Big Data Lab

Pig Programs

1. Write a program to filter data by city.

pig\_script\_filter.pig

student\_details = LOAD '/home/msrit/Downloads/pig/test/Filter/student\_details.txt' USING PigStorage(',') as (id:int, firstname:chararray, lastname:chararray, age:int, phone:chararray, city:chararray);

filter\_data = FILTER student\_details BY city == 'Chennai';

Dump filter\_data;

2. Write a program to group students by age.

pig\_script\_grouping.pig

student = LOAD '/home/msrit/Downloads/pig/test/Grouping/student\_details.txt' USING PigStorage(',') as (id:int, firstname:chararray, lastname:chararray, age:int, phone:chararray, city:chararray);

group\_data = GROUP student by age;

Dump group\_data;

3. Write a program to join 2 tables by id and order by customer\_id

pig\_script\_joining.pig

customers = LOAD '/home/msrit/Downloads/pig/test/Joining/customer.txt' USING PigStorage(',') as (id:int, name:chararray, age:int, address:chararray, salary:int);

orders = LOAD '/home/msrit/Downloads/pig/test/Joining/order.txt' USING PigStorage(',') as (oid:int, date:chararray, customer\_id:int, amount:int);

join\_result = JOIN customers BY id, orders BY customer\_id;

Dump join\_result

4. Write a program to sort student by age in descending order and limit the results to a maximum of 4 records.

pig\_script\_sorting.pig

student = LOAD '/home/msrit/Downloads/pig/test/Sorting/student\_details.txt' USING PigStorage(',') as (id:int, firstname:chararray, lastname:chararray, age:int, phone:chararray, city:chararray);

student\_order = ORDER student BY age DESC;

student\_limit = LIMIT student\_order 4;

Dump student\_limit;

5. Write a program to create a union between 2 tables.

pig\_script\_union.pig

cust1 = LOAD '/home/msrit/Downloads/pig/test/Union/customer1.txt' USING PigStorage(',') as (id:int, firstname:chararray, lastname:chararray, phone:chararray, city:chararray);

cust2 = LOAD '/home/msrit/Downloads/pig/test/Union/customer2.txt' USING PigStorage(',') as (id:int, firstname:chararray, lastname:chararray, phone:chararray, city:chararray);

cust = UNION cust1, cust2;

Dump cust