



Um die Databricks Associate-Developer-Apache-Spark-3.5 Zertifizierungsprüfung zu bestehen, brauchen Sie viel Zeit und Energie. Dabei müssen Sie auch großes Risiko tragen. Wenn Sie ExamFragen wählen, können Sie mit wenigem Geld die Databricks Associate-Developer-Apache-Spark-3.5 Prüfung einmal bestehen. Ich meine, dass ExamFragen heutzutage die beste Wahl für Sie ist, wo die Zeit sehr geschätzt wird. Außerdem ist ExamFragen eine der vielen Websites, die Ihnen einen bestmöglichen Garant bietet. Wenn Sie ExamFragen wählen, kommt der Erfolg auf Sie zu.

per-Apache-Spark-3.5 Deutsch & Associate

Um Ihre Databricks Associate-Developer-Apache-Spark-3.5 Zertifizierungsprüfungen reibungslos erfolgreich zu meistern, brauchen Sie nur unsere Prüfungsfragen und Antworten zu Databricks Associate-Developer-Apache-Spark-3.5 Dumps (Databricks Certified Associate Developer for Apache Spark 3.5 - Python) auswendigzulernen. Viel Erfolg!

Databricks Certified Associate Developer for Apache Spark 3.5 - Python Associate-Developer-Apache-Spark-3.5 Prüfungsfragen mit Lösungen (Q56-Q61):

56. Frage

A data engineer is working on the DataFrame:

Id	Name	count	timestamp
4	Washington	10	2024-09-19T10:10:40.000+00:00
1	Delhi	20	2024-09-19T10:10:10.000+00:00
2	London	50	2024-09-19T10:10:20.000+00:00
1	Delhi	50	2024-09-19T10:10:50.000+00:00
3	Paris	20	2024-09-19T10:11:20.000+00:00
1	Delhi	10	2024-09-19T10:11:10.000+00:00
3	Paris	30	2024-09-19T10:10:30.000+00:00
4	Washington	40	2024-09-19T10:11:00.000+00:00

(Referring to the table image: it has columns Id, Name, count, and timestamp.) Which code fragment should the engineer use to extract the unique values in the Name column into an alphabetically ordered list?

- A. `df.select("Name").orderBy(df["Name"].asc())`
- B. `df.select("Name").distinct().orderBy(df["Name"].desc())`
- C. `df.select("Name").distinct()`
- D. `df.select("Name").distinct().orderBy(df["Name"])`

Antwort: D

Begründung:

To extract unique values from a column and sort them alphabetically:

`distinct()` is required to remove duplicate values.

`orderBy()` is needed to sort the results alphabetically (ascending by default).

Correct code:

`df.select("Name").distinct().orderBy(df["Name"])`

This is directly aligned with standard DataFrame API usage in PySpark, as documented in the official Databricks Spark APIs.

Option A is incorrect because it may not remove duplicates. Option C omits sorting. Option D sorts in descending order, which doesn't meet the requirement for alphabetical (ascending) order.

57. Frage

34 of 55.

A data engineer is investigating a Spark cluster that is experiencing underutilization during scheduled batch jobs.

After checking the Spark logs, they noticed that tasks are often getting killed due to timeout errors, and there are several warnings about insufficient resources in the logs.

Which action should the engineer take to resolve the underutilization issue?

- A. Increase the number of executor instances to handle more concurrent tasks.
- B. Increase the executor memory allocation in the Spark configuration.
- C. Set the `spark.network.timeout` property to allow tasks more time to complete without being killed.
- D. Reduce the size of the data partitions to improve task scheduling.

Antwort: A

Begründung:

Underutilization with timeout warnings often indicates insufficient parallelism - meaning there aren't enough executors to process all tasks concurrently.

Solution:

Increase the number of executors to allow more parallel task execution and better resource utilization.

Example configuration:

--conf spark.executor.instances=8

This distributes the workload more effectively across cluster nodes and reduces idle time for pending tasks.

Why the other options are incorrect:

A: Extending timeouts hides the symptom, not the root cause (lack of executors).

B: More memory per executor won't fix scheduling bottlenecks.

