

# **Databricks-Certified-Data-Analyst-Associate Exams Torrent & Actual Databricks-Certified-Data-Analyst- Associate Tests**



P.S. Free & New Databricks-Certified-Data-Analyst-Associate dumps are available on Google Drive shared by Exams-boost:  
<https://drive.google.com/open?id=1fWrzAnWwQ1d2X8qP0vI70AAIn6FCTV6p>

There are totally three versions of Databricks-Certified-Data-Analyst-Associate practice materials which are the most suitable versions for you: PDF, Software and APP online versions. We promise ourselves and exam candidates to make these Databricks-Certified-Data-Analyst-Associate learning materials top notch. So if you are in a dark space, our Databricks-Certified-Data-Analyst-Associate Exam Questions can inspire you make great improvements. Just believe in our Databricks-Certified-Data-Analyst-Associate training guide and let us lead you to a brighter future!

The Databricks Databricks-Certified-Data-Analyst-Associate exam questions pdf is properly formatted to give candidates the asthetic and unformatted information they need to succeed in the Databricks-Certified-Data-Analyst-Associate exam. In addition to the comprehensive material, a few basic and important questions are highlighted and discussed in the Databricks-Certified-Data-Analyst-Associate Exam Material file. These questions are repeatedly seen in past Databricks Certified Data Analyst Associate Exam exam papers. The Databricks Certified Data Analyst Associate Exam practice questions are easy to access and can be downloaded anytime on your mobile, laptop, or MacBook.

>> **Databricks-Certified-Data-Analyst-Associate Exams Torrent** <<

## **Quiz Databricks-Certified-Data-Analyst-Associate Exams Torrent - Realistic Actual Databricks Certified Data Analyst Associate Exam Tests**

The Databricks Databricks-Certified-Data-Analyst-Associate online exam is the best way to prepare for the Databricks Databricks-Certified-Data-Analyst-Associate exam. Exams-boost has a huge selection of Databricks-Certified-Data-Analyst-Associate dumps and topics that you can choose from. The Databricks-Certified-Data-Analyst-Associate Exam Questions are categorized into specific areas, letting you focus on the Databricks Databricks-Certified-Data-Analyst-Associate subject areas you need to work on.

## **Databricks Databricks-Certified-Data-Analyst-Associate Exam Syllabus Topics:**

Topic	Details

Topic 1	<ul style="list-style-type: none"> <li>• Databricks SQL: This topic discusses key and side audiences, users, Databricks SQL benefits, complementing a basic Databricks SQL query, schema browser, Databricks SQL dashboards, and the purpose of Databricks SQL endpoints</li> <li>• warehouses. Furthermore, the delves into Serverless Databricks SQL endpoint</li> <li>• warehouses, trade-off between cluster size and cost for Databricks SQL endpoints</li> <li>• warehouses, and Partner Connect. Lastly it discusses small-file upload, connecting Databricks SQL to visualization tools, the medallion architecture, the gold layer, and the benefits of working with streaming data.</li> </ul>
Topic 2	<ul style="list-style-type: none"> <li>• Analytics applications: It describes key moments of statistical distributions, data enhancement, and the blending of data between two source applications. Moreover, the topic also explains last-mile ETL, a scenario in which data blending would be beneficial, key statistical measures, descriptive statistics, and discrete and continuous statistics.</li> </ul>
Topic 3	<ul style="list-style-type: none"> <li>• SQL in the Lakehouse: It identifies a query that retrieves data from the database, the output of a SELECT query, a benefit of having ANSI SQL, access, and clean silver-level data. It also compares and contrasts MERGE INTO, INSERT TABLE, and COPY INTO. Lastly, this topic focuses on creating and applying UDFs in common scaling scenarios.</li> </ul>
Topic 4	<ul style="list-style-type: none"> <li>• Data Visualization and Dashboarding: Sub-topics of this topic are about of describing how notifications are sent, how to configure and troubleshoot a basic alert, how to configure a refresh schedule, the pros and cons of sharing dashboards, how query parameters change the output, and how to change the colors of all of the visualizations. It also discusses customized data visualizations, visualization formatting, Query Based Dropdown List, and the method for sharing a dashboard.</li> </ul>
Topic 5	<ul style="list-style-type: none"> <li>• Data Management: The topic describes Delta Lake as a tool for managing data files, Delta Lake manages table metadata, benefits of Delta Lake within the Lakehouse, tables on Databricks, a table owner's responsibilities, and the persistence of data. It also identifies management of a table, usage of Data Explorer by a table owner, and organization-specific considerations of PII data. Lastly, the topic it explains how the LOCATION keyword changes, usage of Data Explorer to secure data.</li> </ul>

## Databricks Certified Data Analyst Associate Exam Sample Questions (Q56-Q61):

### NEW QUESTION # 56

Consider the following two statements:

Statement 1:

Statement 2:

Which of the following describes how the result sets will differ for each statement when they are run in Databricks SQL?

