File : student\_details.txt —-> In local location

001,Rajiv,Reddy,21,9848022337,Hyderabad,89

002,siddarth,Battacharya,22,9848022338,Kolkata,78

003,Rajesh,Khanna,22,9848022339,Delhi,90

004,Preethi,Agarwal,21,9848022330,Pune,93

005,Trupthi,Mohanthy,23,9848022336,Bhuwaneshwar,75

006,Archana,Mishra,23,9848022335,Chennai,87

grunt> Dump student\_details;

( 1,Rajiv,Reddy,21,9848022337,Hyderabad,89 )

( 2,siddarth,Battacharya,22,9848022338,Kolkata,78 )

( 3,Rajesh,Khanna,22,9848022339,Delhi,90 )

( 3,Rajesh,Khanna,22,9848022339,Delhi,90)

( 4,Preethi,Agarwal,21,9848022330,Pune,93 )

( 5,Trupthi,Mohanthy,23,9848022336,Bhuwaneshwar,75 )

( 6,Archana,Mishra,23,9848022335,Chennai,87 )

grunt> Describe student\_details;

student\_details: {id: int, firstname: chararray, lastname: chararray, age: int,

phone: chararray, city: chararray, gpa: int}

//CONCAT//

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

grunt> Dump student\_name\_concat;

O/P :

(RajivReddy)

(siddarthBattacharya)

(RajeshKhanna)

(PreethiAgarwal)

(TrupthiMohanthy)

(ArchanaMishra)

//SUM//

grunt> student\_details1= Group student\_data all;

(all,{(,Rajiv,Reddy,21,9848022337,Hyderabad,89),

( 2,siddarth,Battacharya,22,9848022338,Kolkata,78 )

( 3,Rajesh,Khanna,22,9848022339,Delhi,90 )

( 4,Preethi,Agarwal,21,9848022330,Pune,93 )

( 5,Trupthi,Mohanthy,23,9848022336,Bhuwaneshwar,75 )

( 6,Archana,Mishra,23,9848022335,Chennai,87

grunt> student\_sum = foreach student\_data Generate ( student\_details1.fname,student\_details1.age) , SUM(student\_details1.age)

grunt> Dump student\_details1;

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

{ (89) , (78) , (90) ,(93) , (75) , (87) }), 512}

//STORE//

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

//LIMIT//

limit\_data = LIMIT student\_details 4

DUMP limit\_data

o/p

(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;

O/P

Dump distinct\_data;

(1,Rajiv,Reddy,21,9848022337,Hyderabad,89 )

( 2,siddarth,Battacharya,22,9848022338,Kolkata,78 )

( 3,Rajesh,Khanna,22,9848022339,Delhi,90 )

(4,Preethi,Agarwal,21,9848022330,Pune,93 )

( 5,Trupthi,Mohanthy,23,9848022336,Bhuwaneshwar,75 )

( 6,Archana,Mishra,23,9848022335,Chennai,87 )

//TOKENIZE//

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

DUMP student\_name\_tokenize;

O/P

({(Rajaiv),(Reddy)})

({(siddarth),(Battacharya)})

({(Rajesh),(Khanna)})

({(Preethi),(Agarwal)})

({(Trupthi),(Mohanthy)})

({(Archana),(Mishra)})