Assignment 11.1

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

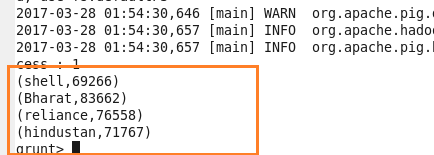
CODE :

A = LOAD ‘petrol\_dataset.txt’ USING PigStorage (‘,’) AS (DistributorID:chararray, name:chararray,buyrate:chararray,sellrate:chararray,volin:int,volout:int,year:int);

B = GROUP A by name;

C = FOREACH B generate group,SUM(A.volout);

DUMP C;



1. Which are the top 10 distributors ID's for selling petrol? Also display the amount of petrol sold in volume.

CODE :

A = LOAD ‘petrol\_dataset.txt’ USING PigStorage (‘,’) AS (DistributorID:chararray, name:chararray,buyrate:chararray,sellrate:chararray,volin:int,volout:int,year:int);

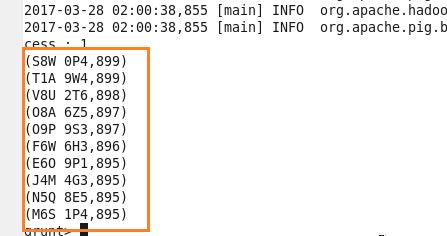
B = GROUP A by DistributorID;

C = FOREACH B generate group,SUM(A.volout) as total;

D = ORDER C by total DESC;

E = LIMIT D 10;

DUMP E;



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

CODE :

A = LOAD ‘petrol\_dataset.txt’ USING PigStorage (‘,’) AS (DistributorID:chararray, name:chararray,buyrate:chararray,sellrate:chararray,volin:int,volout:int,year:int);

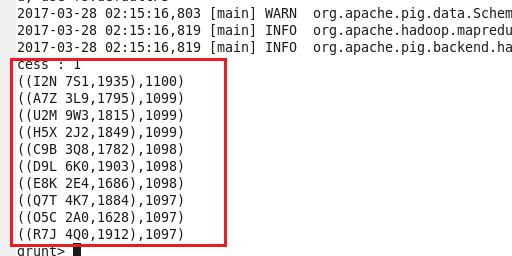
B = GROUP A by (DistributorID,year);

C = FOREACH B generate group,FLATTEN(A.volin) as vol;

D = ORDER C by vol DESC;

E = LIMIT D 10;

DUMP E;



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

CODE :

A = LOAD ‘petrol\_dataset.txt’ USING PigStorage (‘,’) AS (DistributorID:chararray, name:chararray,buyrate:chararray,sellrate:chararray,volin:int,volout:int,year:int);

B = GROUP A by name;

C = FOREACH B generate group,SUM(A.volout) as vol1;

D = ORDER C by vol1;

E = LIMIT D 1;

