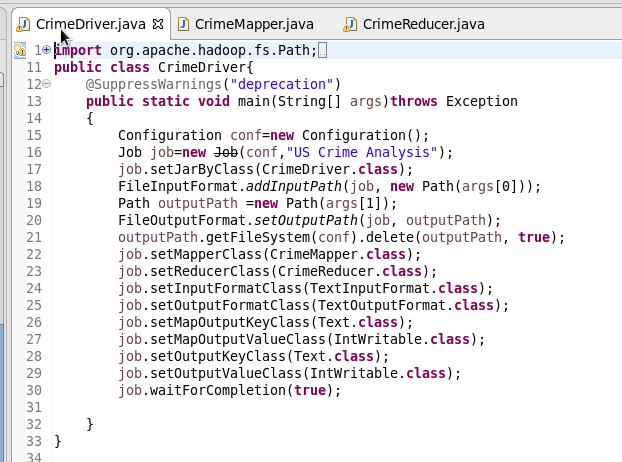
**Project 1**

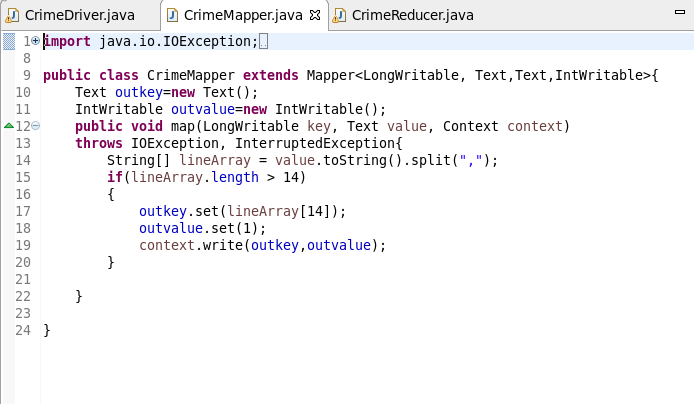
1. **Write a mapreduce and pig program to calculate the number of cases investigated under each FBI code.**

**Mapreduce program:**

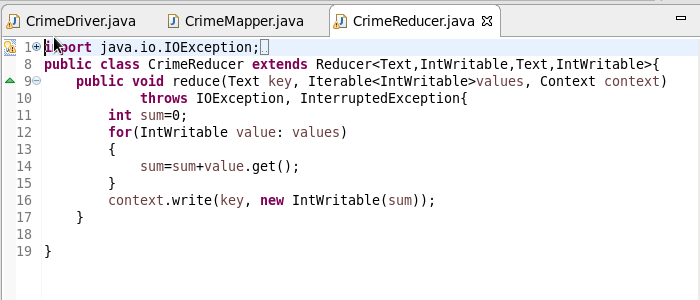
**Driver:**



**Mapper:**

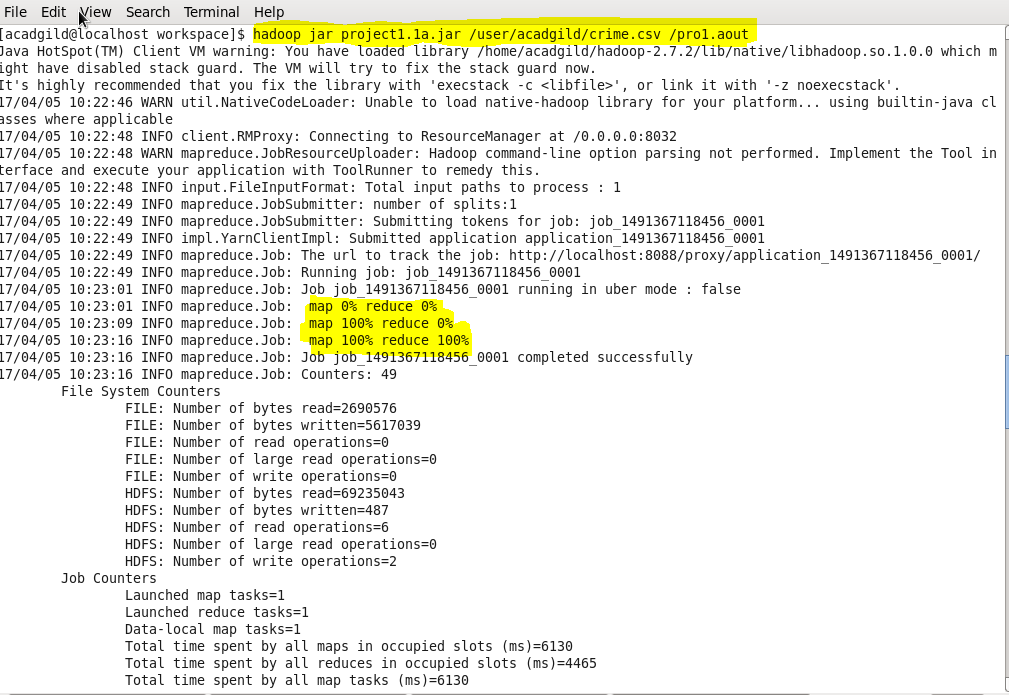


In the mapper class we are splitting the input by the separator **“,”.**We also setting the FBI code as the key and setting the value as 1;

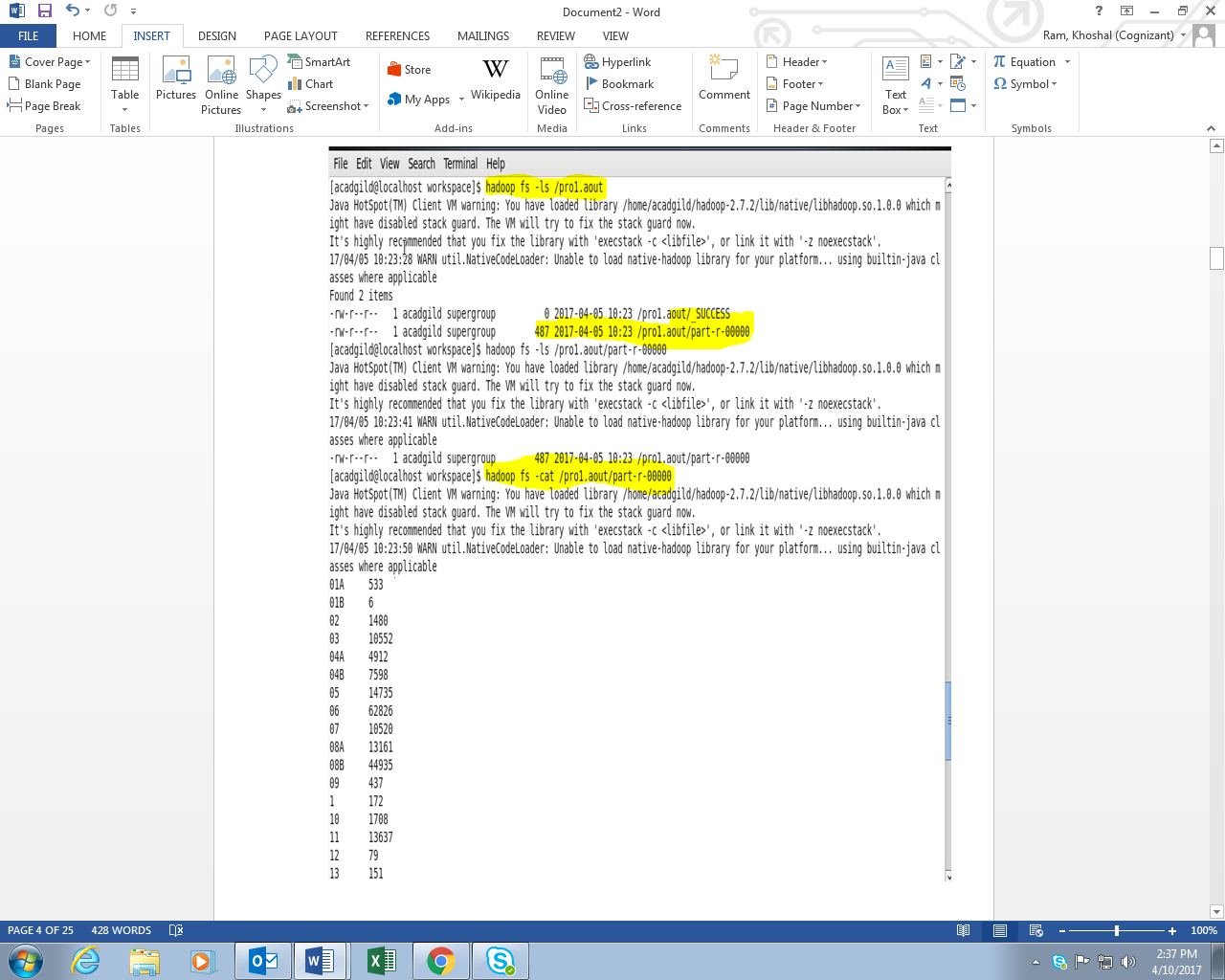
**Reducer:**

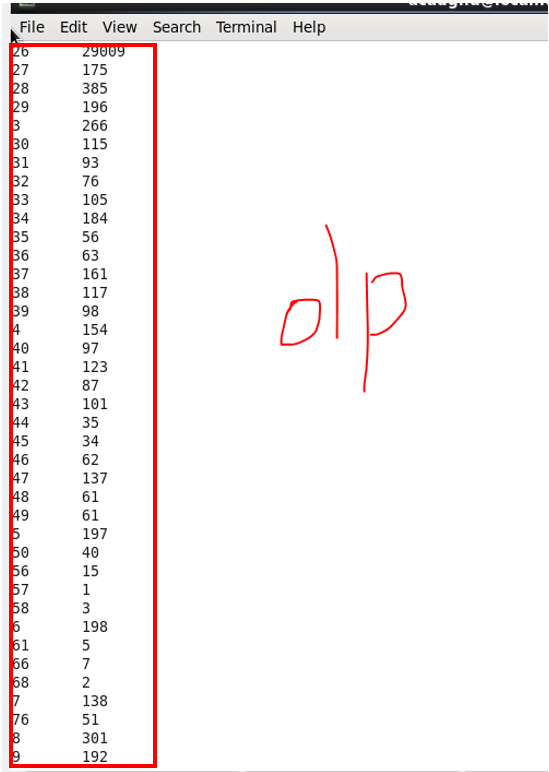
In reducer class we are counting the values corresponding to the keys.

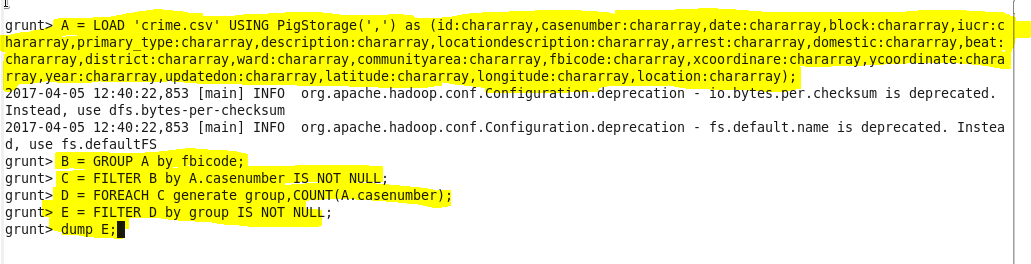
**Running the jar:**



**Output:**





**Pig Script:** 

We are loading the dataset by giving the suitable datatypes.

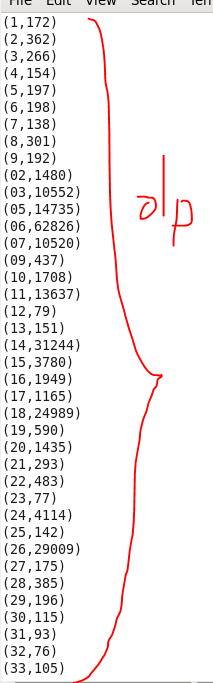
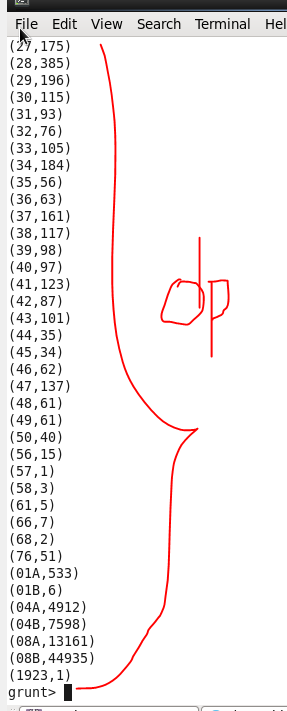
Then grouping it by fbi code.

Then counting them by suing the case number.

Then filtering it by not null.

Finally we are dumping it to get the output.

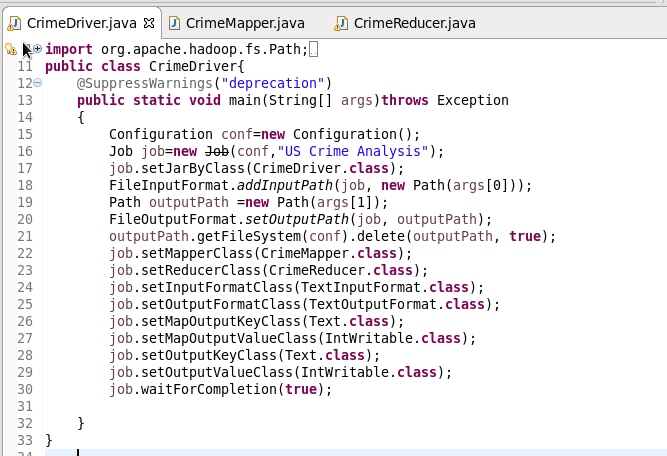
**Output:**

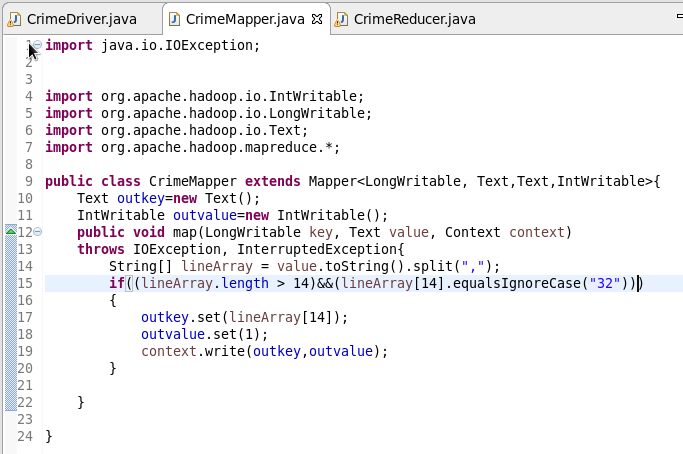


1. **Write a mapreduce and pig program to calculate the number of cases investigated under FBI code 32.**

