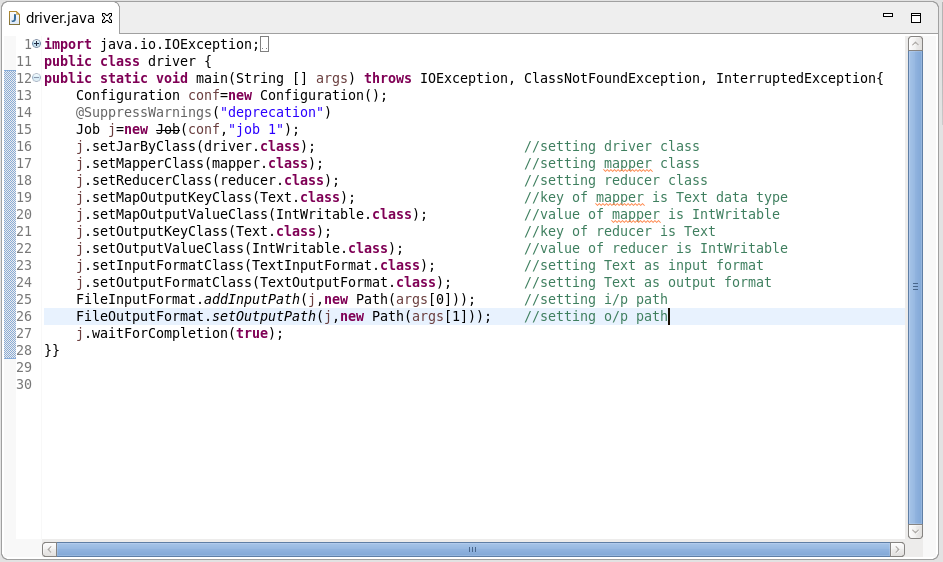
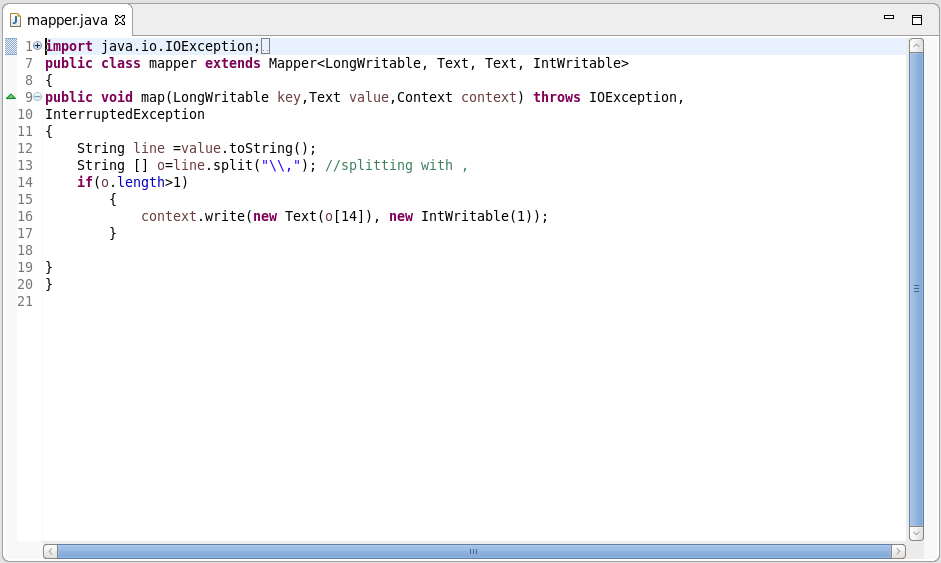
1. **WRITE A MAPREDUCE AND PIG PROGRAM TO CALCULATE THE NUMBER OF CASES INVESTIGATED UNDER EACH FBI CODE.**

**DRIVER CLASS:**

****

**MAPPER CLASS:**

****

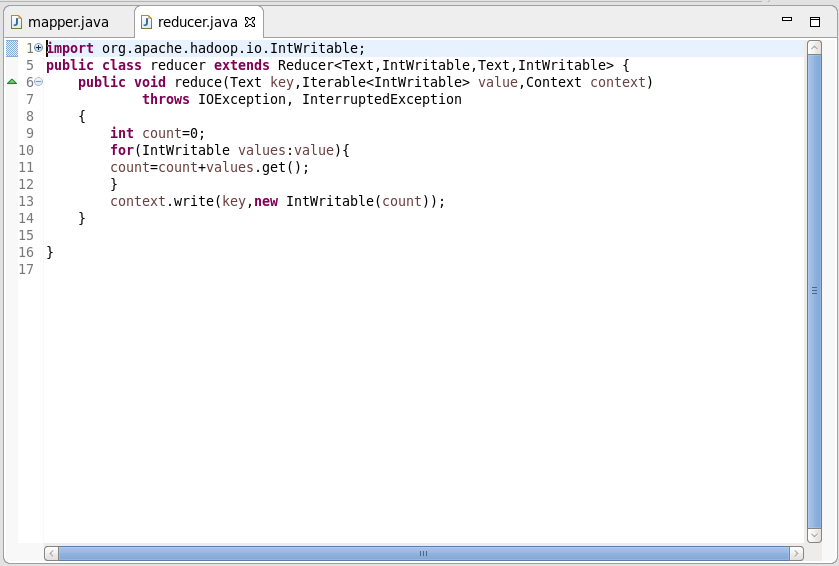
**MAPPER LOGIC:**

*LINE 9 : Text offset from beginning is set as key and the entire dataset is set as value*

*LINE 13 : Separating the dataset with comma and storing it in array o.*

*LINE 16: Setting FBI CODE as key and 1 as value to be further processed by reducer.*

**REDUCER CLASS:**

****

**REDUCER LOGIC:**

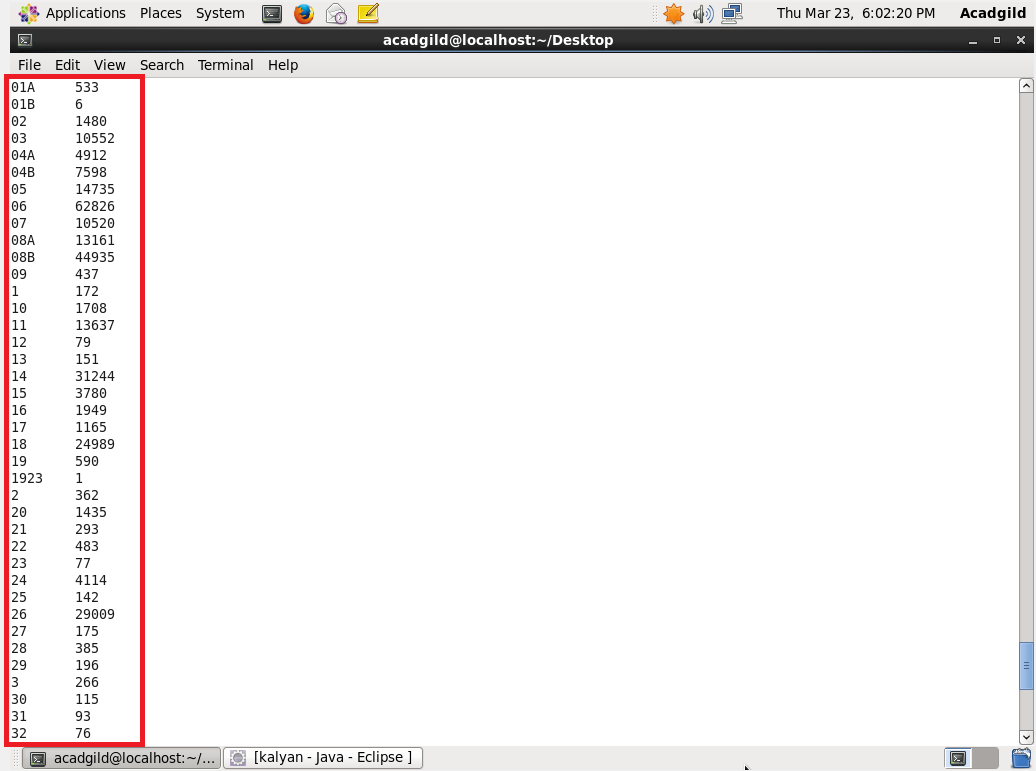
*LINE 9: Initializing the count as 0*

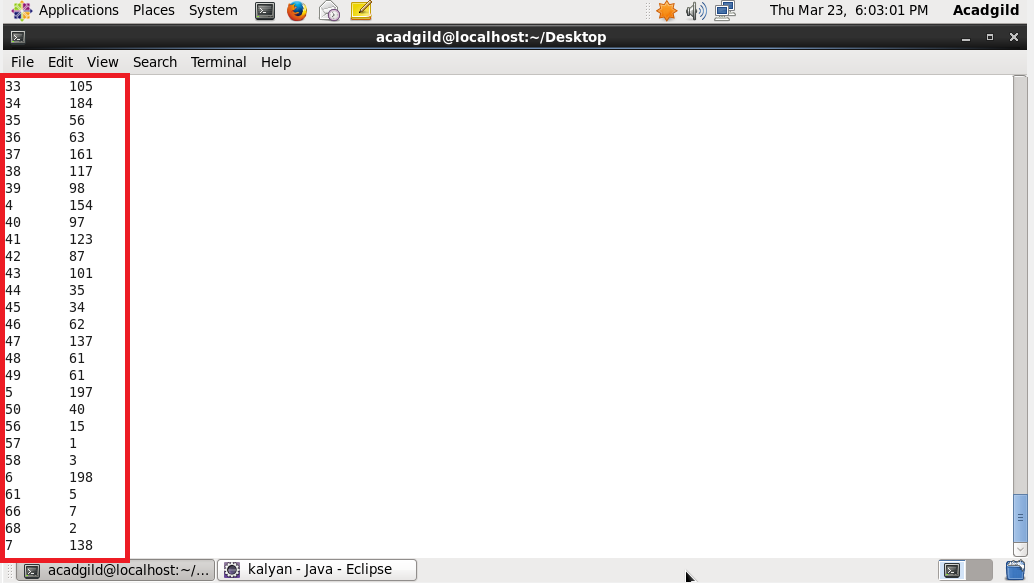
*LINE 10 : Running an infinite for loop*

*LINE 11: count will add the iterative values*

*LINE 13: Finally the same key is used as the key and the count is set as value.*

**OUTPUT:**

****

****

**USING PIG:**

**LOGIC:**

*1) Loading the Dataset into the grunt shell using LOAD command*

*2) Grouping by FBI code (column 15) using GROUP BY command*

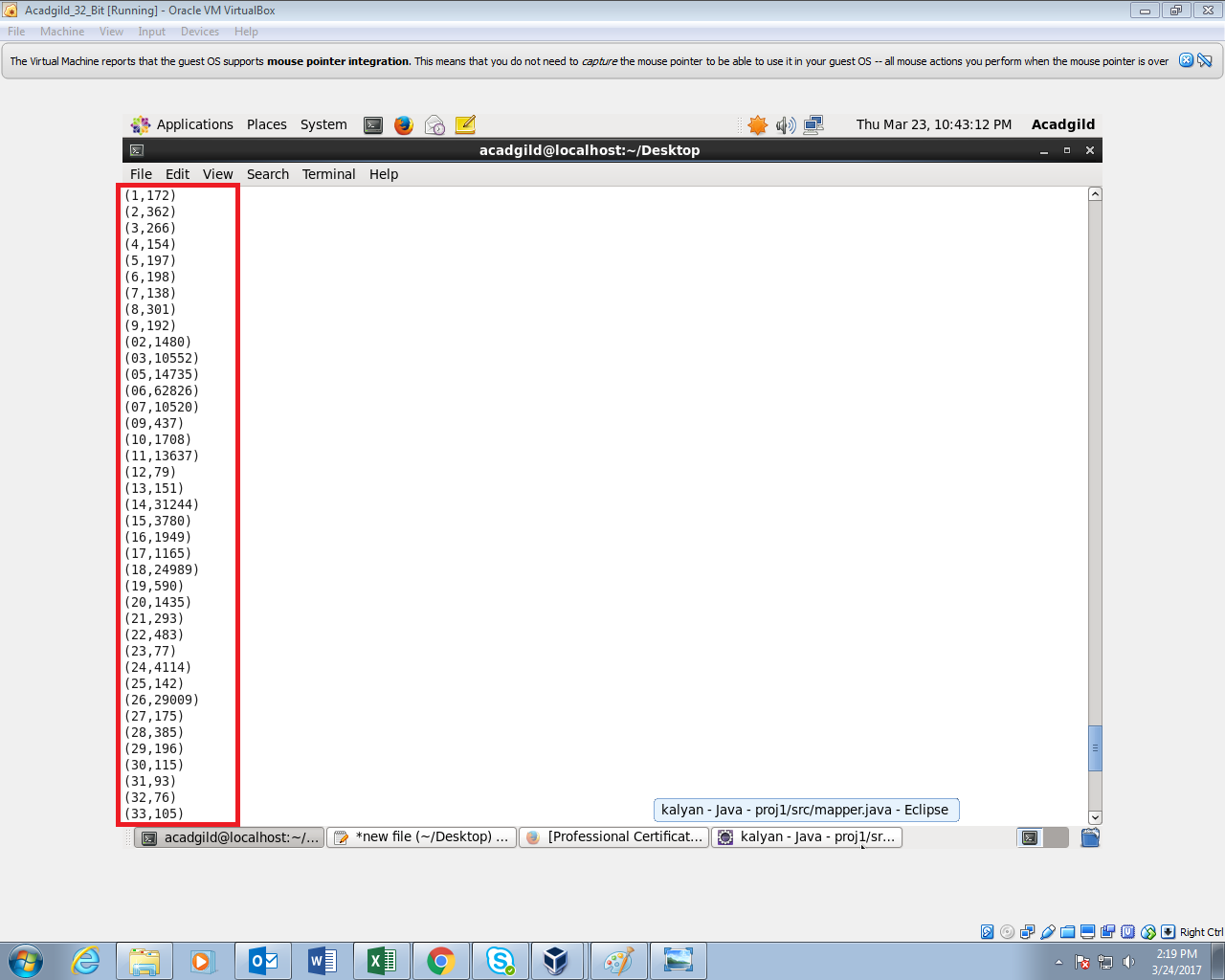
*3) Generating count of cases for each FBI CODE is done by COUNT command*

*4) Dumping the ouput to get the count.*

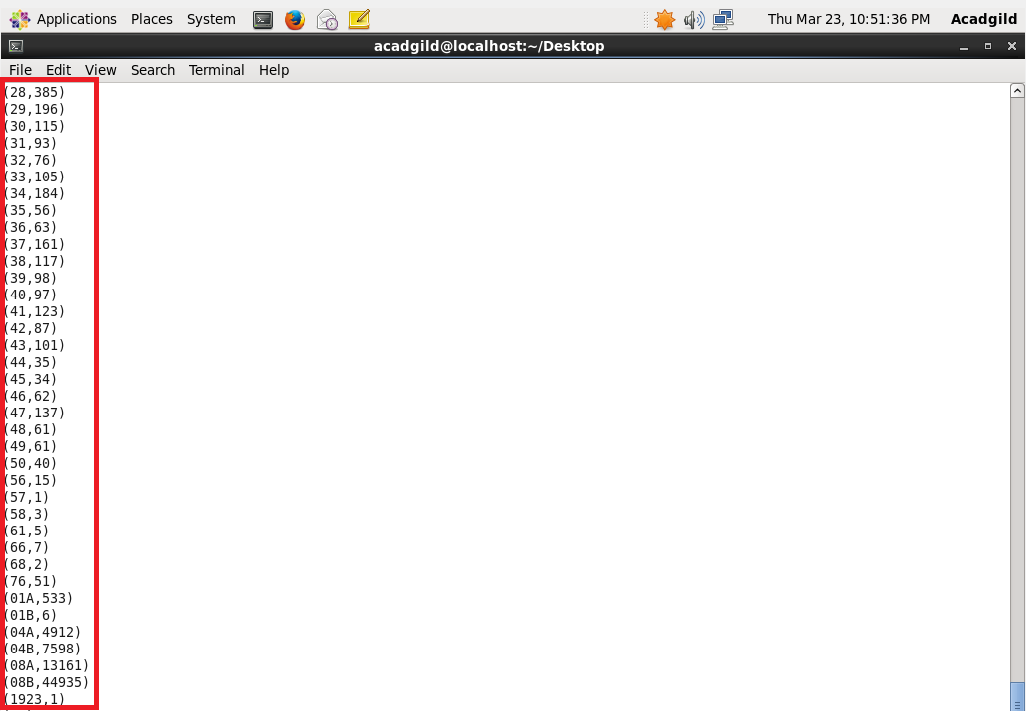
**PIG CODE:**

****

**OUTPUT:**

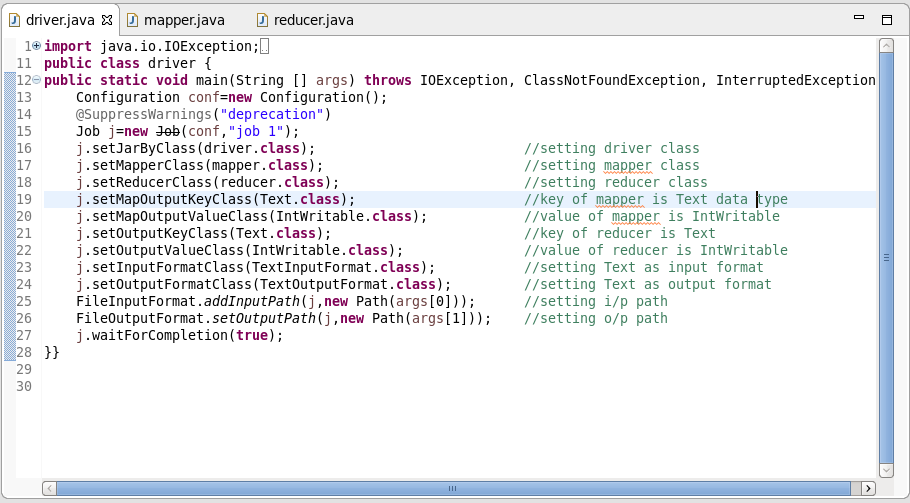
****

**OUTPUT CONTINUATION:**

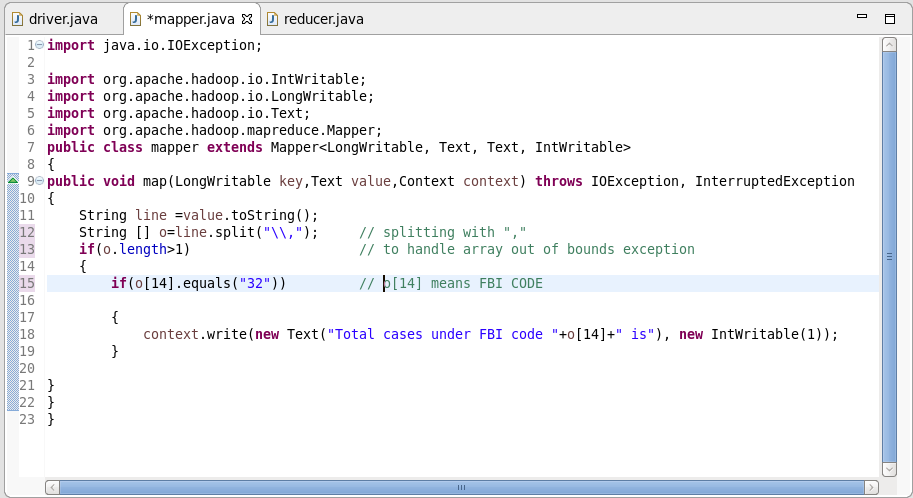
****

**2) WRITE A MAPREDUCE AND PIG PROGRAM TO CALCULATE THE NUMBER OF CASES INVESTIGATED UNDER FBI CODE 32.**

**DRIVER CLASS:**

****

**MAPPER CLASS:**

****

**MAPPER LOGIC:**

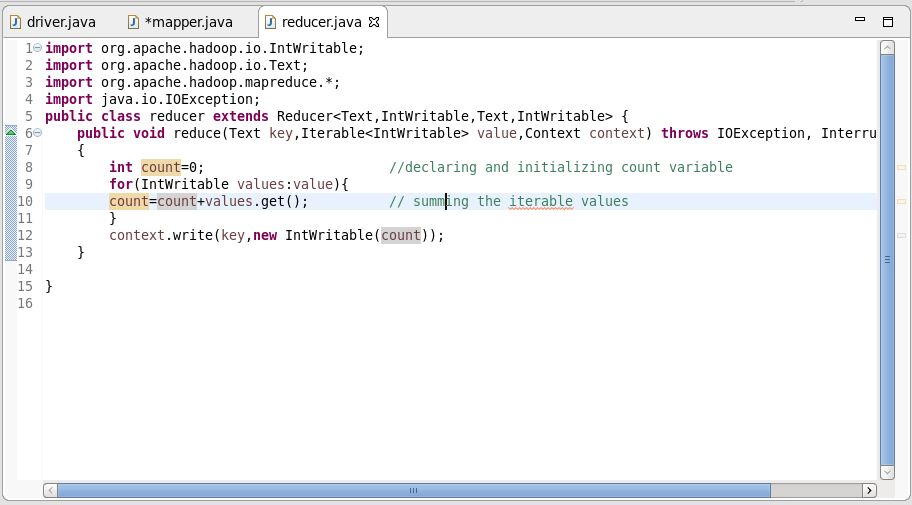
*LINE 9 : Text offset from beginning is set as key and the entire dataset is set as value*

*LINE 12 : Separating the dataset with comma and storing it in array o*

*LINE 15: checking if FBI CODE O[14] is 32(Filtering the FBI CODE=32)*

*LINE 18: Setting FBI CODE as key and 1 as value*

**REDUCER CLASS:**

****

**REDUCER LOGIC:**

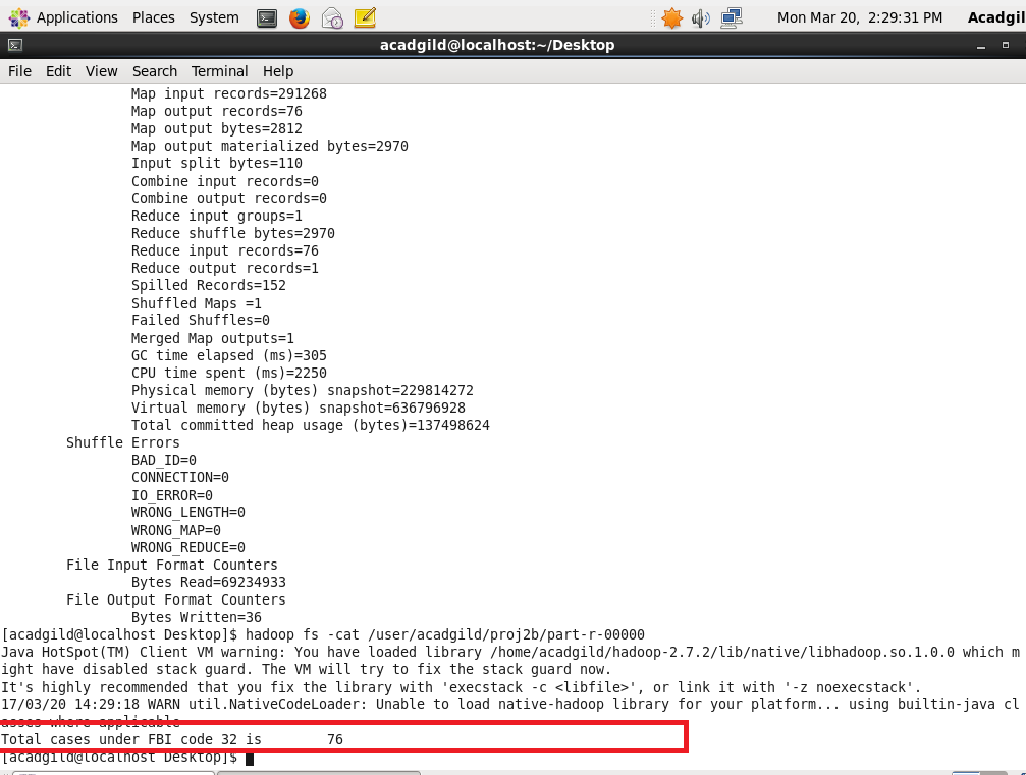
*LINE 8: Initializing the count as 0*

*LINE 9 : Running an infinite for loop*

*LINE 10: count will add the iterative values*

*LINE 12: Finally the same key is used as the key and the count is set as value.*

**OUTPUT:**

****

**USING PIG:**

**LOGIC:**

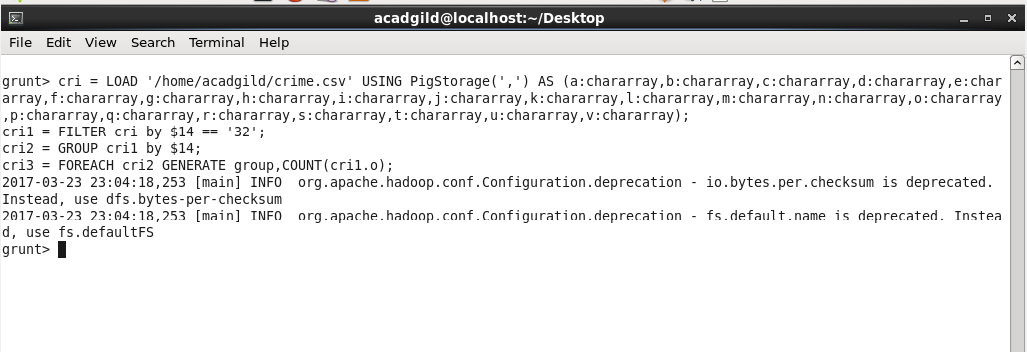
1. *Loading the Dataset into grunt shell using LOAD command.*

*2) Filtering the Dataset by FBICODE=32 using FILTER command.*

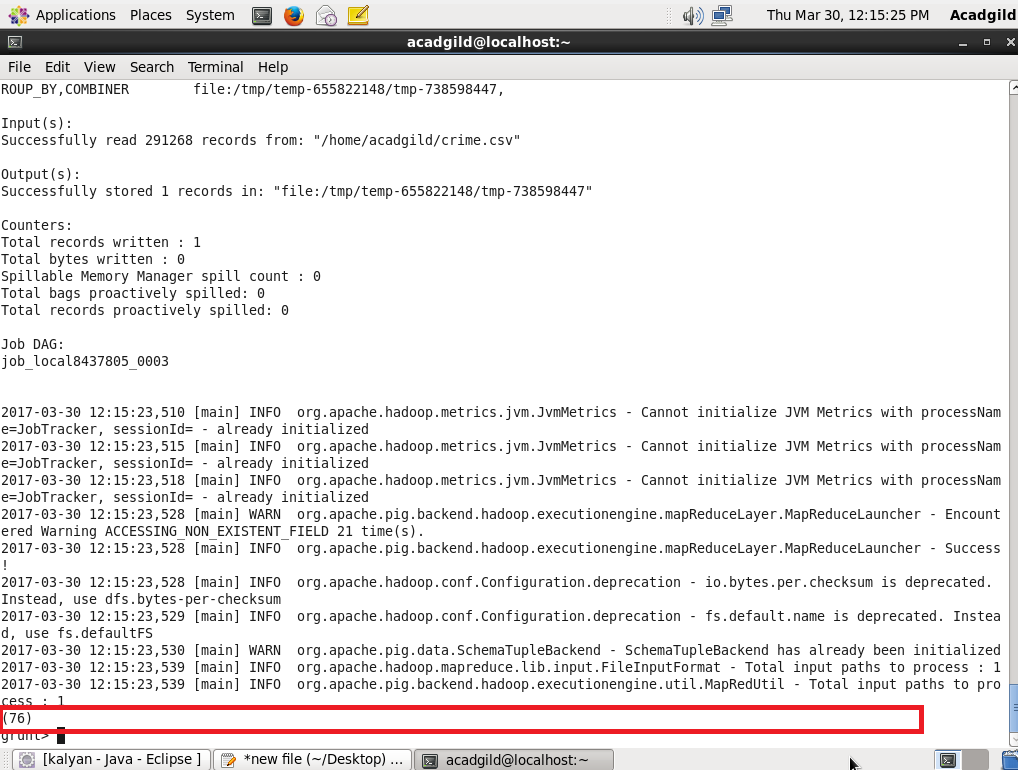
*3) Grouping by FBIcode (column 15) using GROUP BY command.*

