**ASSIGNMENT 16.5:**

**Problem Statement:**

Write a map reduce program to join the below given two datasets using map side join.

**Input Data Set:**

You can download the datasets from the below link:

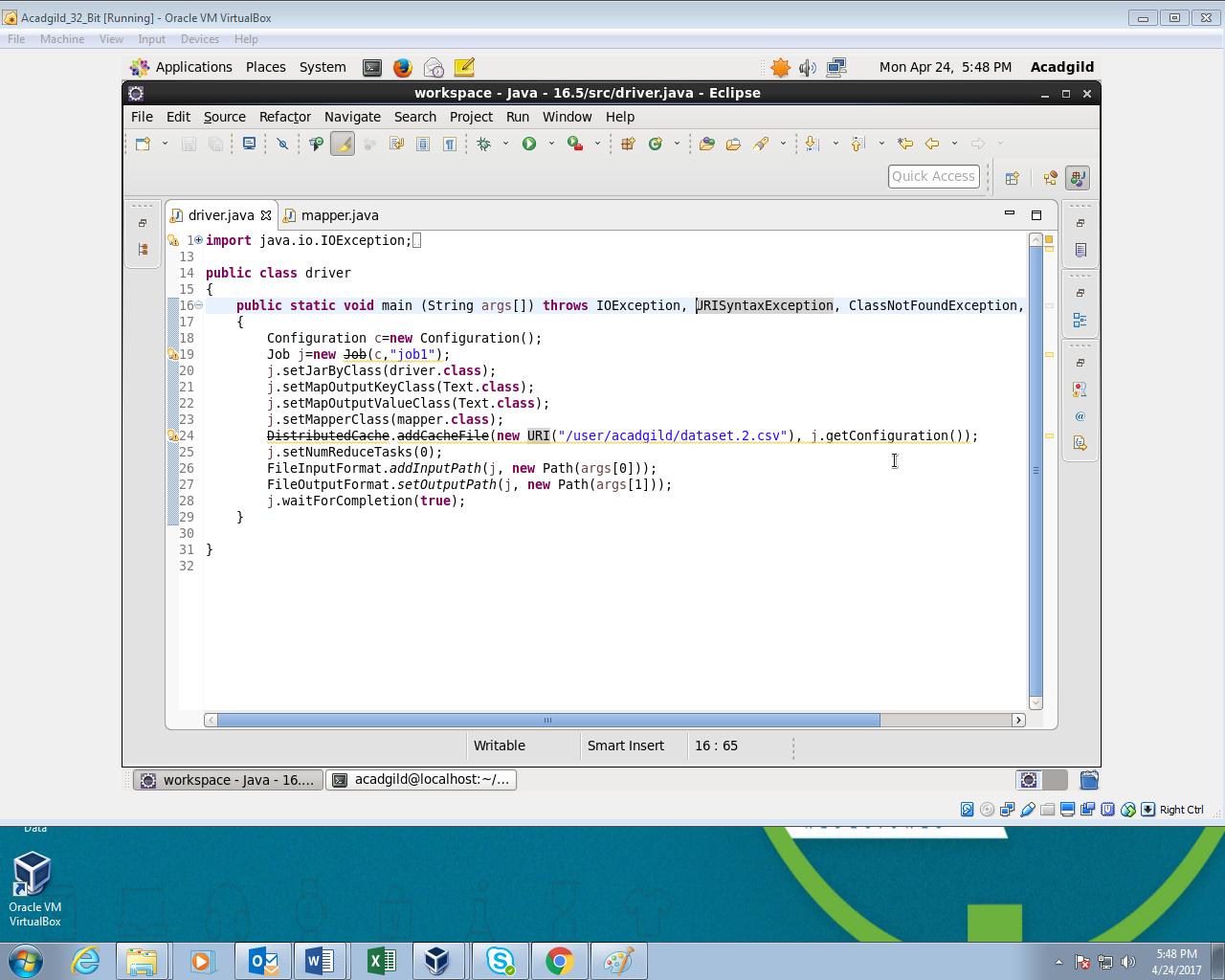
<https://drive.google.com/open?id=0ByJLBTmJojjzSWhfXzI5NGh4RGM>

Dataset description of dataset1.csv is as follows: id,first\_name,last\_name,email,gender,ip\_address Dataset description of tdataset2.csv is as follows: id,first\_name,last\_name

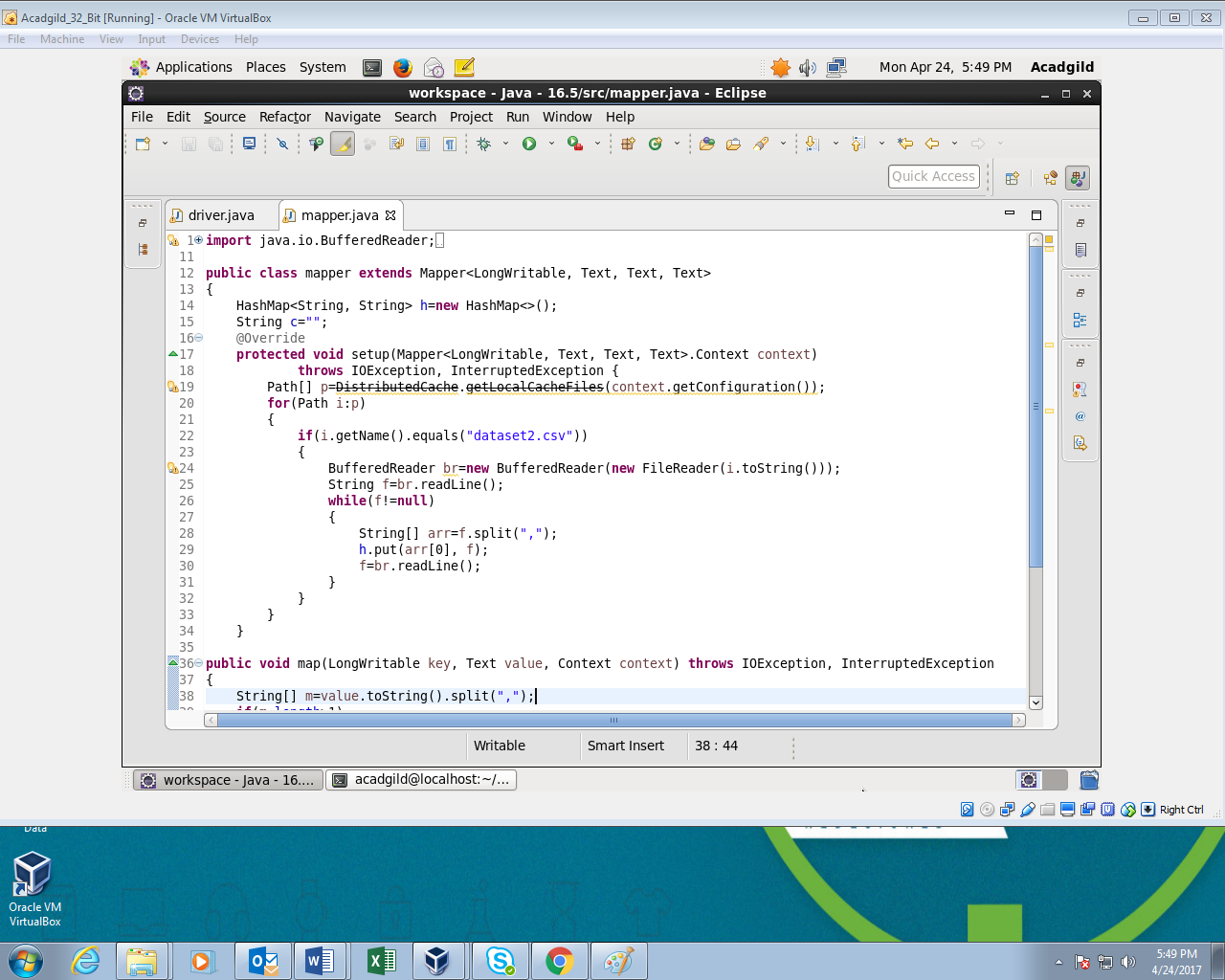
Note: You can just join the datasets using map side join.

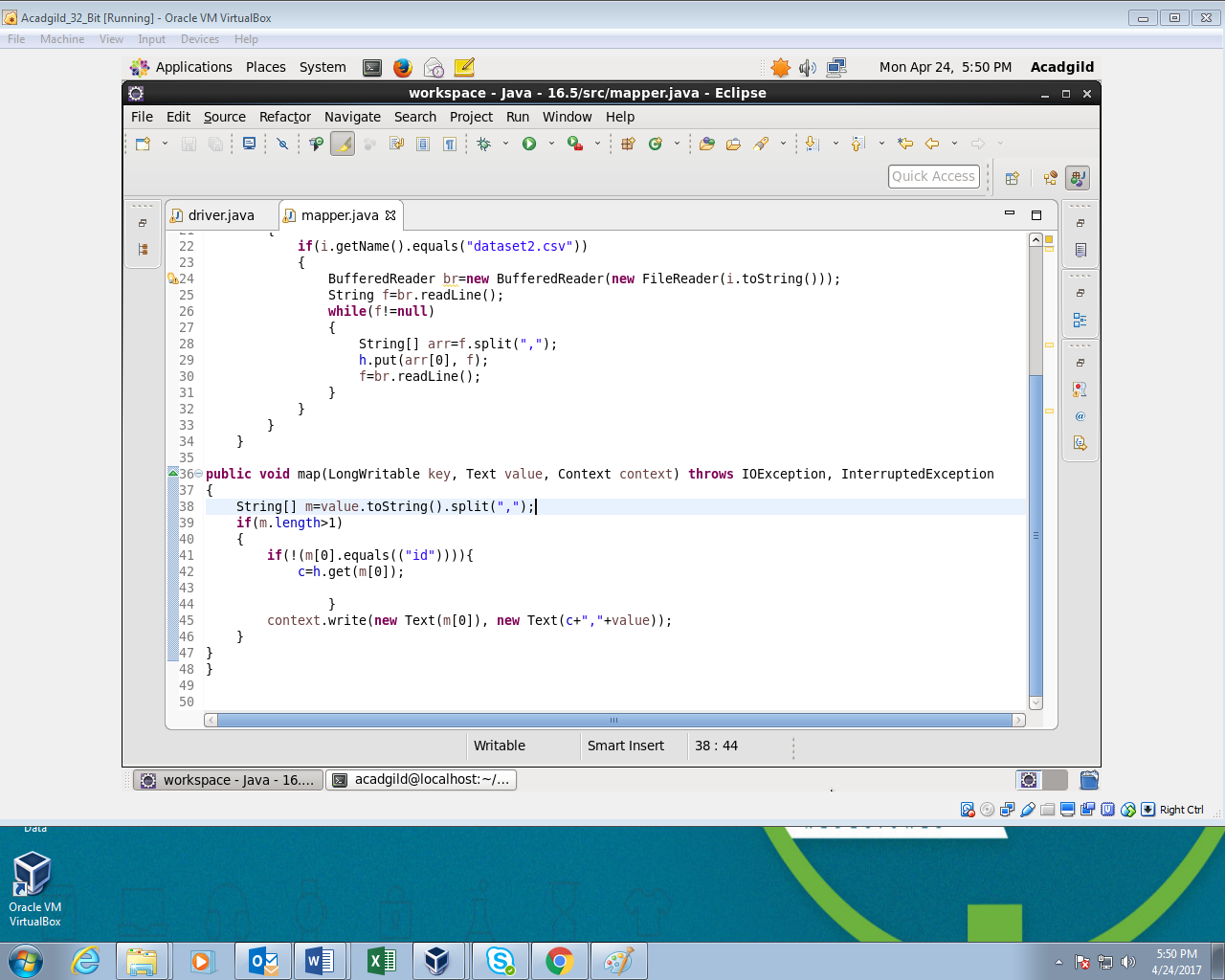
**Program:**

**Driver Class:**



**Mapper Class:**

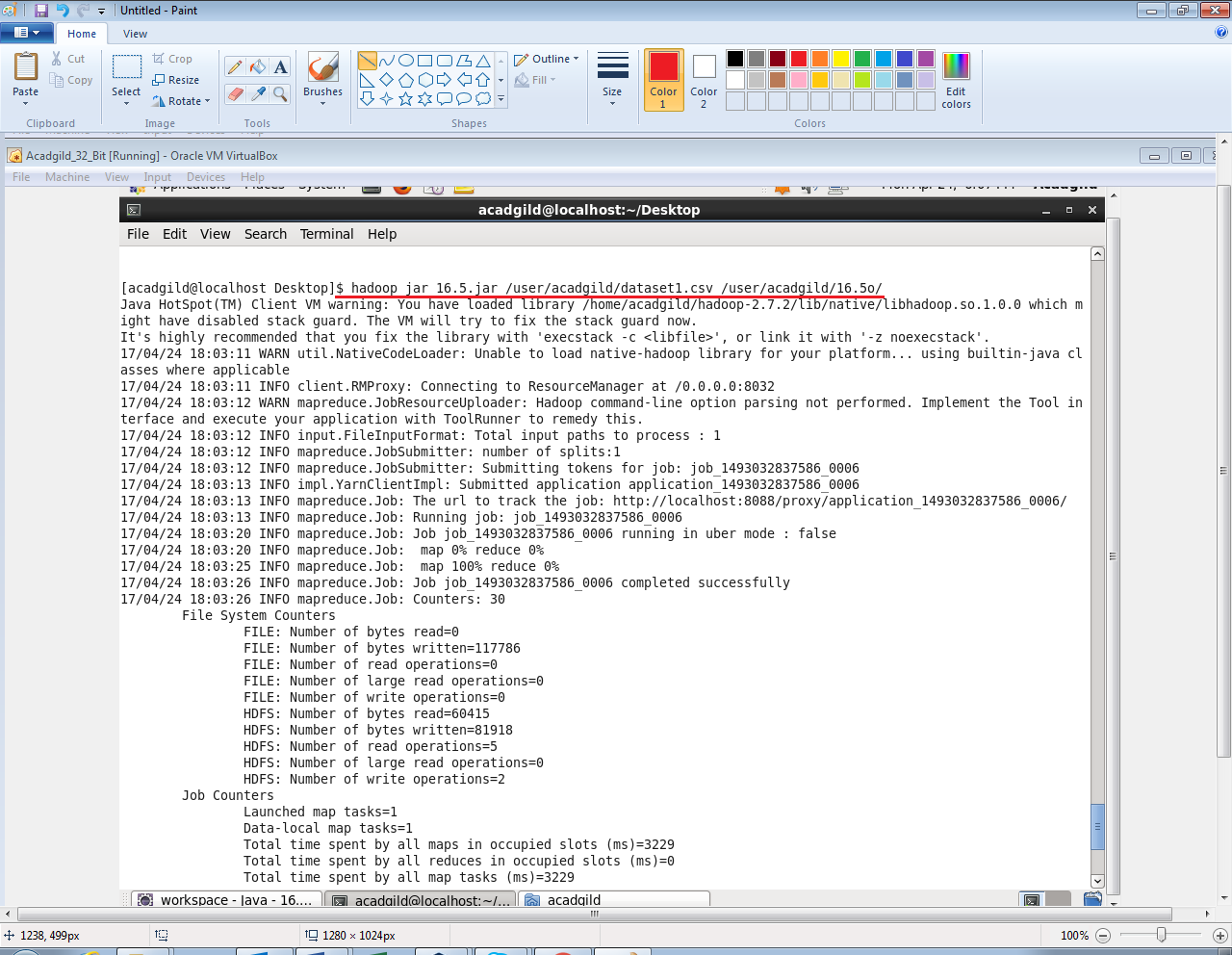


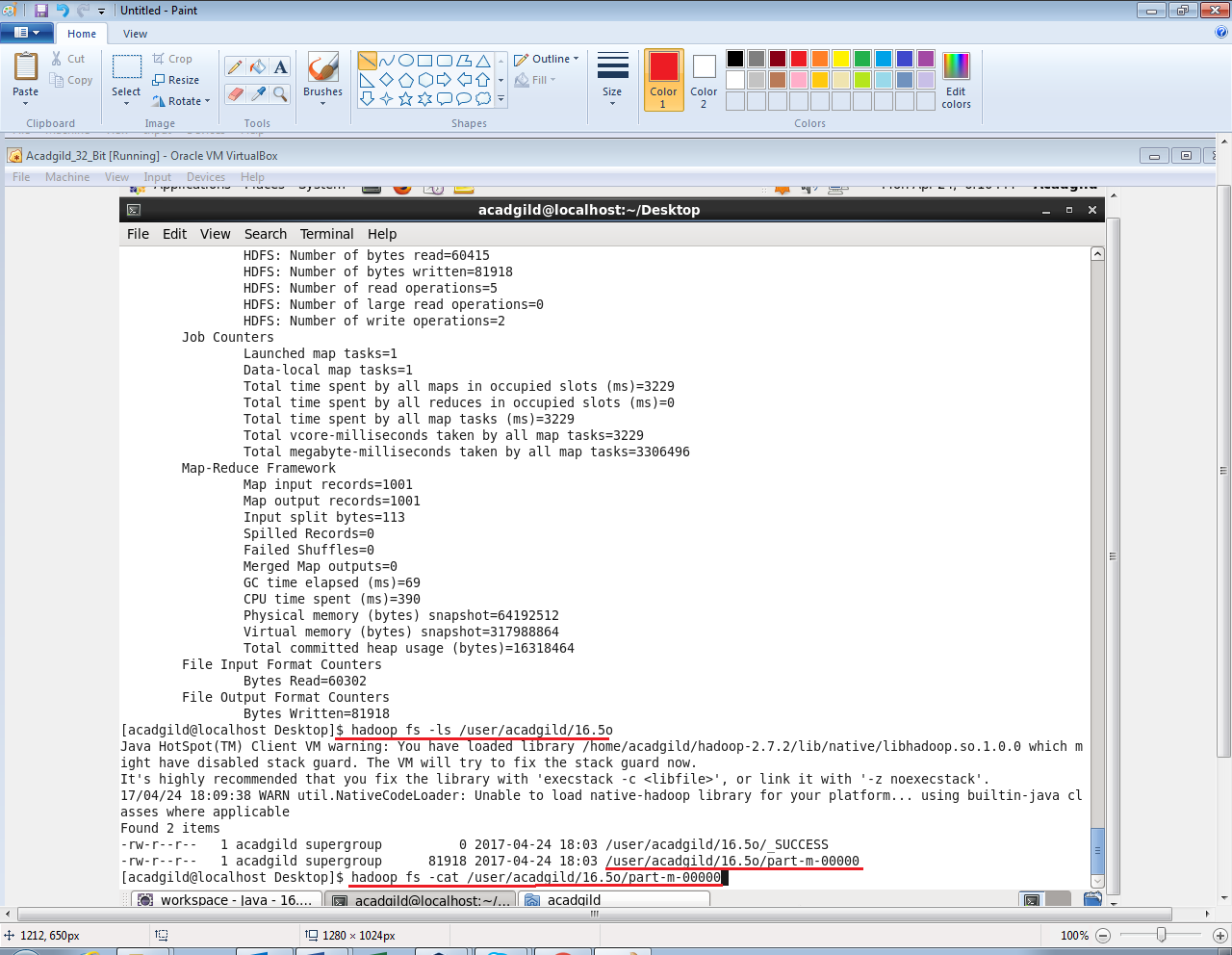


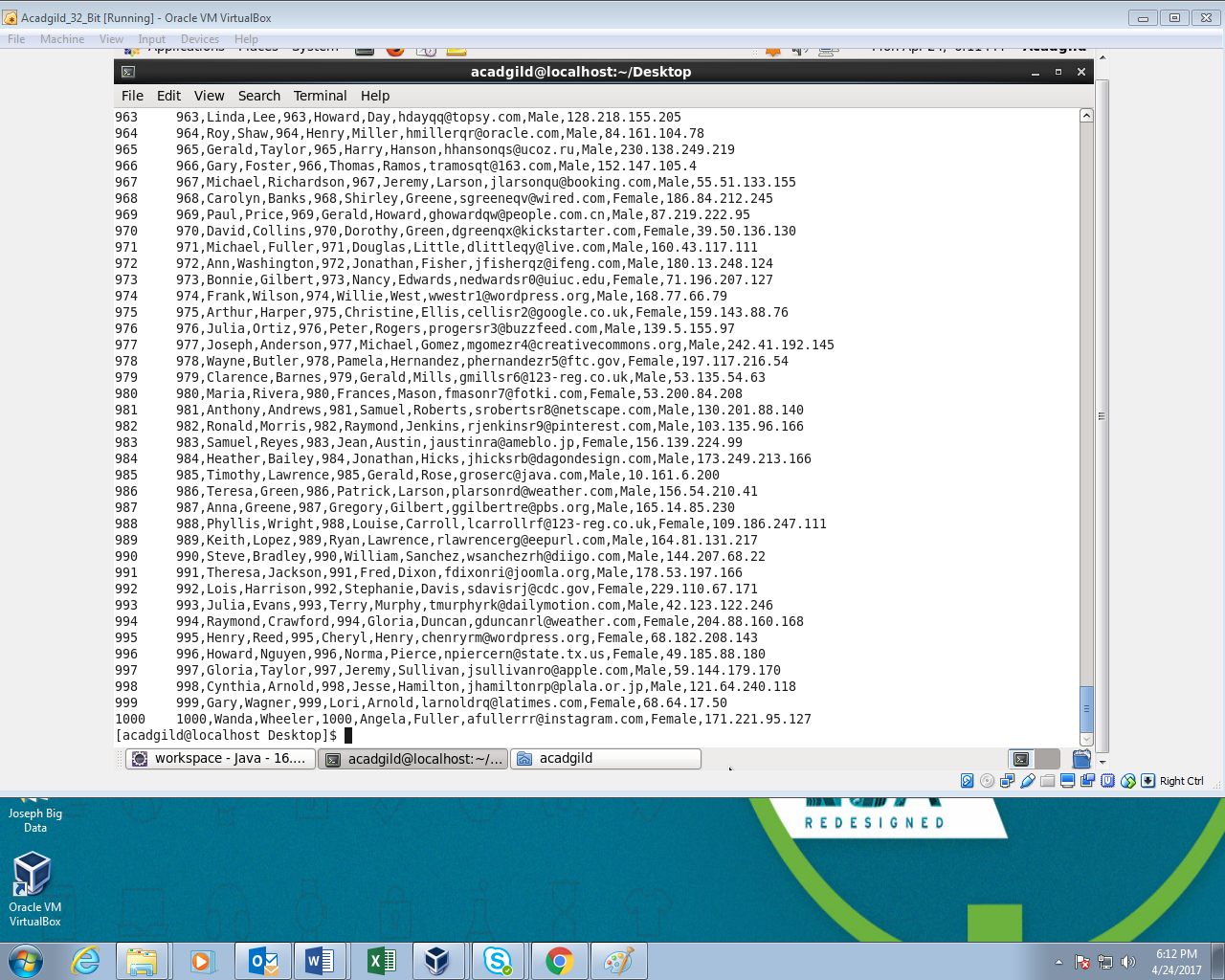
**Output:**

The records from dataset1.csv and dataset2.csv are joined using common field in both the datasets which is id.

As the output consists of large number of records screenshot of last few records is attached below:







Since the output Obtained is very large, I attached the Screenshot of the last part of the output.