**Creation of Table in Hive and Loading of data**

**create table petrol (distributer\_id STRING,distributer\_name STRING,amt\_IN STRING,amy\_OUT STRING,vol\_IN INT,vol\_OUT INT,year INT) row format delimited fields terminated by ‘,’ stored as textfile;**

**load data local inpath ‘/home/hduser/petrol.txt’ into table petrol;**

1)In real life what is the total amount of petrol in volume sold by every distributor?

2)Which are the top 10 distributors ID’s for selling petrol and also display the amount of petrol sold in volume by them individually?

3)Find real life 10 distributor name who sold petrol in the least amount.

4) The constraint to this query is the difference between volumeIN and volumeOuT is illegal in real life if greater than 500. As we see all distributors are receiving patrols on every next cycle.

**List all distributors who have this difference, along with the year and the difference which they have in that year.**

Hint: (vol\_IN-vol\_OUT)>500