# **Session 4: MR -INTRODUCTION**

## **Assignment 4**

**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

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.

**Task 2:**

* Write a Map Reduce program to calculate the total units sold for each Company.

**Task 3:**

* Write a Map Reduce program to calculate the total units sold in each state for Onida

company.