**Elevate Labs Project**

**Project Title : E-commerce Rate Reduction Data Analysis Project**

**INTRODUCTION**

In the highly competitive landscape of e-commerce, customer satisfaction and operational efficiency are key drivers of business success. However, one of the persistent challenges faced by online retailers is the high rate of product returns, which can significantly impact profitability, inventory management, and customer loyalty.This analysis focuses on identifying the underlying causes of product returns, uncovering patterns and trends through data, and recommending actionable strategies to reduce return rates. By leveraging historical order data, return reasons, customer profiles, and product details, the goal is to gain a deeper understanding of what drives returns and how they can be minimized.Reducing return rates not only leads to cost savings but also improves the customer experience, enhances brand trust, and contributes to sustainable business practices. This project aims to empower decision-makers with insights that are both strategic and practical, leading to long-term business growth.

**ABSTRACT**

This project tackles the critical issue of high e-commerce return rates. By analyzing historical order and customer data, the analysis pinpoints the primary drivers of returns, including high-risk products and specific customer behaviors. The resulting insights inform data-driven strategies, such as predictive modeling and improved product descriptions, to significantly reduce returns, enhance efficiency, and boost customer satisfaction.

**TOOLS USED**

* Python,
* Power BI,
* SQL

**STEPS INVOLVED IN BUILDING THE PROJECT**

Here is a step-by-step explanation of the project :

**Step 1: Data Integration with SQL**

Combine data from orders, returns, products, and customers tables using a single SQL query. The key is to create a master dataset where every order is included, with a flag indicating if it was returned (1) or not (0).

**Step 2: Analysis and Insights in Python**

Load the data into Python (Pandas). Clean it by handling missing values and correcting data types. Then, perform Exploratory Data Analysis (EDA) to answer key questions:

* What are the top reasons for returns?
* Which product categories and suppliers have the highest return rates?
* Are there patterns based on customer location or marketing channel?

**Step 3: Predictive Modeling**

Using Scikit-learn in Python, build a Logistic Regression model to predict the probability of an item being returned. Use the trained model to assign a "Return Risk Score" to every order in your dataset.

**Step 4: Interactive Dashboard in Power BI**

Import the final dataset (including the risk scores) into Power BI. Create an interactive dashboard with key visuals:

* KPIs for overall return rate.
* Charts showing returns by category, reason, and location.
* A table displaying the products with the highest risk scores.

**Step 5: Final Deliverables**

Package the project outputs for stakeholders:

1. The complete and commented Python codebase.
2. The interactive Power BI dashboard file or web link.
3. A CSV file listing the high-risk products identified by the model.

**CONCLUSION**

This analysis identifies the root causes of product returns by pinpointing high-risk products, customer segments, and operational inefficiencies. Implementing the recommended data-driven strategies such as optimizing product descriptions, improving quality control, and using predictive models-significantly reduces return rates. This approach boosts profitability and efficiency while strengthening customer trust, leading to a more cost-effective and customer-friendly shopping experience**.**