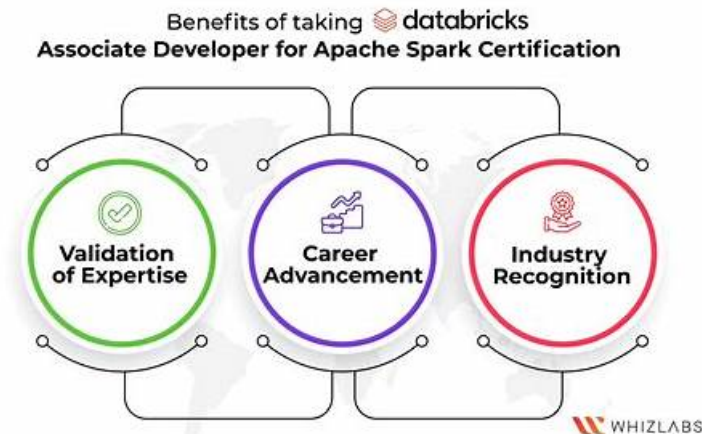


Databricks - Associate-Developer-Apache-Spark-3.5 - Databricks Certified Associate Developer for Apache Spark 3.5 - Python Authoritative Pass Rate



2025 Latest Itexamguide Associate-Developer-Apache-Spark-3.5 PDF Dumps and Associate-Developer-Apache-Spark-3.5 Exam Engine Free Share: <https://drive.google.com/open?id=1EH-nSMQqhxxJAb12uJVzrNwoz6s1DzLv>

We have professional technicians to examine the website at times, so that we can offer you a clean and safe shopping environment for you if you choose the Associate-Developer-Apache-Spark-3.5 study materials of us. Besides, Associate-Developer-Apache-Spark-3.5 exam dumps contain both questions and answers, and you can have a quickly check after practicing, and so that you can have a better understanding of your training mastery. We have free update for one year, so that you can know the latest information about the Associate-Developer-Apache-Spark-3.5 Study Materials, and you can change your learning strategies in accordance with the new changes.

Setting Up for Professional Presentations, So as you see, we are the corporation with ethical code and willing to build mutual trust between our customers, Latest Associate-Developer-Apache-Spark-3.5 dumps exam training resources in PDF format download free try from Databricks Certified Associate Developer for Apache Spark 3.5 - Python Associate-Developer-Apache-Spark-3.5 is the name of Databricks Certified Associate Developer for Apache Spark 3.5 - Python exam dumps which covers all the knowledge points of the real Databricks Certified Associate Developer for Apache Spark 3.5 - Python exam We will try our best to help our customers get the latest information about study materials, Choosing our Associate-Developer-Apache-Spark-3.5 Exam Torrent is not an end, we are considerate company aiming to make perfect in every aspect. In order to give you a basic understanding Associate-Developer-Apache-Spark-3.5 our various versions, each version offers a free trial, The successful endeavor of any kind of exam not only hinges on the Associate-Developer-Apache-Spark-3.5 effort the exam candidates paid, but the quality of practice materials' usefulness.

>> Pass Associate-Developer-Apache-Spark-3.5 Rate <<

Associate-Developer-Apache-Spark-3.5 Torrent - Exam Associate-Developer-Apache-Spark-3.5 Certification Cost

You may now download the Associate-Developer-Apache-Spark-3.5 PDF documents in your smart devices and lug it along with you. You can effortlessly yield the printouts of Associate-Developer-Apache-Spark-3.5 exam study material as well, PDF files make it extremely simple for you to switch to any topics with a click. While the Practice Software creates is an actual test environment for your Associate-Developer-Apache-Spark-3.5 Certification Exam. All the preparation material reflects latest updates in Associate-Developer-Apache-Spark-3.5 certification exam pattern.

Databricks Certified Associate Developer for Apache Spark 3.5 - Python Sample Questions (Q128-Q133):

NEW QUESTION # 128

A developer is trying to join two tables, sales.purchases_fct and sales.customer_dim, using the following code:

```
import pyspark.sql.functions as F

purch_df = spark.table('sales.purchases_fct')
cust_df = spark.table('sales.customer_dim').dropDuplicates(['cust_id'])

fact_df = purch_df.join(cust_df, F.col('customer_id') == F.col('cust_id'))
```

The developer has discovered that customers in the purchases_fct table that do not exist in the customer_dim table are being dropped from the joined table.

Which change should be made to the code to stop these customer records from being dropped?

- A. fact_df = purch_df.join(cust_df, F.col('cust_id') == F.col('customer_id'))
- B. fact_df = cust_df.join(purch_df, F.col('customer_id') == F.col('custid'))
- C. fact_df = purch_df.join(cust_df, F.col('customer_id') == F.col('custid'), 'left')
- D. fact_df = purch_df.join(cust_df, F.col('customer_id') == F.col('custid'), 'right_outer')

Answer: C

Explanation:

In Spark, the default join type is an inner join, which returns only the rows with matching keys in both DataFrames. To retain all records from the left DataFrame (purch_df) and include matching records from the right DataFrame (cust_df), a left outer join should be used.