C: Reducing partition size may increase overhead and does not fix resource imbalance.

Reference:

Databricks Exam Guide (June 2025): Section "Troubleshooting and Tuning Apache Spark DataFrame API Applications" - tuning executors and cluster utilization.

Spark Configuration - executor instances and resource scaling.

58. Frage

How can a Spark developer ensure optimal resource utilization when running Spark jobs in Local Mode for testing?

Options:

- A. Use the spark.dynamicAllocation.enabled property to scale resources dynamically.
- B. Configure the application to run in cluster mode instead of local mode.
- C. Set the spark.executor.memory property to a large value.
- **D. Increase the number of local threads based on the number of CPU cores.**

Antwort: D

Begründung:

When running in local mode (e.g., local[4]), the number inside the brackets defines how many threads Spark will use.

Using local[*] ensures Spark uses all available CPU cores for parallelism.

Example:

spark-submit --master local[*]

Dynamic allocation and executor memory apply to cluster-based deployments, not local mode.

59. Frage

A data engineer wants to process a streaming DataFrame that receives sensor readings every second with columns sensor_id, temperature, and timestamp. The engineer needs to calculate the average temperature for each sensor over the last 5 minutes while the data is streaming.

Which code implementation achieves the requirement?

Options from the images provided:

• A.

```
df.groupBy("sensor_id").avg("temperature")  
df.withWatermark("timestamp", "5 minutes")
```

• **B.**

```
.groupBy("sensor_id", window("timestamp", "5 minutes"))  
.agg(avg("temperature").alias("avg_temp"))
```

• C.

```
df.groupBy("sensor_id", "timestamp")  
.agg(avg("temperature").alias("avg_temp"))
```

• D.

```
df.withColumn("avg_temp", avg("temperature")  
.over(Window.partitionBy("sensor_id"))
```

Antwort: B

Begründung:

The correct answer is D because it uses proper time-based window aggregation along with watermarking, which is the required pattern in Spark Structured Streaming for time-based aggregations over event-time data.

From the Spark 3.5 documentation on structured streaming:

"You can define sliding windows on event-time columns, and use `groupBy` along with `window()` to compute aggregates over those windows. To deal with late data, you use `withWatermark()` to specify how late data is allowed to arrive." (Source: Structured Streaming Programming Guide) In option D, the use of:

python

CopyEdit

```
.groupBy("sensor_id", window("timestamp", "5 minutes"))
```

```
.agg(avg("temperature").alias("avg_temp"))
```

ensures that for each `sensor_id`, the average temperature is calculated over 5-minute event-time windows. To complete the logic, it is assumed that `withWatermark("timestamp", "5 minutes")` is used earlier in the pipeline to handle late events.

Explanation of why other options are incorrect:

Option A uses `Window.partitionBy` which applies to static DataFrames or batch queries and is not suitable for streaming aggregations.

Option B does not apply a time window, thus does not compute the rolling average over 5 minutes.

Option C incorrectly applies `withWatermark()` after an aggregation and does not include any time window, thus missing the time-based grouping required.

Therefore, Option D is the only one that meets all requirements for computing a time-windowed streaming aggregation.

60. Frage

A data engineer is running a Spark job to process a dataset of 1 TB stored in distributed storage. The cluster has 10 nodes, each with 16 CPUs. Spark UI shows:

Low number of Active Tasks

Many tasks complete in milliseconds

Fewer tasks than available CPUs

Which approach should be used to adjust the partitioning for optimal resource allocation?

- A. Set the number of partitions equal to the number of nodes in the cluster
- B. Set the number of partitions equal to the total number of CPUs in the cluster
- **C. Set the number of partitions by dividing the dataset size (1 TB) by a reasonable partition size, such as 128 MB**
- D. Set the number of partitions to a fixed value, such as 200

Antwort: C

Begründung:

Comprehensive and Detailed Explanation From Exact Extract:

Spark's best practice is to estimate partition count based on data volume and a reasonable partition size - typically 128 MB to 256 MB per partition.

With 1 TB of data: $1 \text{ TB} / 128 \text{ MB} \approx 8000$ partitions

This ensures that tasks are distributed across available CPUs for parallelism and that each task processes an optimal volume of data.

