

Assignment

Task1:

Command used: a) create database custom; → database created is stored as directory in hive and metadata in mysql.

b) `dfs -ls /user/hive/warehouse` → check database

```
acagild@localhost:~/install/hive/apache-hive-2.3.2-bin/conf
Wed Mar 28 21:05:14 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.
Wed Mar 28 21:05:15 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.
Wed Mar 28 21:05:15 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.
OK
Time taken: 9.847 seconds
hive> dfs -ls /user/hive
      > ;
Found 1 items
drwxrwxrwx - acagild supergroup          0 2018-03-28 21:05 /user/hive/warehouse
hive> dfs -ls /user/hive/warehouse
      > ;
Found 10 items
drwxrwxrwx - acagild supergroup          0 2018-02-24 17:45 /user/hive/warehouse/acagilddb
drwxrwxrwx - acagild supergroup          0 2018-02-17 15:12 /user/hive/warehouse/advance.db
drwxrwxrwx - acagild supergroup          0 2018-02-17 15:15 /user/hive/warehouse/advanced.db
drwxrwxrwx - acagild supergroup          0 2018-02-17 16:28 /user/hive/warehouse/buck_user
drwxrwxrwx - acagild supergroup          0 2018-03-28 21:05 /user/hive/warehouse/custom.db
drwxrwxrwx - acagild supergroup          0 2018-02-18 15:06 /user/hive/warehouse/json_table
drwxrwxrwx - acagild supergroup          0 2018-02-17 15:46 /user/hive/warehouse/locations
drwxrwxrwx - acagild supergroup          0 2018-02-11 16:17 /user/hive/warehouse/retail.db
drwxrwxrwx - acagild supergroup          0 2018-03-28 20:04 /user/hive/warehouse/train_sales_stage
drwxrwxrwx - acagild supergroup          0 2018-02-17 15:46 /user/hive/warehouse/users
```

c) create table temperature_data →

create table temperature_data

(temp_data String,

Zipcode INT,

Temp INT) row format delimited fields terminated by ',';

```
acagild@localhost:~/install/hive/apache-hive-2.3.2-bin/conf
> ;
Found 1 items
drwxrwxrwx - acagild supergroup 0 2018-03-28 21:05 /user/hive/warehouse
hive> dfs -ls /user/hive/warehouse
> ;
Found 10 items
drwxrwxrwx - acagild supergroup 0 2018-02-24 17:45 /user/hive/warehouse/acagilddb
b.db
drwxrwxrwx - acagild supergroup 0 2018-02-17 15:12 /user/hive/warehouse/advance.d
b
drwxrwxrwx - acagild supergroup 0 2018-02-17 15:15 /user/hive/warehouse/advanced.
db
drwxrwxrwx - acagild supergroup 0 2018-02-17 16:28 /user/hive/warehouse/buck_user
s
drwxrwxrwx - acagild supergroup 0 2018-03-28 21:05 /user/hive/warehouse/custom.db
drwxrwxrwx - acagild supergroup 0 2018-02-18 15:06 /user/hive/warehouse/json_tabl
e
drwxrwxrwx - acagild supergroup 0 2018-02-17 15:46 /user/hive/warehouse/locations
drwxrwxrwx - acagild supergroup 0 2018-02-11 16:17 /user/hive/warehouse/retail.db
drwxrwxrwx - acagild supergroup 0 2018-03-28 20:04 /user/hive/warehouse/train_sal
es_stage
drwxrwxrwx - acagild supergroup 0 2018-02-17 15:46 /user/hive/warehouse/users
hive> create table temperature_data
> (temp_date String,
> zipcode INT,
> temp INT) row format delimited fields terminated by ',';
OK
Time taken: 1.189 seconds
hive> select * from temperature_data
> ;
OK
Time taken: 2.793 seconds
hive> show tables;
OK
buck_users
json_table
locations
store_view_1
temperature_data
train_sales_stage
users
users partitioned
Time taken: 0.09 seconds, Fetched: 8 row(s)
hive>
```