**Mapreduce program:**

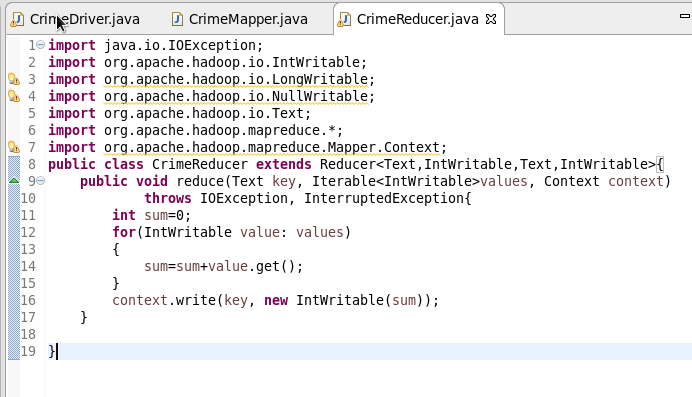
**Driver:**

****

**Mapper: **

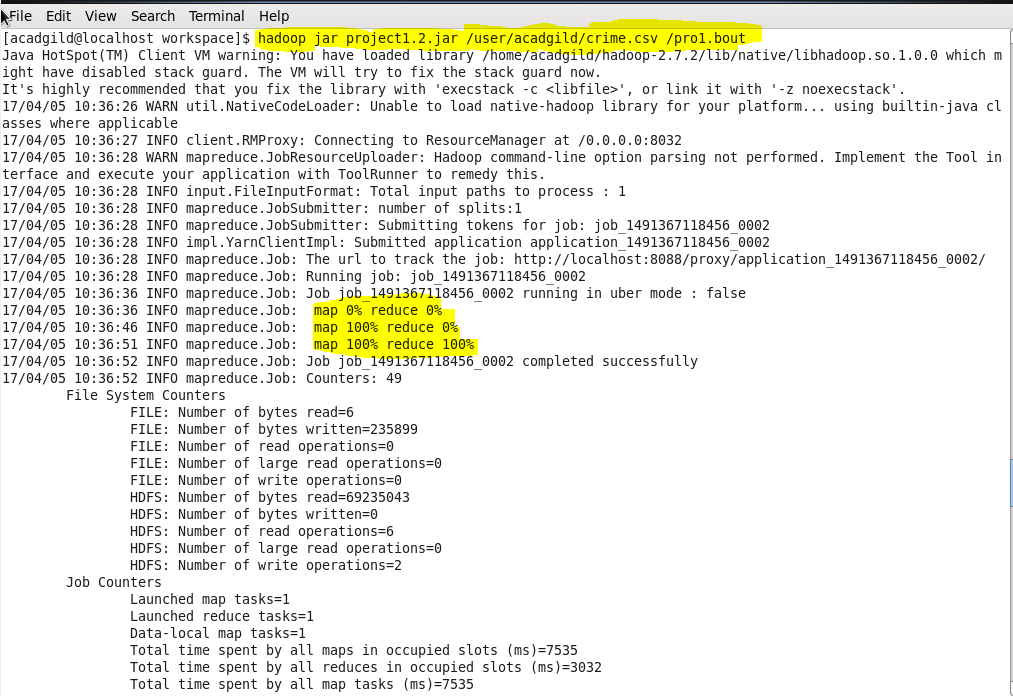
In the mapper class we are filtering the records by considering the fbicode “32” and then setting the key to as fbi code which equlas to “32”.

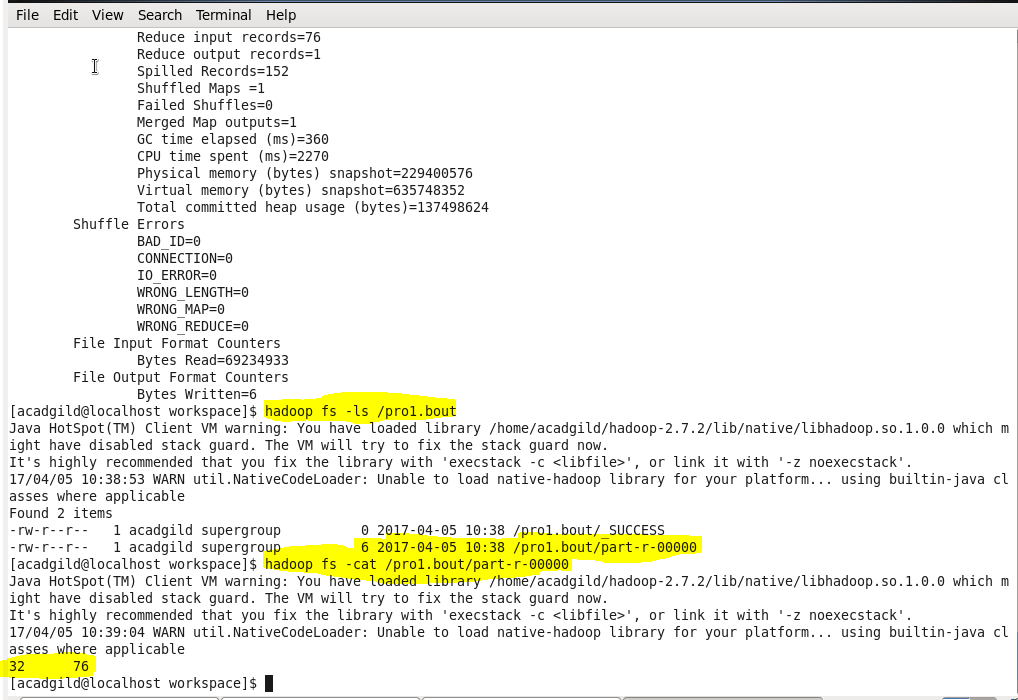
**Reducer:**

****

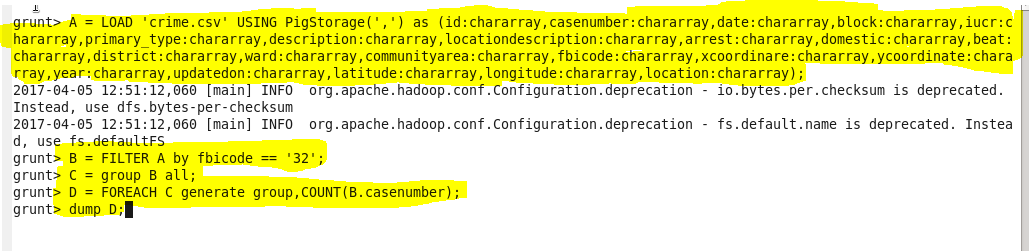
In the reducer class we are counting the number of arrests where the fbi code equals to “32”.

**Output:**

****

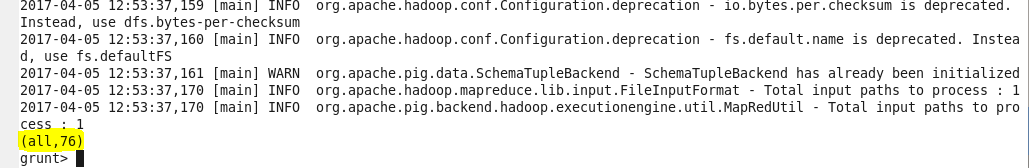
****

**Pig Script:**



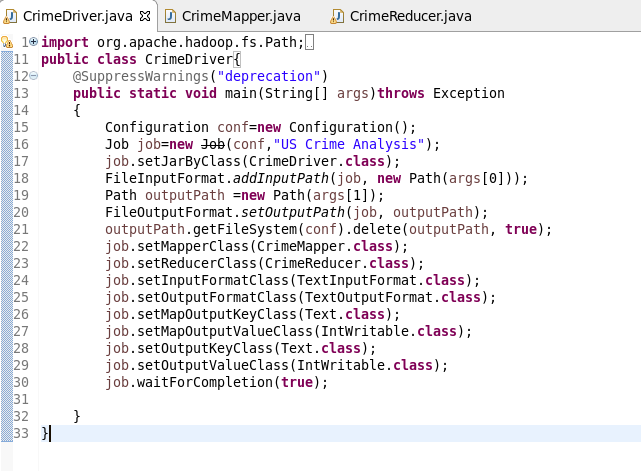
Firstly we are loading the data then we are filtering the data where the fbi code = **“32”** and then grouping it and then we are counting using case number and finally we are dumping to get the output.

**Output:**

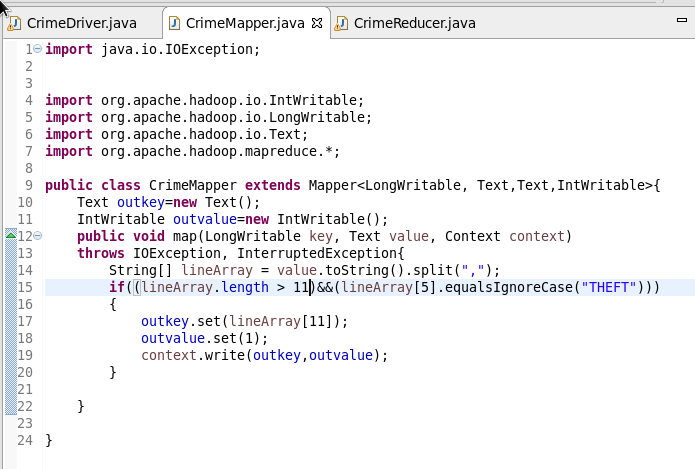


1. **Write a mapreduce and pig program to calculate the number of arrests in theft district wise.**

**Mapreduce program:**

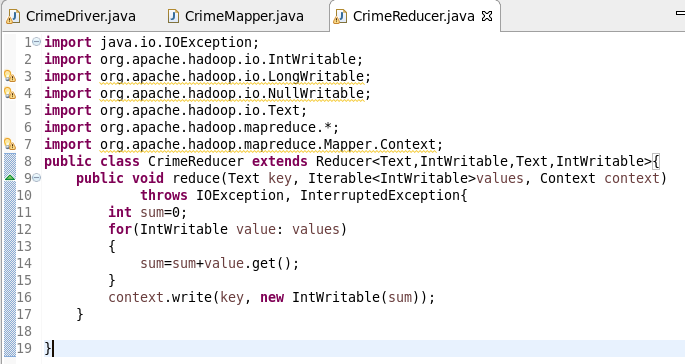
**Driver: **

**Mapper:**

****

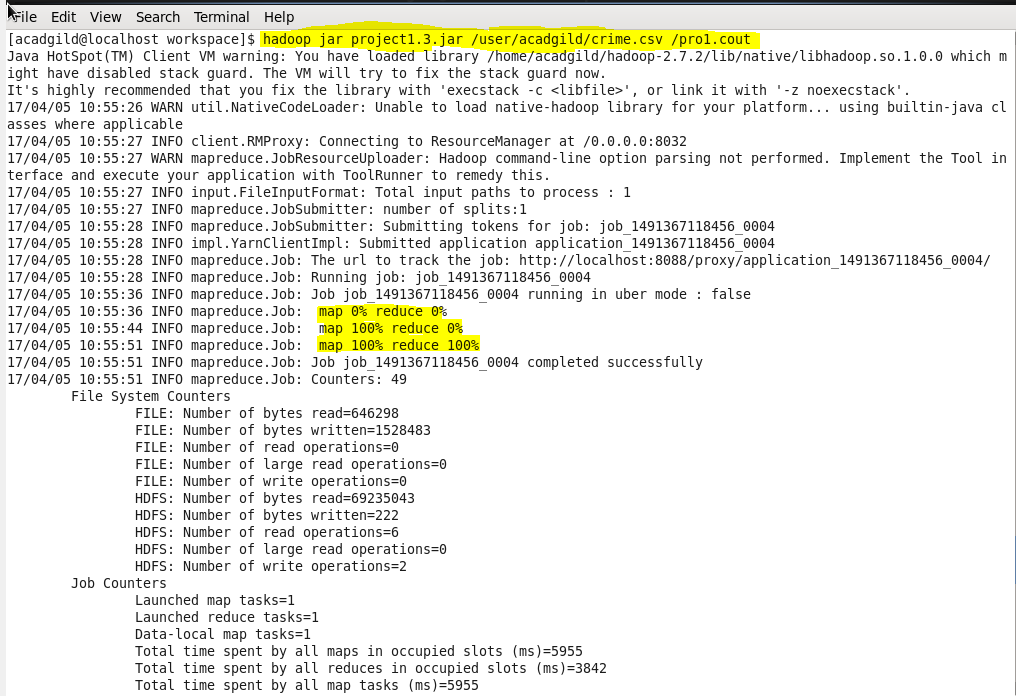
In the Mapper class we are filtering the data by having the primary type as **“Theft”.** Then setting the key as district and the setting the value as 1 to it.

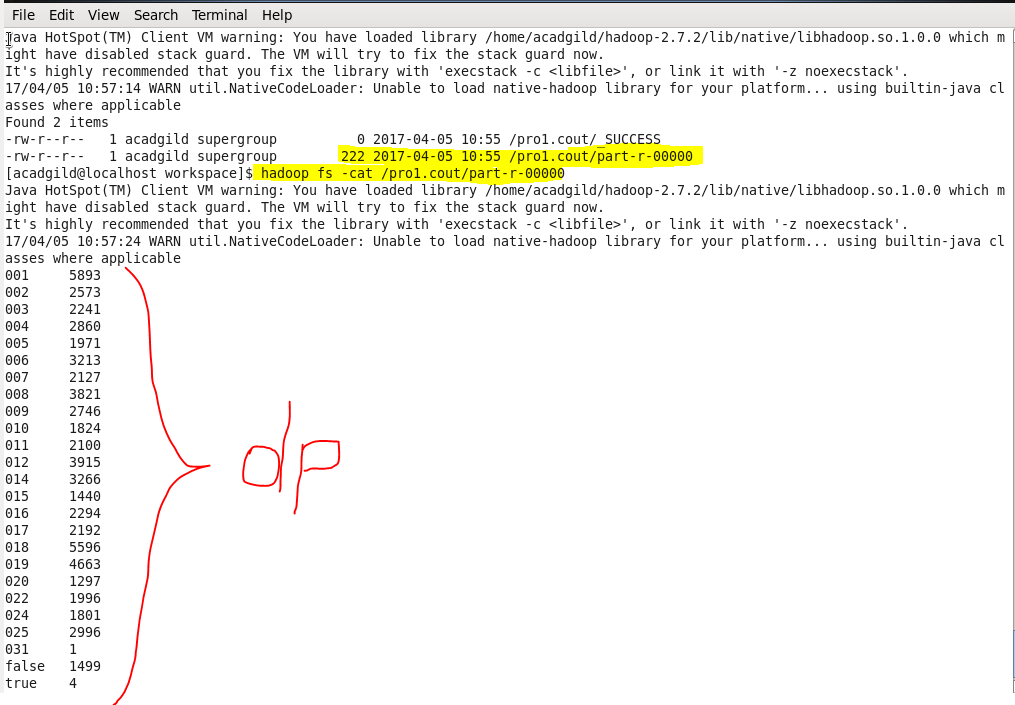
**Reducer:**

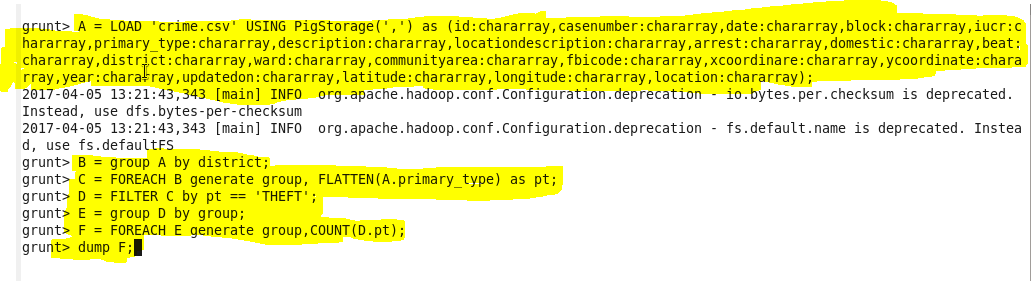
****

In Reducer class we are counting the number of arrests where primary type as “THEFT” and counting them district wise.

