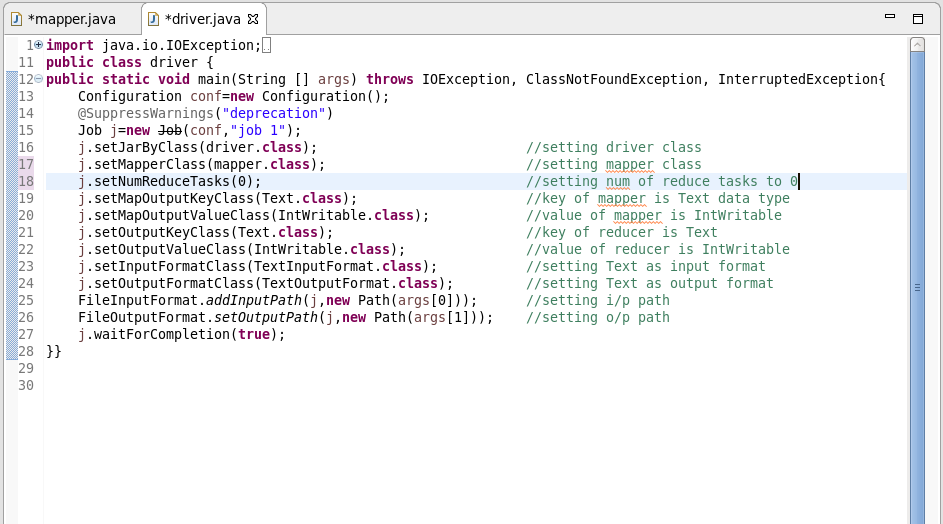
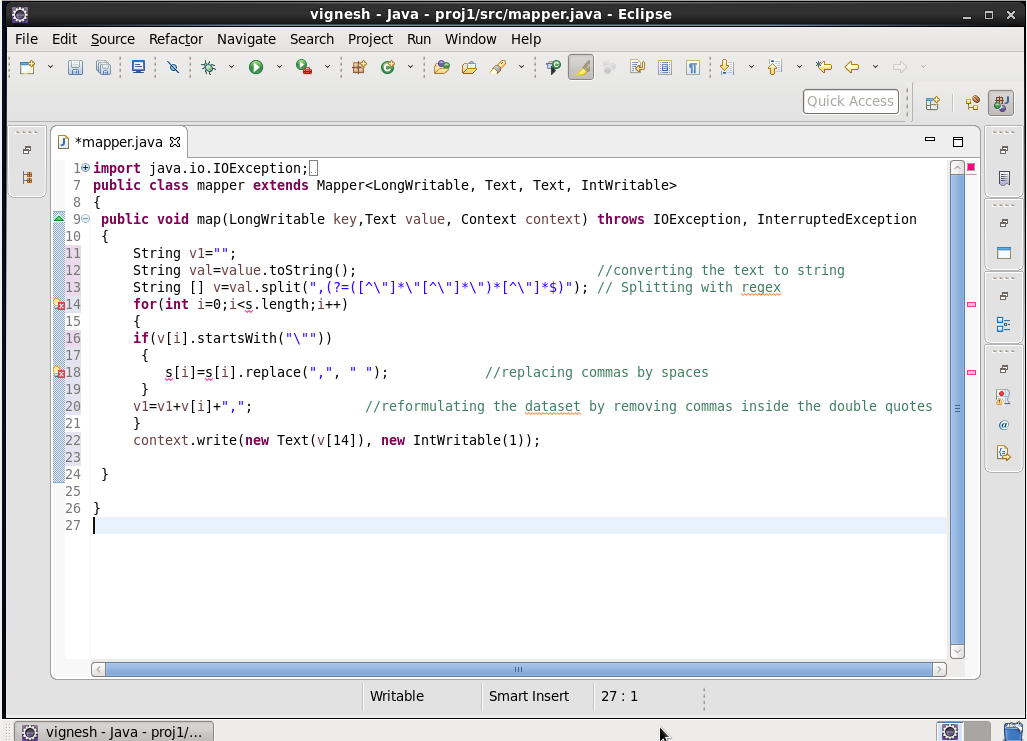
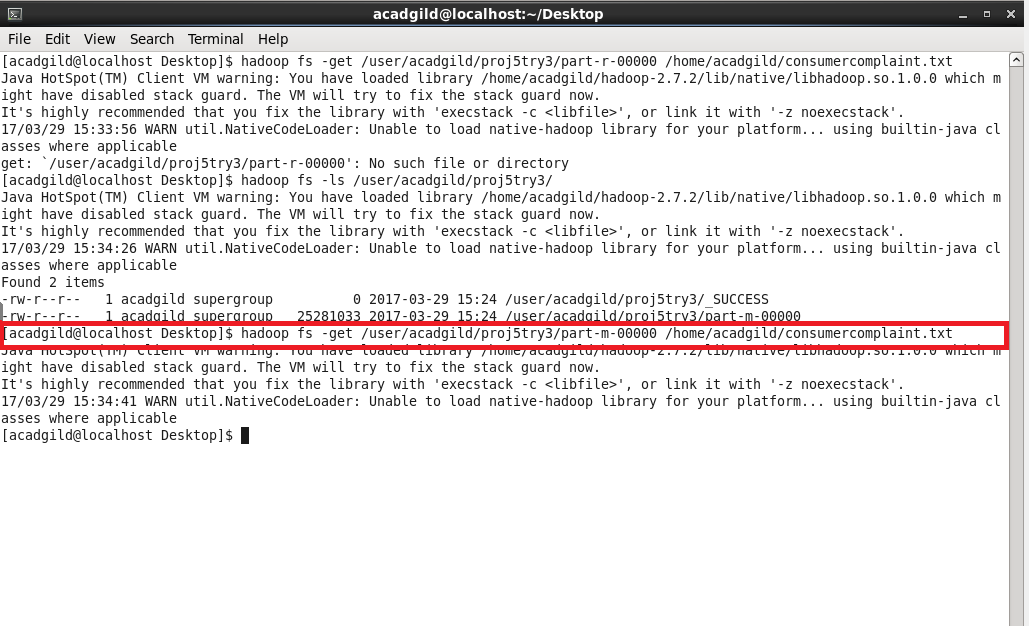
1. **MAP REDUCE PROGRAM TO REPLACE ALL THE COMMAS INSIDE THE DOUBLE QUOTES WITH SPACE:**

