**ASSIGNMENT - 28.3**

**Problem Statement:**

**Problem Statement 1**:

Link:<https://github.com/prateekATacadgild/DatasetsForCognizant/blob/master/emp_details.txt>

Use Hive Scripts to run the below queries.

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

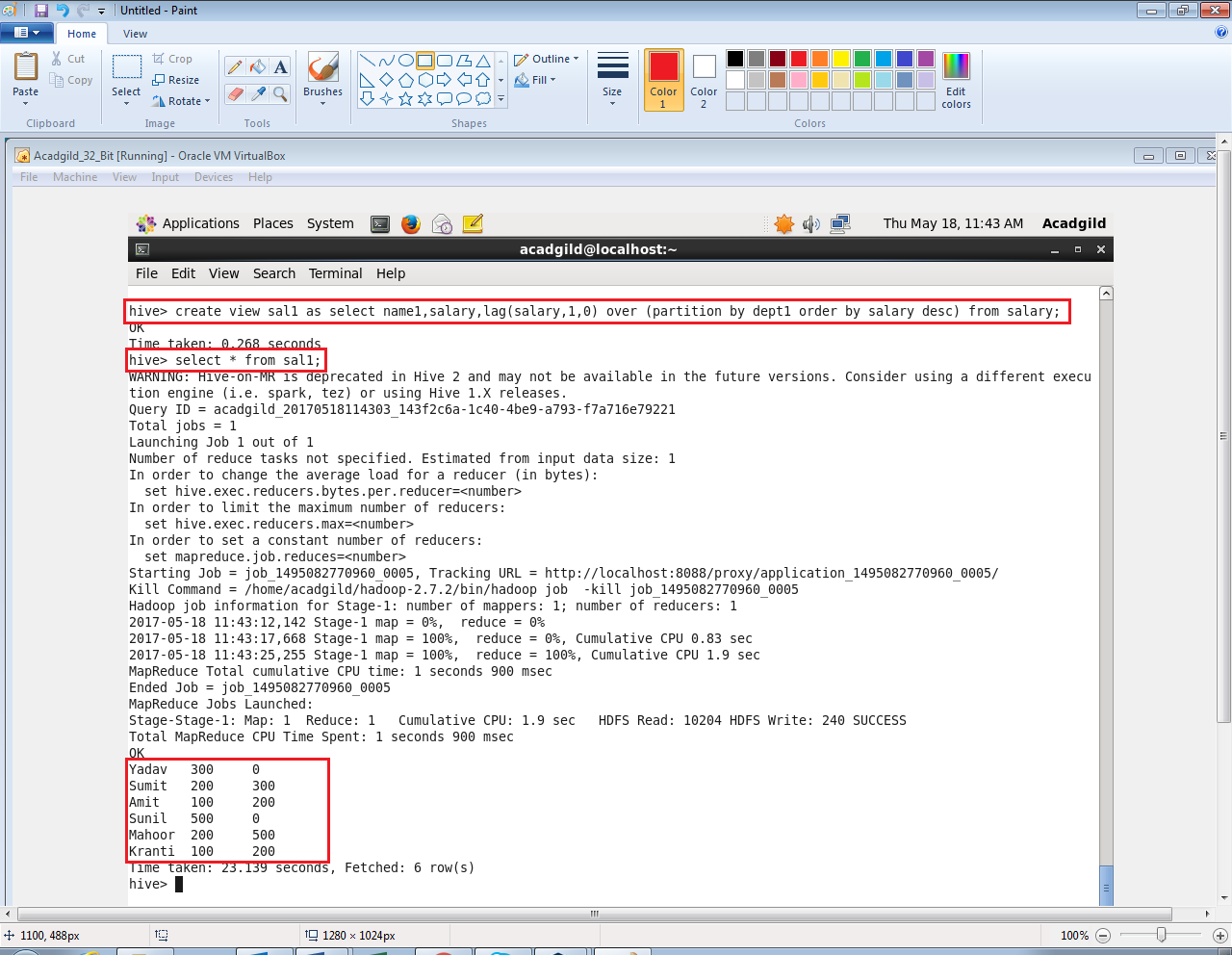
● List of all employees who draw higher salary than the average salary of that department.

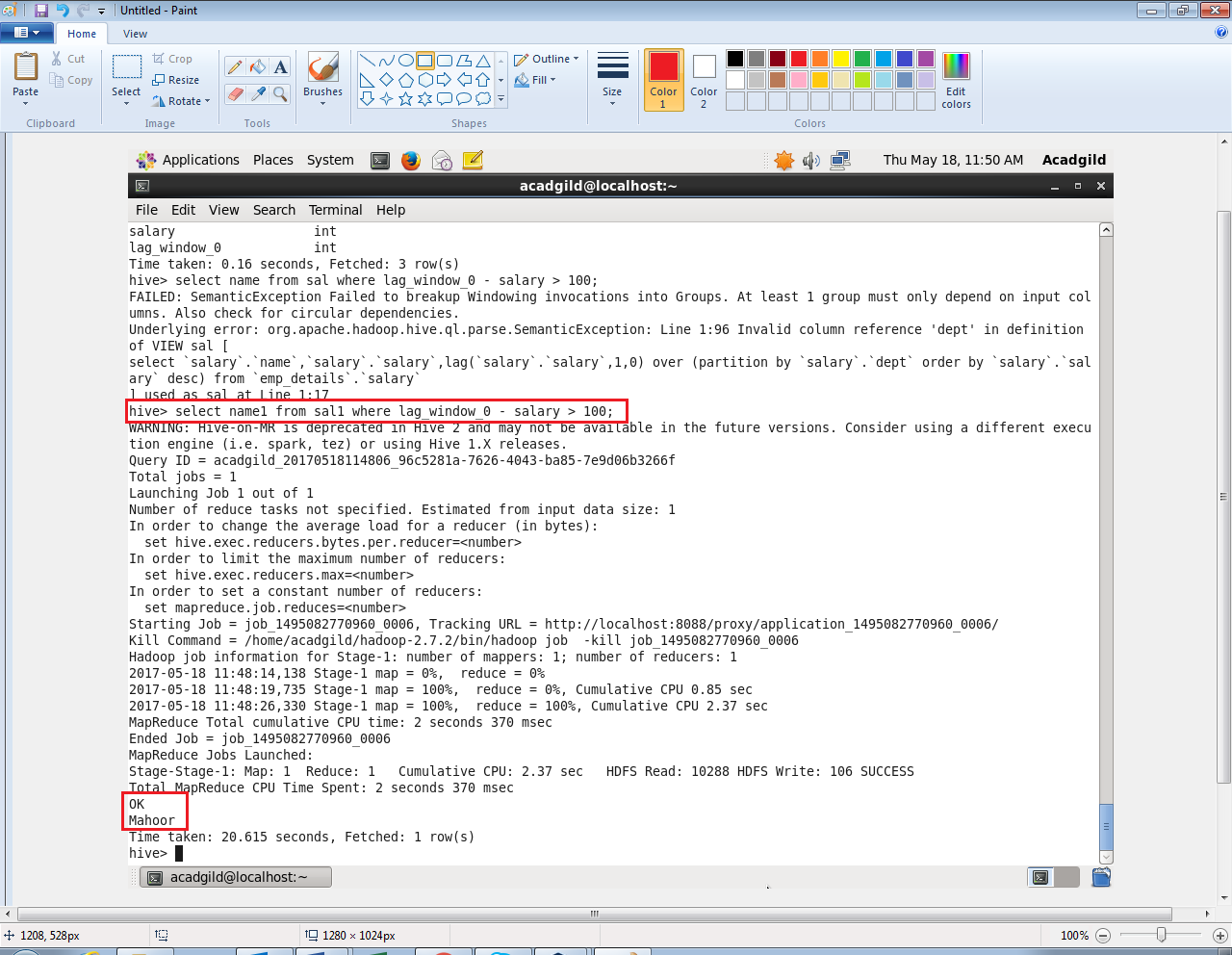
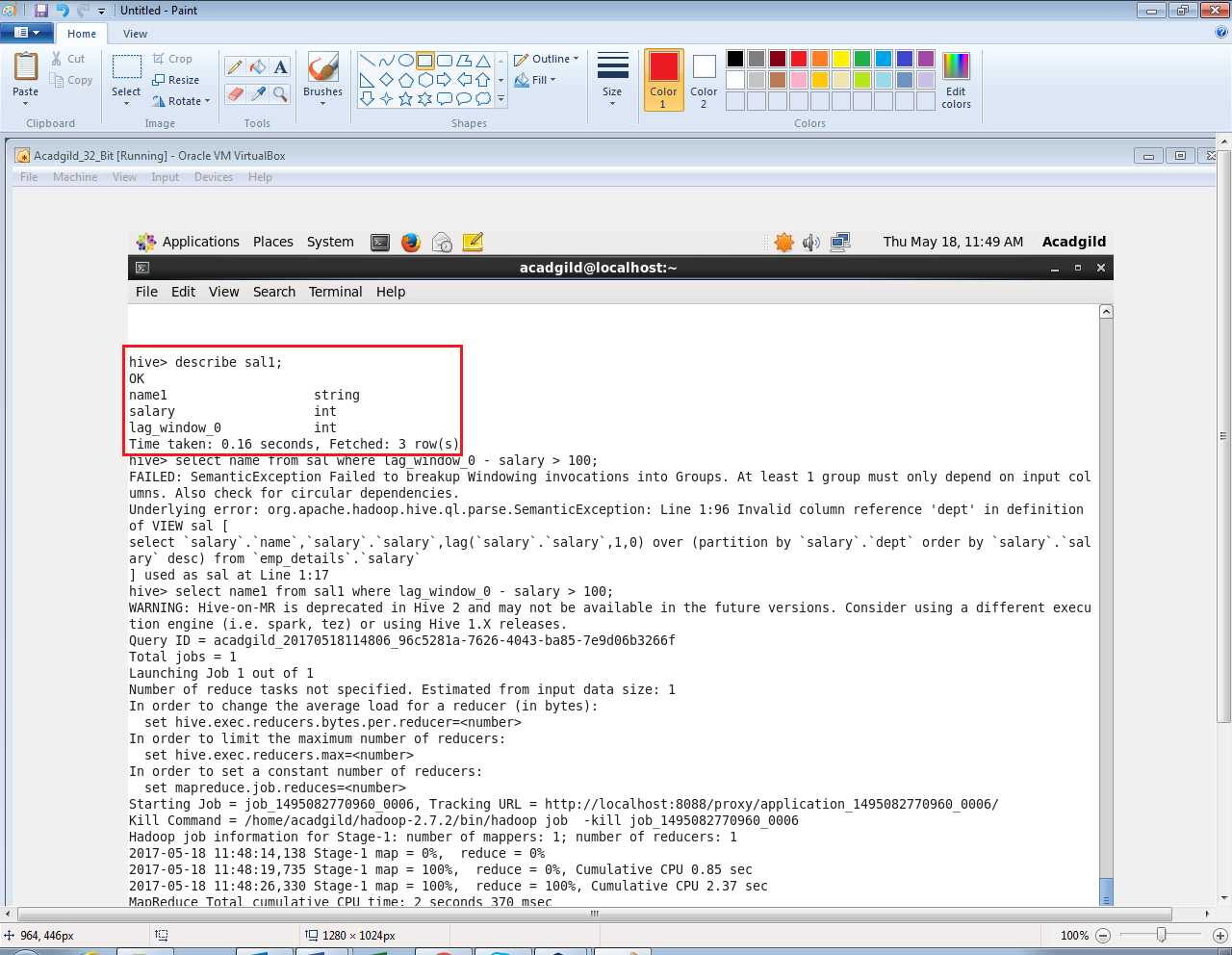
**Solution:**

**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.

**Output:**

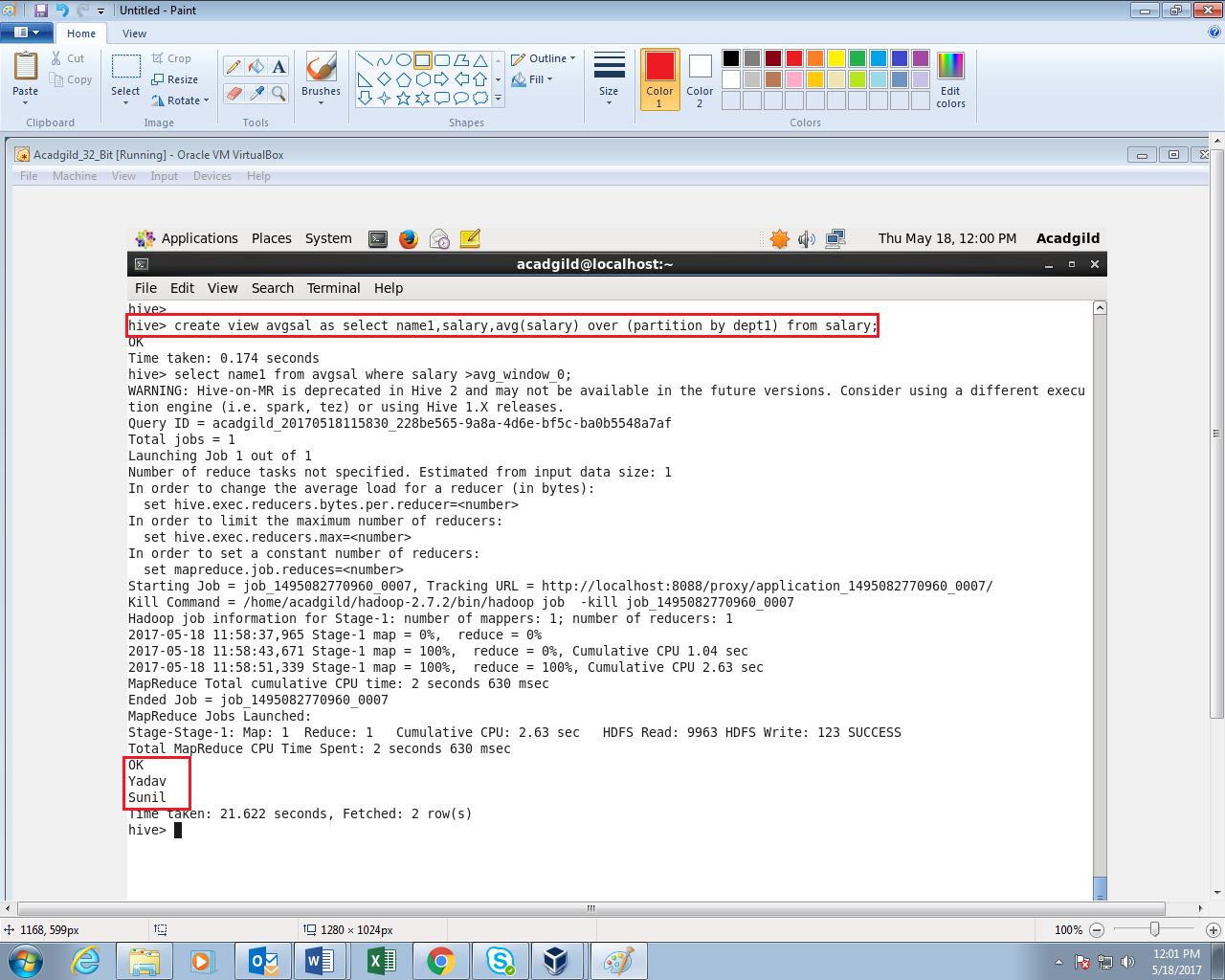




**Problem 2:**

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

**Output:**



**Problem Statement 2:**

Link:<https://github.com/prateekATacadgild/DatasetsForCognizant/blob/master/olympic_data.csv>

Store the above dataset into the following file formats.

● SEQUECEFILE

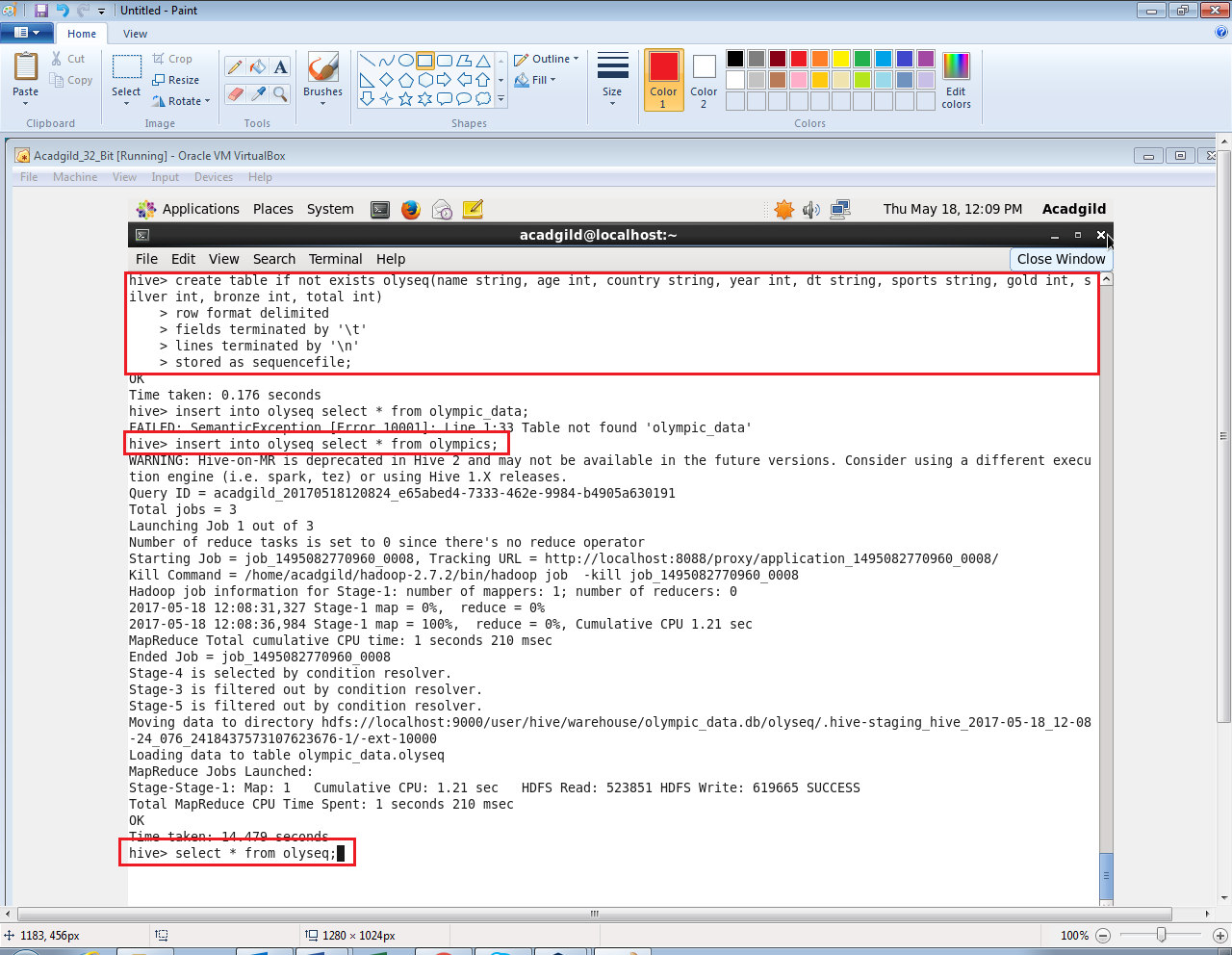
● RCFILE

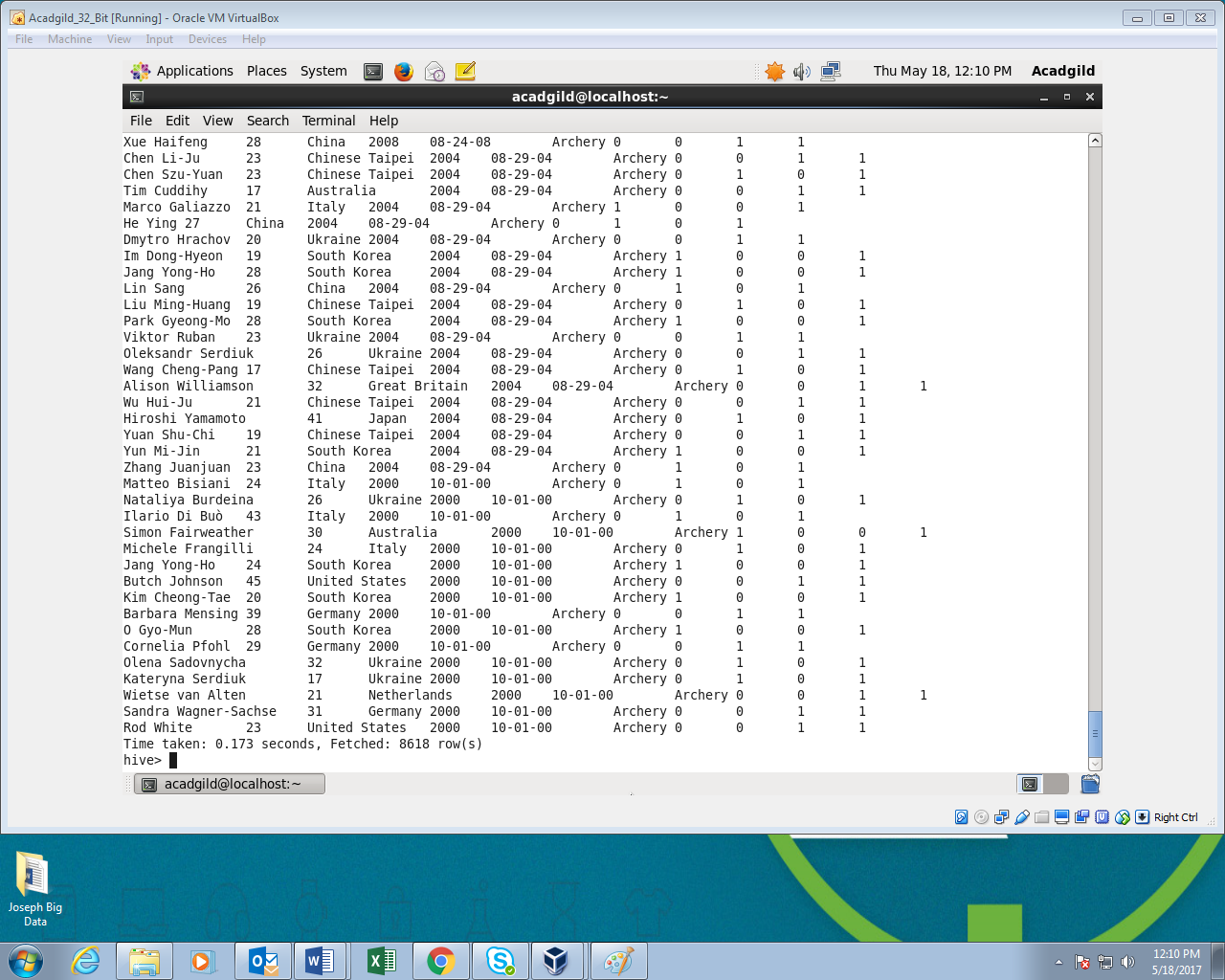
● ORC FILE

Share the commands used with their results.

**Solution:**

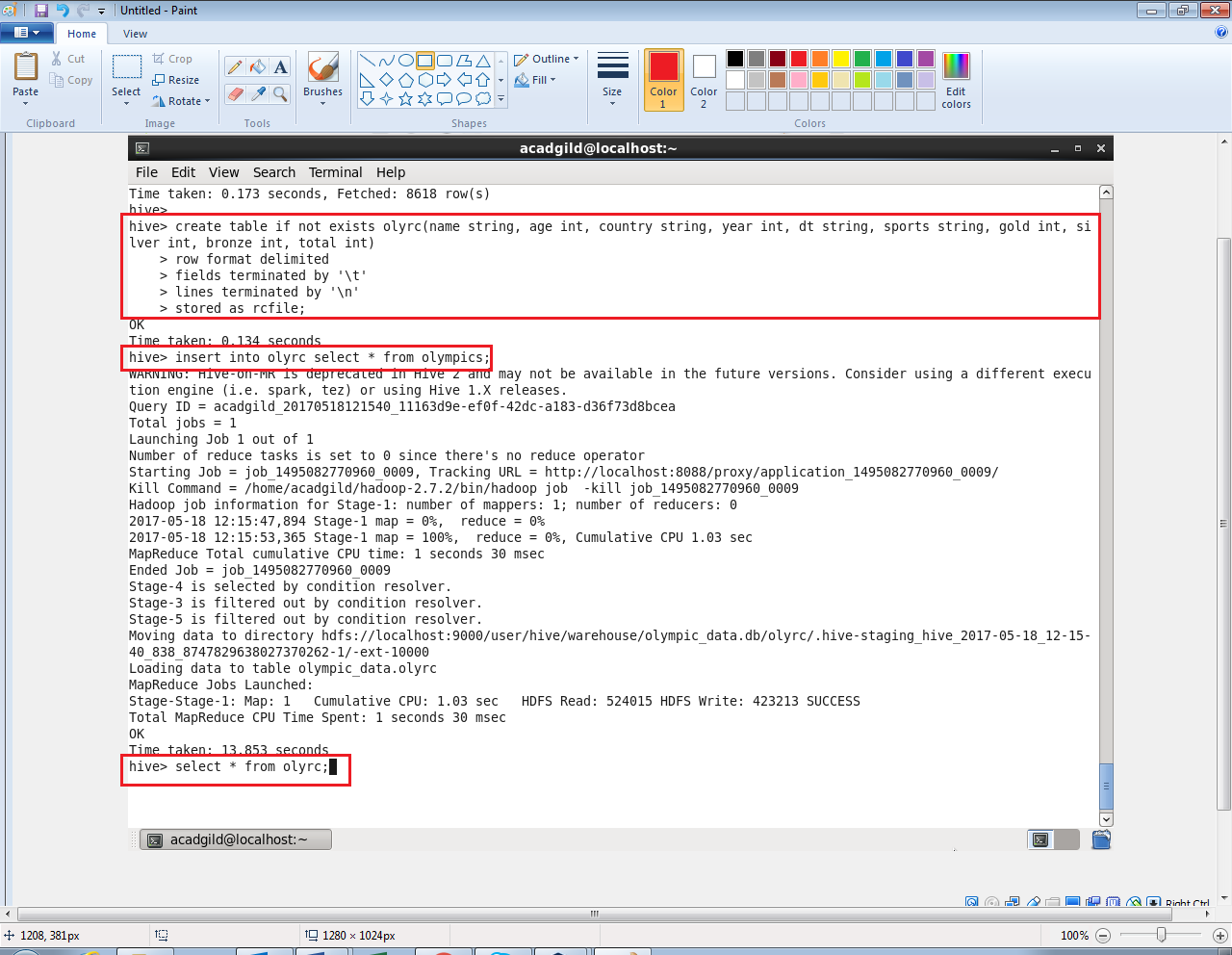
1. **SEQUENCEFILE**:

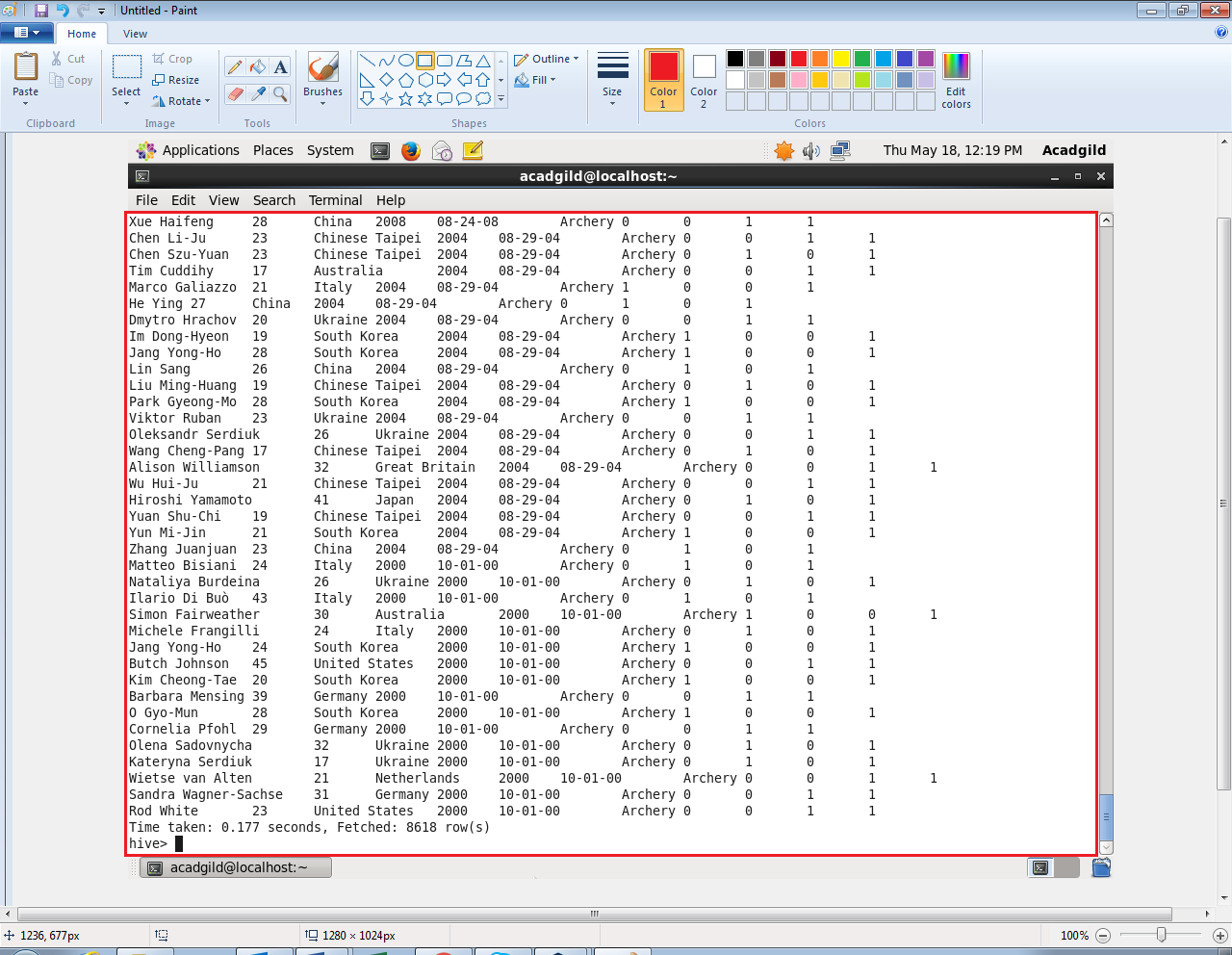




**Since the records obtained are very large, I’m attaching the end of the result which includes how rows have been fetched.**

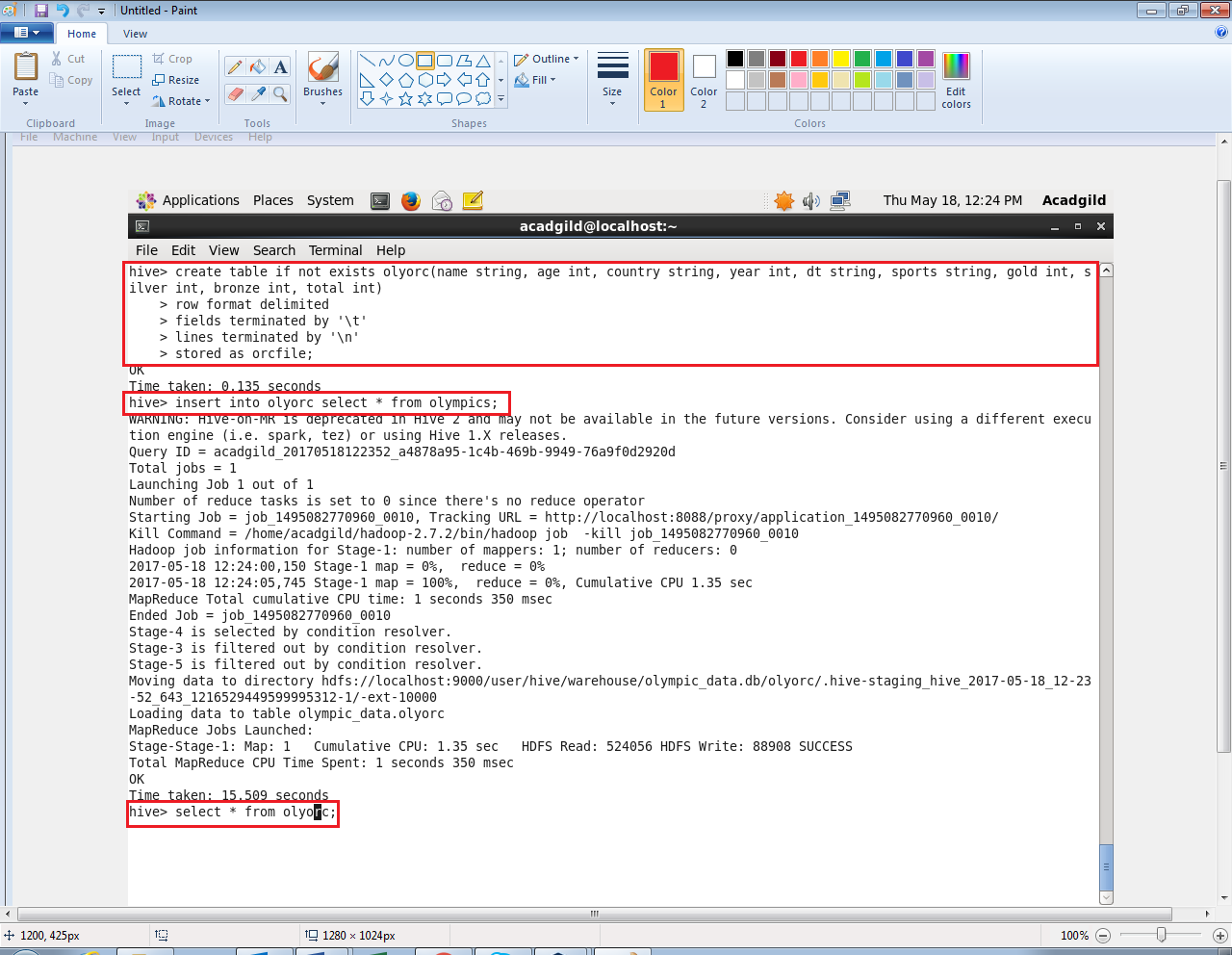
1. **RCFILE:**

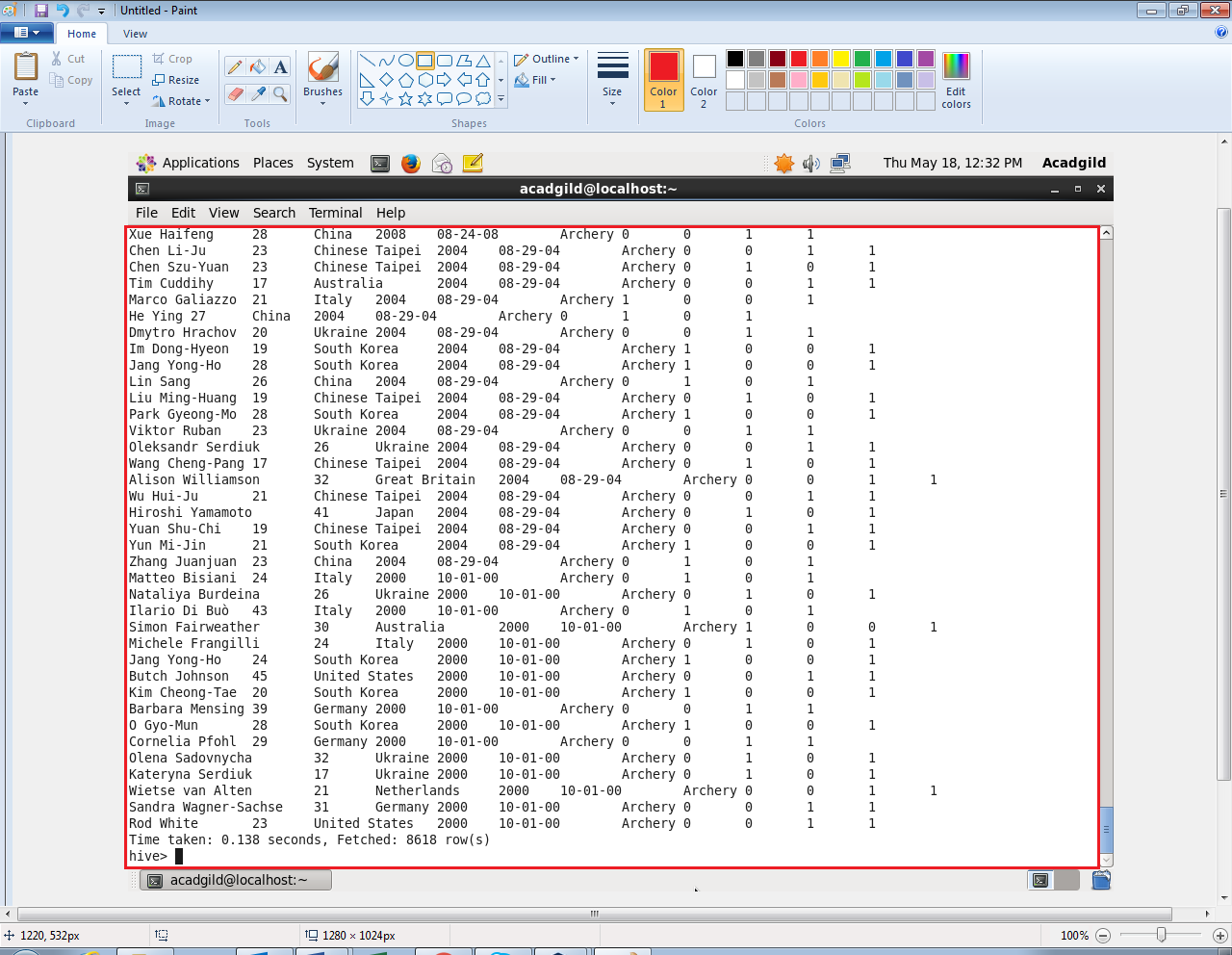




**Since the records obtained are very large, I’m attaching the end of the result which includes how rows have been fetched.**

1. **ORC FILE:**





**Since the records obtained are very large, I’m attaching the end of the result which includes how rows have been fetched.**