DATABRICKS-CERTIFIED-ASSOCIAT

Q&As

Databricks Certified Associate Developer for Apache Spark 3.0

Pass Databricks DATABRICKS-CERTIFIED-ASSOCIATE-DEVELOPER-FOR-APACHE-SPARK Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.geekcert.com/databricks-certified-associate-developer-for-apache-spark.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Databricks Official Exam Center

QUESTION 1

Which of the following code blocks can be used to save DataFrame transactionsDf to memory only, recalculating partitions that do not fit in memory when they are needed?

A. from pyspark import StorageLevel transactionsDf.cache(StorageLevel.MEMORY_ONLY)

B. transactionsDf.cache()

C. transactionsDf.storage_level(\\'MEMORY_ONLY\\')

D. transactionsDf.persist()

E. transactionsDf.clear_persist()

F. from pyspark import StorageLevel transactionsDf.persist(StorageLevel.MEMORY_ONLY)

Correct Answer: F

QUESTION 2

The code block displayed below contains an error. The code block should return the average of rows in

column value grouped by unique storeId. Find the error.

Code block:

transactionsDf.agg("storeId").avg("value")

A. Instead of avg("value"), avg(col("value")) should be used.

B. The avg("value") should be specified as a second argument to agg() instead of being appended to it.

C. All column names should be wrapped in col() operators.

D. agg should be replaced by groupBy.

E. "storeId" and "value" should be swapped.

Correct Answer: D

QUESTION 3

The code block displayed below contains an error. The code block should arrange the rows of DataFrame transactionsDf using information from two columns in an ordered fashion, arranging first by column value, showing smaller numbers at the top and greater numbers at the bottom, and then by column predError, for which all values should be arranged in the inverse way of the order of items in column value. Find the error.

Code block:

transactionsDf.orderBy(\\'value\\', asc_nulls_first(col(\\'predError\\')))

A. Two orderBy statements with calls to the individual columns should be chained, instead of having both columns in one orderBy statement.

B. Column value should be wrapped by the col() operator.

C. Column predError should be sorted in a descending way, putting nulls last.

D. Column predError should be sorted by desc_nulls_first() instead.

E. Instead of orderBy, sort should be used.

Correct Answer: C

QUESTION 4

Which of the following code blocks creates a new one-column, two-row DataFrame dfDates with column date of type timestamp?

A. 1.dfDates = spark.createDataFrame(["23/01/2022 11:28:12","24/01/2022 10:58:34"], ["date"]) 2.dfDates = dfDates.withColumn("date", to_timestamp("dd/MM/yyyy HH:mm:ss", "date"))

B. 1.dfDates = spark.createDataFrame([("23/01/2022 11:28:12",),("24/01/2022 10:58:34",)], ["date"]) 2.dfDates = dfDates.withColumnRenamed("date", to_timestamp("date", "yyyy-MM-ddHH:mm:ss"))

C. 1.dfDates = spark.createDataFrame([("23/01/2022 11:28:12",),("24/01/2022 10:58:34",)], ["date"]) 2.dfDates = dfDates.withColumn("date", to_timestamp("date", "dd/MM/yyyy HH:mm:ss"))

D. 1.dfDates = spark.createDataFrame(["23/01/2022 11:28:12","24/01/2022 10:58:34"], ["date"]) 2.dfDates = dfDates.withColumnRenamed("date", to_datetime("date", "yyyy-MM-ddHH:mm:ss"))

E. 1.dfDates = spark.createDataFrame([("23/01/2022 11:28:12",),("24/01/2022 10:58:34",)], ["date"])

Correct Answer: C

QUESTION 5

Which of the following code blocks returns a DataFrame with a single column in which all items in column

attributes of DataFrame itemsDf are listed that contain the letter i?

Sample of DataFrame itemsDf:

1.+------+-------------------------------+--------------------------+-----------------+

2.|itemId|itemName |attributes |supplier |

DATABRICKS-CERTIFIED-ASSOCIATE-DEVELOPER-FOR-APACHE-SPARK PDF Dumps | DATABRICKS-
CERTIFIED-ASSOCIATE-DEVELOPER-FOR-APACHE-SPARK Study Guide | DATABRICKS-CERTIFIED-
ASSOCIATE-DEVELOPER-FOR-APACHE-SPARK Exam Questions

4 / 5

3.+------+-------------------------------+--------------------------+------------------+

4.|1 |Thick Coat for Walking in the Snow|[blue, winter, cozy] |Sports Company Inc.|

5.|2 |Elegant Outdoors Summer Dress |[red, summer, fresh, cooling]|YetiX |

6.|3 |Outdoors Backpack |[green, summer, travel] |Sports Company Inc.|

7.+------+-------------------------------+--------------------------+------------------+

A. itemsDf.select(explode("attributes").alias("attributes_exploded")).filter(attributes_exploded.c ontains ("i"))

B. itemsDf.explode(attributes).alias("attributes_exploded").filter(col("attributes_exploded").con tains("i"))

C. itemsDf.select(explode("attributes")).filter("attributes_exploded".contains("i"))

D. itemsDf.select(explode("attributes").alias("attributes_exploded")).filter(col("attributes_explo ded").contains("i"))

E. itemsDf.select(col("attributes").explode().alias("attributes_exploded")).filter(col("attributes_e xploded").contains("i"))

Correct Answer: D

DATABRICKS-CERTIFIED-ASSOCIATE-DEVELOPER-FOR-APACHE-SPARK PDF Dumps | DATABRICKS-
CERTIFIED-ASSOCIATE-DEVELOPER-FOR-APACHE-SPARK Study Guide | DATABRICKS-CERTIFIED-
ASSOCIATE-DEVELOPER-FOR-APACHE-SPARK Exam Questions

5 / 5