

Uploaded NYSE_daily_File.csv to my Linux local system using FTP :

```
$ cd Assignment_NYSE_Daily_File/
```

```
$ ls
```

```
NYSE_daily_File.csv
```

Making Stock Data directory in HDFS:

```
$ hdfs dfs -mkdir /user/satish.patil678_gmail/Stock_Data
```

Loading file from my local to HDFS directory:

```
$ hdfs dfs -put NYSE_daily_File.csv /user/satish.patil678_gmail/Stock_Data
```

```
$ hdfs dfs -ls /user/satish.patil678_gmail/Stock_Data
```

Found 1 items

```
-rw-r--r-- 3 satish.patil678_gmail hadoop 3196181 2018-05-26 13:55
/user/satish.patil678_gmail/Stock_Data/NYSE_daily_File.csv
```

Creating Data Base in Hive:

```
> create database Stock678;
```

Creating temporary Table in Hive to load .csv text file (We cannot load text file directly into Partition table with ORC Format):

```
> use Stock678;
```

```
> create table temp_Stock_Data(stock_exchange string,company_symbol string,stock_date
string,day_open_price float,day_high_price float,day_low_price float,day_close_price float,volume
int,day_adjusted_close_price float) row format delimited fields terminated by ',' stored as textfile;
```

```
> desc temp_Stock_Data;
```

Loading .csv file data into above created temporary table:

```
> load data inpath '/user/satish.patil678_gmail/Stock_Data/' into table temp_Stock_Data;
```

```
> select * from temp_Stock_Data;
```

Time taken: 0.546 seconds, Fetched: 57391 row(s)

- 1) **NYSE_Partition table based on the company symbol field.**
- 2) **Table needs to store the data as ORC file.**

Setting parameters for Partition:

```
> set hive.exec.dynamic.partition=true;  
> set hive.exec.dynamic.partition.mode=nonstrict;
```

Creating partition table based on company symbol:

```
> create table Stock_NYSE(stock_exchange string,stock_date date,day_open_price  
float,day_high_price float,day_low_price float,day_close_price float,volume  
int,day_adjusted_close_price float) partitioned by (company_symbol string) stored as orcfile;  
> desc Stock_NYSE;  
> set hive.exec.max.dynamic.partitions.pernode=1000;
```

- 3) **Load data of only those records where open price of the day is greater than 68 and high price of the day is less than 70.**

Loading data from temporary table into partitioned table.

```
> from temp_Stock_Data insert into table Stock_NYSE partition (company_symbol) select  
stock_exchange,to_date(from_unixtime(UNIX_TIMESTAMP(stock_date,'dd-MM-  
yyyy'))),day_open_price,day_high_price,day_low_price,day_close_price,volume,day_adjusted_close_  
price,company_symbol where day_open_price > 68.0 and day_high_price < 70;  
> select * from Stock_NYSE;
```

Time taken: 0.239 seconds, Fetched: 79 row(s)