SmartMart Data Analysis Report

# Part – 1 (Done by – Contribution of the whole team as per requirement)

# Data Wrangling

## Overview

This part of the project focused on preparing and cleaning the raw datasets (customers, products, stores, sales, and stock levels) before loading them into the database and visualization tools.

## Key Features

• Removed duplicates and inconsistent entries across datasets.  
• Handled missing values (e.g., unknown gender, empty city names).  
• Standardized column names and formats for smooth integration.  
• Ensured consistency of IDs (customer\_id, product\_id, store\_id) across all datasets.  
• Exported clean versions of all datasets for SQL and Power BI use.  
• Organised and cleaned files in a dedicated project folder.

## Purpose

Data wrangling ensures that all datasets are reliable, consistent, and analysis-ready. This step improves the accuracy of SQL queries, Python modelling, and Power BI dashboards, supporting high-quality business insights.

# Part – 2 (Done By – Aniket Ramesh Patil)

# Python Analyst

## Overview

This part of the project focused on building and managing the SmartMart database using Jupyter. It involved setting up tables, ensuring data integrity, and performing in-depth analysis on sales, products, customers, and stock levels.

## Key Features

• Created a database schema with proper joins between tables and created a combined worksheet.  
• Imported cleaned datasets (customers, products, stores, sales, stock levels).  
• Removed duplicates and handled missing values using duplicated.sum() and dropna(), and converted into proper data types.  
• Analyzed sales by product, customer, and store.  
• Applied the Apriori Algorithm and Association Rule for data mining.  
• Identified items, products, and categories for optimization.  
• Found product pairs frequently bought together.

## Purpose

The Python analysis ensures clean, structured data and delivers actionable insights into sales performance, customer behavior, and inventory needs, supporting both business strategy and Jupyter Python analyst reporting.

## Project conclusion

The Smart Mart analysis project successfully leveraged the Apriori algorithm and association rule mining to transform raw transactional data into actionable business intelligence. By uncovering hidden purchasing patterns, the project provided a clear, data-driven foundation for strategic decision-making, enabling the business to move beyond general assumptions and toward targeted, effective retail strategies.

# Part – 3 (Done By – Amarjit Kumar Natraj)

# SQL Analysis

## Overview

This part of the project focused on building and managing the SmartMart database in PostgreSQL. It involved setting up tables, ensuring data integrity, and performing in-depth analysis on sales, products, customers, and stock levels.

## Key Features

• Created database schema with proper primary and foreign keys.  
• Imported cleaned datasets (customers, products, stores, sales, stock levels).  
• Removed duplicates and handled missing values using SQL queries.  
• Analyzed sales by product, customer, and store.  
• Identified low-stock products for optimization.  
• Found product pairs frequently bought together.

## Purpose

The SQL analysis ensures clean, structured data and delivers actionable insights into sales performance, customer behavior, and inventory needs, supporting both business strategy and Power BI reporting.

# Part – 4 (Done By – Debashish Parida)

# Revenue Tracker Report (Dashboard)

## Overall Performance

- Revenue: 2.51M  
- Profit: 657.46K  
- Total Customers: 1171  
- Total Transactions: 4462  
- Average Order Value (AOV): 562.50  
  
The company has demonstrated strong revenue growth, with a healthy balance between revenue and profit. The AOV indicates consistent spending patterns across transactions.

## Customer Insights

- By City: Lucknow and Bangalore contribute the highest number of customers, followed by Kolkata and Pune. These cities are the strongest markets, and targeted campaigns can boost loyalty and retention.  
- By Gender: The customer base is evenly distributed across Male, Female, and Other, each holding around one-third share. This indicates a diverse and balanced demographic.  
- By Loyalty: Customers are nearly evenly split between loyalty program members and non-members. This shows strong adoption potential, and more promotions can expand the loyalty base.  
- By Age: Middle-aged customers form the majority of the base, showing higher purchasing activity. Young customers contribute less, indicating room for growth through targeted marketing.

## Revenue Insights

- City Revenue: Metro and tier-1 cities in India dominate overall revenue contribution. Locations like Lucknow, Bangalore, and southern cities are key revenue drivers.  
- Store Revenue: A few stores generate significantly higher revenue, highlighting the importance of scaling successful store models to other areas.  
- Revenue Trends: Revenue has steadily increased with minor fluctuations, reflecting consistent growth. Peaks indicate potential seasonal demand or campaign effectiveness.

## Category Performance

- Revenue & Profit by Category: Dairy, Bakery, and Frozen products generate the strongest profits. Meat, Beverages, and Snacks also contribute substantially but with comparatively lower margins.  
- Treemap View: Dairy and Frozen categories dominate revenue share, followed by Bakery, Meat, Beverages, Produce, and Snacks.

## Product-Level Insights

- Top Revenue Generators: Product\_44, Product\_26, Product\_50, and Product\_39 lead in revenue generation.  
- Top Quantity Sellers: Product\_40, Product\_22, Product\_15, Product\_3, and Product\_9 sell in large volumes, though not always aligned with revenue contribution.

## Recommendations

Strengthen loyalty program adoption with targeted rewards in top cities.  
- Expand high-performing store formats to underperforming regions.  
- Focus on Dairy, Frozen, and Bakery categories for profit maximization.  
- Launch cross-sell campaigns combining high-demand products with high-margin items.  
- Monitor seasonal spikes to align marketing campaigns and inventory planning.

## Conclusion

The company is performing strongly with solid revenue growth, balanced customer demographics, and profitable categories. By expanding loyalty, leveraging high-margin categories, and optimizing product strategies, the business can significantly enhance profitability and sustain long-term growth.