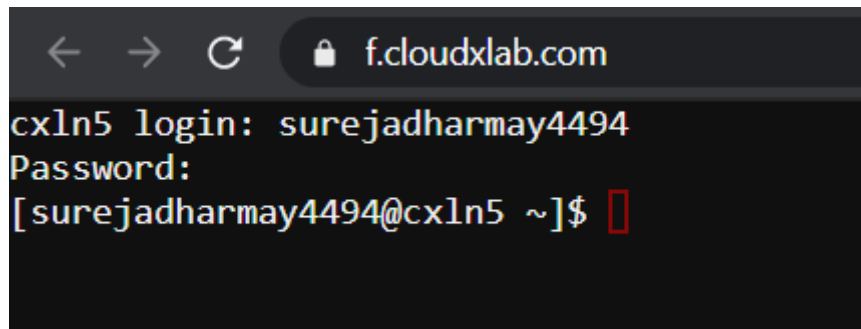


## Practical-12

### Apache Hive

#### STEPS :

- Login to the web console using your CloudXLab Username & password .



- Launch Hive by Typing **hive** in the Web console

```
[surejadharmay4494@cxln5 ~]$ hive
log4j:WARN No such property [maxFileSize] in org.apache.log4j.DailyRollingFileAppender.

Logging initialized using configuration in file:/etc/hive/2.6.2.0-205/0/hive-log4j.properties
hive>
```

- TO see the list of all databases type command .

```
hive> show databases;
OK
01shobha6634
10angshumanbiswas8246
116george5662
14sep
14sep20
19karele9500
90skidshopping2729
9883joy
aaamudha6332
aaayntemp
aakash78959060
aakashyadav1120007745
aamitkumar20131070
aankitsaxena903903
aanzagira3600
aartimohapatra834410
show databases ;
```

- To see the list of all tables type command .  
`show tables ;`

```
hive> show tables;
OK
19_march
a
aac
ab
abhig
abhinav_quick_testing
abhinav_test
ack
ack1
ad_data
ad_data1
ad_data_combined
ad_data_combined_text
ad_emp
adclicks
adclicks1
adclicks2
adclicks_06
adclicks_6
age_1
age_2
age_3
age_4
age_5
airports_pavan
airports_pavan_bucket
airports_pavan_p_b
alive_events
alive_record
amazon_food_review
amazon_food_review_new1
amazon_food_review_updated
ash
ashpart
automobiles
automobiles_bucket
avg_ratings
```

- To create your own database run below command .  
\${enc:USER} gets replaced by your username automatically  
create database \${env:USER} ;

```
hive> create database ${env:USER} ;  
OK  
Time taken: 0.082 seconds  
hive> █
```

- To see metadata of your database run below commands .  
\${enc:USER} gets replaced by your username automatically  
describe database \${env:USER} ;

```
hive> describe database ${env:USER} ;  
OK  
surejadharmay4494      hdfs://cxln1.c.thelab-240901.internal:8020/apps/hive/warehouse/surejadharmay4494.db  surejadharmay4494  USER  
Time taken: 0.122 seconds, Fetched: 1 row(s)
```

- To use your database run below commands .  
\${enc:USER} gets replaced by your username automatically  
use \${env:USER} ;

```
hive> use ${env:USER};  
OK  
Time taken: 0.28 seconds  
hive> █
```

- To create a table x in your database type commands .  
create table x (a int) ;

```
hive> create table x (a int) ;  
OK  
Time taken: 0.225 seconds  
hive> █
```

- To view the data of table x type commands .

select \* from x ;

```
hive> select * from x ;
OK
Time taken: 0.519 seconds
hive>
```

- To view the metadata(structure) of table type commands .

describe x ;

```
hive> describe x ;
OK
a                                int
Time taken: 0.374 seconds, Fetched: 1 row(s)
hive>
```

- To see the metadata and low-level details type commands .

