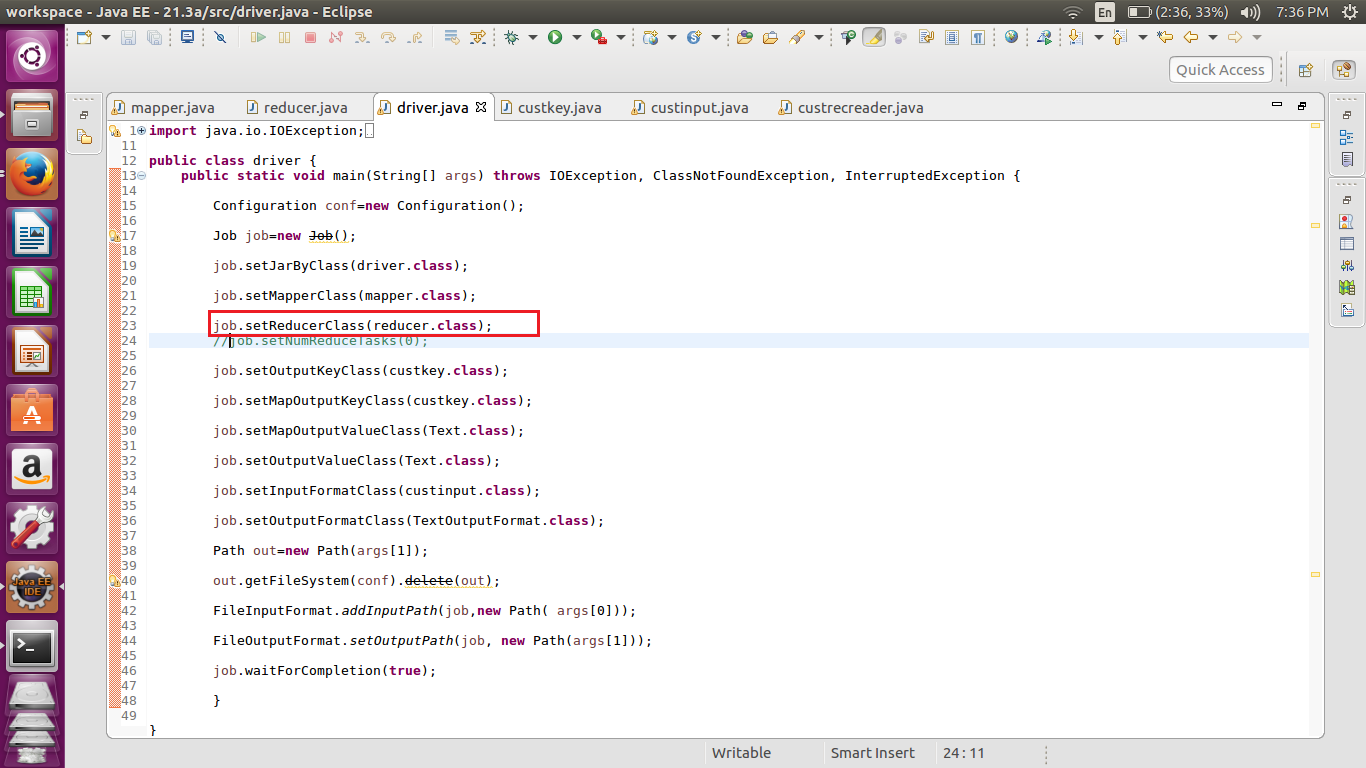
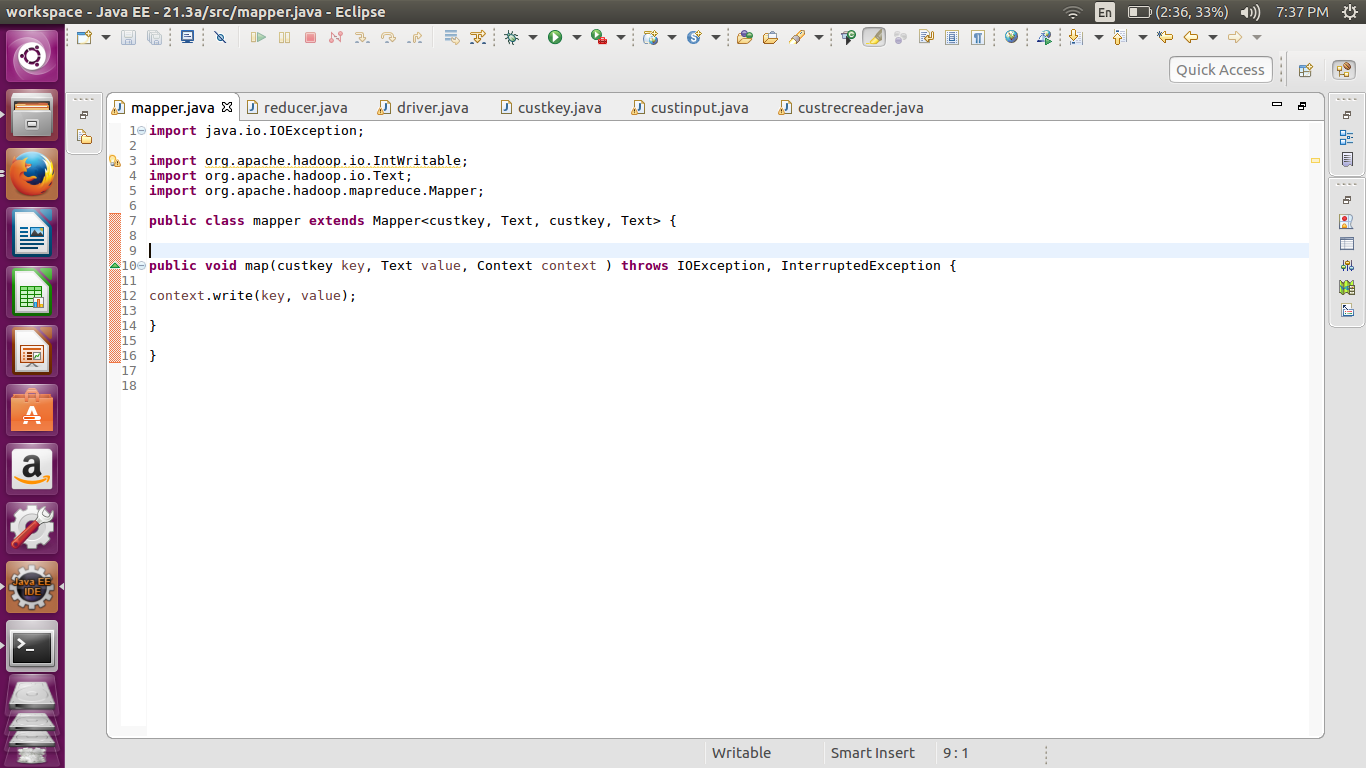
This dataset consists of two columns i,e., first\_name & last\_name &city use custom input format to read the key as the combination of first\_name and last\_name and arrange the output of the program should be the key sorted by last\_name.

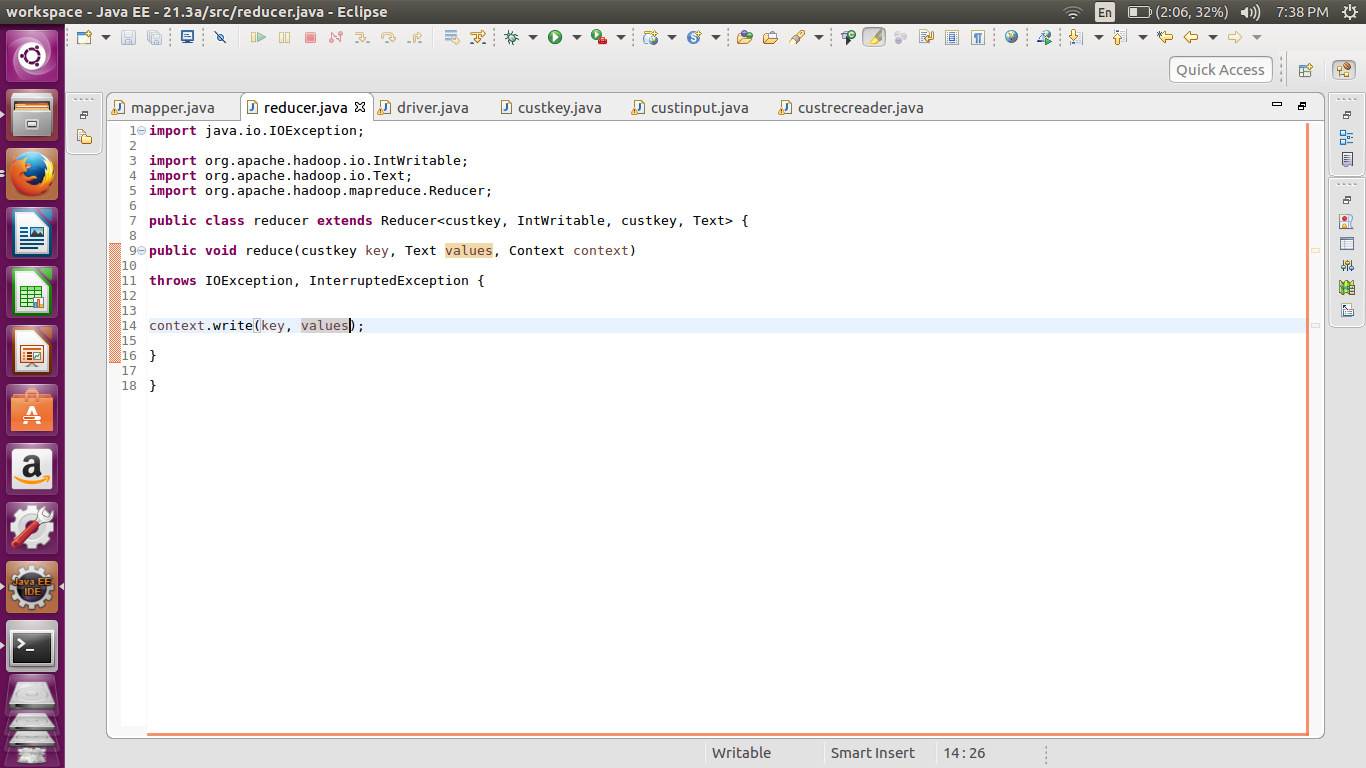
//Driver



//mapper class

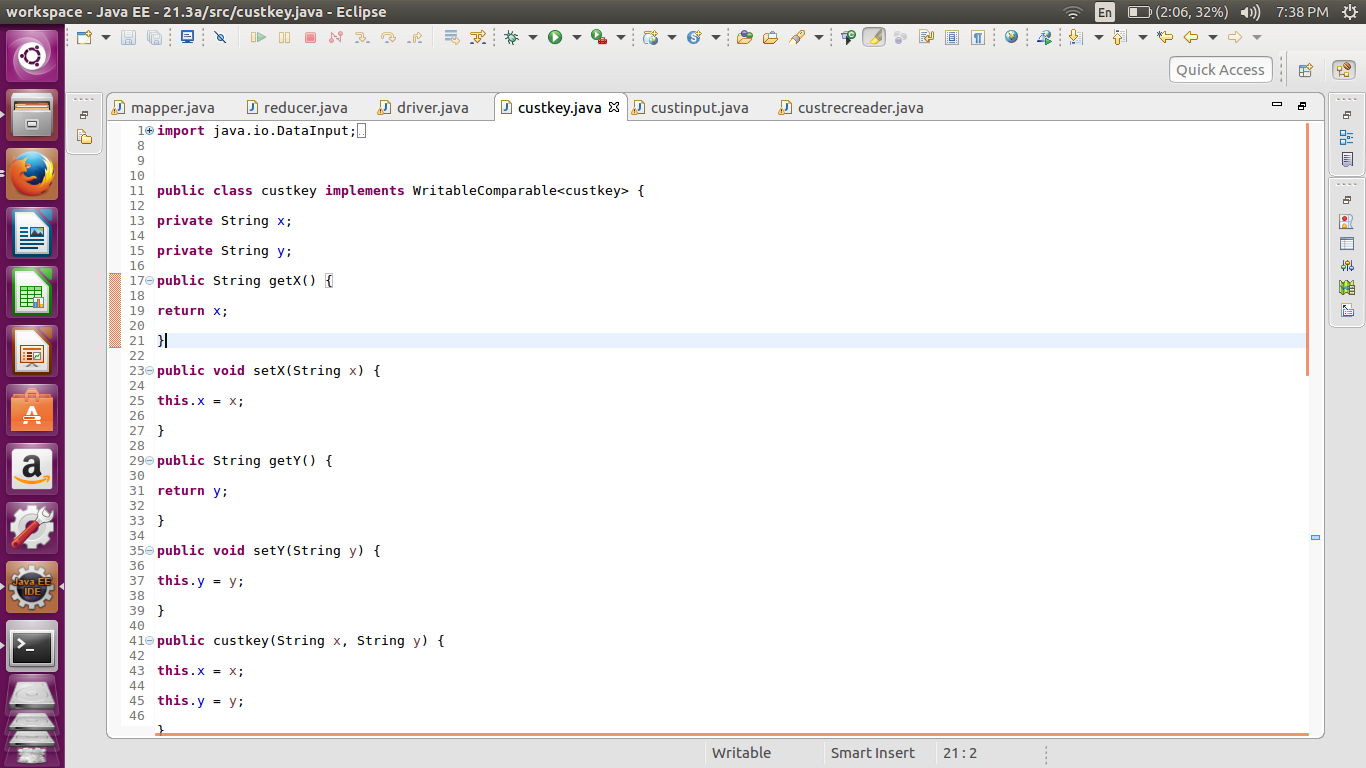


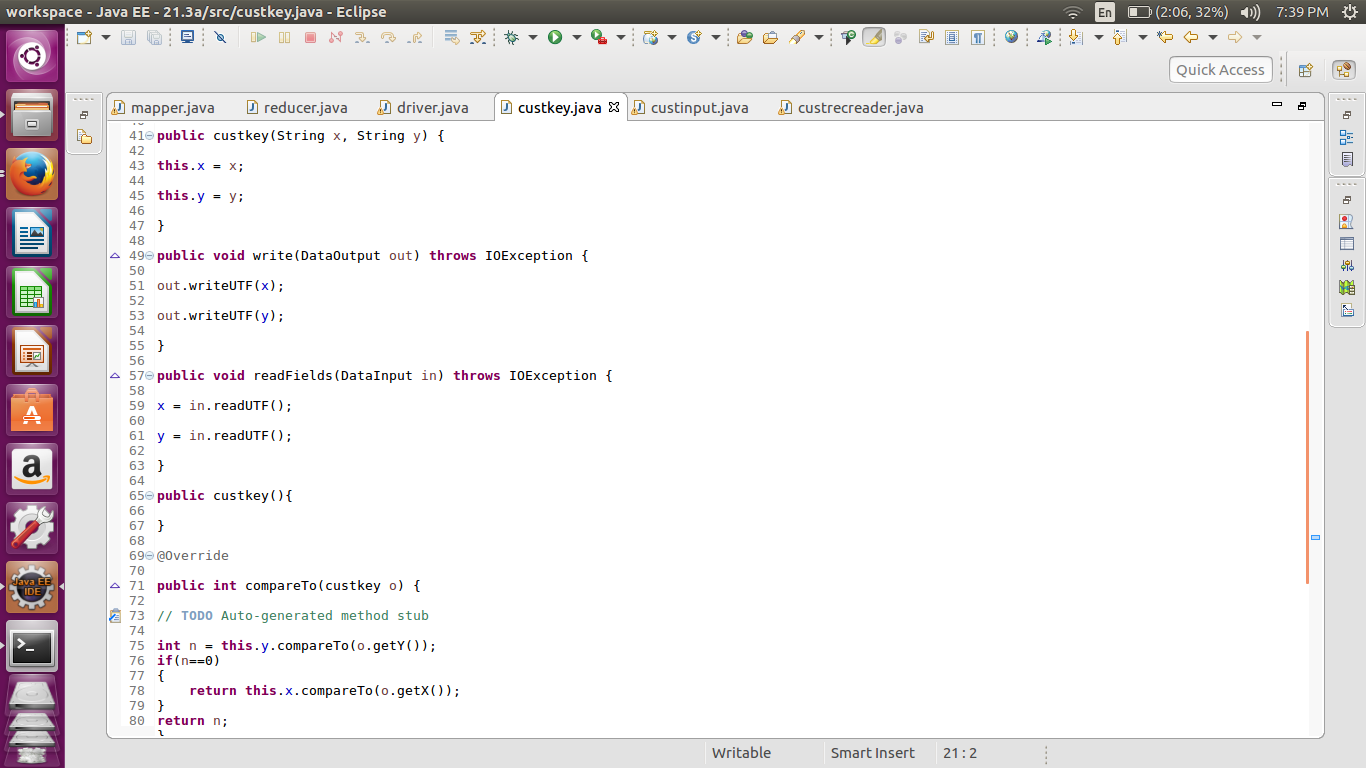
//Reducer

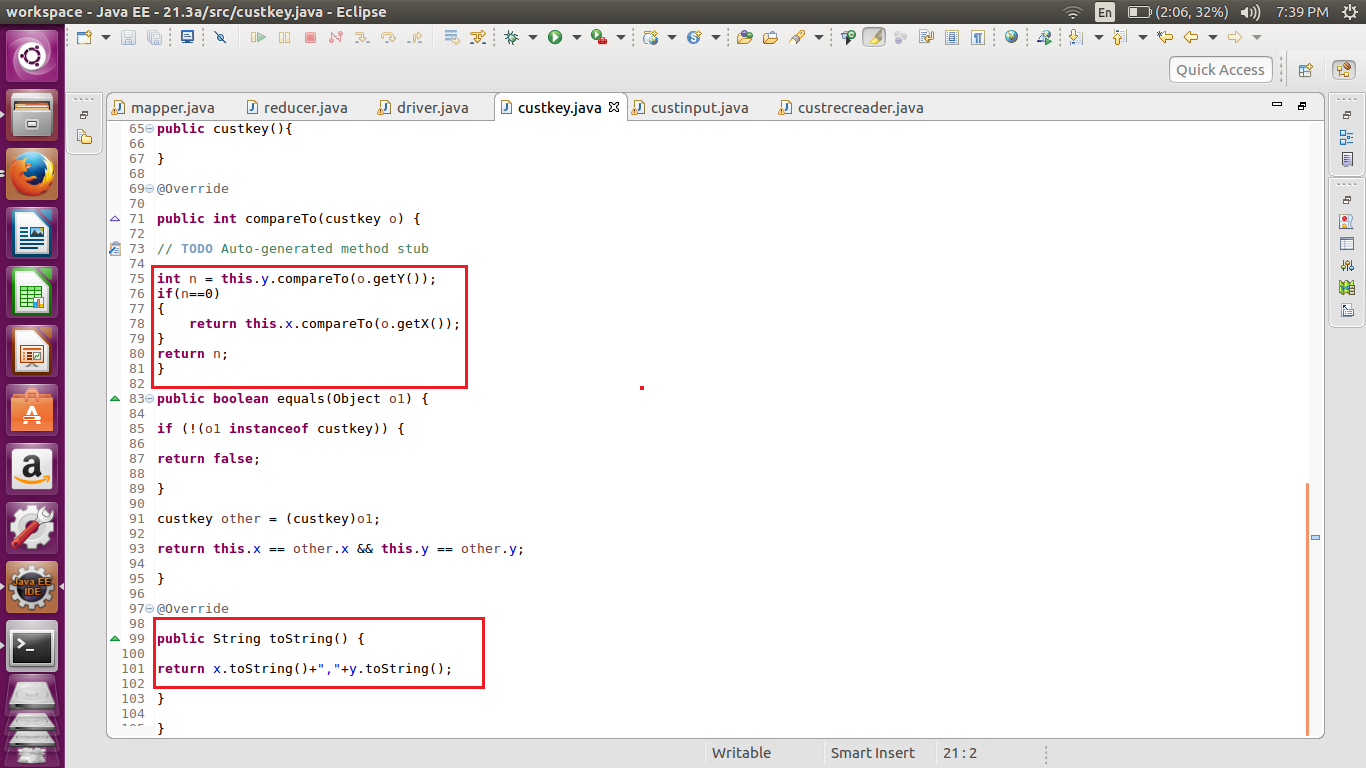


Here the mapper and the reducer are doing the default functions.

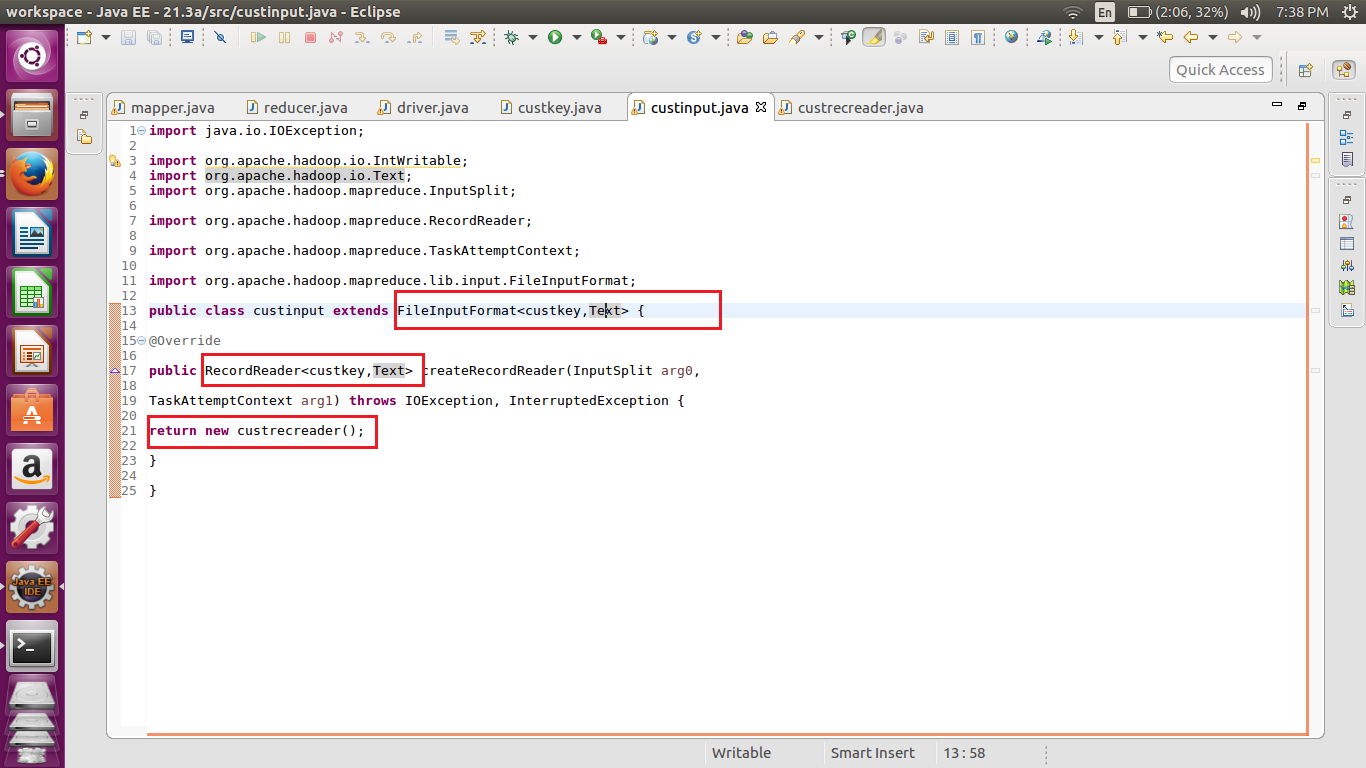
//Custom Key





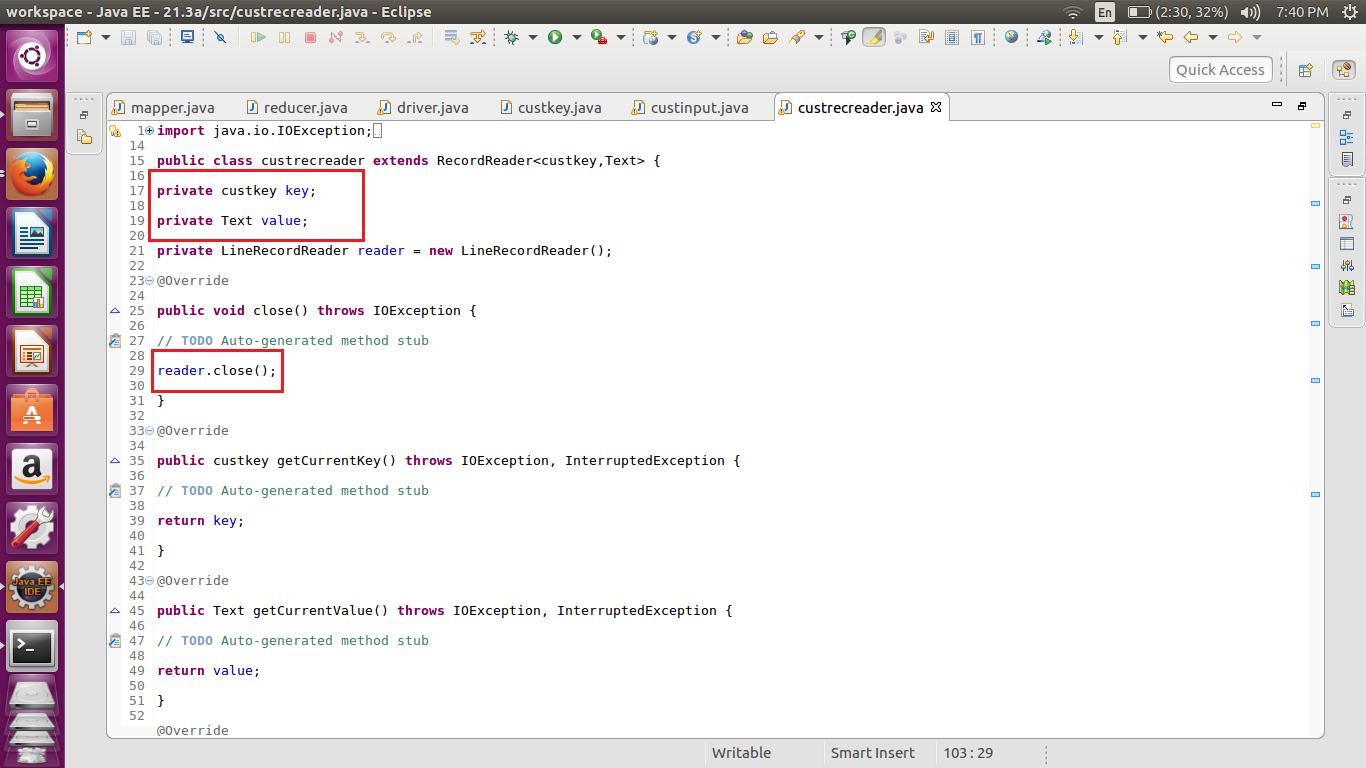


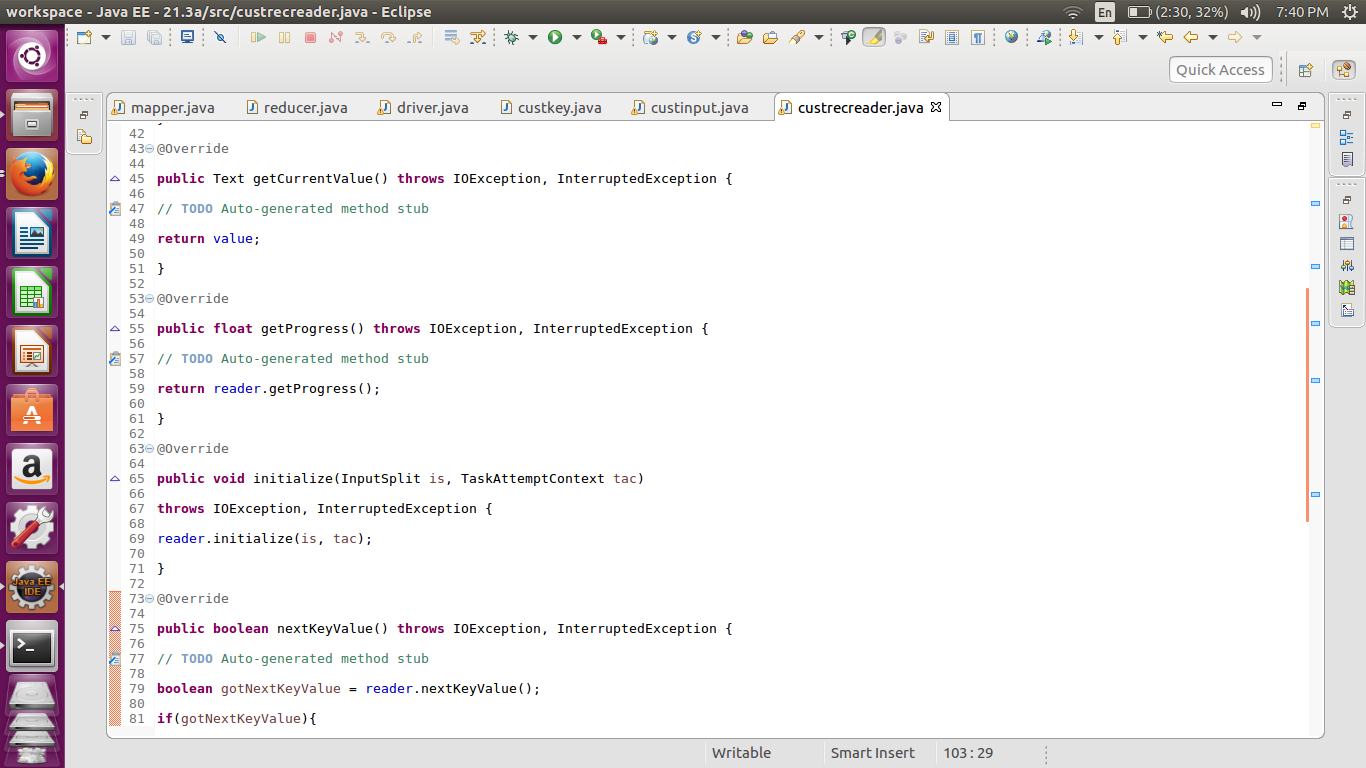
We have created the custom key for first name and the last name . the compare to method has been overridden so as to compare the last name first and then the first name(i.e if the last name is same). The toString method is also customized to get the output in the format “Firstname , Lastname”.

