BIG DATA ASSIGNMENT-2

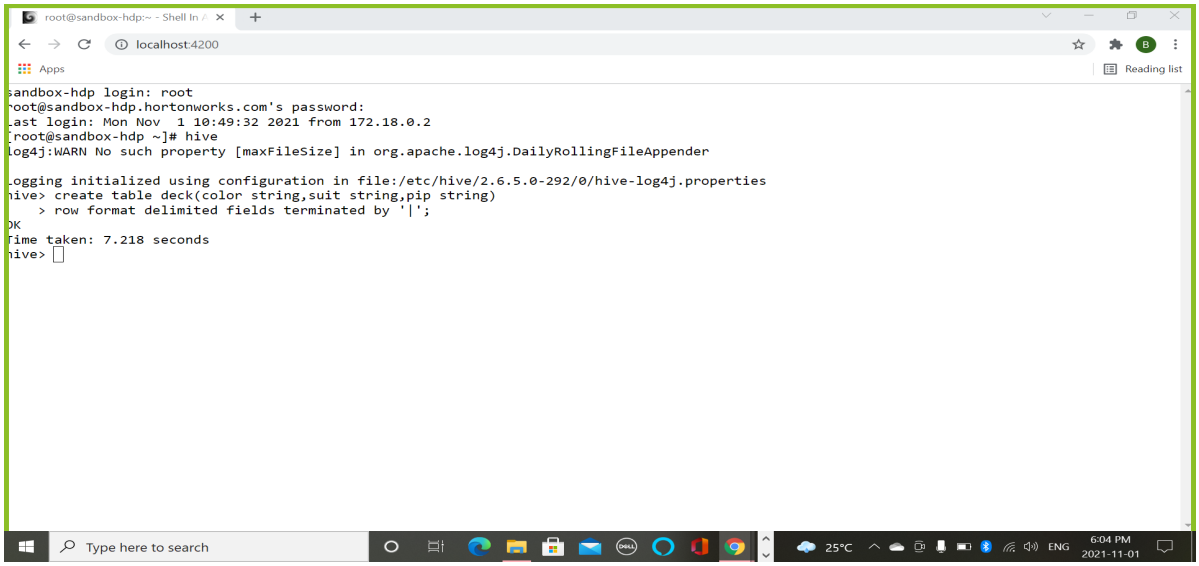
T.Divya Reddy
AP18110010118
CSE-B

Dataset:

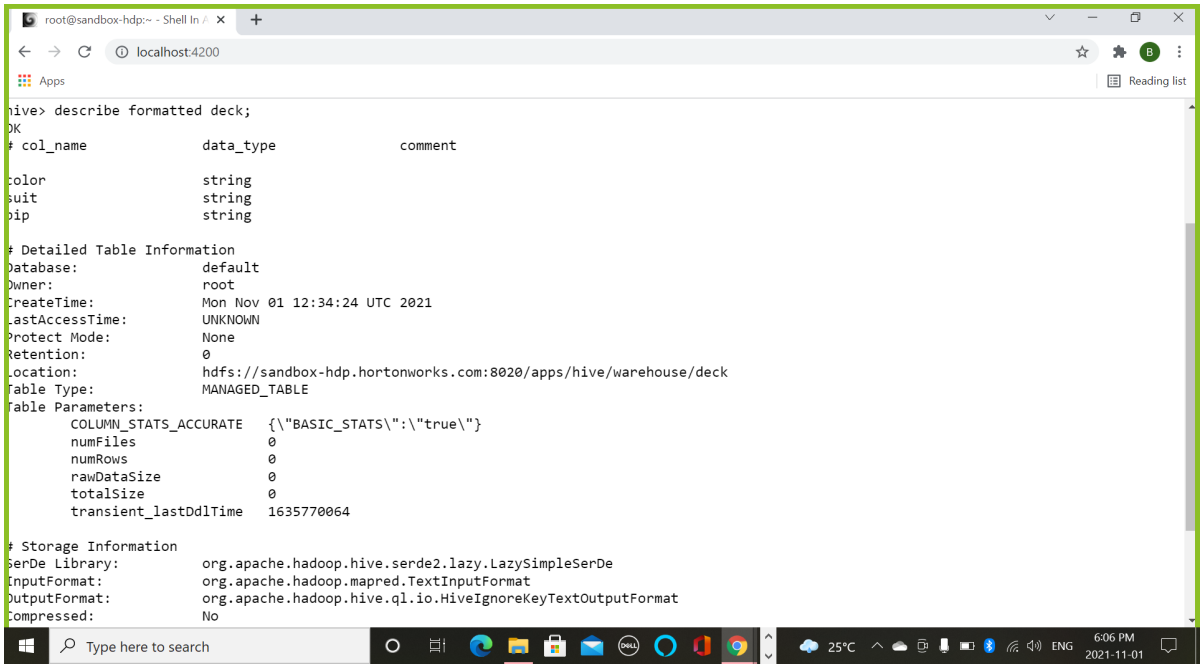
BLACK|SPADE|2
BLACK|SPADE|3
BLACK|SPADE|4
BLACK|SPADE|5
BLACK|SPADE|6
BLACK|SPADE|7
BLACK|SPADE|8
.....

1.) Create a managed table and load the data from LFS

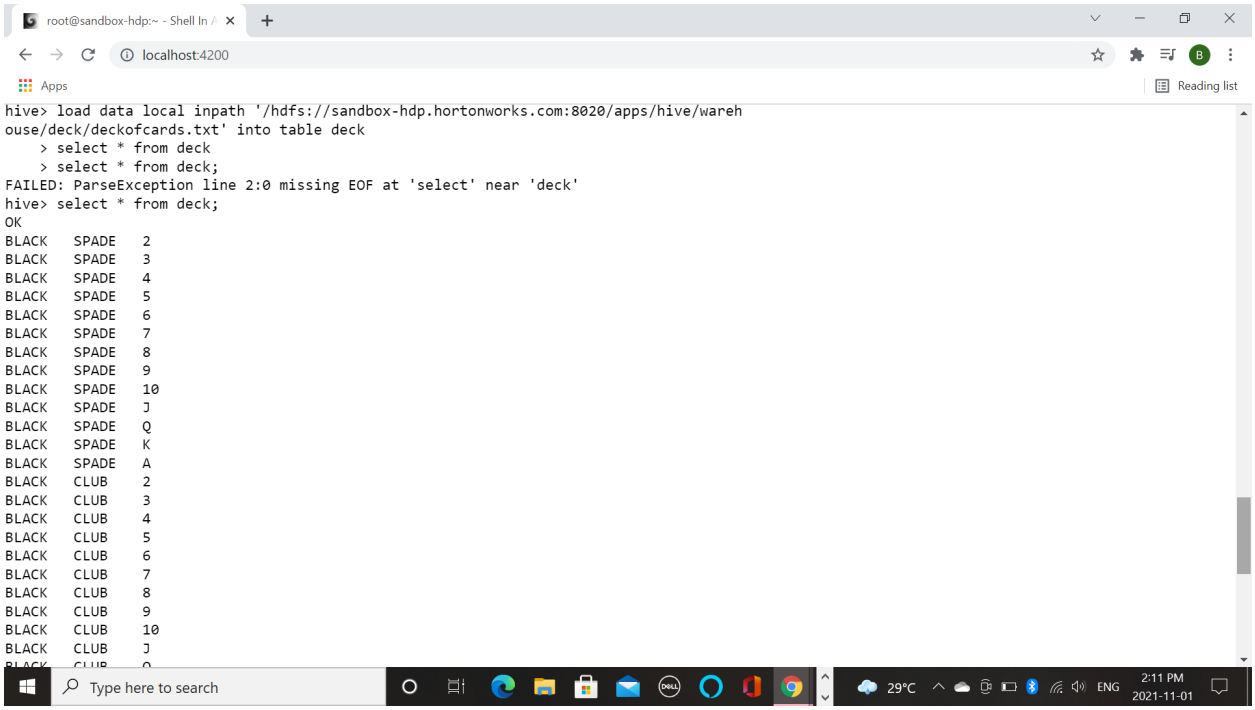
Step-1: Create a managed table deck in Hive having column names as color,suit and pip

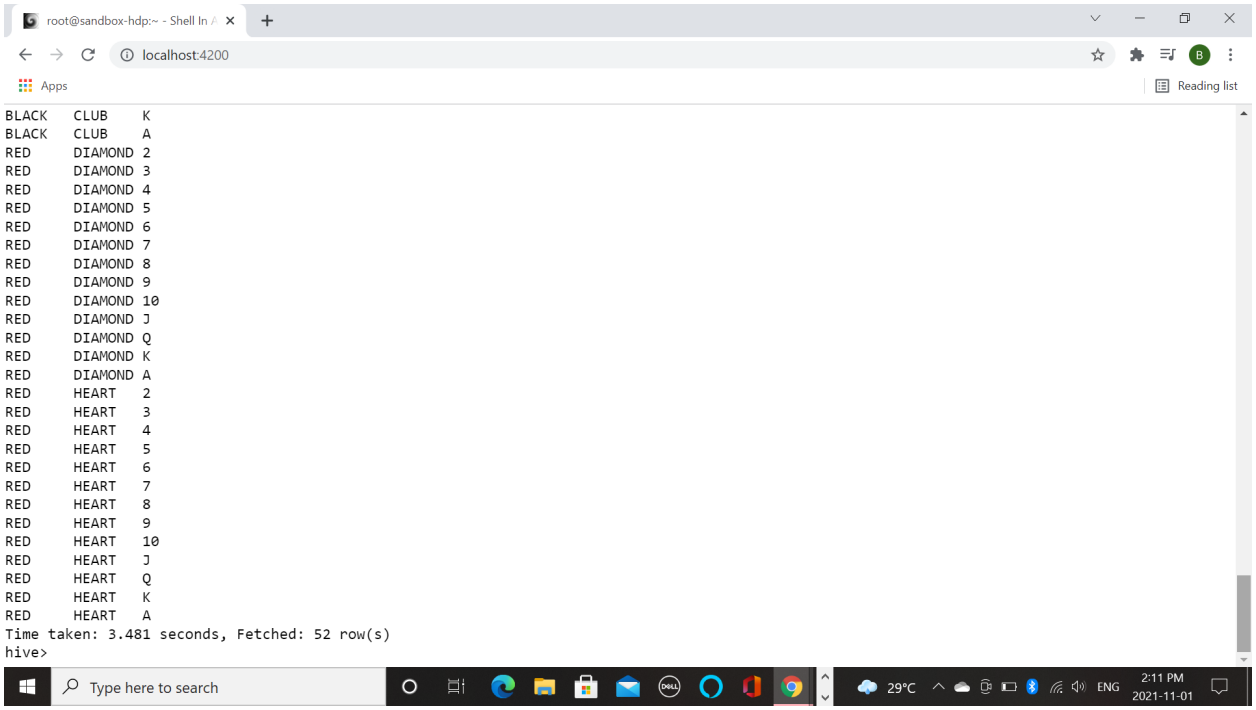


Step-2: Describe the table to know location of it

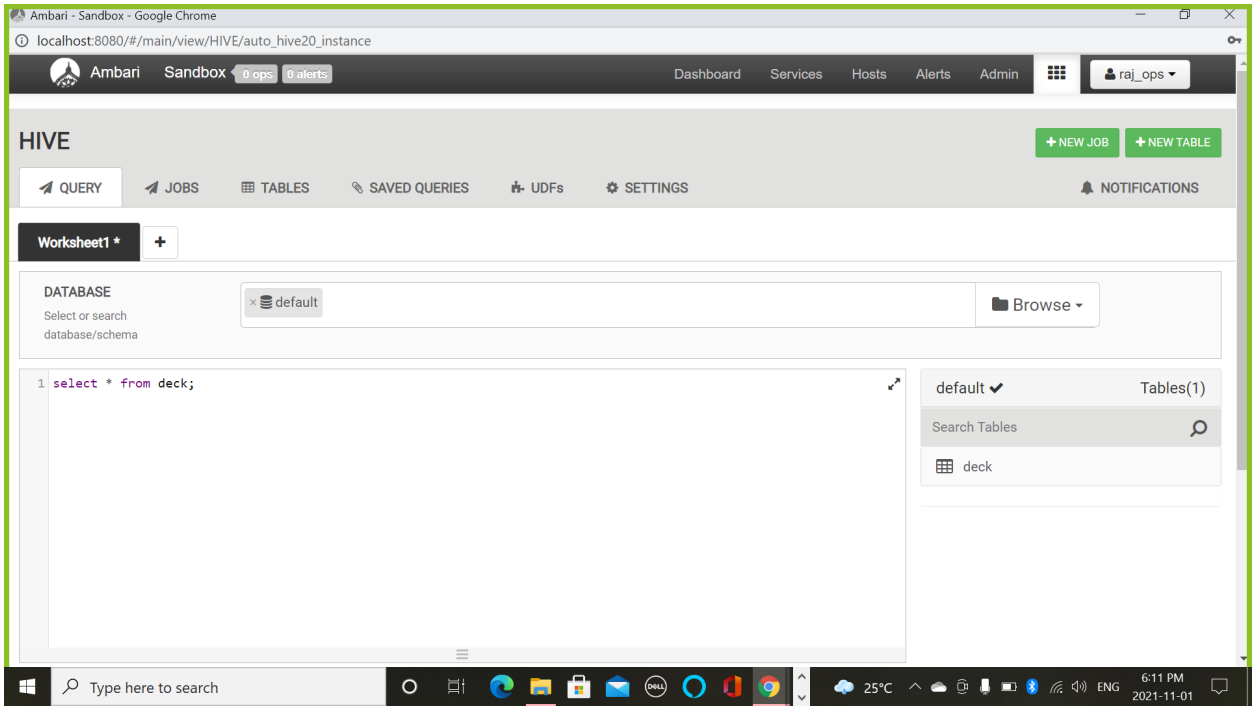


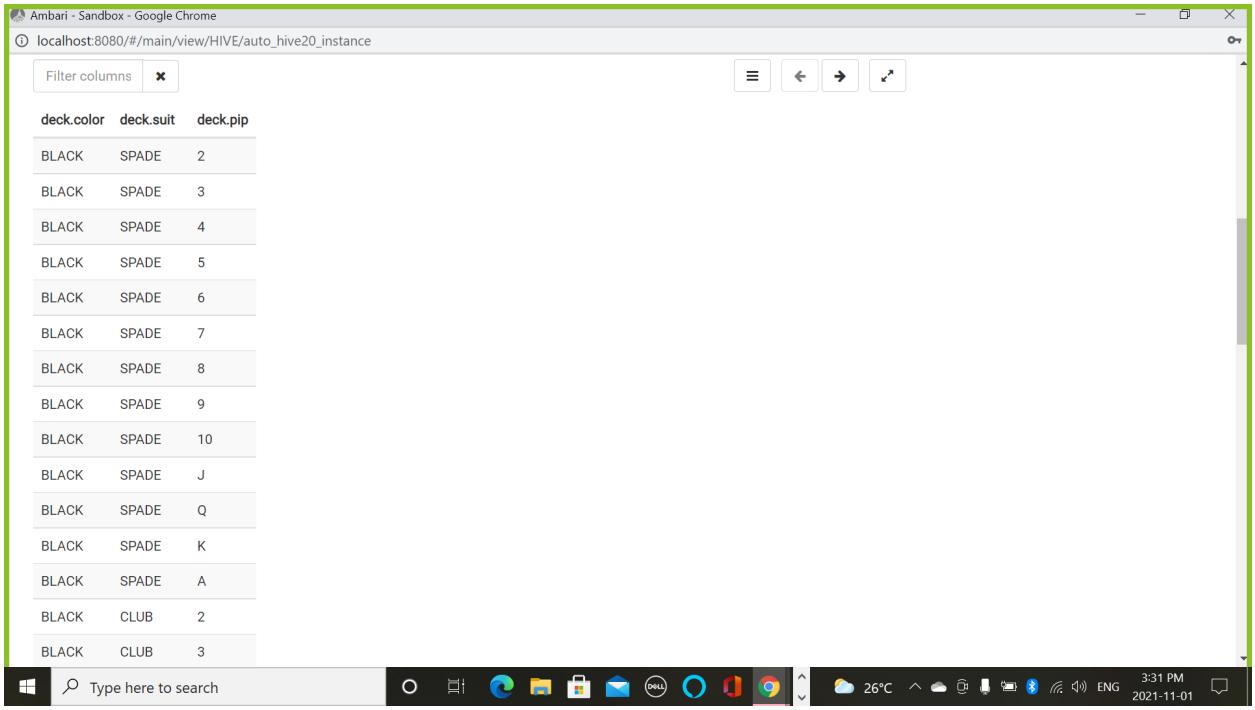
Step-3: Load cards data into the table and display data using select command

