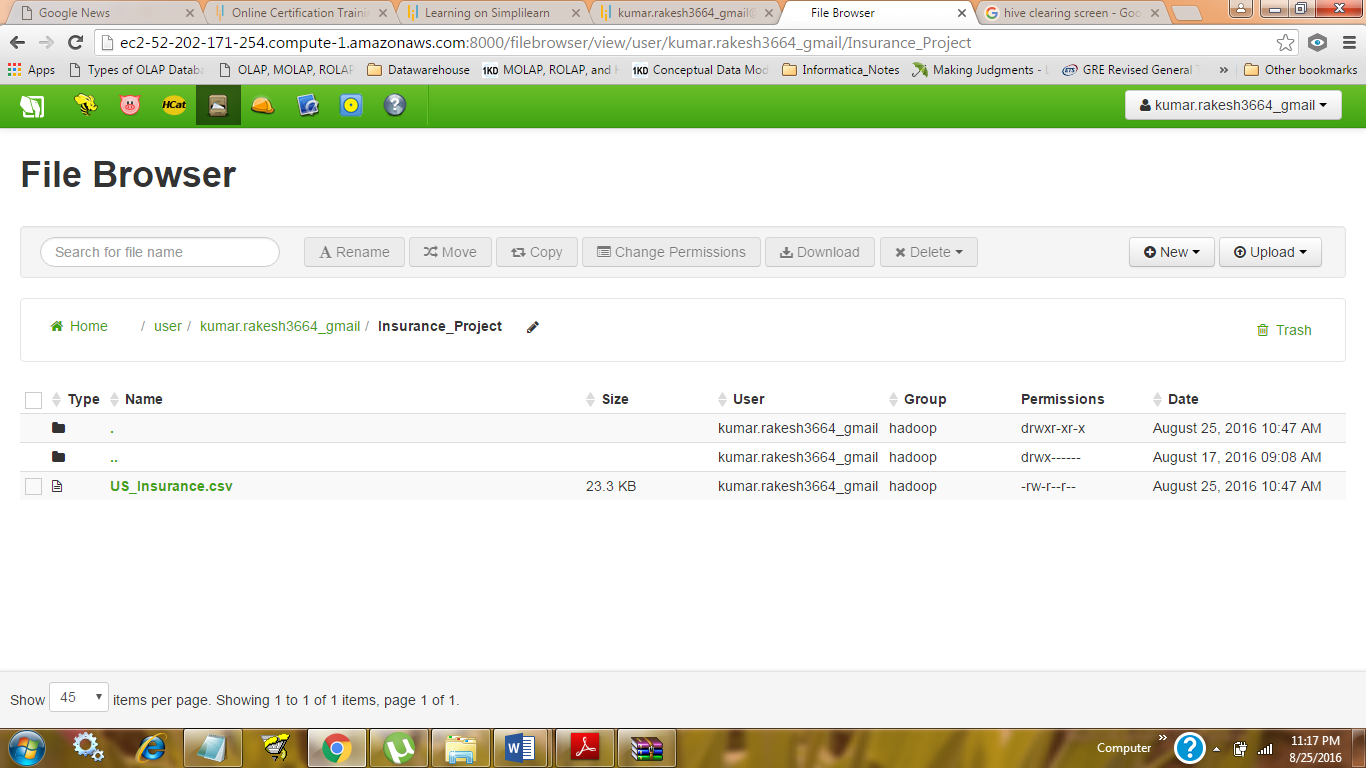
Project Solutions

1)Download Insurance data from the link below:

<http://www.census.gov/hhes/www/hlthins/data/utilization/tables.html>

2) Preparing Insurance file and uploading into Hue:

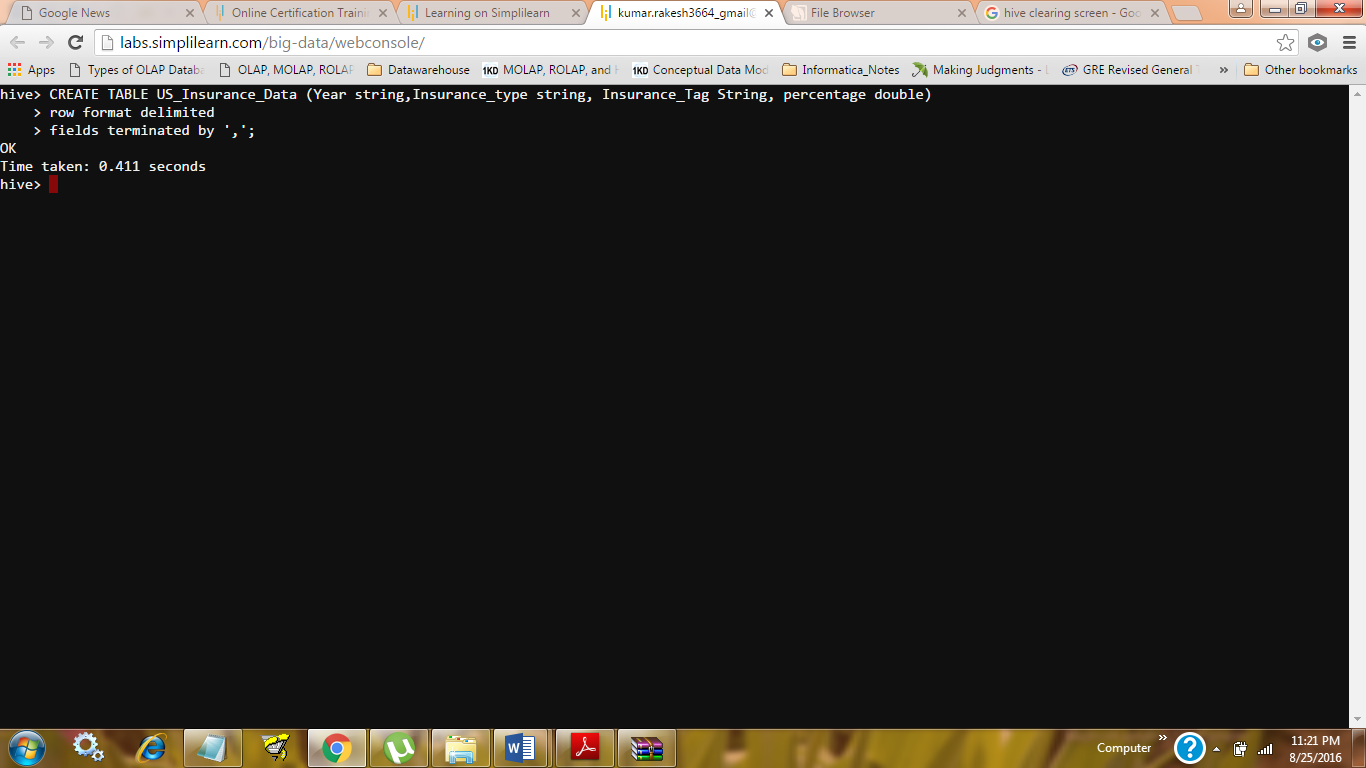


3) Creating US\_Insurance\_Data table:

CREATE TABLE US\_Insurance\_Data (Year string, Insurance\_type string, Insurance\_Tag String, percentage double)