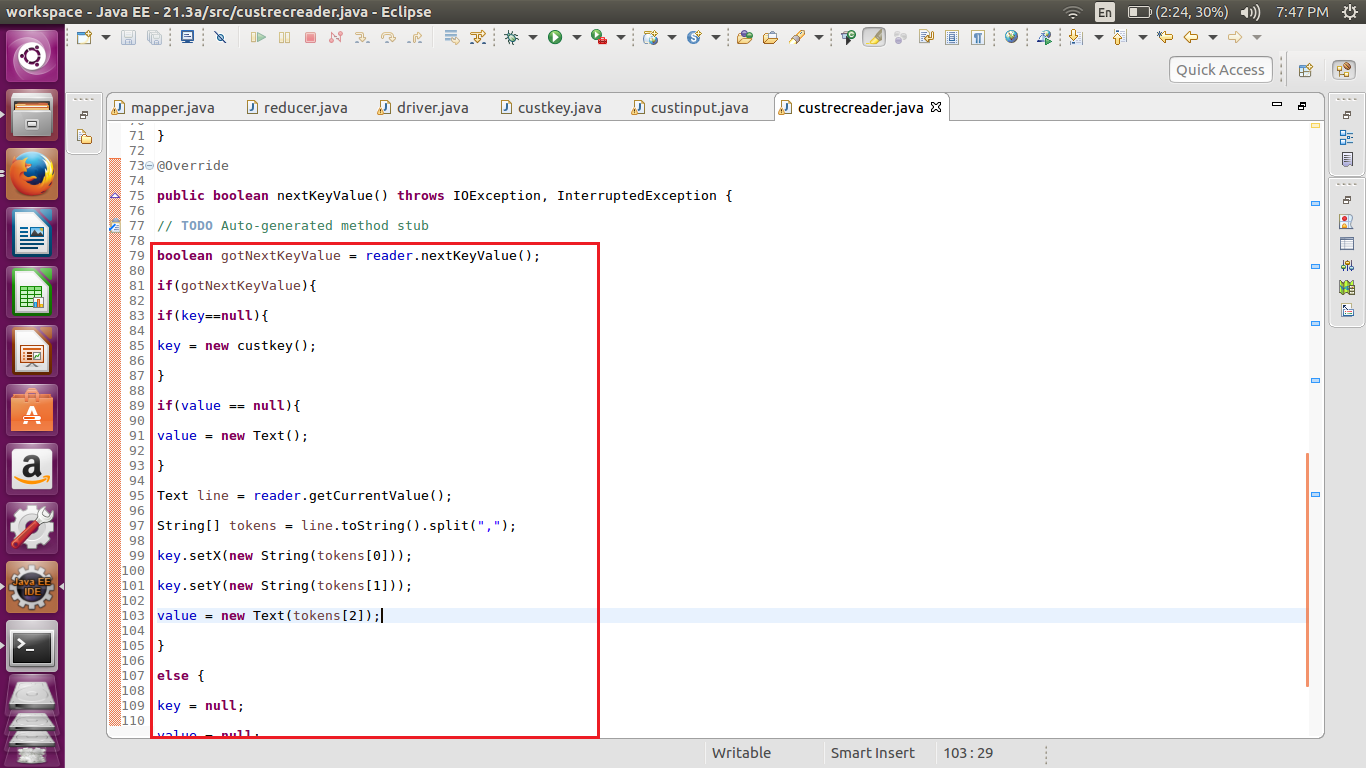
//Custom file format.

Here we are extending the fileinputformat class to return the custom record reader so as to get the desired value as the key and value.

//Custom Record Reader

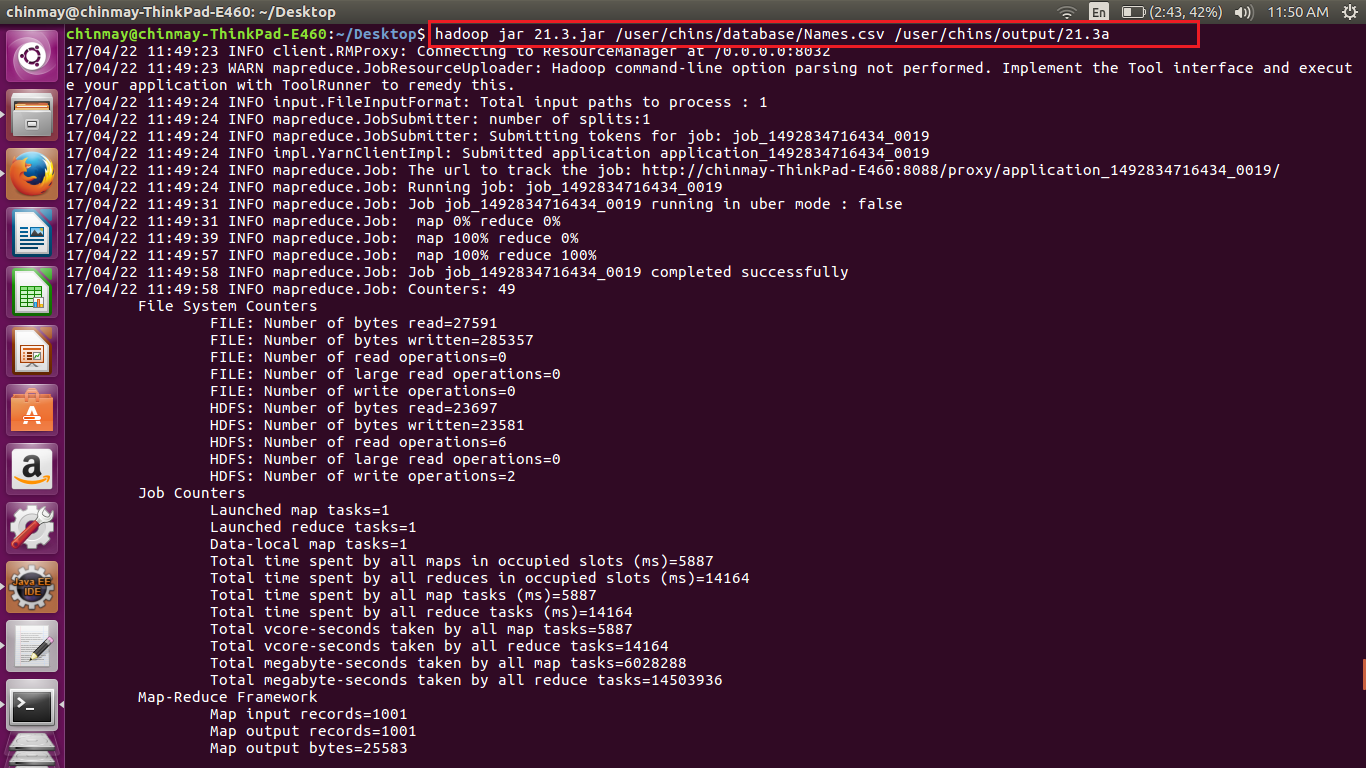






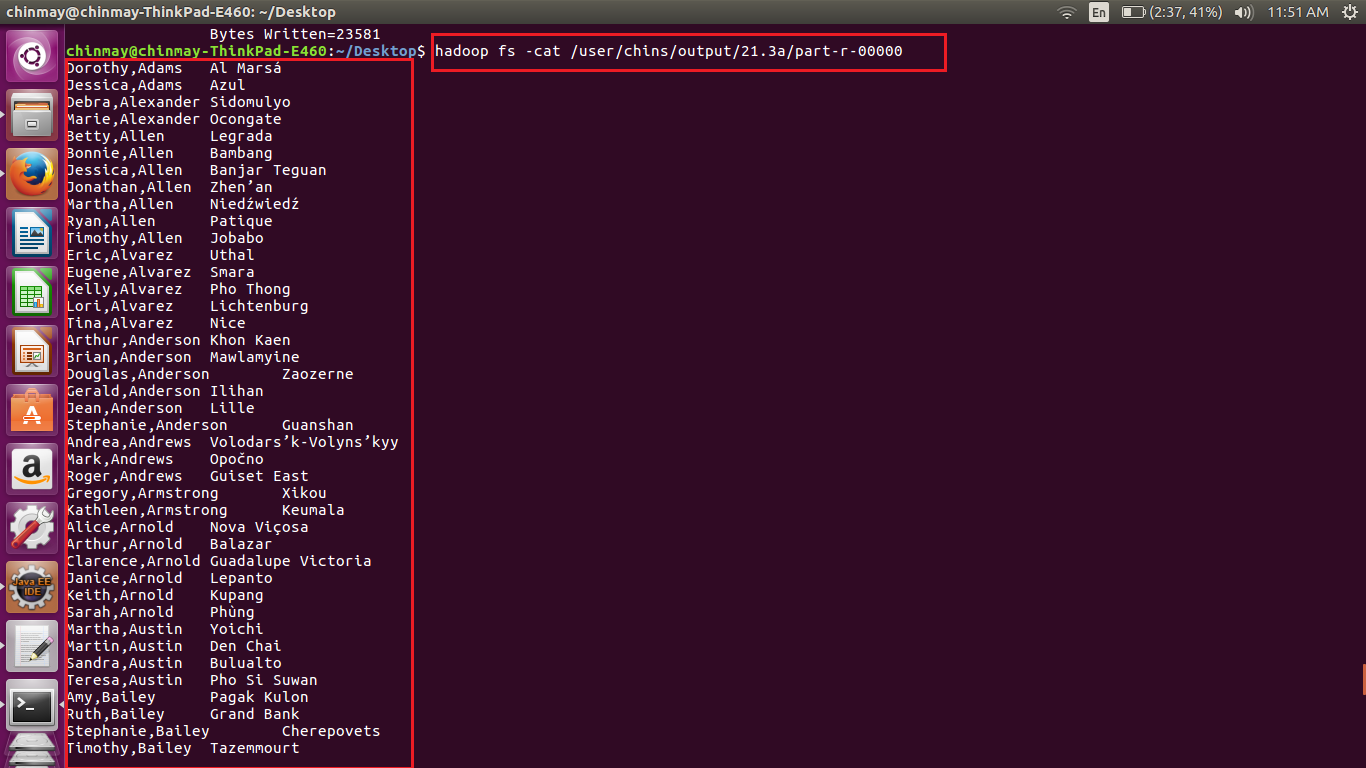
In the custom record reader we are checking if the earlier key value has been set or not and the proceeding to set the next key vale pair. In this example we have set first name and the second name as the key and city as the value.

//Implementation



Here we have set the input path and output path of the HDFS.

//Output



This is the sample output where we can see that the output is arranged with respect to the last name.

***Note – The output file is also attached.***