**Assignment 8.1**

hive> **create table if not exists emp**

(

emp\_id string,

emp\_name string,

salary int,

unit string

)

row format delimited fields terminated by '\t';

OK

Time taken: 0.689 seconds

hive>

hive> **LOAD DATA LOCAL INPATH '/home/acadgild/hive/Datafile.txt'INTO TABLE emp;**

Loading data to table custom.emp

Table custom.emp stats: [numFiles=1, totalSize=291]

OK

Time taken: 2.611 seconds

hive>

* **TASK 1**

hive> **select emp\_name, salary, lead(emp\_name) over (partition by unit order by salary), lead(salary) over (partition by unit order by salary) from emp where salary < 100;**

Query ID = acadgild\_20170924140202\_982fced5-fe3f-44e6-8de9-c2ee387a7497

Total jobs = 1

Launching Job 1 out of 1

Number of reduce tasks not specified. Estimated from input data size: 1

In order to change the average load for a reducer (in bytes):

set hive.exec.reducers.bytes.per.reducer=<number>

In order to limit the maximum number of reducers:

set hive.exec.reducers.max=<number>

In order to set a constant number of reducers:

set mapreduce.job.reduces=<number>

Starting Job = job\_1506241398427\_0001, Tracking URL = http://localhost:8088/proxy/application\_1506241398427\_0001/

Kill Command = /home/acadgild/hadoop-2.6.0/bin/hadoop job -kill job\_1506241398427\_0001

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2017-09-24 14:03:32,054 Stage-1 map = 0%, reduce = 0%

2017-09-24 14:03:54,270 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.98 sec

2017-09-24 14:04:18,053 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 7.38 sec

2017-09-24 14:04:19,400 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 8.95 sec

MapReduce Total cumulative CPU time: 8 seconds 950 msec

Ended Job = job\_1506241398427\_0001

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 8.95 sec HDFS Read: 513 HDFS Write: 76 SUCCESS

Total MapReduce CPU Time Spent: 8 seconds 950 msec

OK

**Smitha 80 NULL NULL**

**Pankaj 85 Supriya 90**

**Supriya 90 Arpitha 95**

**Arpitha 95 NULL NULL**

Time taken: 85.973 seconds, Fetched: 4 row(s)

* **TASK 2**

hive> **select emp\_id, emp\_name, salary, unit, avg\_sal from (select emp\_id, emp\_name, salary, unit,AVG(salary) over (partition by unit order by salary ROWS BETWEEN UNBOUNDED PRECEDING AND UNBOUNDED FOLLOWING) as avg\_sal from emp) emp2 where salary > avg\_sal ;**

Query ID = acadgild\_20170924142727\_60dae807-c1d1-4eff-8f36-e7a91e23399a

Total jobs = 1

Launching Job 1 out of 1

Number of reduce tasks not specified. Estimated from input data size: 1

In order to change the average load for a reducer (in bytes):

set hive.exec.reducers.bytes.per.reducer=<number>

In order to limit the maximum number of reducers:

set hive.exec.reducers.max=<number>

In order to set a constant number of reducers:

set mapreduce.job.reduces=<number>

Starting Job = job\_1506241398427\_0004, Tracking URL = http://localhost:8088/proxy/application\_1506241398427\_0004/

Kill Command = /home/acadgild/hadoop-2.6.0/bin/hadoop job -kill job\_1506241398427\_0004

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2017-09-24 14:27:47,602 Stage-1 map = 0%, reduce = 0%

2017-09-24 14:28:02,935 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.7 sec

2017-09-24 14:28:22,808 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.2 sec

MapReduce Total cumulative CPU time: 7 seconds 200 msec

Ended Job = job\_1506241398427\_0004

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.2 sec HDFS Read: 513 HDFS Write: 139 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 200 msec

OK

**004 Arpitha 95 Data Engineer 94.0**

**010 Siddharath 100 Data Engineer 94.0**

**009 Emma 100 Data Engineer 94.0**

**008 Vihan 120 Data Scientist 115.0**

Time taken: 55.527 seconds, Fetched: 4 row(s)

hive>