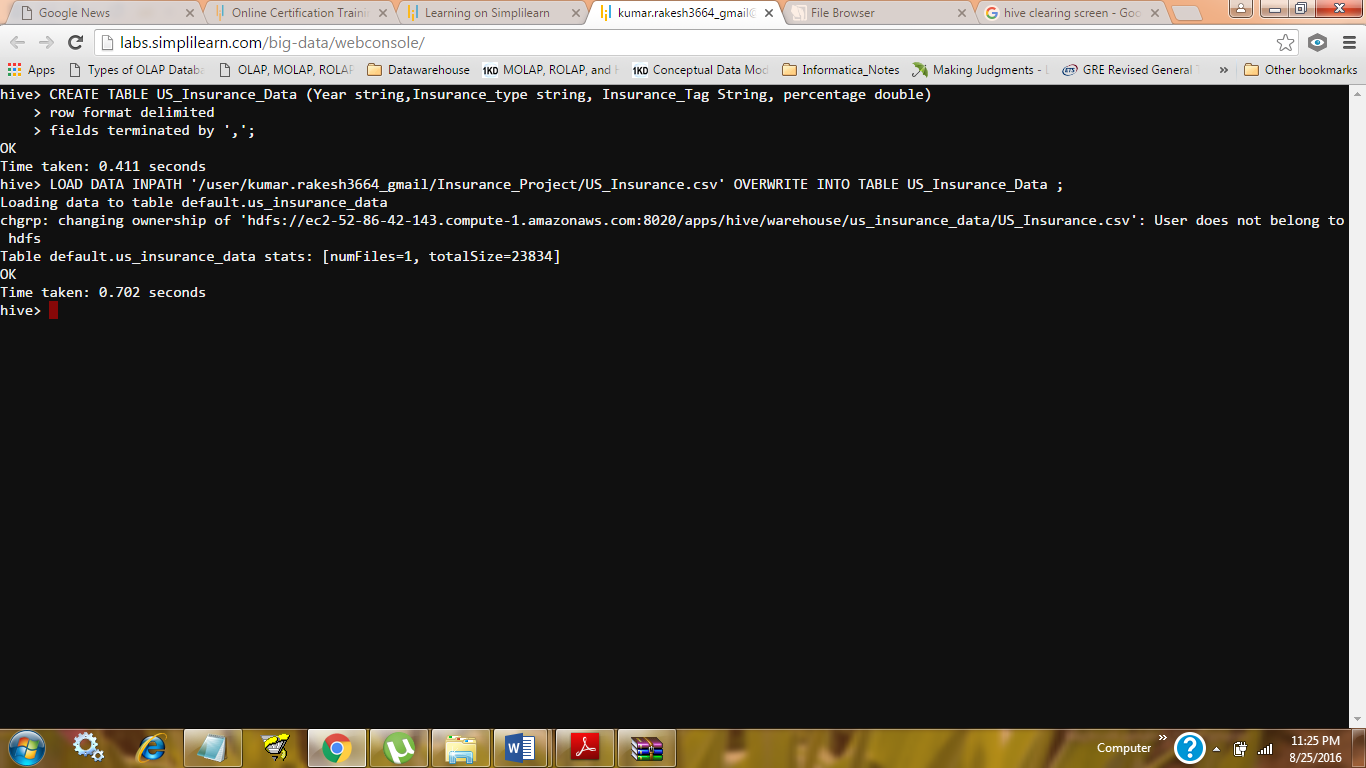
row format delimited

fields terminated by ',';



4) Loading data in US\_Insurance\_Data table:

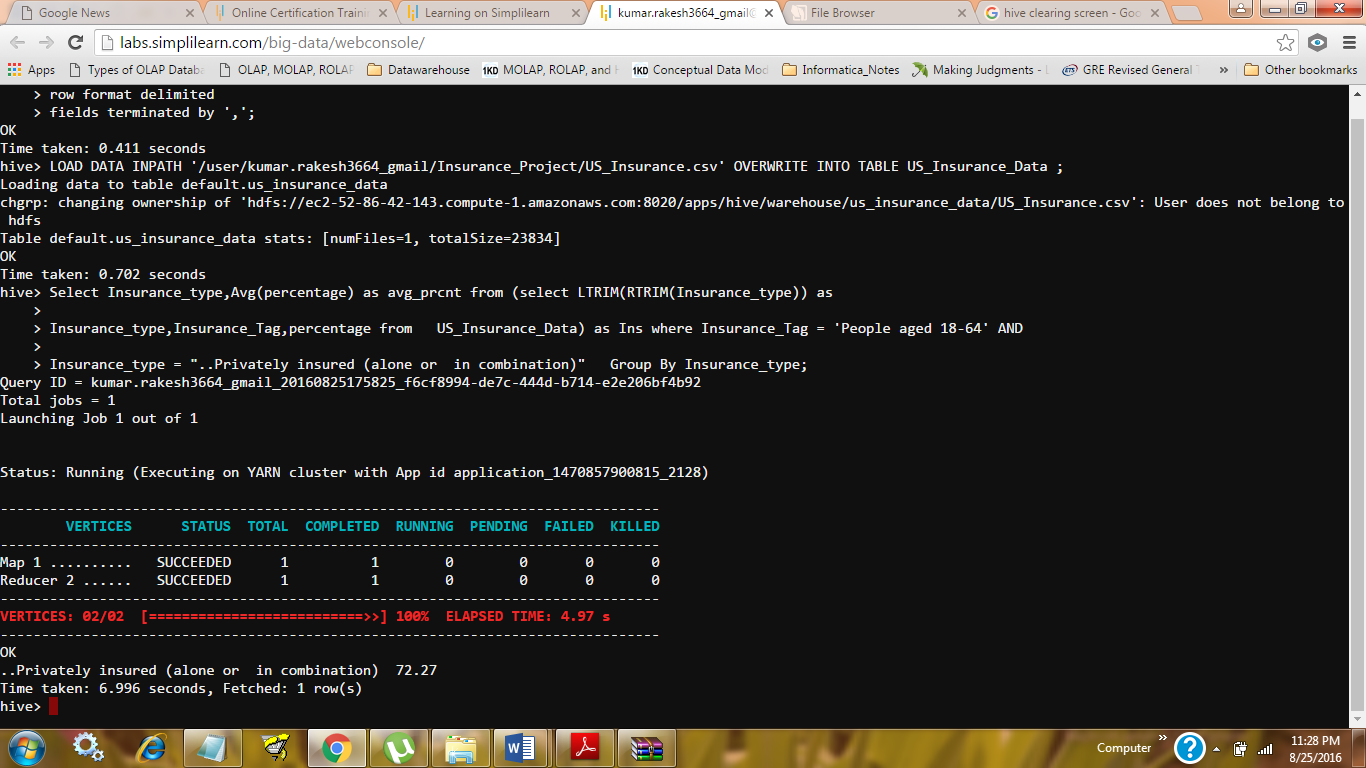
LOAD DATA INPATH '/user/kumar.rakesh3664\_gmail/Insurance\_Project/US\_Insurance.csv' OVERWRITE INTO TABLE US\_Insurance\_Data;



5) Hive query to find out average % of people, aged between 18 and 64, who have obtained insurance from private players from 2001 to 2011.

Select Insurance\_type, Avg (percentage) as avg\_prcnt from (select LTRIM (RTRIM (Insurance\_type)) as Insurance\_type, Insurance\_Tag, percentage from US\_Insurance\_Data) as Ins where Insurance\_Tag = 'People aged 18-64' AND Insurance\_type = "..Privately insured (alone or in combination)" Group By Insurance\_type;

Output:



6) Hive query to find out the average percentage of people, aged 65 years or more, who are solely covered by public insurance from the year 2001 to 2011.

Select Insurance\_type, Avg (percentage) as avg\_prcnt from (select LTRIM (RTRIM (Insurance\_type)) as

Insurance\_type, Insurance\_Tag, percentage from US\_Insurance\_Data) as Ins where Insurance\_Tag = 'People aged 65 and older' AND Insurance\_type = "..Publically insured (no private)" Group by Insurance\_type

Output:

