1.Introduction

This dataset contains attributes related to crimes taking place in various areas like type of crime, FBI code related to that criminal case, arrest frequency, location of crime etc.

2. Objective

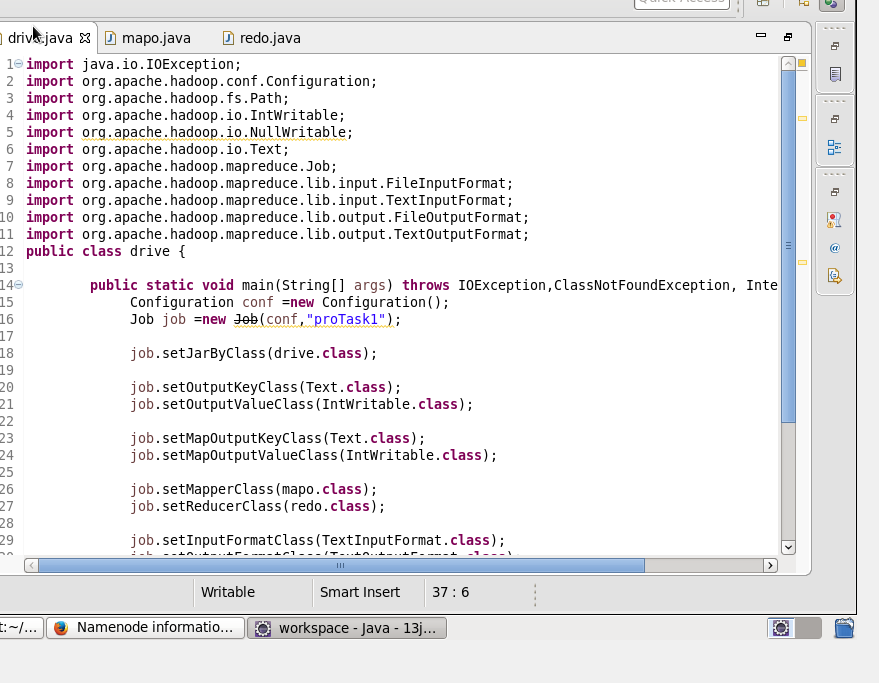
3. Prerequisites

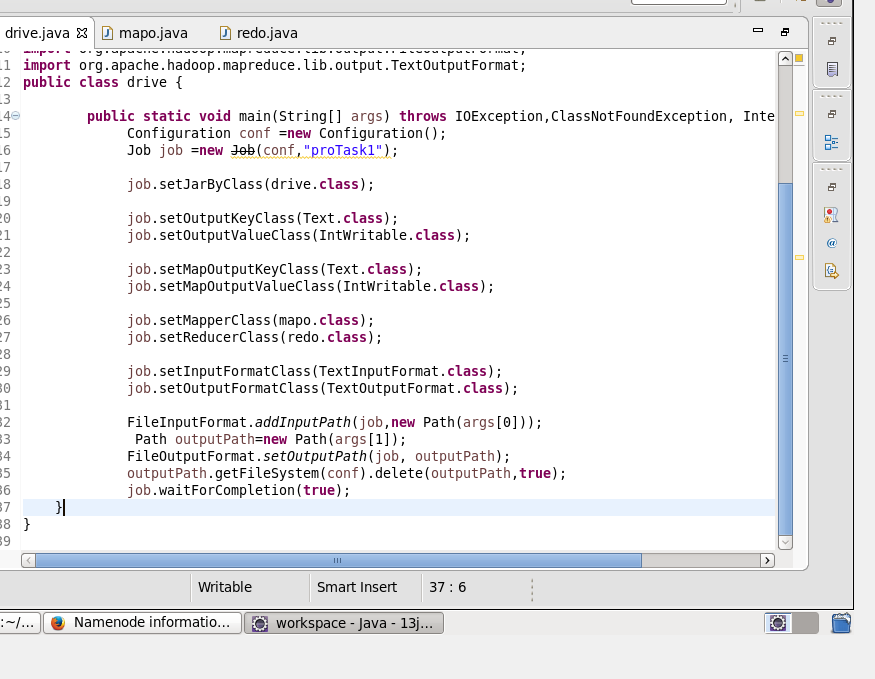
You should have Hadoop cluster installed in your system.

4. Associated Data Files Description: ID,Case Number,Date,Block,IUCR,Primary Type,Description,Location Description,Arrest,Domestic,Beat,District,Ward,Community Area,FBICode,X Coordinate,Y Coordinate,Year,Updated On,Latitude,Longitude,Location 5. Problem Statement

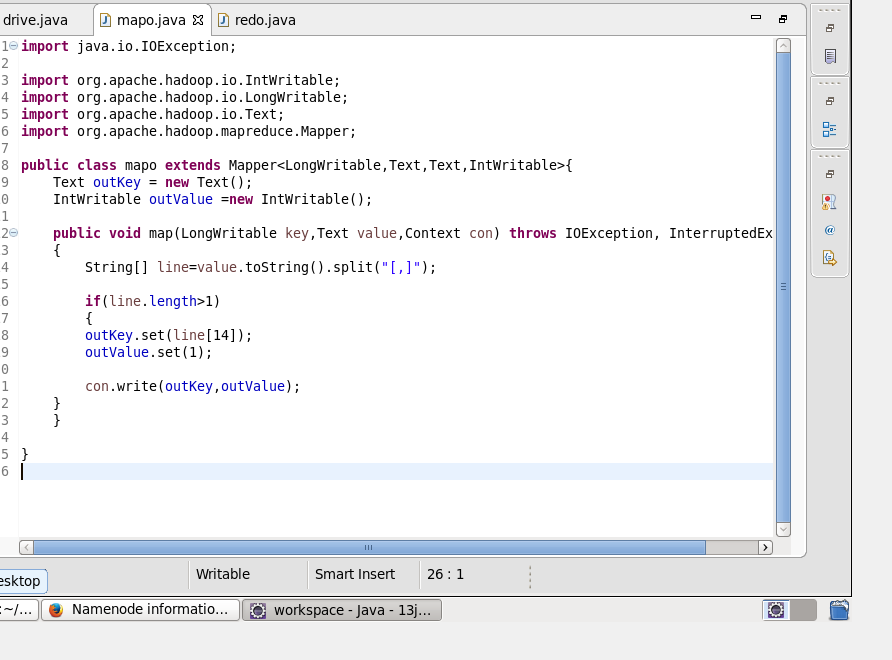
1. Write a mapreduce and pig program to calculate the number of cases investigated under each FBI code

task1driver

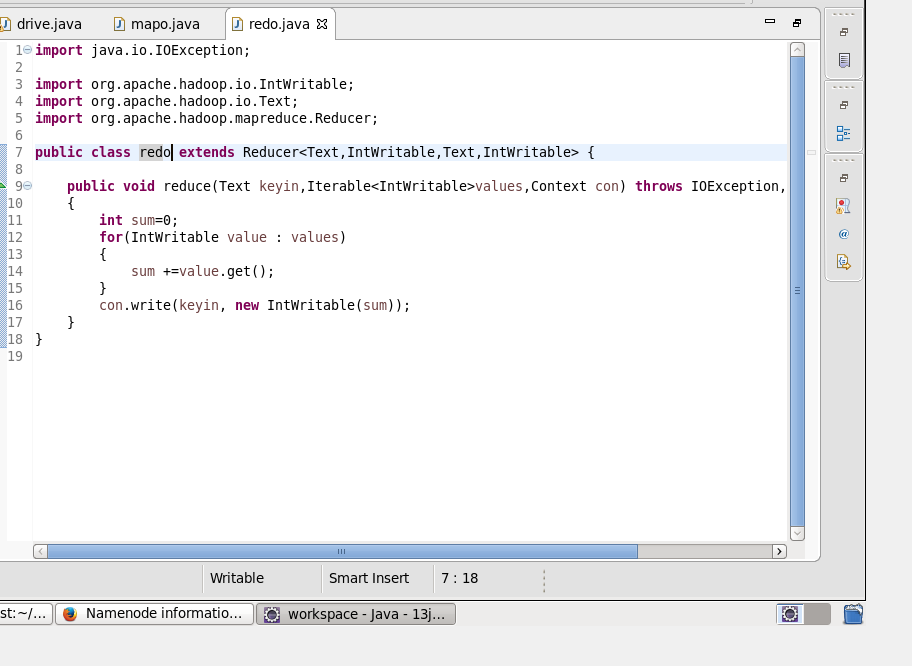




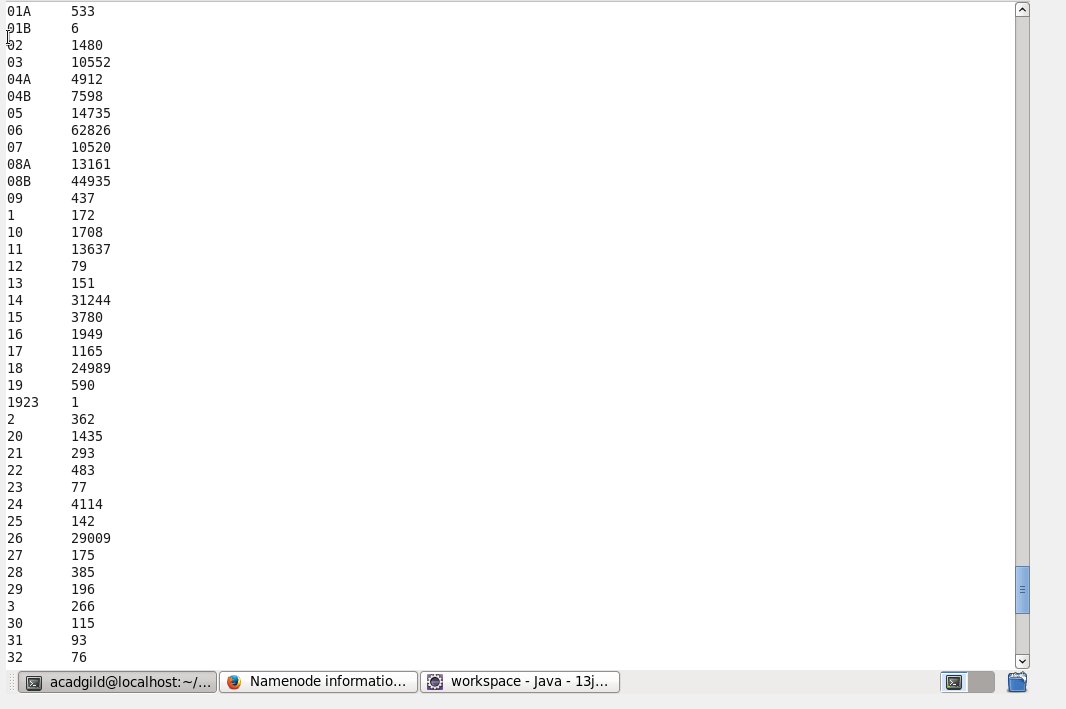
Mapper:

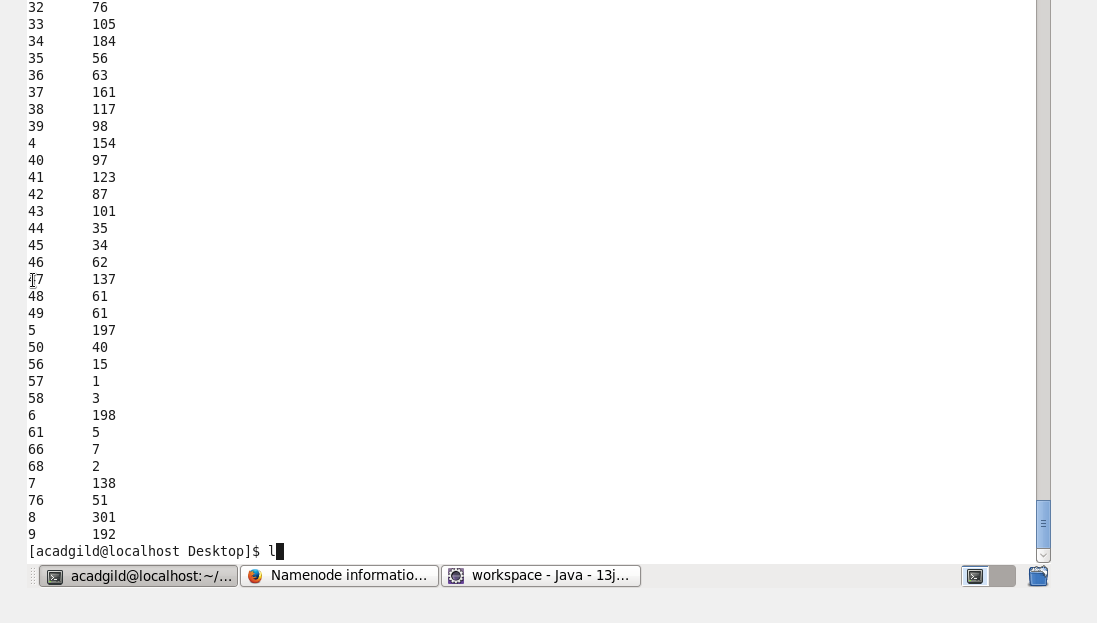


Reducer:



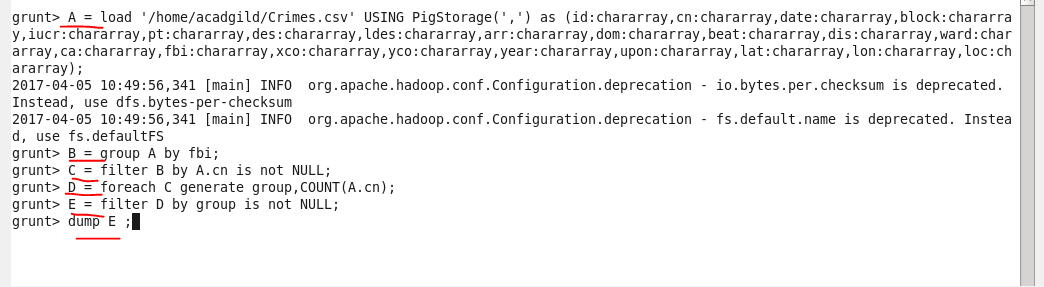
Result:

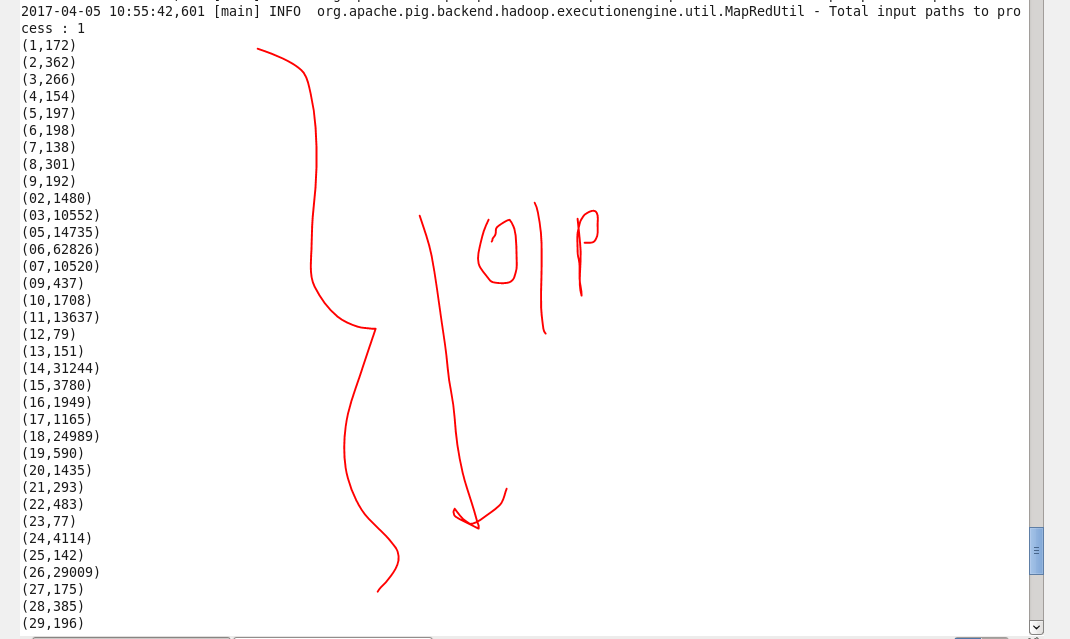


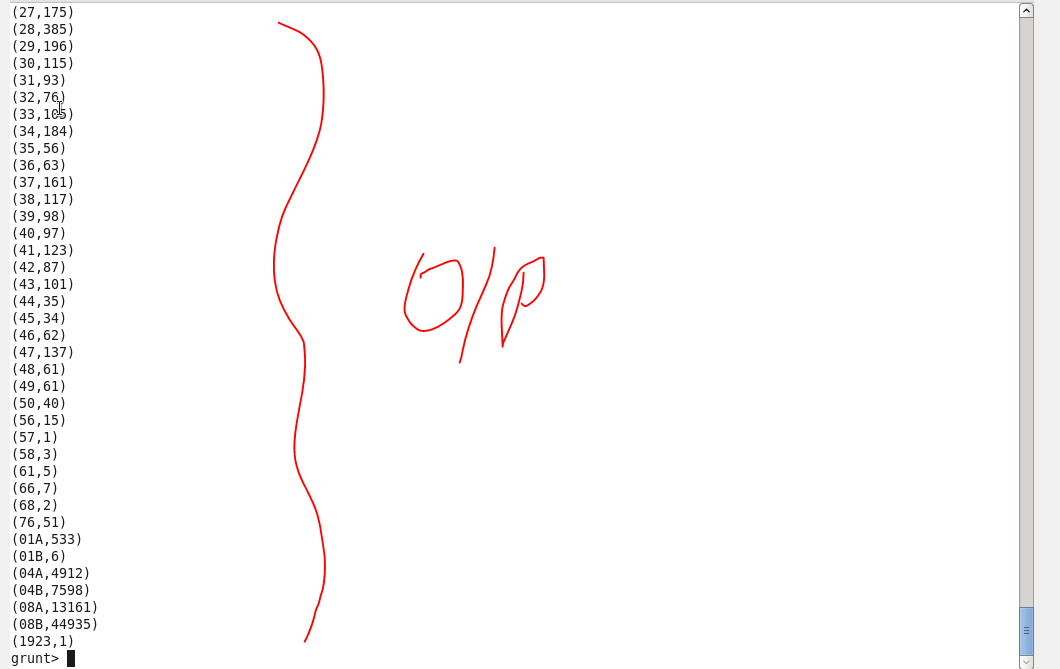


In pig:

Result:

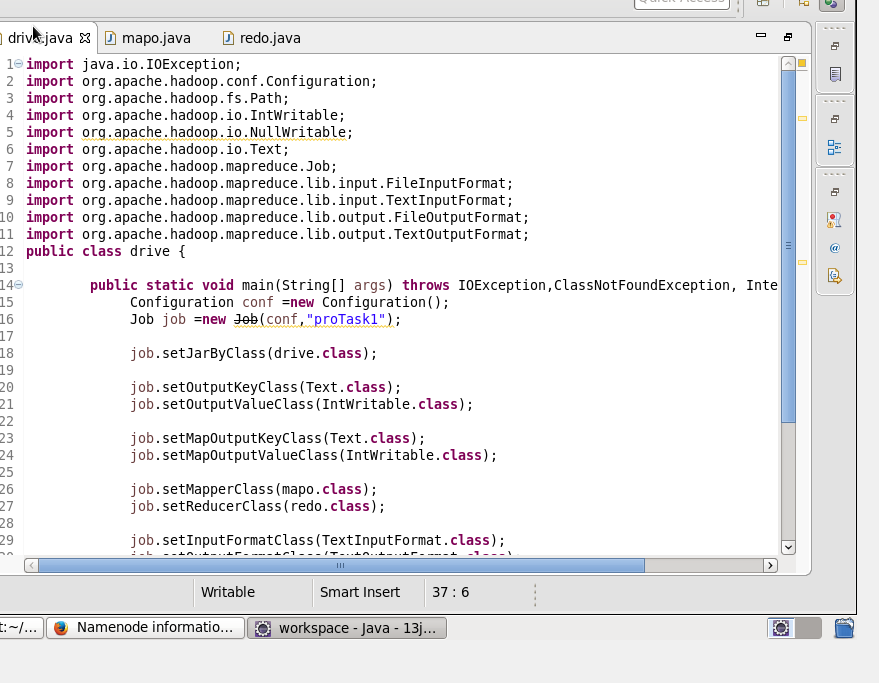


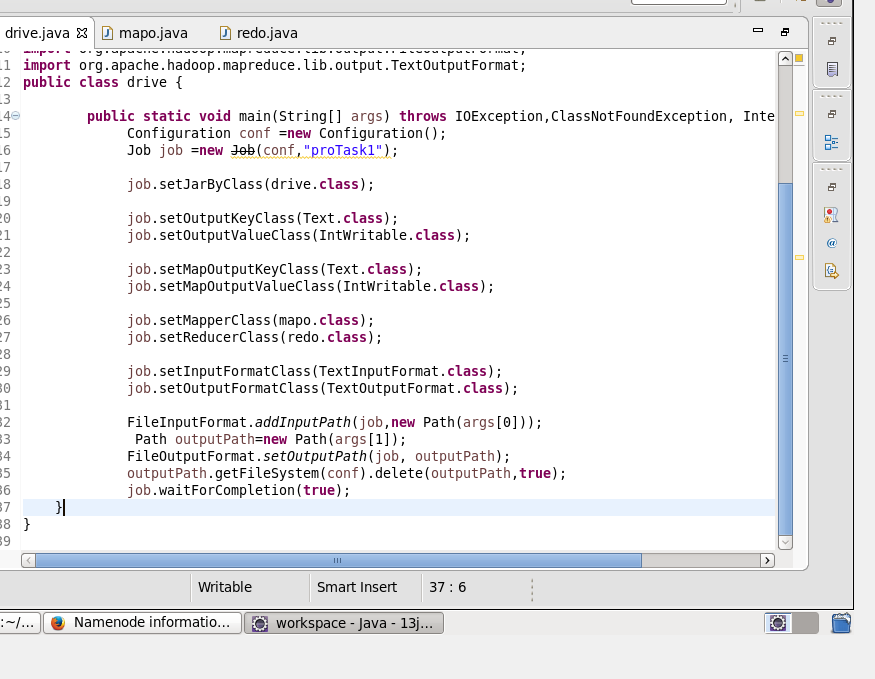




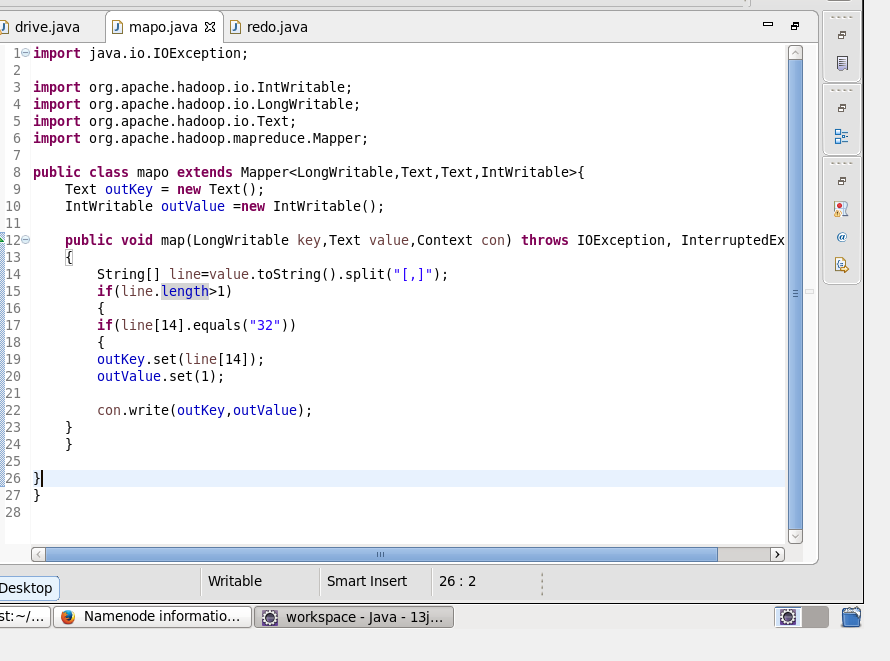
2. Write a mapreduce and pig program to calculate the number of cases investigated under FBI code 32.

\*\*\*\*Driver remains the same.

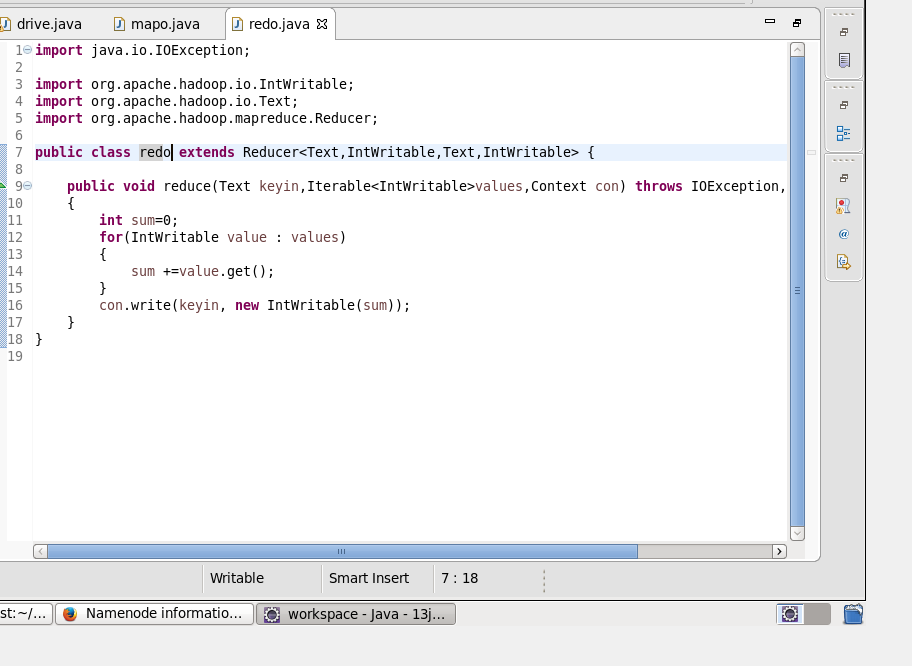




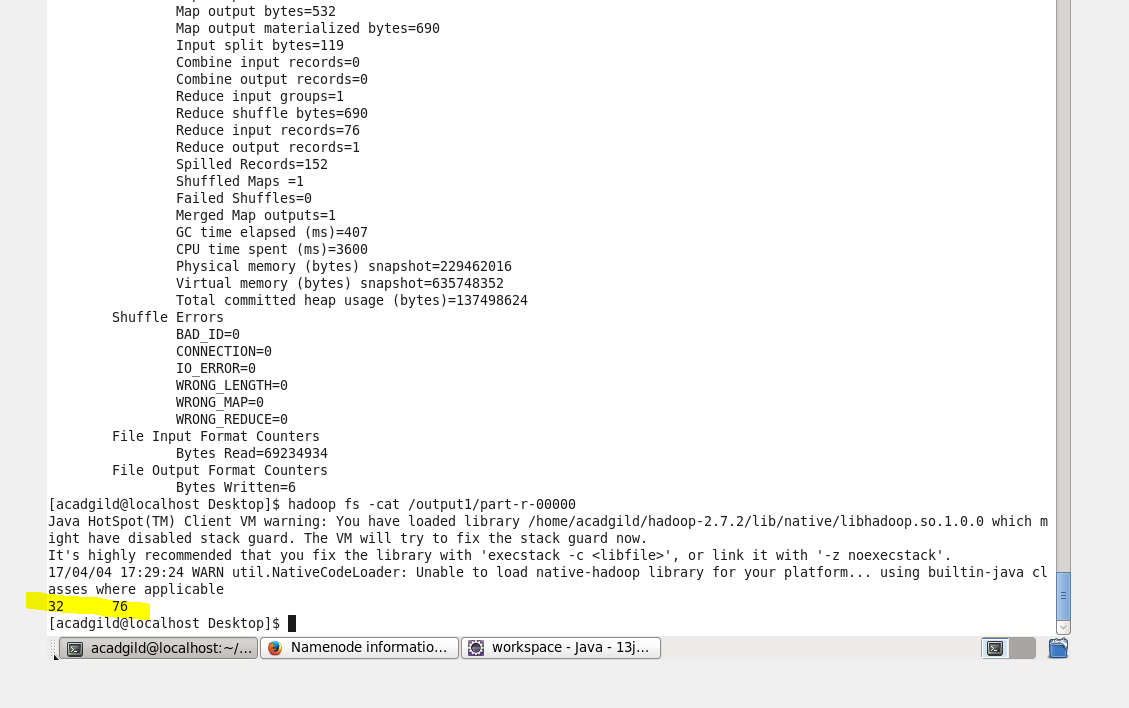
Mapper:



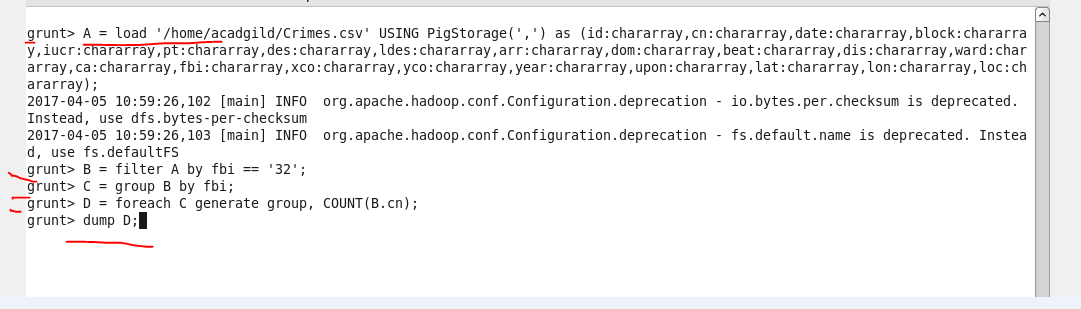
Reducer:



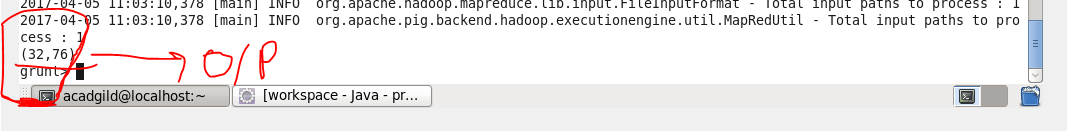
Result:



Pig:



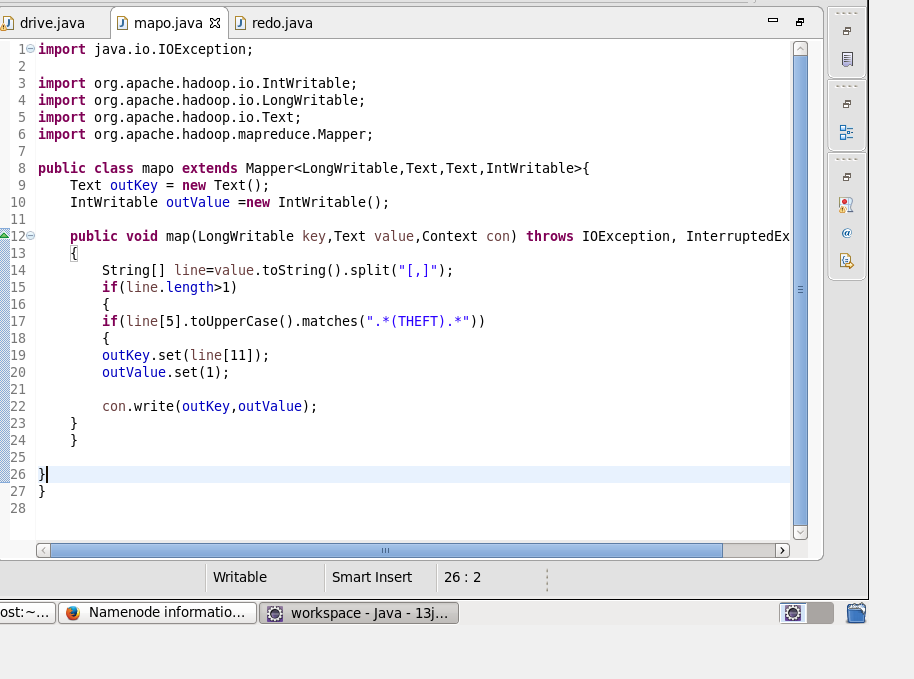
Result:



3. Write a mapreduce and pig program to calculate the number of arrests in theft district wise.

Driver remains the same.

Mapper:

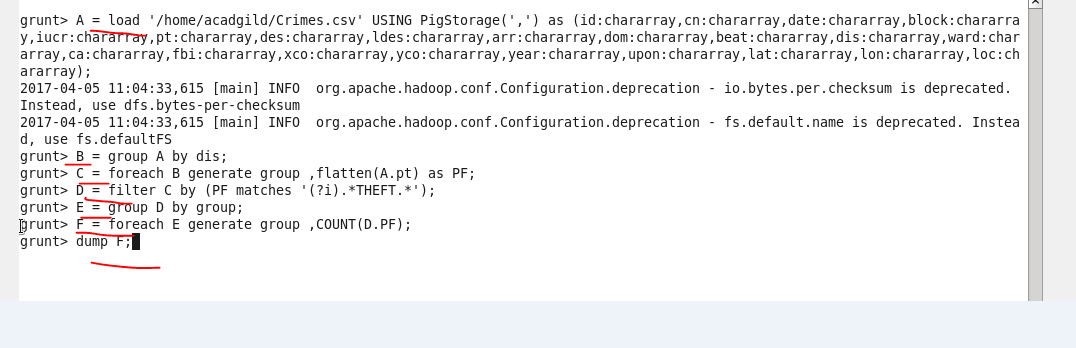


Reducer remains the same.

Result:



In pig:

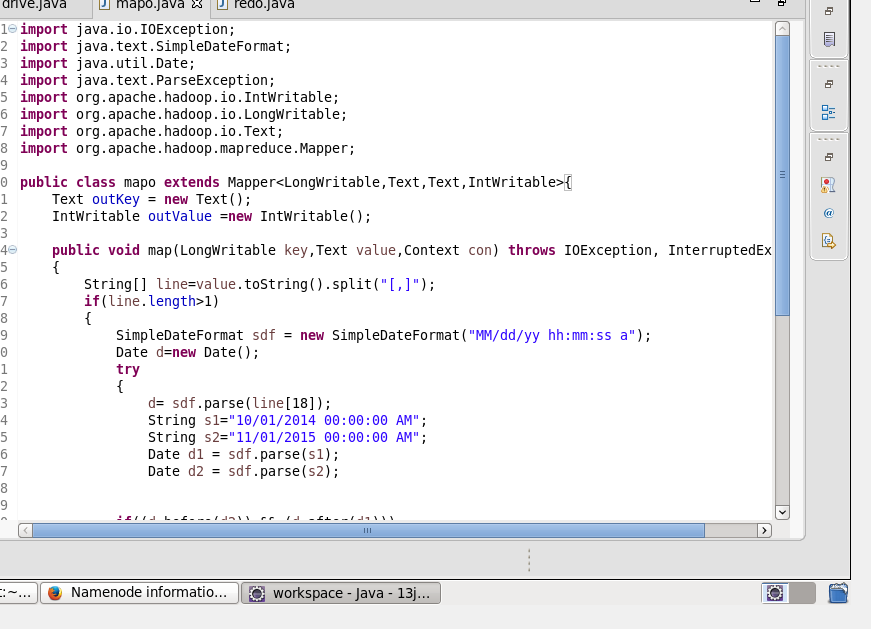


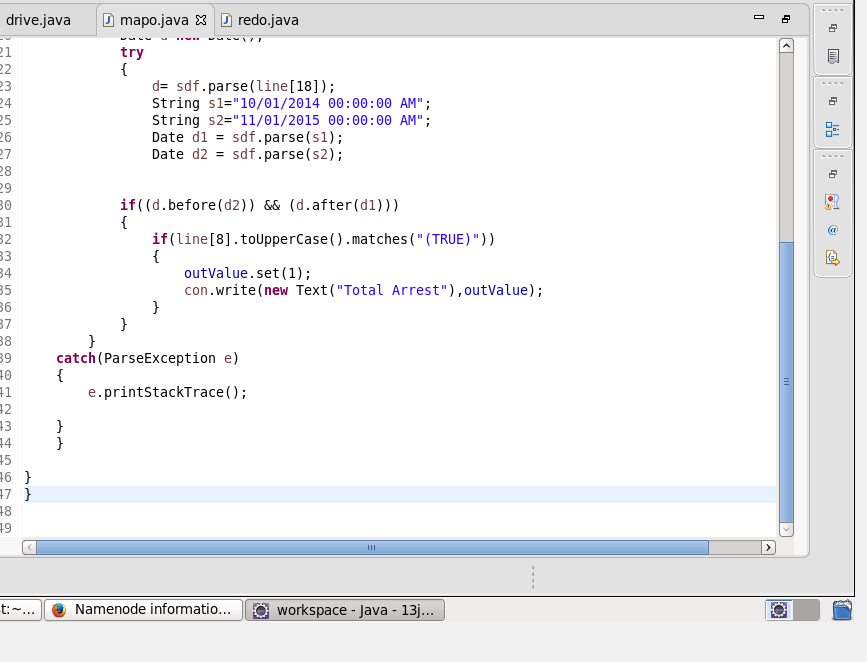


4. Write a mapreduce and pig program to calculate the number of arrests done between October 2014 and October 2015.

Driver will be same as above.

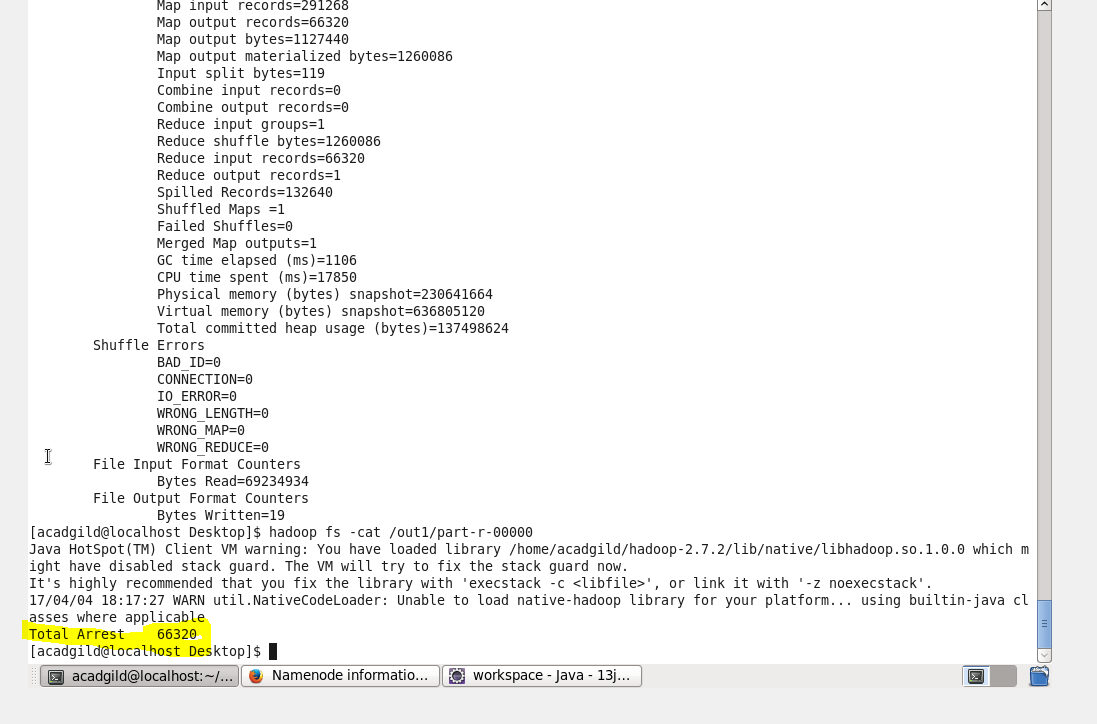
Mapper:





Reducer is the same.

Result:



In pig:

