3)Which industry(SOC\_NAME) has the most number of Data Scientist positions?

[certified]

register /usr/local/hive/lib/hive-exec-1.2.1.jar

register /usr/local/hive/lib/hive-common-1.2.1.jar

data1 = LOAD '/home/hduser/h1b\_final/\*' USING PigStorage('\t') as (s\_no:long,case\_status:chararray,employer\_name:chararray,soc\_name:chararray,job\_title:chararray,full\_time\_position:chararray,prevailing\_wage:double,year:chararray,worksite:chararray,longitude,latitude);

data2 = FOREACH data1 GENERATE soc\_name,case\_status,job\_title;

data3 = FILTER data2 BY case\_status == 'CERITFIED' AND job\_title == ‘DATA SCIENTIST’;

data4 = group data3 by (soc\_name);

Count1 = foreach data4 generate group, COUNT(data3.jobtitle);

Desc1 = limit (order Count1 by $1 desc) 1;