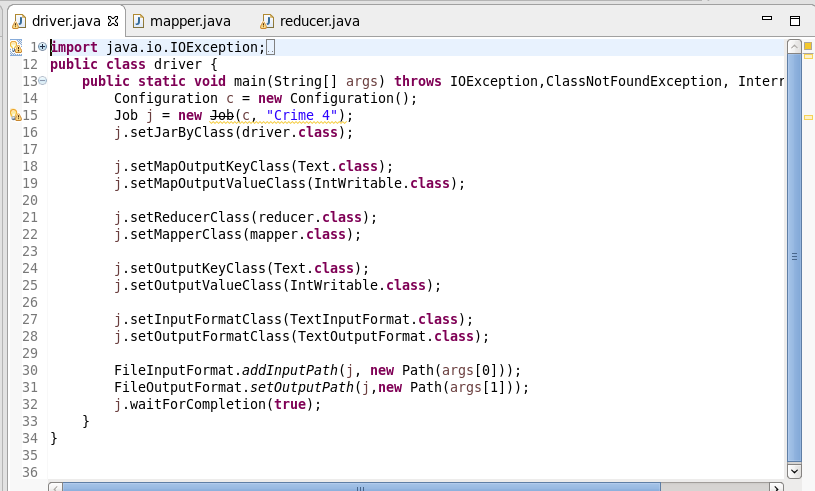
4. Write a mapreduce and pig program to calculate the number of arrests done between October 2014 and October 2015.

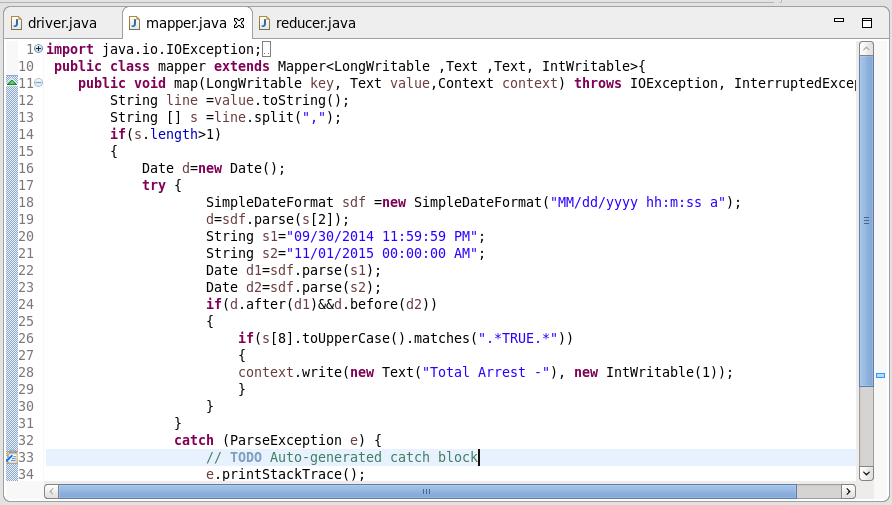
Driver Class:



Here we are setting the job configurations in the Driver Class and setting the Output and Input File Formats for the Mapper in the line 18 and 19. In the line 21 and 22 we are setting the Mapper and reducer class.

Here I have taken the input and output format as the “Text Format class” in the line 27, 28.

Mapper Class:



In this mapper we have using simpledateformat. First selected the date and parsed the date to check the format and convert the given string into the date format.

The Date considered is

Starting Date– 1st Oct 2014 and Ending Date– 1st Nov 2015

Since it is said between the October 14 and October 15. Hence the October moth is considered.

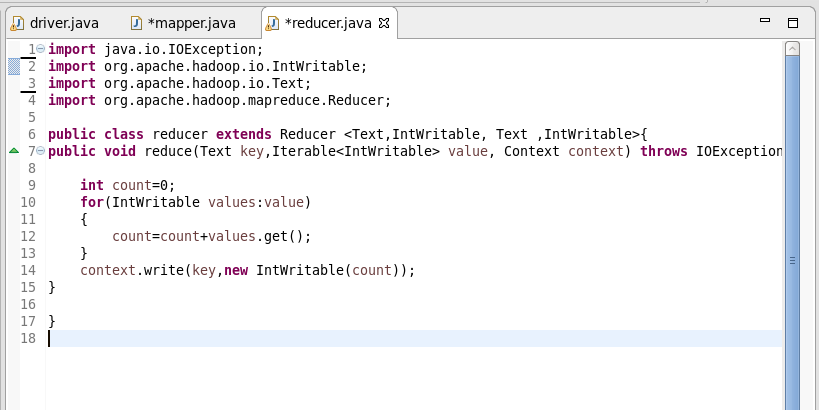
The Format of the date is Month/day/year.

By using the condition of before and after functions date. We can check if the given date is between the specified date. If the date is between the mentioned date then we are passing simple Text as key

“Total Arrest -”

The value passed is intwritable 1.

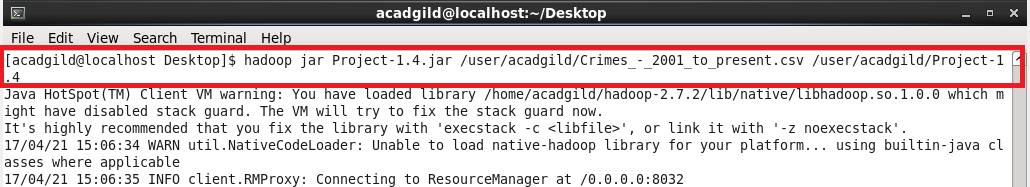
Reducer Class:



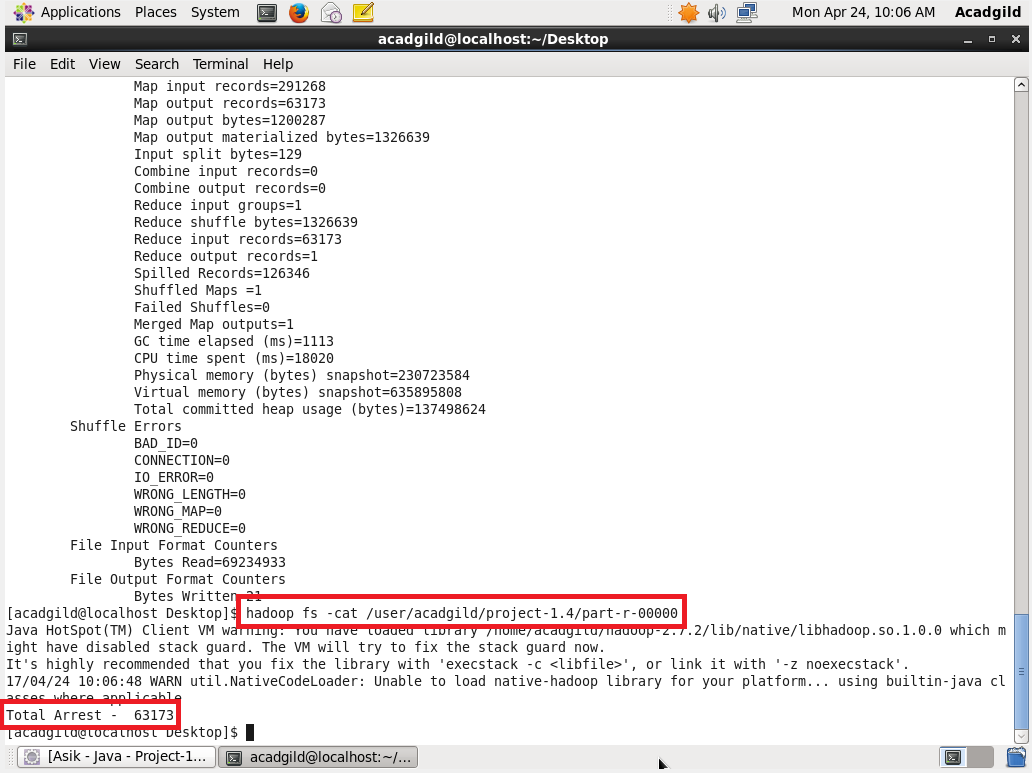
In the reducer we are summing up the values which are related each key and passing the key and the sum as the value in the intwritable format.

Here as we have the single key all the values related to that key will be added.

Running the Jar Files:

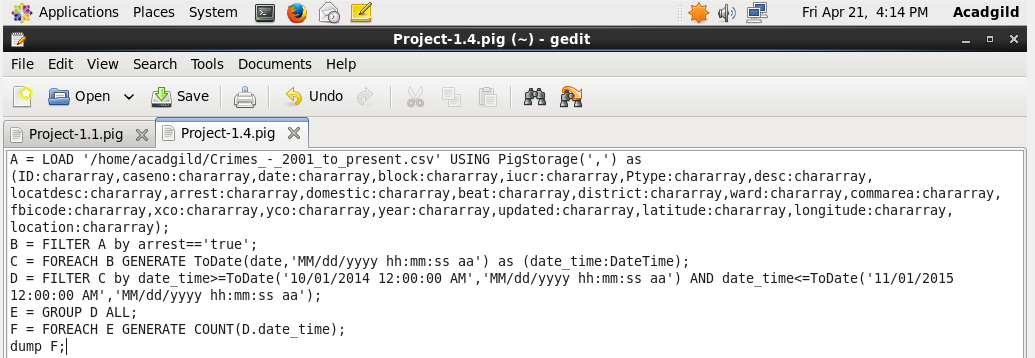


Output:



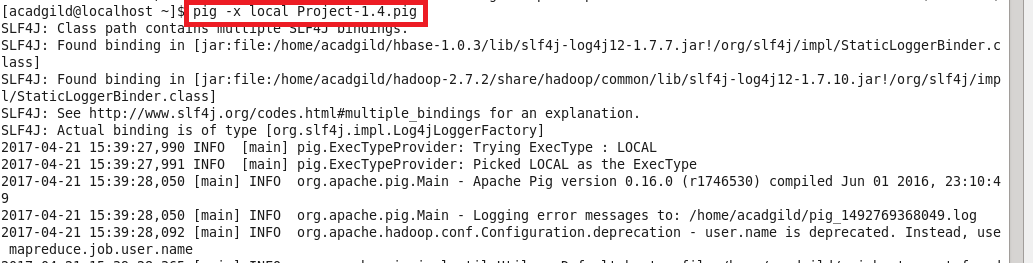
Pig Program:

Pig Script:



In this pig script first loaded the fields with respective data type. Filtered the data if the arrests are true.Then as we have to use the date function so convert the date fields into the date format using the TO DATE fuction in the pig. Then filtered the dates with the AND operator to check before and after of the mentioned date.

Running the Script:



Output:

