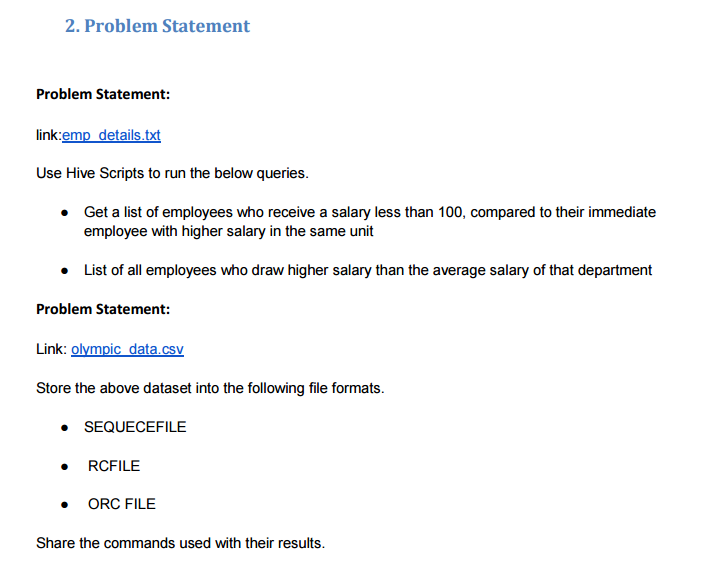
****

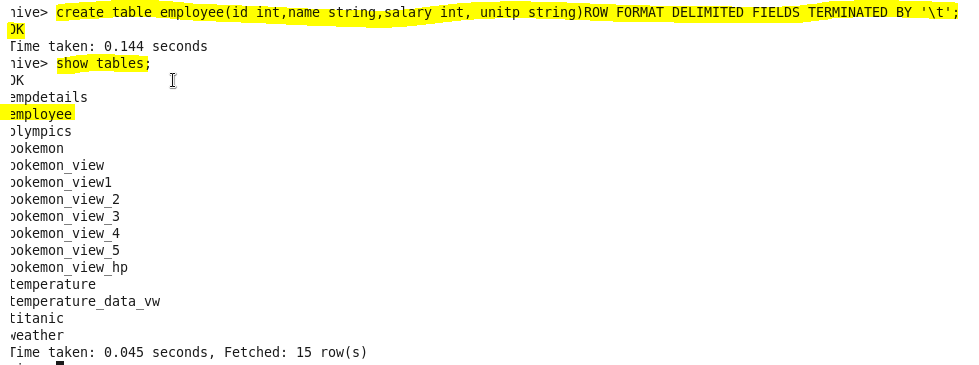
1. Creating table

2. Loading data into the table

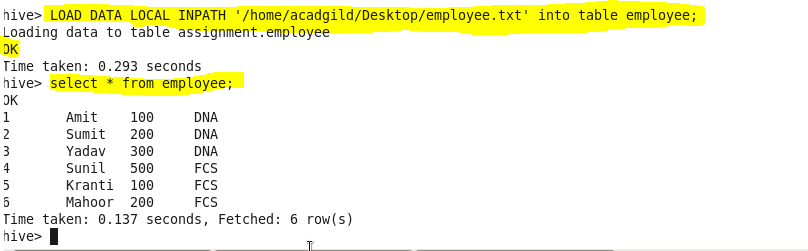
3. Displaying all the data of table

4. Then, according to question finding list of employees who receive a salary less than 100, compared to their immediate employee with higher salary in the same unit using lag and partition keywords.

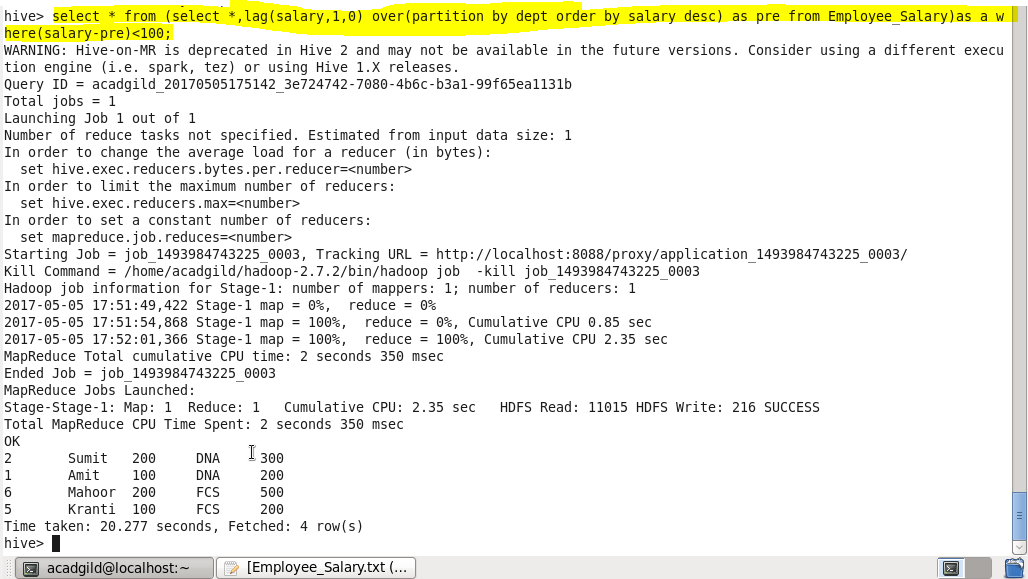
**First we will create table employee using create command**



**Now we will load our dataset employee.txt using LOAD command**

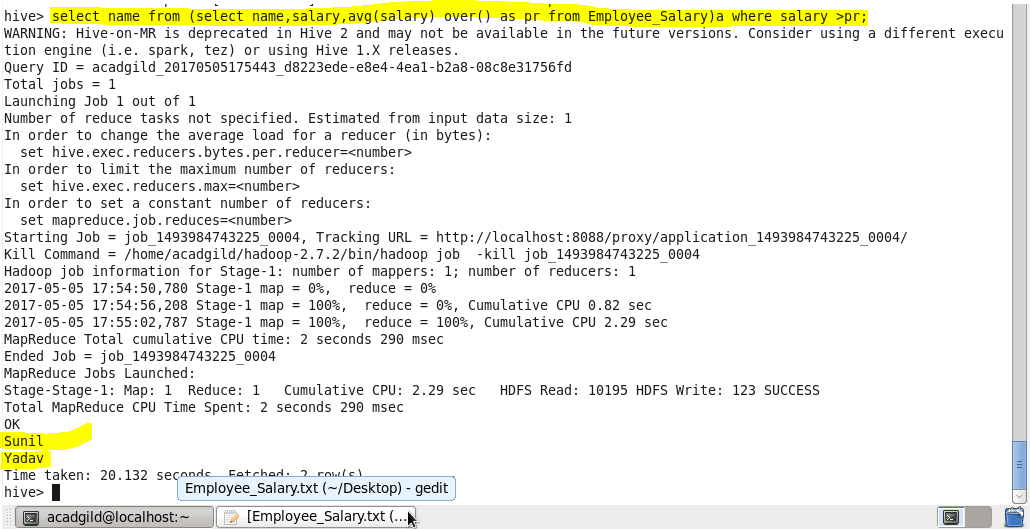


**Q. Get a list of employees who receive a salary less than 100, compared to their immediate employee with higher salary in the same unit**

****

**Q. List of all employees who draw higher salary than the average salary of that department.**

In this particular problem we will check for the employees whose salary is greater than average salary of that department using avg function for that we will select name from a subquery I that we will select name, salary and average salary from the same table and will filter by salary should greater than average salary.



