

Project Report: E-commerce Return Rate Reduction Analysis

1. Introduction

The rise of e-commerce has provided convenience to customers but has also led to a growing challenge—product returns. High return rates impact business profitability and customer satisfaction. This project focuses on analyzing the factors driving returns and developing strategies to reduce return rates.

2. Abstract

This project analyzes product return behavior using order and return datasets. Key objectives include:

- Measuring return percentages across categories, suppliers, and geographies.
- Building a predictive model to estimate the probability of return using logistic regression.
- Creating an interactive Power BI dashboard that provides a return risk score, enabling businesses to identify and address high-risk products.

By integrating SQL, Python, and Power BI, this project delivers actionable insights to improve decision-making and reduce avoidable returns.

3. Tools Used

- **SQL** – For data extraction, cleaning, and aggregation.
- **Python** – For preprocessing, statistical analysis, and predictive modeling (Logistic Regression).
- **Power BI** – For building an interactive dashboard with drill-through filters and return risk scores.

4. Steps Involved in Building the Project

1. **Data Cleaning & Preparation:** Cleaned and structured the return and order datasets, removed duplicates, handled missing values, and standardized fields.
2. **Exploratory Data Analysis (EDA):** Calculated return % by product category, supplier, and geography. Identified top categories with higher-than-average return rates.
3. **Predictive Modeling:** Used Python and logistic regression to model the probability of return. Generated probability scores for each product to identify high-risk items.
4. **Dashboard Development:** Designed a Power BI dashboard with KPIs (Return %, Total Orders, Returned Orders), category-wise return %, supplier risk distribution, and return share by geography. Added drill-through filters for deeper insights.
5. **Deliverables:** Interactive Power BI dashboard, Python codebase for probability prediction, and CSV of high-risk products.

5. Conclusion

The project successfully demonstrates how e-commerce businesses can use data-driven methods to understand and reduce product return rates. By combining SQL, Python, and Power BI, the analysis provides a holistic view of return behavior. The predictive model helps identify risky products, while the dashboard supports decision-makers with actionable insights.

Outcome: Businesses can improve operational efficiency, reduce costs, and enhance customer satisfaction by proactively addressing return-related issues.