**Session 28**

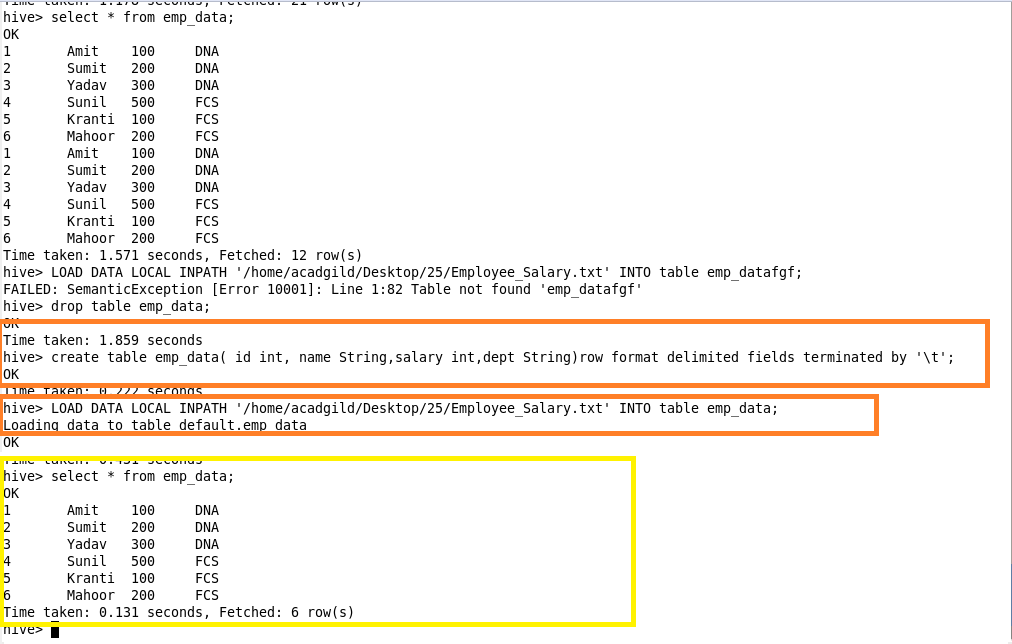
**Assignment 3**

**Part 1:** Problem Statement:

**Link:** emp\_details.txt

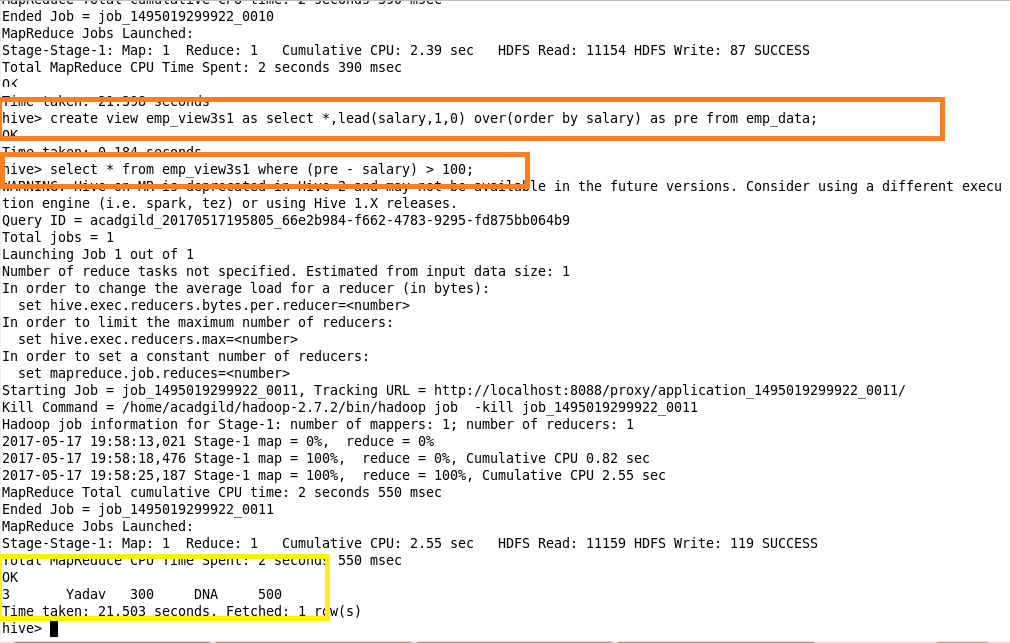
Use Hive Scripts to run the below queries.

**Step 1:** Creating table and loading the data from the local host into the table.



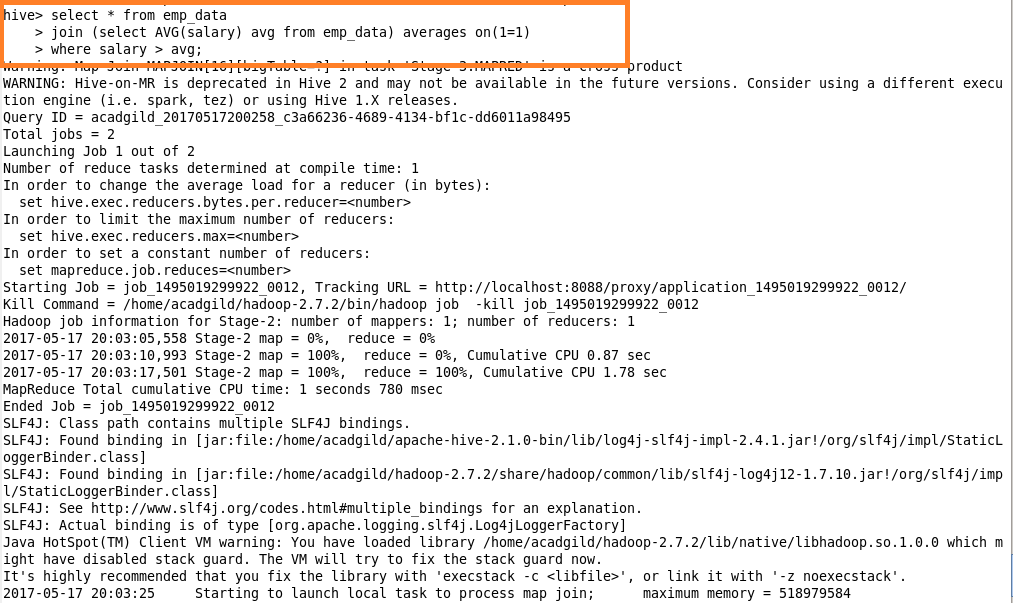
Get a list of employees who receive a salary less than 100, compared to their immediate employee with higher salary in the same unit