describe formatted x ;

```
hive> describe formatted x ;
OK
# col_name          data_type          comment
a                   int
# Detailed Table Information
Database:           surejadharmay4494
Owner:              surejadharmay4494
CreateTime:         Thu Nov 12 20:03:27 UTC 2020
LastAccessTime:     UNKNOWN
Protect Mode:       None
Retention:          0
Location:           hdfs://cx1n1.c.thelab-240901.internal:8020/apps/hive/warehouse/surejadharmay4494.db/x
Table Type:         MANAGED_TABLE
Table Parameters:
  COLUMN_STATS_ACCURATE  {\\"BASIC_STATS\\":\\"true\\"}
  numFiles                0
  numRows                 0
  rawDataSize             0
  totalSize               0
  transient_lastDdlTime   1605211407
# Storage Information
SerDe Library:       org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe
InputFormat:         org.apache.hadoop.mapred.TextInputFormat
OutputFormat:        org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat
Compressed:          No
Num Buckets:         -1
Bucket Columns:      []
Sort Columns:        []
Storage Desc Params:
  serialization.format    1
Time taken: 0.378 seconds, Fetched: 31 row(s)
hive>
```

Create a table called **nyse** in hive and load data from your home folder in

The web console

- Login to the web console using your CloudXLab Username & password .
- Copy NYSE data from HDFS to your local.  
**hadoop fs -copyToLocal /data/NYSE\_daily**

```
[surejadharmay4494@cxln5 ~]$ hadoop fs -copyToLocal /data/NYSE_daily
```

- Launch Hive with typing in hive on the console. Run the below command in Hive

```
[surejadharmay4494@cxln5 ~]$ hive
log4j:WARN No such property [maxFileSize] in org.apache.log4j.DailyRollingFileAppender.
Logging initialized using configuration in file:/etc/hive/2.6.2.0-205/0/hive-log4j.properties
hive>
```

- Use your database by using below command .  
\${enc:USER} gets replaced by your username automatically  
use \${env:USER} ;

```
hive> use ${env:USER} ;
OK
Time taken: 3.693 seconds
hive>
```

- Create table nyse using below command

```
CREATE TABLE nyse(
  exchange1 STRING,
  symbol1 STRING,
  ymd STRING,
  price_open FLOAT,
  price_high FLOAT,
  price_low FLOAT,
  price_close FLOAT,
  volume INT,
  price_adj_close FLOAT
)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t';
```

```
hive> CREATE TABLE nyse(  
  >     exchange1 STRING,  
  >     symbol1 STRING,  
  >     ymd STRING,  
  >     price_open FLOAT,  
  >     price_high FLOAT,  
  >     price_low FLOAT,  
  >     price_close FLOAT,  
  >     volume INT,  
  >     price_adj_close FLOAT  
  > )  
  > ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t';  
OK  
Time taken: 0.336 seconds  
hive> █
```

- See metadata of table using below command  
DESCRIBE nyse ;

```
hive> DESCRIBE nyse ;  
OK  
exchange1          string  
symbol1            string  
ymd                 string  
price_open          float  
price_high          float  
price_low           float  
price_close         float  
volume              int  
price_adj_close     float  
Time taken: 0.374 seconds, Fetched: 9 row(s)  
hive> █
```

- To See more low-level details , type below command .

**DESCRIBE FORMATTED nyse ;**

```
hive> DESCRIBE FORMATTED nyse ;
OK
# col_name          data_type          comment
exchange1          string
symbol1            string
ymd                 string
price_open          float
price_high          float
price_low           float
price_close         float
volume              int
price_adj_close     float

# Detailed Table Information
Database:           surejadharmay4494
Owner:              surejadharmay4494
CreateTime:         Thu Nov 12 20:06:20 UTC 2020
LastAccessTime:     UNKNOWN
Protect Mode:       None
Retention:          0
Location:           hdfs://cxln1.c.thelab-240901.internal:8020/apps/hive/warehouse/surejadharmay4494.db/nyse
Table Type:         MANAGED_TABLE
Table Parameters:
  COLUMN_STATS_ACCURATE {\"BASIC_STATS\": \"true\"}
  numFiles              0
  numRows              0
  rawDataSize          0
  totalSize            0
  transient_lastDdlTime 1605211580

# Storage Information
SerDe Library:      org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe
InputFormat:        org.apache.hadoop.mapred.TextInputFormat
OutputFormat:       org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat
Compressed:         No
Num Buckets:        -1
Bucket Columns:     []
Sort Columns:       []
Storage Desc Params:
  field.delim         \t
  serialization.format \t

Time taken: 0.409 seconds, Fetched: 40 row(s)
hive>
```