**DRIVER CLASS:**

****

**MAPPER CLASS:**

****

****

* The command : **“ hadoop fs –ls /user/acadgild/proj5try/ “** is used to list the files inside the proj5try.
* By using get command, we can copy the file from hdfs to local.
* Part-m-00000 is copied to the local in the name of consumercomplaint.txt with all the commas inside the double quotes replaced with space.
* The processed file is used as the Dataset for executing pig commands.

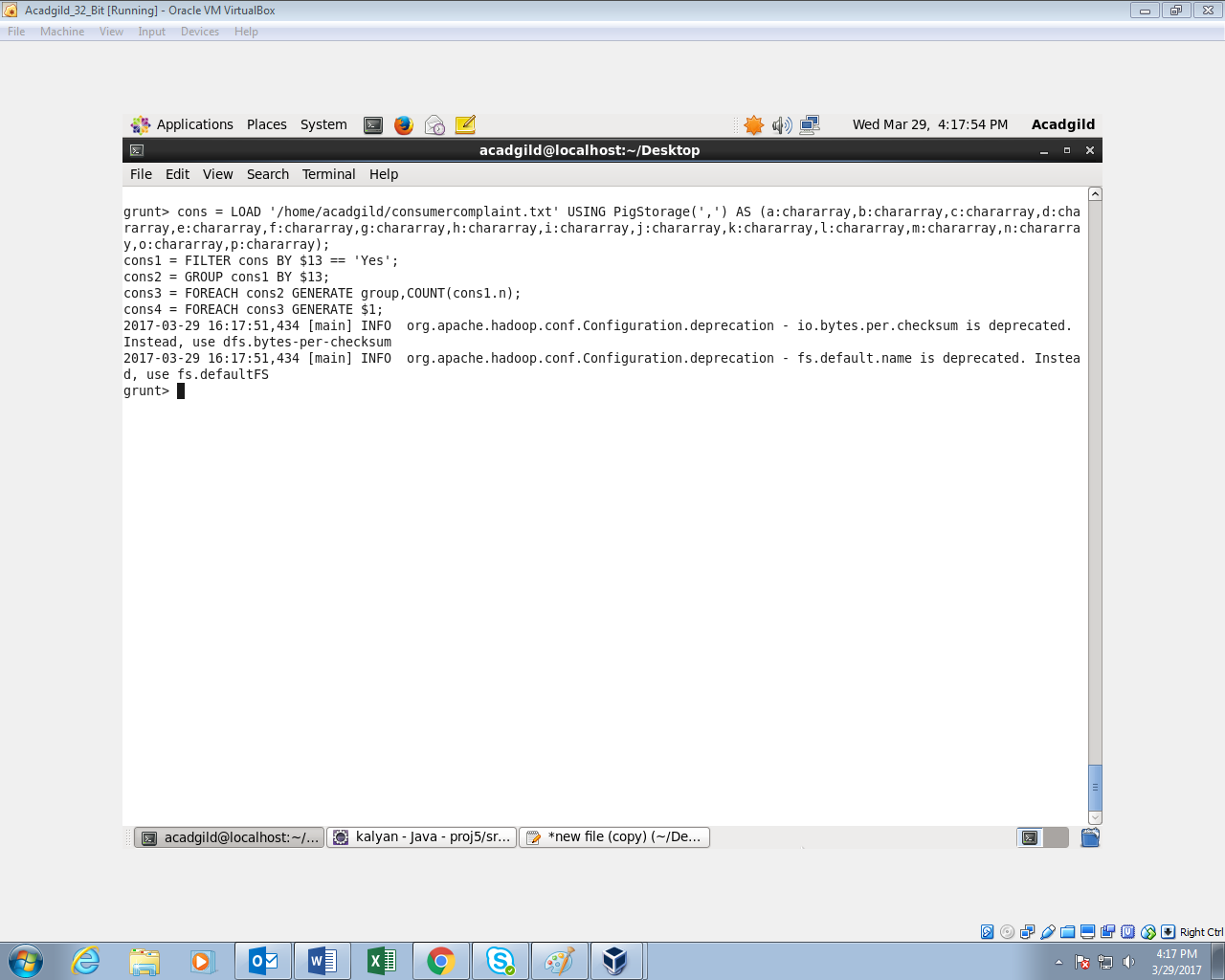
1. **1. WRITE A PIG SCRIPT TO FIND NO OF COMPLAINTS WHICH GOT TIMELY RESPONSE.**

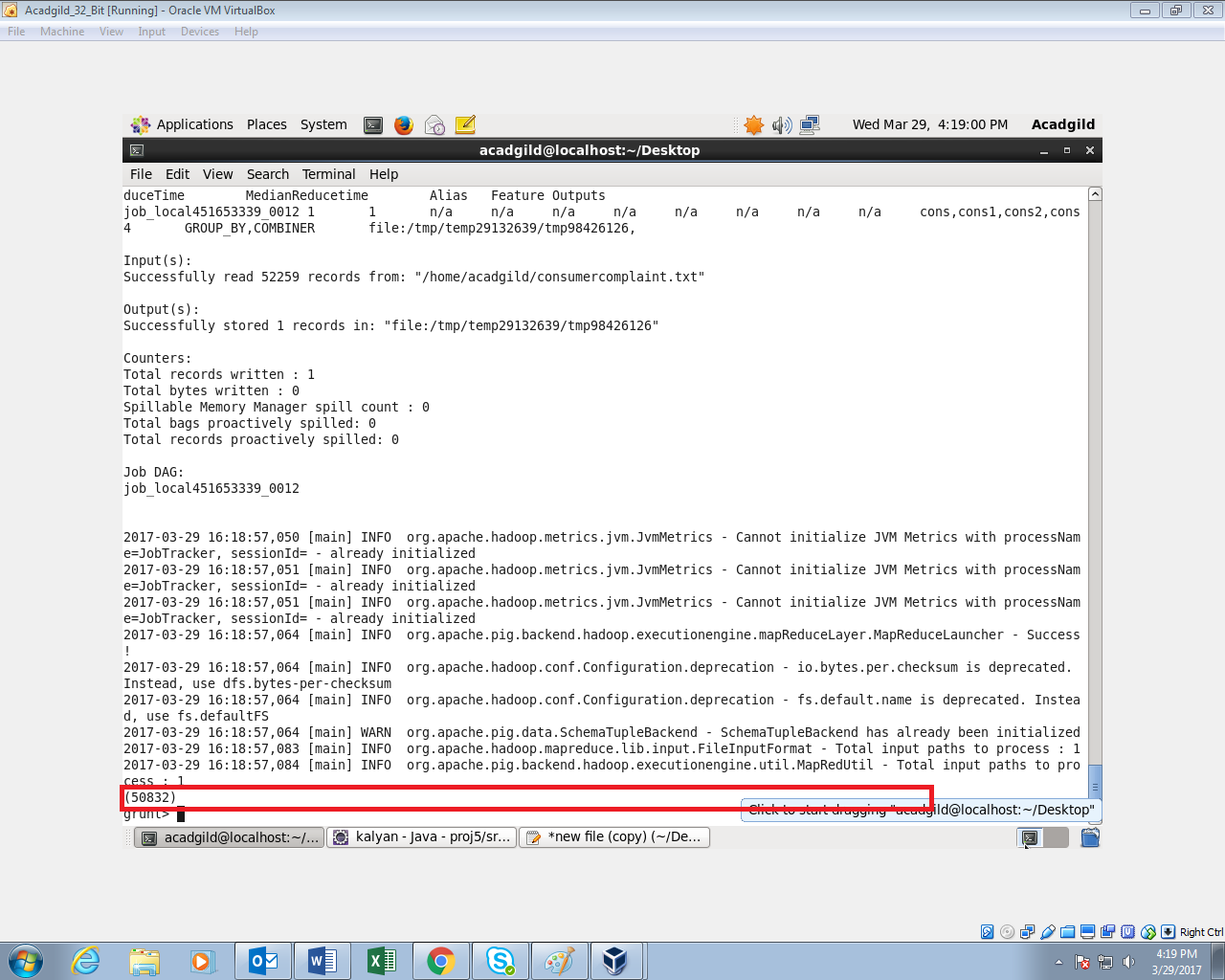
***LOGIC:***

1. Loading the Dataset into the grunt shell using LOAD command.
2. Filtering the Dataset by checking the column 14 is yes.
3. Grouping based on column 14 using Group By command.
4. Counting the number of complaints using COUNT command.

At this stage o/p will be like (Yes, 50832).

1. Generating only the count using GENERATE command by eliminating the yes.

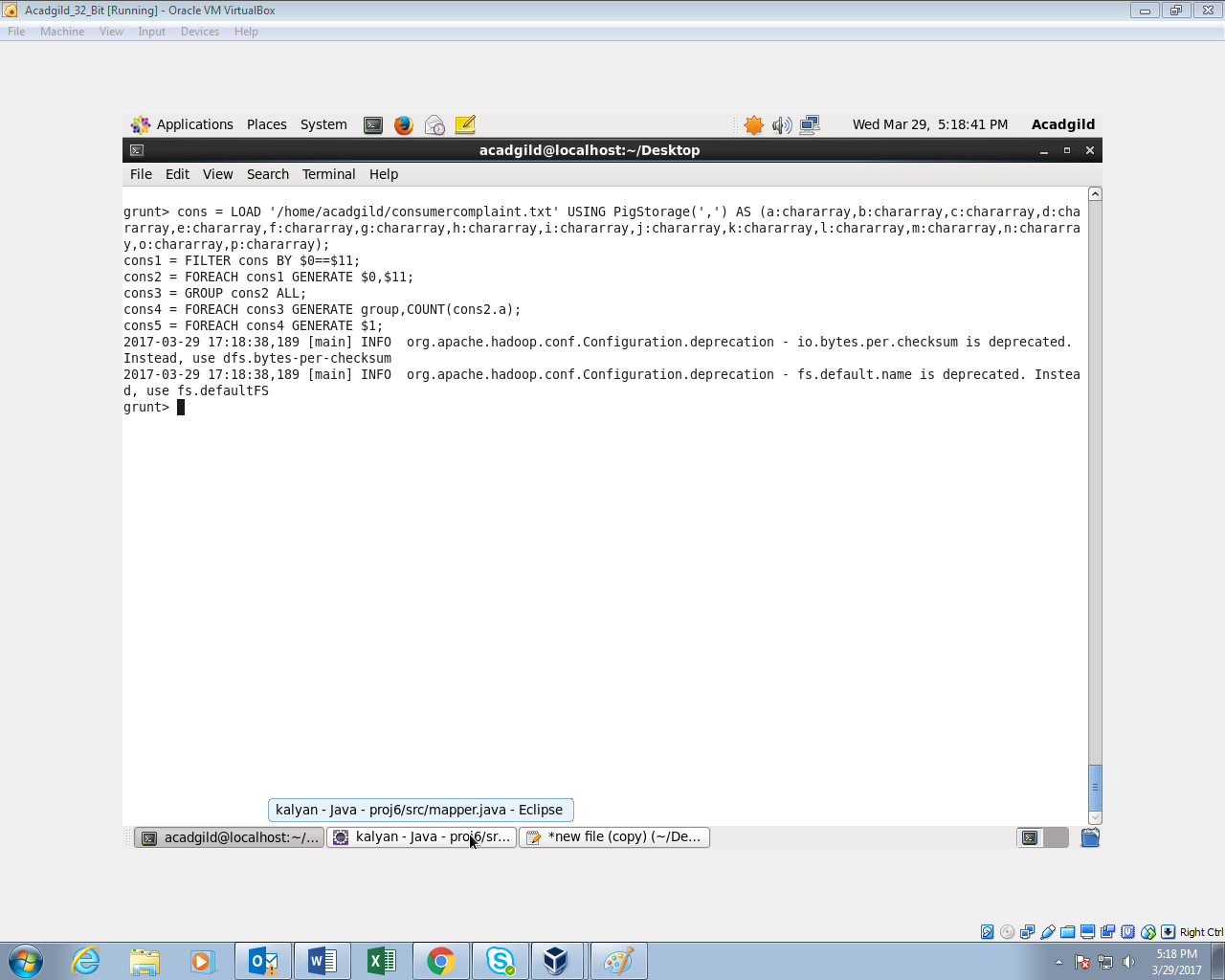


**OUTPUT:**

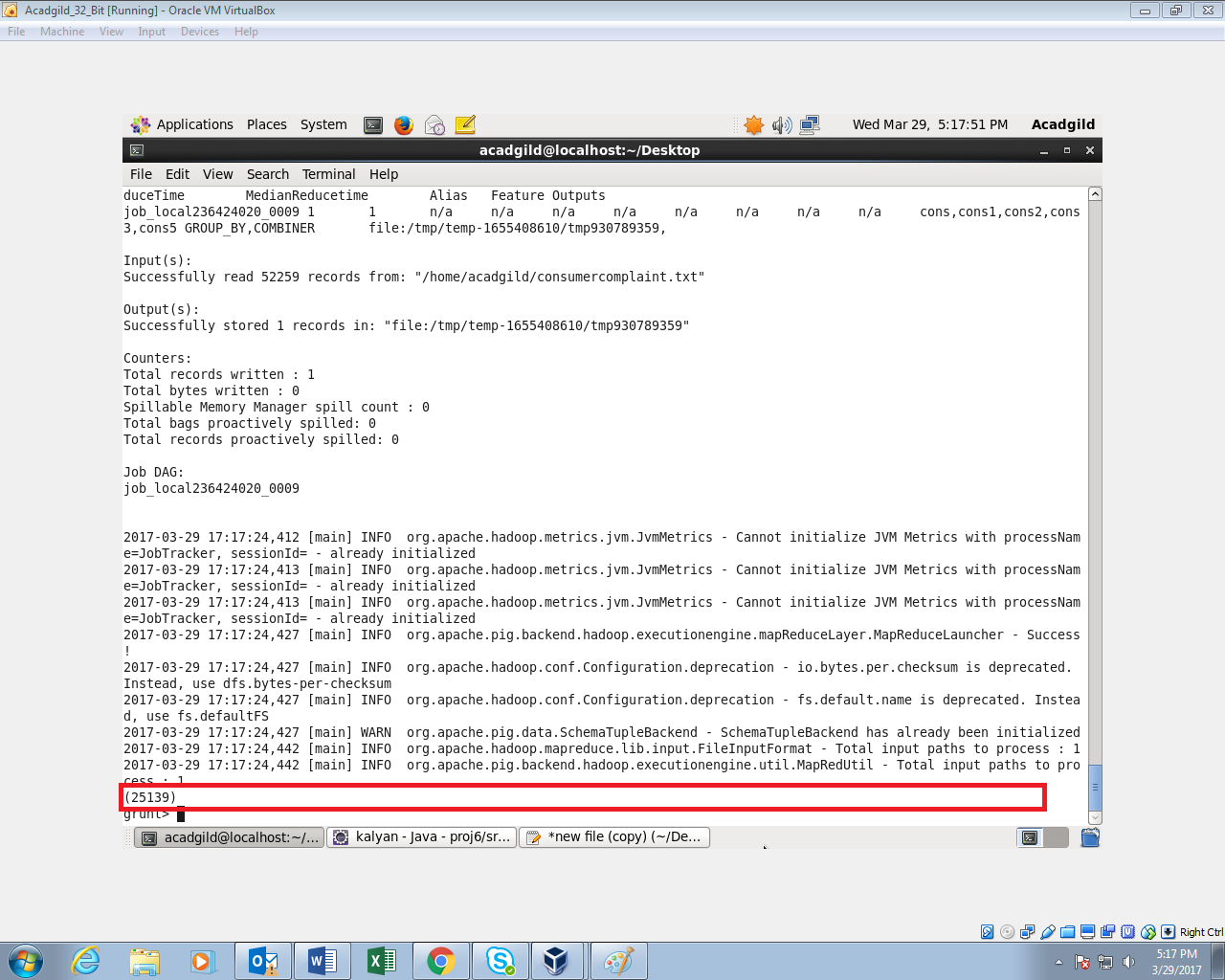
1. **WRITE A PIG SCRIPT TO FIND NO OF COMPLAINTS WHERE CONSUMER FORUM FORWARDED THE COMPLAINT SAME DAY THEY RECEIVED TO RESPECTIVE COMPANY**