**Problem Statement: Link: olympic\_data.csv Store the above dataset into the following file formats.**

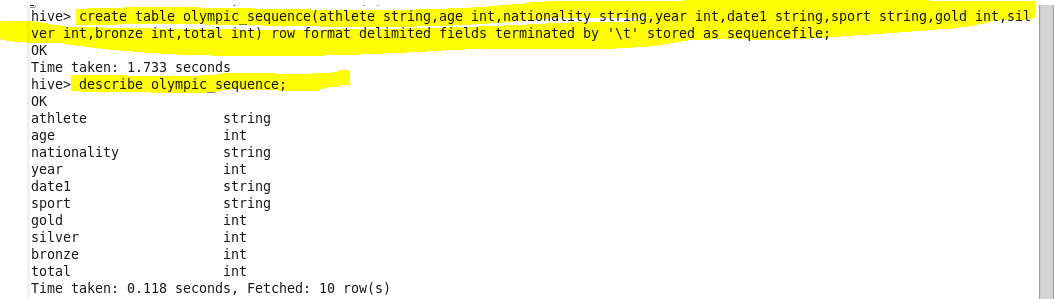
**SEQUENCEFILE**

a) Creating table in **sequencefile format**

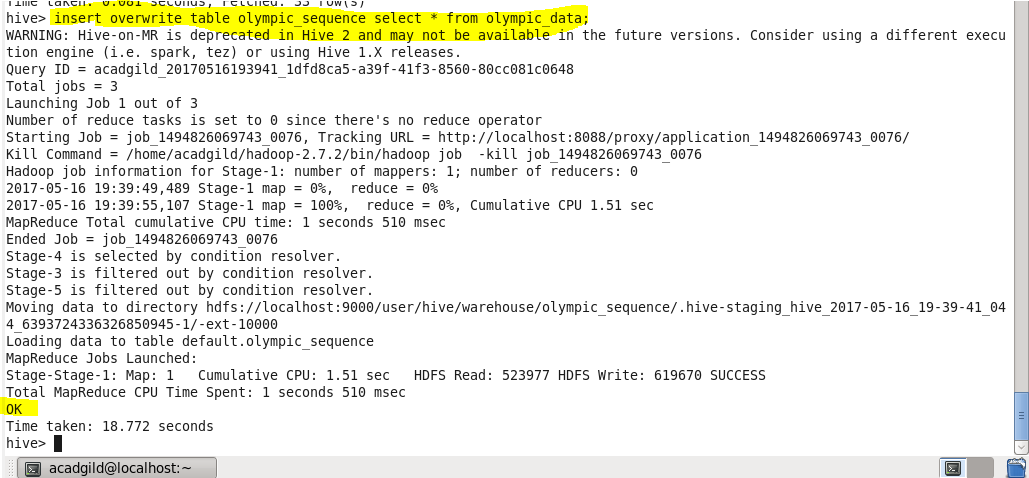
b) Describing the contents of it

c) Loading dataset into file

d) Displaying all the data using ‘select’ keyword.



