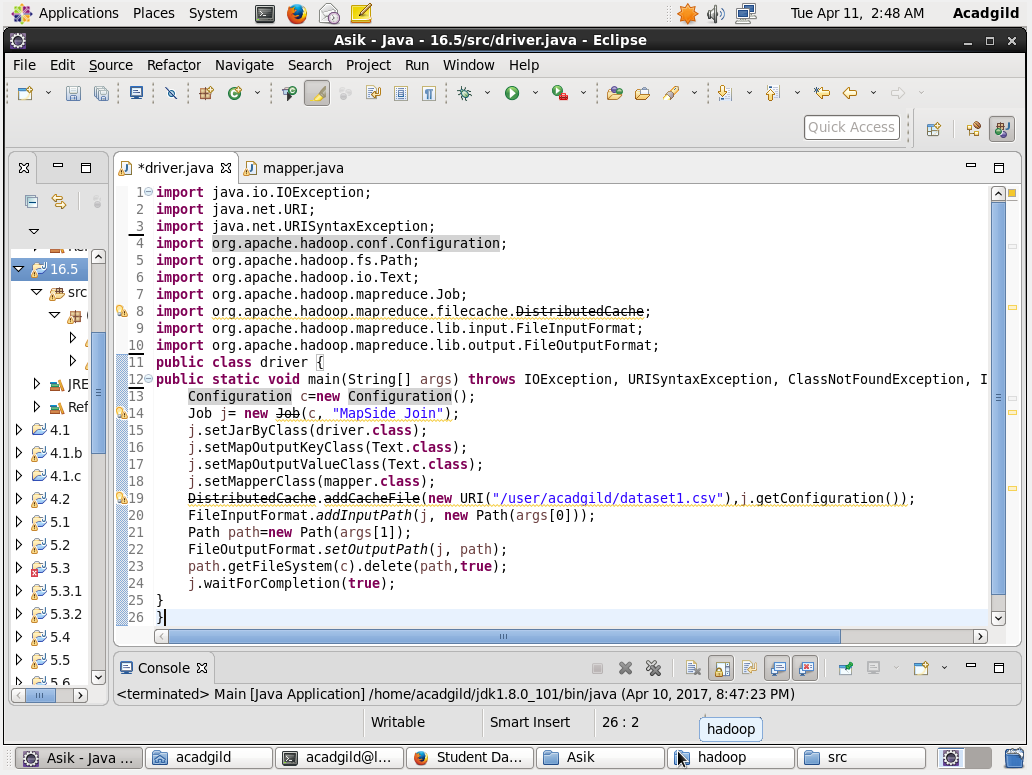
Assignment-16.5:

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

Driver Class:

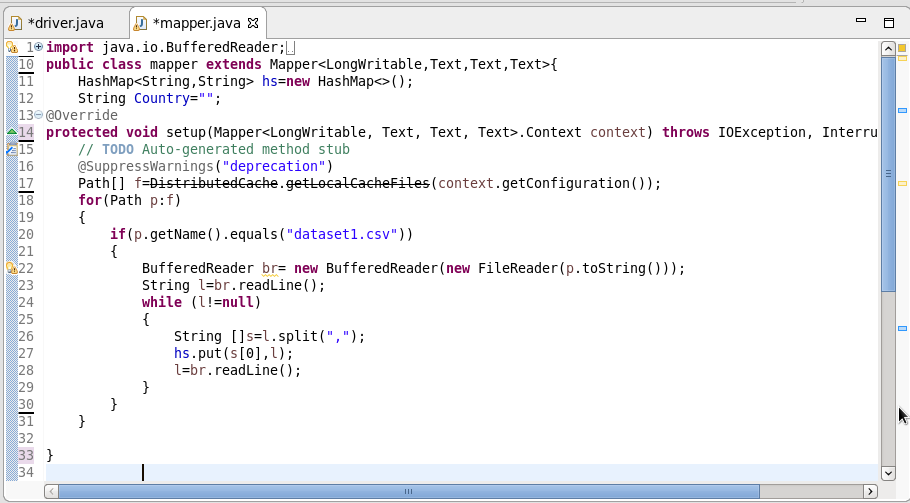


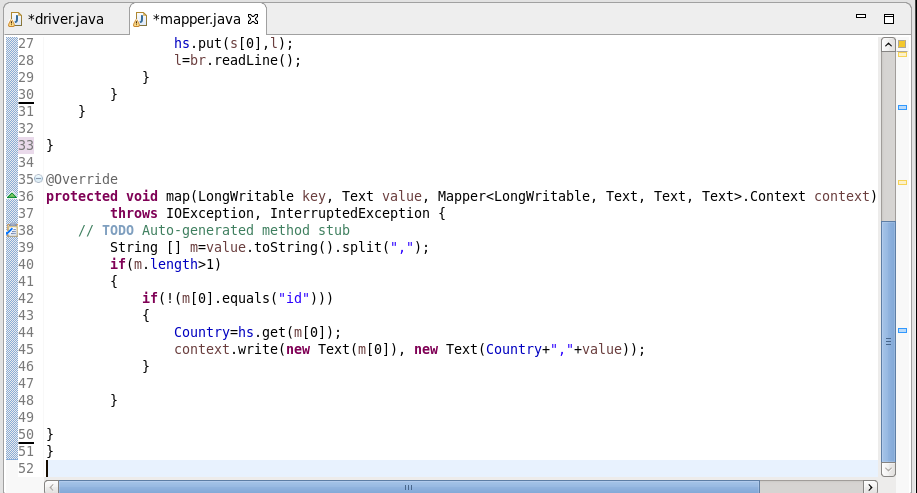
Mapper Class-1:

Overriding the setup method:

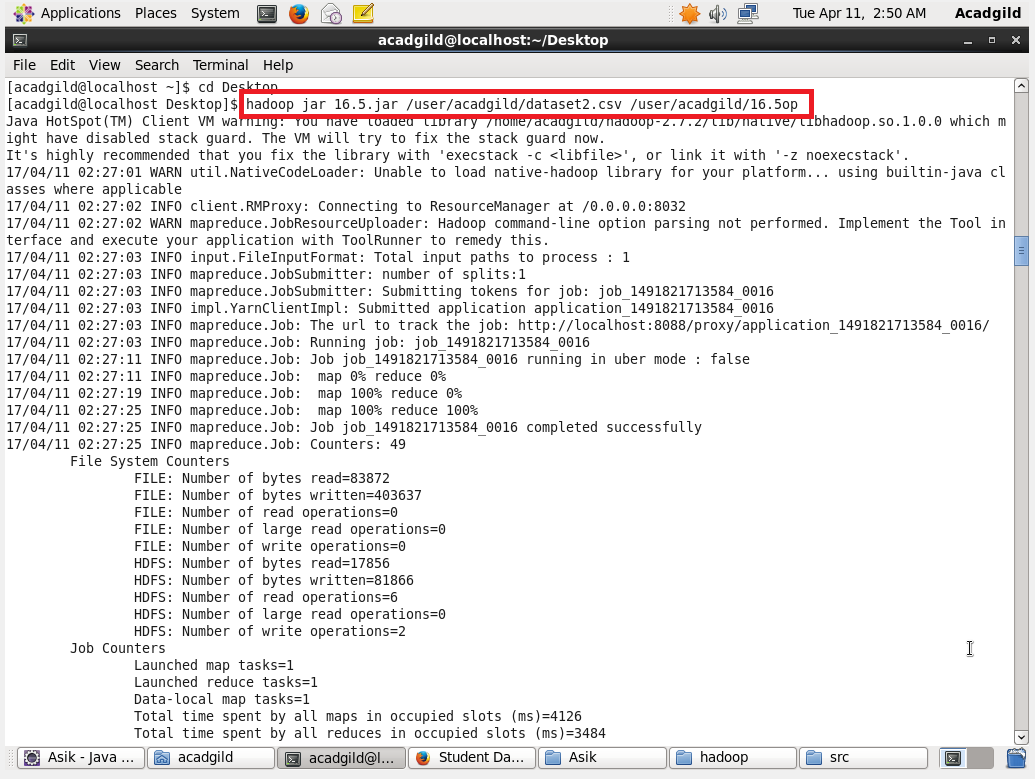
In the mapper class we are overriding setup method. It is used to get the dataset from distributed cache and then I used hashmap and storing the id as key of hashmap and total line as value

In map method I’m sending the id as key and joining both values separated by comma





Running the Jar File:



Output:

This is a screenshot of the sample output file. I could not take the whole file from the virtual machine.

