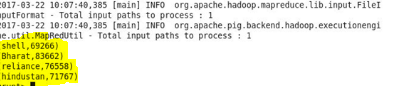
1. What is the total amount of petrol in volume sold by every distributor?

a=load ‘/home/acadgild/Desktop/petrol-dataset.txt’ USING PigStorage(‘,’) AS (ID:chararray, name:chararray, buy:chararray, sell:chararray, in:int ,year:int);

b=group a by name;

c= foreach b GENERATE group,SUM(a.out)



2. Which are the top 10 distributors ID's for selling petrol? Also display the amount of petrol

sold in volume.

a=load ‘/home/acadgild/Desktop/petrol-dataset.txt’ USING PigStorage(‘,’) AS (ID:chararray, name:chararray, buy:chararray, sell:chararray, in:int ,year:int);

b=group a by ID;

c= foreach b GENERATE group,SUM(a.out) as sum;

d= order c by sum DESC;

e=limit d 10;



3. List 10 years where consumption of petrol is more with the distributer id who sold it.

a=load ‘/home/acadgild/Desktop/petrol-dataset.txt’ USING PigStorage(‘,’) AS (ID:chararray, name:chararray, buy:chararray, sell:chararray, in:int ,year:int);

b=group a (ID, year);

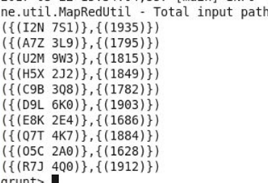
c= foreach b GENERATE a.ID as id, a.year as year, SUM(a.in) as cons;

d= order c by cons DESC;

e=foreach d GENERATE id,year;

e=foreach d GENERATE flatten(id), flatten(year);

f= limit e 10;



4. Find the distributer name who sold petrol in least amount.

a=load ‘/home/acadgild/Desktop/petrol-dataset.txt’ USING PigStorage(‘,’) AS (ID:chararray, name:chararray, buy:chararray, sell:chararray, in:int ,year:int);

b=group a by year;

c= foreach b {

temp = foreach a GENERATE in-out;

generate a.name as name, a.year as year, sum(temp) as sale;};

d=order c by sale DESC;

e=foreach c generate flatten(name);

f=limit e 1;

