

C2090-101^{Q&As}

IBM Big Data Engineer

Pass IBM C2090-101 Exam with 100% Guarantee

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

<https://www.leads4pass.com/c2090-101.html>

100% Passing Guarantee
100% Money Back Assurance

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

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



QUESTION 1

When embedding SPSS models within InfoSphere Streams, what SPSS product must be installed on the same machine with InfoSphere Streams?

- A. SPSS Modeler
- B. SPSS Solution Publisher
- C. SPSS Accelerator for InfoSphere Streams
- D. None, the SPSS software runs remotely to the Streams machine

Correct Answer: B

QUESTION 2

Which of the following Hadoop command flags are used to copy data to the HDFS? (Choose two.)

- A. cp
- B. put
- C. get
- D. copyToLocal
- E. copyFromLocal

Correct Answer: DE

Reference: https://www.ibm.com/support/knowledgecenter/SSPT3X_4.2.0/com.ibm.swg.im.infosphere.biginsights.import.doc/doc/data_rest_shell.html

QUESTION 3

When we create a new table in Hive, which clause can be used in HiveSQL to indicate the storage file format?

- A. SAVE AS
- B. MAKE AS
- C. FORMAT AS
- D. STORED AS

Correct Answer: D

QUESTION 4

What is the primary purpose of Flume in the Hadoop ecosystem?

- A. To stream data from Hadoop
- B. To move static files from the local file system into HDFS
- C. To import data from a relational database or data warehouse into HDFS
- D. To capture log data as it is written to log files and move them into HDFS

Correct Answer: C

Reference: https://blogs.apache.org/flume/entry/flume_ng_architecture

QUESTION 5

Which one of the following statements is TRUE?

- A. Spark SQL does not support HiveQL
- B. Spark SQL does not support ANSI SQL
- C. To use Spark with Hive, HiveQL queries have to be rewritten in Scala
- D. Spark SQL allows relational queries expressed in SQL, HiveQL, or Scala

Correct Answer: D

Reference: <https://spark.apache.org/docs/1.2.1/sql-programming-guide.html> (overview)

QUESTION 6

PCI compliance requirements allow the use of real customer data during testing and development only when:

- A. The data is only processed in memory
- B. Customer data is never allowed in testing or development
- C. The data is never stored longer than 24 hours in the test system
- D. The data is only stored in volatile storage that expires on power loss

Correct Answer: D

QUESTION 7

For a customer satisfaction application, a large eCommerce company is loading click stream data into MySQL and running Cognos Reports which are accessed by 50 plus customer service reps (CSR). The IT team decided to move the application to the new BigInsights-based Hadoop cluster and load the data into a Hive Warehouse and run the queries. Which one of the following would enable them to reuse the SQL queries without extensive rewrite?

- A. Pig
- B. NoSQL
- C. HiveQL
- D. Big SQL

Correct Answer: D

QUESTION 8

Which component of BigInsights is able to mask data items so as restrict viewing of sensitive data?

- A. Flume
- B. HDFS
- C. Oozie
- D. Big SQL

Correct Answer: D

Reference: <https://developer.ibm.com/hadoop/docs/biginsights-ibm-open-platform/biginsights-value-add/big-sql/maximize-hadoop-data-security-ibm-infosphere-biginsights/>

QUESTION 9

One option to import data to your BigInsights cluster is to use Sqoop.

Which of the following examples shows a data exchange with a DB2 database by using the built-in DB2 connector?

- A. `$SQOOP_HOME/bin/sqoop export --connect jdbc:db2://db2.my.com:50000/myDB --username db2user --password db2pwd --table db2tbl --export-dir /sqoop/dataFile.csv`
- B. `$SQOOP_HOME/bin/sqoop import -connect jdbc:db2://db2.my.com:50000/testdb -username db2user password db2pwd -table db2tbl -split-by tbl_primarykey -target-dir sqoopimports`
- C. `$SQOOP_HOME/bin/sqoop import --connect jdbc:db2://db2.my.com:50000/testdb --username db2user --password db2pwd --table db2tbl --split-by tbl_primarykey --target-dir sqoopimports`
- D. `$BIGINSIGHTS_HOME/bin/sqoop import --connect jdbc:db2://db2.my.com:50000/testdb -- username db2user --password db2pwd --table db2tbl --split-by tbl_primarykey --target-dir sqoopimports`

Correct Answer: C

Reference: https://www.ibm.com/support/knowledgecenter/SSPT3X_3.0.0/com.ibm.swg.im.infosphere.biginsights.import.doc/doc/data_warehouse_sqoop.html

QUESTION 10

A large bank was planning to offload existing data from a data warehouse into Hadoop and use SQL queries to access historical data. Which one of the following statements is true for using HiveQL?

- A. It supports four logical operators in query predicates: IN, NOT IN, EXISTS, and NOT EXISTS
- B. It does not support nested sub-queries
- C. Hive supports all ANSI SQL 2011 syntax
- D. All of the above

Correct Answer: A

Reference: https://docs.hortonworks.com/HDPDocuments/HDP2/HDP-2.6.5/bk_data-access/content/hive013-feature-subqueries-in-where-clauses.html

QUESTION 11

Which statement about the Jaqi Programming Language is TRUE?

- A. Jaqi always produces a MapReduce job, but Combiner functionality is optional
- B. Jaqi includes the following operators: filter, extend, groupby, combine, and transform
- C. Data that is read from multiple blocks (splits) is always processed in parallel by MapReduce
- D. The read operator loads data from different source and formats, and then converts this data into JSON format for internal processing by the Jaqi interpreter

Correct Answer: C

QUESTION 12

Which of the following statements regarding Sqoop is TRUE? (Choose two.)

- A. All columns in a table must be imported
- B. Sqoop bypasses MapReduce for enhanced performance
- C. Each row from a source table is represented as a separate record in HDFS
- D. When using a password file, the file containing the password must reside in HDFS
- E. Multiple options files can be specified when invoking Sqoop from the command line

Correct Answer: CE

Reference: <https://data-flair.training/blogs/apache-sqoop-tutorial/>

QUESTION 13

You have social media data in a plain text file. You want to figure out the number of times the word "Watson" is mentioned in it.

Which of the following steps will correctly allow you to visualize that data?

- A. Create a master workbook, add a filter sheet, run the workbook, add chart
- B. Create a master workbook, create a child workbook, add a filter sheet, add chart
- C. Create a master workbook, edit the workbook, add a filter sheet, run the workbook, add chart
- D. Create a master workbook, create a child workbook, add a filter sheet, run the workbook, add chart

Correct Answer: C

QUESTION 14

Which statement is TRUE about Big SQL?

- A. The table definition can include other attributes such as the primary key or check constraints
- B. When using Big SQL, the CREATE TABLE statement cannot be embedded in an application program
- C. If a sub-table is being defined, the authorization ID can be either the same as the owner of the root table or an equivalent
- D. When defining a staging table associated with a materialized query table, the privileges held by the authorization ID of the statement only works with DBADM authority

Correct Answer: D

Reference: https://www.ibm.com/support/knowledgecenter/fi/SSCRJT_5.0.2/com.ibm.swg.im.bigsq.commsql.doc/doc/r0000927.html

QUESTION 15

Which of the following is NOT a characteristic of a consistent region?

- A. Support for non-deterministic applications
- B. Support for single-threaded operators only
- C. Allows "at-least-once" processing by using replayable data sources
- D. Allows "exactly-once" processing if using an external system that can detect duplicate tuples or can restore its state after a failure

Correct Answer: B

Reference: <https://developer.ibm.com/streamsdev/2015/02/20/processing-tuples-least-infosphere-streamsconsistent-regions/>

[C2090-101 PDF Dumps](#)

[C2090-101 Exam Questions](#)

[C2090-101 Braindumps](#)