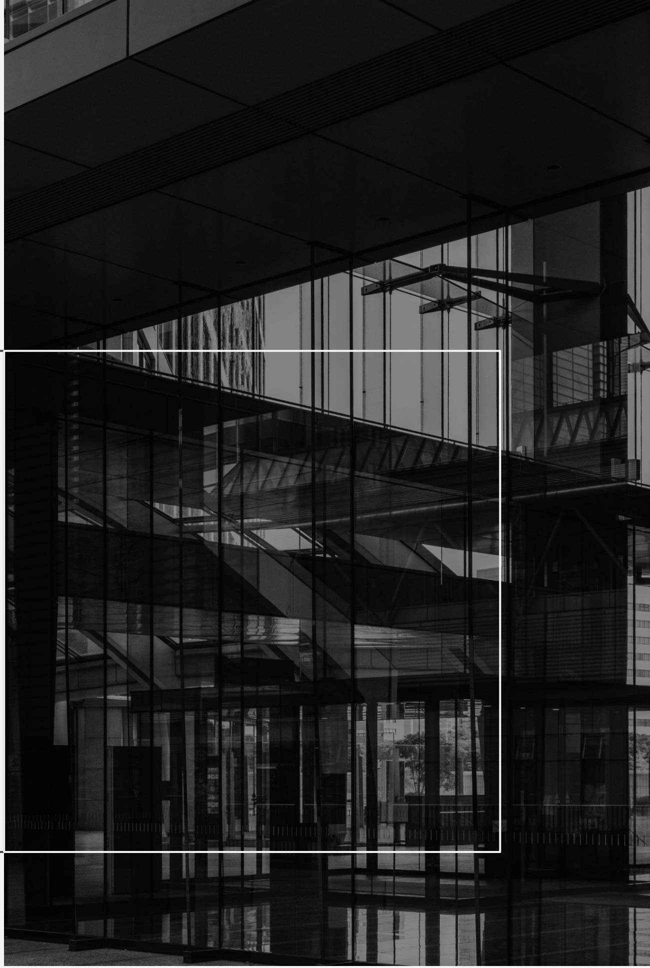
Second-hand Luxury Market Analysis

Vestiaire Collective

Vestiaire Collective

Vestiaire Collective is a global platform for buying and selling pre-owned luxury and designer fashion.

- Revenue Streams: Commission-Based Model (once sold, the platform deducts a commission (e.g., 10%-40%) before paying the seller.)
- Market Capacity: The global secondhand luxury market value stood at US$34.39 billion in 2023, and is expected to reach US$60.55 billion by 2029.
- Problems:$0.84 billion GMV in 2023, significantly lower than the competitor’s $1.83 billion; low conversion rate per listing



Transaction rate

A manifesto for agile data science

Action
Strategy & Recommendation

Prediction
ML & Causal Inference

Reports
Market Segmentation

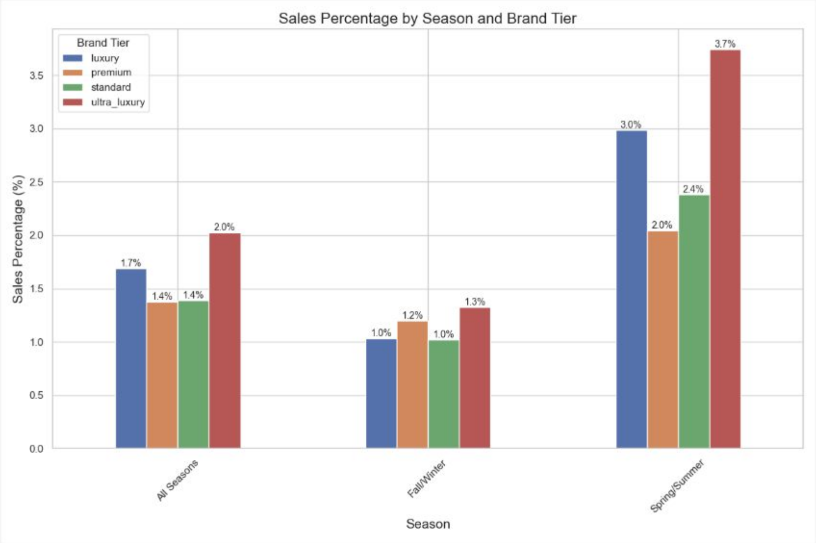
Charts
Clean, Aggregate, Visualize

Records
Collet dataset: Kaggle

Market Segmentation:

1. Sales Percentage by Season and Brand Tier

- Spring/Summer is the peak selling season for all brand categories, with ultra-luxury brands achieving the highest conversion rates (~3.7%).
- Ultra-luxury brands consistently outperform premium brands, highlighting the need for distinct promotional strategies and informing inventory and marketing planning.

