



# 70-775<sup>Q&As</sup>

Perform Data Engineering on Microsoft Azure HDInsight

## Pass Microsoft 70-775 Exam with 100% Guarantee

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

<https://www.geekcert.com/70-775.html>

100% Passing Guarantee  
100% Money Back Assurance

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

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





**QUESTION 1**

You have an Apache HBase cluster in Azure HDInsight. The cluster has a table named sales that contains a column family named customerfamily.

You need to add a new column family named customeraddr to the sales table.

How should you complete the command? To answer, drag the appropriate values to the correct targets. Each value may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Select and Place:

**Values**

- alter
- customeraddr
- disable
- enable
- create
- customerfamily
- drop

**Answer area**

```
Hbase > Value 'sales'
Hbase > value 'sales'
' value ',
NAME=> ' Value ',
IN_MEMORY => 'false' }
Hbase > Value 'sales'
```

Correct Answer:



### Values

Empty text boxes for values: create, drop

### Answer area

```
Hbase > disable 'sales'  
Hbase > alter 'sales'  
    'customerfamily',  
    {NAME=> 'customeraddr',  
      IN_MEMORY => 'false' }  
Hbase > enable 'sales'
```

### QUESTION 2

You have an Apache Hive cluster in Azure HDInsight.

You need to tune a Hive query to meet the following requirements:

How should you complete the query? To answer, drag the appropriate values to the correct targets. Each value may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Select and Place:



**Values**

hive.execution.engine

hive.prewarm.enabled

hive.tez.auto.reducer.parallelism

hive.tez.container.size

hive.vectorized.execution.enabled

tez.runtime.compress.codec

**Answer area**

Set Value = Tez

Set Value = true

SELECT \* FROM sales;

Correct Answer:

**Values**

hive.prewarm.enabled

hive.tez.auto.reducer.parallelism

hive.tez.container.size

tez.runtime.compress.codec

**Answer area**

Set hive.execution.engine = Tez

Set hive.vectorized.execution.enabled = true

SELECT \* FROM sales;

**QUESTION 3**

You use YARN to manage the resources for a Spark Thrift Server running on a Linux- based Apache Spark cluster in Azure HDInsight.

You discover that the cluster does not fully utilize the resources. You want to increase resource allocation.

You need to increase the number of executors and the allocation of memory to the Spark Thrift Server driver.

Which two parameters should you modify? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A. spark.dynamicAllocation.maxExecutors



- B. spark.cores.max
- C. spark.executor.memory
- D. spark\_thrift\_cmd\_opts
- E. spark.executor.instances

Correct Answer: AC

References: <https://stackoverflow.com/questions/37871194/how-to-tune-spark-executor-number-cores-and-executor-memory>

---

#### QUESTION 4

Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series.

Information and details provided in a question apply only to that question.

You are implementing a batch processing solution by using Azure HDInsight.

You have data stored in Azure.

You need to ensure that you can access the data by using Azure Active Directory (Azure AD) identities.

What should you do?

- A. Use a shuffle join in an Apache Hive query that stores the data in a JSON format.
- B. Use a broadcast join in an Apache Hive query that stores the data in an ORC format.
- C. Increase the number of spark.executor.cores in an Apache Spark job that stores the data in a text format.
- D. Increase the number of spark.executor.instances in an Apache Spark job that stores the data in a text format.
- E. Decrease the level of parallelism in an Apache Spark job that stores the data in a text format.
- F. Use an action in an Apache Oozie workflow that stores the data in a text format.
- G. Use an Azure Data Factory linked service that stores the data in Azure Data Lake.
- H. Use an Azure Data Factory linked service that stores the data in an Azure DocumentDB database.

Correct Answer: G

References: <https://docs.microsoft.com/en-us/azure/data-factory/concepts-datasets-linked-services>

---

#### QUESTION 5

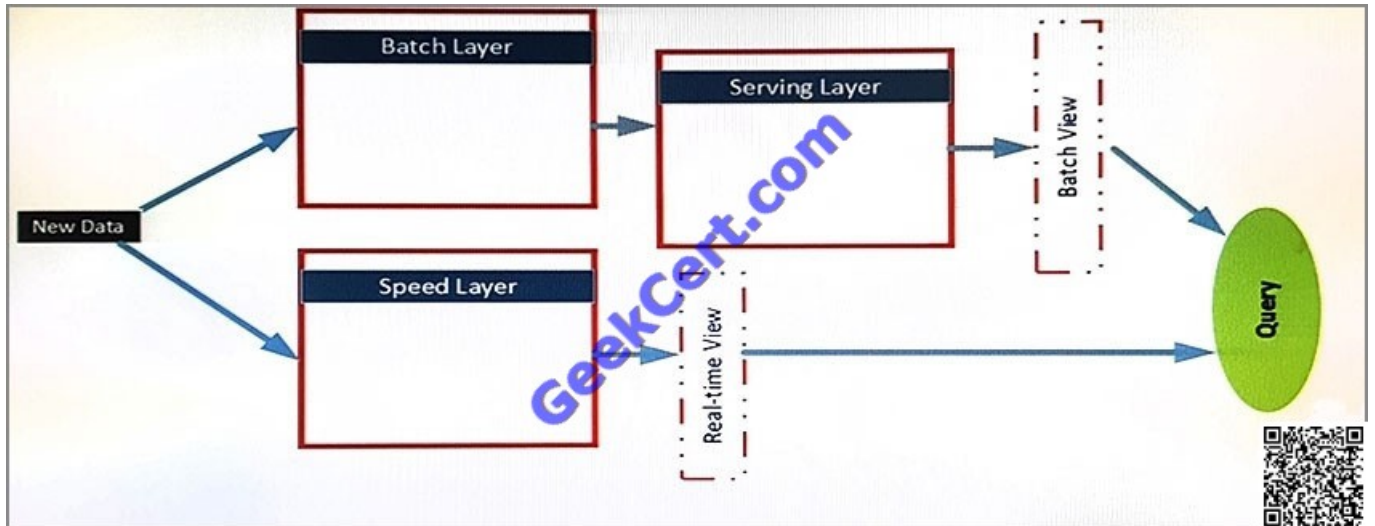
Note: This question is part of a series of questions that use the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is



exactly the same in each question in this series.

You are planning a big data infrastructure by using an Apache Spark cluster in Azure HDInsight. The cluster has 24 processor cores and 512 GB of memory.

The architecture of the infrastructure is shown in the exhibit. (Click the Exhibit button.)



The architecture will be used by the following users:

The data sources in the batch layer share a common storage container. The following data sources are used:

You need to ensure that the support analysts can develop embedded analytics applications by using the least amount of development effort.

Which technology should you implement?

- A. Zeppelin
- B. Jupyter
- C. Apache Ambari
- D. Livy

Correct Answer: D

References: <https://docs.microsoft.com/en-us/azure/hdinsight/hdinsight-apache-spark-livy-rest-interface>





To Read the [Whole Q&As](#), please purchase the [Complete Version](#) from [Our website](#).

## Try our product !

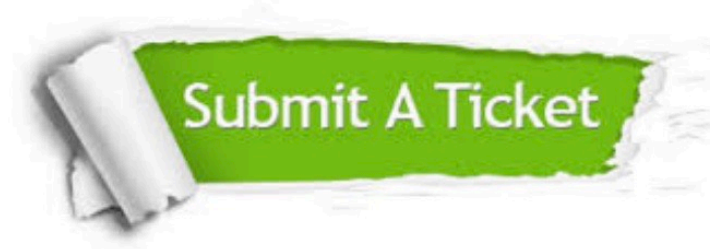
100% Guaranteed Success  
100% Money Back Guarantee  
365 Days Free Update  
Instant Download After Purchase  
24x7 Customer Support  
Average 99.9% Success Rate  
More than 800,000 Satisfied Customers Worldwide  
Multi-Platform capabilities - [Windows](#), [Mac](#), [Android](#), [iPhone](#), [iPod](#), [iPad](#), [Kindle](#)

We provide exam PDF and VCE of Cisco, Microsoft, IBM, CompTIA, Oracle and other IT Certifications.  
You can view Vendor list of All Certification Exams offered:

<https://www.geekcert.com/allproducts>

## Need Help

Please provide as much detail as possible so we can best assist you.  
To update a previously submitted ticket:



 <p><b>One Year Free Update</b> Free update is available within One Year after your purchase. After One Year, you will get 50% discounts for updating. And we are proud to boast a 24/7 efficient Customer Support system via Email.</p>	 <p><b>Money Back Guarantee</b> To ensure that you are spending on quality products, we provide 100% money back guarantee for 30 days from the date of purchase.</p>	 <p><b>Security &amp; Privacy</b> We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information &amp; peace of mind.</p>
---	---	--

Any charges made through this site will appear as Global Simulators Limited.  
All trademarks are the property of their respective owners.  
Copyright © geekcert, All Rights Reserved.