**Step 2:** Creating a view and storing the result of employees who receive a salary less than 100.

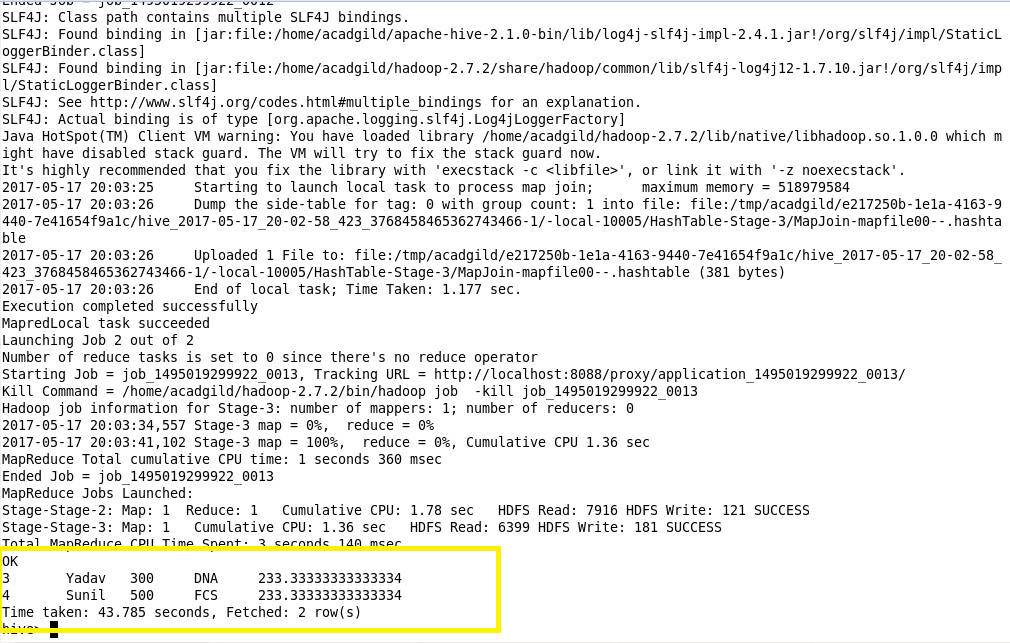


List of all employees who draw higher salary than the average salary of that department

**Step 3:** Selecting the list of top employees from the table whose salary is greater than the average salary of corresponding departments.



