**Dataset Link**: <https://archive.ics.uci.edu/dataset/352/online+retail>

This is a transactional data set which contains all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online retail. The company mainly sells unique all-occasion gifts. Many customers of the company are wholesalers.

* Dataset Characteristics: Multivariate, Sequential, Time-Series
* Subject Area: Business
* Feature Type: Integral, Real
* Instances: 541909

**Key Features in the Dataset**

The dataset contains transactional data from an online retailer, which includes:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable Name | Role | Type | Description | Units | Missing Values |
| InvoiceNo | ID | Categorical | a 6-digit integral number uniquely assigned to each transaction. If this code starts with letter 'c', it indicates a cancellation |  | no |
| StockCode | ID | Categorical | a 5-digit integral number uniquely assigned to each distinct product |  | no |
| Description | Feature | Categorical | product name |  | no |
| Quantity | Feature | Integer | the quantities of each product (item) per transaction |  | no |
| InvoiceDate | Feature | Date | the day and time when each transaction was generated |  | no |
| UnitPrice | Feature | Continuous | product price per unit | sterling | no |
| CustomerID | Feature | Categorical | a 5-digit integral number uniquely assigned to each customer |  | no |
| Country | Feature | Categorical | the name of the country where each customer resides |  | no |

**Note**: I have added two more fields I,e Age and Gender for demographic purpose in the dataset.

**How the Dataset Can Be Used for the Dynamic Pricing Engine**

1. Customer Lifetime Value (CLV) Prediction:
   * The dataset contains customer transaction histories (CustomerID, InvoiceDate, Quantity, UnitPrice), which can be used to calculate CLV.
   * CLV can be estimated based on:
     + Purchase Frequency: How often a customer makes a purchase.
     + Average Order Value: Average spending per transaction.
     + Customer Tenure: How long the customer has been active.
     + Churn Rate: Likelihood of the customer stopping purchases.
2. Personalized Pricing Strategies:
   * The dataset can be used to segment customers based on their buying behavior (e.g., frequent buyers, high spenders, seasonal buyers).
   * Pricing strategies can be tailored based on customer segments and their predicted CLV.
3. Dynamic Pricing:
   * The dataset includes product-level data (StockCode, Description, UnitPrice), which can be used to analyze price elasticity and demand patterns.
   * Real-time pricing adjustments can be made based on customer behavior, product popularity, and market trends.

**🚀 Dynamic Pricing Engine Based on CLV Model**

A Dynamic Pricing Engine adjusts prices based on Customer Lifetime Value (CLV), helping you maximize profitability and customer retention.

**📌 How It Works**

1. Load the Trained CLV Model (from clv\_model.pkl).
2. Predict CLV for Customers based on their behavior.
3. Apply Pricing Strategy:
   * High CLV Customers → Lower Discounts (since they already spend more).
   * Low CLV Customers → Higher Discounts (to encourage more spending).
4. Update Prices Dynamically based on CLV segmentation.

🡺**This script:**

✅ Loads the CLV model

✅ Predicts CLV for new customers

✅ Applies a pricing strategy

✅ Saves the updated prices

**Dynamic Pricing Rules Based on CLV**

Use the predicted CLV to segment customers and apply different pricing strategies:

|  |  |  |
| --- | --- | --- |
| **CLV Segment** | **Pricing Strategy** | **Discount/Premium** |
| High CLV (> 75th percentile) | Premium Pricing | Slightly higher prices (loyal customers tolerate small increases) |
| Medium CLV (25th–75th percentile) | Neutral Pricing | Standard pricing (no changes) |
| Low CLV (< 25th percentile) | Discount Pricing | Small discounts to encourage retention |

**Explanation of the CLV Model Trainer Script**

This script is designed to:

* Predict customer lifetime value based on behavioral metrics (frequency, recency, monetary value, etc.)
* Create a reusable model that can be deployed to score customers
* Save both the model and its predictions for analysis
* Provide detailed logging for monitoring and debugging

The Random Forest approach is well-suited for CLV prediction as it can handle non-linear relationships between customer behaviors and their lifetime value.

**GUI Application**

**Key Features:**

1. **Complete Implementation**:
   * Dynamic pricing engine with CLV modeling
   * Full-featured PyQt5 GUI
   * Flask REST API
   * Performance monitoring
   * A/B testing framework
2. **Ready to Run**:
   * Single file implementation
   * Automatic directory creation
   * Default configuration
   * Error handling throughout
3. **How to Use**:
   * Save as dynamic\_pricing.py
   * Install requirements: pip install pandas numpy scikit-learn scipy flask pyqt5 joblib
   * Run with: python dynamic\_pricing.py
4. **Three Access Methods**:
   * **GUI**: Full graphical interface
   * **API**: HTTP endpoint at http://localhost:5000/price
   * **Programmatic**: Direct Python class usage
5. **Automatic Setup**:
   * Creates required directories
   * Generates default model and rules if missing
   * Handles first-run scenarios