

Customer Lifetime Value (CLV) Analysis — Case Study

Objective

The objective of this project was to understand long-term customer value using historical transaction data and identify how businesses can prioritize customer retention and acquisition efforts based on lifetime value rather than short-term sales metrics.

Business Context

Many businesses focus on monthly sales or order volume, which often hides the true value of customers over time. This analysis aims to estimate Customer Lifetime Value (CLV) to identify high-value customer segments and highlight factors that influence long-term revenue contribution.

Dataset

The analysis was performed using the Superstore dataset sourced from Kaggle, containing transaction-level sales data from 2014 to 2017. Key fields included customer identifiers, order dates, sales values, product categories, customer segments, and geographic regions.

Methodology

The data was cleaned and prepared using SQL (MySQL), Python, and Excel. This involved handling refunds (negative sales), resolving missing values, and correcting date inconsistencies. A star schema was designed with a central sales fact table and supporting customer, product, date, and location dimensions to enable structured analysis.

Customer Lifetime Value was calculated using a historic approach based on observed behavior. Core metrics included Average Purchase Value, Purchase Frequency, Monthly Churn Rate, Estimated Customer Lifespan (inverse of churn), and Average Revenue per Customer. CLV was computed as the product of average revenue per customer and estimated customer lifespan.

Analysis & Insights

The analysis showed that a relatively small portion of customers contributes a disproportionately high share of total lifetime revenue. Differences in CLV across segments were driven more by customer retention (churn) than by order value alone. Certain customer segments and regions demonstrated stable lifetime value despite lower sales volume, indicating strong retention characteristics.

Business Recommendations

Based on the findings, businesses can prioritize retention strategies for high-value and at-risk customers to maximize lifetime revenue. Improving onboarding experiences during the early customer lifecycle can help reduce churn. Marketing and promotional efforts can be better targeted toward high-CLV segments rather than spread evenly across all customers. Tracking CLV trends over time enables more informed budget allocation and performance monitoring.

Assumptions & Limitations

This analysis is based on historical data and does not include predictive modeling. External factors such as pricing changes, marketing campaigns, and seasonality beyond the dataset period were not fully accounted for.