

## Week 2 Graded Quiz

### Question 1

Which statement creates a database named **mydatabase**?

Ans: CREATE DATABASE mydatabase;

### Question 2

A new table is created using the following statement. The database used is in the default storage location in the Hive warehouse. Which statements describe the expected outcomes of this statement? Check all that apply.

**CREATE TABLE thisdb.thistable (id TINYINT, name STRING);**

Ans: The table is in the database **thisdb**

The table is configured to store data in a directory named **thistable**

The table's storage directory is a subdirectory of **/user/hive/warehouse/thisdb.db**

### Question 3

A data file specifies a song (such as "Bohemian Rhapsody") on an album (in this case A Night At The Opera) by an artist or group (Queen). The file uses the pipe character (|) to separate values, so the example would look like this row:

**Bohemian Rhapsody|A Night At The Opera|Queen**

Which statement is appropriate to define a table using data in this format?

Ans: CREATE TABLE songs (song STRING, album STRING, artist STRING)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY '|';

### Question 4

The data for a table to be called **weblogs** is provided in Parquet files (with format **PARQUET**) and are placed in S3 in a directory named **weblogs** in the bucket named **training-coursera1**. Which statement correctly creates this table? (Assume the column list is correct.)

Ans: CREATE EXTERNAL TABLE weblogs (...)

STORED AS PARQUET

LOCATION 's3a://training-coursera1/weblogs/';

### Question 5

Which statement correctly uses a SerDe for reading and writing **newtable**'s data?

Ans: CREATE TABLE newtable (col1 STRING, col2 INT)

ROW FORMAT SERDE 'org.apache.hadoop.hive.serde2.OpenCSVSerde';

### Question 6

An alternative to using **CREATE EXTERNAL TABLE** when creating an externally managed table is to set the table property **EXTERNAL** to **TRUE**. Which of the following would correctly do this?

Ans: CREATE TABLE table\_with\_header (col1 INT, col2 STRING)

TBLPROPERTIES ('EXTERNAL'='TRUE');

Question 7

Which commands are valid ways to change existing table schemas using Apache Impala? Check all that apply.

Ans: ALTER TABLE investors DROP COLUMN share;  
ALTER TABLE investors CHANGE amount quantity INT;

Question 8

Which statements describe the differences between dropping **int\_table**, which was created using **CREATE TABLE** (with no later alterations), and **ext\_table**, which was created using **CREATE EXTERNAL TABLE** (with no later alterations)?

Ans: Dropping **int\_table** might delete the data for the table, but dropping **ext\_table** will not drop the data for that table.

Dropping **int\_table** might delete the directory in which its data is stored, but dropping **ext\_table** will not.

Question 9

For which of the following is Impala a better choice than Hive? Check all that apply.

Ans: Dropping a column using the **ALTER TABLE ... DROP COLUMN** command  
Running queries on tables that are available in both engines and for which you want a fast response

Question 10

Suppose you have been querying a table named **mytable** using Impala. You then added data to **mytable** using an **hdfs dfs** command, and you want to query the table in Impala again, with the new data. What is your best course of action?

Ans: Run **REFRESH newtable;** then query as usual