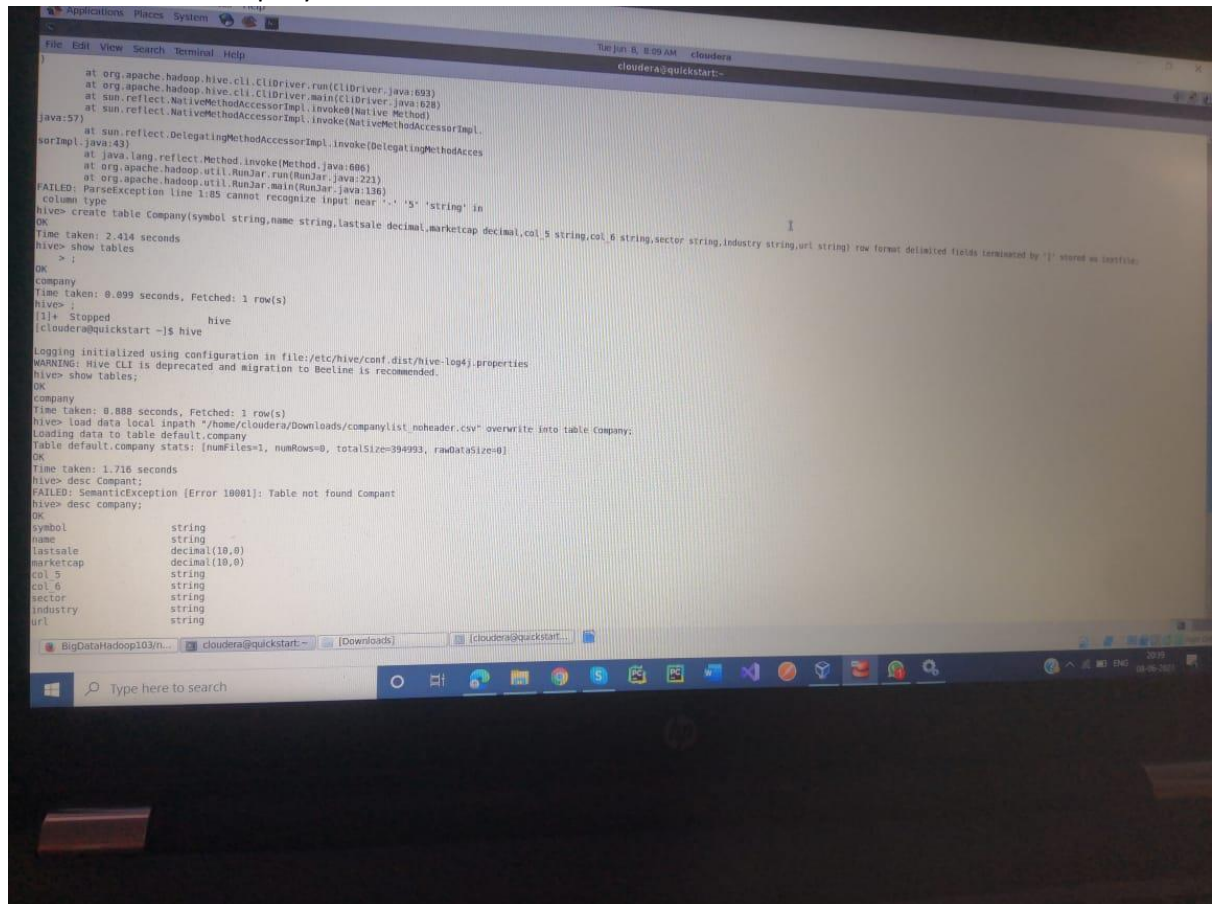


Hive Assignment:

1. Create table for Company.

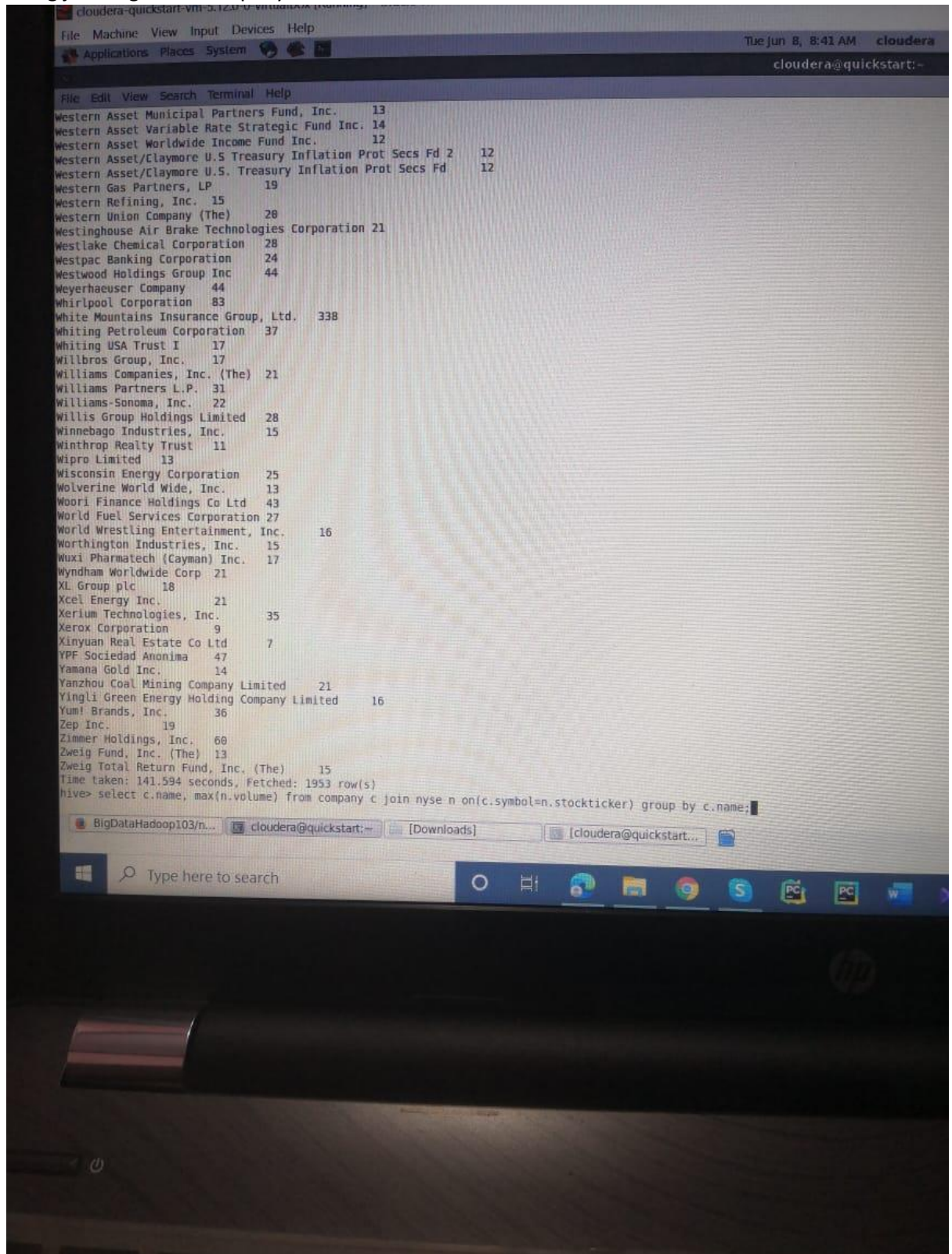


```
at org.apache.hadoop.hive.cli.CliDriver.run(CliDriver.java:693)
at org.apache.hadoop.hive.cli.CliDriver.main(CliDriver.java:678)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccess
sorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:886)
at org.apache.hadoop.util.RunJar.run(RunJar.java:221)
at org.apache.hadoop.util.RunJar.main(RunJar.java:136)
FAILED: ParseException line 1:85 cannot recognize input near '.' 's' 'string' in
column type
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 delimited fields terminated by '|' stored as textfile;
Time taken: 2.414 seconds
hive> show tables
+
OK
company
Time taken: 0.099 seconds, Fetched: 1 row(s)
hive> ;
[1]+  Stopped                  Hive
[cloudera@quickstart ~]$ hive
Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties
WARNING: Hive CLI is deprecated and migration to Beeline is recommended.
hive> show tables;
OK
company
Time taken: 0.888 seconds, Fetched: 1 row(s)
hive> load data local inpath "/home/cloudera/Downloads/companylist_noheader.csv" overwrite into table Company;
Loading data to table default.company
Table default.company stats: [numFiles=1, numRows=0, totalSize=394993, rawDataSize=0]
OK
Time taken: 1.716 seconds
hive> desc Company;
FAILED: SemanticException [Error 18001]: Table not found Company
hive> desc company;
OK
symbol          string
name            string
lastsale        decimal(10,0)
marketcap       decimal(10,0)
col 5           string
col 6           string
sector          string
industry        string
url             string
```

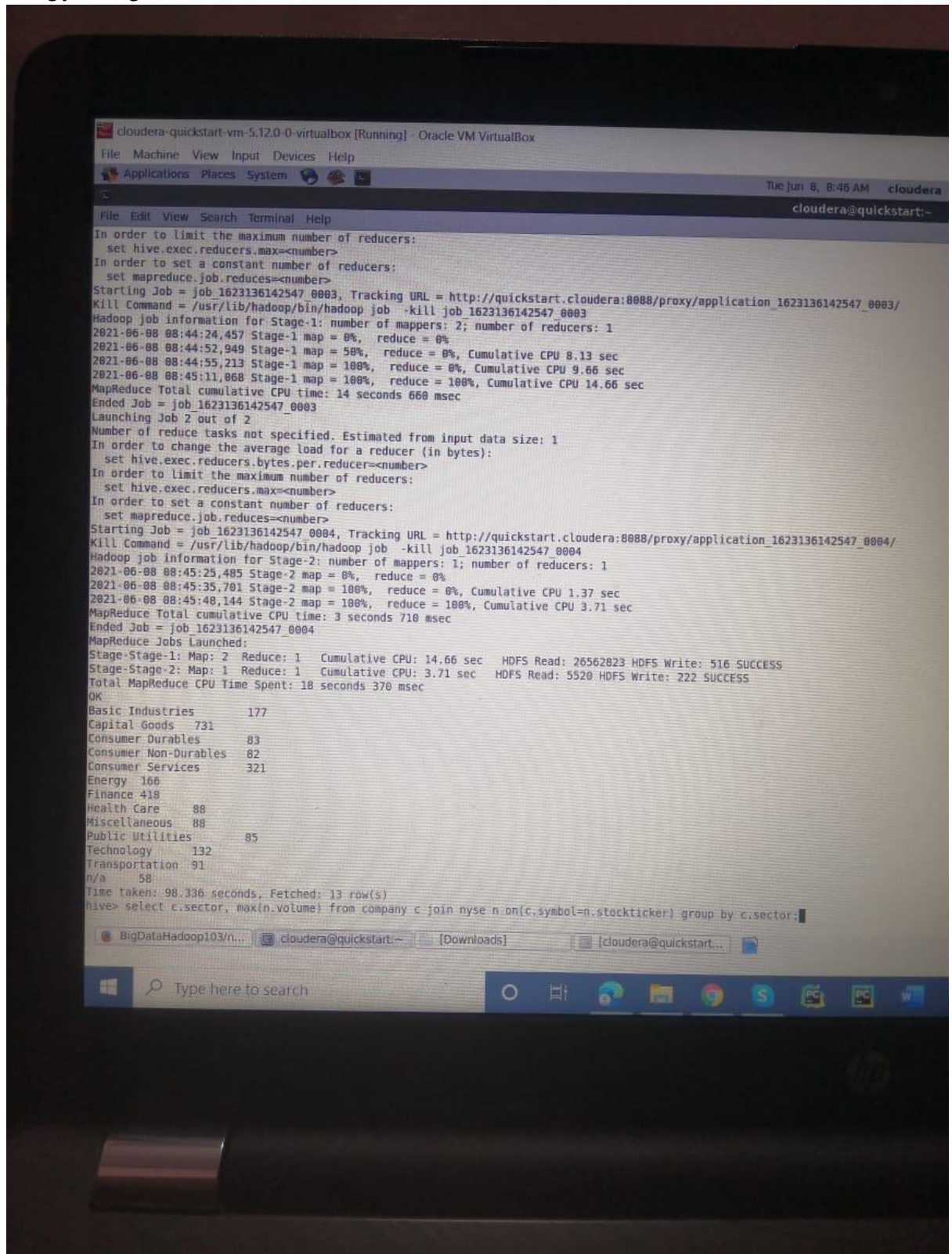
2. Create table for nesy.

```
Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties
WARNING: Hive CLI is deprecated and migration to Beeline is recommended.
hive> show tables;
OK
company
Time taken: 0.888 seconds, Fetched: 1 row(s)
hive> load data local inpath "/home/cloudera/Downloads/companylist noheader.csv" overwrite into table company;
Loading data to table default:company
Table default:company stats: [numFiles=1, numRows=0, totalSize=394993, rawDataSize=0]
OK
Time taken: 1.716 seconds
hive> desc company;
FAILED: SemanticException [Error 10001]: Table not found company
hive> desc company;
OK
symbol          string
name             string
lastsale         decimal(10,0)
marketcap        decimal(10,0)
col_5            string
col_6            string
sector           string
industry         string
url              string
Time taken: 0.158 seconds, Fetched: 9 row(s)
hive> create table nyse(stockticker string, tradingdate string, open price float, low price float, close price float, volume bigint) row format delimited fields terminated by ',' stored as textfile;
Time taken: 0.894 seconds
hive> show tables;
OK
company
nyse
Time taken: 0.077 seconds, Fetched: 2 row(s)
hive> load data local inpath "/home/cloudera/Downloads/nyse_2009.csv" overwrite into table nyse;
Loading data to table default:nyse
Table default:nyse stats: [numFiles=1, numRows=0, totalSize=26154288, rawDataSize=0]
OK
Time taken: 2.861 seconds
hive> select c.name, max(n.volume) from company c join nyse n on (c.symbol=n.stockticker) group by c.name;
FAILED: SemanticException [Error 10002]: Line 1:61 Invalid column reference 'symbol'
hive> select c.name, max(n.volume) from company c join nyse n on (c.symbol=n.stockticker) group by c.name;
Query ID = cloudera_28210608889404_d1b92d0b-4f2b-498e-bb45-93fe1e0d05cf
Total jobs = 1
FAILED: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.mr.MapredLocalTask
hive> select c.name, max(n.volume) from company c join nyse n on (c.symbol=n.stockticker) group by c.name;
Query ID = cloudera_28210608889707_cc4c4ca6-52e4-426a-aff7-f3a531cf7f15
Total jobs = 1
FAILED: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.mr.MapredLocalTask
hive>
```

3. Using join to get the company list.



4. Using join to get sector wise.



5. Using join to get industry wise.

