VCE & PDF
PassApply.com
https://www.passapply.com

CCA175^{Q&As}

CCA Spark and Hadoop Developer Exam

Pass Cloudera CCA175 Exam with 100% Guarantee

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

https://www.passapply.com/cca175.html

100% Passing Guarantee
100% Money Back Assurance

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

🔧 **Instant Download** After Purchase

🔧 **100% Money Back** Guarantee

🔧 **365 Days** Free Update

🔧 **800,000+** Satisfied Customers

QUESTION 1

Problem Scenario 20 : You have been given MySQL DB with following details. user=retail_dba password=cloudera database=retail_db table=retail_db.categories jdbc URL = jdbc:mysql://quickstart:3306/retail_db Please accomplish following activities.

1. Write a Sqoop Job which will import "retaildb.categories" table to hdfs, in a directory name "categories_targetJob".

Correct Answer: See the explanation for Step by Step Solution and configuration.

Solution :

Step 1 : Connecting to existing MySQL Database mysql -user=retail_dba -password=cloudera retail_db

Step 2 : Show all the available tables show tables;

Step 3 : Below is the command to create Sqoop Job (Please note that - import space is

mandatory)

sqoop job -create sqoopjob \ -- import \

-connect "jdbc:mysql://quickstart:3306/retail_db" \

-username=retail_dba \

-password=cloudera \

-table categories \

-target-dir categories_targetJob \

-fields-terminated-by \\'|\\' \ -lines-terminated-by \\'\n\\' Step 4 : List all the Sqoop Jobs sqoop job --list Step 5 : Show details of the Sqoop Job sqoop job --show sqoopjob Step 6 : Execute the sqoopjob sqoopjob --exec sqoopjob Step 7 : Check the output of import job hdfs dfs -ls categories_target_job hdfs dfs -cat categories_target_job/part*

QUESTION 2

Problem Scenario 84 : In Continuation of previous question, please accomplish following activities.

1.

 Select all the products which has product code as null

2.

 Select all the products, whose name starts with Pen and results should be order by Price descending order.

3.

Select all the products, whose name starts with Pen and results should be order by Price descending order and quantity ascending order.

4.

Select top 2 products by price

Correct Answer: See the explanation for Step by Step Solution and configuration.

Solution : Step 1 : Select all the products which has product code as null val results = sqlContext.sql(......SELECT\\' FROM products WHERE code IS NULL......) results. showQ val results = sqlContext.sql(......SELECT * FROM products WHERE code = NULL ",,M) results.showQ Step 2 : Select all the products , whose name starts with Pen and results should be order by Price descending order. val results = sqlContext.sql(......SELECT * FROM products WHERE name LIKE \\'Pen %\\' ORDER BY price DESC......) results. showQ Step 3 : Select all the products , whose name starts with Pen and results should be order by Price descending order and quantity ascending order. val results = sqlContext.sql(\\'.....SELECT * FROM products WHERE name LIKE \\'Pen %\\' ORDER BY price DESC, quantity......) results. showQ Step 4 : Select top 2 products by price val results = sqlContext.sql(......SELECT\\' FROM products ORDER BY price desc LIMIT2......} results. show()

QUESTION 3

Problem Scenario 96 : Your spark application required extra Java options as below. XX:+PrintGCDetails-XX:+PrintGCTimeStamps Please replace the XXX values correctly ./bin/spark-submit --name "My app" --master local[4] --conf spark.eventLog.enabled=talse -conf XXX hadoopexam.jar

Correct Answer: See the explanation for Step by Step Solution and configuration.

Solution

XXX: Mspark.executoi\extraJavaOptions=-XX:+PrintGCDetails -XX:+PrintGCTimeStamps"

Notes: ./bin/spark-submit \

--class

--master \

--deploy-mode \

-conf = \

other options

\

[application-arguments]

Here, conf is used to pass the Spark related contigs which are required for the application

to run like any specific property(executor memory) or if you want to override the default

property which is set in Spark-default.conf.

QUESTION 4

Problem Scenario 92 : You have been given a spark scala application, which is bundled in

jar named hadoopexam.jar.

Your application class name is com.hadoopexam.MyTask

You want that while submitting your application should launch a driver on one of the cluster

node.

Please complete the following command to submit the application.

spark-submit XXX -master yarn \

YYY SSPARK HOME/lib/hadoopexam.jar 10

Correct Answer: See the explanation for Step by Step Solution and configuration.

Solution

XXX: -class com.hadoopexam.MyTask YYY : --deploy-mode cluster

---

QUESTION 5

Problem Scenario 67 : You have been given below code snippet.

lines = sc.parallelize([\\\'Its fun to have fun,\\\',\\\'but you have to know how.\\\'])

M = lines.map(lambda x: x.replace(\\\',7 \\\').replace(\\\'.\\\',\\\' \\\'J.replaceC-V \\\').lower())

r2 = r1.flatMap(lambda x: x.split())

r3 = r2.map(lambda x: (x, 1))

operation1

r5 = r4.map(lambda x:(x[1],x[0]))

r6 = r5.sortByKey(ascending=False)

r6.take(20)

Write a correct code snippet for operationl which will produce desired output, shown below.

[(2, \\\'fun\\\'), (2, \\\'to\\\'), (2, \\\'have\\\'), (1, its\\\'), (1, \\\'know1), (1, \\\'how1), (1, \\\'you\\\'), (1, \\\'but\\\')]

Correct Answer: See the explanation for Step by Step Solution and configuration.

Solution : r4 = r3.reduceByKey(lambda x,y:x+y)