Assignment 2 - Sudhanva

1.

sales.csv

2.

```
grunt> sales = load 'sales.csv' using PigStorage(',')
>> as
>> (
>> CName : chararray,
>> Shop : chararray,
>> Price : double
>> );
```

3.

```
grunt> dump sales;
(Amar,Big Bazaar,450.9)
(Amar,Amazon,8900.0)
(Armaan,Big Bazaar,8900.0)
(Bulbul,Big Bazaar,6789.0)
(Bulbul,Pantaloon,6767.9)
(Balbir,Amazon,89000.0)
(Candy,Walmart,8900.0)
(Chandana,Amazon,6755.99)
(Dolly,Flipkart,20000.7)
(Dolly,Amazon,8023.78)
(Daniel,Amazon,9000.0)
(Francis,Shopper?s Stop,5000.0)
(Ghanshyam,Pantaloon,12000.0)
(Ghanshyam,Mark & Spencer,7580.9)
grunt> describe sales;
sales: {CName: chararray,Shop: chararray,Price: double}
```

4.

```
grunt> result = foreach sales_sub generate group, SUM(sales.Price);
grunt> dump result;
(Amazon,121679.76999999999)
(Walmart,8900.0)
(Flipkart,20000.7)
(Pantaloon,18767.9)
(Big Bazaar,16139.9)
(Mark & Spencer,7580.9)
(Shopper?s Stop,5000.0)
```

5.

```
grunt> sales_sub = group sales by CName;
grunt> dump sales_sub;
(Amar,{(Amar,Big Bazaar,450.9),(Amar,Amazon,8900.0)})
(Candy,{(Candy,Walmart,8900.0)})
(Dolly,{(Dolly,Amazon,8023.78),(Dolly,Flipkart,20000.7)})
(Armaan,{(Armaan,Big Bazaar,8900.0)})
(Balbir,{(Balbir,Amazon,89000.0)})
(Bulbul,{(Bulbul,Pantaloon,6767.9),(Bulbul,Big Bazaar,6789.0)})
(Daniel,{(Daniel,Amazon,9000.0)})
(Francis,{(Francis,Shopper?s Stop,5000.0)})
(Chandana,{(Chandana,Amazon,6755.99)})
(Ghanshyam,{(Ghanshyam,Pantaloon,12000.0),(Ghanshyam,Mark & Spencer,7580.9)})
grunt> result = foreach sales_sub generate group, SUM(sales.Price);
grunt> dump result;
(Amar,9350.9)
(Candy,8900.0)
(Dolly,28024.48)
(Armaan,8900.0)
(Balbir,89000.0)
(Bulbul,13556.9)
(Daniel,9000.0)
(Francis,5000.0)
(Chandana,6755.99)
(Ghanshyam,19580.9)
```

6. **emp.csv**

7. **dept.txt**

8.

```
grunt> emp = load 'emp.csv' using PigStorage(',')
>> as
>> (
>> Emp_No : int,
>> EName : chararray,
>> Job : chararray,
>> Mgr_id : int,
>> Hiredate : datetime,
>> Sal : int,
>> Commision : int,
>> Dept_no : int
>> );
grunt> dept = load 'dept.csv' using PigStorage(',')
>> as
>> (
>> Dept_no : int,
>> DepartmentName : chararray,
>> Area : chararray
>> );
```

9.

```
grunt> emp_sub = filter emp by Job == 'Manager';
grunt> dump emp_sub;
(7234,Jones,Manager,7839,,89000,,20)
(7698,Blake,Manager,7839,,78000,,30)
(7782,Clark,Manager,7839,,56000,,10)
```

10.

```
grunt> result = order emp by Dept_no ASC, Job DESC;
grunt> dump result;
(7839,King,President,,,145000,,10)
(7782,Clark,Manager,7839,,56000,,10)
(8980,Smith,Clerk,8000,,35000,,10)
(7234,Jones,Manager,7839,,89000,,20)
(7788,Scott,Analyst,7566,,78000,,20)
(8769,Martin,Salesman,7698,,45000,73000,30)
(7521,Ward,Salesman,7698,,48000,40000,30)
(6789,Allen,Salesman,7698,,40000,3500,30)
(7698,Blake,Manager,7839,,78000,,30)
```

11.

```
grunt> emp_sub = filter emp by Commision > Sal;
grunt> dump emp_sub;
(8769,Martin,Salesman,7698,,45000,73000,30)
```

12.

```
grunt> emp_sub = filter emp by Commision is not null;
grunt> dump emp_sub;
(6789,Allen,Salesman,7698,,40000,3500,30)
(7521,Ward,Salesman,7698,,48000,40000,30)
(8769,Martin,Salesman,7698,,45000,73000,30)
```