**Step 4:** Displaying the output.

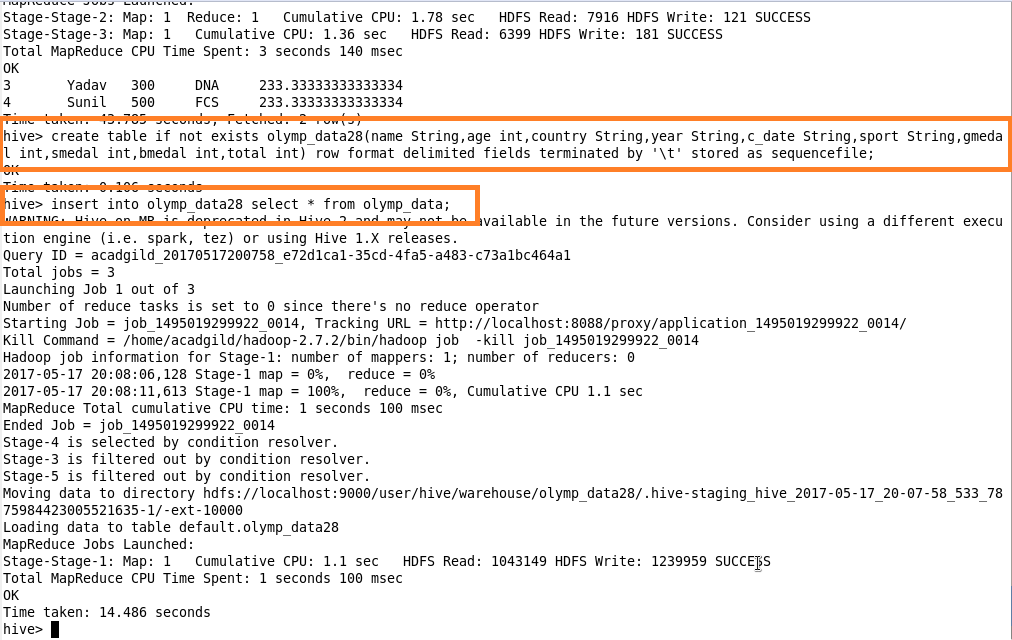


**Part 2:** Problem Statement:

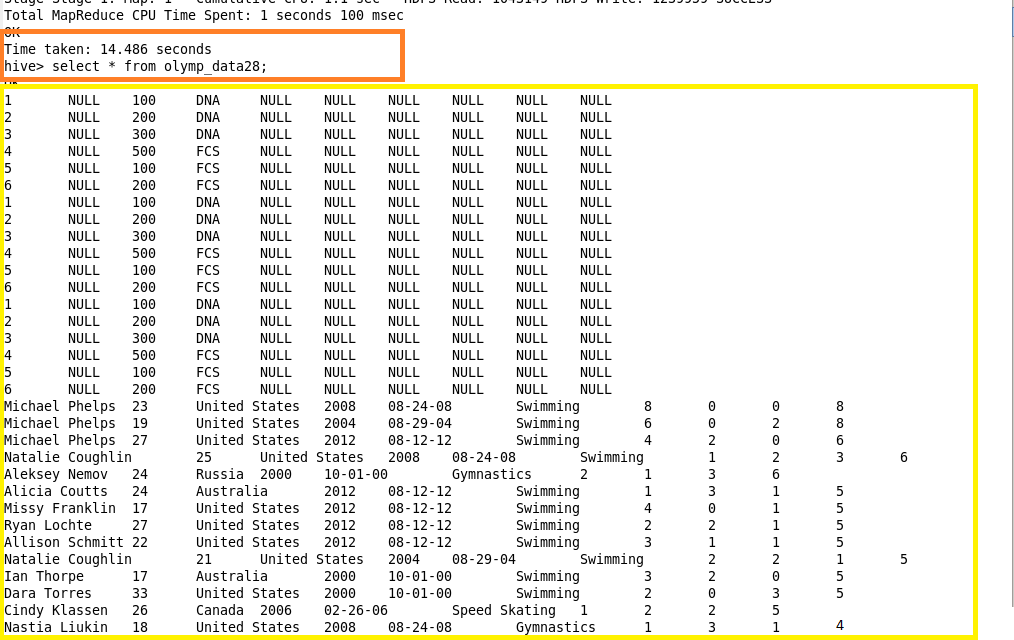
**Link:** olympic\_data.csv

Store the above dataset into the following file formats.