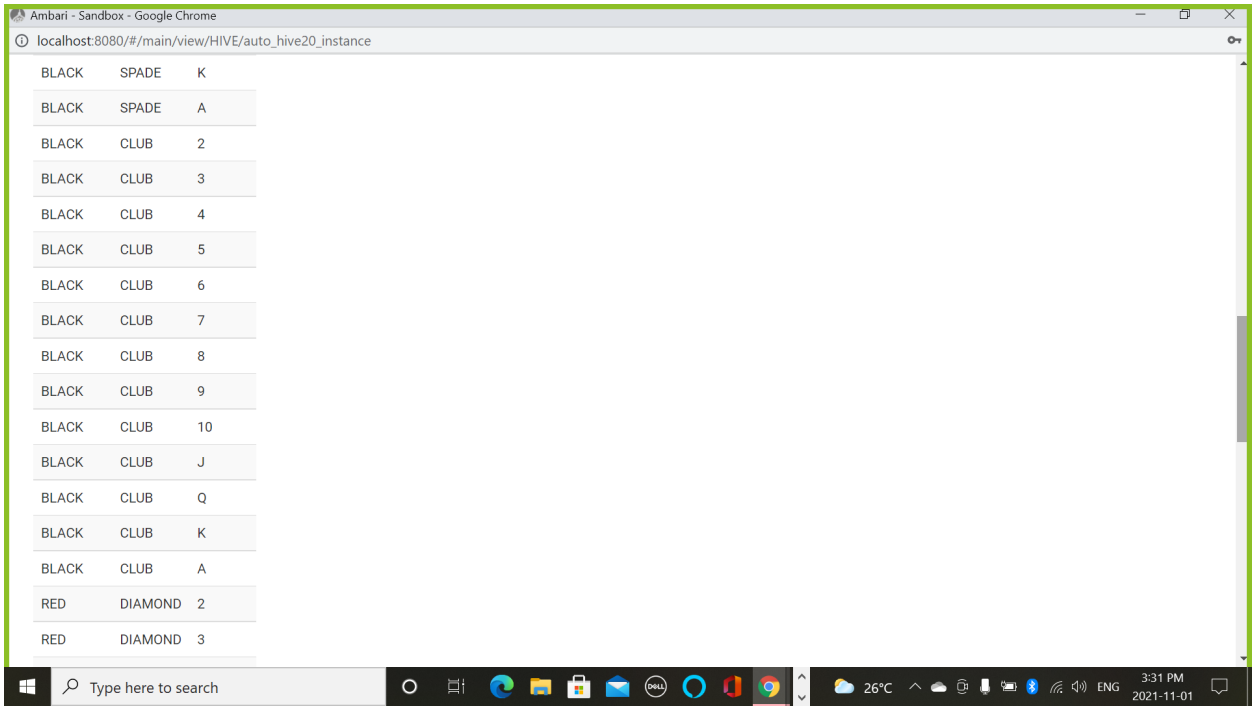




Step-4: Using select command display the table in hive 2.0

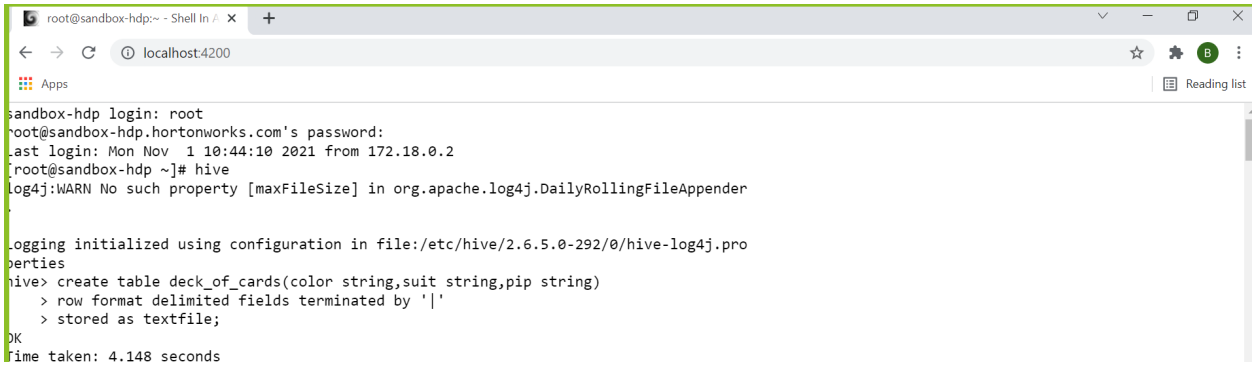




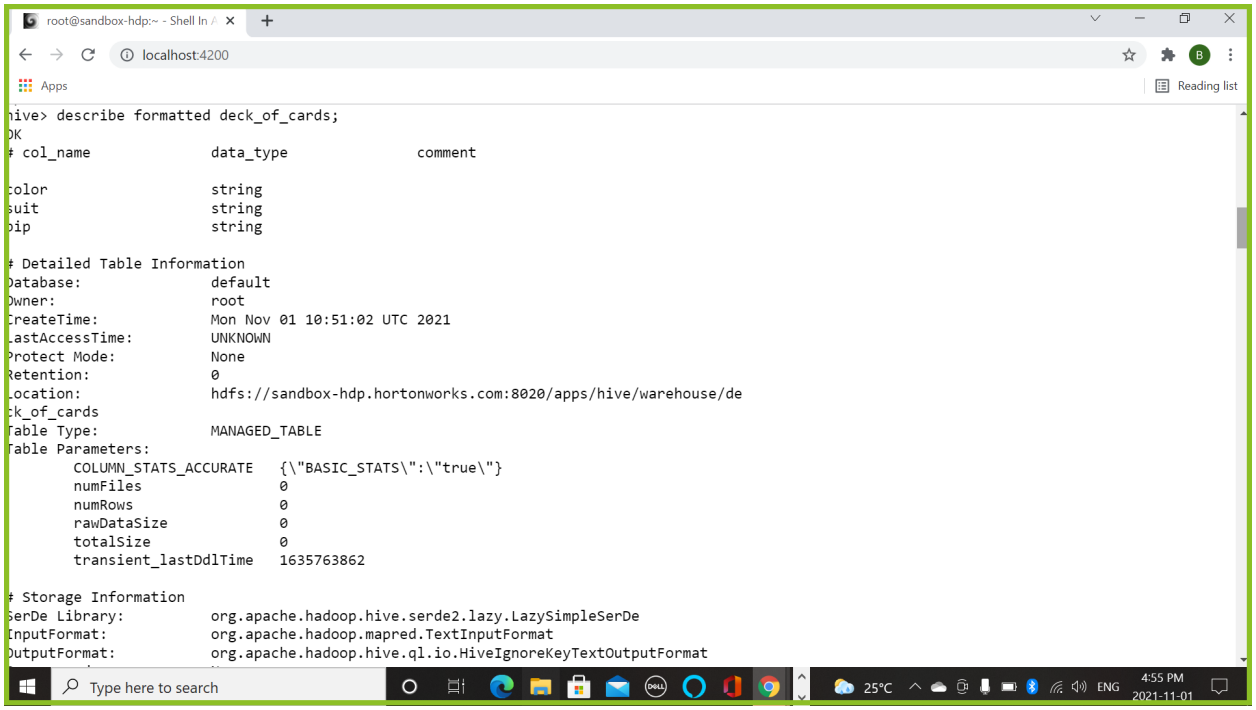


2.) Create a managed table and load the data from HDFS

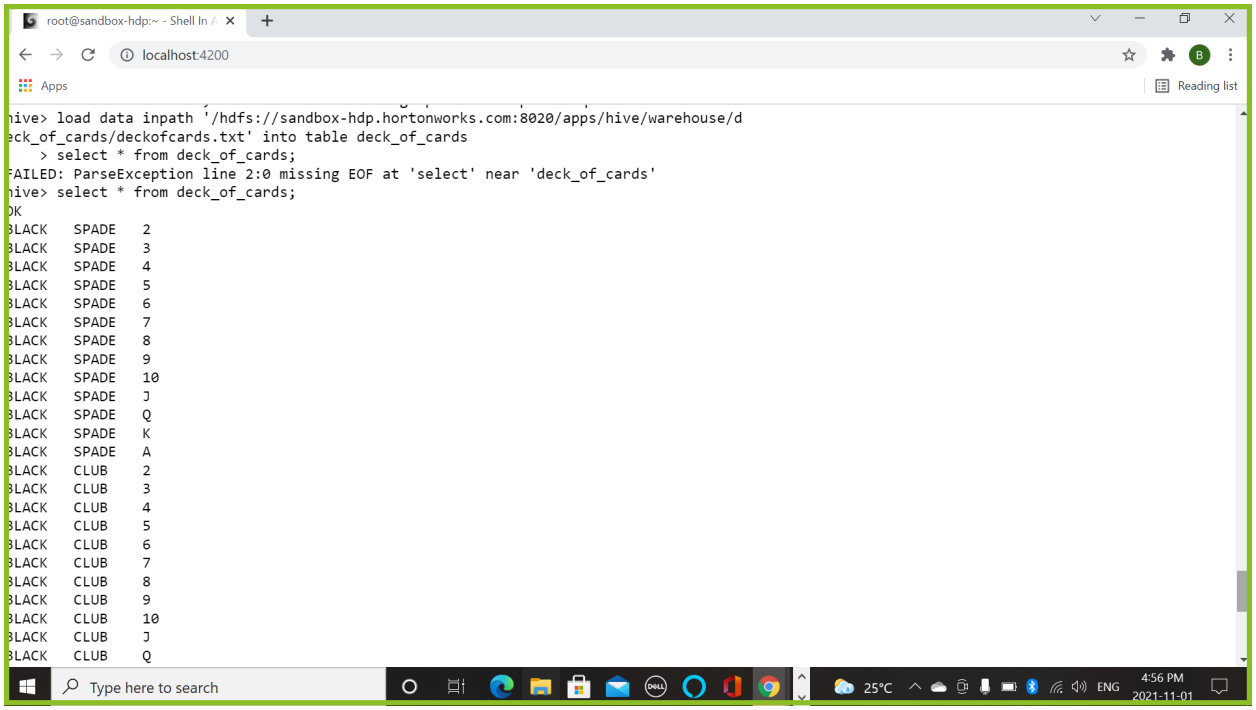
Step-1: Create managed table deck_of_cards in Hive having column names as color,suit and pip

