



# HDPCD<sup>Q&As</sup>

Hortonworks Data Platform Certified Developer

## Pass Hortonworks HDPCD Exam with 100% Guarantee

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

<https://www.pass4itsure.com/hdpcd.html>

100% Passing Guarantee  
100% Money Back Assurance

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

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



**QUESTION 1**

Review the following data and Pig code: What command to define B would produce the output (M,62,95102) when invoking the DUMP operator on B?

```
M, 38, 95111
F, 29, 95060
F, 45, 95192
M, 62, 95102
F, 56, 95102
```

```
A = LOAD 'data' USING PigStorage(',')
AS (gender:chararray, age:int, zip:chararray);
```

- A. B = FILTER A BY (zip == \"95102\" AND gender == \"M\");
- B. B= FOREACH A BY (gender == \"M\" AND zip == \"95102\");
- C. B = JOIN A BY (gender == \"M\" AND zip == \"95102\");
- D. B= GROUP A BY (zip == \"95102\" AND gender == \"M\");

Correct Answer: A

---

**QUESTION 2**

You've written a MapReduce job that will process 500 million input records and generated 500 million key-value pairs. The data is not uniformly distributed. Your MapReduce job will create a significant amount of intermediate data that it needs to transfer between mappers and reduces which is a potential bottleneck. A custom implementation of which interface is most likely to reduce the amount of intermediate data transferred across the network?

- A. Partitioner
- B. OutputFormat
- C. WritableComparable
- D. Writable
- E. InputFormat
- F. Combiner

Correct Answer: F

---



Explanation: Combiners are used to increase the efficiency of a MapReduce program. They are used to aggregate intermediate map output locally on individual mapper outputs. Combiners can help you reduce the amount of data that needs to be transferred across to the reducers. You can use your reducer code as a combiner if the operation performed is commutative and associative.

Reference: 24 Interview Questions and Answers for Hadoop MapReduce developers, What are combiners? When should I use a combiner in my MapReduce Job?

---

### QUESTION 3

On a cluster running MapReduce v1 (MRv1), a TaskTracker heartbeats into the JobTracker on your cluster, and alerts the JobTracker it has an open map task slot.

What determines how the JobTracker assigns each map task to a TaskTracker?

- A. The amount of RAM installed on the TaskTracker node.
- B. The amount of free disk space on the TaskTracker node.
- C. The number and speed of CPU cores on the TaskTracker node.
- D. The average system load on the TaskTracker node over the past fifteen (15) minutes.
- E. The location of the InputSplit to be processed in relation to the location of the node.

Correct Answer: E

Explanation: The TaskTrackers send out heartbeat messages to the JobTracker, usually every few minutes, to reassure the JobTracker that it is still alive. These messages also inform the JobTracker of the number of available slots, so the JobTracker can stay up to date with where in the cluster work can be delegated. When the JobTracker tries to find somewhere to schedule a task within the MapReduce operations, it first looks for an empty slot on the same server that hosts the DataNode containing the data, and if not, it looks for an empty slot on a machine in the same rack.

Reference: 24 Interview Questions and Answers for Hadoop MapReduce developers, How JobTracker schedules a task?

---

### QUESTION 4

Which best describes how TextInputFormat processes input files and line breaks?

- A. Input file splits may cross line breaks. A line that crosses file splits is read by the RecordReader of the split that contains the beginning of the broken line.
- B. Input file splits may cross line breaks. A line that crosses file splits is read by the RecordReaders of both splits containing the broken line.
- C. The input file is split exactly at the line breaks, so each RecordReader will read a series of complete lines.
- D. Input file splits may cross line breaks. A line that crosses file splits is ignored.
- E. Input file splits may cross line breaks. A line that crosses file splits is read by the RecordReader of the split that contains the end of the broken line.



Correct Answer: A

Reference: How Map and Reduce operations are actually carried out

---

#### QUESTION 5

Which one of the following statements is false about HCatalog?

- A. Provides a shared schema mechanism
- B. Designed to be used by other programs such as Pig, Hive and MapReduce
- C. Stores HDFS data in a database for performing SQL-like ad-hoc queries
- D. Exists as a subproject of Hive

Correct Answer: C

[Latest HDPCD Dumps](#)

[HDPCD PDF Dumps](#)

[HDPCD Study Guide](#)