Assignment 9.1

petrol_data = LOAD 'petrol_dataset' USING PigStorage(',') AS (dist_id:chararray,dist_name:chararray,buy_rate:chararray, sell_rate:chararray, volume_in:int, volume_out:int,year:int);
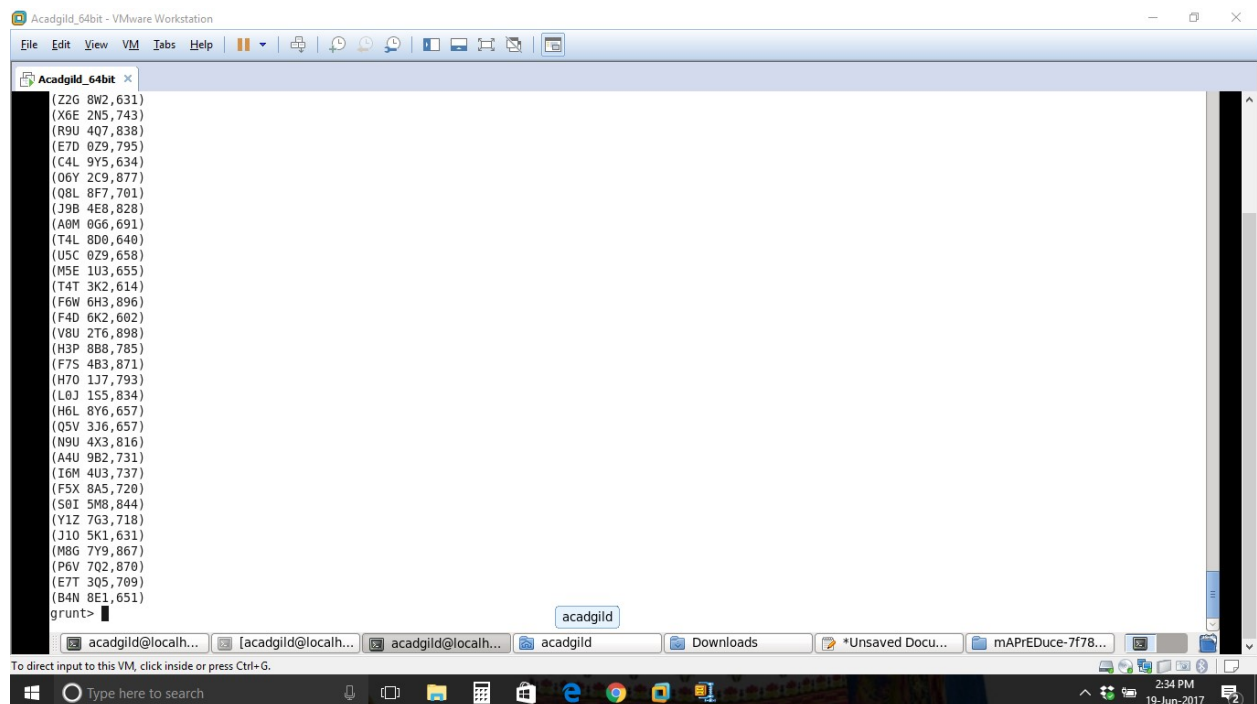
dump petrol_data;

1.Problem Statement

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

petrol_volume_sold = foreach petrol_data generate dist_id, volume_out;

dump petrol_volume_sold;

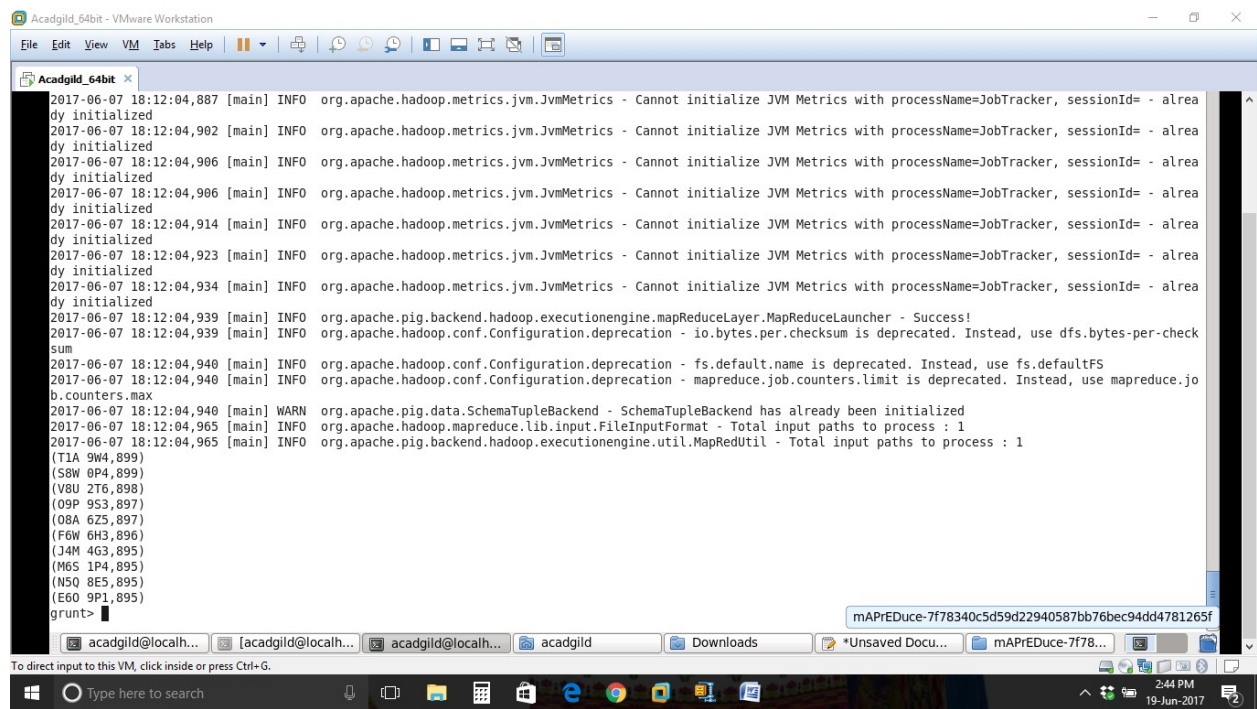**Distributor with petrol volume sold**

2.Problem Statement

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

petrol_volume_top10 = ORDER petrol_volume_sold by volume_out desc;

top_10_distributor = limit petrol_volume_top10 10;

dump top_10_distributor;

Top 10 distributor with volume sold details

3.Problem Statement

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

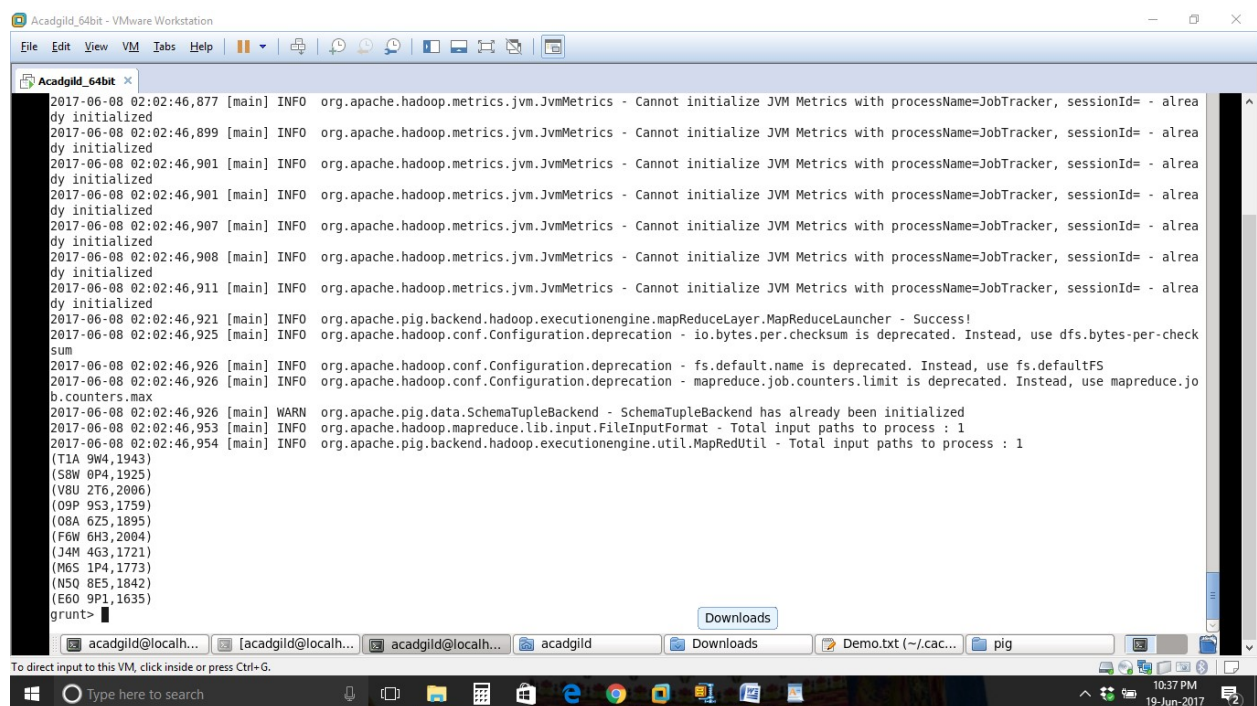petrol_volume_sold_year = foreach petrol_data generate dist_id, volume_out,year;

petrol_volume_top10_year = ORDER petrol_volume_sold_year by volume_out desc;

top_10_distributor_year = limit petrol_volume_top10_year 10;

top_consumption_year = foreach top_10_distributor_year generate dist_id,year;

dump top_consumption_year;

Top 10 years where petrol consumption was more

4.Problem Statement

Find the distributor name who sold petrol in least amount.

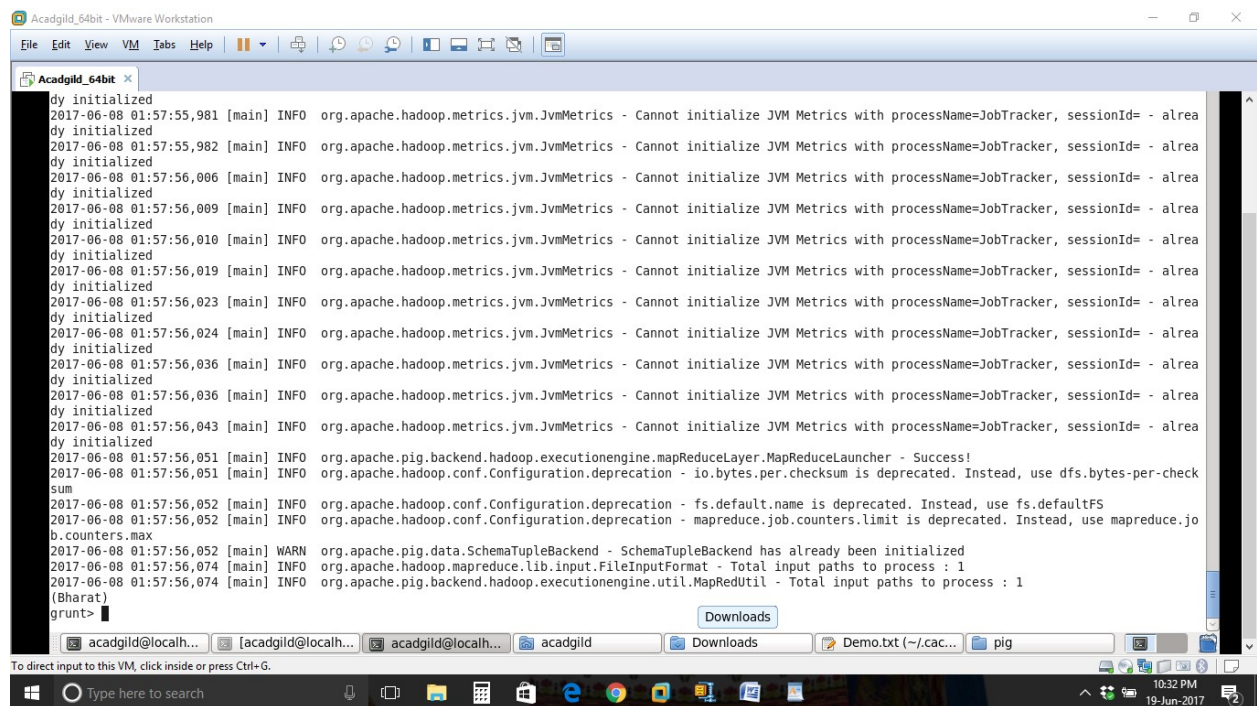petrol_volume_sold = foreach petrol_data generate dist_id,dist_name ,volume_out;

petrol_volume_ascending = ORDER petrol_volume_sold by volume_out asc;

distributor_with_least_sales = limit petrol_volume_ascending 1;

distributor_least_sale_name = foreach distributor_with_least_sales generate dist_name;

dump distributor_least_sale_name;

Name of distributor who sold least amount of Petrol