- A. There is no difference between the result sets for both statements.
- B. The first statement will return all data from the customers table and matching data from the orders table. The second statement will return all data from the orders table and matching data from the customers table. Any missing data will be filled in with NULL.
- C. Both statements will fail because Databricks SQL does not support those join types.
- D. When the first statement is run, all rows from the customers table will be returned and only the customer\_id from the orders table will be returned. When the second statement is run, only those rows in the customers table that do not have at least one match with the orders table on customer\_id will be returned.
- E. When the first statement is run, only rows from the customers table that have at least one match with the orders table on customer\_id will be returned. When the second statement is run, only those rows in the customers table that do not have at least one match with the orders table on customer\_id will be returned.

**Answer: E**

Explanation:

Based on the images you sent, the two statements are SQL queries for different types of joins between the customers and orders tables. A join is a way of combining the rows from two table references based on some criteria. The join type determines how the rows are matched and what kind of result set is returned. The first statement is a query for a LEFT SEMI JOIN, which returns only

the rows from the left table reference (customers) that have a match with the right table reference (orders) on the join condition (customer\_id). The second statement is a query for a LEFT ANTI JOIN, which returns only the rows from the left table reference (customers) that have no match with the right table reference (orders) on the join condition (customer\_id). Therefore, the result sets for the two statements will differ in the following way:

The first statement will return a subset of the customers table that contains only the customers who have placed at least one order. The number of rows returned will be less than or equal to the number of rows in the customers table, depending on how many customers have orders. The number of columns returned will be the same as the number of columns in the customers table, as the LEFT SEMI JOIN does not include any columns from the orders table.

The second statement will return a subset of the customers table that contains only the customers who have not placed any order. The number of rows returned will be less than or equal to the number of rows in the customers table, depending on how many customers have no orders. The number of columns returned will be the same as the number of columns in the customers table, as the LEFT ANTI JOIN does not include any columns from the orders table.

The other options are not correct because:

A) The first statement will not return all data from the customers table, as it will exclude the customers who have no orders. The second statement will not return all data from the orders table, as it will exclude the orders that have a matching customer. Neither statement will fill in any missing data with NULL, as they do not return any columns from the other table.

C) There is a difference between the result sets for both statements, as explained above. The LEFT SEMI JOIN and the LEFT ANTI JOIN are not equivalent operations and will produce different outputs.

D) Both statements will not fail, as Databricks SQL does support those join types. Databricks SQL supports various join types, including INNER, LEFT OUTER, RIGHT OUTER, FULL OUTER, LEFT SEMI, LEFT ANTI, and CROSS. You can also use NATURAL, USING, or LATERAL keywords to specify different join criteria.

E) The first statement will not return only the customer\_id from the orders table, as it will return all columns from the customers table. The second statement is correct, but it is not the only difference between the result sets.

## NEW QUESTION # 57

Which of the following should data analysts consider when working with personally identifiable information (PII) data?

- A. None of these considerations
- B. Legal requirements for the area in which the data was collected
- C. Organization-specific best practices for PII data
- D. All of these considerations
- E. Legal requirements for the area in which the analysis is being performed

### Answer: D

Explanation:

Data analysts should consider all of these factors when working with PII data, as they may affect the data security, privacy, compliance, and quality. PII data is any information that can be used to identify a specific individual, such as name, address, phone number, email, social security number, etc. PII data may be subject to different legal and ethical obligations depending on the context and location of the data collection and analysis. For example, some countries or regions may have stricter data protection laws than others, such as the General Data Protection Regulation (GDPR) in the European Union. Data analysts should also follow the organization-specific best practices for PII data, such as encryption, anonymization, masking, access control, auditing, etc. These best practices can help prevent data breaches, unauthorized access, misuse, or loss of PII data. Reference:

How to Use Databricks to Encrypt and Protect PII Data

Automating Sensitive Data (PII/PHI) Detection

Databricks Certified Data Analyst Associate

## NEW QUESTION # 58

What describes Partner Connect in Databricks?

