

### Assignment-10.3

We have employee\_details and employee\_expenses files. Use local mode while running Pig and write Pig Latin script to get below results:

- (a) Top 5 employees (employee id and employee name) with highest rating. (In case two employees have same rating, employee with name coming first in dictionary should get preference)

```
A = load '/home/acadgild/emp_details.txt' using PigStorage(',') as (id:int,name:chararray,salary:long,rating:int);
B = ORDER A by rating DESC,name;
C = limit B 5;
D = foreach C generate id,name;
Store D into '/home/acadgild/10.3a/';
```

```
[acadgild@localhost ~]$ pig -x local /home/acadgild/10.3.1.pig
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/hbase-1.0.3/lib/slf4j-log4j12-1.7.7.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/hadoop-2.7.2/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
```

```
[acadgild@localhost ~]$ ls /home/acadgild/10.3a/
part-r-000000 SUCCESS
[acadgild@localhost ~]$ cat /home/acadgild/10.3a/part-r-000000
105 Pawan
110 Priyanka
104 Anubhav
109 Katrina
103 Akshay
[acadgild@localhost ~]$ █
```

- (b) Top 3 employees (employee id and employee name) with highest salary, whose employee id is an odd number. (In case two employees have same salary, employee with name coming first in dictionary should get preference)

```
A = load '/home/acadgild/emp_details.txt' using PigStorage(',') as (id:int,name:chararray,salary:long,rating:int);
B = filter A by (id%2==1);
C = ORDER B by salary DESC,name;
D = limit C 3;
E = foreach D generate id,name;
Store E into '/home/acadgild/10.3b/';
```

```
[acadgild@localhost ~]$ pig -x local /home/acadgild/10.3.2.pig
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/hbase-1.0.3/lib/slf4j-log4j12-1.7.7.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/hadoop-2.7.2/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
```

```
[acadgild@localhost ~]$ ls /home/acadgild/10.3b/
part-r-000000 SUCCESS
[acadgild@localhost ~]$ cat /home/acadgild/10.3b/part-r-000000
101 Amitabh
107 Salman
103 Akshay
[acadgild@localhost ~]$
```

- (c) Employee (employee id and employee name) with maximum expense (In case two employees have same expense, employee with name coming first in dictionary should get preference)

```
A = load '/home/acadgild/emp_details.txt' using PigStorage(',') as (emp_id:int,name:chararray,salary:long,rating:int);
B = load '/home/acadgild/emp_expenses.txt' as (id:int,exp:int);
C = group B by id;
D = foreach C generate group as id,SUM(B.exp) as sum;
E = JOIN A by emp_id,D by id;
F = foreach E generate emp_id,name,sum;
G = ORDER F by sum DESC;
H = limit G 1;
I = foreach H generate emp_id,name;
Store I into '/home/acadgild/10.3c/';
```

```
[acadgild@localhost ~]$ pig -x local /home/acadgild/10.3.3.pig
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/hbase-1.0.3/lib/slf4j-log4j12-1.7.7.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/hadoop-2.7.2/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
```

```
2017-03-17 17:44:02,831 [main] INFO org.apache.pig.Main - Pig script completed in 8 seconds and 173 milliseconds (8173 ms)
[acadgild@localhost ~]$ ls /home/acadgild/10.3c/
part-r-000000 SUCCESS
[acadgild@localhost ~]$ cat /home/acadgild/10.3c/part-r-000000
102 Shahrukh
[acadgild@localhost ~]$
```

(d) List of employees (employee id and employee name) having entries in employee\_expenses file.

```
A = load '/home/acadgild/emp_details.txt' using PigStorage(',') as (emp_id:int,name:chararray,salary:long,rating:int);
B = load '/home/acadgild/emp_expenses.txt' as (id:int,exp:int);
C = JOIN A by emp_id,B by id;

D = foreach C generate emp_id,name;
E = DISTINCT D;
Store E into '/home/acadgild/10.3D/';
```

```
acadgild@localhost ~]$ pig -x local /home/acadgild/10.3.4.pig
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/hbase-1.0.3/lib/slf4j-log4j12-1.7.7.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/hadoop-2.7.2/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
```

```

[acadgild@localhost ~]$ ls /home/acadgild/10.3D/
part-r-000000 SUCCESS
[acadgild@localhost ~]$ cat /home/acadgild/10.3D/part-r-000000
101 Amitabh
102 Shahrukh
104 Anubhav
105 Pawan
110 Priyanka
114 Madhuri
[acadgild@localhost ~]$ █

```

- (e) List of employees (employee id and employee name) having no entry in employee\_expenses file.

```

A = load '/home/acadgild/emp_details.txt' using PigStorage(',') as (emp_id:int,name:chararray,salary:long,rating:int);
B = load '/home/acadgild/emp_expenses.txt' as (id:int,exp:int);
C = JOIN A by emp_id LEFT OUTER,B by id;
D = filter C by B::id is null;
E = foreach D generate A::emp_id,name;
F = DISTINCT E;
store F into '/home/acadgild/10.3e/';|

```

```

[acadgild@localhost ~]$ pig -x local /home/acadgild/10.3.5.pig
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/hbase-1.0.3/lib/slf4j-log4j12-1.7.7.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/hadoop-2.7.2/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]

```

```

[acadgild@localhost ~]$ ls /home/acadgild/10.3e/
part-r-000000 SUCCESS
[acadgild@localhost ~]$ cat /home/acadgild/10.3e/part-r-000000
103 Akshay
106 Aamir
107 Salman
108 Ranbir
109 Katrina
111 Tushar
112 Ajay
113 Jubeen
[acadgild@localhost ~]$ █

```