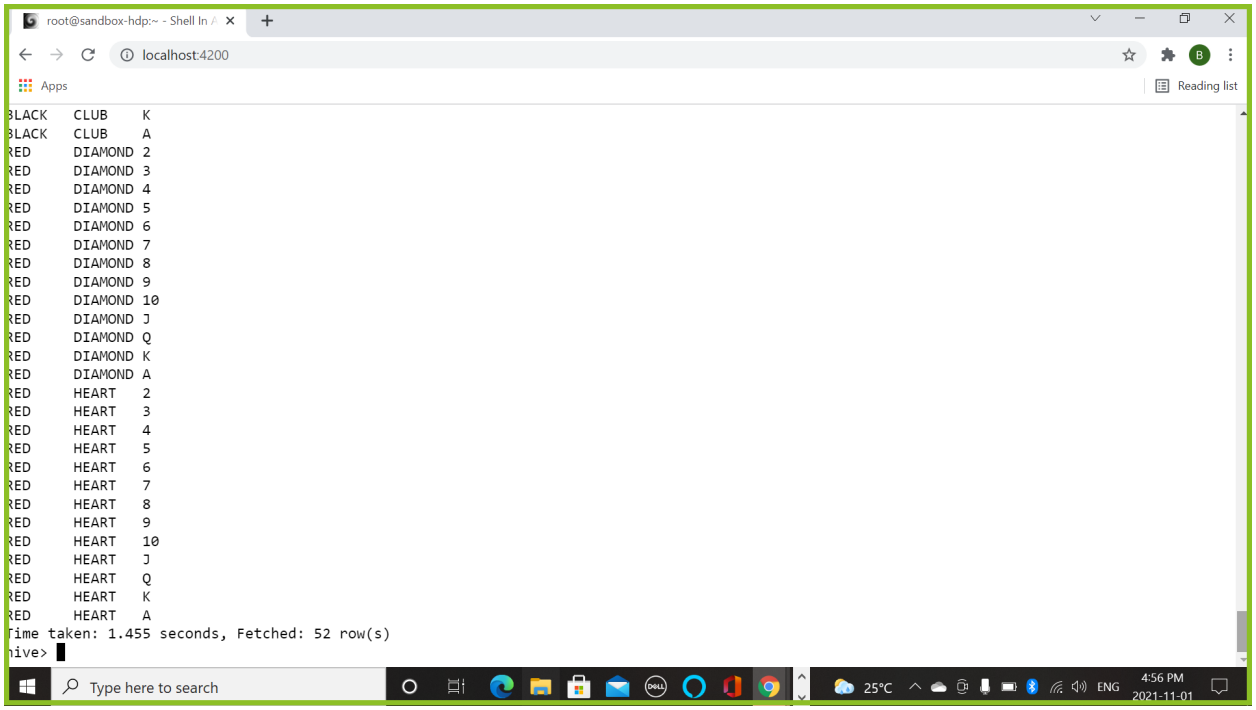


Step-2: Describe table to know location of it

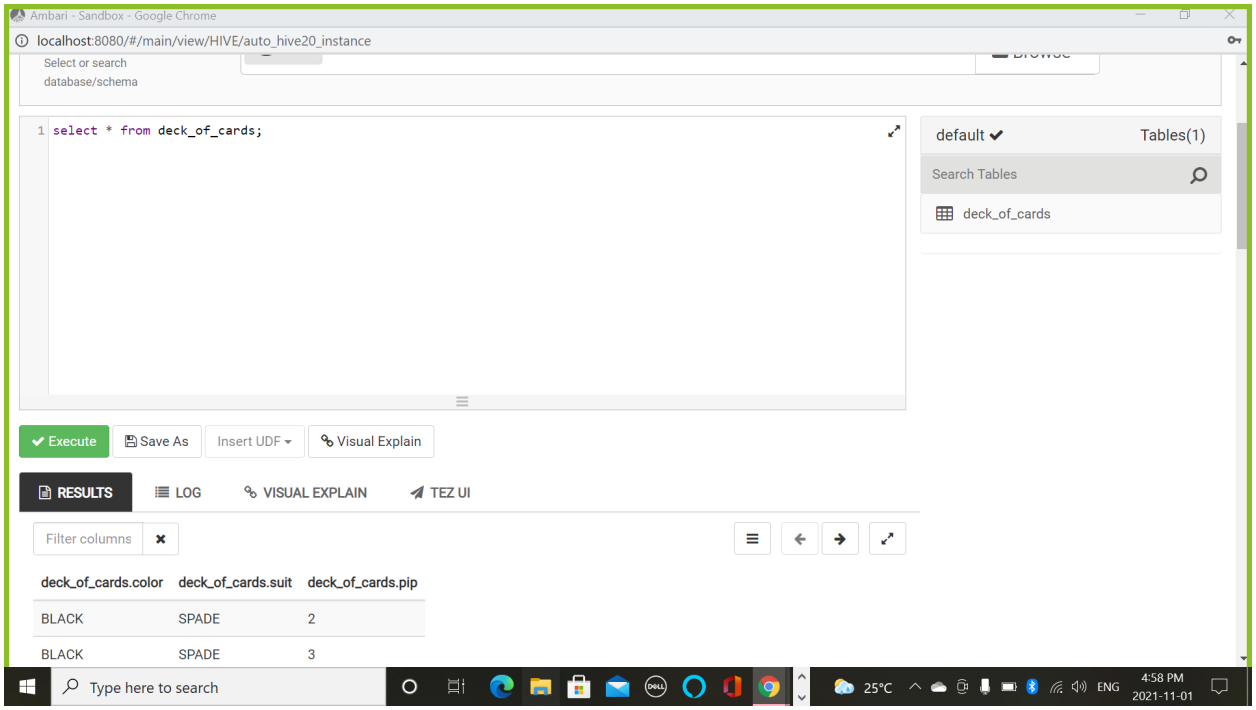


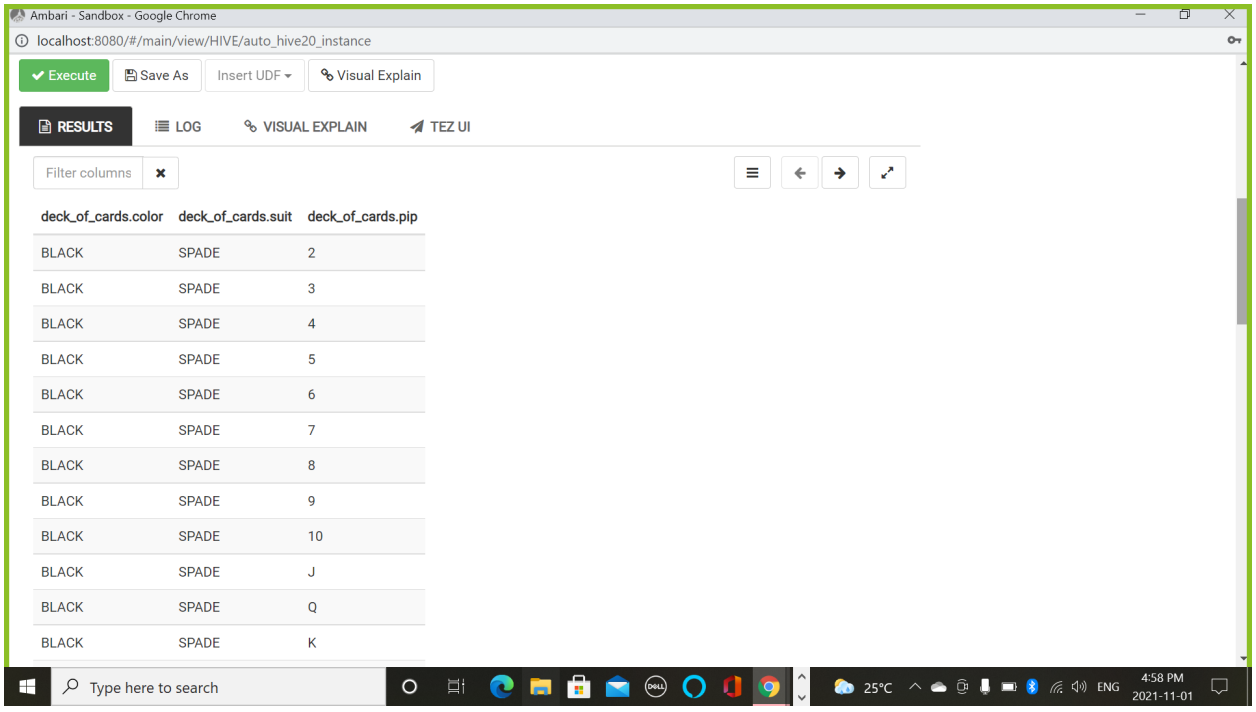
Step-3: Load cards data into table and display it using select command

