student\_details = LOAD 'student\_details.txt' USING PigStorage(',')

as (id:int, firstname:chararray, lastname:chararray, age:int, phone:chararray, city:chararray, gpa:int);

CONCAT:-

student\_name\_concat = foreach student\_details Generate CONCAT (firstname, lastname);

Output:-

({(Rajaiv),(Reddy)})

({(siddarth),(Battacharya)})

({(Rajesh),(Khanna)})

({(Preethi),(Agarwal)})

({(Trupthi),(Mohanthy)})

({(Archana),(Mishra)})

({(Komal),(Nayak)})

({(Bharathi),(Nambiayar)})

**TOKENIZE():-**

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

Dump student\_name\_tokenize;

Output:-

({(Rajaiv),(Reddy)})

({(siddarth),(Battacharya)})

({(Rajesh),(Khanna)})

({(Preethi),(Agarwal)})

({(Trupthi),(Mohanthy)})

({(Archana),(Mishra)})

({(Komal),(Nayak)})

({(Bharathi),(Nambiayar)})

SUM:-

student\_group = Group student all;

student\_workpages\_sum = foreach student\_group Generate (student\_data.name,student\_data.daily\_typing\_pages),SUM(student\_data.daily\_typing\_pages);

Output:-

(({ (Zara), (Zara), (Jill) ,(Jack) , (Jack) , (Ram) , (John) },

{ (350) , (300) , (220) ,(100) , (170) , (220) , (250) }),1610)

Min:-

student\_group\_all = Group student\_details All;

student\_gpa\_min = foreach student\_group\_all Generate

(student\_details.firstname, student\_details.gpa), MIN(student\_details.gpa);

Dump student\_gpa\_min;

Output:-

(({(Bharathi),(Komal),(Archana),(Trupthi),(Preethi),(Rajesh),(siddarth),(Rajiv) } ,

{ (72) , (83) , (87) , (75) , (93) , (90) , (78) , (89) }) ,72)

MAX:-

student\_group\_all = Group student\_details All;

student\_gpa\_max = foreach student\_group\_all Generate

(student\_details.firstname, student\_details.gpa), MAX(student\_details.gpa);

Dump student\_gpa\_max;

Output:-

(({(Bharathi),(Komal),(Archana),(Trupthi),(Preethi),(Rajesh),(siddarth),(Rajiv) } ,

{ (72) , (83) , (87) , (75) , (93) , (90) , (78) , (89) }) ,93)

**IsEmpty() :-**

grunt> emp\_sales = LOAD 'emp\_sales.txt' USING PigStorage(',')

as (sno:int, name:chararray, age:int, salary:int, dept:chararray);

grunt> emp\_bonus = LOAD 'mp\_bonus.txt' USING PigStorage(',')

as (sno:int, name:chararray, age:int, salary:int, dept:chararray);

grunt> cogroup\_data = COGROUP emp\_sales by age, emp\_bonus by age;

grunt> isempty\_data = filter cogroup\_data by IsEmpty(emp\_sales);

**Dump isempty\_data;**

Output:-

(30,{},{(6,Omar,30,30000,admin)})

LIMIT():-

limit\_data = LIMIT student\_details 4;

Dump limit\_data;

Output:-

(1,Rajiv,Reddy,21,9848022337,Hyderabad)

(2,siddarth,Battacharya,22,9848022338,Kolkata)

(3,Rajesh,Khanna,22,9848022339,Delhi)

(4,Preethi,Agarwal,21,9848022330,Pune)

Distinct:-

distinct\_data = DISTINCT student\_details;

Dump distinct\_data;

Output:-

(1,Rajiv,Reddy,9848022337,Hyderabad)

(2,siddarth,Battacharya,9848022338,Kolkata)

(3,Rajesh,Khanna,9848022339,Delhi)

(4,Preethi,Agarwal,9848022330,Pune)

(5,Trupthi,Mohanthy,9848022336,Bhuwaneshwar)

(6,Archana,Mishra,9848022335,Chennai)

**STORE:-**

student = LOAD 'hdfs://localhost:9000/pig\_data/student\_data.txt' USING PigStorage(',')

as ( id:int, firstname:chararray, lastname:chararray, phone:chararray, city:chararray );

STORE student INTO ' hdfs://localhost:9000/pig\_Output/ ' USING PigStorage (',');

**Output :-**

**hdfs dfs -cat 'hdfs://localhost:9000/pig\_Output/part-m-00000'**

1,Rajiv,Reddy,9848022337,Hyderabad

2,siddarth,Battacharya,9848022338,Kolkata

3,Rajesh,Khanna,9848022339,Delhi

4,Preethi,Agarwal,9848022330,Pune

5,Trupthi,Mohanthy,9848022336,Bhuwaneshwar

6,Archana,Mishra,9848022335,Chennai