- Load data to your nyse table  
use `${env:USER}` ;  
load data local inpath ' NYSE\_daily ' overwrite into table nyse ;

```
hive> load data local inpath 'NYSE_daily' overwrite into table nyse;
Loading data to table surejadharmay4494.nyse
Table surejadharmay4494.nyse stats: [numFiles=1, numRows=0, totalSize=3194099, rawDataSize=0]
OK
Time taken: 1.081 seconds
hive>
```

- Select rows from the table ( in HIVE console ) .  
select \* from nyse ;

```
hive> select * from nyse;
OK
NYSE    CLI      2009-12-31      35.39    35.7      34.5      34.57    890100    34.12
NYSE    CLI      2009-12-30      35.22    35.46    34.96    35.4     516900    34.94
NYSE    CLI      2009-12-29      35.69    35.95    35.21    35.34    556500    34.88
NYSE    CLI      2009-12-28      35.67    36.23    35.49    35.69    565000    35.23
NYSE    CLI      2009-12-24      35.38    35.6     35.19    35.47    230200    35.01
NYSE    CLI      2009-12-23      35.13    35.51    35.07    35.21    520200    34.75
NYSE    CLI      2009-12-22      34.76    35.04    34.71    35.04    564600    34.58
NYSE    CLI      2009-12-21      34.65    34.74    34.41    34.73    428400    34.28
NYSE    CLI      2009-12-18      34.11    34.38    33.73    34.22    1152600   33.77
NYSE    CLI      2009-12-17      34.18    34.53    33.84    34.21    1082600   33.76
NYSE    CLI      2009-12-16      34.79    35.1     34.48    34.66    1007900   34.21
NYSE    CLI      2009-12-15      34.6     34.91    34.39    34.84    813200    34.39
NYSE    CLI      2009-12-14      34.21    34.9     33.86    34.82    987700    34.37
NYSE    CLI      2009-12-11      33.55    34.08    33.4     34.0     836500    33.56
NYSE    CLI      2009-12-10      33.61    33.8     33.09    33.26    1296300   32.83
NYSE    CLI      2009-12-09      33.25    33.71    33.01    33.25    863900    32.82
NYSE    CLI      2009-12-08      32.57    33.8     32.52    33.18    890000    32.75
NYSE    CLI      2009-12-07      33.63    33.77    32.78    33.0     961800    32.57
NYSE    CLI      2009-12-04      33.46    34.12    33.16    33.69    1412400   33.25
NYSE    CLI      2009-12-03      32.55    33.43    32.48    32.65    1170600   32.22
NYSE    CLI      2009-12-02      31.72    32.62    31.41    32.47    975400    32.05
NYSE    CLI      2009-12-01      31.14    31.74    30.79    31.69    1284400   31.28
NYSE    CLI      2009-11-30      29.59    30.86    29.29    30.69    1411200   30.29
NYSE    CLI      2009-11-27      29.48    30.19    29.24    29.47    453600    29.09
NYSE    CLI      2009-11-25      30.38    30.58    30.14    30.4     574500    30.0
NYSE    CLI      2009-11-24      30.48    30.56    30.04    30.15    685900    29.76
NYSE    CLI      2009-11-23      31.06    31.42    30.35    30.57    1541600   30.17
NYSE    CLI      2009-11-20      31.1     31.15    30.39    30.54    1721700   30.14
NYSE    CLI      2009-11-19      31.73    31.79    31.26    31.29    861800    30.88
NYSE    CLI      2009-11-18      31.74    32.26    31.56    32.16    715000    31.74
NYSE    CLI      2009-11-17      32.07    32.59    31.51    31.66    936400    31.25
NYSE    CLI      2009-11-16      32.18    32.73    32.0     32.31    798600    31.89
NYSE    CLI      2009-11-13      31.38    31.82    31.24    31.8     556100    31.39
NYSE    CLI      2009-11-12      31.67    31.73    31.1     31.15    599900    30.74
NYSE    CLI      2009-11-11      31.43    31.75    31.0     31.69    689800    31.28
NYSE    CLI      2009-11-10      31.3     31.44    30.58    30.97    593500    30.57
NYSE    CLI      2009-11-09      30.21    31.59    30.0     31.52    1105300   31.11
```

