EDUSUM
#1 Online Certification Guide

IBM C2090-101

IBM Big Data Engineer Certification Questions & Answers

Exam Summary – Syllabus –Questions

C2090-101
IBM Certified Data Engineer - Big Data
53 Questions Exam – 64% Cut Score – Duration of 75 minutes

Table of Contents:

Know Your C2090-101 Certification Well:

The C2090-101 is best suitable for candidates who want to gain knowledge in the IBM Data and AI - Platform Analytics. Before you start your C2090-101 preparation you may struggle to get all the crucial Big Data Engineer materials like C2090-101 syllabus, sample questions, study guide.

But don't worry the C2090-101 PDF is here to help you prepare in a stress free manner.

The PDF is a combination of all your queries like-
- What is in the C2090-101 syllabus?
- How many questions are there in the C2090-101 exam?
- Which Practice test would help me to pass the C2090-101 exam at the first attempt?

Passing the C2090-101 exam makes you IBM Certified Data Engineer - Big Data. Having the Big Data Engineer certification opens multiple opportunities for you. You can grab a new job, get a higher salary or simply get recognition within your current organization.

IBM C2090-101 Big Data Engineer Certification Details:

Exam Name	IBM Certified Data Engineer - Big Data
Exam Code	C2090-101
Exam Price	$200 (USD)
Duration	75 mins
Number of Questions	53
Passing Score	64%
Books / Training	**IBM Knowledge Center**
Schedule Exam	**Pearson VUE**
Sample Questions	**IBM Big Data Engineer Sample Questions**
Practice Exam	**IBM C2090-101 Certification Practice Exam**

C2090-101 Syllabus:

Topic	Details	Weights
Data Loading	- Load unstructured data into InfoSphere BigInsights - Import streaming data into Hadoop using InfoSphere Streams - Create a BigSheets workbook - Import data into Hadoop and create Big SQL table definitions - Import data to HBase - Import data to Hive - Use Data Click to load from relational sources into InfoSphere BigInsights with a self-service process - Extract data from a relational source using Sqoop - Load log data into Hadoop using Flume - Insert data via IBM General Parallel File System (GPFS) Posix file system API - Load data with Hadoop command line utility	34%
Data Security	- Keep data secure within PCI standards - Uses masking (e.g. Optim, Big SQL), and redaction to protect sensitive data	8%
Architecture and Integration	- Implement MapReduce - Evaluate use cases for selecting Hive, Big SQL, or HBase - Create and/or query a Solr index - Evaluate use cases for selecting potential file formats (e.g. JSON, CSV, Parquet, Sequence, etc..) - Utilize Apache Hue for search visualization	17%
Performance and Scalability	- Use Resilient Distributed Dataset (RDD) to improve MapReduce performance - Choose file formats to optimize performance of Big SQL, JAQL, etc. - Make specific performance tuning decisions for Hive and HBase - Analyze performance considerations when using Apache Spark	15%
Data Preparation, Transformation, and Export	- Use Jaql query methods to transform data in InfoSphere BigInsights - Capture and prep social data for analytics - Integrating SPSS model scoring in InfoSphere Streams - Implement entity resolution within a Big Data platform (e.g. Big Match) - Utilize Pig for data transformation and data manipulation - Use Big SQL to transform data in InfoSphere BigInsights - Export processing results out of Hadoop (e.g. DataClick, DataStage, etc.) - Utilize consistent regions in InfoSphere Streams to ensure at least once processing	26%

IBM C2090-101 Sample Questions:

Question: 1

Use of Bulk Load in HBase for loading large volume of data will result in which of the following?

    a) It will use less CPU but will use more network resource
    b) It will use less network resource but more CPU
    c) It will behave same way as using HBase API for loading large volume of data
    d) None of the above

Answer: d

Question: 2

A Resilient Distributed Dataset supports which of the following?

    a) Creating a new dataset from an old one
    b) Returning a computed value to the driver program
    c) Both Returning a computed value to the driver program and Creating a new dataset from an old one
    d) Neither Creating a new dataset from an old one nor Returning a computed value to the driver program

Answer: d

Question: 3

How are insights derived from Big Match moved to an MDM system?

    a) Extract insights from HBase and load into MDM through an API call
    b) Extract insights from Hive and load into MDM using standard tooling
    c) Extract insights from HDFS and load into MDM by simulating delta load
    d) Extract insights from HBase and load into MDM using standard MDM batch processing tool

Answer: c

Question: 4

In order for an SPSS Modeler stream to be incorporated for use in an InfoSphere Streams application leveraging SPSS Modeler Solution Publisher, you need to:

a) add a Type node
b) insert any Output node
c) add a Table node as the terminal node
d) Make the terminal node a scoring branch

Answer: d

Question: 5

Which of the following are capabilities of the Apache Spark project?

a) Large scale graph processing
b) Large scale machine learning
c) Live data stream processing
d) All of the above

Answer: b

Question: 6

Which ONE of the following statements regarding Sqoop is TRUE?
a) HBase is not supported as an import target
b) Data imported using Sqoop is always written to a single Hive partition
c) Sqoop can be used to retrieve rows newer than some previously imported set of rows
d) Sqoop can only append new rows to a database table when exporting back to a database

Answer: c

Question: 7

The number of partitions created by Dynamic Partitions in Hive can be controlled by which of the following?
a) hive.exec.max.dynamic.partitions
b) hive.exec.max.dynamic.partitions.pernode
c) hive.exec.max.created.files
d) All of the above

Answer: b

Question: 8

The GPFS implementation of Data Management API is compliant to which Open Group storage management Standard?

a) XSH
b) XBD
c) XDSM
d) X/Open

Answer: c

Question: 9

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

Answer: d

Question: 10

What is a method for loading RDBMS data into an HBase table?

a) HDFSLOAD
b) SQOOP IMPORT
c) LOAD HADOOP USING
d) hadoop jar hbase-VERSIONjar importtsv

Answer: c

Study Guide to Crack IBM Big Data Engineer C2090-101 Exam:

- Getting details of the C2090-101 syllabus, is the first step of a study plan. This pdf is going to be of ultimate help. Completion of the syllabus is must to pass the C2090-101 exam.
- Making a schedule is vital. A structured method of preparation leads to success. A candidate must plan his schedule and follow it rigorously to attain success.
- Joining the IBM provided training for C2090-101 exam could be of much help. If there is specific training for the exam, you can discover it from the link above.
- Read from the C2090-101 sample questions to gain your idea about the actual exam questions. In this PDF useful sample questions are provided to make your exam preparation easy.
- Practicing on C2090-101 practice tests is must. Continuous practice will make you an expert in all syllabus areas.

Reliable Online Practice Test for C2090-101 Certification

Make EduSum.com your best friend during your IBM Big Data Engineer exam preparation. We provide authentic practice tests for the C2090-101 exam. Experts design these online practice tests, so we can offer you an exclusive experience of taking the actual C2090-101 exam. We guarantee you 100% success in your first exam attempt if you continue practicing regularly. Don't bother if you don't get 100% marks in initial practice exam attempts. Just utilize the result section to know your strengths and weaknesses and prepare according to that until you get 100% with our practice tests. Our evaluation makes you confident, and you can score high in the C2090-101 exam.

Start Online practice of C2090-101 Exam by visiting URL
https://www.edusum.com/ibm/c2090-101-ibm-big-data-engineer