**Project 1.1:**

**Checking list of files in Hadoop.**

**Hadoop fs –ls /hadoopdata.**



**Created folder acadgild mini project .**

**Hadoop fs –ls /hadoopdata/aminiproject.**



**Imported Crimes.csv file into agminiproject.**

**Hadoop fs –ls /hadoopdata/aminiproject/Crimes.csv**



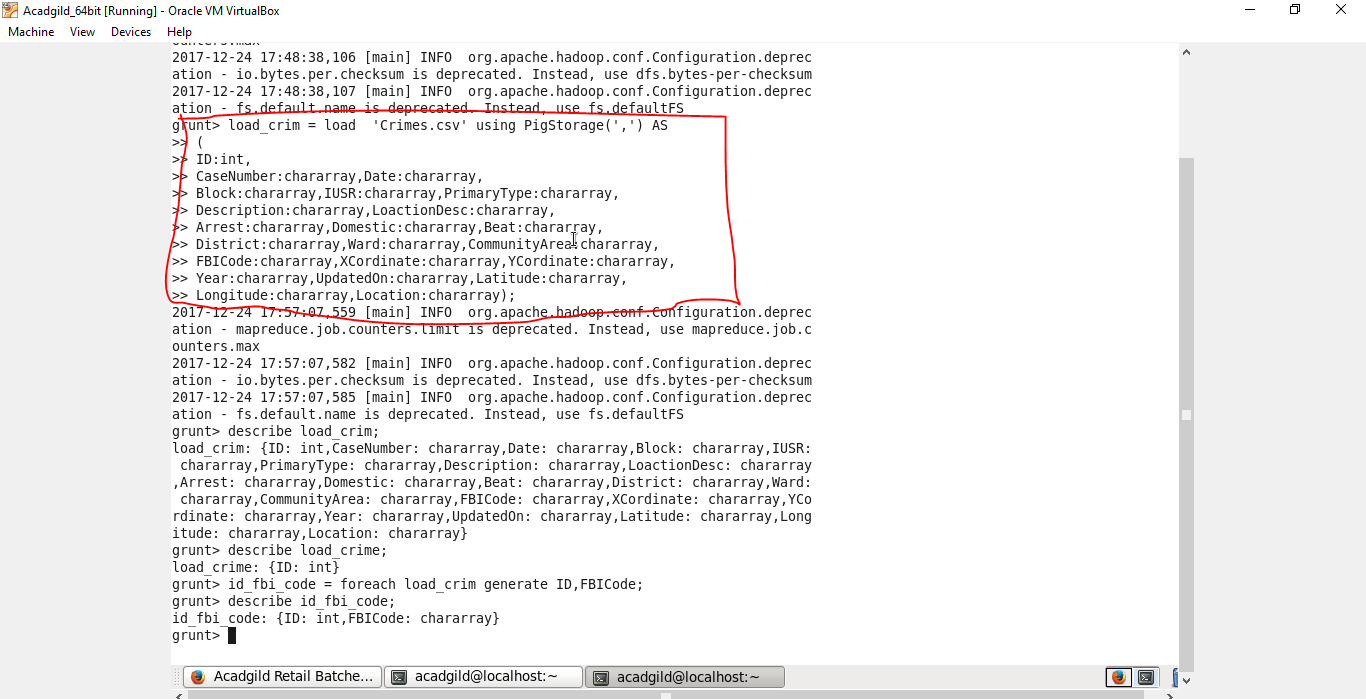
**For showing data in Crimes.csv file**

**Hadoop fs –cat /hadoopdata/aminiproject/Crimes.csv**

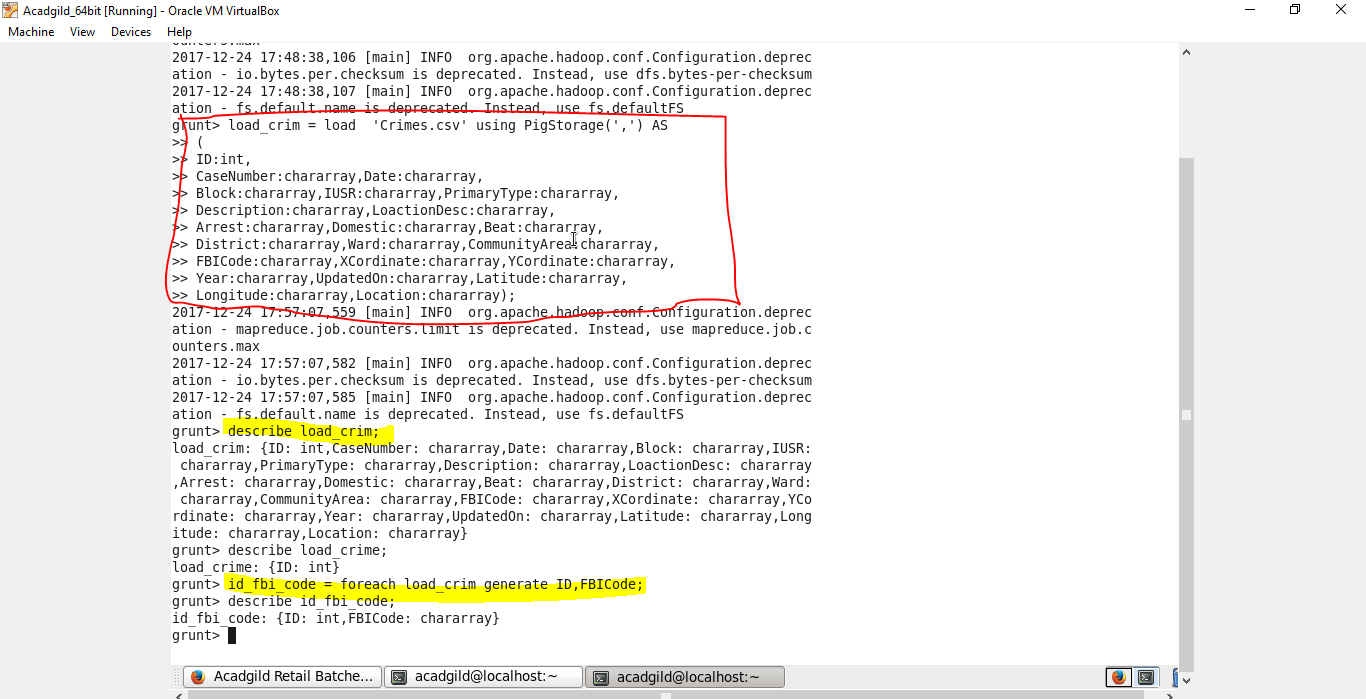
**1. Write a MapReduce/Pig program to calculate the number of cases investigated under each**

**FBI code**

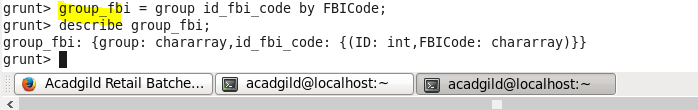
**Using Pig :**



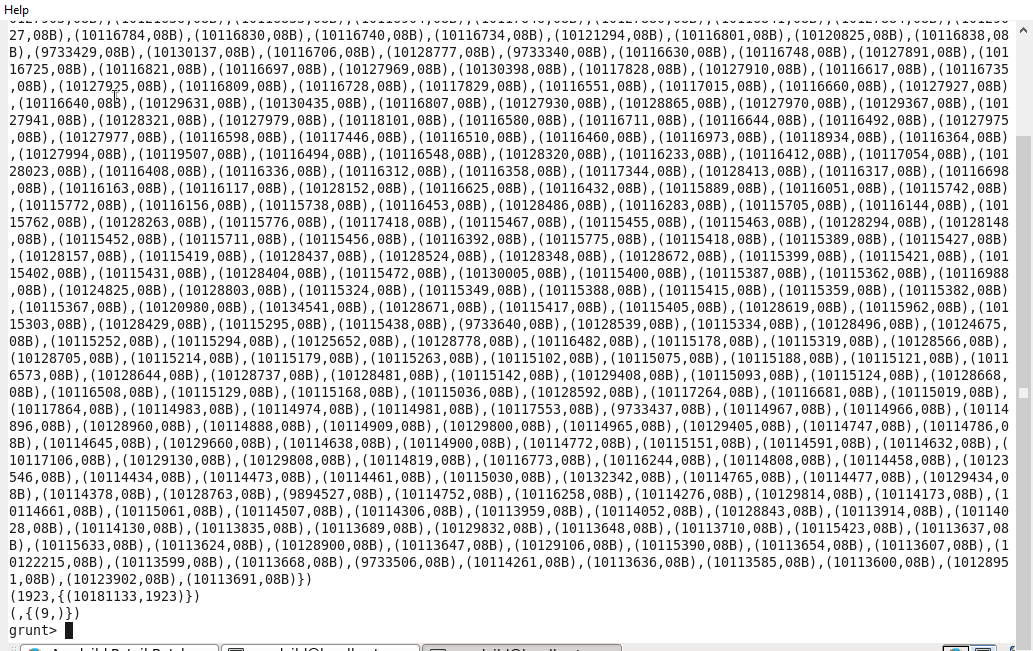
**Loading in fbi code:**

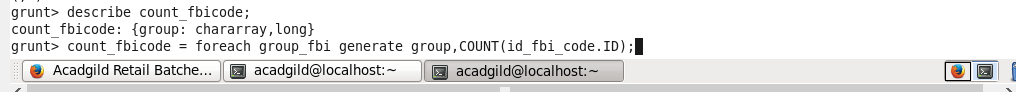


**Grouping the fbi code by using fbi code:**



**After dumping data above group\_fbi:**







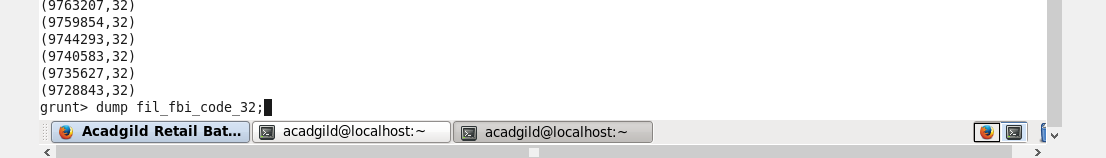


**2. Write a MapReduce/Pig program to calculate the number of cases investigated under FBI**

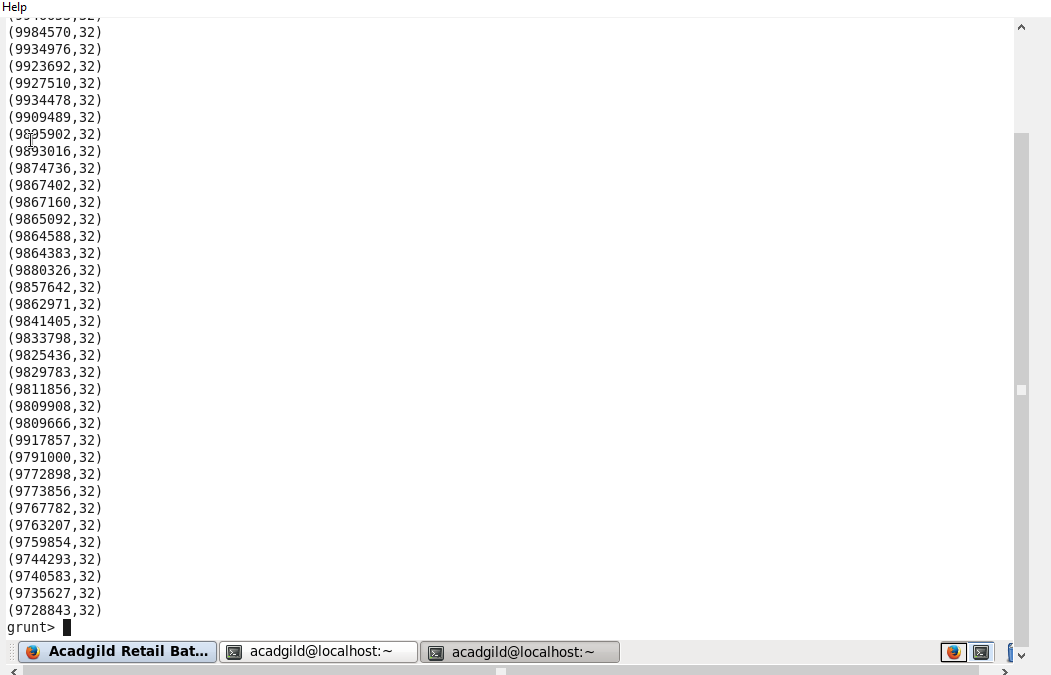
**code 32.**

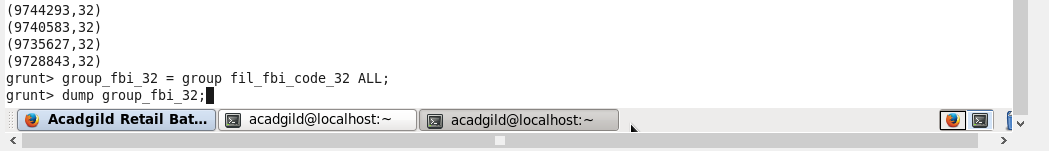


**Dumping data in**

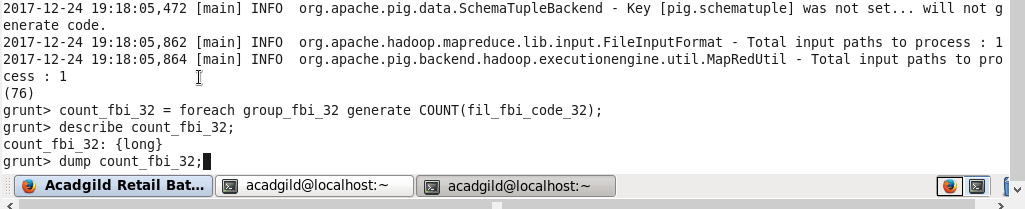


**After dumping data:**



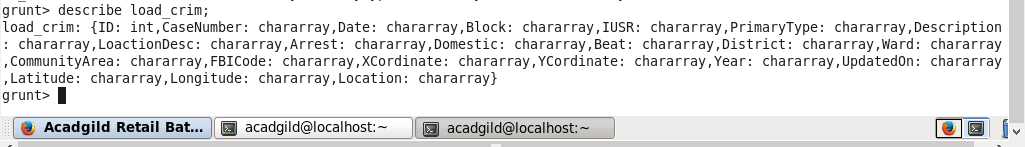


**For counting investigated FBI code under 32.**

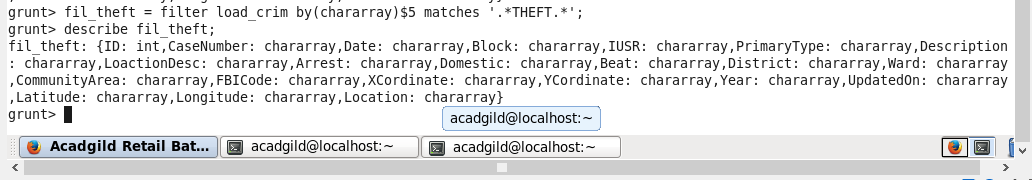


**3. Write a MapReduce/Pig program to calculate the number of arrests in theft district wise.**

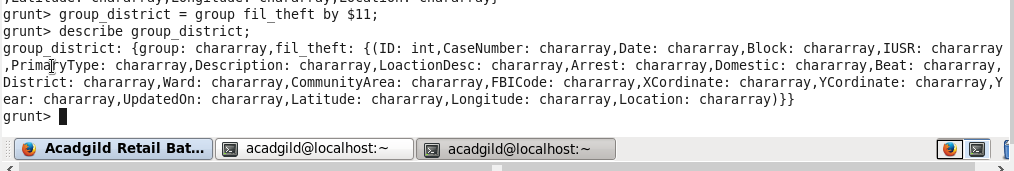
Step - 1



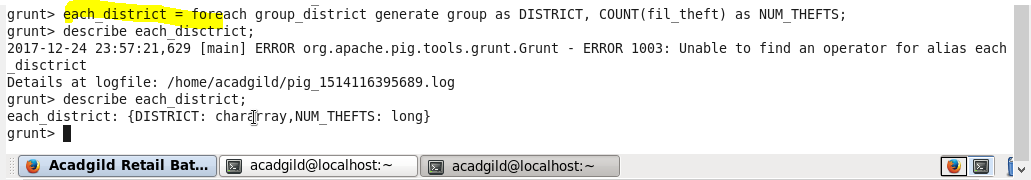
Step – 2 : filter only THEFTS records.

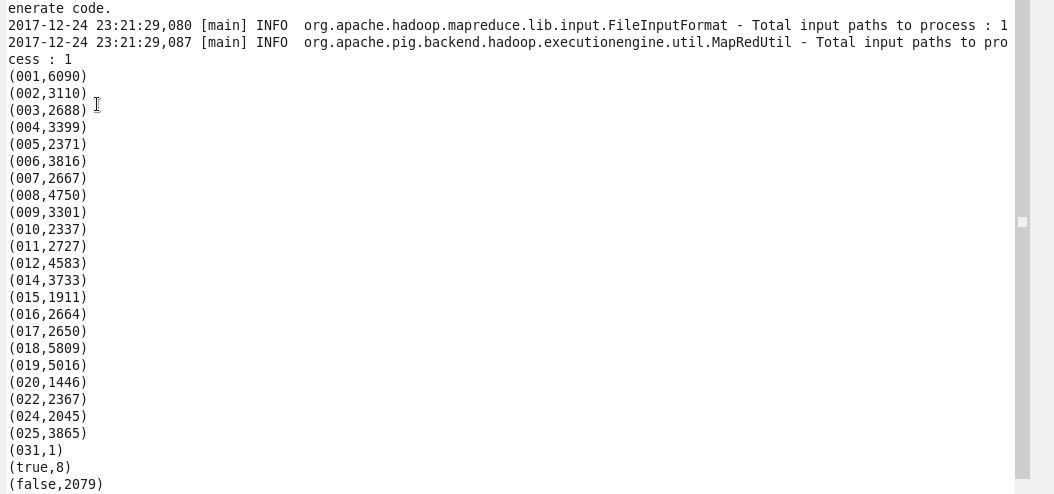


Step – 3 : grouping with respect to district:



Step – 4 grouping each district wise:

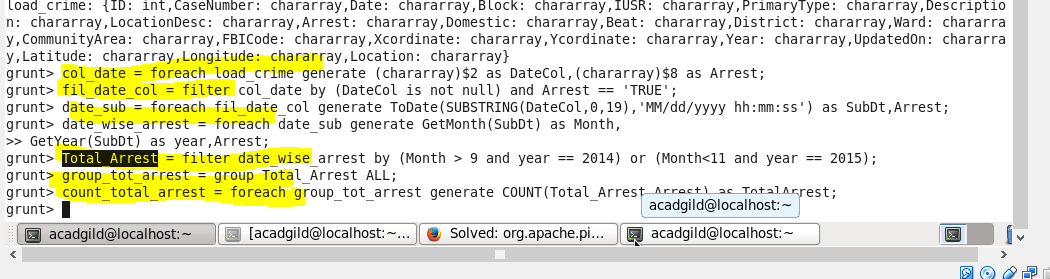




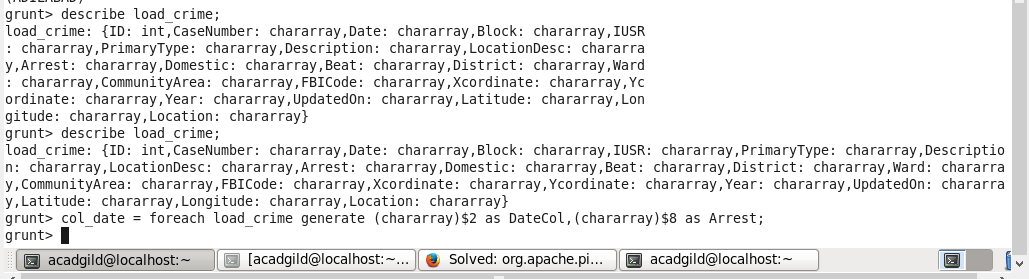
**4. Write a MapReduce/Pig program to calculate the number of arrests done between October**

**2014 and October 2015.**

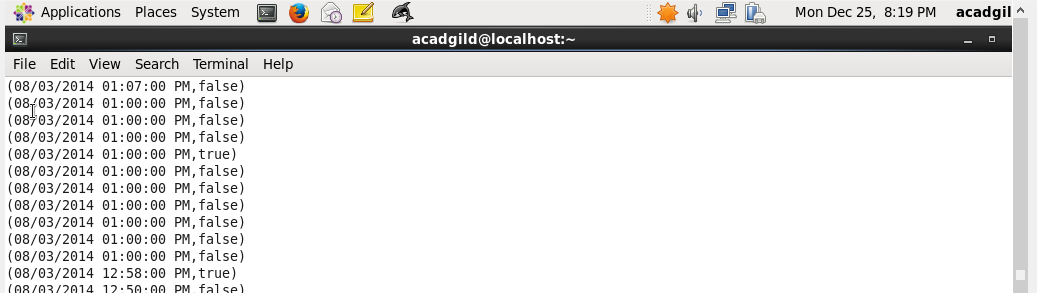
**Filtering records in between 2014 and 2015;**



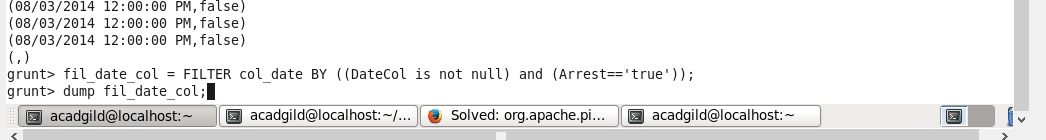
**Selecting columns:**



**Data dumped on col\_date:**

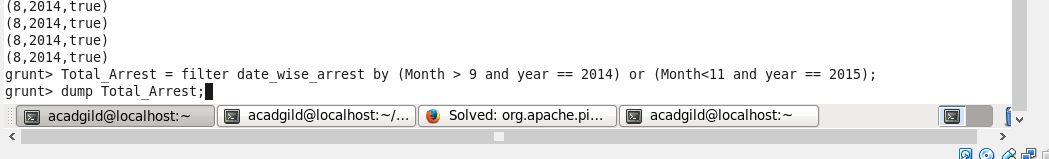


**Filtering from above selected columns:**

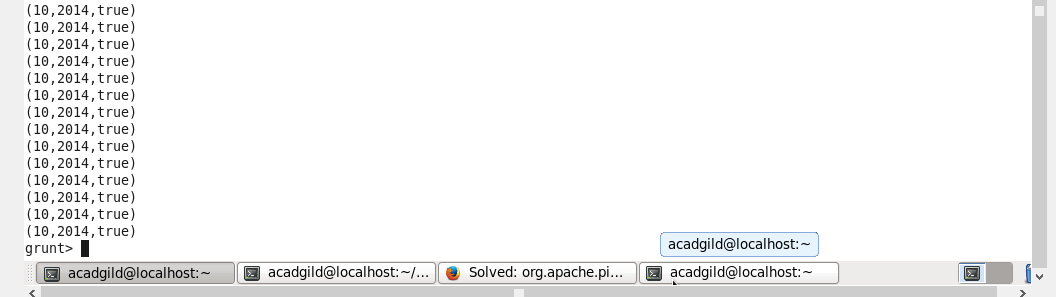


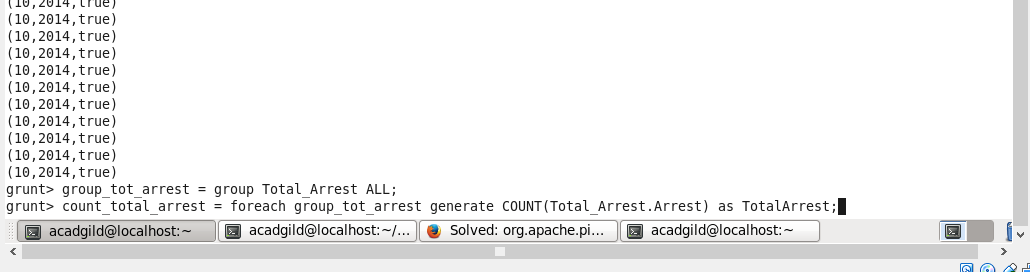
**After dumping :**





**After dumping data:**

**Grouping by using Arrest :**



**After counting total arrest :**

