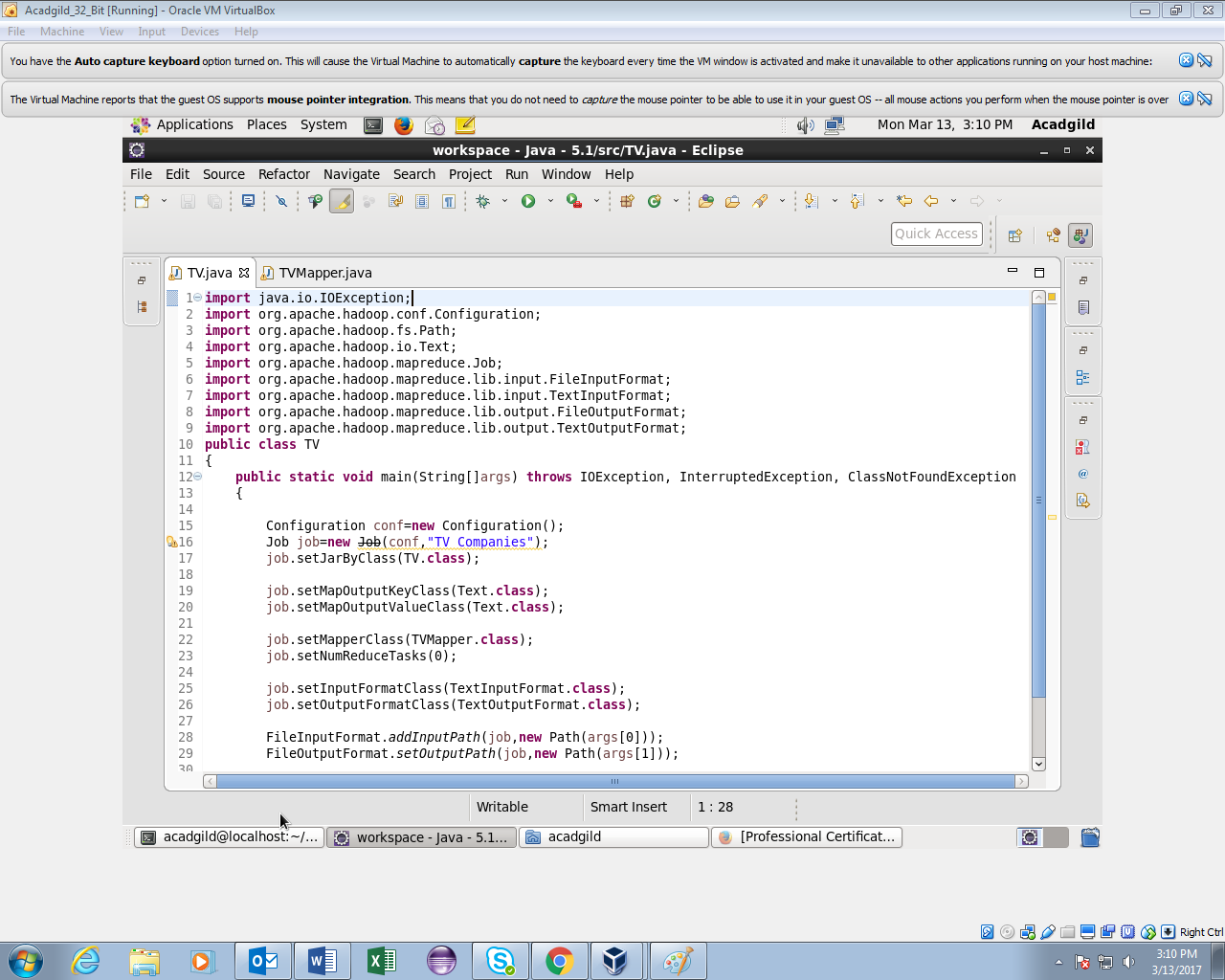
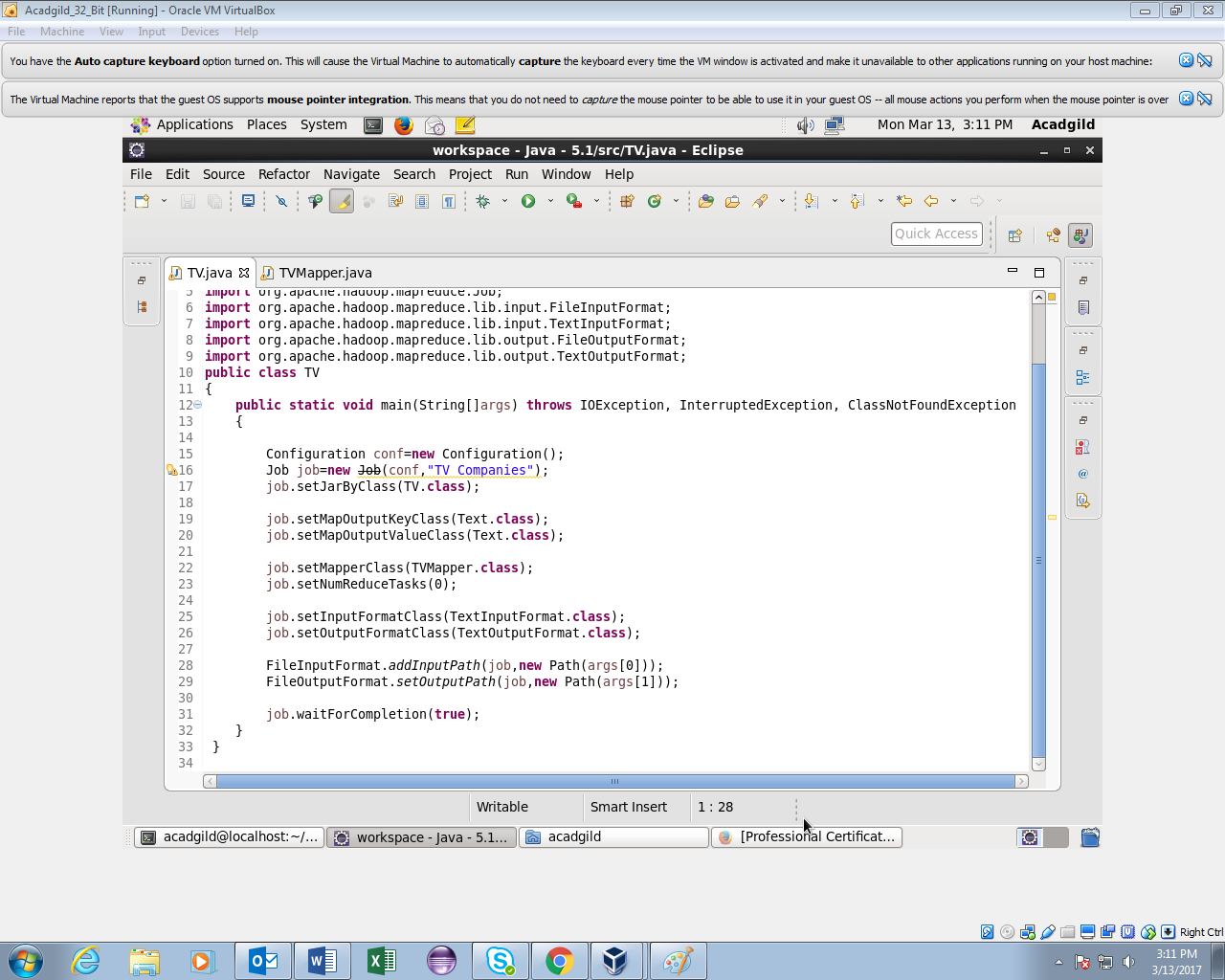
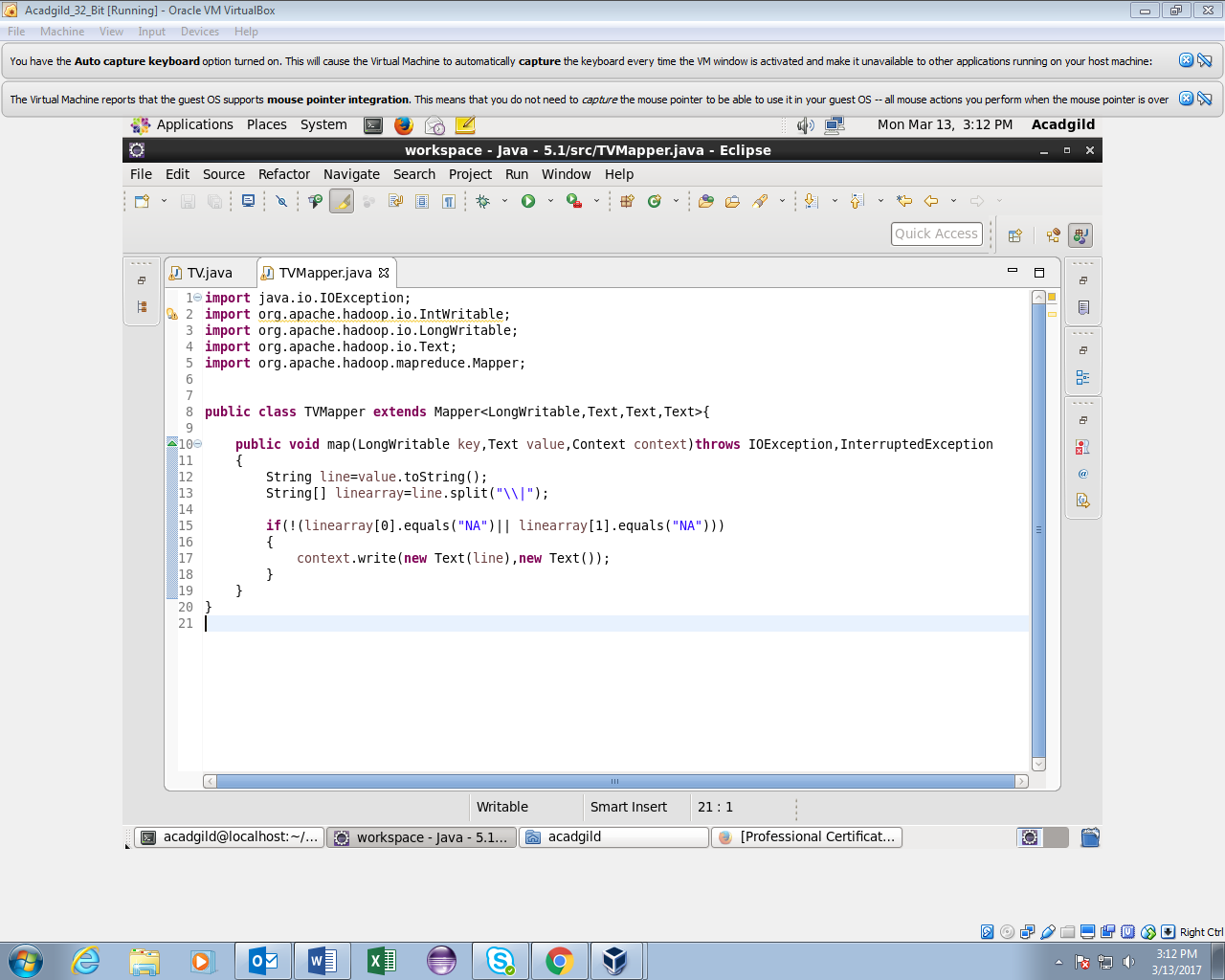
**Assignment 5.1**

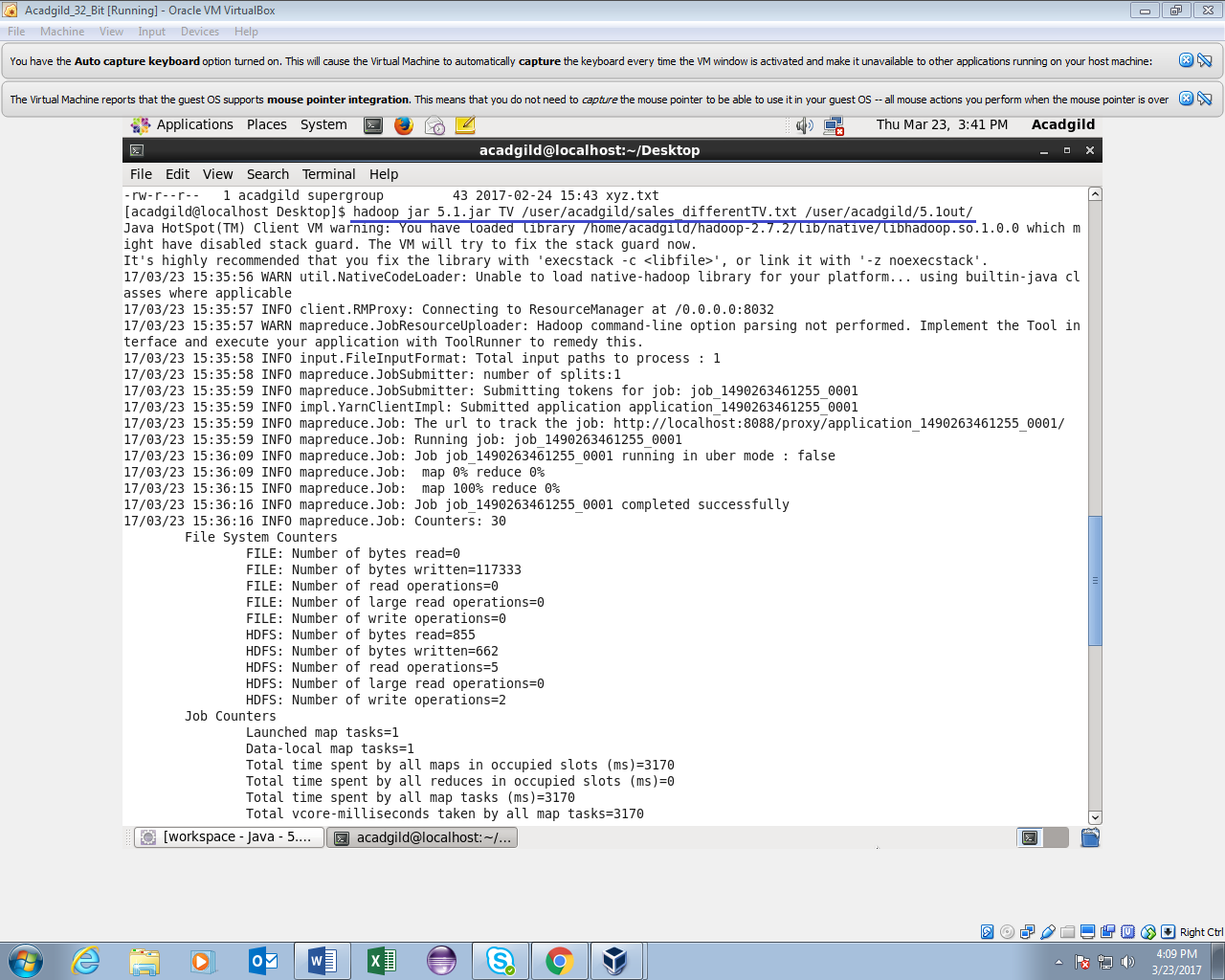
We have a dataset of sales of different TV sets across different locations. Records look like: Samsung|Optima|14|Madhya Pradesh|132401|14200 The fields are arranged like: Company Name|Product Name|Size in inches|State|Pin Code|Price sales\_differentTV : https://github.com/prateekATacadgild/DatasetsForCognizant/blob/master/sales\_differentTV.tx t There are some invalid records which contain 'NA' in either Company Name or Product Name.

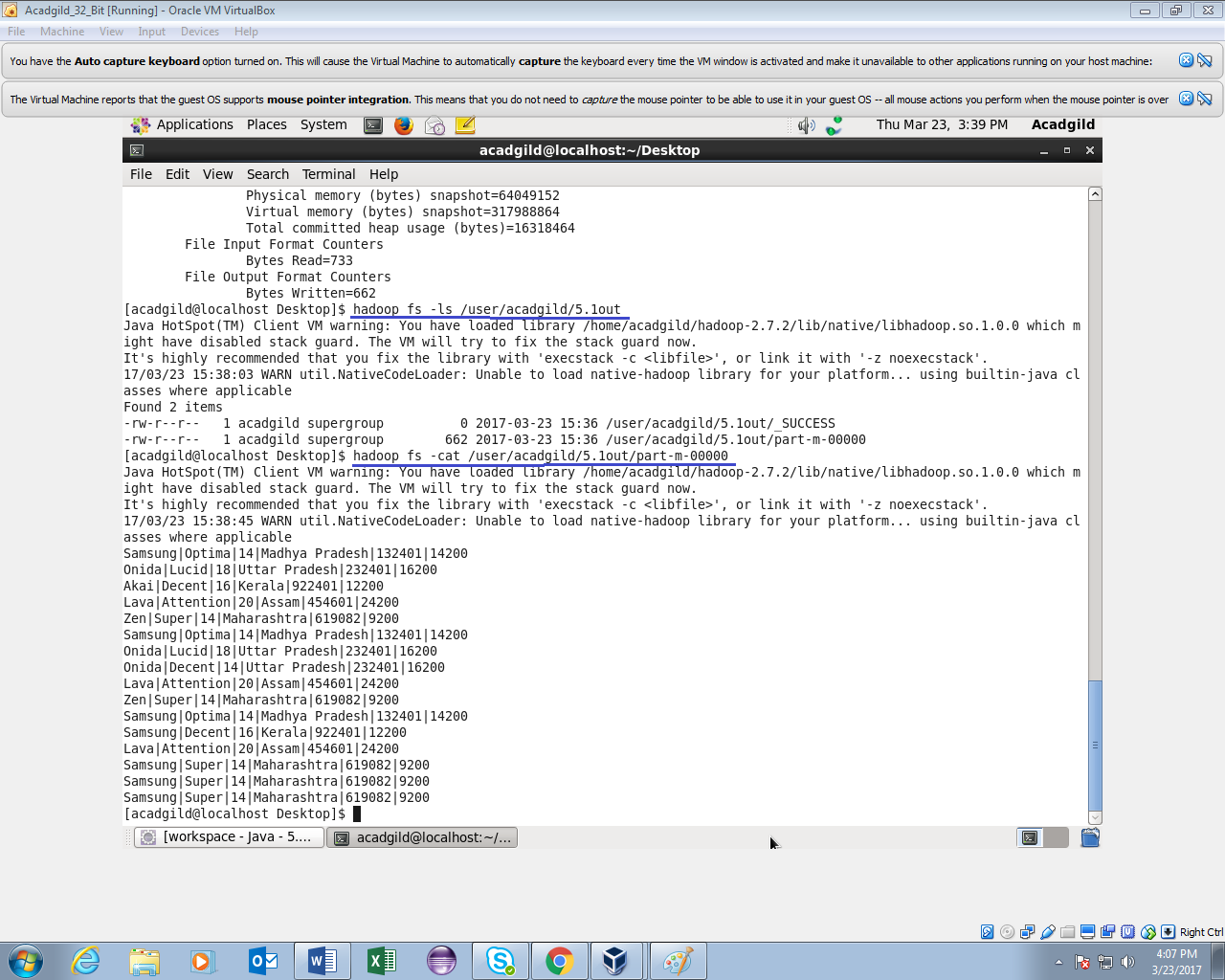
Sales of different TV Task 1: Write a MapReduce program to filter out the invalid records. The output of this program will act as input for subsequent tasks. Map only job will fit for this context.

****

****

****

****

****