

Stock data Analysis using Hive.

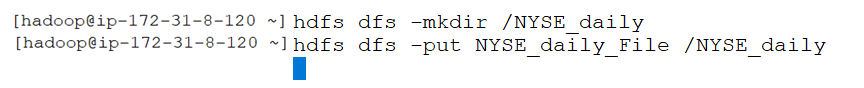
Solution:

1. Create a directory in HDFS to copy this data in HDFS

You can create a directory in HDFS by executing below command and then copy give NYSE file into HDFS folder

hdfs dfs -mkdir /NYSE\_daily

hdfs dfs -put NYSE\_daily\_File /NYSE\_daily

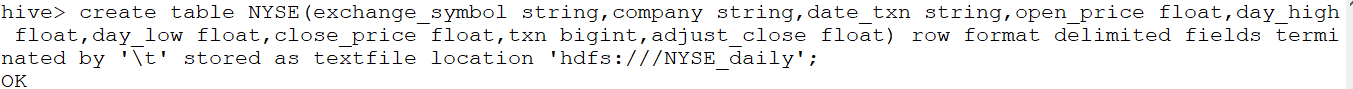


2. To Load data into partitioned table, we need to load data into non-partition table and then by using SQL Query we can load the data into partitioned table.

Lets create "NYSE" as non partition table

hive>

create table NYSE(exchange\_symbol string,company string,date\_txn string,open\_price float,day\_high float,day\_low float,close\_price float,txn bigint,adjust\_close float) row format delimited fields terminated by '\t' stored as textfile location 'hdfs:///NYSE\_daily';



3.

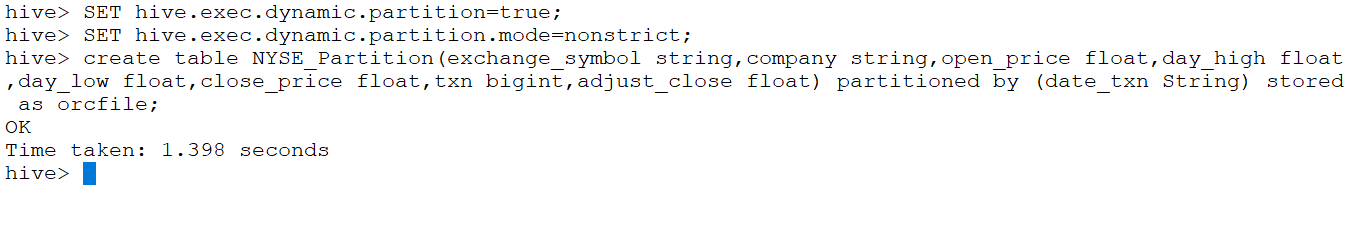
Let’s create a partitioned table as ORC File which is partitioned on date field.

hive>

SET hive.exec.dynamic.partition=true;

SET hive.exec.dynamic.partition.mode=nonstrict;

create table NYSE\_Partition(exchange\_symbol string,company string,open\_price float,day\_high float,day\_low float,close\_price float,txn bigint,adjust\_close float) partitioned by (date\_txn String) stored as orcfile;



4.

Now, we can load data into partition table using Hive SQL Query

hive>

from NYSE insert into NYSE\_Partition partition(date\_txn) select exchange\_symbol,company,open\_price,day\_high,day\_low,close\_price,txn,adjust\_close,date\_txn where open\_price > 68 AND close\_price <70;

