SESSION 8 ASSIGNMENT 1

Problem 1: Get a list of employees who receive a salary less than 100, compared to their immediate employee with higher salary in the same unit.

CREATE DATABSE session\_8;

USE session\_8;

--CREATING TABLE employees in which I will put input data and employees\_salary where I will put the data obtained after using rank query.

CREATE TABLE employees(

id int,

name string,

salary int,

unit string)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY '\t';

LOAD DATA LOCAL INPATH 'input' INTO TABLE employees;

CREATE TABLE employees\_salary(

id int,

name string,

salary int,

unit string,

rnk int) ROW FORMAT DELIMITED

FIELDS TERMINATED BY '\t'

tblproperties(“skip.header.line.count” = “1”);

--Assigning rank unit wise in ascending order of salary.

INSERT OVERWRITE TABLE employees\_salary

SELECT id, name, salary, unit , dense\_rank() over (partition by unit order by salary) rnk

from employees;

SET hive.auto.convert.join = false;

--Comparing employee with the subsequent employee having a higher salary in same unit

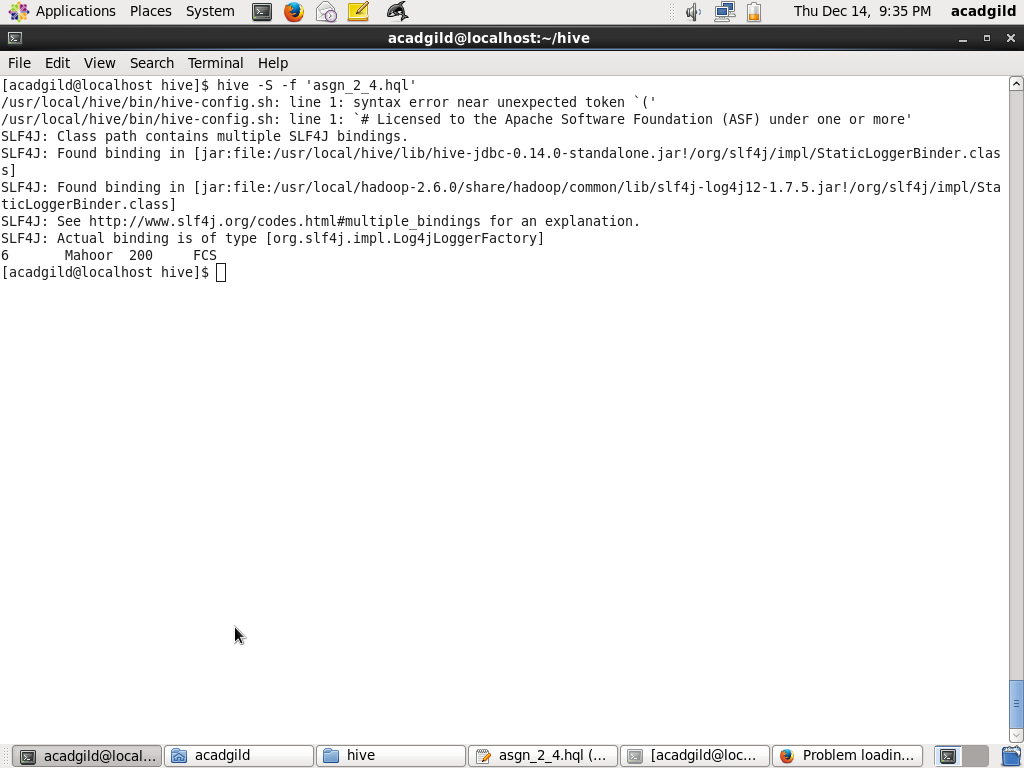
SELECT T1.id, T1.name, T1.salary, T1.unit

from employees\_salary T1

LEFT JOIN employees\_salary T2

ON T1.unit = T2.unit and (T1.rnk+1) = (T2.rnk)

where (T2.salary-T1.salary) >100;



Problem 2: List of all employees who draw higher salary than the average salary of that department.

--Using the same employees table which was used in Problem 1.

--Creating a new table employees\_salary\_avg so as to store unit and its average salary.

use session\_7;

set hive.auto.convert.join = false;

create table employees\_salary\_avg

(unit string,

salary int)

row format delimited

fields TERMINATED by '\t'

tblproperties("skip.header.line.count"="1");

--Inserting Into employees\_salary\_avg table.

insert overwrite table employees\_salary\_avg

SELECT unit,avg(salary) salary

FROM employees

GROUP BY unit;

--Selecting only those records whose salary is higher that unit’s average salary.

select e.\*

from employees e

left join employees\_salary\_avg es

on e.unit = es.unit

where e.salary > es.salary;

