

Project Title:

Fashion Retail Intelligence: Return Prediction

1. Background & Business Objective

In the dynamic world of fashion retail, understanding customer behavior and sales performance is essential for sustaining profitability, optimizing inventory, and enhancing customer satisfaction. Return rates are a major concern in online and offline fashion retail, often driven by size mismatches, quality issues, and shifting customer expectations.

The dataset under consideration represents real-world fashion retail sales and returns data from 2024–2025. It spans eight leading brands, covers four seasonal collections, and includes pricing, discounting, inventory, and customer feedback details.

Business Objective:

Use data analytics and machine learning to improve sales forecasting, minimize return rates, optimize pricing, and drive informed decision-making across product categories and seasons.

2. Business Problem

Despite having access to seasonal trends and pricing strategies, fashion retailers face critical challenges:

- High return rates lead to increased reverse logistics costs and customer dissatisfaction.
- Inefficient inventory stocking results in overstocking or stockouts.
- Lack of understanding of markdown effectiveness impairs pricing strategies.
- Customer satisfaction insights are often not integrated into product planning.

3. Goal

Develop data-driven models and insights that:

- Predict whether a product will be returned at the time of purchase.
- Analyze the effectiveness of markdown strategies on sales.
- Assess customer satisfaction using product ratings and return reasons.
- Deliver actionable visual dashboards for retail stakeholders.

4. Scope of Work

In Scope:

- Data cleaning, preparation, and exploratory data analysis
- Feature engineering for model building
- Classification model to predict return likelihood

- Correlation between ratings and returns (customer satisfaction)
- Inventory recommendations based on historical trends
- Business Intelligence dashboard using Power BI

5. Tools & Technologies

- **SQL (BigQuery):** Data modeling, cleaning, transformations
- **Python (Pandas, NumPy, Seaborn, Scikit-learn):** EDA & Machine Learning
- **Power BI:** Interactive dashboards and storytelling
- **Jupyter Notebook:** Data analysis and modeling workflow
- **GitHub:** Code and documentation repository

6. Deliverables

- Cleaned dataset and Snowflake schema.
- Python-based EDA and logistic regression model.
- Power BI dashboard showing insights and KPIs
- GitHub repository with README, Project Proposal, and PACE document.
- Executive Summary for stakeholders.

Data Analyst & Consultant: Shijin Ramesh

Date: 8th August, 2025