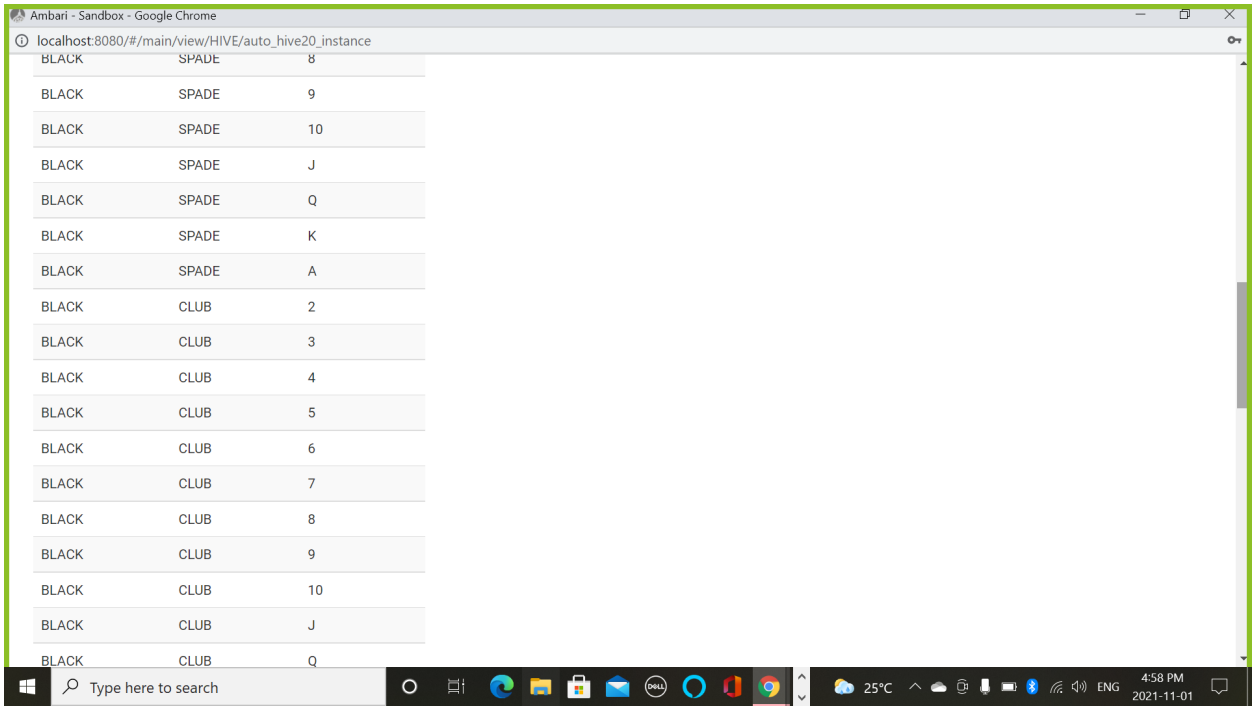




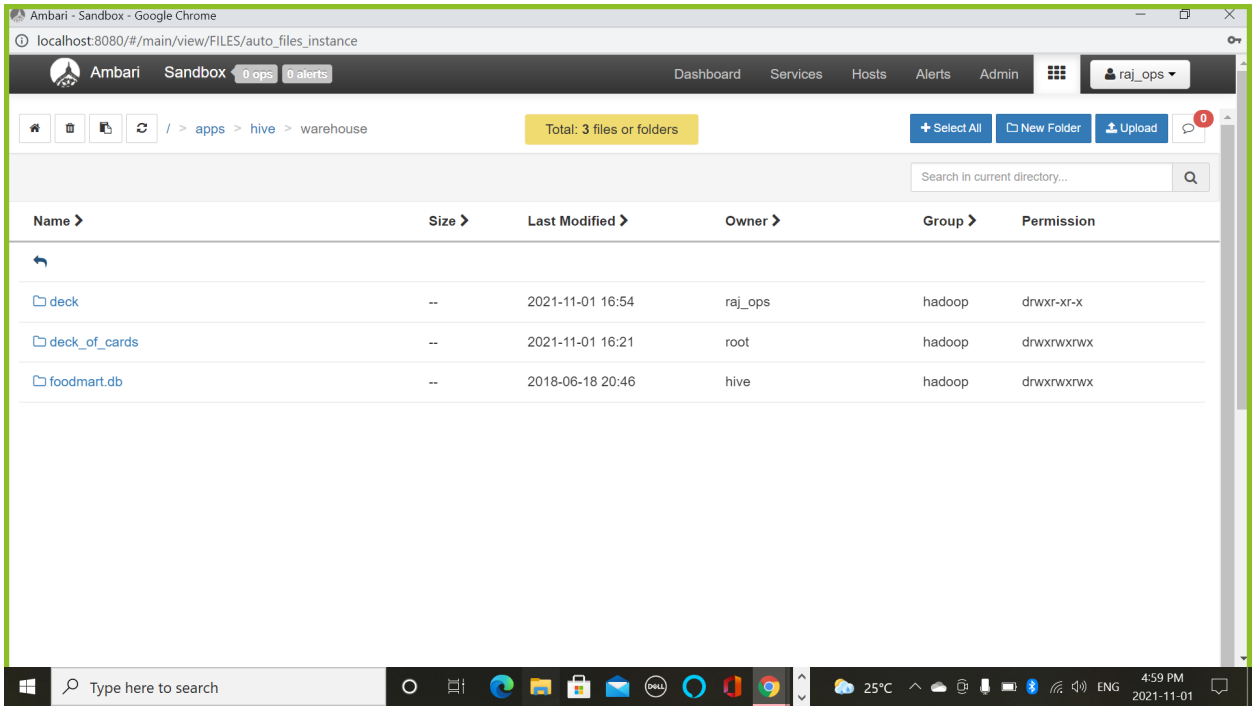
Step-4: Using select command display contents of table in hive 2.0


Ambari - Sandbox - Google Chrome
localhost:8080/#/main/view/HIVE/auto_hive20_instance
Select or search database/schema
1 select * from deck_of_cards;
default
Tables(1)
Search Tables
deck_of_cards
Execute
Save As
Insert UDF
Visual Explain
RESULTS
LOG
VISUAL EXPLAIN
TEZ UI
Filter columns
deck_of_cards.color
deck_of_cards.suit
deck_of_cards.pip
BLACK SPADE 2
BLACK SPADE 3
Type here to search
25°C
ENG
4:58 PM
2021-11-01


Ambari - Sandbox - Google Chrome
localhost:8080/#/main/view/HIVE/auto_hive20_instance
Execute
Save As
Insert UDF
Visual Explain
RESULTS
LOG
VISUAL EXPLAIN
TEZ UI
Filter columns
deck_of_cards.color
deck_of_cards.suit
deck_of_cards.pip
BLACK SPADE 2
BLACK SPADE 3
BLACK SPADE 4
BLACK SPADE 5
BLACK SPADE 6
BLACK SPADE 7
BLACK SPADE 8
BLACK SPADE 9
BLACK SPADE 10
BLACK SPADE J
BLACK SPADE Q
BLACK SPADE K
Type here to search
25°C
ENG
4:58 PM
2021-11-01

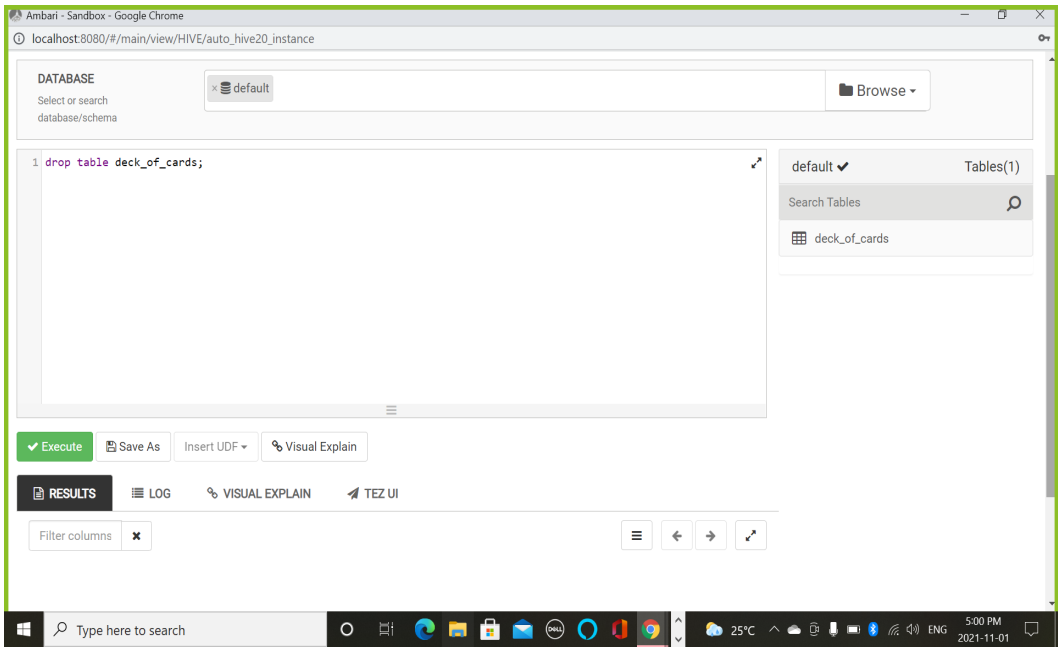


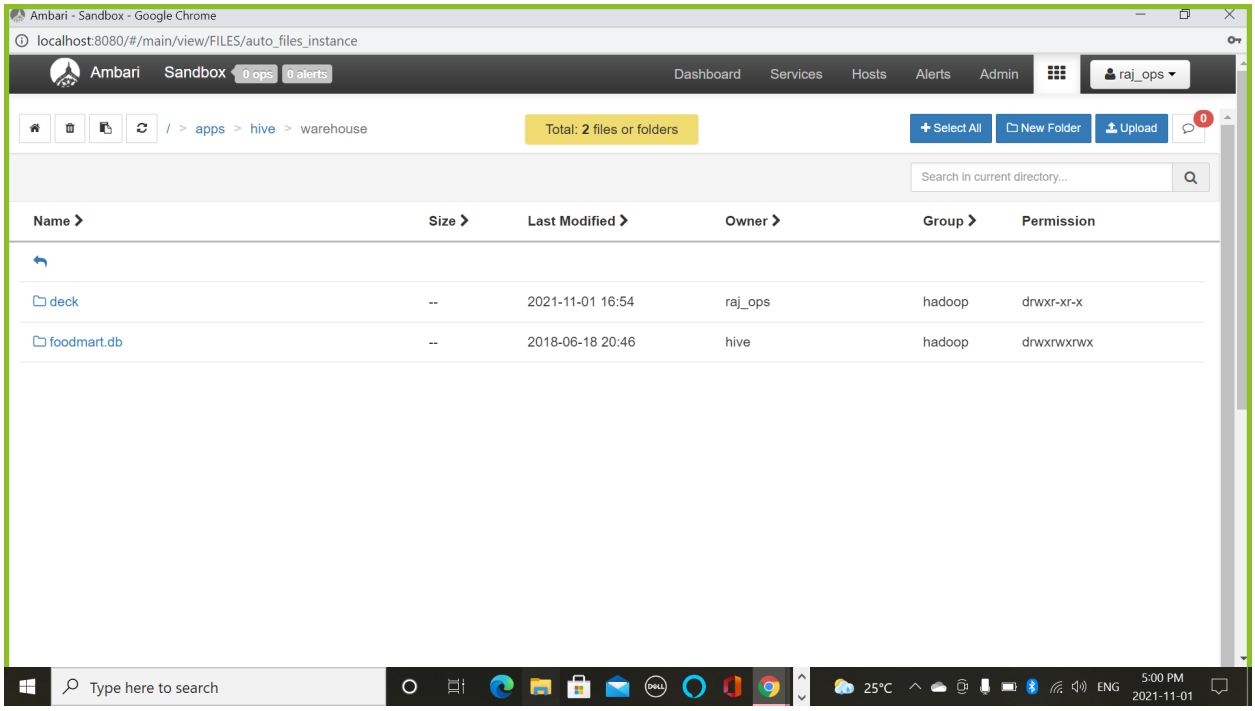
Step-5: Checking managed tables in sandbox