**Insert data :**



**Result :**



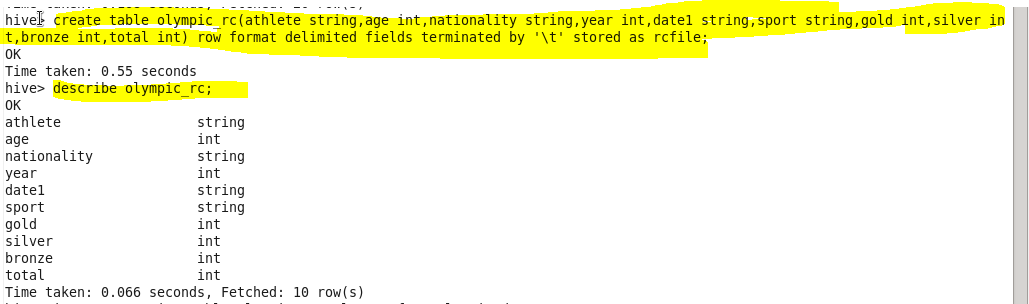
**RCFILE**

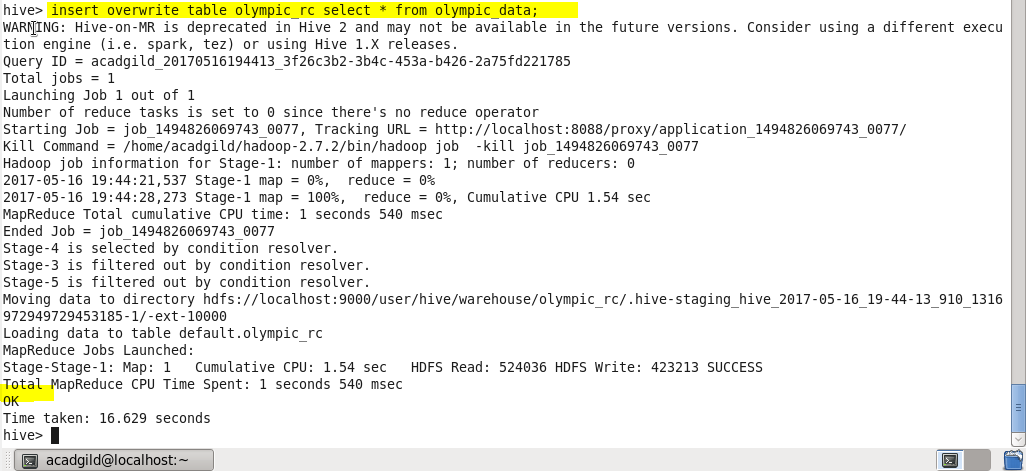
a) Creating table in rcfile format.

b) Describing the contents of it.

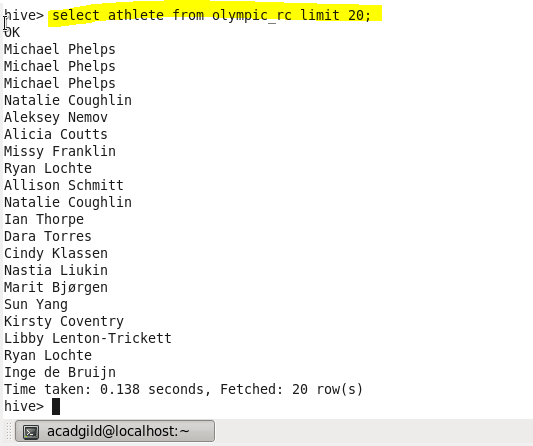
c) Inserting the data into created table.

d) Displaying all the data using ‘select’ keyword.





**Result**

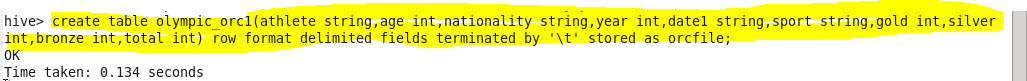


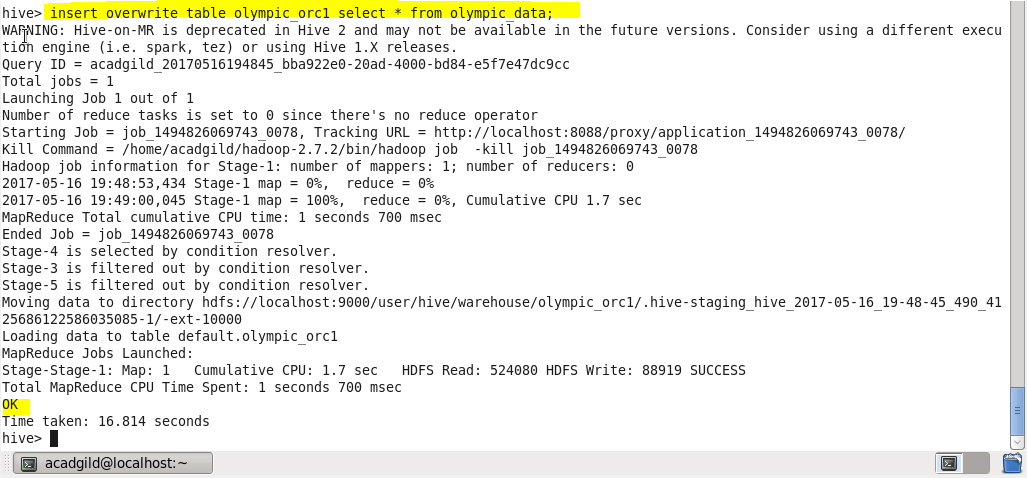
**ORC FILE**

a) Creating table in orcfile format.

b) Inserting the data into created table.

c) Displaying all the data using ‘select’ keyword.





**Result:**