**Output:**

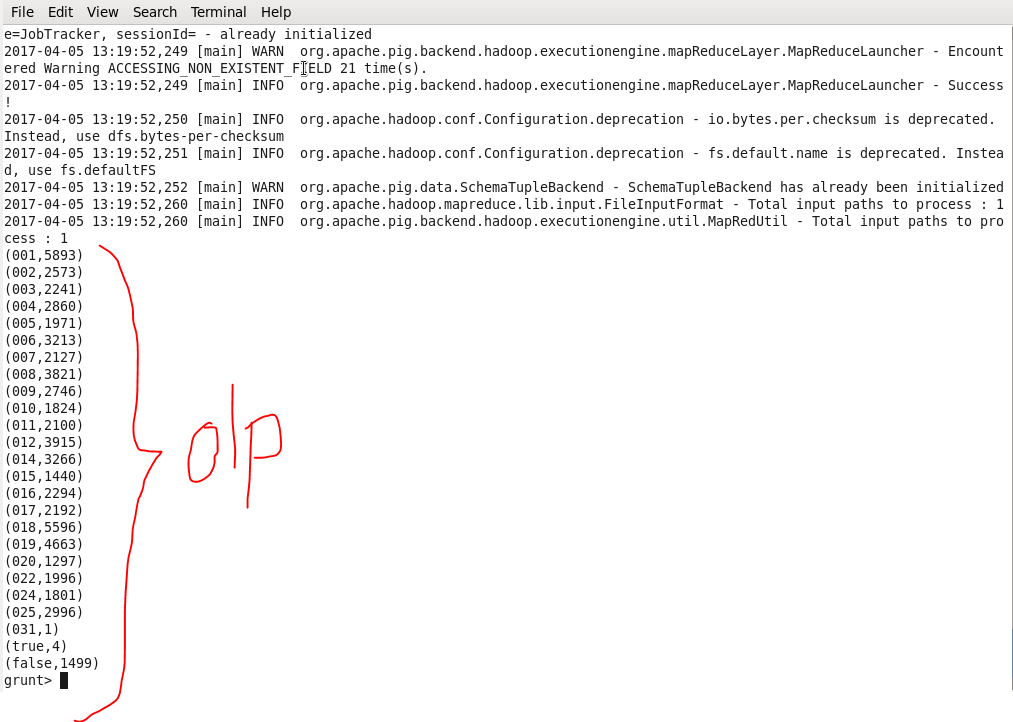




**Pig Script:** 

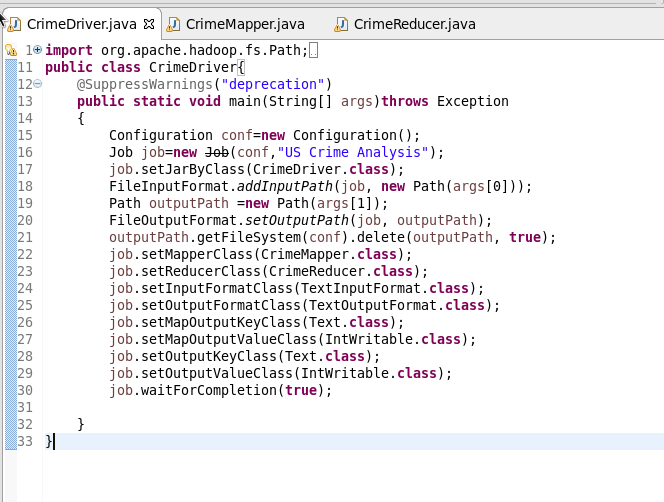
We are loading the data and then grouping it by district and then filtering it by primary type =”THEFT” and grouping it and then count the number of arrests and then finally dumping it to get the output.

**Output:**

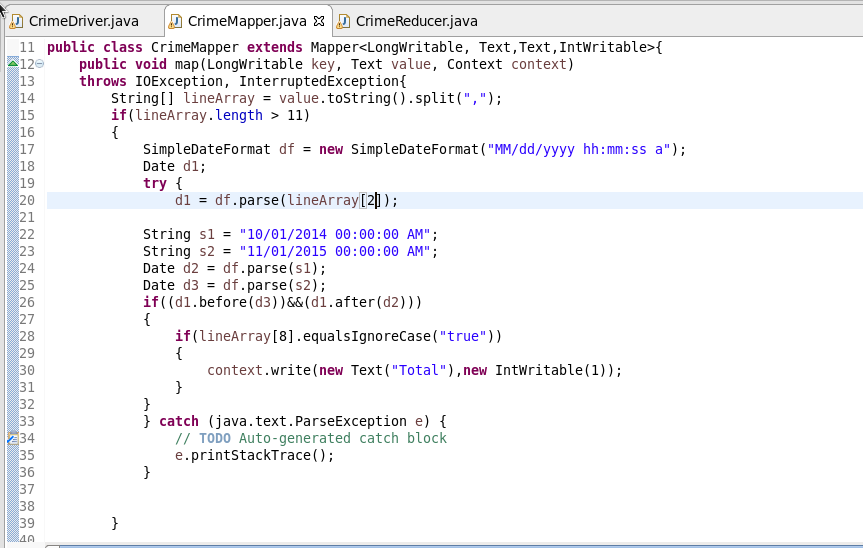


1. **Write a mapreduce and pig program to calculate the number of arrests done between October 2014 and October 2015.**