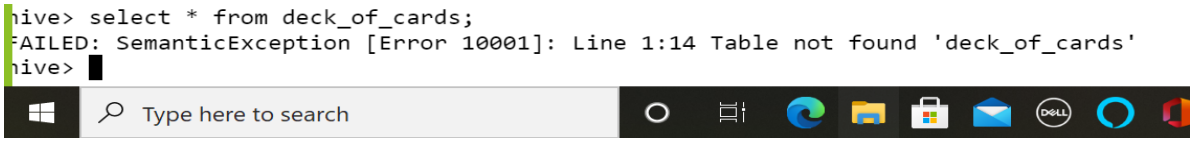
3.) Drop a managed table and check the result in HDFS

Step-1: Drop managed table in hive 2.0



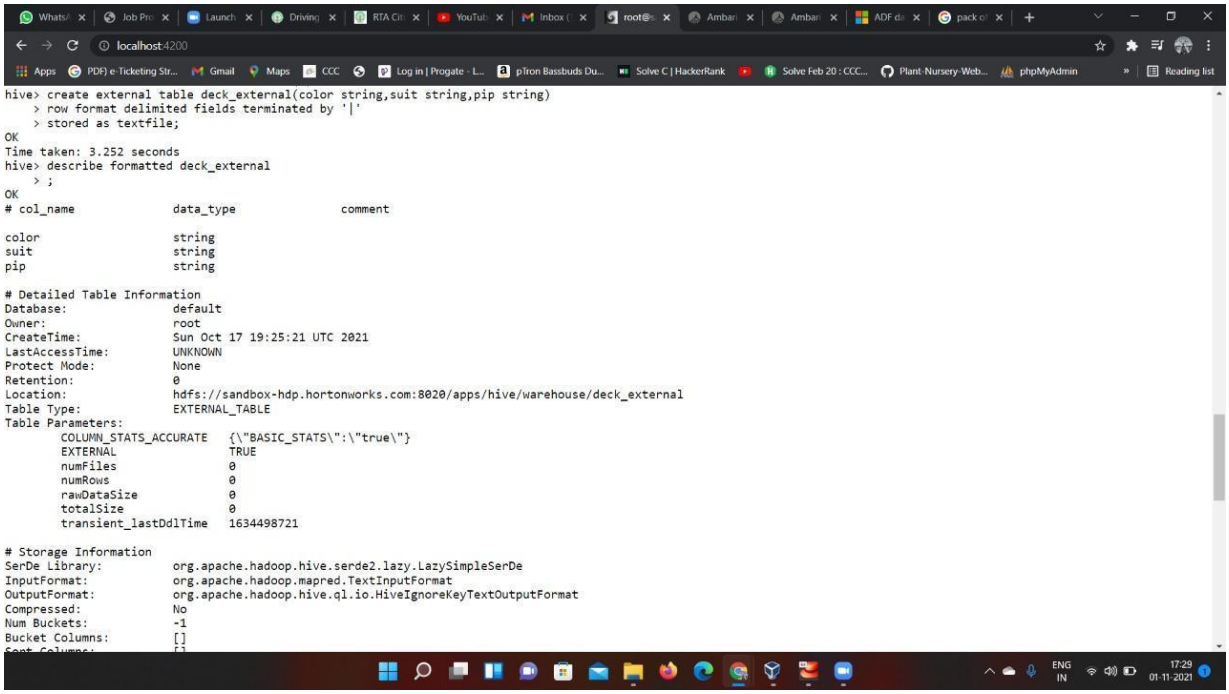


Step-2: Checking whether table exist or not



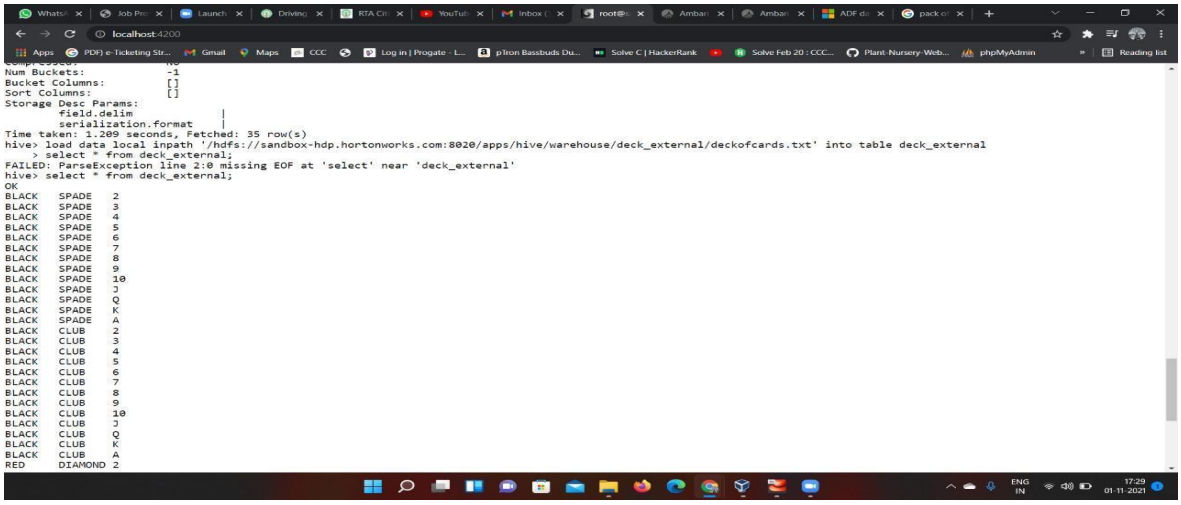
4.) Create an external table and load the data from LFS

Step-1: Create an external table deck_external and describe table to get location of it



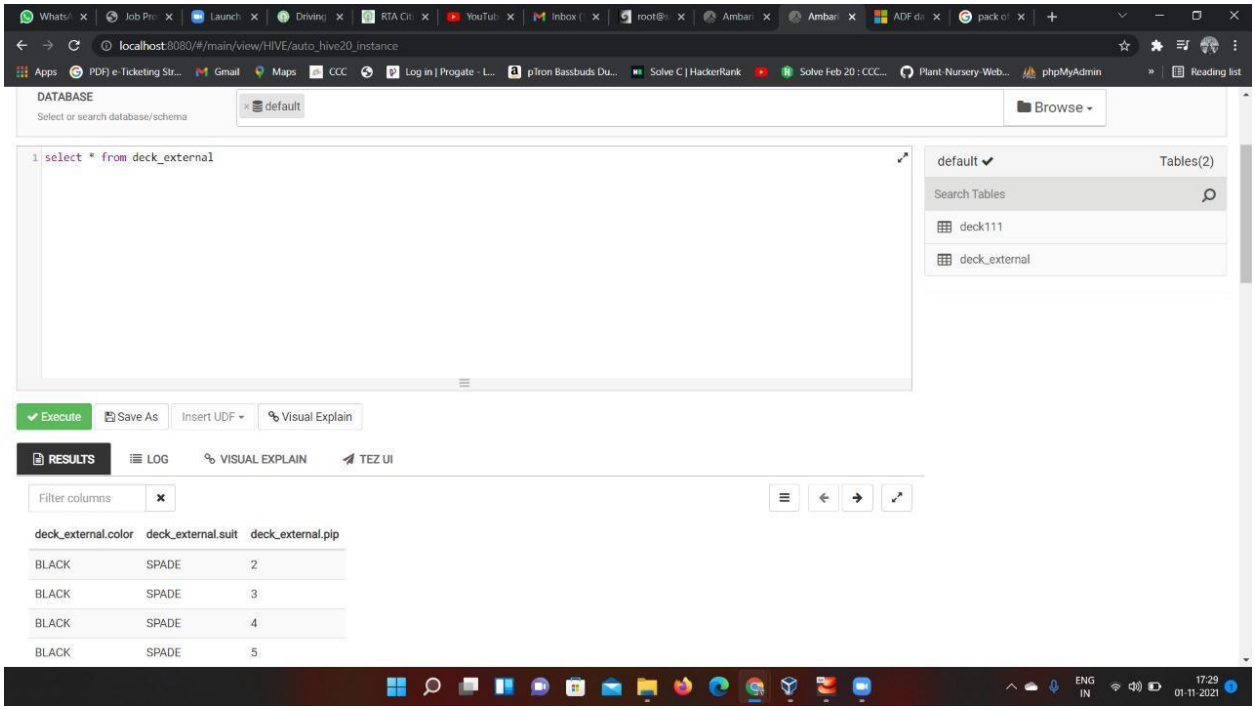
```
hive> create external table deck_external(color string,suit string,pip string)
    > row format delimited fields terminated by '|'
    > stored as textfile;
OK
Time taken: 3.252 seconds
hive> describe formatted deck_external
    > ;
OK
# col_name              data_type               comment

color                   string
suit                    string
pip                     string

# Detailed Table Information
Database:               default
Owner:                  root
CreateTime:             Sun Oct 17 19:25:21 UTC 2021
LastAccessTime:         UNKNOWN
Protect Mode:           None
Retention:              0
Location:               hdfs://sandbox-hdp.hortonworks.com:8020/apps/hive/warehouse/deck_external
Table Type:             EXTERNAL_TABLE
Table Parameters:
        COLUMN_STATS_ACCURATE   {\"BASIC_STATS\":\"true\"}
        EXTERNAL                TRUE
        numFiles                0
        numRows                 0
        rawDataSize             0
        totalSize               0
        transient_lastDdlTime   1634498721

# 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:         []
```

