

## HIVE

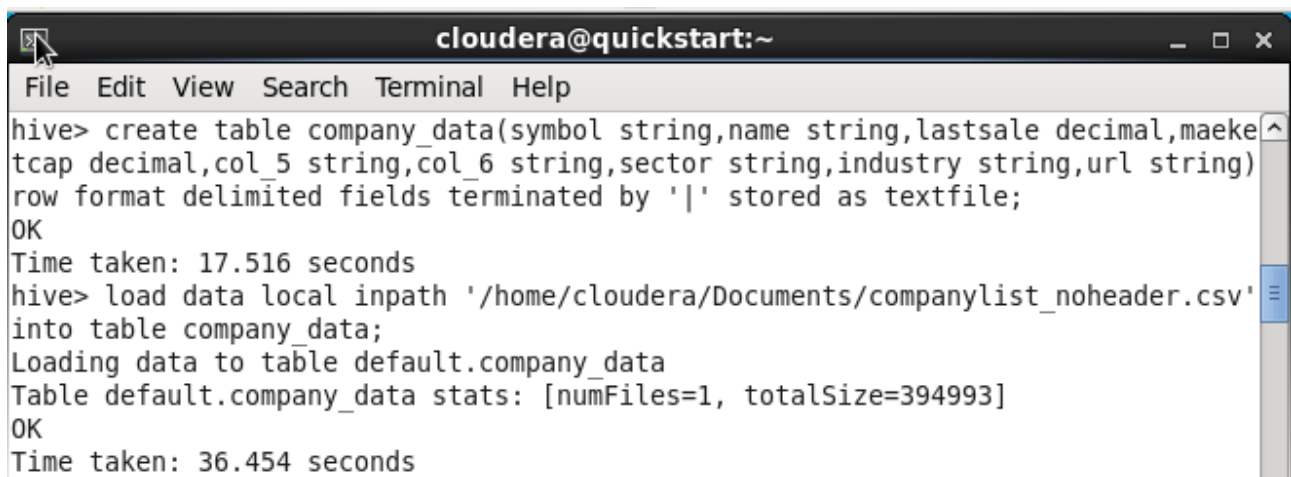
1. Write a query to calculate the maximum number of shares traded in 2009

- i) companywise
- ii) industrywise
- iii) sectorwise

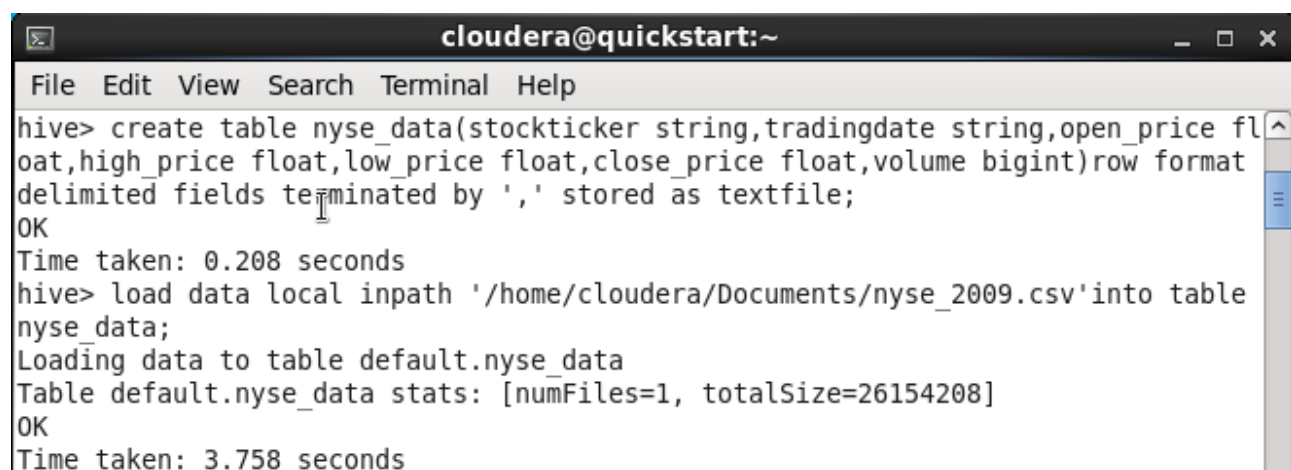
STEPS :