**Mapreduce program:**

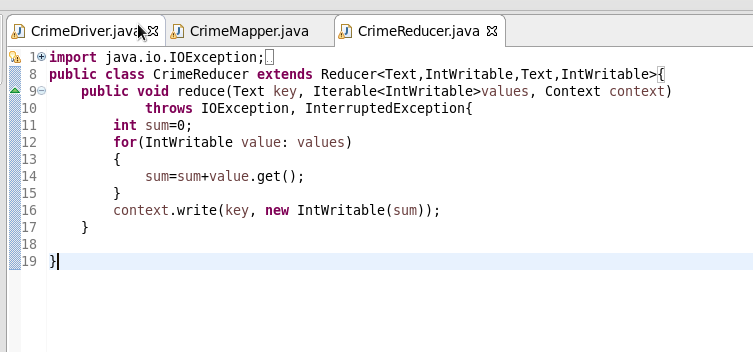
**Driver: **

**Mapper:**

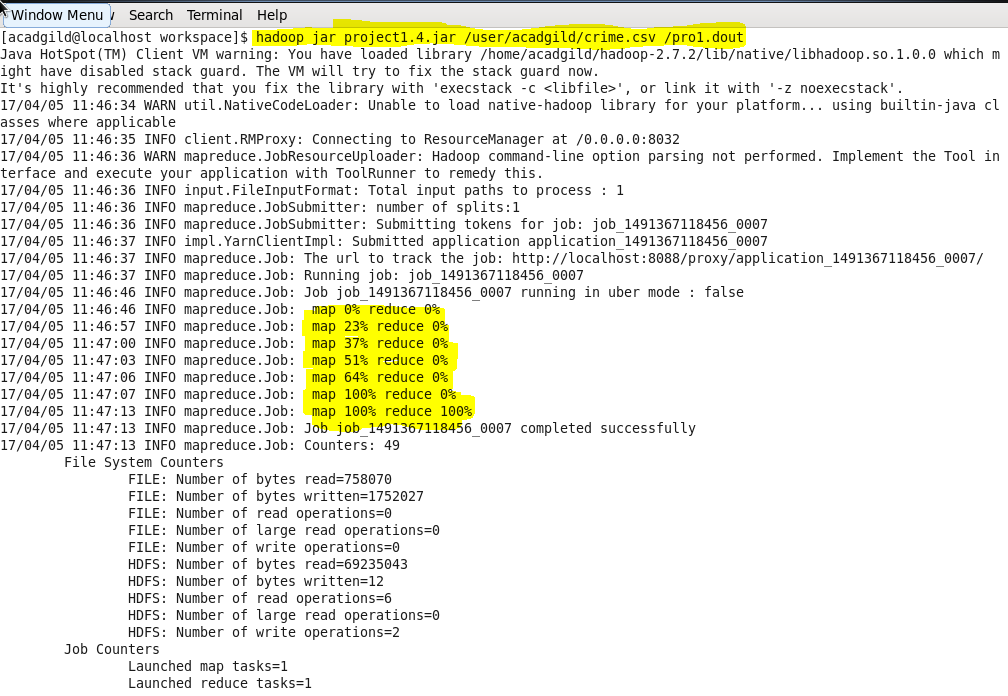
****

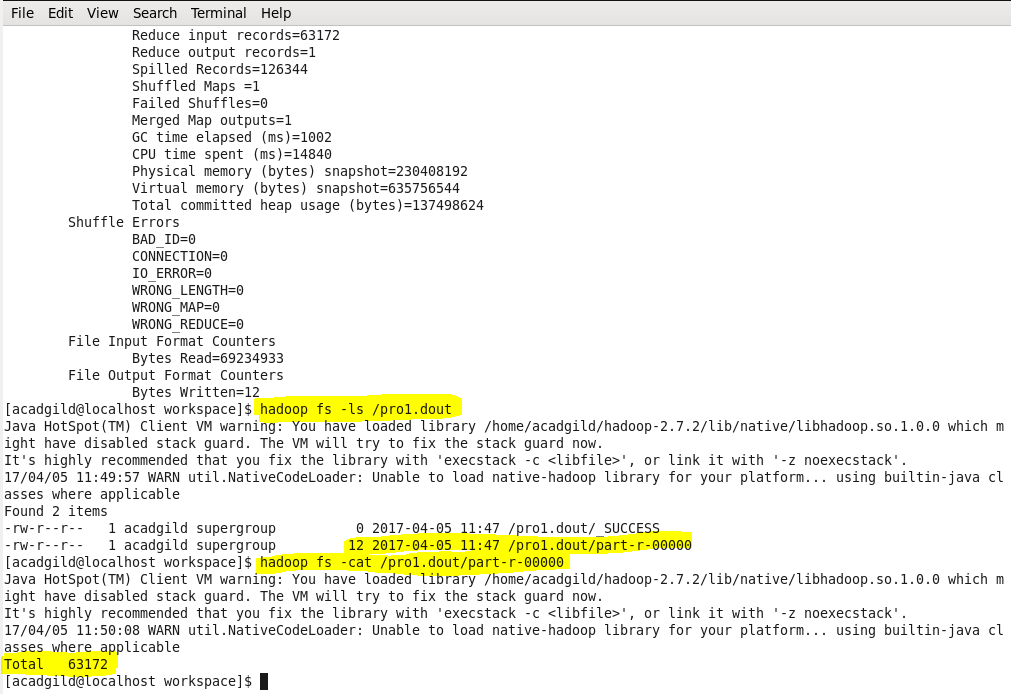
In the Mapper class we are filtering the records by considering the date which is between october 2014 and october 2015.

**Reducer:**

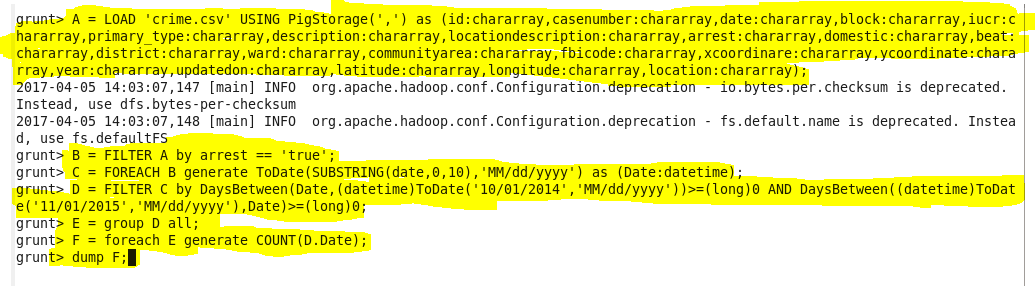
****

**We** are counting the number of arrests between October 2014 and October 2015.

**Output:** 

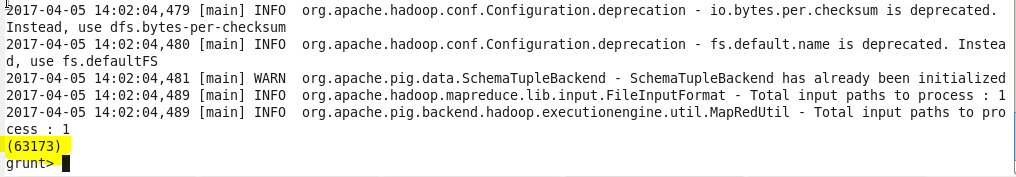


**Pig Script:**

******

First we are loading the data and then filtering the arrests where they are true and filtering the data where are arrests are between October 2014 and October 2015.

**Output:**

******