Step-2: Load cards data into table and display contents of it using select command



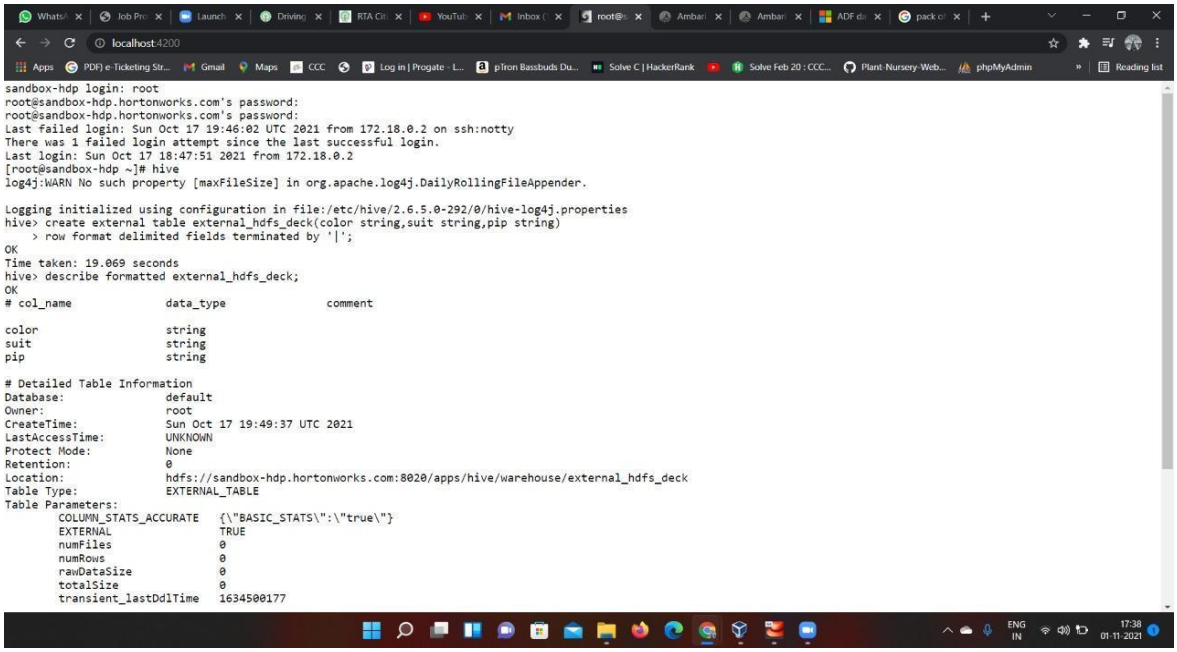
```
Num Buckets:            -1
Bucket Columns:         []
Sort Columns:           []
Storage Desc Params:
        field.delim             |
        serialization.format    |
Time taken: 1.209 seconds, Fetched: 35 row(s)
hive> load data local inpath '/hdfs://sandbox-hdp.hortonworks.com:8020/apps/hive/warehouse/deck_external/deckofcards.txt' into table deck_external
    > select * from deck_external;
FAILED: ParseException line 2:0 missing EOF at 'select' near 'deck_external'
hive> select * from deck_external;
OK
BLACK   SPADE   2
BLACK   SPADE   3
BLACK   SPADE   4
BLACK   SPADE   5
BLACK   SPADE   6
BLACK   SPADE   7
BLACK   SPADE   8
BLACK   SPADE   9
BLACK   SPADE   10
BLACK   SPADE   J
BLACK   SPADE   Q
BLACK   SPADE   K
BLACK   SPADE   A
BLACK   CLUB    2
BLACK   CLUB    3
BLACK   CLUB    4
BLACK   CLUB    5
BLACK   CLUB    6
BLACK   CLUB    7
BLACK   CLUB    8
BLACK   CLUB    9
BLACK   CLUB    10
BLACK   CLUB    J
BLACK   CLUB    Q
BLACK   CLUB    K
BLACK   CLUB    A
RED     DIAMOND 2
```

Step-3: Display table in hive 2.0 using select command

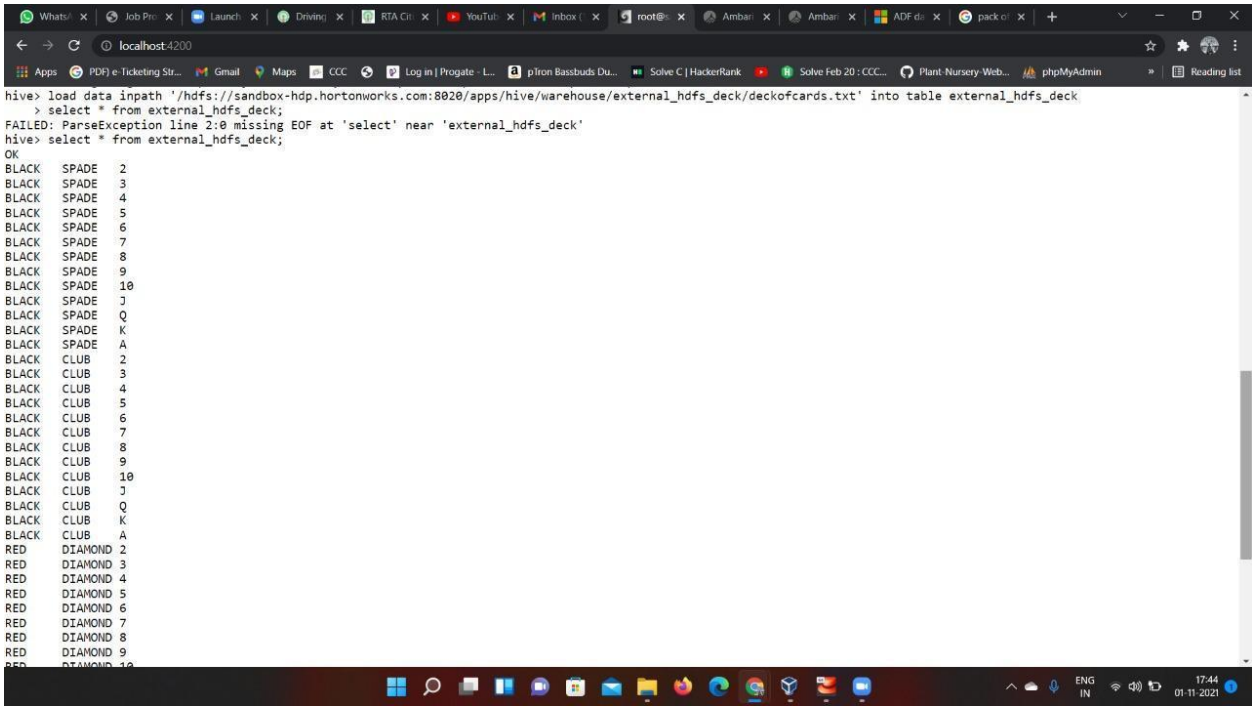


5.) Create an external table and load the data from HDFS

Step-1: Create an external table external_hdfs_deck and describe it to get location of it