*4) Generating the count using COUNT command*

**PIG COMMANDS:**

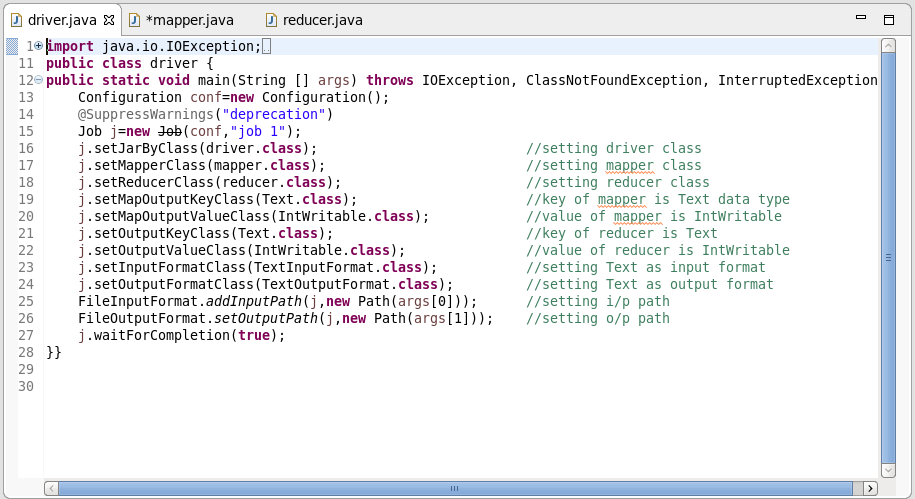
****

**PIG OUTPUT:**

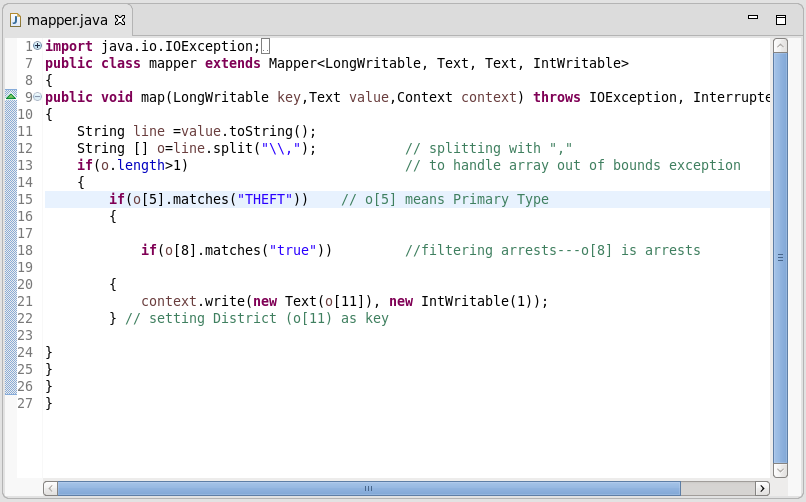
****

**3) WRITE A MAPREDUCE AND PIG PROGRAM TO CALCULATE THE NUMBER OF ARRESTS IN THEFT DISTRICT WISE.**

**DRIVER CLASS:**

****

**MAPPER CLASS:**

****

**MAPPER LOGIC:**

*LINE 11: converting the value to string*

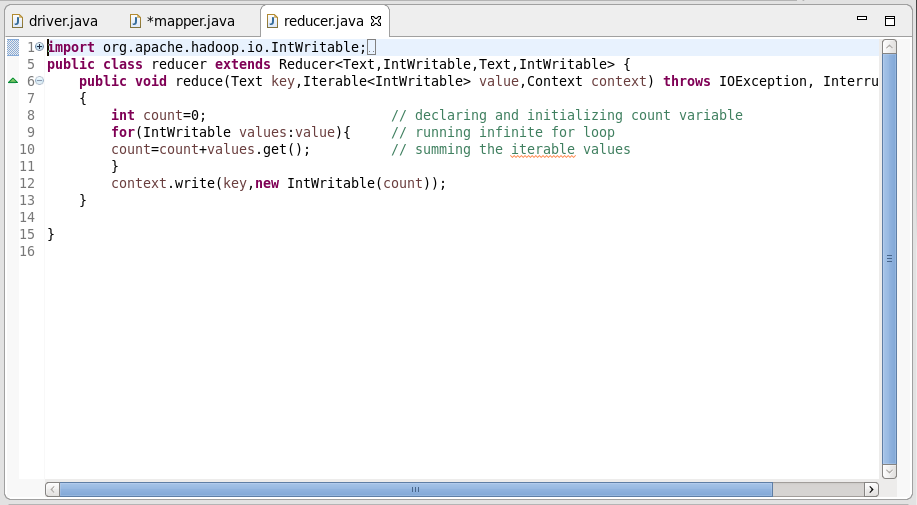
*LINE 12: splitting the string using comma and storing it in array o*

*LINE 15 : checking whether the primary type is THEFT*

*LINE 18: checking whether arrest is done*

*LINE 21 : Setting District as key and 1 as value*

**REDUCER CLASS:**

****

**REDUCER LOGIC:**

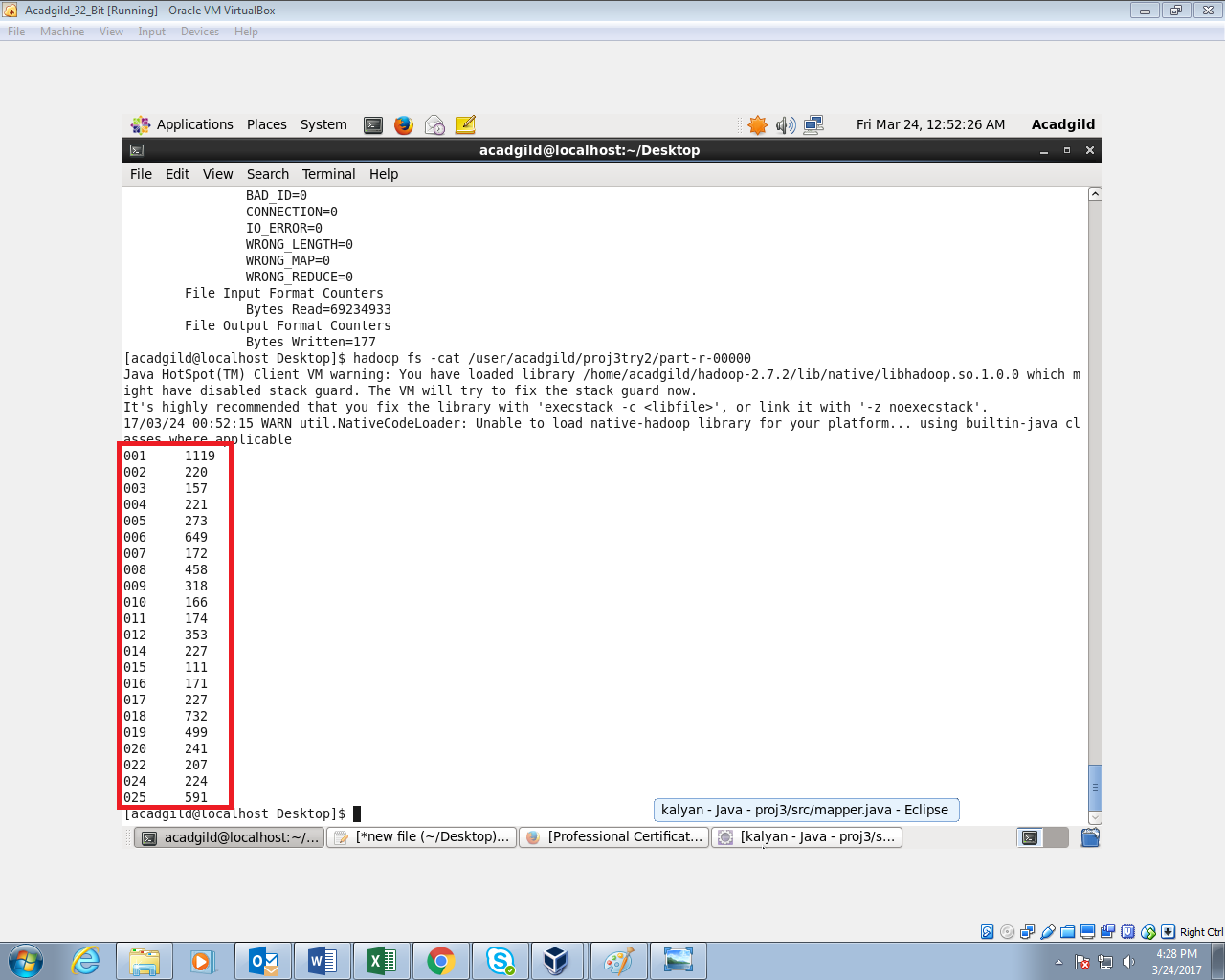
*LINE 8: Initializing the count as 0*

*LINE 9 : Running an infinite for loop*

*LINE 10: count will add the iterative values*

*LINE 12: Finally the same key is used as the key and the count is set as value.*

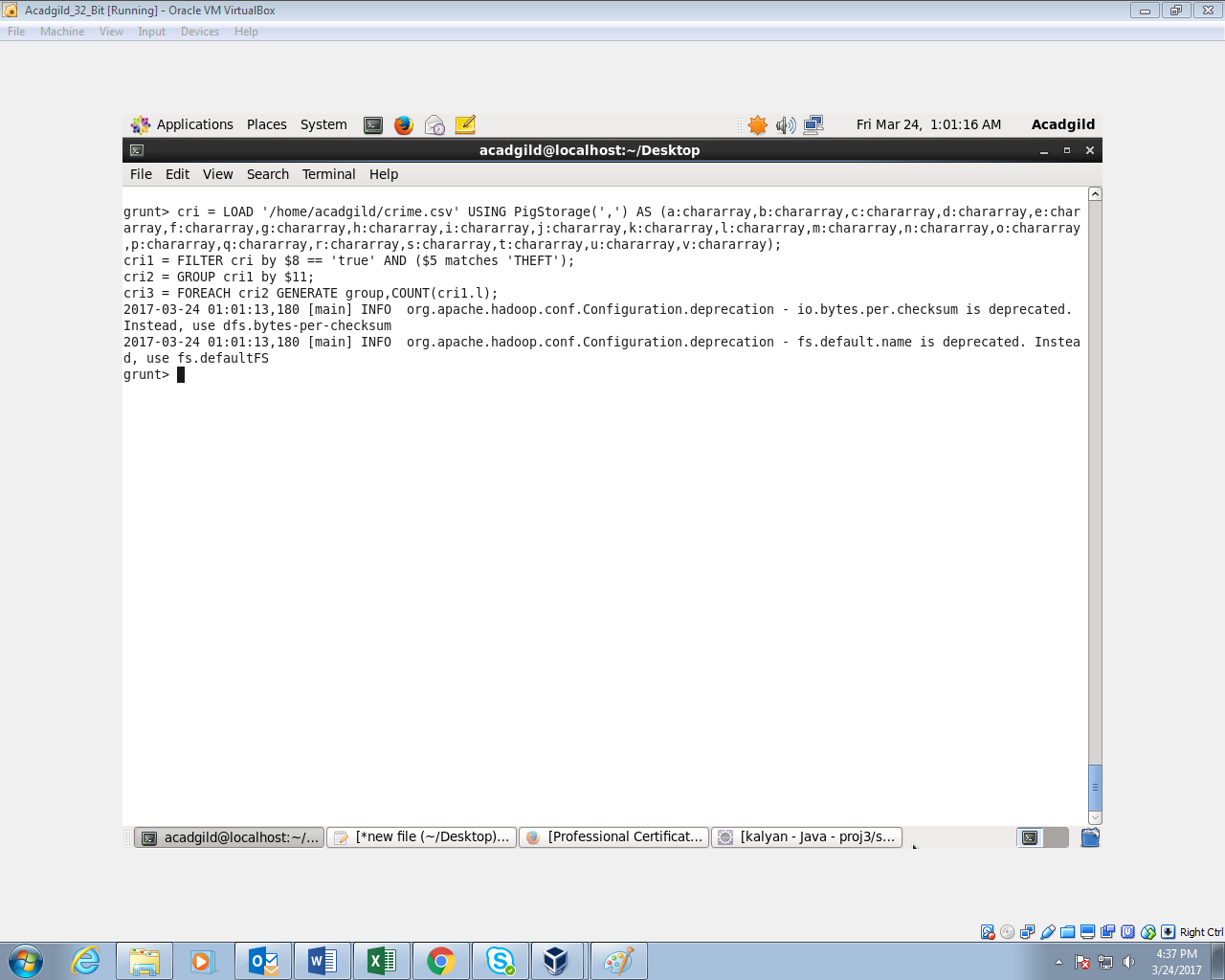
**OUTPUT:**

****

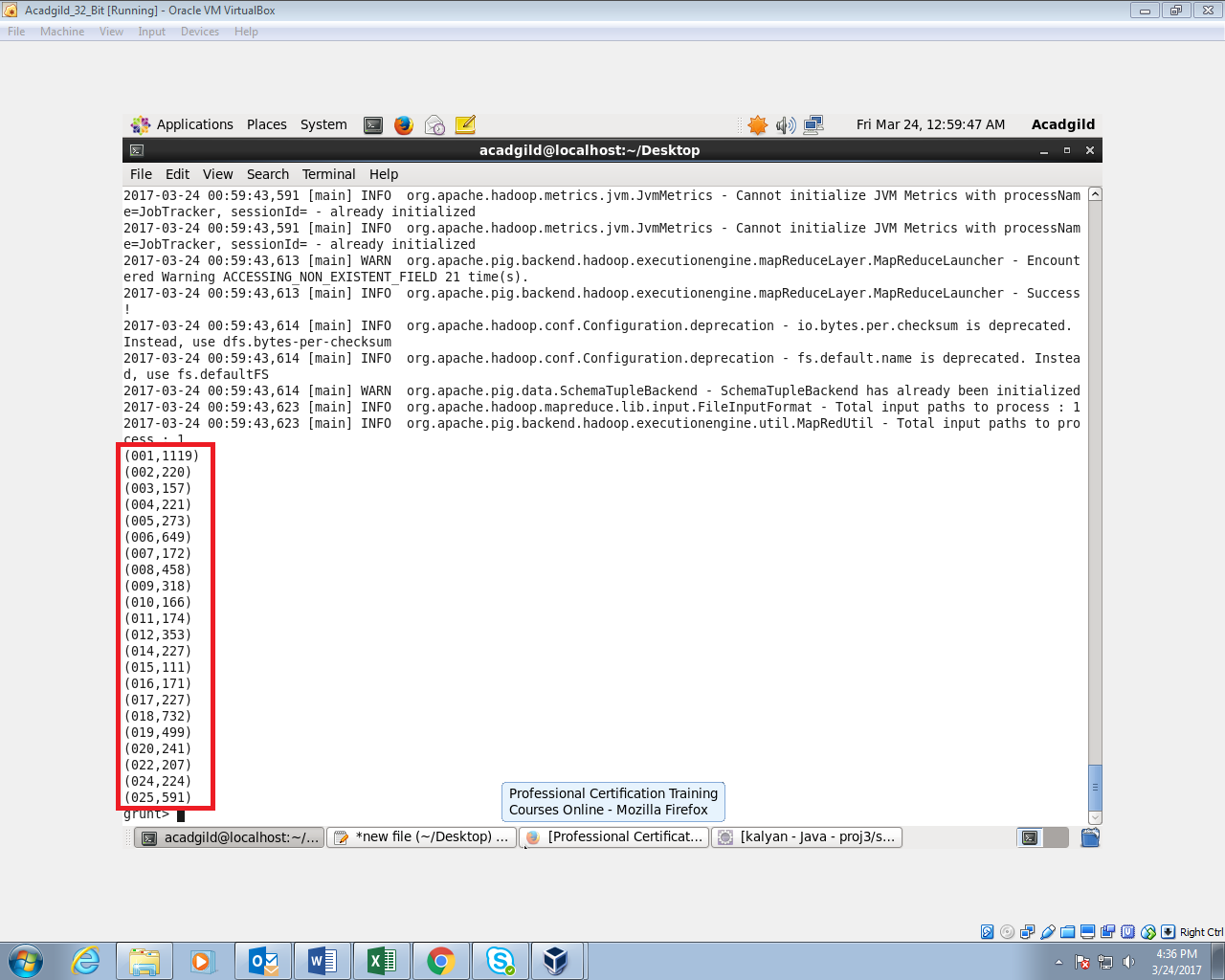
**USING PIG:**

1. *Loading the Dataset into the Grunt shell using LOAD command.*
2. *Filter the Dataset by checking whether column 9 is true and column 6 is theft.*
3. *Grouping by District ID using GROUP BY command.*
4. *Generating the count using COUNT command.*

**PIG COMMAND:**

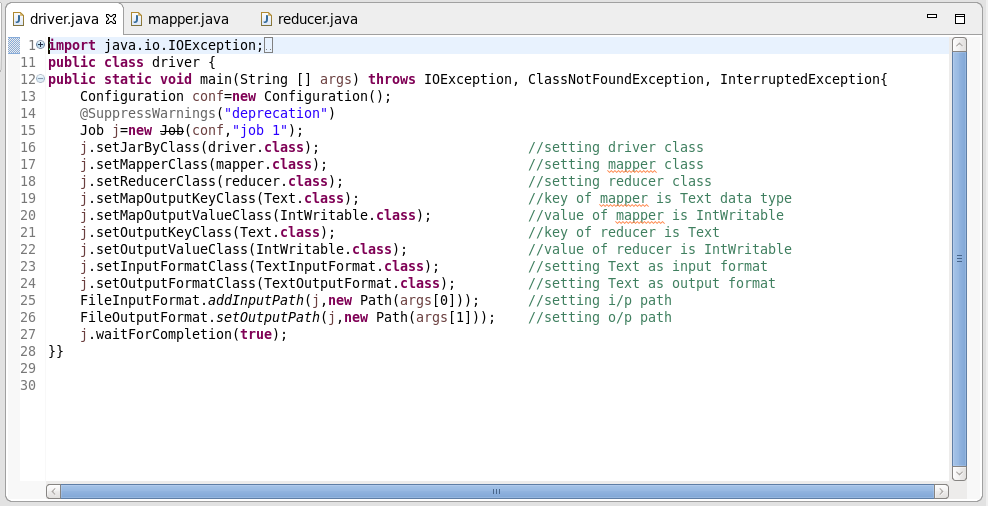
****

**PIG OUTPUT:**

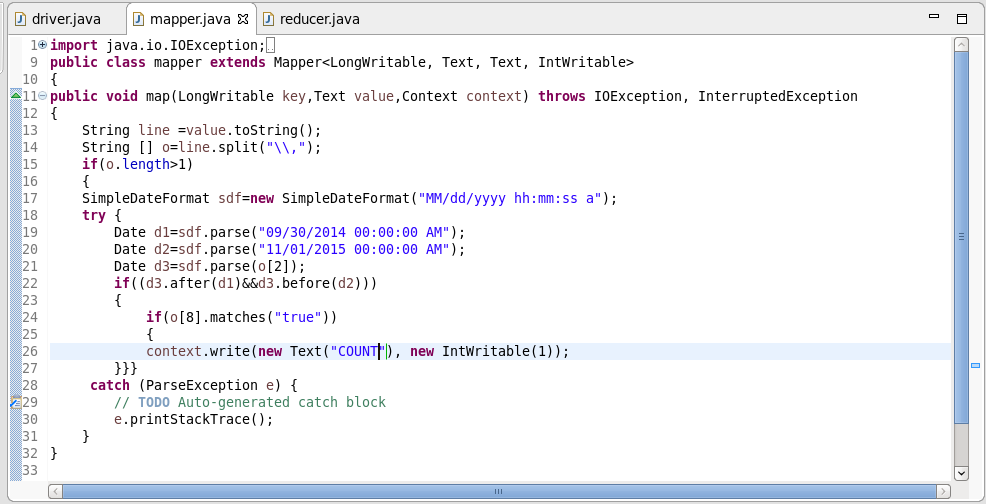
****

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

**DRIVER CLASS:**

****

**MAPPER CLASS:**

****

**MAPPER LOGIC:**

*LINE 13: converting the value to string and storing it in string line*

*LINE 14: splitting the dataset with comma and storing it in array o*

*LINE 17 : using simple date format we are specifying the date format*

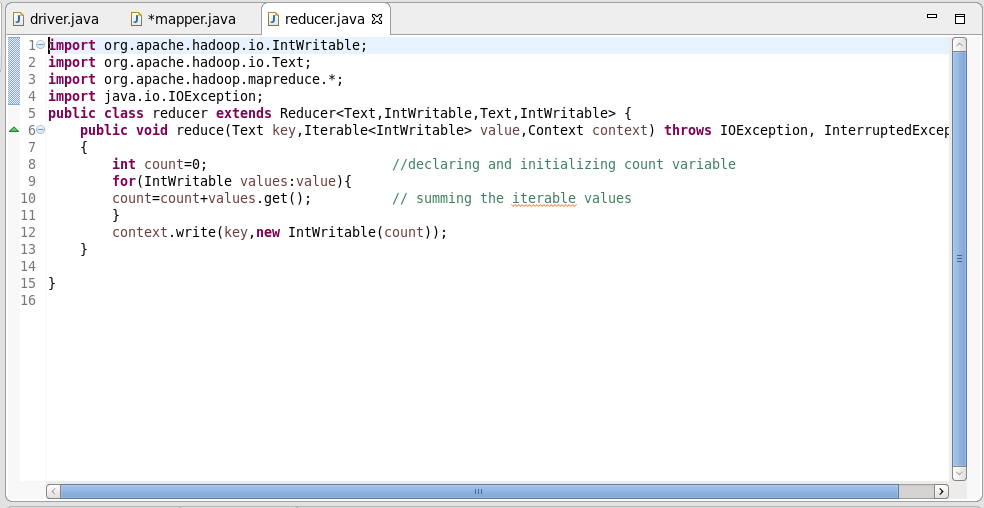
*LINE 19 AND 20: using parse we are converting the string to date d1 and d2.*

