Query 10

Which are the job positions along with the number of petitions who have the success rate more that 70% in petitions and total petitions filed more that 1000?

Platform used: Mapreduce

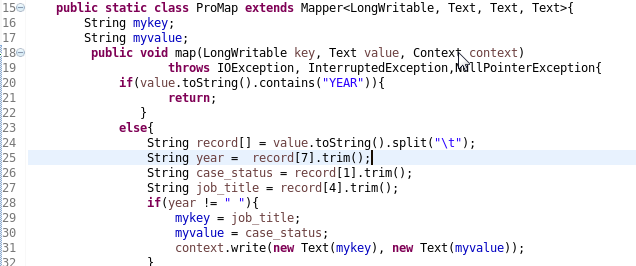
Inut file:

/user/hive/warehouse/project.db/h1b\_app2

Explanation:

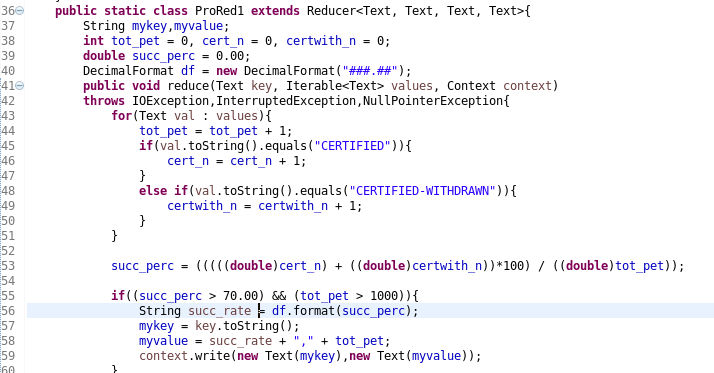
Program- query10.java

Mapper



Here I take the job\_title as the key and the case\_status as the value to count petitions for each case status for each job.

Reducer



Again we have used three separate variable for CERTIFIED, CERTIFIED-WITHDRAWN and total petitions. We calculate by adding values for each case status (lined 44 to 49).Not to mentions the tot\_pet is incremented for each loop no matter what. That represents the total count.

In line 53 we calculate the success pecent(succ\_perc) by adding CERTIFIED and CERTIFIED-WITHDRAWN petitions and dividing that sum by total petitions.

Program file:



Output file:

