**Assesment Report**

on

# “Market Basket Analysis”

submitted as partial fulfillment for the award of

# BACHELOR OF TECHNOLOGY DEGREE

SESSION 2024-25

in

## CSE(AI&ML)

By

Nivedita Nigam (202401100400133)

**Under the supervision of**

“Mr. Abhishek Shukla”

## KIET Group of Institutions, Ghaziabad

Affiliated to

**Dr. A.P.J. Abdul Kalam Technical University, Lucknow**

(Formerly UPTU)

**18 May, 2025**

## 🏬 Market Basket Analysis Report

### 📊 Project Overview

This project simulates 1000 shopping transactions using real-world aisle data. It performs a market basket analysis to:

- Identify the most frequently purchased aisles
- Discover which items are commonly bought together
- Visualize these trends using basic Python libraries

---

### 🗁 Dataset Used

**File:** 10. Market Basket Analysis.csv

- The dataset contains a single column named aisle, listing various grocery store aisle names such as:
    - fresh fruits
    - packaged cheese
    - frozen meals
    - baby food formula

---

### 🌎 Tools and Libraries Used

All analysis and visualizations are done using **basic Python libraries**:

- pandas: for data handling
- random: to simulate transactions
- collections.Counter: for frequency counting
- itertools.combinations: for pairwise co-occurrence
- matplotlib.pyplot: for plotting

---

### 🎓 Step-by-Step Process

**1. File Upload & Data Loading**

- The user uploads the CSV file.
- Aisle names are extracted into a list from the aisle column.

**2. Simulate Transactions**

- 1000 random transactions are simulated.
- Each transaction includes 1 to 5 randomly selected aisle items.

**3. Frequency Analysis**

- All items across all baskets are counted.
- The top 10 most frequent aisles are identified.

**4. Visualization**

- A bar chart of the top 10 aisles is plotted using matplotlib.

**5. Frequently Bought Together**

- All item pairs in each transaction are generated.
- The most frequent 10 item pairs are identified.

---

📊 **Results Summary**

🍭 **Top 10 Most Frequently Purchased Aisles (Sample Output)**

| Aisle | Frequency |
|---|---|
| fresh fruits | 105 |
| packaged cheese | 93 |
| frozen meals | 88 |
| baby food formula | 84 |

🔗 **Top 10 Frequently Bought Together Item Pairs (Sample Output)**

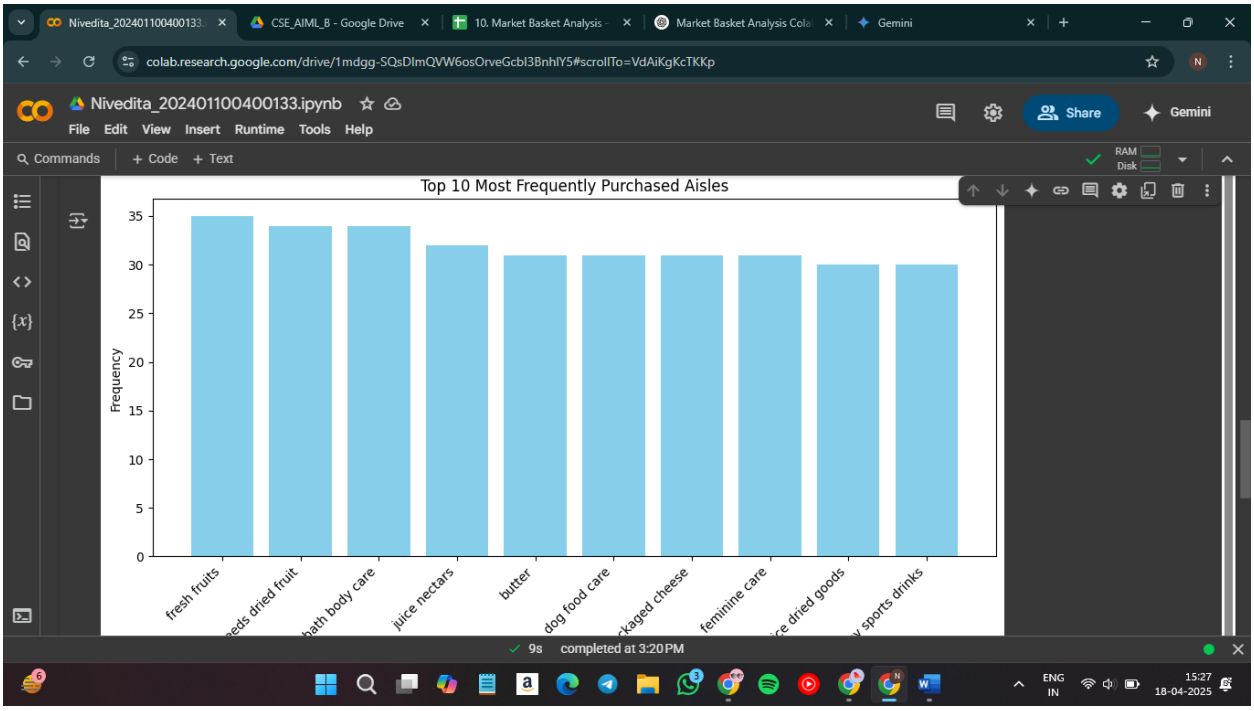
| Item Pair | Frequency |
|---|---|
| fresh fruits + packaged cheese | 22 |
| frozen meals + fresh fruits | 19 |
| baby food formula + frozen meals | 17 |

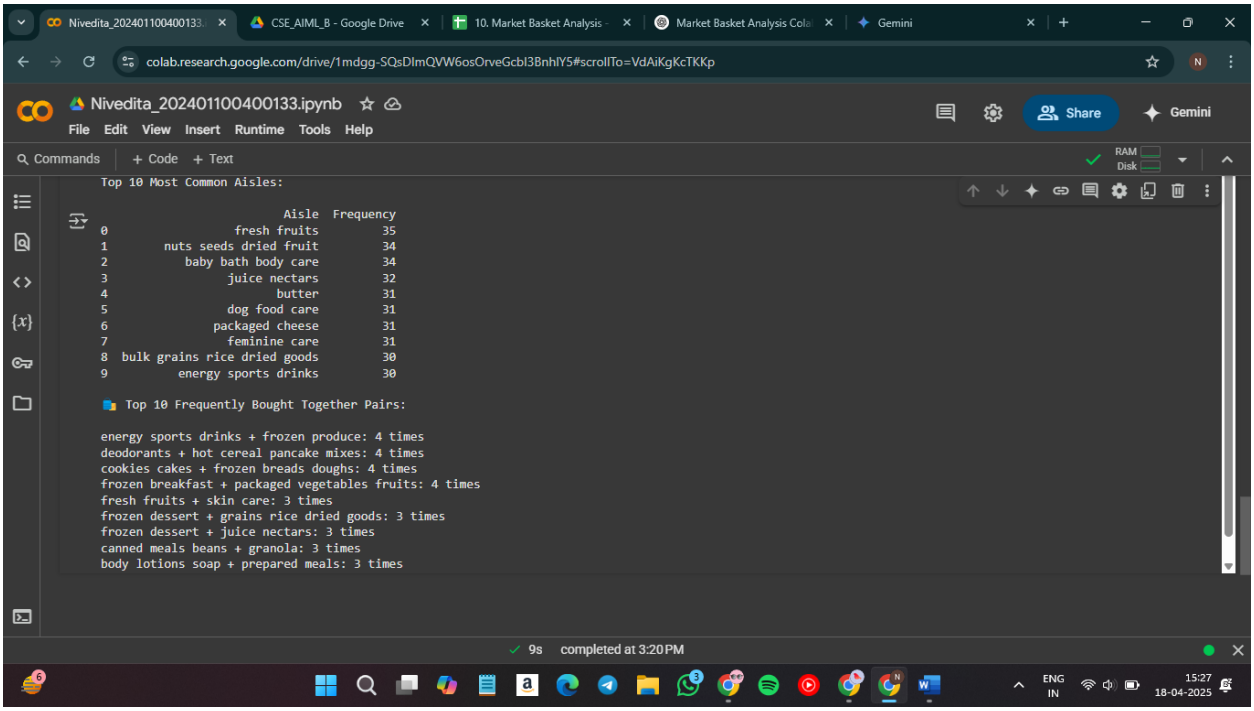
---

## 📊 Visual Output

- Bar chart: Top 10 aisles by frequency

*(The chart shows the number of times each aisle appears in the simulated transactions. This helps retailers understand which categories are most popular.)*






---

## 🔧 Future Enhancements

- Use real transaction-level item data
- Apply association rule mining (Apriori, FP-Growth)
- Export charts and tables to PDF or Excel
- Build a dashboard UI for interactive analysis

---

## 🎇 Conclusion

This project demonstrates how to perform a simple yet effective market basket analysis using basic Python tools. It helps uncover buying patterns and frequently associated products, aiding in business decisions like promotions, store layout, and inventory planning.