Assignment 4.1

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

There are some invalid records which contain 'NA' in either Company Name or Product Name.

Task 1: Write a Map Reduce program to filter out the invalid records. Map only job will fit for this context.

Uploaded the Driver(TvSurveyTask1.java), Mapper(TvSurveyTask1Mapper.java) and reducer(TvSurveyTask1Reducer..java) files separately along with code explanation.

Output:

Task 2: Write a Map Reduce program to calculate the total units sold for each Company.

Uploaded the Driver(TvSurveyTask2.java), Mapper(TvSurveyTask2Mapper.java) and reducer(TvSurveyTask2Reducer..java) files separately along with code explanation.

Output:

Task 3: Write a Map Reduce program to calculate the total units sold in each state for

Uploaded the Driver(TvSurveyTask3.java), Mapper(TvSurveyTask3Mapper.java) and reducer(TvSurveyTask3Reducer..java) files separately along with code explanation.

Output:

```
[acadgild@localhost ~]$ hadoop fs -cat /TvSurveyTask3/part-r-00000
18/05/11 11:38:54 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classe
s where applicable
Kerala  1
Uttar Pradesh   3
[acadgild@localhost ~]$
```