E-commerce Return Analysis Project Report

1. Project Overview

This project focuses on analyzing product returns in an e-commerce business.

The goal is to understand return patterns, identify high-risk products, and enable data-driven decisions through visualization and modeling.

2. Tools Used

- Python (Pandas, Scikit-learn)
- Power BI (Visualization)
- CSV Files (orders.csv, returns.csv)

3. Data Preparation

- Loaded order and return datasets using Pandas.
- Cleaned missing values and duplicates.
- Converted date fields and ensured proper data types.
- Merged datasets on 'order_id' and created 'is_returned' flag.

4. Exploratory Analysis

- Grouped data by category, supplier, region, and marketing channel.
- Calculated return rates to identify patterns.
- Identified segments with higher return percentages.

5. Predictive Modeling

- Used Logistic Regression to predict return probability.
- Features: product price, category, supplier.

- Added 'return_probability' to final dataset.
- High-risk products identified with probability > 0.5.

6. Power BI Dashboard

- Visuals created:
  - Bar chart: Return rate by category/supplier.
  - Line chart: Monthly return trends.
  - Table: High-risk product list.
  - KPI Card: Return rate per selected category.
- Interactive features:
  - Slicers for category, region, supplier.
  - Drill-through page for detailed analysis by category.

7. Conclusion

The project successfully analyzed return behavior, built a return prediction model, and delivered a dynamic Power BI dashboard.

This empowers stakeholders to reduce return rates by identifying risky segments and tracking performance visually.