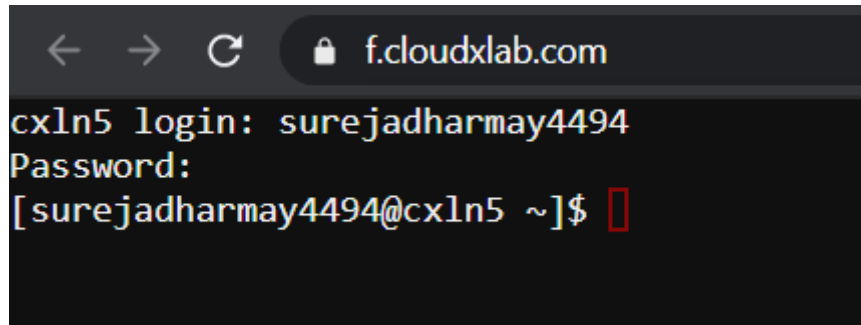


NYSE	CVA	2009-02-27	15.58	15.67	14.29	15.23	3310700	15.23
NYSE	CVA	2009-02-26	17.47	17.64	16.49	16.67	1940500	16.67
NYSE	CVA	2009-02-25	17.47	17.73	17.11	17.4	1550400	17.4
NYSE	CVA	2009-02-24	17.91	18.23	17.2	17.43	2278100	17.43
NYSE	CVA	2009-02-23	17.92	18.1	17.52	17.9	3036200	17.9
NYSE	CVA	2009-02-20	18.25	18.26	17.29	17.89	1650500	17.89
NYSE	CVA	2009-02-19	18.47	19.0	18.29	18.37	761100	18.37
NYSE	CVA	2009-02-18	18.42	19.21	18.4	18.6	1536900	18.6
NYSE	CVA	2009-02-17	18.42	18.57	18.12	18.3	1215800	18.3
NYSE	CVA	2009-02-13	19.15	19.48	18.9	18.99	470900	18.99
NYSE	CVA	2009-02-12	18.64	19.28	18.45	19.28	870200	19.28
NYSE	CVA	2009-02-11	19.0	19.38	18.74	18.99	784700	18.99
NYSE	CVA	2009-02-10	19.5	19.78	18.85	19.0	1520700	19.0
NYSE	CVA	2009-02-09	19.01	19.71	18.86	19.65	1203900	19.65
NYSE	CVA	2009-02-06	18.63	19.1	18.59	18.89	1004300	18.89
NYSE	CVA	2009-02-05	18.19	18.83	17.95	18.63	1013500	18.63
NYSE	CVA	2009-02-04	18.42	18.81	18.38	18.43	1141500	18.43
NYSE	CVA	2009-02-03	18.61	18.82	18.32	18.59	1192600	18.59
NYSE	CVA	2009-02-02	17.74	18.71	17.64	18.45	2453600	18.45
NYSE	CVA	2009-01-30	17.67	17.96	17.1	17.29	1672500	17.29
NYSE	CVA	2009-01-29	17.59	18.04	17.51	17.81	1228000	17.81
NYSE	CVA	2009-01-28	18.22	18.34	17.53	17.83	1739300	17.83
NYSE	CVA	2009-01-27	18.12	18.25	17.52	17.83	1360900	17.83
NYSE	CVA	2009-01-26	18.49	18.84	17.95	18.12	1340900	18.12
NYSE	CVA	2009-01-23	18.29	18.55	17.87	18.36	1258500	18.36
NYSE	CVA	2009-01-22	18.88	18.9	18.29	18.54	1623300	18.54
NYSE	CVA	2009-01-21	19.49	19.83	18.75	19.22	1598400	19.22
NYSE	CVA	2009-01-20	20.92	21.04	19.31	19.48	1435400	19.48
NYSE	CVA	2009-01-16	20.76	21.06	20.19	20.93	966600	20.93
NYSE	CVA	2009-01-15	19.4	20.75	19.3	20.43	1063900	20.43
NYSE	CVA	2009-01-14	19.69	19.95	19.0	19.64	1744100	19.64
NYSE	CVA	2009-01-13	20.11	20.11	19.31	19.65	1261200	19.65
NYSE	CVA	2009-01-12	20.46	20.88	19.91	20.24	700800	20.24
NYSE	CVA	2009-01-09	21.08	21.63	20.3	20.75	723800	20.75
NYSE	CVA	2009-01-08	21.08	21.25	20.43	20.81	1633700	20.81
NYSE	CVA	2009-01-07	22.34	22.5	21.22	21.32	1169500	21.32
NYSE	CVA	2009-01-06	22.85	22.92	22.25	22.5	1058700	22.5
NYSE	CVA	2009-01-05	21.79	22.89	21.49	22.68	981100	22.68
NYSE	CVA	2009-01-02	21.76	22.8	21.46	21.71	1344600	21.71

