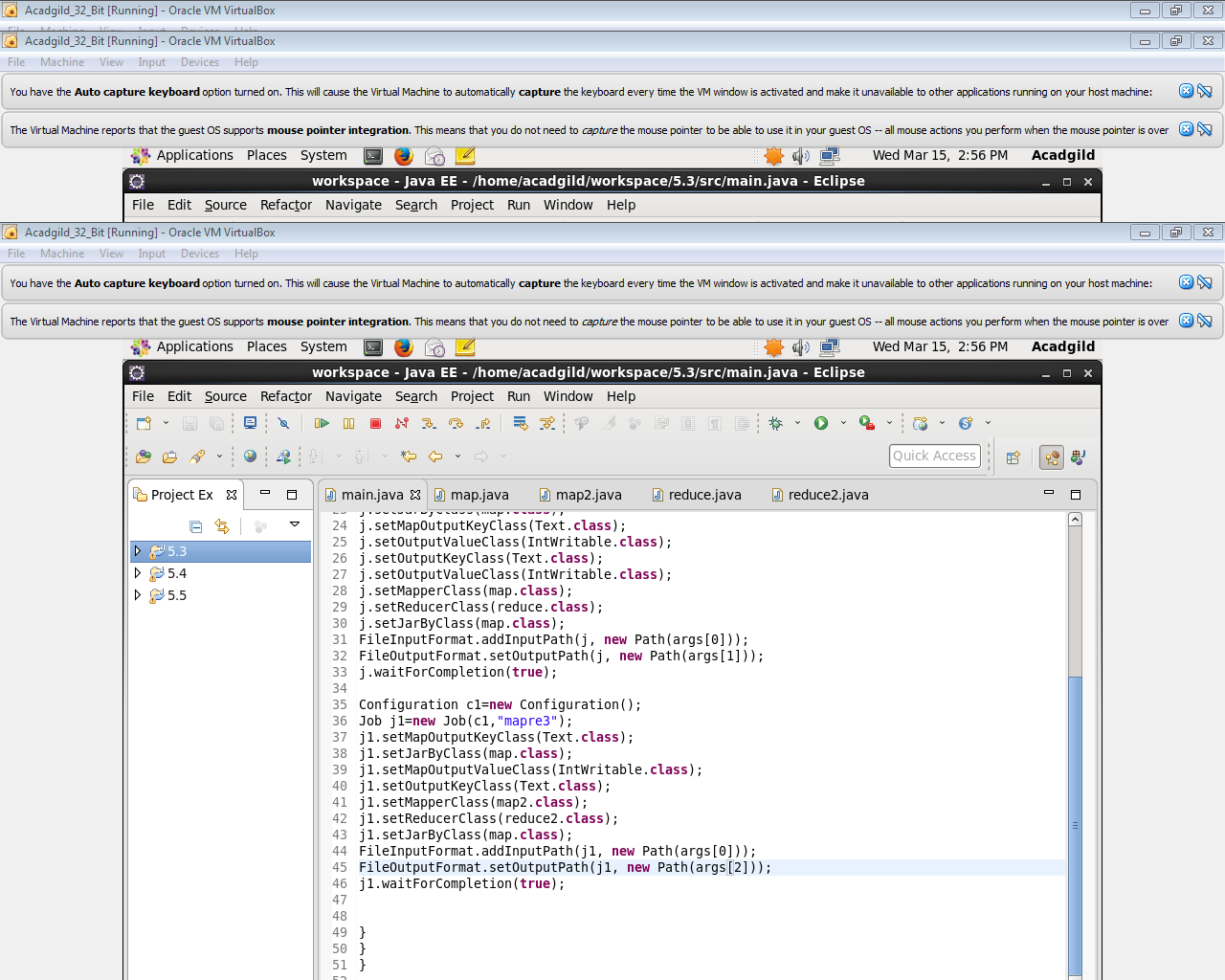
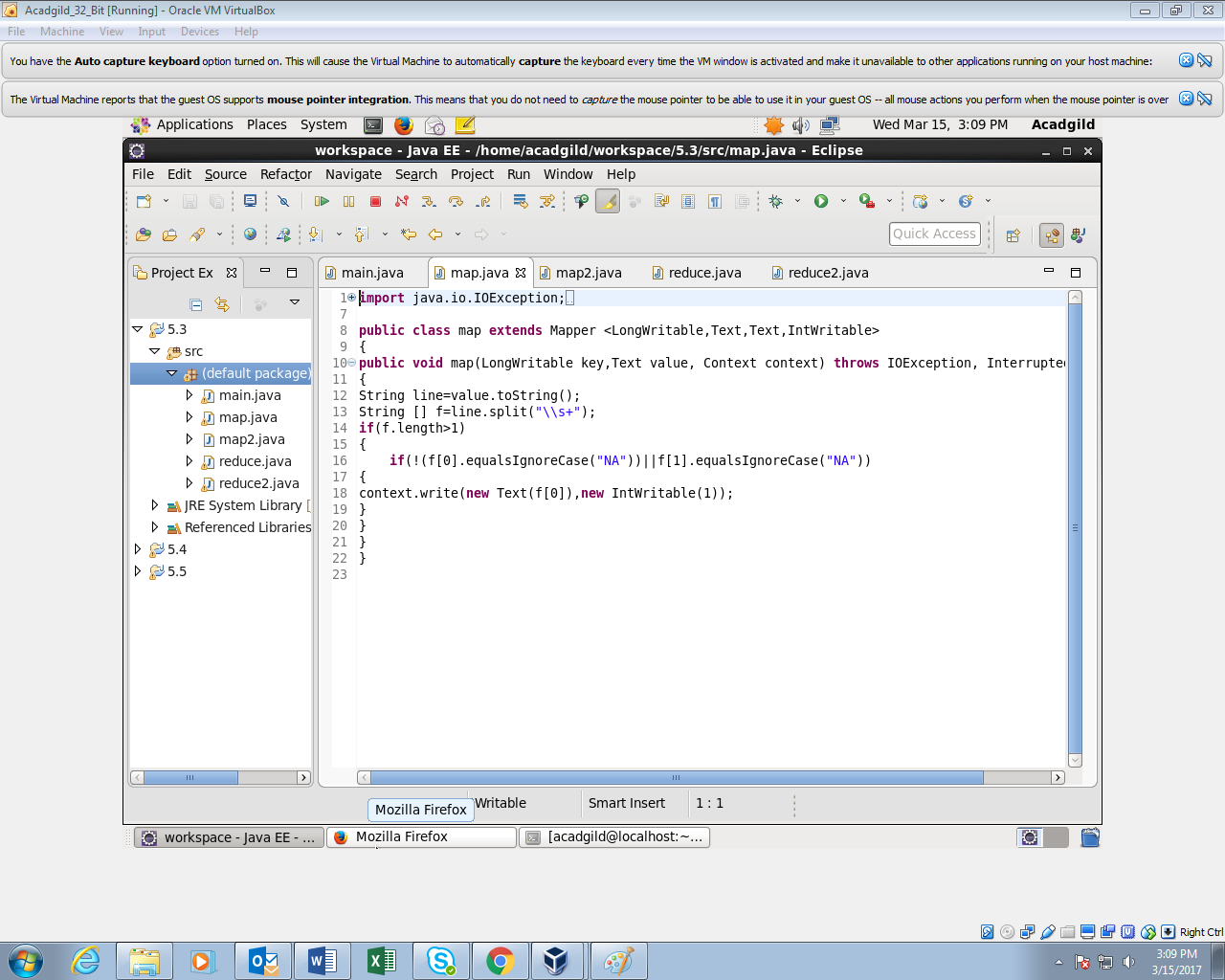
1. Introduction This assignment will help you to consolidate the concepts learnt in the session. 2. Problem Statement We have a dataset of movie data analysis Column Name: Column 1: User ID Column 2: Movie ID Column 3: Rating Column 4: Time stamp The input file can be downloaded from the below link Movie data analysis Problem statement 1: Find the number of ratings given by each user. Problem statement 2: Find the number of ratings received by each movie.

