark, Spark and the Spark logo are trademarks of the [Apache Software Foundation](#).

[Privacy Policy](#) | [Terms of Use](#) | [Support](#)

1.2.13