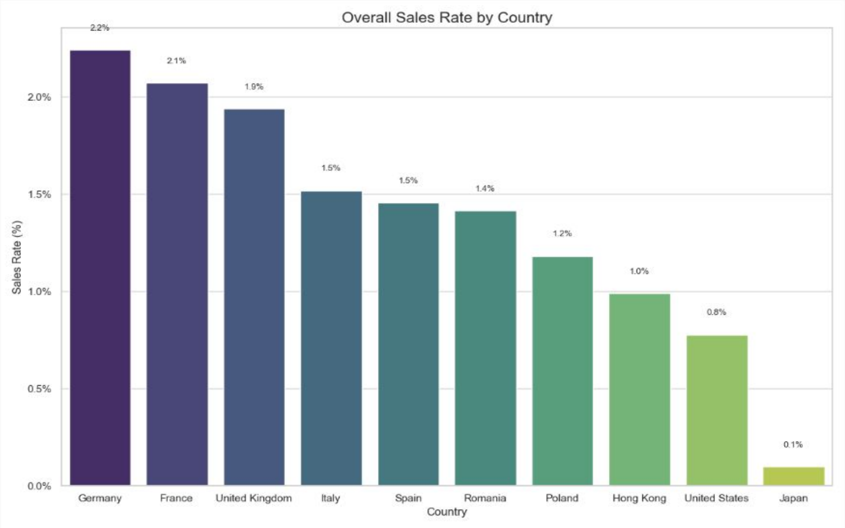


Market Segmentation:

2. Country Sales Rates Comparison (Bar Chart)

- European markets, such as Germany (2.24%) and France (2.07%), have significantly higher conversion rates than Japan (0.10%).

- These disparities highlight market traction differences and inform resource allocation for strategic growth.

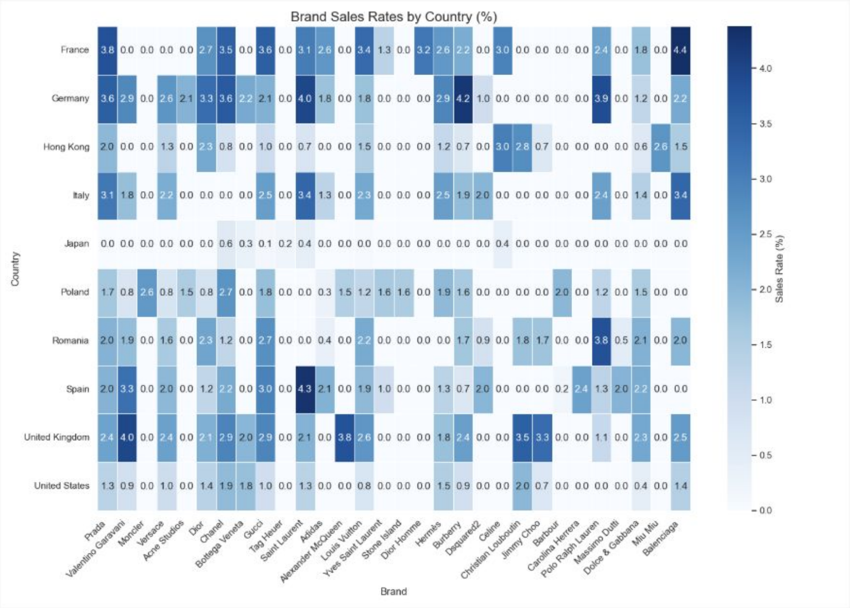


Market Segmentation:

3. Brand Sales Rates by Country (Heatmap)

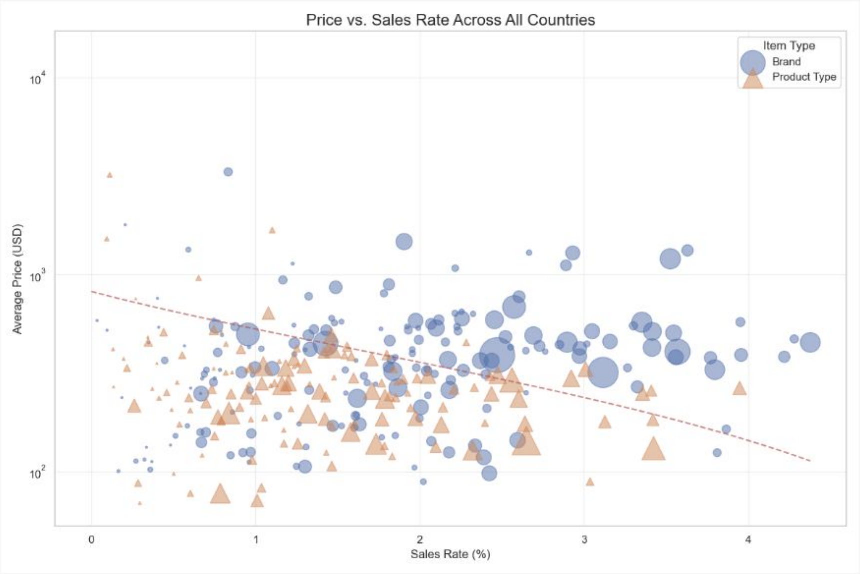
- Luxury brand performance varies by region, with Balenciaga excelling in France (4.38%) and Christian Louboutin resonating in the U.S. (1.98%).

- These cultural differences emphasize the need for localized marketing and geographic segmentation in merchandising and promotions.



Market Segmentation:

4. Price vs Sales Rate Scatter Plot



- Higher prices do not necessarily lower conversion rates in luxury markets, with some ultra-luxury brands seeing higher conversion despite premium pricing.

- This insight highlights exclusivity as a key driver of demand, reshaping pricing strategies to optimize conversion and revenue.

02

Buyer Segmentation

Product(Sold) Segmentation

Data pre-processing

- EDA
- Handle missing value
- Feature selection (only product related info & correlation matrix)

Feature Engineering

- Standardization /Normalization
- One-Hot Encoding
- Create brands representative column with average price
- Frequency representation of materials and color

```
------------------------------------------------
Column: brand_name
Unique Values (8868): ['Miu Miu' 'Barbara Bui' 'Comme Des Garcons
 'Just Cavalli' 'J.Crew' 'Aquascutum' 'Dior']
------------------------------------------------
Column: product_material
Unique Values (70): ['Wool' 'Cotton' 'Polyester' 'Vegan leather'
 'Synthetic' 'Lycra' 'Viscose' 'Denim - Jeans']
------------------------------------------------
Column: product_color
Unique Values (30): ['Grey' 'Navy' 'White' 'Black' 'Beige' 'Red'
 'Multicolour']
------------------------------------------------
```

Clustering Algorithm

- K-Means
- DBSCAN
- Gaussian Mixture Model

Evaluation/Fine tuning

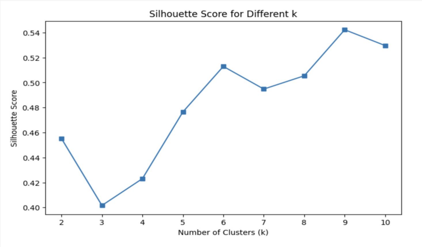
- Elbow Methods
- Silhouette Score

Visualization & Interpretation

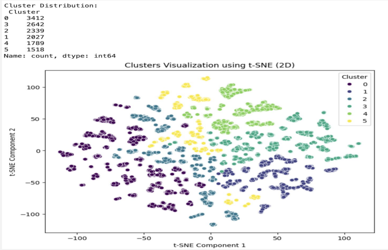
- PCA/t-SNE - Cluster visualization
- Semi-supervised learning: Decision tree

Cluster Results & Insights:

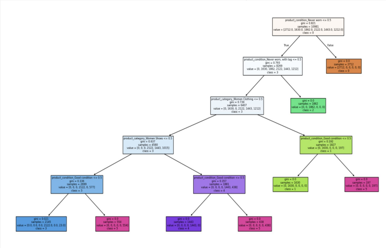
Silhouette score: 0.53



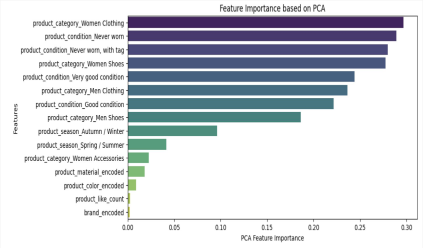
6 Clusters



Decision Tree



Feature importance



Cluster	**Buyer Profile**	**Key Preferences**	**Strategy**
0 & 2 - Quality Buyers	High-end shoppers	Only buy "Never Worn" or "With Tags" items, focus on authenticity & resale value	Offer premium authentication & certified resale options
1 & 5 - Budget Buyers	Price-sensitive or exploring luxury	Prefer "Good Condition", accessories, shoes, and entry-level luxury	Highlight discounts, first-time deals & smart savings sections
3 & 4 - Fashion & Seasonal Buyers	Trend-conscious & occasional shoppers	Focus on Women's & Men's Clothing, buy based on seasonality & trends	Promote seasonal collections, influencer styles & gifting ideas

03 Price Elasticity and Analysis on Sold Items

Price Elasticity Analysis

Price Elasticity measures how demand changes in response to price changes. Typically, demand decreases when price increases (**elastic**), or remains stable (**inelastic**).

However, for **Veblen goods** (luxury/status-driven) and **Giffen goods** (essential but consumed more when prices rise), demand may increase with price.

Analysis Steps

Missing Values

Handled missing values with appropriate imputation (eg: extracting color names from product_description using a predefined list and a regex pattern to match standalone color)

Collinearity Checks

Used VIF Scores and Collinearity Matrix to eliminate variables.

Feature importance

Used XGBRegressor to further filter out important features for training.

Uneven Distribution of Target Variable

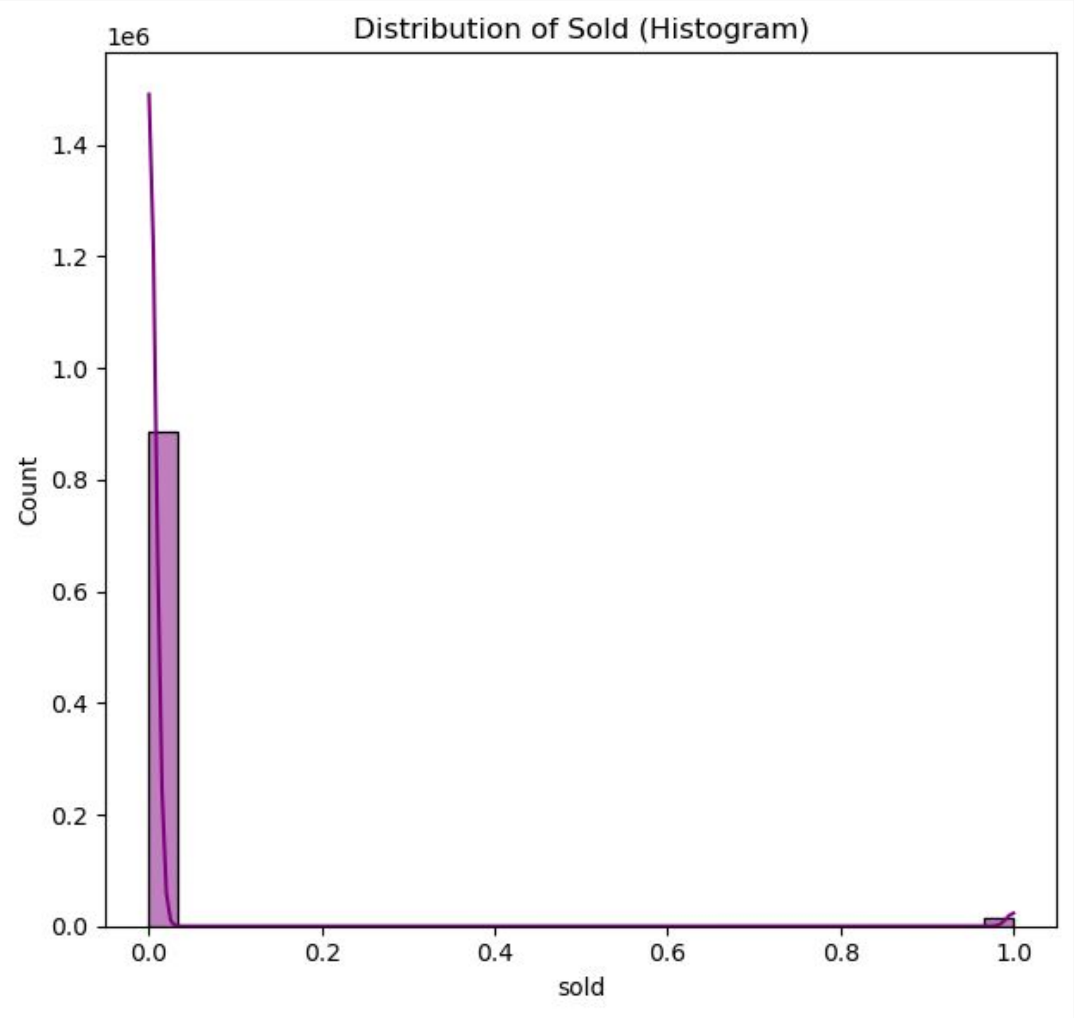
Used SMOTomek to balance the dataset and improve model generalization.

Model training + Learning Curve

Trained three models — XGBoost, LightGBM, and CatBoost — to predict product demand and evaluated their performance using the ROC-AUC score and learning curves.

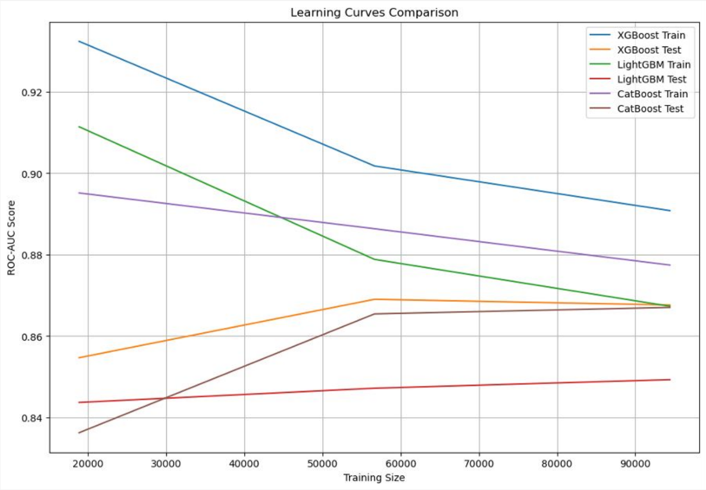
Causal Analysis - Price Elasticity

Examined price elasticity by assessing demand shifts in response to price changes, identifying features that significantly impact purchasing behavior.



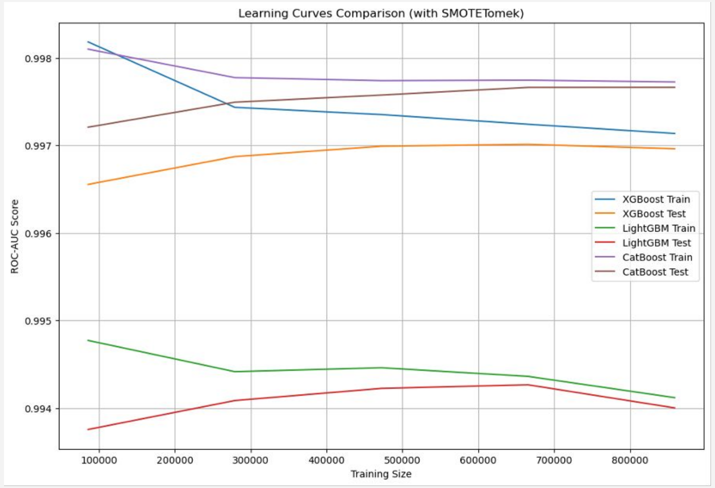
As seen in this figure,
The distribution of the target was extremely skewed, so using sampling techniques like SMOTETETOMEK was a necessity.

