Market basket Analysis

Association Rule Mining

Stephanie Cheng
January 10, 2025

EXECUTIVE SUMMARY

This report presents the results of an Association Rule Mining analysis conducted on the dataset "bread_basket.csv" with using R programming.

CONTENTS

ASSOCIATION RULE MINING

Introduction

Purpose

Association Rule Mining (ARM) is analyzing purchasing patterns and optimizing business strategies. This part shows the application of ARM to a bakery dataset to:

- Identify frequently purchased itemsets
- Generate association rules to uncover relationships between items
- Provide actionable insights for sales optimization

Dataset Description

The dataset "bread_basket.csv" contains transaction data from a bakery, including:
- Transaction ID: Unique identifier for each transaction
- Date and Time: The timestamp of when the transaction occurred
- Item: The product purchased

Methodology

Data Preparation

The dataset consists of **9,465 transactions** and **94 distinct items**. I transformed the dataset into a matrix format, where each row represents a transaction, and each column indicates whether an item was purchased. This matrix was then used for association rule mining, applying the **Apriori Algorithm**.

General Findings

Top 10 & Last 10 Purchased Items
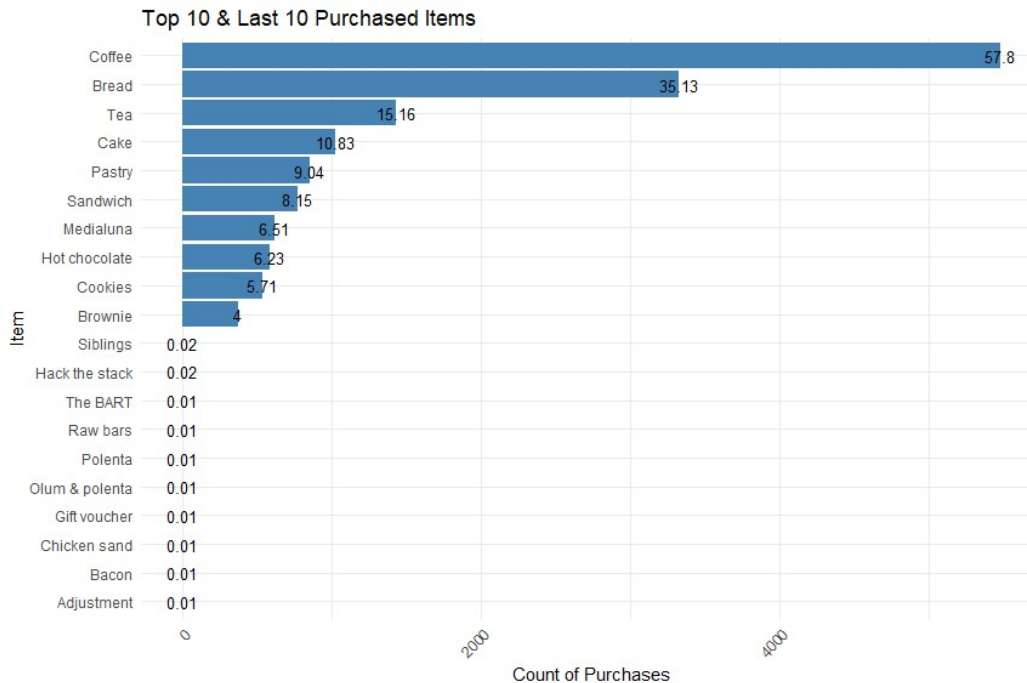
The **top 10 most purchased items** (visualized in a bar chart along with their frequencies) are:
Coffee, Bread, Tea, Cake, Pastry, Sandwich, Medialuna, Hot Chocolate, Cookies, Brownie.

The **10 least purchased items** are:
Siblings, Hack the Stack, Chicken Sand, The BART, Adjustment, Bacon, Olum & Polenta, Polenta, Gift Voucher, Raw Bars.

Association rule finding

The **count** in the rules represents the number of transactions supporting a given association, highlighting frequent purchasing patterns.

The dataset was filtered based on:

- **Support > 0.04** – Association in at least **4% of transactions** are considered

- **Lift > 1** – Indicates a **stronger-than-random** relationship, removing weak associations

R Output

lhs	rhs	support	confidence	coverage	lift	count
{Pastry} => {Coffee}		0.04754358	0.55214724	0.08610671	1.154168	450
{Coffee} => {Pastry}		0.04754358	0.09938163	0.47839408	1.154168	450
{Cake} => {Coffee}		0.05472795	0.52695829	0.10385631	1.101515	518
{Coffee} => {Cake}		0.05472795	0.11439929	0.47839408	1.101515	518

Data Visualization

Conclusion

Customers who purchase **Pastry** have a **55.21%** chance of also buying **Coffee**. Similarly, customers who purchase **Cake** have a **52.70%** chance of also buying **Coffee**. These findings suggest that **Coffee is a popular complementary item for both Pastry and Cake**, making it a strong candidate for bundled promotions.

Customers who purchase **Coffee** have a **9.94%** chance of also buying **Pastry**. Similarly, customers who purchase **Coffee** have an **11.44%** chance of also buying **Cake**. While these percentages are lower than the previous direction of association, they still indicate that **Pastry and Cake are frequently bought alongside Coffee**.

Recommendations

Given that coffee is frequently purchased alongside pastries and cakes, the bakery could implement **bundle discounts or combo deals** to encourage customers to purchase these items together. Such promotions can enhance customer satisfaction while increasing the overall sales volume.

Strategic **product placement** can also optimize sales. Positioning **pastries and cakes near the coffee counter** can increase visibility and encourage customers to make additional purchases.

Targeted promotional strategies may also prove beneficial in boosting sales. For instance, implementing a **"Receive 10% off Coffee while purchase Pastry"** campaign could incentivize cross-purchasing.

Furthermore, adjustments to the pricing and inventory management strategy could improve overall profitability. Given that the least purchased items exhibit lower demand, their **prices could be slightly increased** to maximize revenue from occasional buyers. Additionally, the bakery may consider **reducing import quantities of underperforming items** and reallocating resources toward **increasing the supply of top-selling products** or investing in targeted advertising campaigns to attract more customers.