Customer Lifetime Value (CLV) Prediction Project Report

Objective:

The objective of this project is to analyze customer purchasing behavior and predict Customer Lifetime Value (CLV) using SQL and Python. This analysis helps businesses identify high-value customers, enhance retention, and optimize marketing strategies.

1. Data Source:

Dataset: Online Retail CLV.csv

Records: Approximately 540K transactions

Fields: Invoice_No, Stock_Code, Description, Quantity, Invoice_Date, Unit_Price, Customer_ID, Country

2. SQL-Based Data Analysis:

- Database Creation & Loading
- Profit Calculation and Revenue Analysis
- Customer Insights and RFM Feature Engineering

3. Python-Based LTV Modeling:

Libraries: pandas, sklearn, seaborn, matplotlib

Model: Linear Regression

$R^2 \approx 0.85$ | Low MSE → Good accuracy

4. Business Insights:

- Top 20% of customers generate approximately 80% of revenue.
- Recency, Frequency, and Monetary features strongly influence LTV.

5. Recommendations:

1. Segment customers using RFM.
2. Offer loyalty and re-engagement programs.
3. Target high-AOV customers for upselling.

6. Conclusion:

This project combines SQL, Python, and Machine Learning to build an interpretable

LTV prediction pipeline. Insights support retention and revenue growth strategies.

Prepared By:

Data Analytics Project Team

Customer Lifetime Value Prediction Analysis