**LOGIC:**

* Loading the Dataset into the grunt shell using LOAD command.
* Filtering the Dataset by checking whether column 1 and 12 are same.
* Generating only column 1 and 12 using GENERATE command.
* Grouping by all using GROUP ALL command.
* Generating the count using COUNT command

****

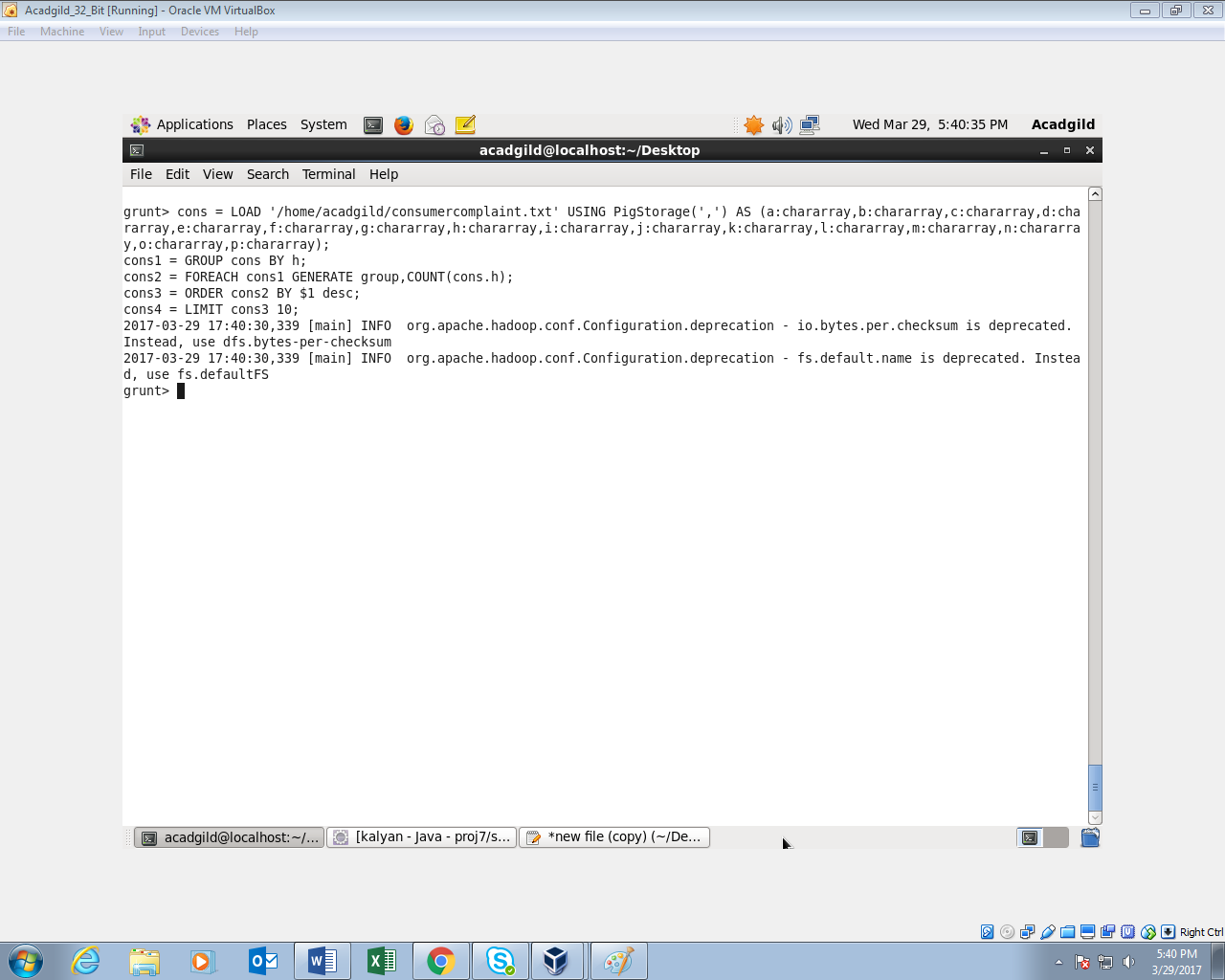
**OUTPUT:**

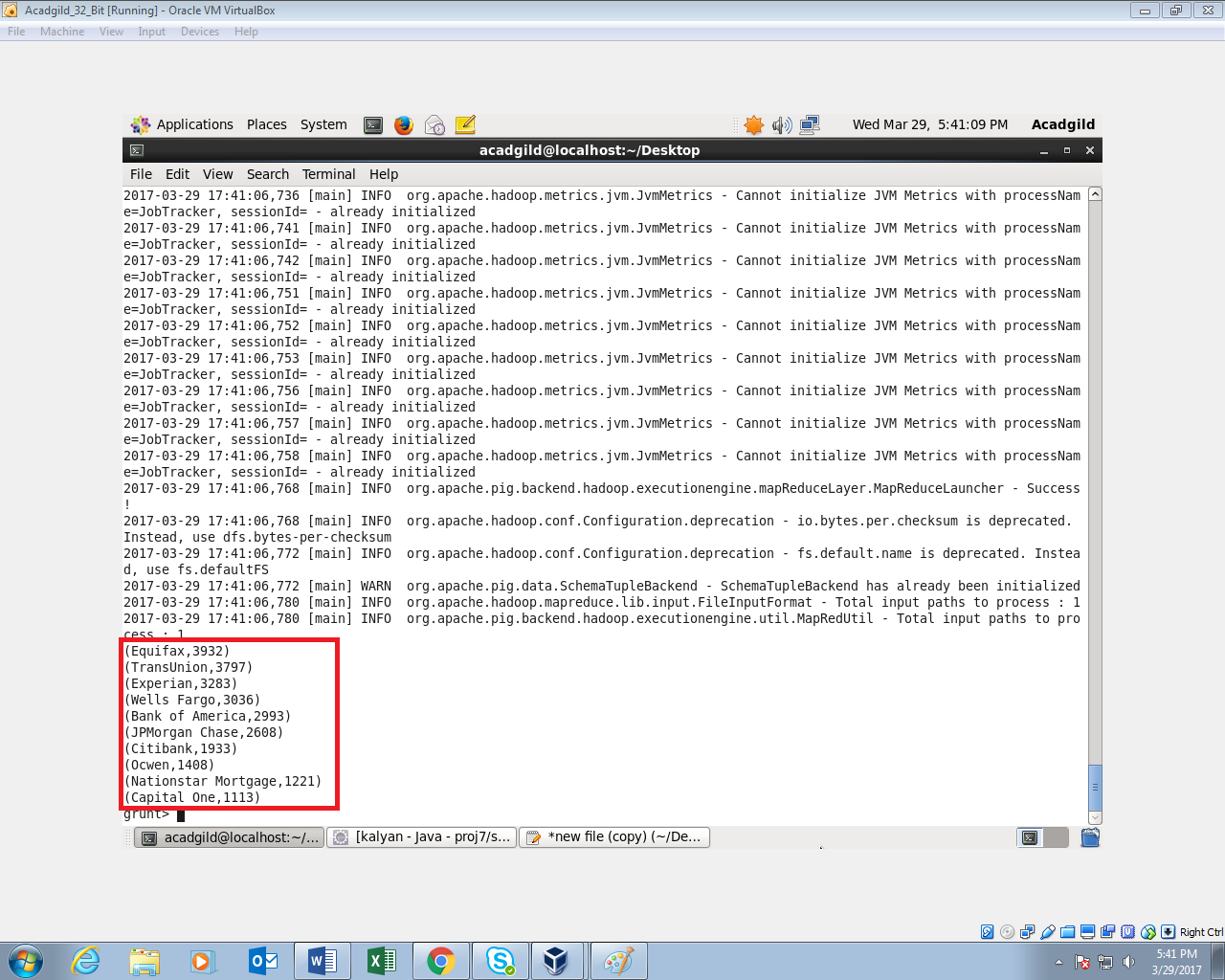
****

1. **WRITE A PIG SCRIPT TO FIND LIST OF COMPANIES TOPPING IN COMPLAINT CHART (COMPANIES WITH MAXIMUM NUMBER OF COMPLAINTS)**

**LOGIC:**

* Loading the Dataset into the grunt shell using LOAD command.
* Grouping the Dataset by Company name column.
* Generating the count using COUNT command.
* Ordering the content based on the count in Descending order using ORDER BY command.
* Displaying only Top 10 companies using LIMIT command (I am displaying only 10 companies)

**

**OUTPUT:**

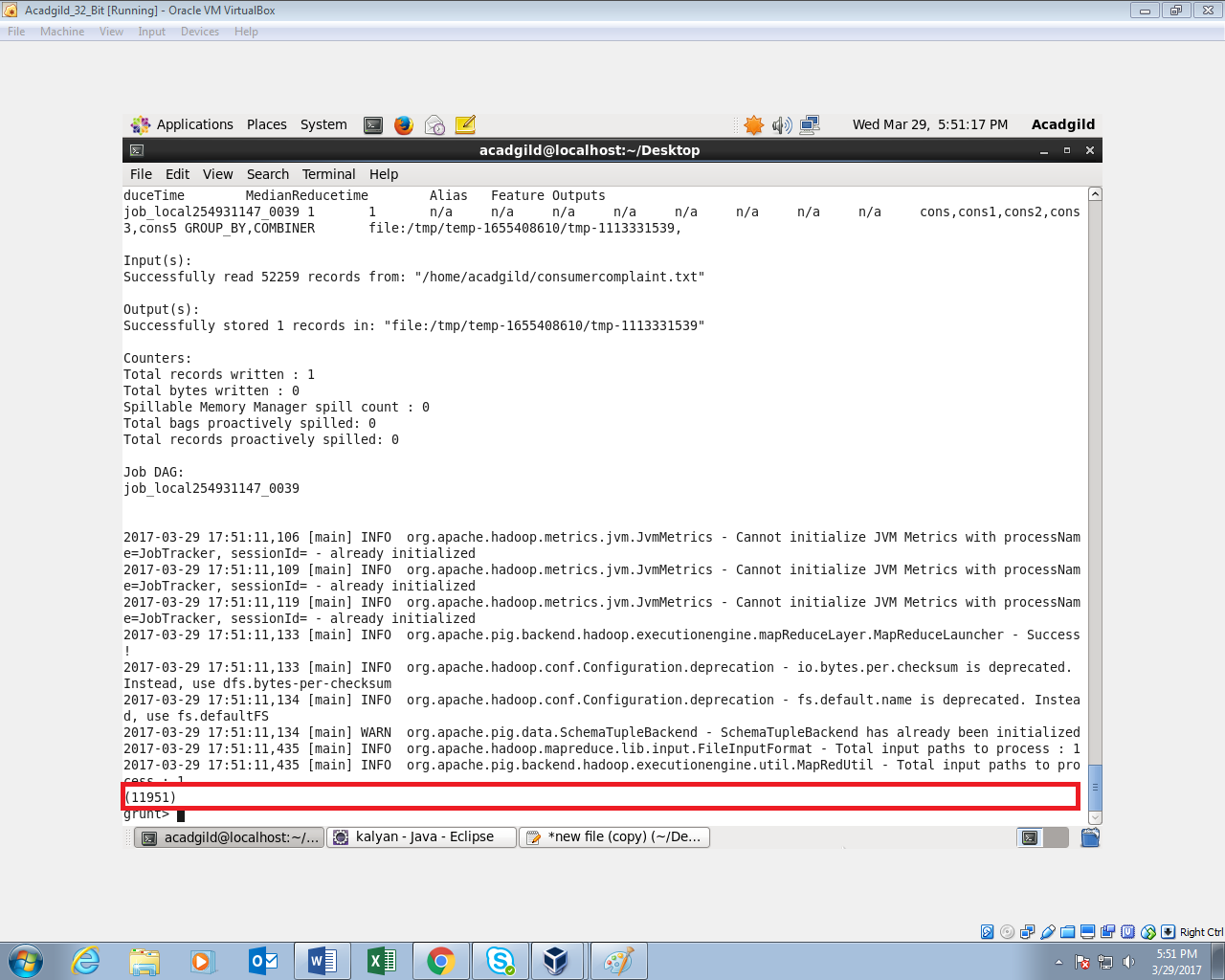
1. **WRITE A PIG SCRIPT TO FIND NO OF COMPLAINTS FILED WITH PRODUCT TYPE HAS "DEBT COLLECTION" FOR THE YEAR 2015**

**LOGIC:**

* + Filtering the Dataset by checking whether column 2 contains Debt collection using FILTER command.
  + Generating the year (yyyy) from the column
  + Generating only the Year using GENERATE and SUBSTRING command
  + Filtering the contents by year =”2015”;
  + Grouping by Debt collection using GROUP BY command.
  + Generating the count using COUNT command.
  + Displaying only the count as output.

****

**OUTPUT:**

****