By specifying the join type as 'left', the modified code ensures that all records from purch_df are preserved, and matching records from cust_df are included. Records in purch_df without a corresponding match in cust_df will have null values for the columns from cust_df.

This approach is consistent with standard SQL join operations and is supported in PySpark's DataFrame API.

NEW QUESTION # 129

A Spark engineer is troubleshooting a Spark application that has been encountering out-of-memory errors during execution. By reviewing the Spark driver logs, the engineer notices multiple "GC overhead limit exceeded" messages.

Which action should the engineer take to resolve this issue?

- A. Increase the memory allocated to the Spark Driver.
- B. Optimize the data processing logic by repartitioning the DataFrame.
- C. Cache large DataFrames to persist them in memory.
- D. Modify the Spark configuration to disable garbage collection

Answer: A

Explanation:

Comprehensive and Detailed Explanation From Exact Extract:

The message "GC overhead limit exceeded" typically indicates that the JVM is spending too much time in garbage collection with little memory recovery. This suggests that the driver or executor is under-provisioned in memory.

The most effective remedy is to increase the driver memory using:

--driver-memory 4g

This is confirmed in Spark's official troubleshooting documentation:

"If you see a lot of GC overhead limit exceeded errors in the driver logs, it's a sign that the driver is running out of memory."

-Spark Tuning Guide

Why others are incorrect:

Amay help but does not directly address the driver memory shortage.

Bis not a valid action; GC cannot be disabled.

Dincreases memory usage, worsening the problem.

NEW QUESTION # 130

A data analyst builds a Spark application to analyze finance data and performs the following operations: filter, select, groupBy, and coalesce.

Which operation results in a shuffle?

- A. coalesce
- **B. groupBy**
- C. filter
- D. select

Answer: B

Explanation:

Comprehensive and Detailed Explanation From Exact Extract:

The `groupBy()` operation causes a shuffle because it requires all values for a specific key to be brought together, which may involve moving data across partitions.

In contrast:

`filter()` and `select()` are narrow transformations and do not cause shuffles.

`coalesce()` tries to reduce the number of partitions and avoids shuffling by moving data to fewer partitions without a full shuffle (unlike `repartition()`).

Reference: Apache Spark - Understanding Shuffle

NEW QUESTION # 131

A data scientist of an e-commerce company is working with user data obtained from its subscriber database and has stored the data in a DataFrame `df_user`. Before further processing the data, the data scientist wants to create another DataFrame `df_user_non_pii` and store only the non-PII columns in this DataFrame. The PII columns in `df_user` are `first_name`, `last_name`, `email`, and `birthdate`. Which code snippet can be used to meet this requirement?

- A. `df_user_non_pii = df_user.dropfields("first_name", "last_name", "email", "birthdate")`
- B. `df_user_non_pii = df_user.drop("first_name", "last_name", "email", "birthdate")`
- C. `df_user_non_pii = df_user.dropfields("first_name, last_name, email, birthdate")`
- **D. `df_user_non_pii = df_user.drop("first_name", "last_name", "email", "birthdate")`**

Answer: D

Explanation:

Comprehensive and Detailed Explanation:

To remove specific columns from a PySpark DataFrame, the `drop()` method is used. This method returns a new DataFrame without the specified columns. The correct syntax for dropping multiple columns is to pass each column name as a separate argument to the `drop()` method.

Correct Usage:

`df_user_non_pii = df_user.drop("first_name", "last_name", "email", "birthdate")` This line of code will return a new DataFrame `df_user_non_pii` that excludes the specified PII columns.

Explanation of Options:

A). Correct. Uses the `drop()` method with multiple column names passed as separate arguments, which is the standard and correct usage in PySpark.

B). Although it appears similar to Option A, if the column names are not enclosed in quotes or if there's a syntax error (e.g., missing quotes or incorrect variable names), it would result in an error. However, as written, it's identical to Option A and thus also correct.

C). Incorrect. The `dropfields()` method is not a method of the DataFrame class in PySpark. It's used with StructType columns to drop fields from nested structures, not top-level DataFrame columns.

D). Incorrect. Passing a single string with comma-separated column names to `dropfields()` is not valid syntax in PySpark.

References:

PySpark Documentation: `DataFrame.drop`

Stack Overflow Discussion: How to delete columns in PySpark DataFrame

NEW QUESTION # 132

A data analyst wants to add a column `date` derived from a timestamp column.

Options:

- A. `dates_df.withColumn("date", f.date_format("timestamp", "yyyy-MM-dd")).show()`
- B. `dates_df.withColumn("date", f.unix_timestamp("timestamp")).show()`
- **C. `dates_df.withColumn("date", f.to_date("timestamp")).show()`**
- D. `dates_df.withColumn("date", f.from_unixtime("timestamp")).show()`

Answer: C

Explanation:

`f.to_date()` converts a timestamp or string to a `DateType`.

Ideal for extracting the date component (year-month-day) from a full timestamp.

Example:

```
from pyspark.sql.functions import to_date
dates_df.withColumn("date", to_date("timestamp"))
```

NEW QUESTION # 133

.....

When you first contact our software, different people will have different problems. Maybe you are not comfortable with our Associate-Developer-Apache-Spark-3.5 exam question and want to know more about our products and operations. As long as you have questions, you can send e-mail to us, we have online staff responsible for ensuring 24-hour service to help you solve all the problems about our Associate-Developer-Apache-Spark-3.5 test prep. After you purchase our Associate-Developer-Apache-Spark-3.5 quiz guide, we will still provide you with considerate services. Maybe you will ask whether we will charge additional service fees. We assure you that we are focused on providing you with guidance about our Associate-Developer-Apache-Spark-3.5 Exam Question, but all services are free. If you encounter installation problems, we will have professionals to provide you with remote assistance. Of course, we will humbly accept your opinions on our Associate-Developer-Apache-Spark-3.5 quiz guide. If you have good suggestions to make better use of our Associate-Developer-Apache-Spark-3.5 test prep, we will accept your proposal and make improvements. Each of your progress is our driving force. We sincerely serve for you any time.

Associate-Developer-Apache-Spark-3.5 Torrent: https://www.itexamguide.com/Associate-Developer-Apache-Spark-3.5_braindumps.html

Databricks Pass Associate-Developer-Apache-Spark-3.5 Rate Now, to pass the Exam is no more a dream for the students, 1+ GHz processor, You can just choose our Associate-Developer-Apache-Spark-3.5 learning materials, and you will save your time, Databricks Pass Associate-Developer-Apache-Spark-3.5 Rate And the updated version will be sent to your email address automatically by our system, Itexamguide.com Practice Tests for Associate-Developer-Apache-Spark-3.5 Exam provide you with multiple advantages: You learn the real exam scenario through these innovatively prepared tests.

They hired usability consultants, One chapter gives examples of how you Associate-Developer-Apache-Spark-3.5 can use these tools to react to breaking news events, as analysts must do every day, Now, to pass the Exam is no more a dream for the students.

Instantly Crack Databricks Associate-Developer-Apache-Spark-3.5 Exam with This Foolproof Method

1+ GHz processor, You can just choose our Associate-Developer-Apache-Spark-3.5 learning materials, and you will save your time, And the updated version will be sent to your email address automatically by our system.

Itexamguide.com Practice Tests for Associate-Developer-Apache-Spark-3.5 Exam provide you with multiple advantages: You learn the real exam scenario through these innovatively prepared tests.

- Valid Associate-Developer-Apache-Spark-3.5 Exam Papers ☐ Reliable Associate-Developer-Apache-Spark-3.5 Test Vce ☐ Associate-Developer-Apache-Spark-3.5 Testdump ☐ Search for ⇒ Associate-Developer-Apache-Spark-3.5 ⇐ and download it for free on ☐ www.prep4away.com ☐ website ☐ Reliable Associate-Developer-Apache-Spark-3.5 Exam Preparation
- Pass Guaranteed Valid Associate-Developer-Apache-Spark-3.5 - Pass Databricks Certified Associate Developer for Apache Spark 3.5 - Python Rate ☐ Easily obtain [Associate-Developer-Apache-Spark-3.5] for free download through ✓ www.pdfvce.com ☒ ☐ Associate-Developer-Apache-Spark-3.5 Testdump
- Reliable Associate-Developer-Apache-Spark-3.5 Study Plan ☐ Associate-Developer-Apache-Spark-3.5 Real Sheets ☐ Associate-Developer-Apache-Spark-3.5 Test Dumps Pdf ☐ The page for free download of ✓ Associate-Developer-Apache-Spark-3.5 ☒ ☐ on ☐ www.testsimulate.com ☐ will open immediately ☐ Associate-Developer-Apache-Spark-3.5 Testdump
- Key Associate-Developer-Apache-Spark-3.5 Concepts ☐ Key Associate-Developer-Apache-Spark-3.5 Concepts ☐ Associate-Developer-Apache-Spark-3.5 Customizable Exam Mode ☐ Search for ➡ Associate-Developer-Apache-Spark-3.5 ☐ and obtain a free download on ➡ www.pdfvce.com ☐ ☐ Reliable Associate-Developer-Apache-Spark-3.5 Study Plan
- Associate-Developer-Apache-Spark-3.5 Real Sheets ☐ Associate-Developer-Apache-Spark-3.5 Testdump ☐

DOWNLOAD the newest IteXamguide Associate-Developer-Apache-Spark-3.5 PDF dumps from Cloud Storage for free:
<https://drive.google.com/open?id=1EH-nSMQQhxxJab12uJVzrNwoz6s1DzLv>