Problem Statement and Insights

Problem Statement

The e-commerce business has a steady customer base and ongoing sales, but lacks clear visibility into customer behavior, service performance, and revenue drivers. While transactions are occurring, there is limited understanding of how to improve customer experience, increase retention, and maximize profitability.

This project analyzes historical e-commerce data to identify trends in sales, revenue, profit, and customer behavior. Key areas of focus include churn patterns, regional performance, purchasing behavior, and periods of growth or decline.

The insights aim to support better decision-making, improve service delivery, strengthen customer retention, and drive sustainable revenue growth.

Business Questions and Insights

---

1. Customer Retention and Segmentation

Question

What is the monthly churn rate, and are there customer attributes linked to higher churn?

Chart 1: Monthly Active Customers Over Time

1. Active customers grew sharply from late 2023 to early 2024, increasing from about 1,179 to over 3,100. This likely reflects a platform launch, marketing campaign, or seasonal demand.

2. From January 2024 to October 2025, customer numbers remained stable between 2,800 and 3,200. This indicates strong replacement of churned customers and overall retention

stability.

3. After October 2025, there is a noticeable decline toward 2,000 customers, signaling a potential operational, competitive, or marketing issue that requires urgent investigation.

Chart 2: Active vs Churned Customers Per Month

1. During the stable period (Jan 2024 – Oct 2025), churned customers closely match active customers. This indicates a “leaky bucket” situation, where new customers replace those leaving rather than long-term retention driving growth.

2. A dip in February 2025 followed by recovery in March 2025 suggests that reactivation or acquisition campaigns can quickly restore customer levels.

Metric	Trend	Business Implication
Growth	Plateau after Jan 2024	Customer acquisition no longer exceeds churn
Retention	Low and volatile	Most customers behave as short-term users
Recovery	Strong rebound in Mar 2025	Campaigns effectively reactivate users
Risk	Elevated late 2025	Retention cycle may be breaking down

Chart 3: Churned Customers Over Time

Churn occurs consistently throughout the timeline, showing that it is an ongoing issue rather than a one-time event. Only a small group of customers remains active toward late 2025, indicating limited long-term loyalty.

Chart 4: Customer Churn by Attributes

1. Customers in lower credit tiers or non-qualified transaction profiles show higher churn, likely due to payment friction and weaker platform commitment.

2. Customers in regions with higher shipping costs or longer delivery times churn more frequently.

3. Age does not appear to be a strong determinant of churn.

4. Newsletter subscribers show better retention and higher likelihood of returning.

Question

Which regions have the highest customer concentration versus highest revenue per capita?

Chart 5: Customer Concentration by Country

1. Revenue is well distributed globally, with Peru and Bahrain generating the highest total revenue. Many countries contribute steady mid-level revenue, creating overall stability.

2. Smaller countries such as Bahrain and Barbados generate high revenue despite smaller populations, indicating high revenue per customer. Larger countries like China and Nigeria have lower per-capita revenue, suggesting upselling opportunities.

Category	Top Countries	Strategic Insight
Highest Revenue	Peru, Bahrain	Strong-performing markets
Highest Per Capita	Bahrain, Barbados	High purchasing power
Volume Markets	China, Nigeria	High user base, low average spend

---

Question

How do loyalty, credit tiers, and subscriptions affect purchase frequency and LTV for top customers?

Chart 6: Top 10% Customers by Total Spend

Top customers have very similar spending levels, indicating revenue stability rather than reliance on a few extreme spenders. Losing any of these customers would significantly impact revenue. A loyalty threshold around this spending level could help retain them.

Metric	Observation	Strategic Action
Spend Range	Very narrow	Introduce a higher premium tier
Diversity	High	Personalize communication
Stability	High	Use for revenue forecasting

Chart 7: Purchase Frequency vs Lifetime Value (LTV)

1. High LTV customers tend to be in Credit Tiers A or B and subscribed to newsletters, even with lower purchase frequency. Value comes from high-ticket purchases rather than frequent buying.

2. High loyalty scores alone do not guarantee high LTV or retention.

3. Customers without credit tiers show low loyalty and negligible LTV.

Attribute	LTV Impact	Frequency Impact	Churn Risk
Tier A/B	Very High	Moderate	Low with newsletter
Newsletter	High	Neutral	Improves retention
High Loyalty	Neutral	Low	Not sufficient alone
Tier D / None	Low	Low	High

---

2. Product Performance and Market Basket

Question

Which products are volume drivers vs profit engines?

Chart 8: Top 10 Products by Revenue and Profit

1. PROD000419 shows extremely high negative revenue, indicating major returns, pricing errors, or data issues. It represents a critical business risk.

2. Products PROD002487 to PROD002126 generate strong revenue with high margins and form the core profit base.

Question

Which products fall into the bottom 5% of sales and ratings?

Chart 9: Bottom 5% Products by Revenue and Rating

1. Products with very low ratings and low revenue contribute little value and should be cleared.

2. PROD000419 remains the most urgent issue due to its massive negative impact.

3. Products with poor ratings but moderate revenue should be reviewed for quality improvements.

Category	Indicator	Recommended Action
Critical Failure	Negative revenue	Pause sales immediately
Quality Risk	Low rating, high revenue	Investigate and redesign
Dead Stock	Low rating, low revenue	Clear inventory

---

3. Operations and Fulfillment

Question

Is delivery speed correlated with customer ratings?