Time taken: 0.349 seconds, Fetched: 57391 row(s)

### Extra Practice :

- Login to the web console using your CloudXLab Username & password .



- Copy NYSE data from HDFS to your local.

**hadoop fs -cp /data/NYSE\_daily**

```
[surejadharmay4494@cxln5 ~]$ hadoop fs -cp /data/NYSE_daily
```

- Launch Hive with typing in hive on console .

```
[surejadharmay4494@cxln5 ~]$ hive
log4j:WARN No such property [maxFileSize] in org.apache.log4j.DailyRollingFileAppender.
Logging initialized using configuration in file:/etc/hive/2.6.2.0-205/0/hive-log4j.properties
hive>
```

- Use your own database by using below command .  
\${enc:USER} gets replaced by your username automatically  
use \${env:USER} ;

```
hive> use ${env:USER} ;
OK
Time taken: 3.693 seconds
hive>
```

- Create table nyse using below command

```
CREATE TABLE nyse_external(  
  exchange1 STRING,  
  symbol1 STRING,  
  ymd STRING,  
  price_open FLOAT,  
  price_high FLOAT,  
  price_low FLOAT,  
  price_close FLOAT,  
  volume INT,  
  price_adj_close FLOAT  
)  
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t';  
LOCATION '/user/${env:USER}/NYSE_daily' ;
```

```
hive> CREATE EXTERNAL TABLE nyse_external(  
  > exchange1 STRING,  
  > symbol1 STRING,  
  > ymd STRING,  
  > price_open FLOAT,  
  > price_high FLOAT,  
  > price_low FLOAT,  
  > price_close FLOAT,  
  > volume INT,  
  > price_adj_close FLOAT  
  > )  
  > ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t'  
  > LOCATION '/user/${env:USER}/NYSE_daily';  
OK  
Time taken: 2.165 seconds  
hive> █
```

- To See more low-level details , type below command .  
**DESCRIBE FORMATTED nyse\_external ;**

```
hive> DESCRIBE FORMATTED nyse_external ;
OK
# col_name          data_type          comment
exchange1          string
symbol1            string
ymd                 string
price_open          float
price_high          float
price_low           float
price_close         float
volume              int
price_adj_close     float

# Detailed Table Information
Database:           surejadharmay4494
Owner:              surejadharmay4494
CreateTime:         Thu Nov 12 20:12:10 UTC 2020
LastAccessTime:     UNKNOWN
Protect Mode:       None
Retention:          0
Location:           hdfs://cxln1.c.thelab-240901.internal:8020/apps/hive/warehouse/surejadharmay4494.db/nyse_external
Table Type:         MANAGED_TABLE
Table Parameters:
  COLUMN_STATS_ACCURATE {\"BASIC_STATS\": \"true\"}
  numFiles              0
  numRows              0
  rawDataSize          0
  totalSize            0
  transient_lastDdlTime 1605211930

# Storage Information
SerDe Library:      org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe
InputFormat:        org.apache.hadoop.mapred.TextInputFormat
OutputFormat:       org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat
Compressed:         No
Num Buckets:        -1
Bucket Columns:     []
Sort Columns:       []
Storage Desc Params:
  field.delim         \t
```

- Check the warehouse directory in HDFS by typing below command in unix console ( press ctrl+c to exit hive console ) .  
**hadoop fs -ls NYSE\_daily ;**

```
[surejadharmay4494@cxln5 ~]$ hadoop fs -ls NYSE_daily ;
```

- To drop the external table type in unix console  
DROP TABLE nyse\_external ;

```
hive> DROP TABLE nyse_external;  
OK  
Time taken: 2.087 seconds  
hive> █
```