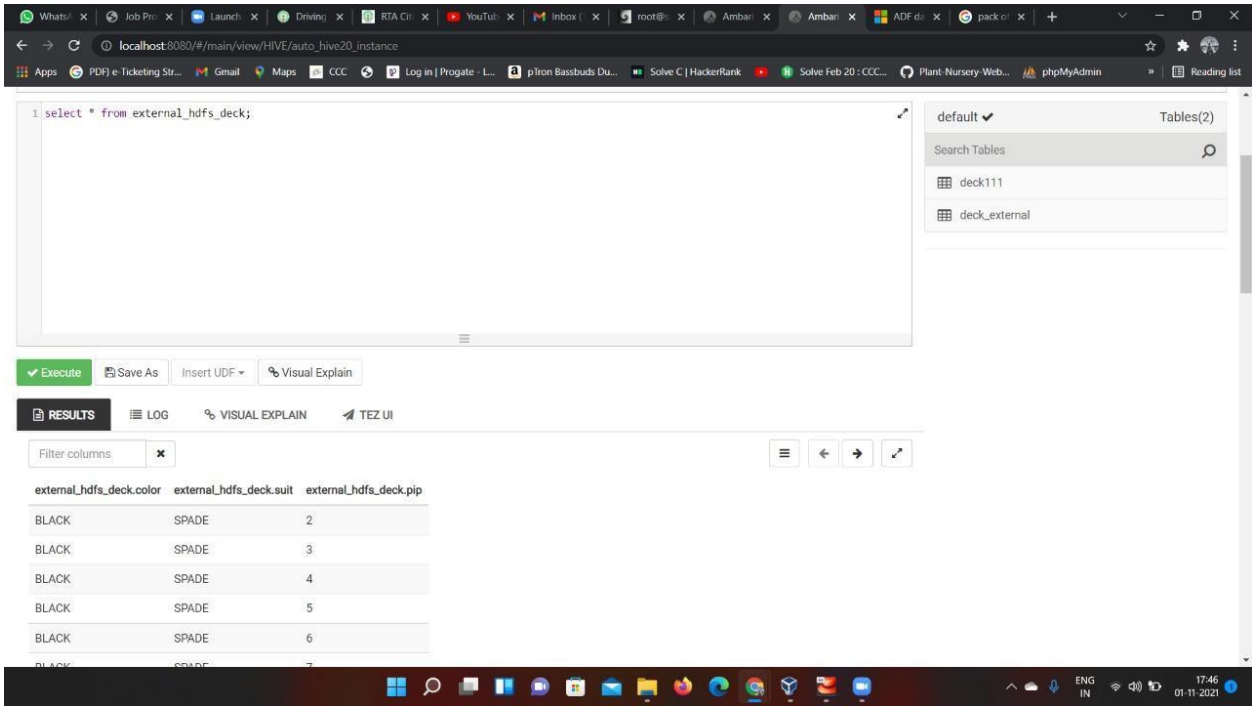


Step-2: Load cards data into table and display contents of it using select command



```
hive> load data inpath '/hdfs://sandbox-hdp.hortonworks.com:8020/apps/hive/warehouse/external_hdfs_deck/deckofcards.txt' into table external_hdfs_deck
    > select * from external_hdfs_deck;
FAILED: ParseException line 2:0 missing EOF at 'select' near 'external_hdfs_deck'
hive> select * from external_hdfs_deck;
OK
BLACK   SPADE   2
BLACK   SPADE   3
BLACK   SPADE   4
BLACK   SPADE   5
BLACK   SPADE   6
BLACK   SPADE   7
BLACK   SPADE   8
BLACK   SPADE   9
BLACK   SPADE   10
BLACK   SPADE   J
BLACK   SPADE   Q
BLACK   SPADE   K
BLACK   SPADE   A
BLACK   CLUB    2
BLACK   CLUB    3
BLACK   CLUB    4
BLACK   CLUB    5
BLACK   CLUB    6
BLACK   CLUB    7
BLACK   CLUB    8
BLACK   CLUB    9
BLACK   CLUB    10
BLACK   CLUB    J
BLACK   CLUB    Q
BLACK   CLUB    K
BLACK   CLUB    A
RED     DIAMOND 2
RED     DIAMOND 3
RED     DIAMOND 4
RED     DIAMOND 5
RED     DIAMOND 6
RED     DIAMOND 7
RED     DIAMOND 8
RED     DIAMOND 9
```

Step-3: Using select command display table in hive 2.0



```
select * from external_hdfs_deck;
```

default ✔ Tables(2)

Search Tables

deck111

deck_external

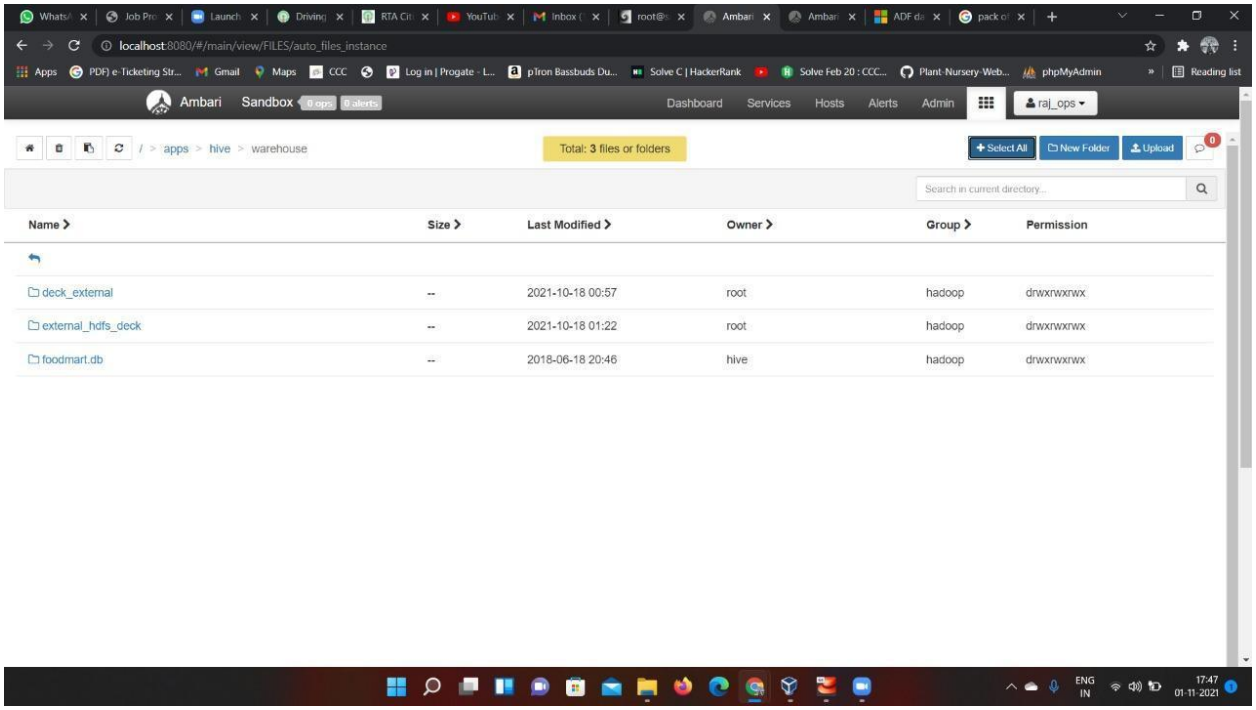
✔ Execute | Save As | Insert UDF | Visual Explain

RESULTS | LOG | VISUAL EXPLAIN | TEZ UI

Filter columns

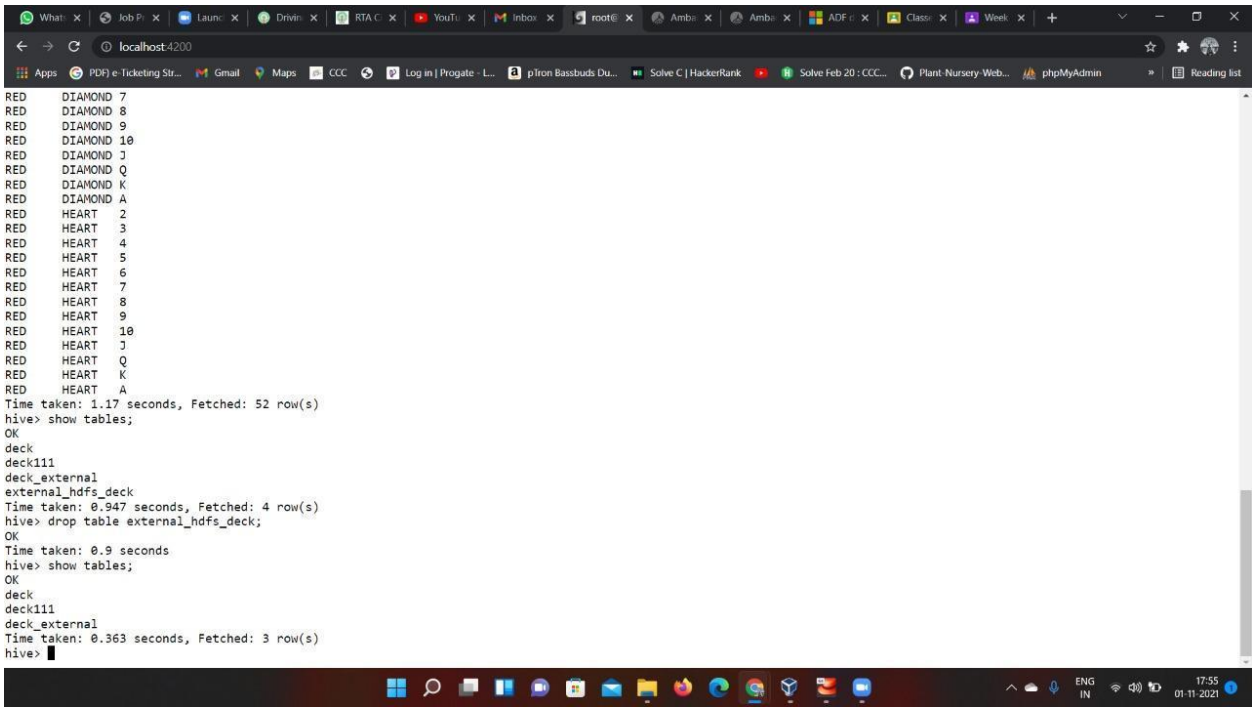
external_hdfs_deck.color	external_hdfs_deck.suit	external_hdfs_deck.pip
BLACK	SPADE	2
BLACK	SPADE	3
BLACK	SPADE	4
BLACK	SPADE	5
BLACK	SPADE	6

Step-4: Checking external tables in sandbox



6.) Drop an external table and check the data from HDFS

Step-1: Drop external table using drop command and using show command checking table



Step-2: Checking status of table in sandbox