Assignment 8.1

For the given dataset Emp_Sal find

1. Get the list of employees who receive a salary less than 100, compared to their immediate employee with higher salary in the same unit.
2. List all employees who receive higher salary more than the average salary of their department.

The given Dataset is as follows

```
       Edit    View    Search    Terminal    Help
hive> select * from emp_sal;
OK
1          Amit      105        Data Minning
2          Pankaj    85         Data Engineer
3          Kiran     110        Data Scientist
4          Arpitha   95         Data Engineer
5          Viraj     105        Data Mining
6          Smitha    80         Data Analyst
7          Supriya   90         Data Engineer
8          Vihan     120        Data Scientist
9          Emma      100        Data Engineer
10         Siddharath         100        Data Engineer
Time taken: 0.882 seconds, Fetched: 10 row(s)
hive>
```

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

Following query is fired to find the employees who has salary less than 100.

```
FAILED: SemanticException [Error 10128]: Line 1:05 Not yet supported place for UDAF max
hive> select eid,ename,esal,eunit from emp_sal group by eid,ename,esal,eunit having esal < 100;
Query ID = acadgild_20180210021919_182f9bb2-5af8-4cab-a687-b29b362f6972
Total jobs = 1
Launching Job 1 out of 1
```

```
Total MapReduce CPU Time Spent: 17 seconds 400 msec
OK
2       Pankaj  85      Data Engineer
4       Arpitha 95      Data Engineer
6       Smitha  80      Data Analyst
7       Supriya 90      Data Engineer
Time taken: 268.222 seconds, Fetched: 4 row(s)
hive> ▮
```

2. **List all employees who receive higher salary more than the average salary of their department .**

To get the desired output , firstly a temporary table is created similar to data set given using temporary keyword as
> **>create temporary table temp as select * from emp_sal;**
> **Then following query is fired**

```
hive> select ename,esal,eunit from ( select avg(esal) over(partition by eunit) as avg_sal,eid,ename,esal,eunit from emp_sal)
temp where esal>avg_sal;
Query ID = acadgild_20180210153434_1e151cc8-75d2-45d3-b267-ace660a3a24b
Total jobs = 1
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 20.8 sec   F
Total MapReduce CPU Time Spent: 20 seconds 800 msec
OK
Siddharath      100     Data Engineer
Emma    100     Data Engineer
Arpitha 95      Data Engineer
Vihan   120     Data Scientist
Time taken: 201.472 seconds, Fetched: 4 row(s)
hive> ▮
```