*LINE 21 :**parsing o[2] by converting it to date d3*

*LINE 22 : checking the condition (from 1st October 2014 to 31st October 2015 )*

*LINE 24: Checking if arrest is TRUE (o[8] is true)*

**REDUCER CLASS:**

****

**REDUCER LOGIC:**

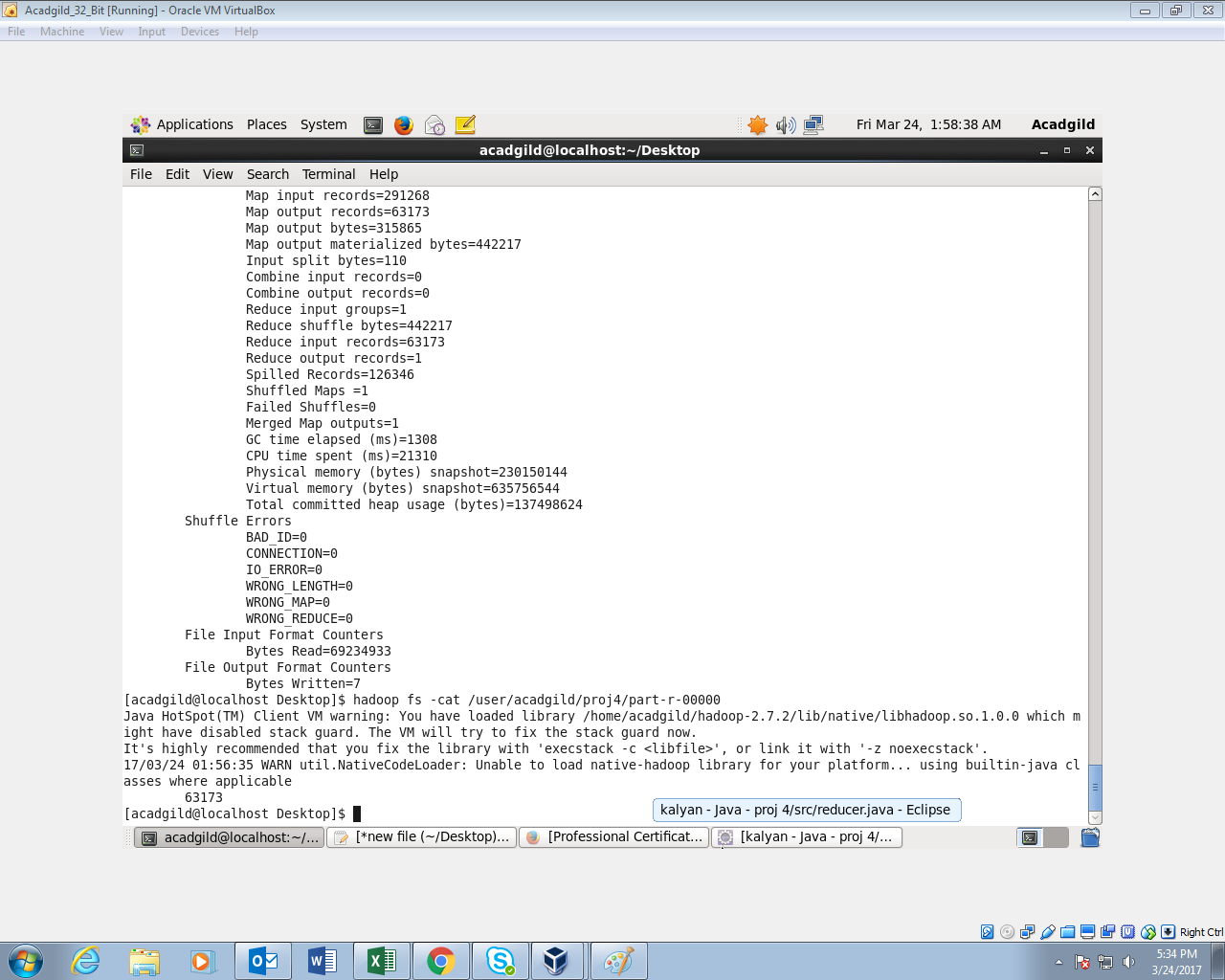
*LINE 8: Initializing the count as 0*

*LINE 9 : Running an infinite for loop*

*LINE 10: count will add the iterative values*

*LINE 12: Finally the same key is used as the key and the count is set as value.*

**OUTPUT:**

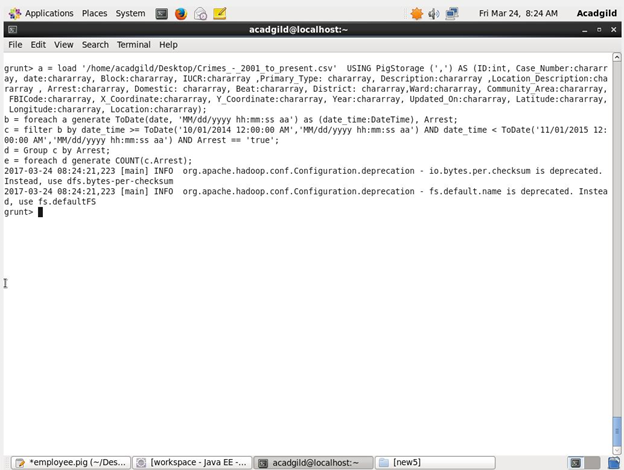
****

**USING PIG:**

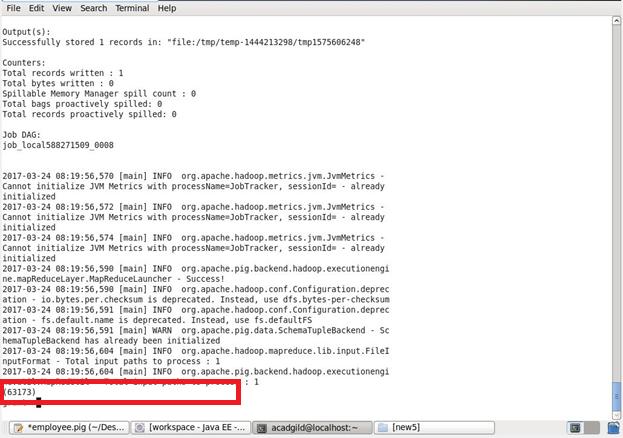
**LOGIC:**

1. *Loading the Dataset into the Grunt shell using LOAD command.*
2. *Converting the date column to date format using ToDate command*
3. *Filtering the dataset by specifying the period (from 1st of October 2014 to 31st of October 2015) and checking whether arrest is done*
4. *Counting the filtered dataset by count command*

**PIG COMMAND:**

****

**PIG OUTPUT:**

****