Chart 10: Delivery Speed vs Customer Rating

1. No strong overall correlation exists, but extremely long delivery times consistently result in low ratings.
2. Outliers with delivery times over 300 days suggest data or fulfillment issues.
3. Fast delivery does not guarantee high ratings, indicating product quality also matters.

Question

What is the late delivery rate and which warehouses are bottlenecks?

Chart 11: Warehouse Delivery Performance

1. All warehouses perform similarly at around 65%, indicating a systemic issue.
2. A 35% failure rate represents significant revenue loss and customer dissatisfaction.
3. No warehouse meaningfully outperforms others.

Question

What are payment method failure rates?

Chart 12: Payment Method Failure Rate

1. Cash and PayPal have the highest failure rates, often due to customer refusal or authentication issues.

2. Digital wallets (Apple Pay, Google Pay) perform best due to lower friction.
3. Even small failure rate differences represent large revenue losses at scale.

---

4. Financial Trends and Forecasting

Chart 13: Sales Trend Over Time

Sales peaked in early 2024 and declined steadily through 2025. There is no clear seasonality, and overall revenue has dropped by roughly 50% from peak levels.

Chart 14: Profit Trend Over Time

Profits mirror the sales decline, falling more than 50% between 2024 and 2025. Margins appear under pressure from rising costs or inefficiencies.

Chart 15: Order Volume Trend

Order volumes have steadily decreased since early 2024, aligning with declining sales and profits. Without intervention, demand may continue to weaken.

Chart 16 & 17: Sales and Profit by Currency

1. USD dominates sales but delivers weaker margins.
2. EUR and GBP generate higher profit relative to sales, indicating stronger efficiency.
3. Overreliance on USD exposes the business to currency and margin risk.

Chart 18: Customer Distribution by Spending Tier

Most customers fall into low or occasional spending tiers, while a small percentage generates most revenue. There is strong opportunity to move customers up tiers through engagement and loyalty programs.

Churn & LTV Modeling Insights

Churn Model

Customer inactivity is the strongest driver of churn. Customers who have not purchased in over 180 days are significantly more likely to disengage, making recency the most influential churn predictor. One-time buyers are particularly vulnerable, indicating that while customer acquisition is strong, post-purchase onboarding and early retention are weak.

Churn is driven by two main customer groups:

- Low-value, low-frequency customers who show little brand attachment.
- High-value customers who disengage over time, representing a major revenue risk despite strong historical spending.

The model identifies substantial revenue at risk, with several high-risk customers having spent over $1,000 historically. However, it also enables early intervention by providing churn probabilities, risk categories, and a ranked list of customers most likely to churn. This allows for targeted re-engagement, improved marketing efficiency, and increased customer lifetime value (CLV).

Marketing and operational insights reveal that:

- Dormant high-value customers are prime candidates for personalized reactivation.
- Discount-driven customers churn quickly and require low-cost engagement.
- Early churn among new customers highlights gaps in the first 30–60 days experience.
- Loyalty programs and email marketing significantly reduce churn, validating their strategic importance.

The churn model achieves perfect classification (Accuracy, Precision, Recall = 1.00), confirming that RFM (Recency, Frequency and Monetery value) features, especially Recency, are highly predictive and reliable for churn prevention strategies.

Customer Lifetime Value (LTV) Model

Customer value is highly concentrated among a small group of high-spending customers. Many top-LTV customers make large, infrequent purchases, often followed by long periods of inactivity. This indicates a big-ticket purchasing behavior, where order size matters more than purchase frequency.

Two high-LTV segments emerge:

- High-value but inactive customers with large historical purchases and long buying cycles.
- High-value repeat buyers who return less frequently but spend heavily each time.

Customer tenure and average order value (AOV) strongly influence LTV, with some single-order customers ranking among the highest predicted lifetime values. The model achieves near-perfect accuracy ($R^2 = 0.99$), making it highly reliable for revenue forecasting and strategic prioritization.

Key business implications include:

- A small number of customers drive the majority of future revenue.
- Inactivity should be interpreted differently for high-value versus low-value customers.
- Predicted LTV enables smarter allocation of retention and account management resources.
- High-potential new customers can be identified early and nurtured proactively.

General Recommendation

Overall, the business demonstrates strong customer acquisition and market reach, but sustainable growth is limited by high churn, weak retention, operational inefficiencies, and declining financial performance. Growth is currently driven by customer replacement rather than loyalty.

To stabilize performance and drive long-term growth, the company should:

- **Shift focus from acquisition to retention and customer lifetime value**, prioritizing high-value and high-potential customers using churn and LTV insights
- **Segment inactivity management**, applying high-touch engagement for high-value customers and cost-efficient interventions for low-value customers
- **Address fulfillment and delivery failures** to improve reliability, customer satisfaction, and trust
- **Eliminate or restructure loss-making products** and refocus efforts on profitable offerings
- **Promote high-performing payment methods** to reduce transaction failures and checkout friction
- **Invest in loyalty programs, personalization, and customer tier migration** to strengthen engagement and repeat purchasing
- **Expand into higher-margin regions and currencies** to improve overall profitability