# **Hadoop Spark Assignment 7.3**

## **Hive Data Definitions**

It is a subset of Hive SQL statements that describe the data structure in Hive by creating, deleting, or altering schema objects such as databases, tables, views, partitions, and buckets. Most Hive DDL statements start with the keywords CREATE, DROP, or ALTER. The syntax of Hive DDL is very similar to the DDL in SQL. The comments in Hive start from --.

### Create the database without checking whether the database already exists:

*CREATE DATABASE akshs;*

### Create the database and check whether the database already exists:

*CREATE DATABASE IF NOT EXISTS akshs;*

### Create the database with location, comments, and metadata information:

*CREATE DATABASE IF NOT EXISTS akshs*

*COMMENT 'hive database demo'*

*LOCATION '/hdfs/directory'*

*WITH DBPROPERTIES ('creator'=’abdcd,'date'='2015-01-01');*

### Describe database:

Describe akshs;

### Use the database:

Use database\_name;

### Drop the empty database:

Drop database if exists akshs;

Hive Data Manipulations/   
 HiveQL Manipulations

the *data manipulation language* parts that are used to put data into tables and to extract data from tables to the filesystem.

### Loading files into tables:

Hive does not do any transformation while loading data into tables. Load operations are currently pure copy/move operations that move datafiles into locations corresponding to Hive tables.

Syntax:

|  |
| --- |
| LOAD DATA [LOCAL] INPATH 'filepath' [OVERWRITE] INTO TABLE tablename  [PARTITION (partcol1=val1, partcol2=val2 ...)] |

### Inserting data into Hive Tables from queries

INSERT OVERWRITE TABLE tablename1 [PARTITION (partcol1=val1, partcol2=val2 ...) [IF NOT EXISTS]] select\_statement1 FROM from\_statement;

INSERT INTO TABLE tablename1 [PARTITION (partcol1=val1, partcol2=val2 ...)] select\_statement1 FROM from\_statement;

### Inserting values into tables from SQL

INSERT INTO TABLE tablename [PARTITION (partcol1[=val1], partcol2[=val2] ...)] VALUES values\_row [, values\_row ...]

Where values\_row is:

( value [, value ...] )

where a value is either null or any valid SQL literal

### Update

Updates can only be performed on tables that support ACID.

UPDATE table name SET column = value [, column = value ...] [WHERE expression]

### Delete:

Deletes can only be performed on tables that support ACID

DELETE FROM tablename [WHERE expression]

### Merge

### Merge allows actions to be performed on a target table based on the results of a join with a source table.

MERGE INTO <target table> AS T USING <source expression/table> AS S

ON <boolean expression1>

WHEN MATCHED [AND <boolean expression2>] THEN UPDATE SET <set clause list>

WHEN MATCHED [AND <boolean expression3>] THEN DELETE

WHEN NOT MATCHED [AND <boolean expression4>] THEN INSERT VALUES<value list>