HIVE SCRIPT…

create database nyse;

use nyse;

create table nyse\_data(exchang string , stock\_symbol string , date string , stock\_price\_open float , stock\_price\_high float , stock\_price\_low float , stock\_price\_close float , stock\_volume bigint , stock\_price\_adj\_close float) row format delimited fields terminated by ‘,’ tblproperties(“skip.header.line.count”=”1”);

load data local inpath ‘home/lavish/Desktop/nyse\_data.csv’ into table nyse\_data;

create table nyse\_join(stock\_symbol\_a string , stock\_symbol\_b string , date string , stock\_price\_high\_a float , stock\_price\_high\_b float) row format delimited fields terminated by ‘,’ ;

insert into table nyse\_join select a.stock\_symbol , b.stock\_symbol , a.date , a.stock\_price\_high , b.stock\_price\_high from nyse\_data a join nyse\_data b on a.date=b.date;

create table nyse\_result(stock\_symbol\_a string, stock\_symbol\_b string , month string , covariance float) row format delimited fields terminated by ‘,’ ;

insert into table nyse\_result select stock\_symbol\_a , stock\_symbol\_b , month(date) , (avg(stock\_price\_high\_a \* stock\_price\_high\_b) – (avg(stock\_price\_high\_a) \* avg(stock\_price\_high\_b))) from nyse\_join where stock\_symbol\_a < stock\_symbol\_b and year(date)=2008 group by stock\_symbol\_a , stock\_symbol\_b , month(date);

Export table nyse\_result to local as a csv file..

insert overwrite local directory ‘home/lavish/Desktop’ row format delimited fields terminated by ‘,’ select \* from nyse\_result;

Rename the exported file as nyse\_result.csv..