Problem Statement:

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 2: Write a MapReduce program to calculate the total units sold for each Company.

Samsung|Optima|14|Madhya Pradesh|132401|14200

Onida|Lucid|18|Uttar Pradesh|232401|16200

Akai|Decent|16|Kerala|922401|12200

Lava|Attention|20|Assam|454601|24200

Zen|Super|14|Maharashtra|619082|9200

Samsung|Optima|14|Madhya Pradesh|132401|14200

Onida|Lucid|18|Uttar Pradesh|232401|16200

Onida|Decent|14|Uttar Pradesh|232401|16200

Onida|NA|16|Kerala|922401|12200

Lava|Attention|20|Assam|454601|24200

Zen|Super|14|Maharashtra|619082|9200

Samsung|Optima|14|Madhya Pradesh|132401|14200

NA|Lucid|18|Uttar Pradesh|232401|16200

Samsung|Decent|16|Kerala|922401|12200

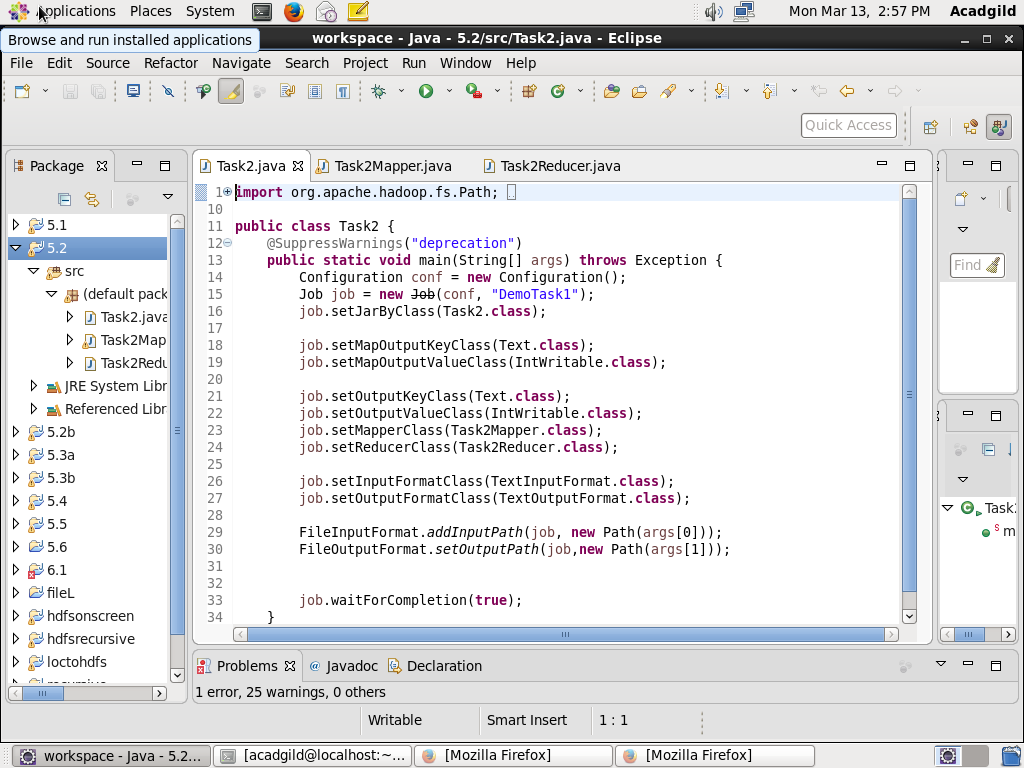
Lava|Attention|20|Assam|454601|24200

Samsung|Super|14|Maharashtra|619082|9200

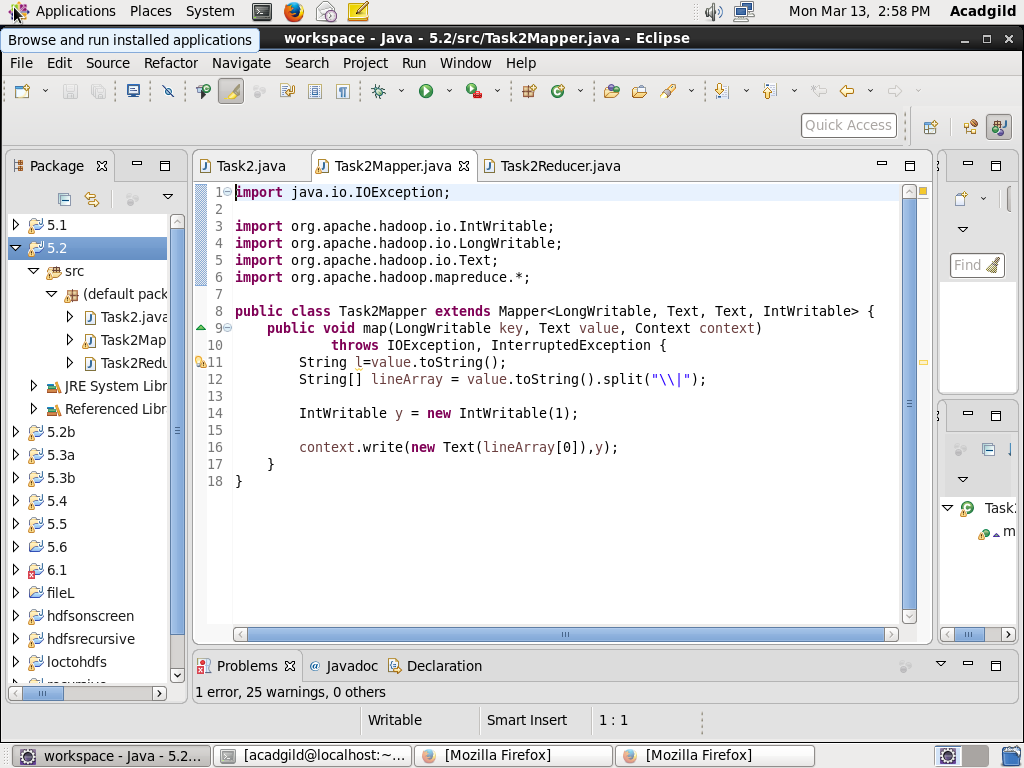
Samsung|Super|14|Maharashtra|619082|9200

Samsung|Super|14|Maharashtra|619082|9200

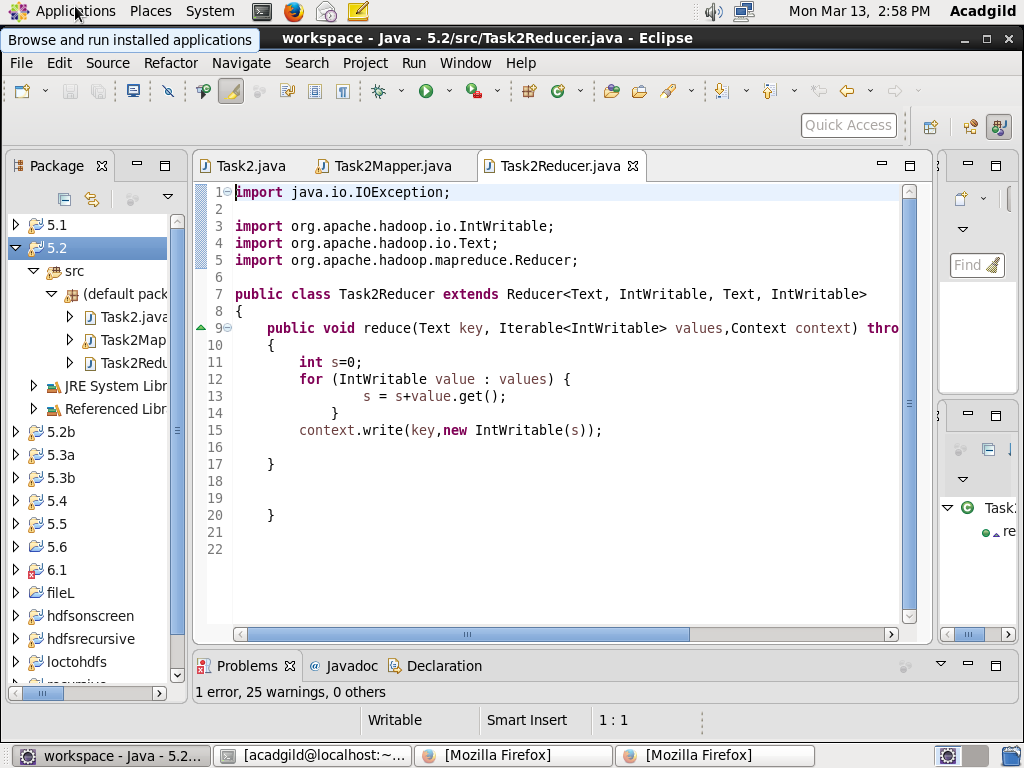
Driver Class:



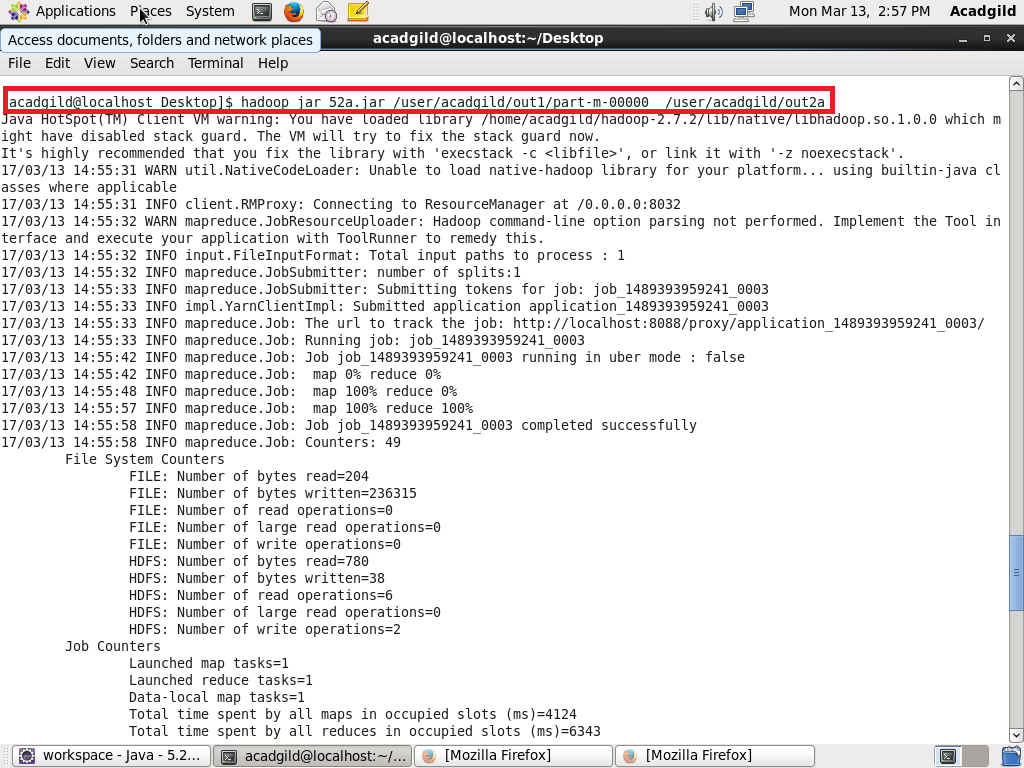
Mapper Class:

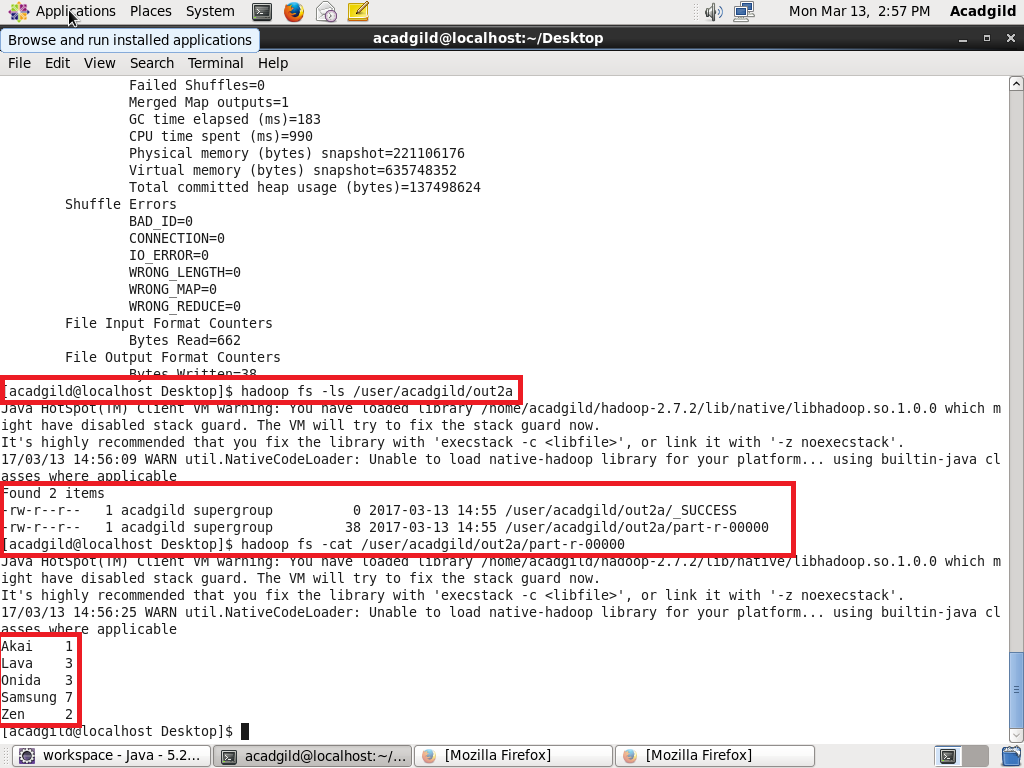


Reducer Class:



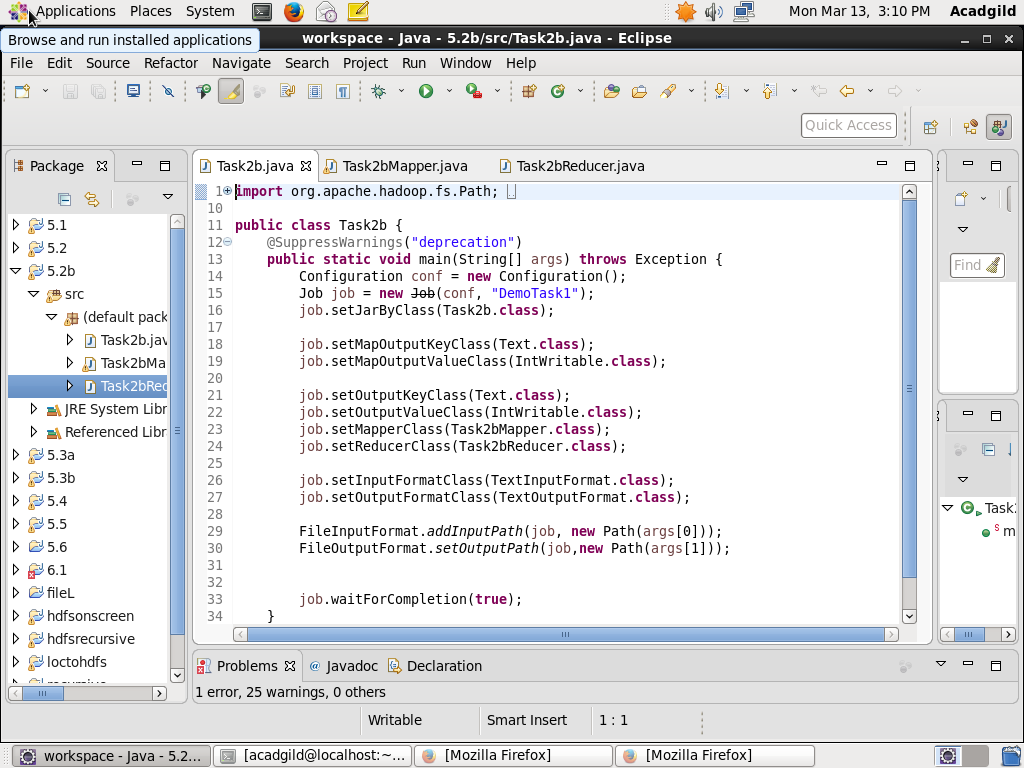
Output:



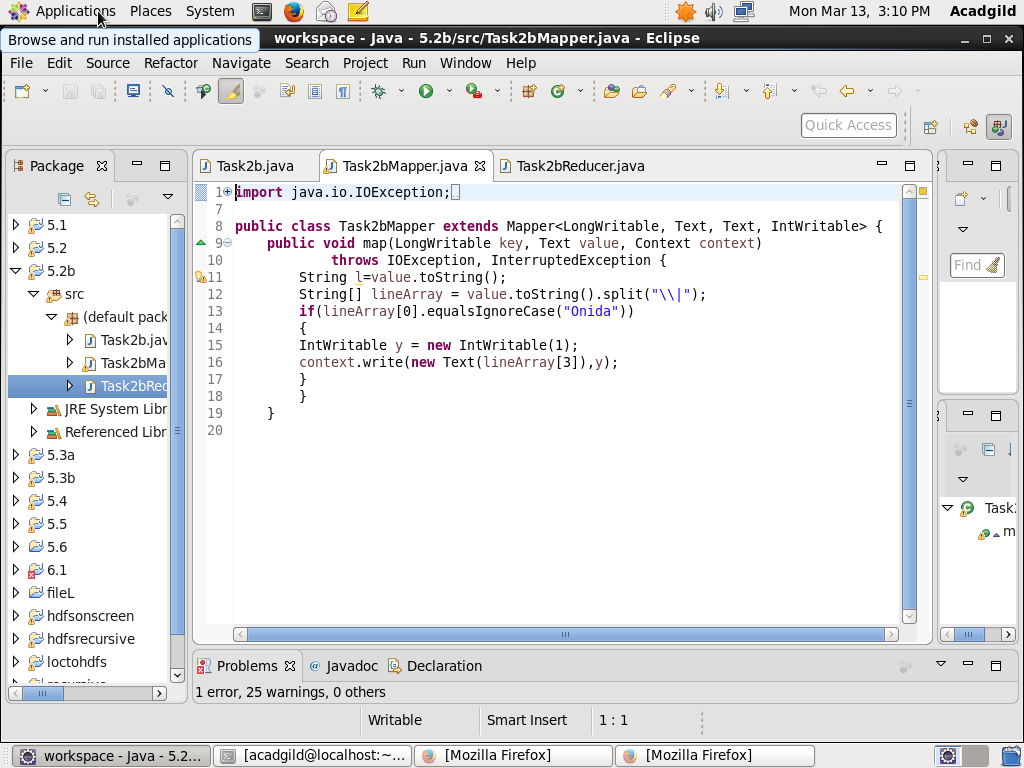


Sales of different TV Task 3: Write a MapReduce program to calculate the total units sold in each state for Onida company.

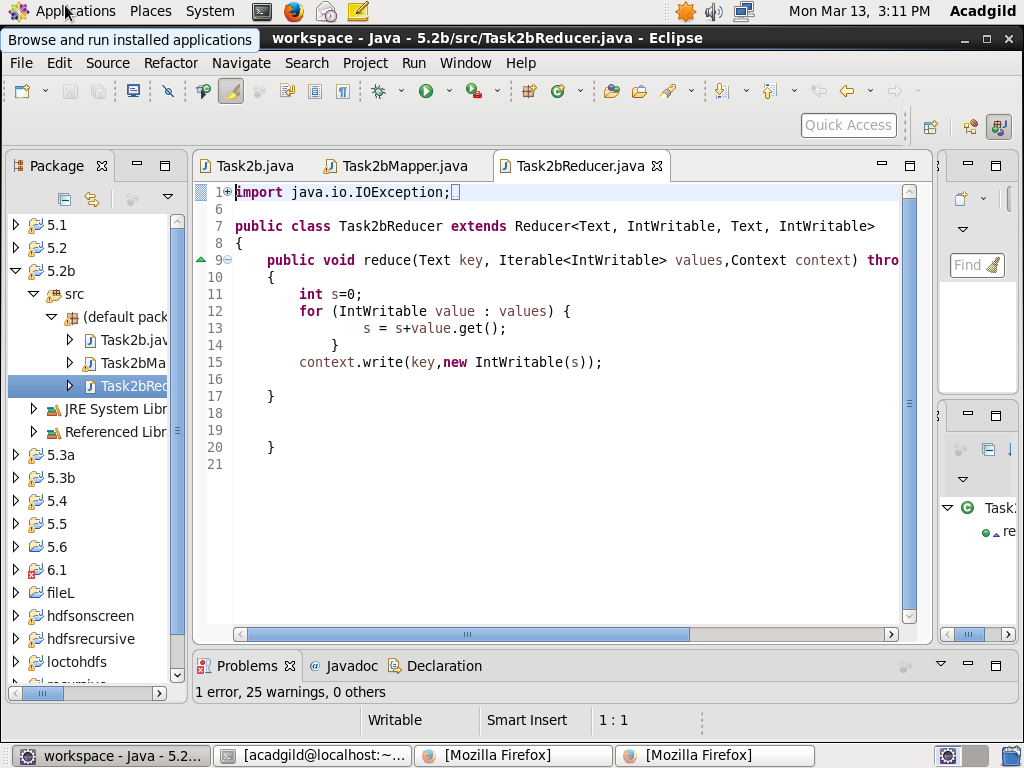
Driver Class:



Mapper class:



Reducer Class:



Output:

