**🧠 Customer Lifetime Value (LTV) Prediction Model**

**Objective:**

To develop a machine learning model that predicts the future lifetime value (LTV) of customers based on their past purchase behavior. This aids in customer segmentation and enables more effective targeted marketing and resource allocation.

**Tools Used**

- Python: For data analysis, model building, and evaluation

- Libraries: pandas, sklearn, xgboost, matplotlib, seaborn

- Excel: For preliminary data inspection and reporting

**Methodology**

**A. Data Collection & Preparation:**

Transaction-level data with Customer ID, Order ID, Order Date, and Order Value.

Cleaned data by removing duplicates and filtering out invalid entries.

Converted date columns to datetime format for accurate temporal analysis.

**B. Feature Engineering:**

Generated customer-level features including:

Recency: Time since last purchase.

Frequency: Number of purchases per customer.

Monetary: Total amount spent.

AOV (Average Order Value) = Monetary / Frequency.

Tenure: Time between first and last purchase.

**C. Modeling Approach:**

Target: Predicted Customer LTV (Total expected spend in next 6 months).

Algorithm: XGBoost Regressor (chosen for its accuracy and feature importance analysis).

Data split into training and test sets.

Evaluation using:

MAE (Mean Absolute Error)

RMSE (Root Mean Squared Error)

Visual inspection via true vs. predicted scatter plot.

**D. Customer Segmentation:**

Predicted LTV values were binned into three groups:

High Value Customers

Mid Value Customers

Low Value Customers

**Results**

The model achieved a reasonable MAE and RMSE on the test set.

Key features influencing LTV: Frequency, Recency, and Monetary value.

Clear segmentation revealed where the company should focus marketing efforts.

**Business Impact**

Helps marketing teams focus on high-LTV customers for promotions.

Optimizes marketing budget by reducing waste on low-value leads.

Offers potential for automation in CRM tools using this model

**Conclusion**

TheCustomerLifetime Value (LTV) prediction model successfully leveraged historical purchase behavior to forecast future customer value using advanced regression techniques. By engineering key features such as frequency, recency, and average order value, the model was able to capture purchasing patterns that strongly correlate with future spend. TheXGBoost regression model performed well, providing reasonably accurate predictions validated using MAE and RMSE metrics. Further, the segmentation of customers based on predicted LTV offers actionable insights for the business, enabling smarter, data-driven marketing strategies.

This model equips the business with the ability to:

- Prioritize high-value customers for retention campaigns

- Re-engage medium-value customers with targeted offers

- Minimize spend on low-value customers, optimizing marketing ROI

With minor extensions, such as incorporating demographic or behavioral data, this LTV framework can evolve into a powerful, production-ready tool for customer relationship management and long-term strategic planning.