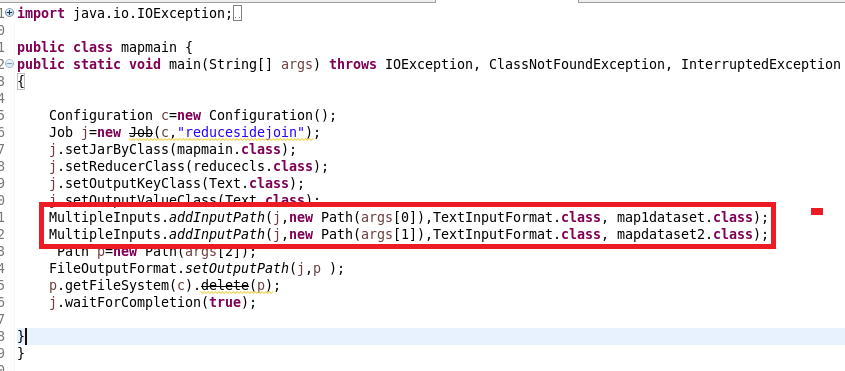
**Use reduce side join to join the two datasets and find What kind of problems are prevalent, and in which state?**

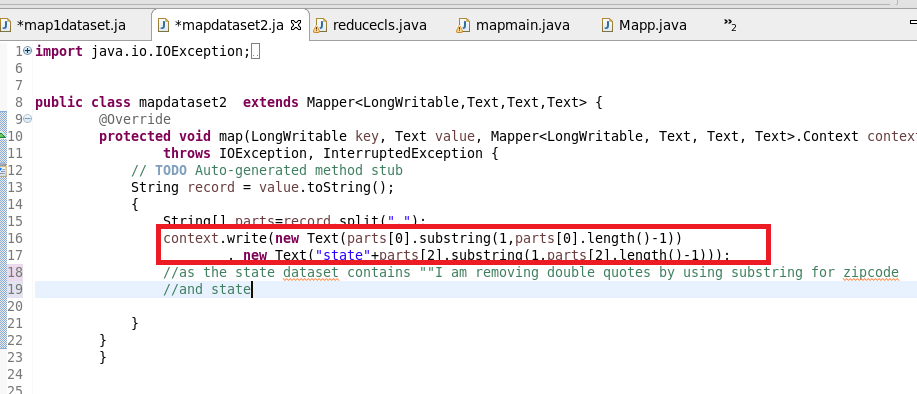
Driver:

Since it is a reduce side join Multiple Input format class is used to get the arguments as shown



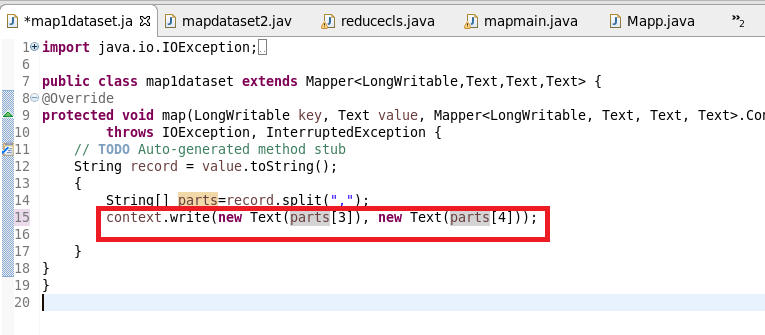
**Mapper1:**

In Mapper I am getting the zip code dataset in which I am splitting it by “,” and sending zipcode as key and state as value which I am appending with state to Identify it in reducer



**Mapper2:**

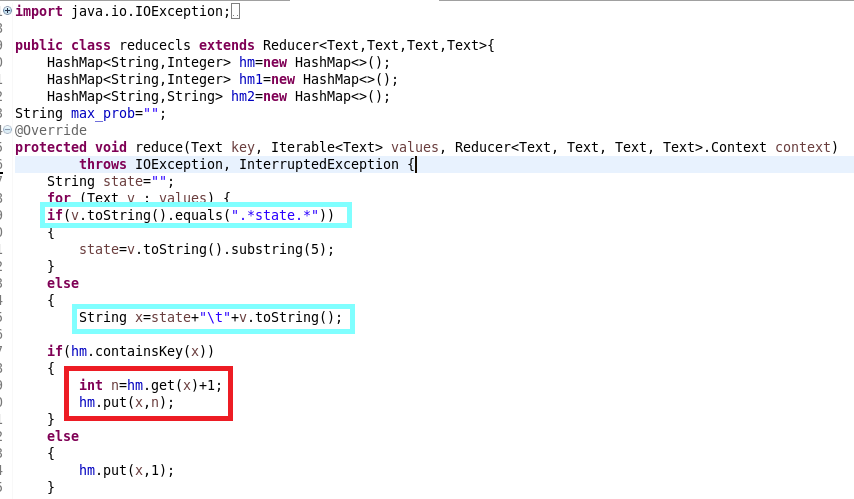
I am splitting the 911 dataset by comma and sending zipcode as key and problems as value.



**Reducer:**

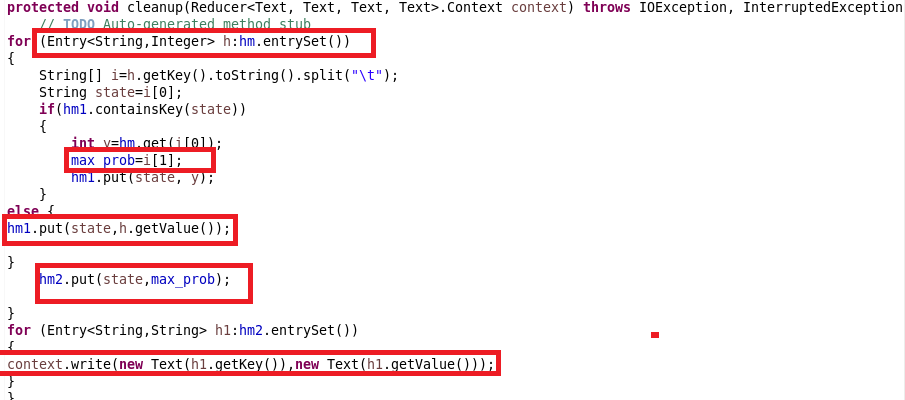
**In reduce Method**

**If it is a state I am storing it in a state variable by using if else I am storing it in a variable x which contains state and corresponding problem and I am feeding it in a hashmap(hm) where I will provide this key and a count will be incremented as we want maximum problem .Now we have all the problems along with their count**

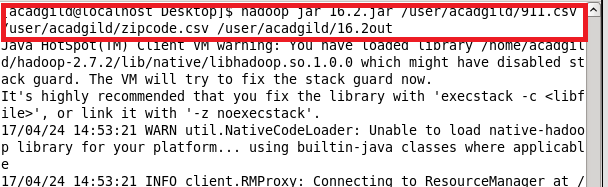


**Clean Up Method**

Now since we want only max problem I am putting a new hashmap(hm2) on which I am sending the state as key so If value as hm ie count of problem I am using a maximum logic to get the maximum value(the problem which occurred max) and I am storing the problem with max value and In the third hashmap(hm2) I am putting State as key and maximum problem as value



**Running Jar**



**Output**

