

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.leads4pass.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 57 : You have been given below code snippet.

```
val a = sc.parallelize(1 to 9, 3) operationl
```

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

```
Array[(String, Seq[Int])] = Array((even,ArrayBuffer(2, 4, G, 8)), (odd,ArrayBuffer(1, 3, 5, 7, 9)))
```

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

Solution :

```
a.groupBy(x => {if (x % 2 == 0) "even" else "odd" }).collect
```

QUESTION 2

Problem Scenario 36 : You have been given a file named spark8/data.csv (type,name). data.csv 1,Lokesh 2,Bhupesh 2,Amit 2,Ratan 2,Dinesh 1,Pavan 1,Tejas 2,Sheela 1,Kumar 1,Venkat

1. Load this file from hdfs and save it back as (id, (all names of same type)) in results directory. However, make sure while saving it should be

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

Solution :

Step 1 : Create file in hdfs (We will do using Hue). However, you can first create in local filesystem and then upload it to hdfs.

Step 2 : Load data.csv file from hdfs and create PairRDDs

```
val name = sc.textFile("spark8/data.csv")
```

```
val namePairRDD = name.map(x=> (x.split(",")(0),x.split(",")(1)))
```

Step 3 : Now swap namePairRDD RDD.

```
val swapped = namePairRDD.map(item => item.swap)
```

Step 4 : Now combine the rdd by key.

```
val combinedOutput = namePairRDD.combineByKey(List(_), (x:List[String], y:String) => y ::
```

```
x, (x:List[String], y:List[String]) => x ::: y)
```

Step 5 : Save the output as a Text file and output must be written in a single file.

```
combinedOutput.repartition(1).saveAsTextFile("spark8/result.txt")
```

QUESTION 3

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(" ", "").replace(".", "").replace("-", "").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 operation1 which will produce desired output, shown below.

```
[(2, 'fun'), (2, 'to'), (2, 'have'), (1, 'its'), (1, 'know'), (1, 'how'), (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)`

QUESTION 4

Problem Scenario 10 : You have been given following mysql database details as well as other info. user=retail_dba password=cloudera database=retail_db jdbc URL = jdbc:mysql://quickstart:3306/retail_db Please accomplish following.

1. Create a database named hadoopexam and then create a table named departments in it, with following fields. department_id int, department_name string

e.g. location should be `hdfs://quickstart.cloudera:8020/user/hive/warehouse/hadoopexam.db/departments`

2.

Please import data in existing table created above from `retaildb.departments` into hive table `hadoopexam.departments`.

3.

Please import data in a non-existing table, means while importing create hive table named `hadoopexam.departments_new`

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

Solution :

Step 1 : Go to hive interface and create database.

hive

create database hadoopexam;

Step 2. Use the database created in above step and then create table in it. use

hadoopexam; show tables;

Step 3 : Create table in it.

create table departments (department_id int, department_name string);

show tables;

desc departments;

desc formatted departments;

Step 4 : Please check following directory must not exist else it will give error, hdfs dfs -ls

/user/cloudera/departments

If directory already exists, make sure it is not useful and then delete the same.

This is the staging directory where Sqoop store the intermediate data before pushing in

hive table.

hadoop fs -rm -R departments

Step 5 : Now import data in existing table

sqoop import \

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

~username=retail_dba \

-password=cloudera \

--table departments \

-hive-home /user/hive/warehouse \

-hive-import \

-hive-overwrite \

-hive-table hadoopexam.departments

Step 6 : Check whether data has been loaded or not.

hive;

use hadoopexam;

show tables;

```
select" from departments;
```

```
desc formatted departments;
```

Step 7 : Import data in non-existing tables in hive and create table while importing.

```
sqoop import \
```

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

```
--username=retail_dba \
```

```
--password=cloudera \
```

```
-table departments \
```

```
-hive-home /user/hive/warehouse \
```

```
-hive-import \
```

```
-hive-overwrite \
```

```
-hive-table hadoopexam.departments_new \
```

```
-create-hive-table
```

Step 8 : Check-whether data has been loaded or not.

```
hive;
```

```
use hadoopexam;
```

```
show tables;
```

```
select" from departments_new;
```

```
desc formatted departments_new;
```

QUESTION 5

Problem Scenario 40 : You have been given sample data as below in a file called spark15/file1.txt

```
3070811,1963,1096,,"US","CA",,1, 3022811,1963,1096,,"US","CA",,1,56 3033811,1963,1096,,"US","CA",,1,23 Below is the code snippet to process this tile. val field= sc.textFile("spark15/f ile1.txt") val mapper = field.map(x=> A) mapper.map(x => x.map(x=> {B})).collect
```

Please fill in A and B so it can generate below final output

```
Array(Array(3070811,1963,109G, 0, "US", "CA", 0,1, 0)
```

```
,Array(3022811,1963,1096, 0, "US", "CA", 0,1, 56)
```

```
,Array(3033811,1963,1096, 0, "US", "CA", 0,1, 23)
```

```
)
```

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

Solution :

A. `x.split(",",-1)`

B. `if (x.isEmpty) 0 else x`

[Latest CCA175 Dumps](#)

[CCA175 Exam Questions](#)

[CCA175 Braindumps](#)