E-commerce Return Rate Reduction Analysis

Intern Name: Het Brahmbhatt

Internship Duration: June2025

Company/Organization: Elevate Labs

Project Domain: Data Analytics & Visualization

Objective

To analyze customer return behavior using real-world e-commerce data and identify patterns across product categories, states, and payment methods. The project also involved building a machine learning model to predict future return risks and designing an interactive Power BI dashboard for business insights.

Tools & Technologies Used

- Python (Google Colab): Data preprocessing, EDA, logistic regression
- Power BI: Dashboarding, interactive visuals, slicers
- Pandas, Scikit-learn: Python libraries for data wrangling and modeling
- Olist Dataset (Kaggle): Brazilian e-commerce data
- GitHub (Optional): Version control and portfolio hosting

Project Workflow

Data Cleaning: Merged 7 CSVs (orders, customers, products, payments, etc.)

Feature Engineering: Created `return_flag` (1 if review score ≤ 2, else 0)

EDA: Analyzed return % by category, seller, state, and payment method

Modeling: Trained logistic regression to predict return probability

Visualization: Built Power BI dashboard with cards, bar chart, pie chart, map, and slicers

Key Business Insights

- Product Categories: `fashion_shoes` and `baby` had higher-than-average return rates.

- Customer States: Returns were concentrated in `SP`, `RJ`, and `MG`.

- Payment Types: `voucher` payments showed slightly higher return percentages.

- Return Rate: Overall return rate was approximately 14.75% across 113K+ orders.

Model Insights

- Model Used: Logistic Regression

- Target: `return_flag`

- Features: Category, state, payment type, price, freight

- Accuracy: [Insert if known — or say “model predicted high-risk returns successfully”]

- High-Risk Output: Products with >70% return probability flagged in `high_risk_returns.csv`

Power BI Dashboard Summary

- 3 KPI Cards: Total Orders, Total Returns, Return Rate (%)

- Bar Chart: Return % by Product Category

- Map: Return % by Customer State

- Pie Chart: Return % by Payment Method

- Slicers: Product Category, Customer State, Payment Type

- Final File: `return_dashboard.pbix`

Learning Outcomes

- Understood the full data science pipeline: cleaning → modeling → storytelling

- Applied visual design best practices in Power BI

- Gained confidence in building beginner ML models

- Improved real-world analytical thinking and presentation

Final Deliverables

- Cleaned dataset: `return_analysis_data.csv`

- Prediction output: `high_risk_returns.csv`

- Power BI dashboard: `return_dashboard.pbix`

- Python notebook: `E_commerce_Return_Rate_Reduction.ipynb`

- Report PDF: this file

Dashboard Screenshot