**PROJECT -1**

Project 1 - USA Crime Analysis

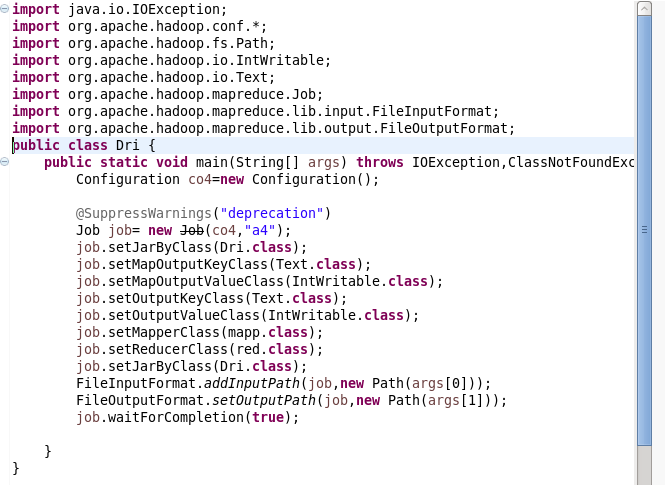
**DATA SET** : ID, Case Number, Date, Block, IUCR, Primary Type, Description, Location Description, Arrest, Domestic, Beat, District, Ward, Community Area, FBI Code, X Coordinate, Y Coordinate, Year, Updated On, Latitude, Longitude, Location

**Problem:**

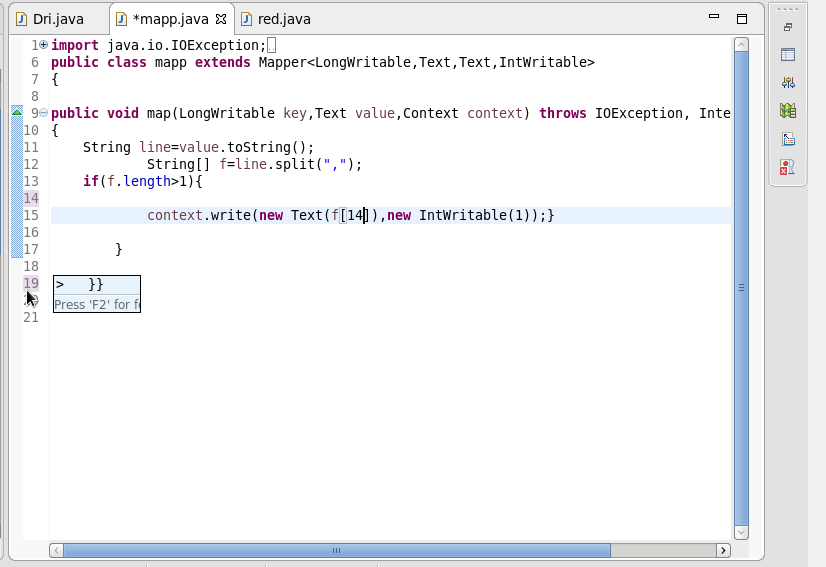
**1. Write a mapreduce and pig program to calculate the number of cases investigated under each FBI code**

Mapreduce program: In this prob, we have to calculate the no of cases in each code so at first we are giving the input data file name and then output path to store the output in the driver class

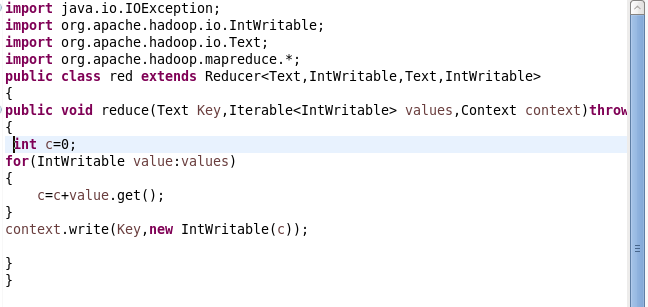
***Driver class:***



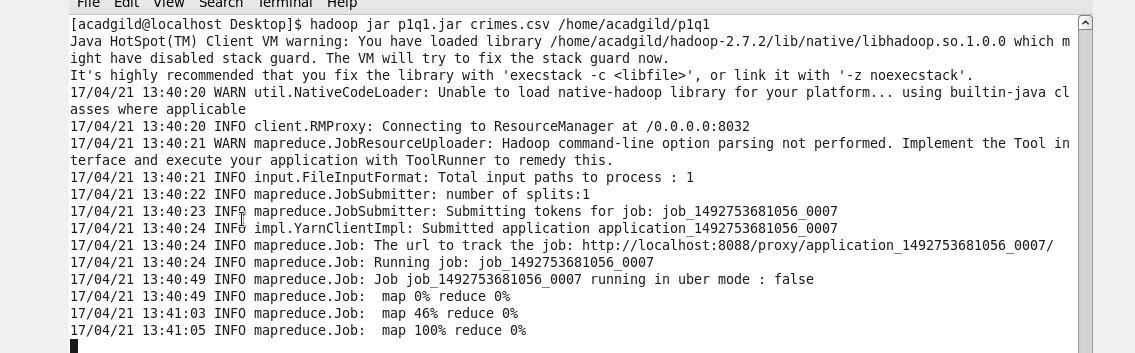
***Mapper class:*** In this the key is given as fbi code and the value is increased in reducer class so the counting of cases can be done



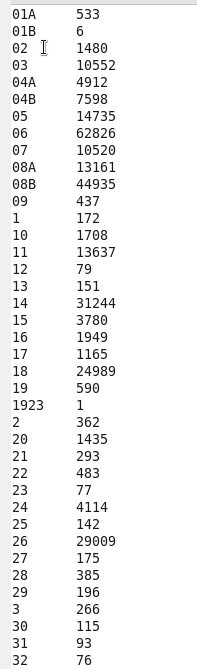
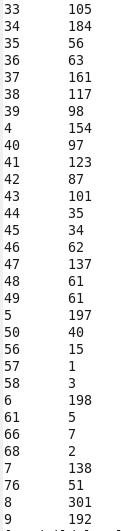
***Reducer class:***



***Run command:***



***Output terminal:***

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

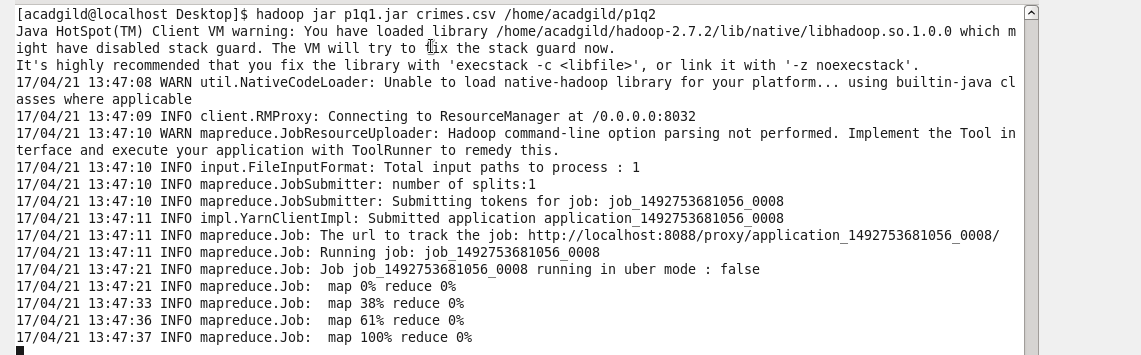
This is similar to the previous question , except that here the specific fbi code=32 is given ,which we can obtain by giving if condition

Driver class and reducer class are same as that of above exam

***Mapper class:***



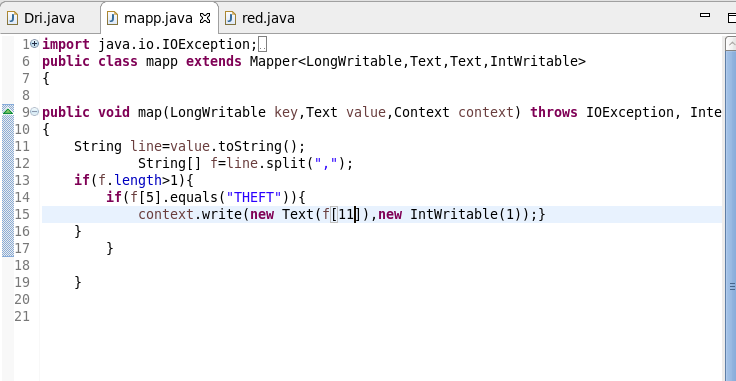
***Run command***:



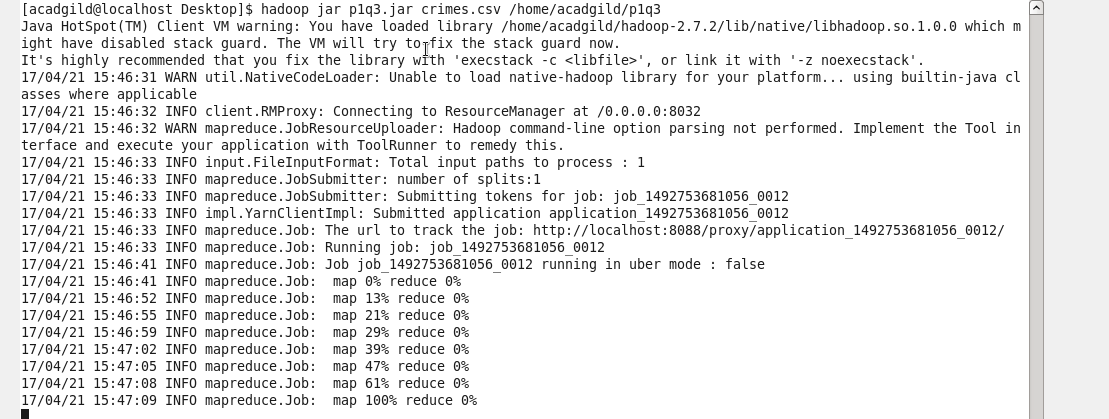
***Output terminal***: C:\Users\612912\Pictures\p1q2 mp op.PNG

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

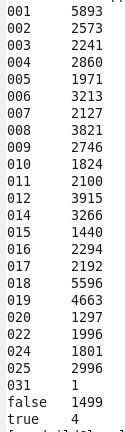
***Mapper class:*** In this ques ,first we are filtering by arrest =theft and then grouping by district code



***Run:***

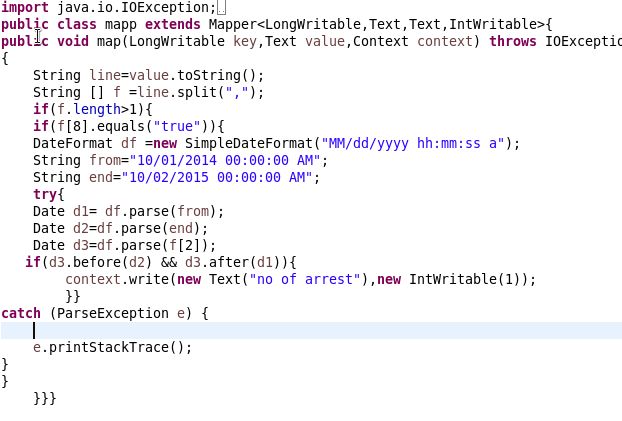


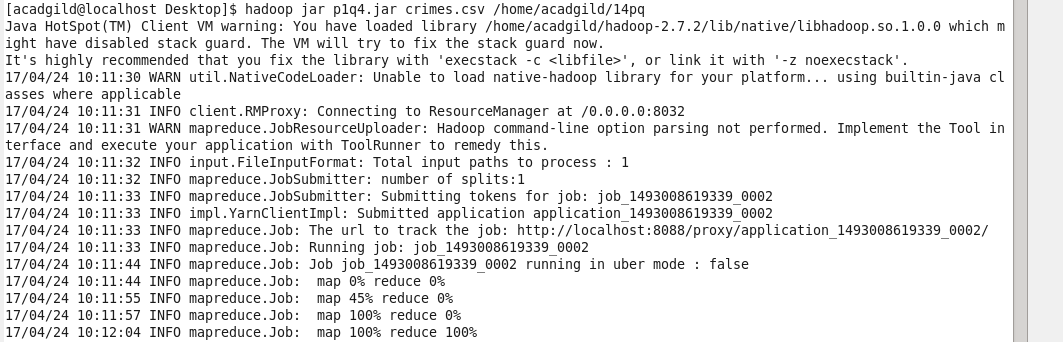
***Output terminal:***



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

***Mapper class***: in this we are filtering by arrest =true and then filtered by date between oct 2014 and oct 2015



***Run:*** 

***Output terminal***:

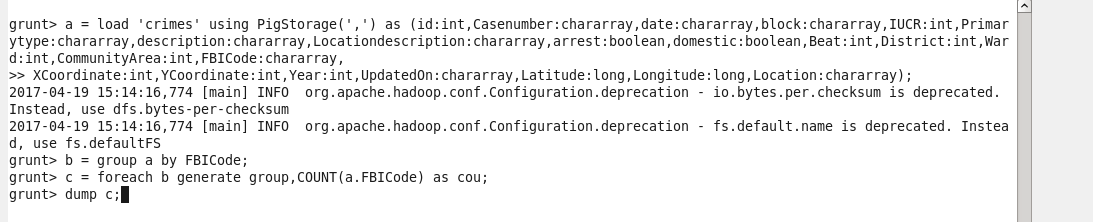
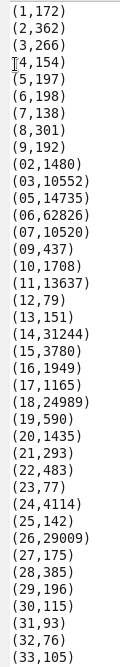
C:\Users\612912\Pictures\proj 1.4 op.PNG

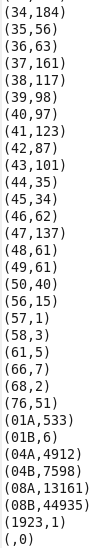
Using Pig commands

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

**Pig command:**

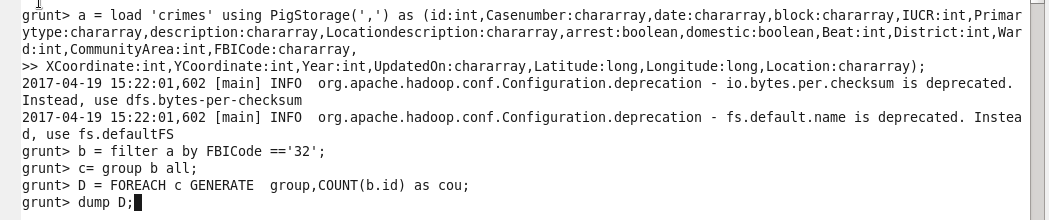
At first ,we are loading the data file into pig and then we grouping by fbi code and then foreach group we are counting the no of cases by using fbi code

**Output terminal:**

****

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

**Pigscript:** Here we are filtering the fbi code of 32 and then counting the complaint id.

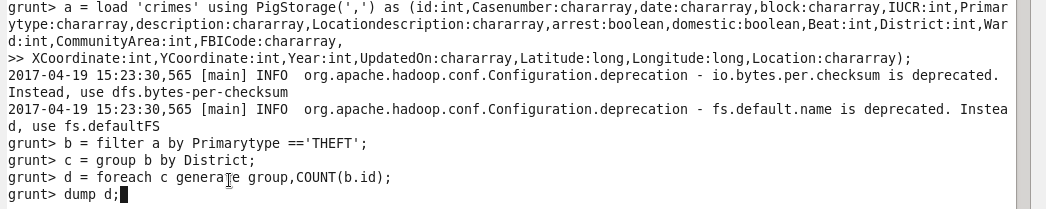
****

**Output terminal:**

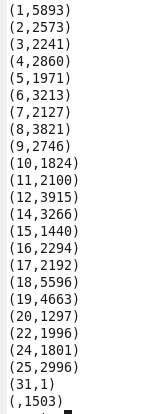
**C:\Users\612912\Pictures\proj 1.2 op.PNG**

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

**Pigscript**: herewe are filtering by arrest and then grouping by district we are counting the no of cases

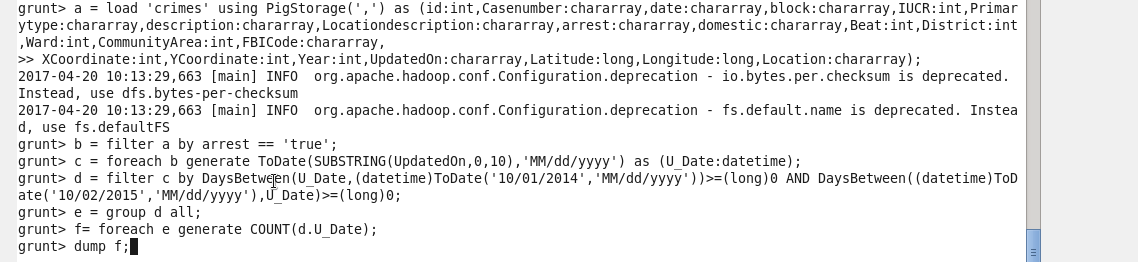
****

**Output terminal:**

****

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

**Pigscript:** here we are first checking whether they are arrested and then the arrested date is obtained which we are storing as U\_date and then we are finding the difference b/w oct 2014 , oct 2015 with U\_date checking whether they are before oct 2015 and after oct 2014

****

**Output terminal:** **C:\Users\612912\Pictures\proj 1.4 op.PNG**