a) create a table for company data and nyse data

b) load data to the table



```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
hive> create table company_data(symbol string,name string,lastsale decimal,maeke  
tcap decimal,col_5 string,col_6 string,sector string,industry string,url string)  
row format delimited fields terminated by '|' stored as textfile;  
OK  
Time taken: 17.516 seconds  
hive> load data local inpath '/home/cloudera/Documents/companylist_noheader.csv'  
into table company_data;  
Loading data to table default.company_data  
Table default.company_data stats: [numFiles=1, totalSize=394993]  
OK  
Time taken: 36.454 seconds
```



```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
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: 0.208 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=26154208]  
OK  
Time taken: 3.758 seconds
```

c) write a query by

Before using JOIN Query we will set that command **SET hive.auto.convert.join=false**

i) companywise

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
[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> SET hive.auto.convert.join=false  
    > ;  
hive> select cd.name,MAX(nd.volume) from company_data cd JOIN nyse_data nd ON cd  
.symbol = nd.stockticker where tradingdate between '01/01/2009' AND '31/12/2009'  
group by name limit 50;  
Query ID = cloudera_20210611042626_a70c268a-1ef5-487f-8d2b-17b56dd86f9c  
Total jobs = 2
```

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
3D Systems Corporation 291000  
3M Company 11541100  
A.H. Belo Corporation 2205300  
AAR Corp. 2193900  
ABB Ltd 10402700  
ABM Industries Incorporated 2040100  
AGCO Corporation 8722300  
AGL Resources, Inc. 2129500  
AK Steel Holding Corporation 21113500  
ALLIANCEBERNSTEIN INCOME FUND INC 1941000  
AMERIPRISE FINANCIAL SERVICES, INC. 34678100  
AMN Healthcare Services Inc 1142500  
AMREP Corporation 57800  
AMTEK, Inc. 3156300  
ANN INC. 10452300  
AOL Inc. 13092100  
ASA Gold and Precious Metals Limited 319800  
AT&T Inc. 61246200  
AU Optronics Corp 11996500  
AVX Corporation 1268700  
AZZ Incorporated 1511300  
Aaron's, Inc. 6602700  
Abbott Laboratories 26743800  
Abercrombie & Fitch Company 13615800
```

## ii)Industrywise

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
hive> select cd.industry,MAX(nd.volume) from company_data cd JOIN nyse_data nd ON  
N cd.symbol = nd.stockticker where tradingdate between '01/01/2009' AND '31/12/2  
009' group by industry limit 50;  
Query ID = cloudera_20210611045454_e625bf35-626d-46f1-85e2-5380912c06a1  
Total jobs = 2  
Launching Job 1 out of 2
```

```
cloudera@quickstart:~  
File Edit View Search Terminal Help
```

Accident &Health Insurance	46322900
Advertising	20133400
Aerospace	43549300
Agricultural Chemicals	35457100
Air Freight/Delivery Services	70268096
Aluminum	2823700
Apparel	25609600
Auto Manufacturing	541015000
Auto Parts:O.E.M.	103526200
Automotive Aftermarket	14966300
Beverages (Production/Distribution)	62086900
Biotechnology: Commercial Physical & Biological Research	90581900
Biotechnology: Electromedical & Electrotherapeutic Apparatus	51810700
Biotechnology: Laboratory Analytical Instruments	7938000
Books	2023400
Broadcasting	49699400
Building Materials	107831400
Building Products	3662100
Building operators	20712000
Business Services	142985300
Catalog/Specialty Distribution	385400
Clothing/Shoe/Accessory Stores	34168600
Coal Mining	25003000
Commercial Banks	143028496

### iii)sectorwise

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
hive> select cd.sector,MAX(nd.volume) from company_data cd JOIN nyse_data nd ON  
cd.symbol = nd.stockticker where tradingdate between '01/01/2009' AND '31/12/200  
9' group by sector limit 50;  
Query ID = cloudera_20210611052424_4221490a-1260-47b5-8e6b-ea6aa432c57f  
Total jobs = 2
```

Basic Industries	125062800
Capital Goods	541015000
Consumer Durables	28364200
Consumer Non-Durables	147438000
Consumer Services	144189000
Energy	752731008
Finance	1226792000
Health Care	210463008
Miscellaneous	142985300
Public Utilities	244426200
Technology	163101600
Transportation	70268096
n/a	43881700