d) load dataset → load data local inpath “/user/acagild/dataset-session 14.txt” into table temperature_data;

```
acagild@localhost:~/install/hive/apache-hive-2.3.2-bin/conf
Time taken: 0.168 seconds
hive> LOAD DATA LOCAL INPATH "Display all 574 possibilities? (y or n)"
hive> LOAD DATA LOCAL INPATH "/user/acagild/dataset_Session 14.txt" INTO TABLE temperature_data;
FAILED: SemanticException Line 1:23 Invalid path ""/user/acagild/dataset_Session 14.txt"": No files m
atching path file:/user/acagild/dataset_Session14.txt
hive> LOAD DATA LOCAL INPATH "/home/acagild/dataset_Session 14.txt" INTO TABLE temperature_data;
Loading data to table custom.temperature_data
OK
Time taken: 1.211 seconds
hive> dfs -ls /user/hive/warehouse;
Found 11 items
drwxrwxrwx - acagild supergroup 0 2018-02-24 17:45 /user/hive/warehouse/acagilddb.db
drwxrwxrwx - acagild supergroup 0 2018-02-17 15:12 /user/hive/warehouse/advance.db
drwxrwxrwx - acagild supergroup 0 2018-02-17 15:15 /user/hive/warehouse/advanced.db
drwxrwxrwx - acagild supergroup 0 2018-02-17 16:28 /user/hive/warehouse/buck_users
drwxrwxrwx - acagild supergroup 0 2018-03-28 21:05 /user/hive/warehouse/custom.db
drwxrwxrwx - acagild supergroup 0 2018-02-18 15:06 /user/hive/warehouse/json_table
drwxrwxrwx - acagild supergroup 0 2018-02-17 15:46 /user/hive/warehouse/locations
drwxrwxrwx - acagild supergroup 0 2018-02-11 16:17 /user/hive/warehouse/retail.db
drwxrwxrwx - acagild supergroup 0 2018-03-28 21:26 /user/hive/warehouse/temperature_data
drwxrwxrwx - acagild supergroup 0 2018-03-28 20:04 /user/hive/warehouse/train_sales_stage
drwxrwxrwx - acagild supergroup 0 2018-02-17 15:46 /user/hive/warehouse/users
hive> dfs -ls /user/hive/warehouse/custom.db;
Found 1 items
drwxrwxrwx - acagild supergroup 0 2018-03-28 21:55 /user/hive/warehouse/custom.db/temperat
ure_data
hive> dfs -ls /user/hive/warehouse/custom.db/temperature_data;
Found 1 items
-rwxrwxrwx 1 acagild supergroup 437 2018-03-28 21:55 /user/hive/warehouse/custom.db/temperat
ure_data/dataset_Session 14.txt
hive> select * from temperature_data limit 10;
OK
10-01-1990 123112 10
14-02-1991 283901 11
10-03-1990 381920 15
10-01-1991 302918 22
12-02-1990 384902 9
10-01-1991 123112 11
14-02-1990 283901 12
10-03-1991 381920 16
10-01-1990 302918 23
12-02-1991 384902 10
Time taken: 0.366 seconds, Fetched: 10 row(s)
hive>
```

Task2

- a) Select * from temperature_data where zipcode>300000 and zipcode<399999

```
acsdgild@localhost:~/install/hive/apache-hive-2.3.2-bin/conf
at org.apache.hadoop.hive.q1.parse.ParseDriver.parse(ParseDriver.java:208)
at org.apache.hadoop.hive.q1.parse.ParseUtils.parse(ParseUtils.java:77)
at org.apache.hadoop.hive.q1.parse.ParseUtils.parse(ParseUtils.java:70)
at org.apache.hadoop.hive.q1.Driver.compile(Driver.java:468)
at org.apache.hadoop.hive.q1.Driver.compileInternal(Driver.java:1317)
at org.apache.hadoop.hive.q1.Driver.runInternal(Driver.java:1457)
at org.apache.hadoop.hive.q1.Driver.run(Driver.java:1237)
at org.apache.hadoop.hive.q1.Driver.run(Driver.java:1227)
at org.apache.hadoop.hive.cli.CliDriver.processLocalCmd(CliDriver.java:233)
at org.apache.hadoop.hive.cli.CliDriver.processCmd(CliDriver.java:184)
at org.apache.hadoop.hive.cli.CliDriver.processLine(CliDriver.java:403)
at org.apache.hadoop.hive.cli.CliDriver.executeDriver(CliDriver.java:821)
at org.apache.hadoop.hive.cli.CliDriver.run(CliDriver.java:759)
at org.apache.hadoop.hive.cli.CliDriver.main(CliDriver.java:686)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:498)
at org.apache.hadoop.util.RunJar.run(RunJar.java:239)
at org.apache.hadoop.util.RunJar.main(RunJar.java:153)
FAILED: ParseException line 1:56 cannot recognize input near '<' '399999' '<EOF>' in expression specification
hive>
> select * from temperature_data where zipcode>300000 and zipcode<399999;
OK
10-03-1990      381920      15
10-01-1991      302918      22
12-02-1990      384902      9
10-03-1991      381920      16
10-01-1990      302918      23
12-02-1991      384902      10
10-03-1993      381920      16
10-01-1994      302918      23
12-02-1991      384902      10
10-03-1991      381920      16
10-01-1990      302918      23
12-02-1991      384902      10
Time taken: 0.311 seconds, Fetched: 12 row(s)
hive>
>
>
>
>
```