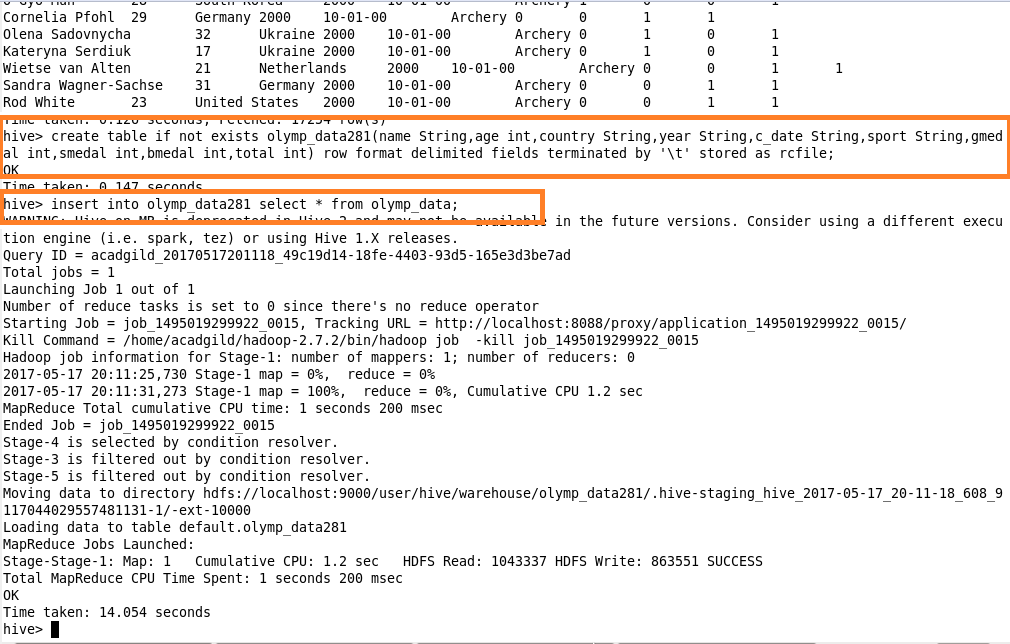
1. **SEQUECEFILE**



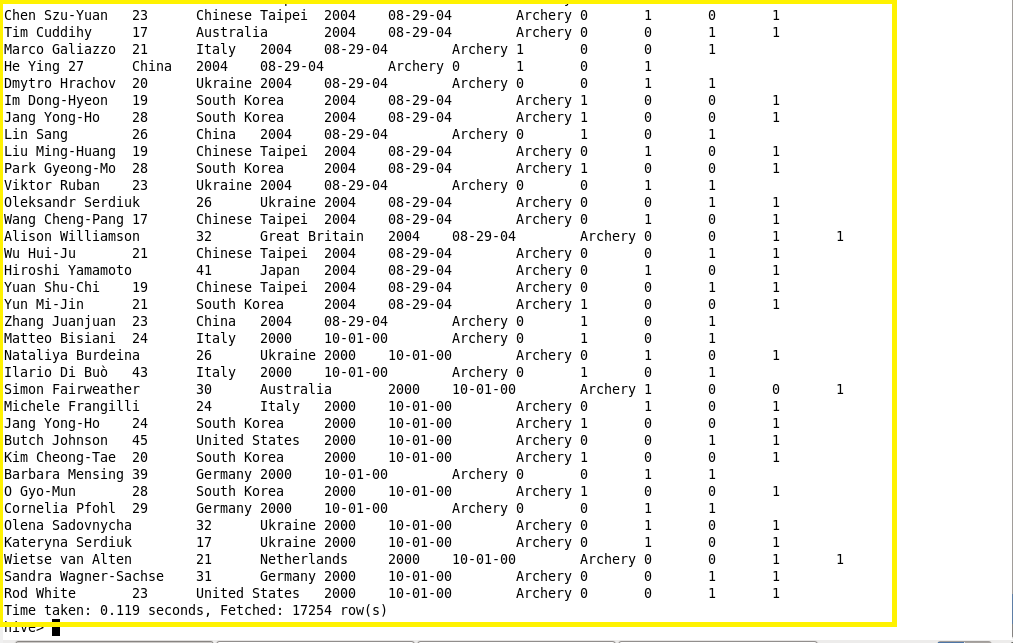
**Output:**



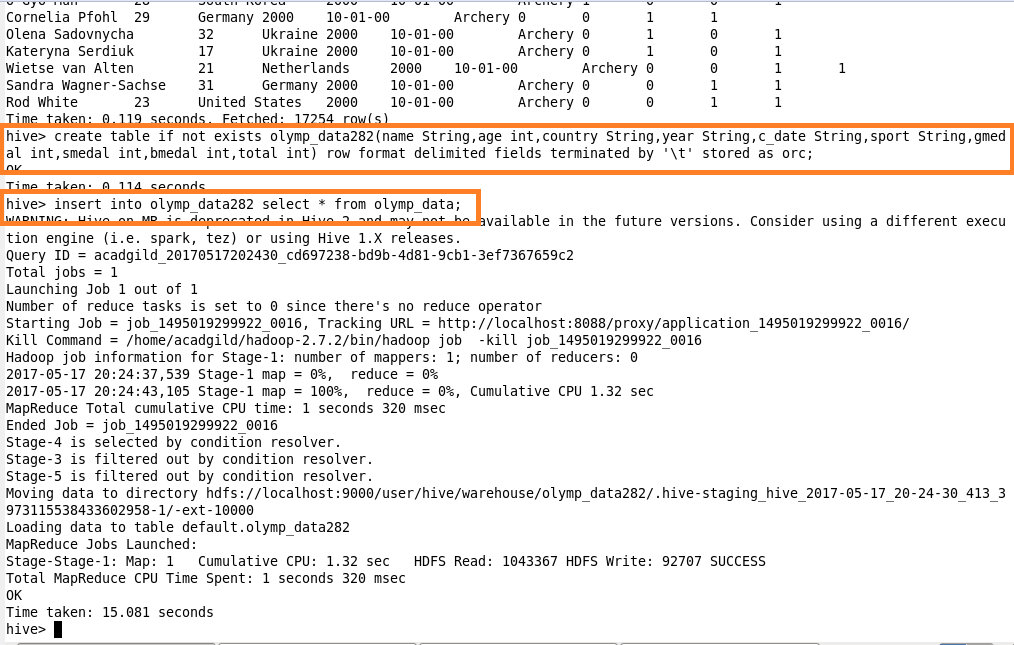
1. **RCFILE**



**Output:**



1. **ORC FILE**



**Output:**

