

## **\*\* create table nyse**


create table nyse\_data(stockticker string,tradingdate string,open\_price float,high\_price float,low\_price float,close\_price float,volume bigint)row format delimited fields terminated by ',' stored as textfile;

## **\*\*load data in nyse**

load data local inpath '/home/cloudera/Documents/nyse\_2009.csv' into table nyse\_data;

## **\*\* create table company**

create table company(symbol string,name string,lastsale decimal,marketcap decimal,col\_5 string,col\_6 string,sector string,industry string, url string)row format delimited fields terminated by '|' stored as textfile;



```
cloudera@quickstart:~$ hive
Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.p
roperties
WARNING: Hive CLI is deprecated and migration to Beeline is recommended.
hive> create table nyse_data(stockticker string,tradingdate string,open_price fl
oat,high_price float,low_price float,close_price float,volume bigint)row format
delimited fields terminated by ',' stored as textfile;
OK
Time taken: 1.065 seconds
hive> load data local inpath '/home/cloudera/Documents/nyse_2009.csv' into table nyse_data;
Loading data to table default.nyse_data
Table default.nyse_data stats: [numFiles=1, totalSize=4527855]
OK
Time taken: 0.908 seconds
hive> create table company(symbol string,name string,lastsale decimal,marketcap decimal,col_5 string,col_6 string,sector string,industry string, url string)row format d
elimited fields terminated by '|' stored as textfile;
OK
Time taken: 0.289 seconds
```

## **\*\* load data in company table**

load data local inpath '/home/cloudera/Documents/companylist\_noheader.csv' into table company;

## **\*\*Sql Query**

### **1. for company wise**

SELECT c.name as company\_name ,max(n.volume) as max\_share\_count FROM nyse\_data n join company as c on n.stockticker=c.symbol group by c.name order by max\_share\_count desc LIMIT 1;

**Result- Bank of America Corporation 1226792000**

### **2. for industry wise**

```
SELECT c.industry as industry_name ,max(n.volume) as max_share_count FROM nyse_data n join  
company as c on n.stockticker=c.symbol group by c.industry order by max_share_count desc LIMIT  
1;
```

**Result- Major Banks    1226792000**

### **3. for sector wise**

```
SELECT c.sector as sector_name ,max(n.volume) as max_share_count FROM nyse_data n join  
company as c on n.stockticker=c.symbol group by c.sector order by max_share_count desc LIMIT 1;
```

**Result- Finance            1226792000**