Hive Queries

Given dataset is:

001 Amit 105 Data Minning

002 Pankaj 85 Data Engineer

003 Kiran 110 Data Scientist

004 Arpitha 95 Data Engineer

005 Viraj 105 Data Mining

006 Smitha 80 Data Analyst

007 Supriya 90 Data Engineer

008 Vihan 120 Data Scientist

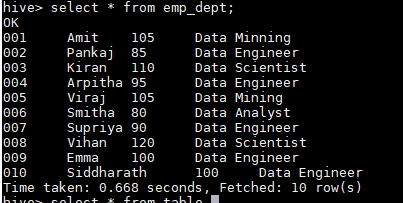
009 Emma 100 Data Engineer

010 Siddharath 100 Data Engineer

Create table and load data.

create table emp\_dept(id string,name string,salary int, dept string) row format delimited fields terminated by '\t';

load data local inpath '/home/acadgild/nikidir/A1\_emp\_dataset.txt' overwrite into table emp\_dept;

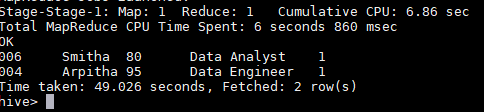


**Q. Get a list of employees who receive a salary less than 100, compared**

**to their immediate employee with higher salary in the same unit**

select id,name,salary,dept, rank from ( SELECT id, name, salary, dept, rank() OVER (PARTITION BY dept ORDER BY salary desc) as rank FROM emp\_dept where salary<100 ) temp where rank=1;

set yarn.nodemanager.resource.memory-mb=49152;



List of all employees who draw higher salary than the average salary

of that department

select id,name,salary,dept, avgsalary from ( SELECT id, name, salary, dept, AVG(salary) OVER (PARTITION BY dept ) as avgsalary FROM emp\_dept ) temp where salary> avgsalary;

