**Big data Analytics Mini Project**

**Simulating MapReduce for Customer Purchase Pattern Analysis: A Management Perspective**

**🎯 Objective:**

To simulate the working of MapReduce using a sample retail transaction dataset to derive meaningful business insights. This project helps you apply your theoretical knowledge of MapReduce to a practical business problem using tools like Excel or Google Sheets.

**🔶 Project Instructions (Follow step-by-step):**

**Step 1: Understand the Business Scenario**

Assume you are a data analyst in a retail/e-commerce company. You are provided with a **customer transaction dataset**. Your job is to analyze the data using MapReduce principles to:

* Identify the **top 5 selling products**
* Determine **revenue generated by each product category**
* Find the **most frequent buyers**

**Step 2: Dataset Preparation**

Use the sample dataset provided OR download any of the following:

* [Online Retail Dataset (UCI)](https://archive.ics.uci.edu/ml/datasets/online+retail)
* Kaggle E-commerce Dataset

Ensure your dataset includes at least the following columns:

* TransactionID
* CustomerID
* Product
* Category
* Quantity
* Price
* Date

📌 **Note:** If data is not provided, simulate at least **50 records** manually in Excel.

**Step 3: Simulate the Map Phase**

In a separate sheet or table:

* Create **key-value pairs** for:
  + Product → Quantity
  + Category → Revenue (Quantity × Price)
  + CustomerID → Frequency

Example:

javascript

CopyEdit

Map Output for Products:

(Product1, 2), (Product2, 1), (Product1, 3)

**Step 4: Simulate the Reduce Phase**

Now, **aggregate the results**:

* For each product, **sum quantities**
* For each category, **sum revenues**
* For each customer, **count total purchases**

**Step 5: Visual Representation (Optional but encouraged)**

Create at least **2 charts**:

* Bar chart of **top 5 products by quantity sold**
* Pie chart of **revenue by category**
* Bar chart of **top 5 customers by frequency**

Use Excel or any visualization tool (Tableau, Power BI).

**Step 6: Report Writing**