Training + Learning Curves



WITHOUT Smottomek

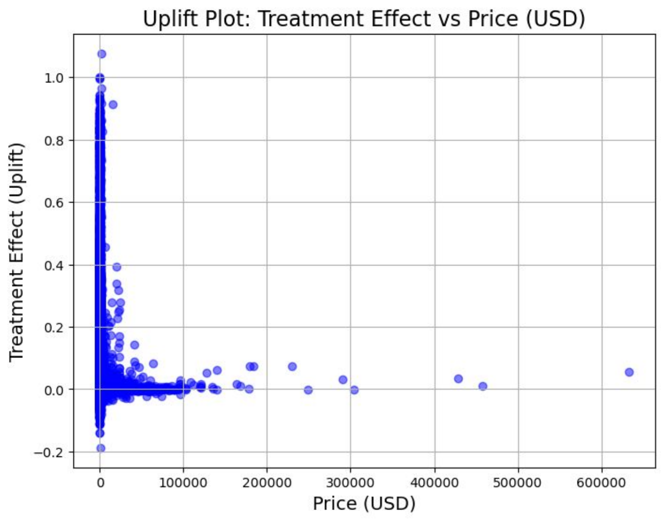
- XGBoost shows the highest training and testing performance but exhibits a clear gap between training and testing curves, indicating **overfitting**.
- LightGBM and CatBoost have lower ROC-AUC scores, with LightGBM performing better on both train and test sets.
- There is a **downward trend** in training performance as the sample size increases, which is expected due to increased complexity in larger datasets.



WITH Smottomek

- The overall ROC-AUC scores for all three models **improve significantly** across both training and testing sets.
- The gap between the training and test curves is **reduced**, particularly for XGBoost and CatBoost, indicating better generalization.
- Performance is more **stable** as the training size increases, suggesting improved handling of class imbalance.

DoubleMLPLR Model

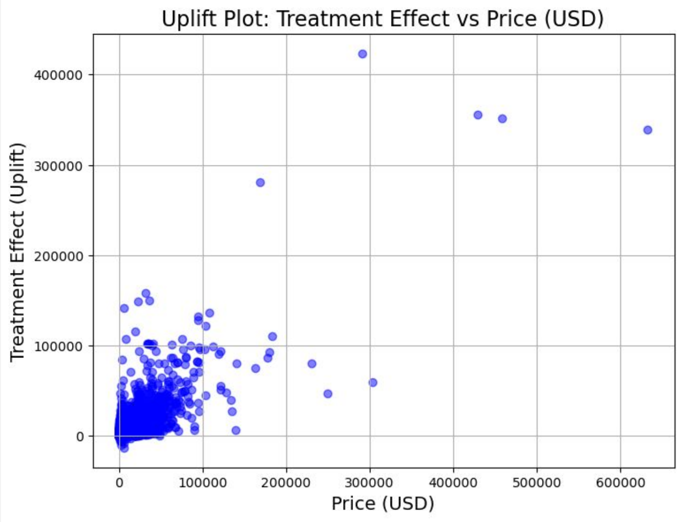


Outcome Model

Left (High Uplift Segment): Certain customer groups respond positively to price increases.

Right (Low Uplift Segment): Most customers show minimal response, indicating limited elasticity.

Focus price increases on high uplift groups to maximize revenue.



Treatment Model

Steep Decline: Clear differentiation in how customer segments react to price changes.

Implement dynamic pricing by targeting segments most likely to convert at higher prices.

Price Elasticity Differences

```
Estimated Price Elasticity: [1.73491026e-07]
Standard Errors: [6.8830299e-08]
Demand increases with price (Inelastic).
```

Inelastic Demand for Listed Price (price_usd)

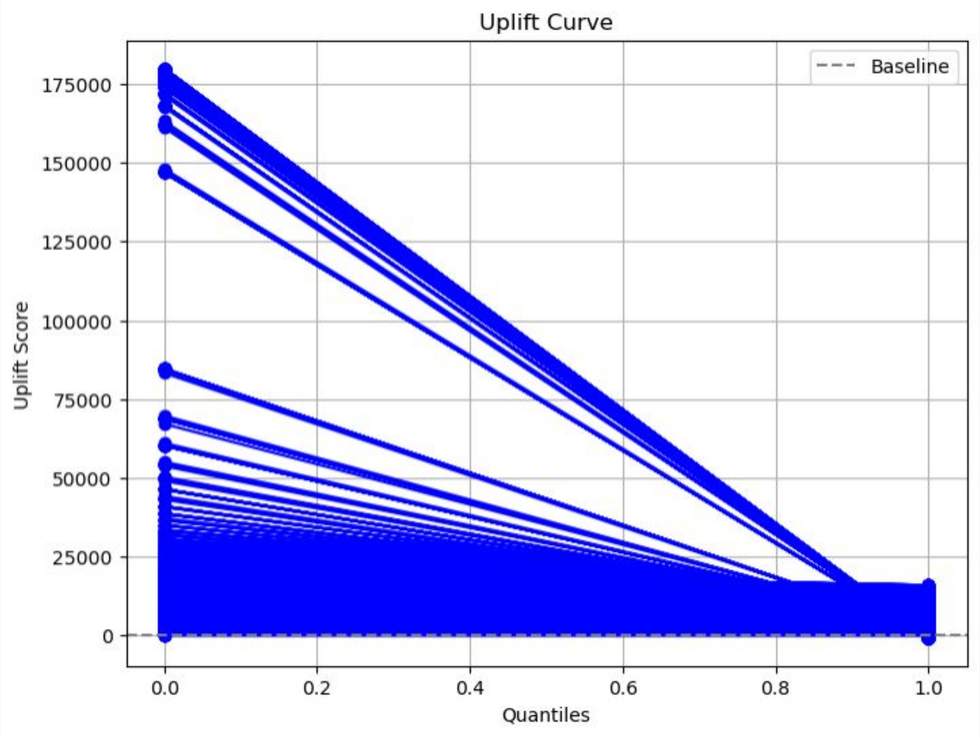
When the **listed price** is high, demand does not decrease significantly—or may even increase slightly. This suggests that customers are **not highly sensitive** to the official listed price. This makes sense because these items are mostly luxury brand items.

```
Estimated Price Elasticity: [-7.41819351e-07]
Standard Errors: [1.14836247e-07]
Demand decreases as price increases (Elastic).
```

Elastic Demand for Seller Price (seller_price)

When the seller’s actual price increases, demand drops noticeably. Customers are **price-sensitive** when it comes to the amount they actually pay.

Uplift curve



Steep Decline from Left to Right:

The leftmost customers (top quantiles) have the **highest uplift**, meaning these groups **respond well to higher prices**.

As you move right, the effect weakens, and **most customers show little to no uplift**.

Large Positive Uplift in Early Quantiles:

Suggests there is a **segment** of the market where **raising prices increases sales** (e.g., luxury buyers or less price-sensitive consumers).

Baseline (Dashed Line at 0):If the uplift score is **above the baseline**, the treatment (high price) is **beneficial**.

Average Uplift Effect = 0.0789: On average, **raising prices increases** the likelihood of a sale by **7.89%**.

Positive value confirms that your customer base exhibits inelastic demand overall.

Business Insights

Premium Positioning

Use higher listed prices for inelastic products to enhance prestige and drive profitability.

Price Optimization

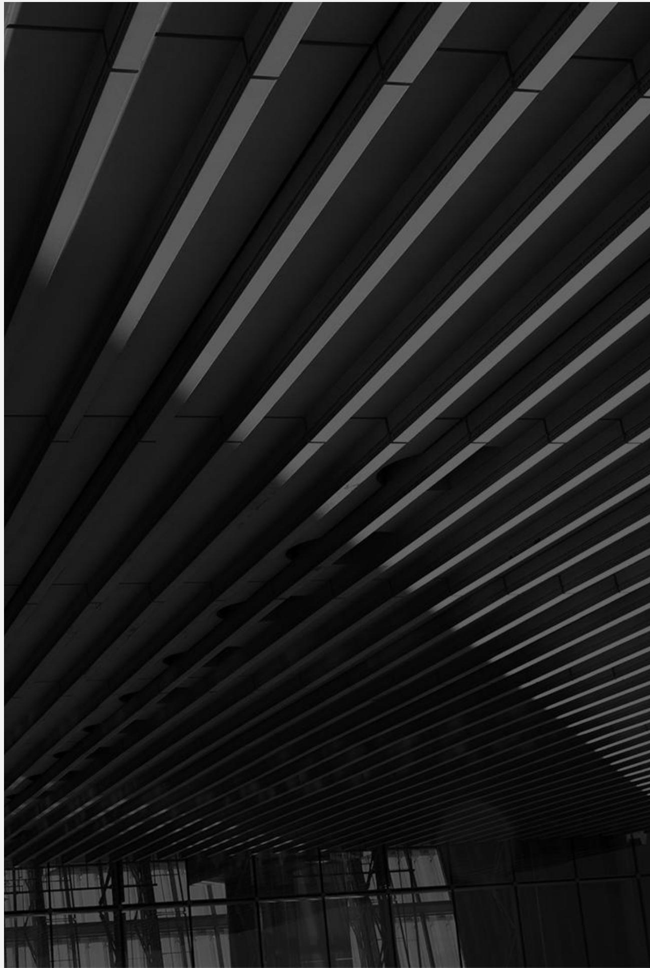
Apply a dynamic pricing model for seller prices to balance competitiveness with profitability.

