

Q: Q. W A Query to perform max no of share traded Industry wise, Sector Wise, Company Wise.

Ans :

Nyse 2009.csv columns

stockticker string,
tradingdate string,
open_price float,
high_price float,
low_price float,
close_price float,
volume bigint

Company_data.csv

symbol string,
name string,
lastsale decimal,
marketcap decimal,
col_5 string,
col_6 string,
sector string,
industry string,
url string

Creating Schema “Comp1” for CompanyData

hive> CREATE TABLE IF NOT EXISTS Comp1 (Symbol String, name String, lastsale decimal, mktcap decimal, col5 string, col6 string, sector string, industry string, urladd string) ROW FORMAT DELIMITED FIELDS TERMINATED BY "|" STORED AS TEXTFILE;

Loading Data in “Comp1”

Hive> load data local inpath '/home/deep/Downloads/BigDataHadoop103-main/sample_data/companylist_noheader.csv' into table Comp1;

describe comp1;

OK
symbol string
name string
lastsale decimal(10,0)
mktcap decimal(10,0)
col5 string
col6 string
sector string
industry string
urladd string

Creating Schema “stock1” for NYSE 2009.csv

Hive> CREATE TABLE IF NOT EXISTS stock1 (StockSticker String, TrdDate String, OpenPrice Float, HighPrice Float, LowPrice Float, close_price float, volume bigint) ROW FORMAT DELIMITED FIELDS TERMINATED BY "," STORED AS TEXTFILE;

Loading Data in “stock1”

Hive> load data local inpath '/home/deep/Downloads/BigDataHadoop103-main/sample_data/nyse_2009.csv' into table stock1;

describe stock1;

OK
stockticker string
trddate string
openprice float
highprice float
lowprice float
close_price float
volume bigint

Query max no of share traded Industry Wise; → Hive> select c.industry, max(s.volume) from comp1 c join stock1 s on (c.symbol = s.stocksticker) group by c.industry;

[illegible][illegible]

```

Jun 12 10:19
deep@deep-inspiron-5579: /usr/local/bin

ds 100 num

0
1941000
34078100

319800

0
311300
410100

You'ds, S.A. 132400

80
6281100

Advantest1 ADS 1166000
and Income Fund 321600
and Income Fund II 367700
e Fund 215700

4418200

```

Query max no of share traded Sector Wise;

Hive> select c.sector, max(s.volume) from comp1 c join stock1 s on (c.symbol = s.stocksticker) group by c.sector;

```
Activities Terminal Jun 12 10:20
deep@deep-inspiron-5579: /usr/local/hive

hive> select c.sector, max(s.volume) from comp1 c join stock1 s on (c.symbol = s.stocksticker) group by c.sector;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
Query ID = deep_20210612095652_b4e215df-4569-4b3a-b4d7-7a3b2bcbcsafd
Total Jobs = 1
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hive/lib/log4j-slf4j-impl-2.6.2.jar/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/hadoop-2.6.0/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
2021-06-12 09:56:57 Starting to launch local task to process map join; maximum memory = 477626368
2021-06-12 09:56:58 Dump the side-table for tag: 0 with group count: 3298 into file: file:/tmp/hive/java/deep/128cf564-bc31-492c-aa60-99cae6457e8a/hive_2021-06-12_09-56-52_449_4207048044881470152-1/-local-10
005/HashTable-Stage-2/MapJoin-napfile20-..hashtable
2021-06-12 09:56:58 Uploaded 1 file to: file:/tmp/hive/java/deep/128cf564-bc31-492c-aa60-99cae6457e8a/hive_2021-06-12_09-56-52_449_4207048044881470152-1/-local-10005/HashTable-Stage-2/MapJoin-napfile20-..has
htable (118152 bytes)
2021-06-12 09:56:58 End of local task; Time Taken: 1.05 sec.
Execution completed successfully
MapReduce task succeeded
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reducers=<number>
Starting Job = job_1623470091288_0003, Tracking URL = http://deep-inspiron-5579:8088/proxy/application_1623470091288_0003/
Kill Command = /opt/hadoop-2.6.0/bin/hadoop job -kill job_1623470091288_0003
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
2021-06-12 09:57:04.494 Stage-2 map = 0%, reduce = 0%
2021-06-12 09:57:09.625 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 4.3 sec
2021-06-12 09:57:15.836 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 6.26 sec
MapReduce Total cumulative CPU time: 6 seconds 260 msec
Ended Job = job_1623470091288_0003
MapReduce jobs launched:
Stage-2: Map: 1 Reduce: 1 Cumulative CPU: 6.26 sec HDFS Read: 26166626 HDFS Write: 548 SUCCESS
Total MapReduce CPU Time Spent: 6 seconds 260 msec
OK
Basic Industries      125862800
Capital Goods         541015000
Consumer Durables     28364200
Consumer Non-Durables 147438000
Consumer Services     144189000
Energy                752731000
Finance               122672000
Health Care           210463000
Miscellaneous          142985300
Public Utilities       24426200
Technology            163101600
Transportation         70268096
N/A                   43881700
Time taken: 24.505 seconds, Fetched: 13 row(s)
hive>
```