- A. it allows multi-directional connection between Databricks and Databricks partners easier.
- B. It is a feature that runs Databricks partner tools on a Databricks SQL Warehouse (formerly known as a SQL endpoint).
- C. it allows for free use of Databricks partner tools through a common API.
- D. It exposes connection information to third-party tools via Databricks partners.

### Answer: A

Explanation:

Databricks Partner Connect is designed to simplify and streamline the integration between Databricks and its technology partners. It

provides a unified interface within the Databricks platform that facilitates the discovery and connection to a variety of data, analytics, and AI tools. By automating the configuration of necessary resources such as clusters, tokens, and connection files, Partner Connect enables seamless, bi-directional data flow between Databricks and partner solutions. This integration enhances the overall functionality of the Databricks Lakehouse by allowing users to easily incorporate external tools and services into their workflows, thereby expanding the platform's capabilities and fostering a more cohesive data ecosystem.[https://www.databricks.com/blog/2021/11/18/how-generally-available-introducing-databricks-partner-connect-to-discover-and-connect-popular-data-and-ai-tools-to-the-lakehouse?utm\\_source=chatgpt.com](https://www.databricks.com/blog/2021/11/18/how-generally-available-introducing-databricks-partner-connect-to-discover-and-connect-popular-data-and-ai-tools-to-the-lakehouse?utm_source=chatgpt.com)

### NEW QUESTION # 59

A data analyst needs to share a Databricks SQL dashboard with stakeholders that are not permitted to have accounts in the Databricks deployment. The stakeholders need to be notified every time the dashboard is refreshed.

Which approach can the data analyst use to accomplish this task with minimal effort?

- A. By granting the stakeholders' email addresses permissions to the dashboard
- B. By granting the stakeholders' email addresses to the SQL Warehouse (formerly known as endpoint) subscribers list
- C. **By adding the stakeholders' email addresses to the refresh schedule subscribers list**
- D. By downloading the dashboard as a PDF and emailing it to the stakeholders each time it is refreshed

#### Answer: C

Explanation:

To share a Databricks SQL dashboard with stakeholders who do not have accounts in the Databricks deployment and ensure they are notified upon each refresh, the data analyst can add the stakeholders' email addresses to the dashboard's refresh schedule subscribers list. This approach allows the stakeholders to receive email notifications containing the latest dashboard updates without requiring them to have direct access to the Databricks workspace. This method is efficient and minimizes effort, as it automates the notification process and ensures stakeholders remain informed of the most recent data insights.

### NEW QUESTION # 60

In which of the following situations should a data analyst use higher-order functions?

- A. **When custom logic needs to be applied at scale to array data objects**
- B. When built-in functions are taking too long to perform tasks
- C. When built-in functions need to run through the Catalyst Optimizer
- D. When custom logic needs to be applied to simple, unnested data
- E. When custom logic needs to be converted to Python-native code

#### Answer: A

Explanation:

Higher-order functions are a simple extension to SQL to manipulate nested data such as arrays. A higher-order function takes an array, implements how the array is processed, and what the result of the computation will be. It delegates to a lambda function how to process each item in the array. This allows you to define functions that manipulate arrays in SQL, without having to unpack and repack them, use UDFs, or rely on limited built-in functions. Higher-order functions provide a performance benefit over user defined functions. Reference: Higher-order functions | Databricks on AWS, Working with Nested Data Using Higher Order Functions in SQL on Databricks | Databricks Blog, Higher-order functions - Azure Databricks | Microsoft Learn, Optimization recommendations on Databricks | Databricks on AWS

### NEW QUESTION # 61

.....

As a professional website, Exams-boost offers you the latest and valid Databricks-Certified-Data-Analyst-Associate test questions and latest learning materials, which are composed by our experienced IT elites and trainers. They have rich experience in the Databricks actual test and are good at making learning strategy for people who want to pass the Databricks-Certified-Data-Analyst-Associate Practice Exam

**Actual Databricks-Certified-Data-Analyst-Associate Tests:** <https://www.exams-boost.com/Databricks-Certified-Data-Analyst-Associate-valid-materials.html>

P.S. Free 2026 Databricks Databricks-Certified-Data-Analyst-Associate dumps are available on Google Drive shared by Exam-boost: <https://drive.google.com/open?id=1fWrzAnWwQ1d2X8qP0vI70AAIn6FCTV6p>