E-commerce Return Rate Reduction Analysis

Introduction

E-commerce platforms often face significant challenges due to high product return rates, which can lead to increased operational costs, customer dissatisfaction, and revenue losses. This project aims to analyze patterns across different categories, geographies, and marketing channels, and develop a predictive model to identify products with a high risk of being returned.

Abstract

This project involved analyzing return and order data using SQL and Python to identify trends and high-risk segments. A logistic regression model was built to predict return probabilities based on historical order attributes. Furthermore, an interactive dashboard was developed in Power BI to visualize return rates and provide a return risk score for products. The end goal was to help business teams proactively address return-prone products and optimize logistics and marketing strategies.

Tools Used

- Python: Data cleaning, EDA, logistic regression model
- Power BI: Dashboard creation and visualization
- Excel/CSV: Data exchange between tools

Steps Involved in Building the Project

1. Data Collection and Cleaning

Cleaned datasets: handled nulls, removed duplicates, standardized formats.

2. Exploratory Data Analysis (EDA)

Analyzed return rates by:

- Brand
- Screen Size
- Return rates vs Price
- RAM

4. Dashboard Creation

Developed a Power BI dashboard:

- Value by base color
- Count of RAM by return rates
- Count of ROM by return rates
- Sum of sales by brand
- Sales price by risk score
- Ratings by model
- Return rates by brand

5. Output Generation

- CSV file containing high-risk products with return probability scores.
- Python scripts saved for retraining model as new data comes in.
- Power BI dashboard published with user-level access.

▪ Conclusion

By visualization techniques, this project provided a scalable solution to reduce return rates. The model enables proactive risk mitigation, while the dashboard empowers stakeholders to monitor trends and take data-driven actions. Future improvements may include using more advanced models and integrating customer feedback data for deeper insights.