

STRATIFY

Business Analytics Bootcamp - 2025

Project : Decoding the Basket

QuickBasket Basket Breakdown

Background:

QuickBasket is a leading online grocery delivery platform that partners with multiple local supermarkets. One of their largest partners, FreshMart, wants to improve its online sales performance through data-driven strategies on QuickBasket.

FreshMart has provided anonymized historical order data collected via GroceryKart's app and website. The dataset contains millions of customer orders, including product details, order timings, and purchase patterns.

Business Problem:

FreshMart wants actionable insights to optimize:

- Inventory management
- Product recommendations
- Promotional strategies
- Online content placement (e.g., featured products, app banners)

By analyzing past customer buying behavior, the goal is to answer the following key questions-

1. What are the most popular products on GroceryKart?
2. Which products are reordered most frequently?
3. When do customers place the most orders?
4. What combinations of products are frequently bought together? (Market Basket Analysis)

5. How can FreshMart design product bundles, promotional campaigns, or content placements based on customer shopping behavior?

Dataset Overview:

File	Description
<code>order_metadata.csv</code>	All customer orders with timestamps and user IDs
<code>order_items_prior.csv</code>	Products purchased in each prior order
<code>order_items_train.csv</code>	Products in each training set order
<code>product_catalog.csv</code>	Product details (product names, IDs, department, aisle)
<code>department_info.csv</code>	Department names for each department ID
<code>aisle_info.csv</code>	Aisle names for each aisle ID

https://drive.google.com/drive/folders/19v0lNW9QnQhvNDgb9gYOnj_qR7FRQQ?usp=sharing

Task 1: Exploratory Data Analysis

This task will help you understand customer purchasing behavior, product popularity, and temporal ordering patterns, which helps in optimizing inventory, personalizing marketing strategies, and improving customer retention through better product recommendations and timing of promotions.:

1. Total number of orders and unique customers
2. Number of unique products, aisles, and departments
3. Top 20 most frequently ordered products
4. Top departments and aisles by order volume
5. Identify most reordered products (products with highest reorder rate)
6. Average basket size (number of items per order)

7. Order trends by hour of the day, day of the week, and days since prior order

Task 2: Understanding Customer Buying Patterns using Product Combination Analysis

The goal here is to answer the question:

What products do customers frequently buy together in the same order?

You will:

- Create a table that shows which products appear together in the same shopping cart (order).
- Identify popular product combinations, like-
 - Customers who buy bread often buy milk.
 - People buying coffee often buy sugar and creamer.
- Learn how often these combinations occur in the dataset.

Task 3: Business Recommendations

Based on inference collected from the data, provide the following:

1. Inventory suggestions for high-demand products
2. Ideas for product bundling (based on association rules)
3. Suggested times for sending promotional notifications (based on peak ordering times)
4. Recommendations for featured products in the app homepage or banners

Optional

Complete the following tasks for extra credit:

1. Predict the number of items in the next customer order (basket size prediction).
2. Visualize association rules using network graphs.
3. Execute the segmentation of customers based on ordering behavior.

Tools & Technologies Required:

- **Python** (Preferred IDE: Jupyter Notebook, Google Colab, or VS Code)
- **Libraries:**
 - a. **pandas, numpy** – For efficient data wrangling, transformation, and numerical computation.

- b. **matplotlib, seaborn, plotly** – For creating static, statistical, and interactive visualizations.
 - c. **mlxtend** – For implementing frequent itemset mining and association rule learning.
 - d. **networkx or graphviz** – For visualizing relationships and patterns using network or graph structures.
-

Expected Final Deliverables:

1. Submit a Jupyter Notebook or Google Colab notebook with the following in a single file:
 - EDA findings
 - Plots and visualizations
 - QuickBasket Analysis results
 - Association rules table
 - Business recommendations
 2. Create a presentation slide deck (PowerPoint / Google Slides) summarizing:
 - The business problem
 - Key insights
 - Visualizations
 - Final recommendations for FreshMart management
-