**Assignment 4.2**

1. /\*Load file/\*

emp\_details = LOAD '/home/acadgild/Desktop/employee.txt' USING PigStorage(',') AS (id:int, name:chararray, age:int, city:chararray, point:int);

output:

(101,Amitabh,23,Delhi,101)

(102,Shahrukh,25,Mumbai,102)

(103,Akshay,26,Kolkata,123)

(104,Anubhav,23,Pune,125)

(105,Pawan,42,Chennai,464)

(106,Aamir,27,Hyderabad,466)

(107,Salman,22,Gurgaon,744)

(108,Ranbir,25,Rohtak,184)

(109,Katrina,27,Bahadurgarh,375)

-------------------------------------------------------------------------

2.

/\* CONCAT \*/

emp\_concat = foreach emp\_details Generate CONCAT (name ,' Is From ', city);

output:

(Amitabh Is From Delhi)

(Shahrukh Is From Mumbai)

(Akshay Is From Kolkata)

(Anubhav Is From Pune)

(Pawan Is From Chennai)

(Aamir Is From Hyderabad)

(Salman Is From Gurgaon)

(Ranbir Is From Rohtak)

(Katrina Is From Bahadurgarh

-------------------------------------------------------------------------

3.

/\* TOKENIZE \*/

emp\_name\_tokenize = foreach emp\_details Generate TOKENIZE(name);

Dump student\_name\_tokenize;

output:

({(Amitabh)})

({(Shahrukh)})

({(Akshay)})

({(Anubhav)})

({(Pawan)})

({(Aamir)})

({(Salman)})

({(Ranbir)})

({(Katrina)})

-------------------------------------------------------------------------

4.

/\* SUM \*/

emp\_group\_all = Group emp\_details all;

emp\_sum\_points = Foreach emp\_group\_all Generate(emp\_details.name,emp\_details.point),SUM(emp\_details.point);

Dump emp\_sum\_points;

output:

(({(Katrina),(Ranbir),(Salman),(Aamir),(Pawan),(Anubhav),(Akshay),(Shahrukh),(Amitabh)},{(375),(184),(744),(466),(464),(125),(123),(102),(101)}),2684)

-------------------------------------------------------------------------

5.

/\* MIN \*/

emp\_sum\_min = FOREACH emp\_group\_all GENERATE(emp\_details.name,emp\_details.point),MIN(emp\_details.point);

Dump emp\_sum\_min;

output:

(({(Katrina),(Ranbir),(Salman),(Aamir),(Pawan),(Anubhav),(Akshay),(Shahrukh),(Amitabh)},{(375),(184),(744),(466),(464),(125),(123),(102),(101)}),101)

-------------------------------------------------------------------------

6.

/\* MAX \*/

emp\_sum\_max = FOREACH emp\_group\_all GENERATE(emp\_details.name,emp\_details.point),MAX(emp\_details.point);

Dump emp\_sum\_max;

output:

(({(Katrina),(Ranbir),(Salman),(Aamir),(Pawan),(Anubhav),(Akshay),(Shahrukh),(Amitabh)},{(375),(184),(744),(466),(464),(125),(123),(102),(101)}),744)

-------------------------------------------------------------------------

7.

/\*LIMIT \*/

emp\_details\_limit = LIMIT emp\_details 3;

Dump emp\_details\_limit;

output:

(101,Amitabh,23,Delhi,101)

(102,Shahrukh,25,Mumbai,102)

(103,Akshay,26,Kolkata,123)

-------------------------------------------------------------------------

8.

/\* STORE \*/

STORE emp\_concat INTO '/home/acadgild/Desktop/out\_file/out/concat' USING PigStorage(',');

STORE emp\_name\_tokenize INTO '/home/acadgild/Desktop/out\_file/out/tokenize' USING PigStorage(',');

-------------------------------------------------------------------------

9.

/\* DISTINCT & FLATTEN\*/

emp\_city\_group = GROUP emp\_details by city;

emp\_city\_ditinct = FOREACH emp\_city\_group{

result =(1,0,$1);

GENERATE FLATTEN(result);

}

DUMP emp\_city\_ditinct;

output:

(1,0,{(104,Anubhav,23,Pune,125)})

(1,0,{(101,Amitabh,23,Delhi,101)})

(1,0,{(102,Shahrukh,25,Mumbai,102)})

(1,0,{(108,Ranbir,25,Rohtak,184)})

(1,0,{(105,Pawan,42,Chennai,464)})

(1,0,{(107,Salman,22,Gurgaon,744)})

(1,0,{(103,Akshay,26,Kolkata,123)})

(1,0,{(106,Aamir,27,Hyderabad,466)})

(1,0,{(109,Katrina,27,Bahadurgarh,375)})

-------------------------------------------------------------------------

10.

/\* IsEmpty \*/

emp\_is\_empty = FILTER emp\_city\_group by not IsEmpty(group);

DUMP emp\_is\_empty;