Personalized Discounts

Offer targeted incentives to capture elastic demand and boost conversions.

A/B Testing

Continuously monitor and test pricing strategies to refine approaches and maximize both demand and revenue.



04 Seller Analysis

Classification

Feature Selection

Targets: Seller type

Seller Group	Group Size
Regular	106976
High Revenue	79840

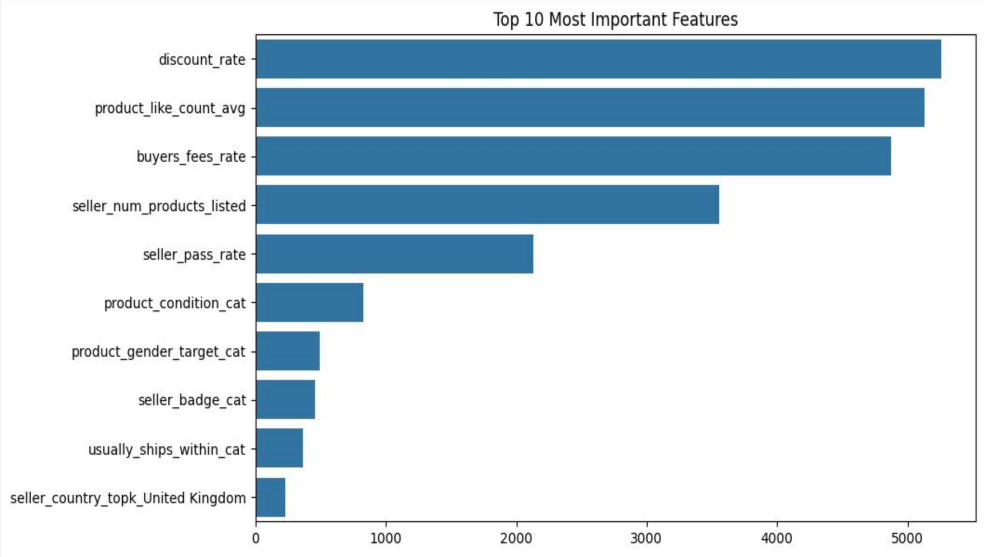
Predictors:

- **Price related:** Discount rate; Buyers’ fee rate
- **Seller related:** Seller pass rate; Seller badge level; Seller country; Avg like counts per product; Avg num listed products
- **Product related:** Product condition; Product gender target; Shipment detail

LightGBM

	Precision	Recall	F1 - score
High Revenue	0.82	0.77	0.80
Regular	0.84	0.87	0.85
Accuracy	—	—	0.83
Macro Avg	0.83	0.82	0.82
Weighted Avg	0.83	0.83	0.83

Feature Importance



- **Top Influencers:**
  Pricing strategy and customer engagement.
- **Moderate Influencers:**
  Seller reliability and inventory.
- **Less Influential Factors:**
  Products details and seller country features


Causal Inference

Uplift Random Forest

Y: Seller Type

Regular vs Top Seller

Treatment: Badge

Common vs Advanced

Uplift	Y(Seller Type)	T(Badge)
21.18%	Regular	Common
21.86%	Top	Common
22.18%	Top	Advanced
21.73%	Regular	Advanced

Business Overview

For Vestiaire Collective Platform

Fair Pricing Strategy

Encourage sellers to adopt fair pricing and build high-quality storefronts.

Badge System

Implement a tiered badge system that rewards quality and fair pricing.

Seller Support

Allocate additional resources to high-revenue potential sellers to drive platform growth

Interactive Atmosphere

Foster a dynamic community by encouraging customer-seller interactions like likes and comments.

Thank You!