

KHULNA UNIVERSITY

190205 || 200240

Groceries Dataset Analysis Report

Submitted to:

Dr.Amit Kumar Mondal

Associate Professor

Computer Science And Engineering Discipline, Khulna University

Submitted By:

Sumaiah Binta Musa, istyaque Ahammed

Student ID:190205.200240

Computer Science And Engineering Discipline,Khulna University

# 1. Domain

The domain of this dataset is Retail and Grocery Transactions. It consists of transaction records from a grocery store where each row contains items purchased together by customers. This dataset can be used for market basket analysis, customer behaviour insights, and more.

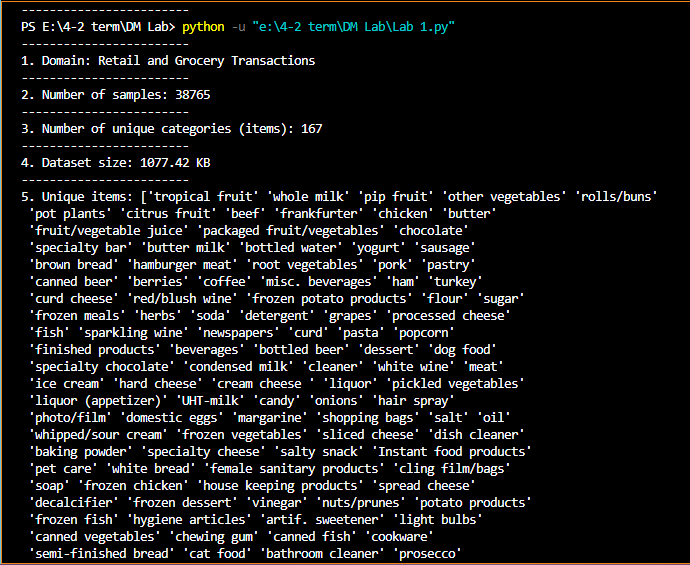
# 2. Number of Samples

The dataset contains 38,765 samples. Each sample represents a unique transaction in which one or more items were purchased.

# 3. Number of Unique Categories (Items)

The dataset contains 167 unique grocery items. These items represent the different products available for purchase in the grocery store.

# 4. Dataset Size

The dataset is 1077.42 KB in size, making it relatively small and easy to work with in terms of both data storage and processing requirements.

# 5. Dataset Value

The intellectual value of this dataset lies in its ability to offer insights into customer purchasing behaviour and product relationships. Some key uses include:

### **Market Basket Analysis**

* **Explanation**: This involves analyzing what products are often purchased together.
* **Example**: If the data shows that customers frequently buy bread and butter together, a retailer can place these items near each other in the store or create combo deals to increase sales.
* **Intellectual Value**: This helps retailers understand product relationships and optimize store layouts, leading to increased sales.

### **Customer Segmentation**

* **Explanation**: Analyzing buying patterns allows businesses to group customers into different segments based on their preferences or habits.
* **Example**: If the data shows that young adults frequently buy snacks and soft drinks, while families buy fruits and vegetables, the retailer can create promotions targeted at each group.
* **Intellectual Value**: This allows for personalized marketing and promotions, increasing the chances of making a sale

### **Pricing Strategy**

* **Explanation**: The dataset can show which items are frequently purchased, helping the retailer create effective pricing strategies.
* **Example**: If a certain brand of cereal is often bought without discounts, the store may not need to offer discounts on that product, but they could offer discounts on less popular items to boost sales.
* **Intellectual Value**: Knowing what customers will buy regardless of price helps retailers design discount strategies more effectively, maximizing profit.

### **Fraud Detection**

* **Explanation**: Unusual purchase patterns can sometimes indicate fraudulent activity.
* **Example**: If a customer who typically buys small quantities suddenly buys a large number of expensive items, this could trigger a check for fraud.
* **Intellectual Value**: By spotting irregular buying patterns, businesses can reduce the chances of fraud, saving money and improving security.