



# DATABRICKS-CERTIFIED-ASSOCIATE-DEVELOPER-FOR-APACHE-SPARK

## 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



VCE & PDF

GeekCert.com

<https://www.geekcert.com/databricks-certified-associate-developer-for-apac>  
2024 Latest geekcert DATABRICKS-CERTIFIED-ASSOCIATE-DEVELOPER-  
FOR-APACHE-SPARK PDF and VCE dumps Download

- ⚙️ **Instant Download** After Purchase
- ⚙️ **100% Money Back** Guarantee
- ⚙️ **365 Days** Free Update
- ⚙️ **800,000+** Satisfied Customers





### QUESTION 1

Which of the following code blocks reads JSON file imports.json into a DataFrame?

- A. `spark.read().mode("json").path("/FileStore/imports.json")`
- B. `spark.read.format("json").path("/FileStore/imports.json")`
- C. `spark.read("json", "/FileStore/imports.json")`
- D. `spark.read.json("/FileStore/imports.json")`
- E. `spark.read().json("/FileStore/imports.json")`

Correct Answer: D

Static notebook | Dynamic notebook: See test 1, 25 (Databricks import instructions) ([https://flrs.github.io/spark\\_practice\\_tests\\_code/#1/25.html](https://flrs.github.io/spark_practice_tests_code/#1/25.html) , [https://bit.ly/sparkpracticeexams\\_import\\_instructions](https://bit.ly/sparkpracticeexams_import_instructions))

### QUESTION 2

Which of the following code blocks returns all unique values across all values in columns value and productId in DataFrame transactionsDf in a one-column DataFrame?

- A. `transactionsDf.select(\value\).join(transactionsDf.select(\productId\), col(\value\)==col(\productId\), \outer\)`
- B. `transactionsDf.select(col(\value\), col(\productId\)).agg({\*\: \count\})`
- C. `transactionsDf.select(\value\, \productId\).distinct()`
- D. `transactionsDf.select(\value\).union(transactionsDf.select(\productId\)).distinct()`
- E. `transactionsDf.agg({\value\: \collect_set\, \productId\: \collect_set\})`

Correct Answer: D

### QUESTION 3

Which of the following code blocks returns a DataFrame with an added column to DataFrame transactionsDf that shows the unix epoch timestamps in column transactionDate as strings in the format month/day/year in column transactionDateFormatted?

Excerpt of DataFrame transactionsDf:

- A. `transactionsDf.withColumn("transactionDateFormatted", from_unixtime("transactionDate", format="dd/ MM/yyyy"))`
- B. `transactionsDf.withColumnRenamed("transactionDate", "transactionDateFormatted", from_unixtime("transactionDateFormatted", format="MM/dd/yyyy"))`



- C. transactionsDf.apply(from\_unixtime(format="MM/dd/yyyy")).asColumn("transactionDateFormatted")
- D. transactionsDf.withColumn("transactionDateFormatted", from\_unixtime("transactionDate", format="MM/dd/yyyy"))
- E. transactionsDf.withColumn("transactionDateFormatted", from\_unixtime("transactionDate"))

Correct Answer: D

#### QUESTION 4

The code block shown below should return a DataFrame with columns transactionId, predError, value, and f from DataFrame transactionsDf. Choose the answer that correctly fills the blanks in the code block to accomplish this.

```
transactionsDf.__1__(__2__)
```

- A. 1. filter  
2. "transactionId", "predError", "value", "f"
- B. 1. select  
2. "transactionId, predError, value, f"
- C. 1. select  
2. ["transactionId", "predError", "value", "f"]
- D. 1. where  
2. col("transactionId"), col("predError"), col("value"), col("f")
- E. 1. select  
2. col(["transactionId", "predError", "value", "f"])

Correct Answer: C

#### QUESTION 5

The code block shown below should set the number of partitions that Spark uses when shuffling data for joins or aggregations to 100. Choose the answer that correctly fills the blanks in the code block to accomplish this.

```
spark.sql.shuffle.partitions
```

```
__1__.__2__.__3__(__4__, 100)
```

- A. 1. spark  
2.  
conf



3.

set

4.

"spark.sql.shuffle.partitions"

B. 1. pyspark

2.

config

3.

set

4.

spark.shuffle.partitions

C. 1. spark

2.

conf

3.

get

4.

"spark.sql.shuffle.partitions"

D. 1. pyspark

2.

config

3.

set

4.

"spark.sql.shuffle.partitions"

E. 1. spark

2.

conf

3.



set

4.

"spark.sql.aggregate.partitions"

Correct Answer: A

Correct code block:

spark.conf.set("spark.sql.shuffle.partitions", 100) The conf interface is part of the SparkSession, so you need to call it through spark and not pyspark. To configure spark, you need to use the set method, not the get method. get reads a property, but does not write it. The correct property to achieve what is outlined in the

[Latest DATABRICKS-CERTIFIED-ASSOCIATE-DEVELOPER-FOR-APACHE-SPARK Dumps](#)

[DATABRICKS-CERTIFIED-ASSOCIATE-DEVELOPER-FOR-APACHE-SPARK VCE Dumps](#)

[DATABRICKS-CERTIFIED-ASSOCIATE-DEVELOPER-FOR-APACHE-SPARK Braindumps](#)