Query 5

Find the most popular top 10 job positions for H1B visa applications for each year?

Platform used: PIG

Source input file:

/home/cloudera/h1b\_data

Query file:



Explanation:

Step1) Filtered the data from h1b\_data into different bags year wise.

Step2) Grouped each from step1 in terms of job\_title.

Step3) Count the number of petitions for the job in each bag of step2.

Step4) Sorted the bags in step3 in descending order.

Strp5) Limited the records in the each bags of step4 to 10.

Output files below:



