Databricks Developer for Apache Spark - Python

Databricks Apache Spark Developer Associate Certification Questions & Answers

Get Instant Access to Vital Exam Acing Materials | Study Guide | Sample Questions | Practice Test

---

DEVELOPER FOR APACHE SPARK - PYTHON

Databricks Certified Associate Developer for Apache Spark
60 Questions Exam – 70% Cut Score – Duration of 120 minutes

Table of Contents:

Discover More about the Databricks Developer for Apache Spark - Python Certification

Are you interested in passing the Databricks Developer for Apache Spark - Python exam? First discover, who benefits from the Developer for Apache Spark - Python certification. The Developer for Apache Spark - Python is suitable for a candidate if he wants to learn about Specialty. Passing the Developer for Apache Spark - Python exam earns you the Databricks Certified Associate Developer for Apache Spark title.

While preparing for the Developer for Apache Spark - Python exam, many candidates struggle to get the necessary materials. But do not worry; your struggling days are over. The Developer for Apache Spark - Python PDF contains some of the most valuable preparation tips and the details and instant access to useful **Developer for Apache Spark - Python study materials just at one click.**

Databricks Developer for Apache Spark - Python Apache Spark Developer Associate Certification Details:

Exam Name	Databricks Certified Associate Developer for Apache Spark
Exam Code	Developer for Apache Spark - Python
Exam Price	$200 (USD)
Duration	120 mins
Number of Questions	60
Passing Score	70%
Schedule Exam	**Kryterion Webassesor**
Sample Questions	**Databricks Developer for Apache Spark - Python Sample Questions**
Practice Exam	**Databricks Developer for Apache Spark - Python Certification Practice Exam**

Databricks Developer for Apache Spark - Python Syllabus:

Topic	Weights
Apache Spark Architecture Concepts	17%
Apache Spark Architecture Applications	11%
Apache Spark DataFrame API Applications	72%

Broaden Your Knowledge with Databricks Developer for Apache Spark - Python Sample Questions:

Question: 1

Which of the following code blocks adds a column predErrorSqrt to DataFrame transactionsDf that is the square root of column predError?

- a) transactionsDf.withColumn("predErrorSqrt", sqrt(col("predError")))
- b) transactionsDf.withColumn("predErrorSqrt", sqrt(predError))
- c) transactionsDf.select(sqrt(predError))
- d) transactionsDf.withColumn("predErrorSqrt", col("predError").sqrt())
- e) transactionsDf.select(sqrt("predError"))

Answer: a

Question: 2

Which of the following statements is NOT true for broadcast variables?

- a) It provides a mutable variable that a Spark cluster can safely update on a per-row basis.
- b) It is a way of updating a value inside of a variety of transformations and propagating that value to the driver node in an efficient and fault-tolerant way.
- c) You can define your own custom broadcast class by extending org.apache.spark.util.BroadcastV2 in Java or Scala or pyspark.AccumulatorParams in Python.
- d) Broadcast variables are shared, immutable variables that are cached on every machine in the cluster instead of serialized with every single task.
- e) The canonical use case is to pass around a small large table that does fit in memory on the executors.

Answer: a, b, c

Question: 3

The code blown down below intends to join df1 with df2 with inner join but it contains an error. Identify the error.

- a) d1.join(d2, "inner", d1.col("id") === df2.col("id"))
- a) The join type is not in right order. The correct query should be d2.join(d1, d1.col("id") === df2.col("id"), "inner")
- b) There should be two == instead of ===. So the correct query is d1.join(d2, "inner", d1.col("id") == df2.col("id"))
- c) Syntax is not correct d1.join(d2, d1.col("id") == df2.col("id"), "inner")
- d) We cannot do inner join in spark 3.0, but it is in the roadmap.

Answer: c

Question: 4

Which of the following are valid execution modes?

- a) Kubernetes, Local, Client
- b) Client, Cluster, Local
- c) Server, Standalone, Client
- d) Cluster, Server, Local
- e) Standalone, Client, Cluster

Answer: b

Question: 5

What command we can use to get the number of partition of a dataframe named df?

- a) df.rdd.getPartitionSize()
- b) df.getPartitionSize()
- c) df.getNumPartitions()
- d) df.rdd.getNumPartitions()

Answer: d

Question: 6

Which of the following DataFrame methods is classified as a transformation?

- a) DataFrame.count()
- b) DataFrame.show()
- c) DataFrame.select()
- d) DataFrame.foreach()
- e) DataFrame.first()

Answer: c

Question: 7

The code block displayed below contains an error. The code block is intended to join DataFrame itemsDf with the larger DataFrame transactionsDf on column itemId. Find the error.

Code block: transactionsDf.join(itemsDf, "itemId", how="broadcast")

a) The syntax is wrong, how= should be removed from the code block.
b) The join method should be replaced by the broadcast method.
c) Spark will only perform the broadcast operation if this behavior has been enabled on the Spark cluster.
d) The larger DataFrame transactionsDf is being broadcasted, rather than the smaller DataFrame itemsDf
e) broadcast is not a valid join type.

Answer: e

Question: 8

Which of the following three DataFrame operations are classified as an action?

(Choose 3 answers)

a) PrintSchema()
b) Show()
c) First()
d) limit()
e) foreach()
f) cache

Answer: b, c, e

Question: 9

If we want to create a constant integer 1 as a new column 'new_column' in a dataframe df, which code block we should select?

a) df.withColumnRenamed('new_column', lit(1))
b) df.withColumn(new_column, lit(1))
c) df.withColumn("new_column", lit("1"))
d) df.withColumn("new_column", 1)
e) df.withColumn("new_column", lit(1))

Answer: e

Question: 10

If spark is running in client mode, which of the following statement about is correct?

a) Spark driver is randomly attributed to a machine in the cluster
b) Spark driver is attributed to the machine that has the most resources
c) Spark driver remains on the client machine that submitted the application
d) The entire spark application is run on a single machine.

Answer: c

Avail the Study Guide to Pass Databricks Developer for Apache Spark - Python Apache Spark Developer Associate Exam:

- Find out about the Developer for Apache Spark - Python syllabus topics. Visiting the official site offers an idea about the exam structure and other important study resources. Going through the syllabus topics help to plan the exam in an organized manner.
- Once you are done exploring the **Databricks Developer for Apache Spark - Python syllabus**, it is time to plan for studying and covering the syllabus topics from the core. Chalk out the best plan for yourself to cover each part of the syllabus in a hassle-free manner.
- A study schedule helps you to stay calm throughout your exam preparation. It should contain your materials and thoughts like study hours, number of topics for daily studying mentioned on it. The best bet to clear the exam is to follow your schedule rigorously.
- The candidate should not miss out on the scope to learn from the **Apache Spark Developer Associate training.** Joining the Databricks provided training for this Databricks certification exam helps a candidate to strengthen his practical knowledge base from the certification.
- Learning about the probable questions and gaining knowledge regarding the exam structure helps a lot. Go through the **Databricks Developer for Apache Spark - Python sample questions** and boost your knowledge
- Make yourself a pro through online practicing the syllabus topics. Developer for Apache Spark - Python practice tests would guide you on your strengths and weaknesses regarding the syllabus topics. Through rigorous practicing, you can improve the weaker sections too. Learn well about time management during exam and become confident gradually with practice tests.

Career Benefits:

Passing the Databricks Developer for Apache Spark - Python exam, helps a candidate to prosper highly in his career. Having the certification on the resume adds to the candidate's benefit and helps to get the best opportunities.

Here Is the Trusted Practice Test for the Databricks Developer for Apache Spark - Python Certification

CertFun.Com is here with all the necessary details regarding the Developer for Apache Spark - Python exam. We provide authentic practice tests for the Developer for Apache Spark - Python exam. What do you gain from these practice tests? You get to experience the real exam-like questions made by industry experts and get a scope to improve your performance in the actual exam. Rely on CertFun.Com for rigorous, unlimited two-month attempts on the **Developer for Apache Spark - Python practice tests**, and gradually build your confidence. Rigorous practice made many aspirants successful and made their journey easy towards grabbing the Databricks Certified Associate Developer for Apache Spark.

Start Online practice of Databricks Developer for Apache Spark - Python Exam by visiting URL

https://www.certfun.com/databricks/databricks-certified-associate-developer-apache-spark-python