



ACADGILD

Session 5: Getting Started with MapReduce

Assignment 2

Session 5: Assignment 2

Table of Contents

- [1. Introduction](#)
- [2. Problem Statement](#)
- [3. Output](#)

1. Introduction

This assignment will help you to consolidate the concepts learnt in the session .

2. 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.txt

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.

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

3. Output

N/A