MAIN



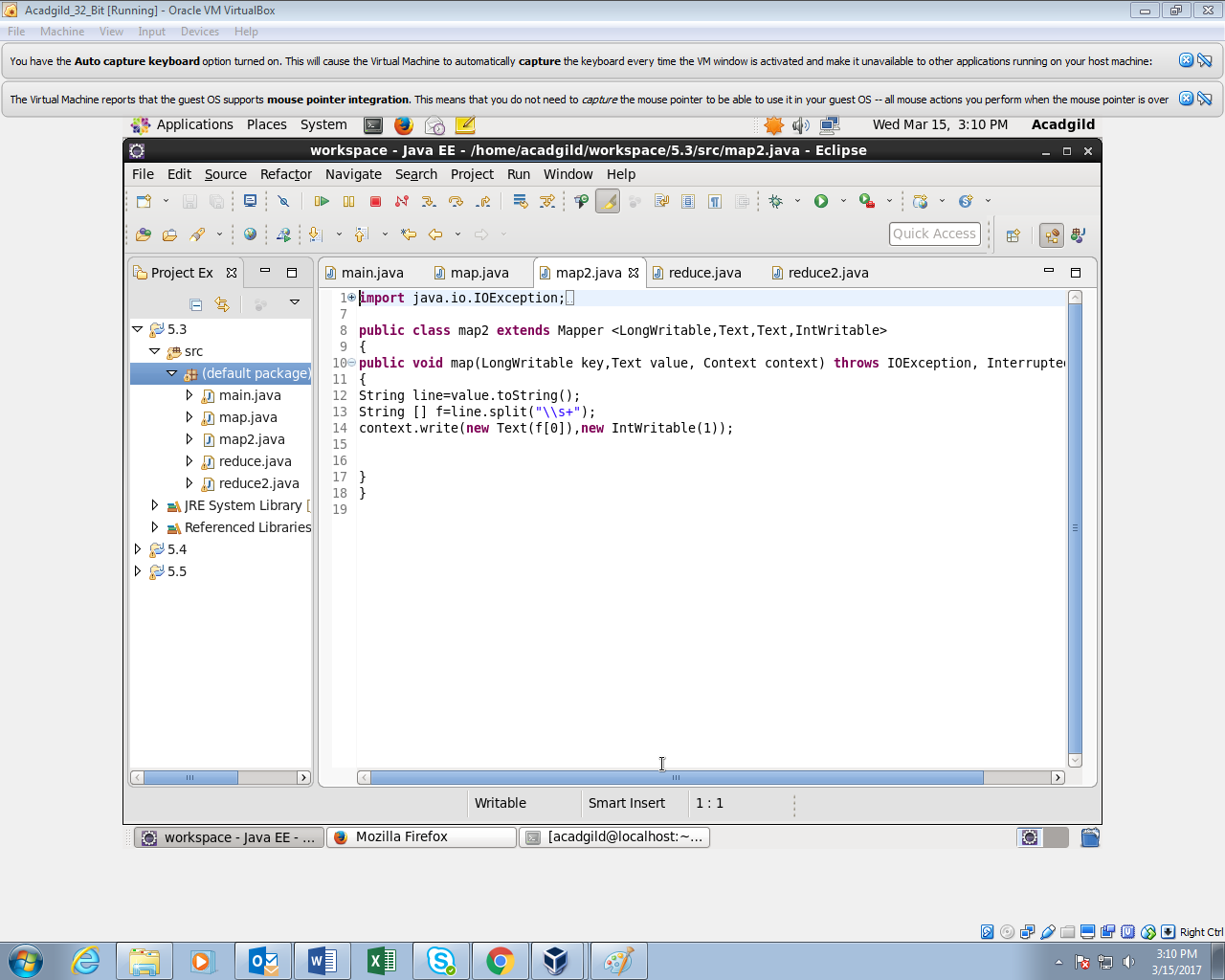
MAPPER 1

Logic : Used to split Input by [\\s+](file:///\\s+) for splitti ng by multiple spaces and send the user Id and a count of 1 for each Rating given by a user for a movie as Key and value



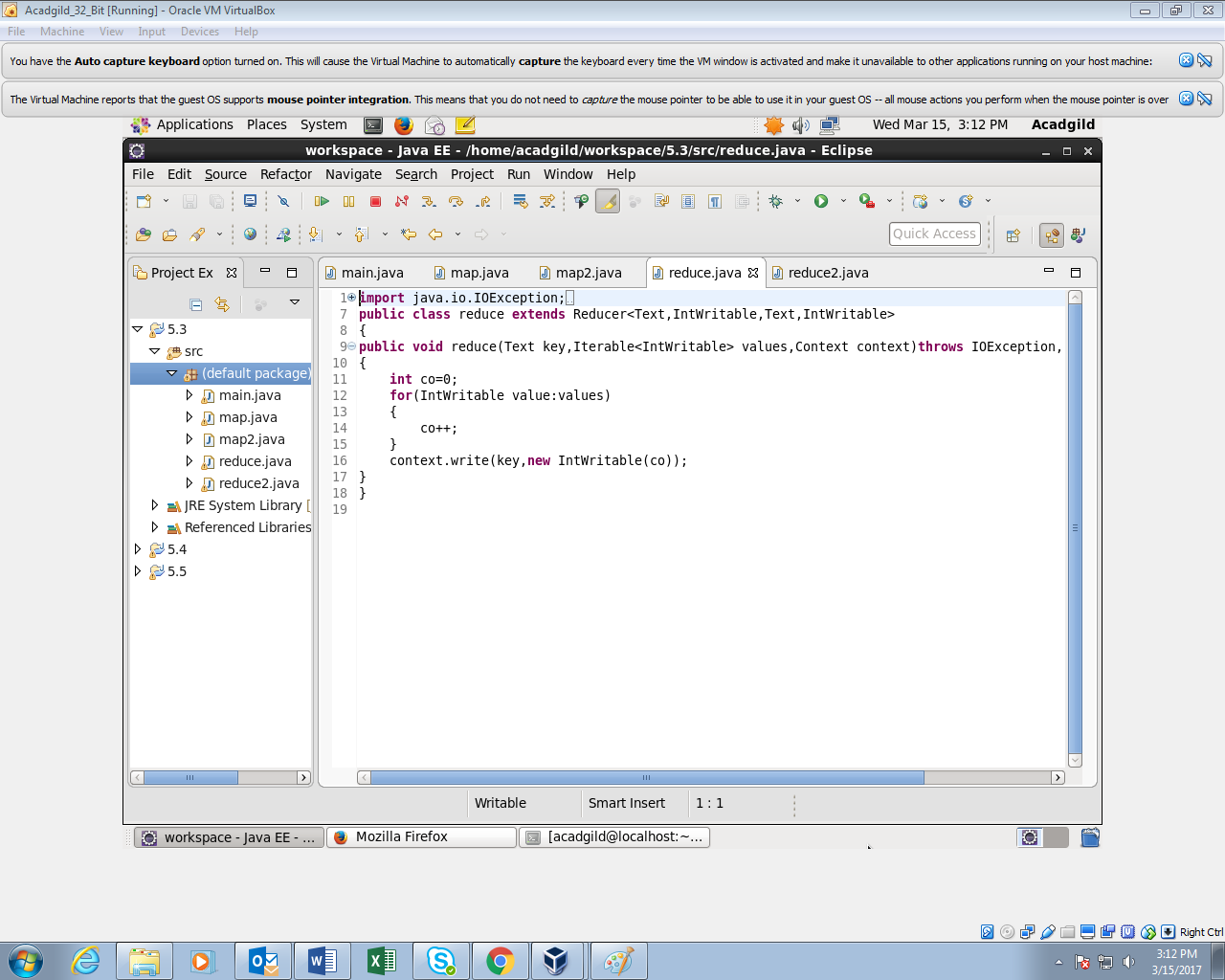
MAPPER 2

Logic : Used to split Input by [\\s+](file:///\\s+) for splitting by multiple spaces and send the Movie Id and a count of 1 for each Rating given by the users for each movie as Key and value



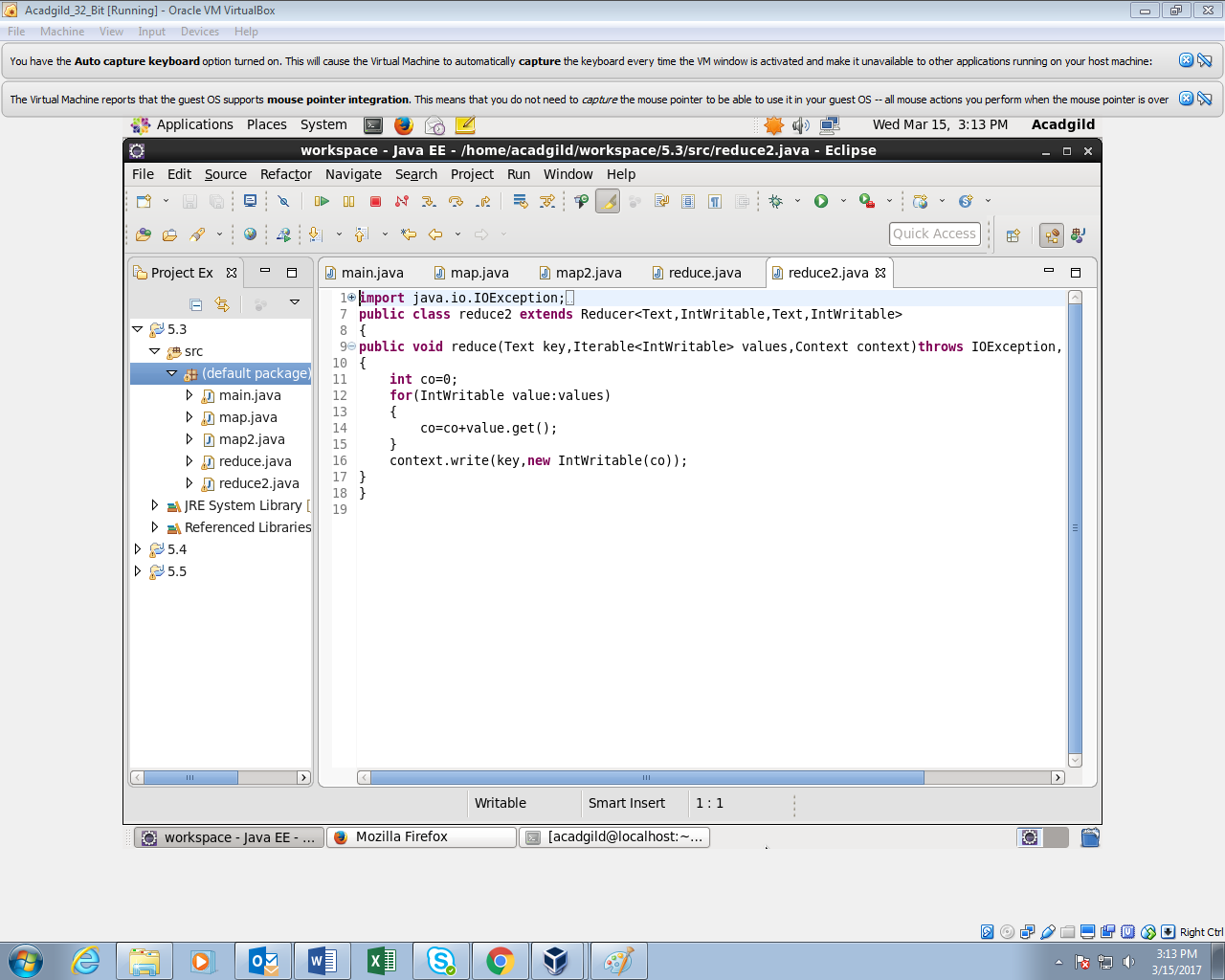
REDUCER 1

Count the no of rating given by each user by iterating through values and adding

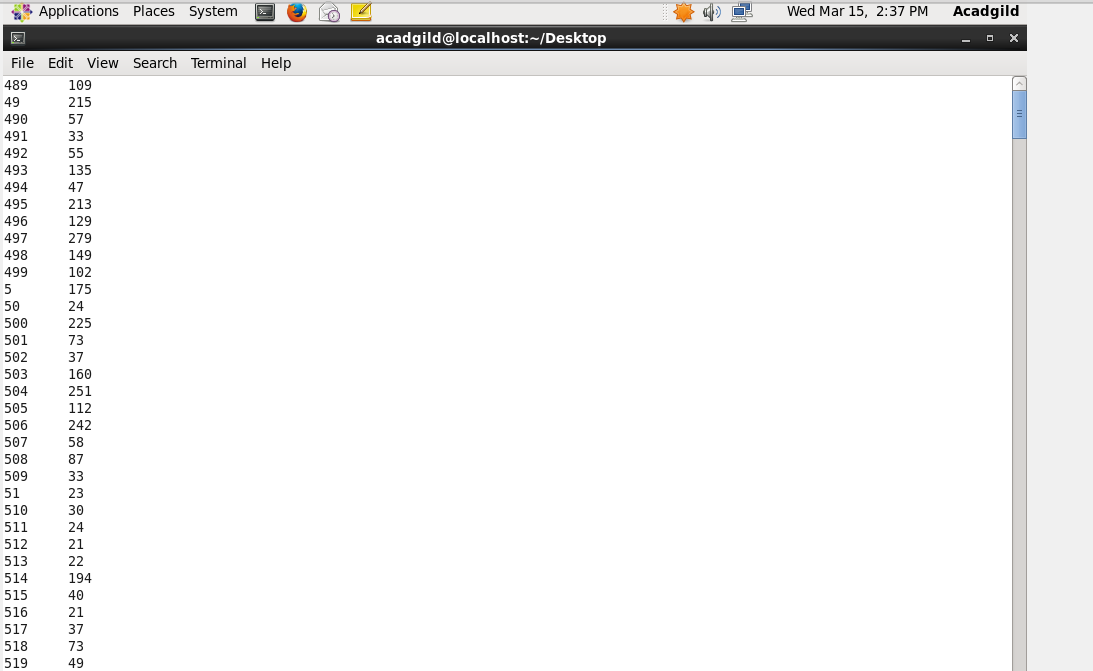


REDUCER 2

Count the no of rating given by each user by iterating through values and adding



OUTPUT 1



OUTPUT 2

