

**Project methodology:** Our project follows the CRISP-DM framework, progressing through key phases for a structured approach. We start by understanding the business problem and identifying objectives. Next, we translate the business problem into a data problem. After that, we prepare the data using SQL to integrate, clean, and analyze datasets. Once the data is ready, we conduct further analysis, applying techniques like EDA, feature engineering. Then, we build and evaluate machine learning model. Finally, we deploy the model. We utilize Streamlit for deploying the trained model, providing an interactive interface for real-time predictions and analysis.

## 1. Business Understanding

We are working with a large offline retailer that sells Fast Moving Consumer Goods (FMCGs) such as pretzels, frozen pizza, boxed cereal, and mouthwash. The retailer acquires products from various manufacturers, stores them in inventory, and supplies them to its stores based on customer demand.

**Business Problem:** The retailer is experiencing **high business losses**.

**Objective:** Reduce business losses by either increasing revenue or cutting costs.

To achieve this, we analyze the formula:

**Profit/Loss = Revenue - Cost**

Both revenue has **decreased** and costs have **increased**, contributing to the losses. Therefore, we need to break down the factors driving both revenue and cost to identify optimization opportunities:

- **Revenue Breakdown:**

Revenue = (#products\_sold) \* (average\_product\_price)

Revenue = (#customers) \* (average\_products\_bought) \* (average\_product\_price)

Revenue = (#old\_customers + #new\_customers) \* (average\_products\_bought) \* (average\_product\_price)

- **Cost Breakdown:**

Cost = (#products\_bought) \* (average\_buying\_price) + (operational\_cost)

Cost includes:

- **Inventory cost** (storage, handling, etc.)
- **Manpower cost** (staffing, wages)
- **Promotional cost** (marketing, campaigns)
- **Other operational costs** (logistics, utilities)

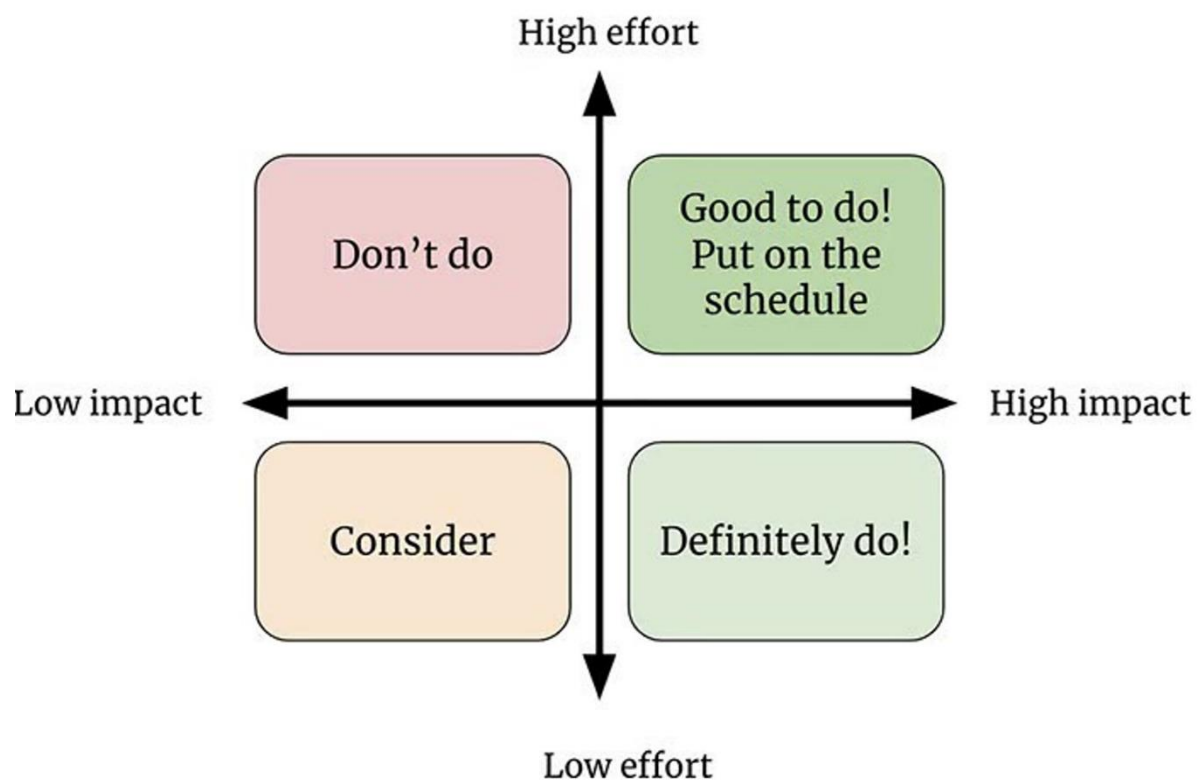
**Revenue Optimization Focus Areas:**

- **Increase the number of old customers** through email targeting, in-store experience improvement, and optimizing wait times.
- **Increase the number of new customers** by optimizing online distribution, customer acquisition strategies, and exploring social media channels.
- **Increase average products bought** by preventing stockouts, optimizing store layouts, and running effective promotions.
- **Optimize average product price** through better pricing strategies, region-based pricing, and maximizing profit margins.

#### Cost Reduction Focus Areas:

- **Reduce buying price** by negotiating better deals with manufacturers and exploring alternative suppliers.
- **Optimize promotional costs** by cutting spending on expensive channels and exploring cost-effective alternatives.
- **Lower operational costs** by optimizing logistics, reducing leakages, and investing in automation.

To prioritize these initiatives, we apply the **Impact-Effort Framework** (as shown in the diagram provided). After applying this, we selected the following key business objective:



**Business Objective Selected:**

**Prevent overstocking and understocking** of goods to reduce costs.

- **Overstocking:** When availability exceeds demand, resulting in excessive storage costs.
- **Understocking:** When availability falls short of demand, leading to lost sales.

Both situations increase the retailer's operational costs and negatively impact revenue, leading to higher losses.

## **2. Converting Business Problem to Data Problem**

**Data Problem:**

To address this business issue, we will focus on **predicting the demand for packed items** in each store for the next week. This will allow the retailer to better manage inventory, ensuring optimal stock levels that match demand and reducing both overstocking and understocking.

This solution will optimize the retailer's supply chain, reduce unnecessary costs, and improve profitability by aligning inventory with actual demand.