Optimizing Online Retail Sales Using Data-Driven Insights

A Proposal report for the BDM capstone Project

Submitted by
Name: Mohit Tewari
Roll number:23f1002364

IITM Online BS Degree Program,
Indian Institute of Technology, Madras, Chennai
Tamil Nadu, India, 600036

Contents

1. Executive Summary and Title
2. Organization Background
3. Problem Statement
4. Background of the Problem
5. Problem Solving Approach
6. Expected Timeline
7. Expected Outcome

Declaration Statement

I am working on a Project titled “Optimizing Online Retail Sales Using Data-Driven Insights”. I extend my appreciation to **UCI Machine Learning Repository** for providing the necessary resources that enabled me to conduct my project.
I hereby assert that the data presented and assessed in this project report is genuine and precise to the utmost extent of my knowledge and capabilities. The data has been gathered through primary sources and carefully analysed to assure its reliability.

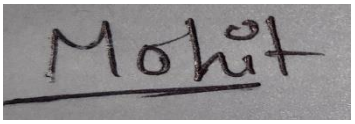
Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through thorough analytical procedures.

I am dedicated to adhering to the information of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. In the event that plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I agree that all the recommendations are business-specific and limited to this project exclusively, and cannot be utilized for any other purpose with an IIT Madras tag. I understand that IIT Madras does not endorse this.

Signature of Candidate:



Name: Mohit Tewari
Date: 07/03/2025

1. Executive Summary and Title

The UK-based online retail company specializes in selling unique homeware and gift items to both individual customers (B2C) and wholesalers (B2B). The company operates globally, with a diverse customer base purchasing through its e-commerce platform.

The online retail business faces challenges related to in demand forecasting, product performance evaluation, and return management. Unpredictable sales trends lead to inventory issues, while a significant portion of revenue comes from a small percentage of products. Additionally, frequent product returns impact profitability, requiring a deeper understanding of return patterns.

To address these challenges, a **data-driven approach** will be employed. Time-series forecasting models will be used to predict demand, helping optimize stock levels. Pareto analysis will identify high-performing products, enabling better sales and marketing strategies. Lastly, return analysis will highlight common return and problematic products, leading to improved quality control and reduced losses. By leveraging these insights, the company can enhance operational efficiency, minimize costs, and improve customer satisfaction.

2. Organizational Background

The company is a UK-based online retail business specializing in the sale of homeware, gift items, and decorative products. It operates through an e-commerce platform, serving both individual customers (B2C) and wholesalers (B2B). The company has a strong international presence, with a significant portion of its customers located across Europe.

Established as a niche retailer, the company has built its reputation by offering a curated selection of unique and aesthetically appealing products. Its business model heavily relies on bulk purchasing and seasonal sales, making inventory planning and demand forecasting crucial for its success.

Despite its growth, the company faces several operational challenges, such as demand unpredictability and high product return rates. To remain competitive and improve profitability, it aims to leverage **data-driven insights** to optimize inventory, identify top-performing products, and reduce losses associated with product returns.

3. Problem Statement

3.1 Demand Forecasting Challenges:

The company struggles with unpredictable sales trends, leading to stockouts or overstocking. A lack of accurate demand forecasting results in revenue loss and increased holding costs.

3.2 Identifying Key Revenue-Generating Products:

A small percentage of products contribute to the majority of sales, but the company lacks data-driven insights to prioritize these items. Understanding product performance using the Pareto Principle can improve sales strategies and inventory planning.

3.3 High Return Rates Impacting Profitability:
Frequent product returns create financial losses and affect customer satisfaction. Analysing return trends, identifying high-return products, and understanding return reasons can help reduce losses and improve product quality.

4. Background of the problem

The online retail company operates in a competitive e-commerce environment where accurate demand forecasting, efficient inventory management, and return reduction are crucial for profitability. However, the company faces significant challenges in these areas due to fluctuating customer demand, seasonal trends, and varying purchasing behaviours.

One of the primary issues is demand forecasting. The inability to predict sales accurately results in stockouts, leading to lost sales opportunities, or overstocking, which increases storage costs and leads to potential product obsolescence. This problem arises due to unpredictable market trends, changing customer preferences, and the lack of advanced predictive modelling techniques.

Another challenge is identifying high-performing products. While a small percentage of products contribute to the majority of revenue, the company lacks a structured approach to prioritize these items. Without proper analysis, resources may be allocated inefficiently, leading to poor marketing and stocking decisions.

The company also struggles with high return rates, affecting operational costs and customer satisfaction. While internal factors such as poor-quality control and inventory mismanagement contribute to the issue, external factors like fraudulent returns and logistical inefficiencies worsen the problem.
Addressing these challenges with data-driven insights will enhance efficiency and profitability.

5. Problem Solving Approach

To address the key business challenges identified in the **Online Retail dataset**, a structured **data-driven approach** will be implemented. This involves multiple analytical and machine learning techniques to extract insights and optimize decision-making. Python and python-libraries will be used and code will be in a google colab notebook for the analysis.

Step 1: Data Preprocessing & Exploration
Before any analysis, data cleaning and preprocessing will be performed:

- **Handling Missing Values:** Missing entries will be either removed or imputed if necessary.
- **Removing Outliers:** Unusual transactions will be analysed and possibly removed.
- **Adding Derived Features:** New columns (if necessary) from existing columns will be added for better insights and interpretability of the data.

Step 2: Pareto Analysis (80-20 Rule) for Customers & Products
The Pareto principle states that **80% of revenue comes from 20% of customers or products**. This will be analysed by:

- Calculating **cumulative revenue contribution** from different products and customers.

- Identifying **top-performing products** and **high-value customers** to focus marketing & inventory efforts.
- Determining **low-selling and frequently returned products** for potential discontinuation.

Step 3: Demand Forecasting

To predict future sales trends and optimize inventory, **time-series forecasting** will be applied:

- We will begin with **ARIMA**, as they are efficient for structured business data. Other models will be used if necessary.
- Sales forecasting will be conducted at **daily and monthly levels**, enabling better inventory planning and demand management.

Step 4: Return Analysis & Optimization

Returns cause revenue loss. The analysis will:

- Identify **products with high return rates**.
- Examine **customer behaviour** (which customers return the most?).
- Suggest **business strategies** to reduce returns, such as better-quality control and targeted discounts for high-return products.

Step 5: Customer Segmentation for Marketing Strategies

Using **clustering techniques (e.g. K-Means)**, customers will be grouped into segments based on:

- **Purchase frequency & spending behaviour**.
- **Likelihood to return products**.
- **Country & regional preferences**.
- This helps in **targeted marketing campaigns** and loyalty programs.

This analytical approach will provide a solid foundation for data-driven decision-making in the online retail business.

6. Expected Timeline

Gantt Chart and Work Breakdown Structure:

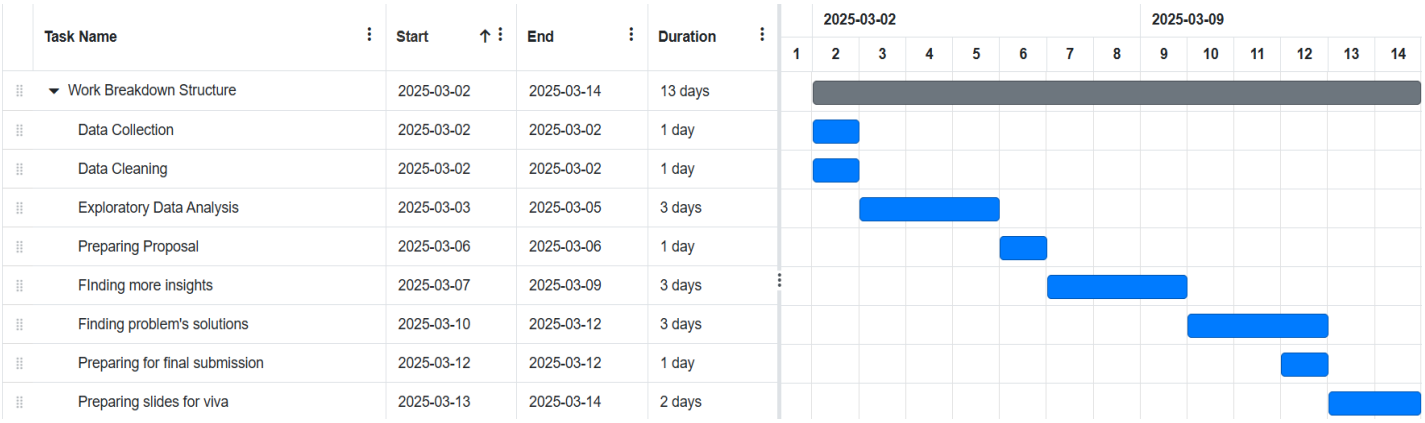


Figure 1. Expected timeline of the Project.

7. Expected Outcome

The analysis and forecasting techniques applied to the **Online Retail dataset** will provide valuable insights to improve business decision-making.

1. **Optimized Inventory Management:**
    - By forecasting demand accurately, the company can **prevent overstocking and stockouts**, ensuring that the right products are available when needed.
2. **Revenue Optimization and Business Growth:**
    - Pareto analysis will help identify **the top 20% of customers and products generating 80% of revenue**, allowing better resource allocation.
3. **Understanding Returns and Their Impact:**
    - A detailed **return analysis** will reveal which products have **high return rates**.
    - This will help in **reducing return-related losses** by improving product quality, refining logistics, and enhancing customer satisfaction.
4. **Customer Segmentation for Targeted Marketing:**
    - By analysing customer behaviour, we can identify **high-value customers**, **one-time buyers**, and **frequent returners**.
    - This enables personalized marketing strategies, promotions, and loyalty programs, increasing customer retention and profitability.

This data-driven approach will lead to **better financial performance, increased efficiency, and improved customer experience**.