Option A (equal to cores) may result in partitions that are too large.

Option B (fixed 200) is arbitrary and may underutilize the cluster.

Option C (nodes) gives too few partitions (10), limiting parallelism.

Reference: Databricks Spark Tuning Guide # Partitioning Strategy

61. Frage

.....

Viele IT-Leute sind sich einig, dass Databricks Associate-Developer-Apache-Spark-3.5 Zertifikat ein Sprungbrett zu dem Höhepunkt der IT-Branche ist. Deshalb kümmern sich viele IT-Experten um die Databricks Associate-Developer-Apache-Spark-3.5 Zertifizierungsprüfung.

Associate-Developer-Apache-Spark-3.5 Deutsch: <https://www.examfragen.de/Associate-Developer-Apache-Spark-3.5-pruefung-fragen.html>

Und der Aktualisierungsdienst der Databricks Associate-Developer-Apache-Spark-3.5 Deutsch Associate-Developer-Apache-

Spark-3.5 Deutsch - Databricks Certified Associate Developer for Apache Spark 3.5 - Python ist innerhalb einem Jahr nach Ihrem Kauf ganz gratis, Databricks Associate-Developer-Apache-Spark-3.5 Prüfungsübungen Wenn Sie wie ich einen IT-Traum haben, Schicken bitte schnell die Produkte von ExamFragen Associate-Developer-Apache-Spark-3.5 Deutsch in den Warenkorb, Databricks Associate-Developer-Apache-Spark-3.5 Prüfungsübungen Das macht doch nichta.

Was ist die Welt, Dann nahm sie die Kohle und hielt sie plötzlich dicht Associate-Developer-Apache-Spark-3.5 vor Toms Gesicht, Und der Aktualisierungsdienst der Databricks Databricks Certified Associate Developer for Apache Spark 3.5 - Python ist innerhalb einem Jahr nach Ihrem Kauf ganz gratis.

Associate-Developer-Apache-Spark-3.5 Prüfungsguide: Databricks Certified Associate Developer for Apache Spark 3.5 - Python & Associate-Developer-Apache-Spark-3.5 echter Test & Associate-Developer-Apache-Spark-3.5 sicherlich-zu-bestehen

Wenn Sie wie ich einen IT-Traum haben, Schicken bitte schnell die Produkte von ExamFragen in den Warenkorb, Das macht doch nichta, Beim Associate-Developer-Apache-Spark-3.5 Lernen makiert das System automatisch die falsch geschriebenen Associate-Developer-Apache-Spark-3.5 Fragenkatalog Übungen, dann erscheinen soche Übungen mehr als die anderen, bis dass Sie richtige Antworten kreuzen.

- Die neuesten Associate-Developer-Apache-Spark-3.5 echte Prüfungsfragen, Databricks Associate-Developer-Apache-Spark-3.5 originale fragen ☐ Geben Sie ☐ www.zertfragen.com ☐ ein und suchen Sie nach kostenloser Download von ✓ Associate-Developer-Apache-Spark-3.5 ☐ ✓ ☐ Associate-Developer-Apache-Spark-3.5 Trainingsunterlagen
- Associate-Developer-Apache-Spark-3.5 Zertifizierung ☐ Associate-Developer-Apache-Spark-3.5 Fragenpool ☐ Associate-Developer-Apache-Spark-3.5 Prüfungs ☐ URL kopieren ☐ www.itzert.com ☐ Öffnen und suchen Sie ➡ Associate-Developer-Apache-Spark-3.5 ☐ Kostenloser Download ☐ Associate-Developer-Apache-Spark-3.5 PDF Testsoftware
- Die neuesten Associate-Developer-Apache-Spark-3.5 echte Prüfungsfragen, Databricks Associate-Developer-Apache-Spark-3.5 originale fragen ☐ Suchen Sie auf ➡ www.zertsoft.com ☐ nach kostenlosem Download von ➡ Associate-Developer-Apache-Spark-3.5 ☐ ☐ Associate-Developer-Apache-Spark-3.5 Dumps
- Associate-Developer-Apache-Spark-3.5 Prüfungen ☐ Associate-Developer-Apache-Spark-3.5 Zertifizierungsantworten ☐ ☐ Associate-Developer-Apache-Spark-3.5 Zertifizierungsantworten ☐ ➡ www.itzert.com ☐ ist die beste Webseite um den kostenlosen Download von ☐ Associate-Developer-Apache-Spark-3.5 ☐ zu erhalten ☐ Associate-Developer-Apache-Spark-3.5 Ausbildungsressourcen
- Associate-Developer-Apache-Spark-3.5 Trainingsmaterialien: Databricks Certified Associate Developer for Apache Spark 3.5 - Python - Associate-Developer-Apache-Spark-3.5 Lernmittel - Databricks Associate-Developer-Apache-Spark-3.5 Quiz ☐ Suchen Sie einfach auf ➡ www.zertfragen.com ☐ nach kostenloser Download von ☀ Associate-Developer-Apache-Spark-3.5 ☐ ☀ ☐ Associate-Developer-Apache-Spark-3.5 Antworten
- Associate-Developer-Apache-Spark-3.5 Zertifizierung ☐ Associate-Developer-Apache-Spark-3.5 PDF Testsoftware ☐ ☐ Associate-Developer-Apache-Spark-3.5 Echte Fragen ☐ Öffnen Sie die Website ➡ www.itzert.com ☐ ☐ Suchen Sie ☀ Associate-Developer-Apache-Spark-3.5 ☐ ☀ ☐ Kostenloser Download ☀ Associate-Developer-Apache-Spark-3.5 Deutsch Prüfungsfragen
- Associate-Developer-Apache-Spark-3.5 Quizfragen Und Antworten ☐ Associate-Developer-Apache-Spark-3.5 Deutsch Prüfungsfragen ☐ Associate-Developer-Apache-Spark-3.5 Prüfungsfrage ☐ Sie müssen nur zu ☐ www.it-pruefung.com ☐ gehen um nach kostenloser Download von ▶ Associate-Developer-Apache-Spark-3.5 ◀ zu suchen ☐ Associate-Developer-Apache-Spark-3.5 Online Tests
- Associate-Developer-Apache-Spark-3.5 Schulungsmaterialien - Associate-Developer-Apache-Spark-3.5 Dumps Prüfung - Associate-Developer-Apache-Spark-3.5 Studienguide ☐ Öffnen Sie die Webseite ➡ www.itzert.com ☐ und suchen Sie nach kostenloser Download von ▷ Associate-Developer-Apache-Spark-3.5 ◁ ☐ Associate-Developer-Apache-Spark-3.5 Examengine
- Associate-Developer-Apache-Spark-3.5 Unterlage ☐ Associate-Developer-Apache-Spark-3.5 Fragenpool ☐ Associate-Developer-Apache-Spark-3.5 Deutsch Prüfungsfragen ☐ Öffnen Sie die Website ☀ www.deutschpruefung.com ☐ ☀ ☐ Suchen Sie (Associate-Developer-Apache-Spark-3.5) Kostenloser Download ☐ ☐ Associate-Developer-Apache-Spark-3.5 PDF Testsoftware
- Neuester und gültiger Associate-Developer-Apache-Spark-3.5 Test VCE Motoren-Dumps und Associate-Developer-Apache-Spark-3.5 neueste Testfragen für die IT-Prüfungen ☐ Öffnen Sie (www.itzert.com) geben Sie { Associate-Developer-Apache-Spark-3.5 } ein und erhalten Sie den kostenlosen Download ☐ Associate-Developer-Apache-Spark-3.5 Dumps
- Associate-Developer-Apache-Spark-3.5 Dumps ☐ Associate-Developer-Apache-Spark-3.5 Quizfragen Und Antworten ☐ Associate-Developer-Apache-Spark-3.5 Quizfragen Und Antworten ☐ Suchen Sie auf der Webseite ☐ www.examfragen.de ☐ nach ➡ Associate-Developer-Apache-Spark-3.5 ☐ und laden Sie es kostenlos herunter

