Ecommerce Customer Churn Report

This comprehensive report provides an in-depth analysis of e-commerce customer churn. Leveraging strategic ETL pipeline analysis, we identify key drivers, segment customers by risk, and deliver actionable, data-driven recommendations to significantly enhance retention.

PROJECT OVERVIEW

Project Objective

This analysis is designed to precisely identify the underlying causes of customer churn, segment the customer base by risk profile, and formulate data-driven recommendations to significantly enhance retention.

The project adheres to a standard ETL (Extract, Transform, Load) pipeline, utilizing Python for meticulous data cleansing and SQL for insightful strategic analysis.

01

Extract

Efficiently gather raw customer data from diverse sources.

02

Transform

Rigorously clean and prepare data for analytical processing.

03

Load

Execute comprehensive analysis to generate actionable insights.

PHASE I

Data Cleaning & Preparation

Leveraging Python (Pandas) for robust data integrity prior to database ingestion. This meticulous process yielded a refined dataset of over 486,000 records, primed for advanced strategic analysis.

1

Standardization

To ensure seamless SQL compatibility across all database operations, column names were rigorously converted to `snake_case`.

2

Data Integrity

Rigorous checks confirmed no duplicate entries. Null values in critical identifiers, such as Customer ID, were meticulously addressed, and all records containing negative prices or quantities were systematically removed to uphold data quality.

3

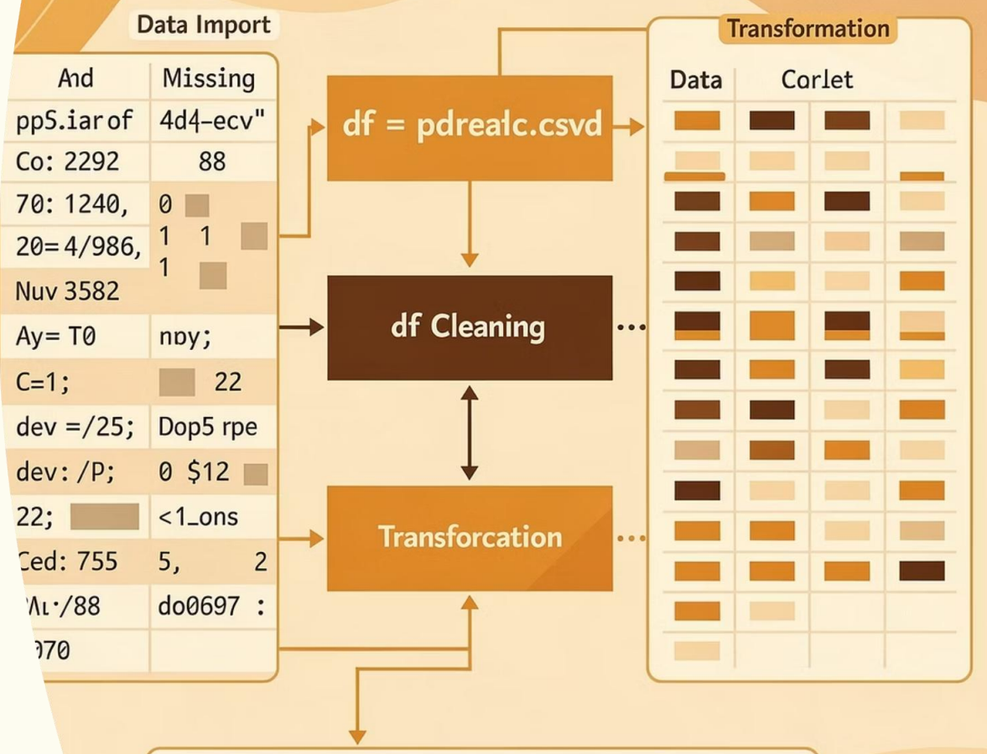
Logic Checks

Comprehensive logic checks were implemented to validate `tenure_days`, proactively filtering out anomalous records where the last login date preceded the signup date, thereby ensuring all tenure values are logically sound.

4

Feature Engineering

Key features were engineered, including the calculation of `tenure_days` (derived from Last Login Date - Signup Date). This metric is pivotal for enabling granular, time-based churn analysis and comprehensive customer lifecycle tracking.




```
import pandas as pd
import(pandas.esvd) {
df = pd.read.csv "data.csv" }
df = df dropna?
df = df(df/"value" > 0)
df "date" = pad-to-diixetm(df("date'))
df.clean = df reset indexx.drop.Truie"}
}
```

PHASE II

Strategic Analysis with SQL

Leveraging PostgreSQL for an in-depth exploration of customer behavior and retention patterns, this phase unveiled critical insights across four strategic pillars.

Acquisition & Marketing

Optimizing customer acquisition channels and identifying impactful seasonal trends.

Pricing & Operations

Assessing the efficacy of pricing strategies and mitigating operational friction, particularly concerning delivery.

Customer Segmentation

Refining customer segmentation to identify at-risk VIPs and predict potential churn.

Behavioral Insights

Uncovering key customer purchase patterns to drive enhanced engagement strategies.

ACQUISITION & MARKETING

Optimizing Acquisition Quality

Analyzing Churn Rate by Marketing Channel

The sheer volume of acquired users does not inherently signify high quality. For instance, if "Paid Ads" yields 10,000 new users but exhibits a 60% churn rate, its long-term value is significantly lower than a "Referrals" channel that acquires 1,000 users with a mere 10% churn rate. This clearly demonstrates the superior value of the referral channel despite its lower volume.

Strategic Action: Reallocate marketing expenditures from channels with high churn rates to those demonstrating superior retention to significantly enhance Return on Ad Spend (ROAS) and customer lifetime value.

60%

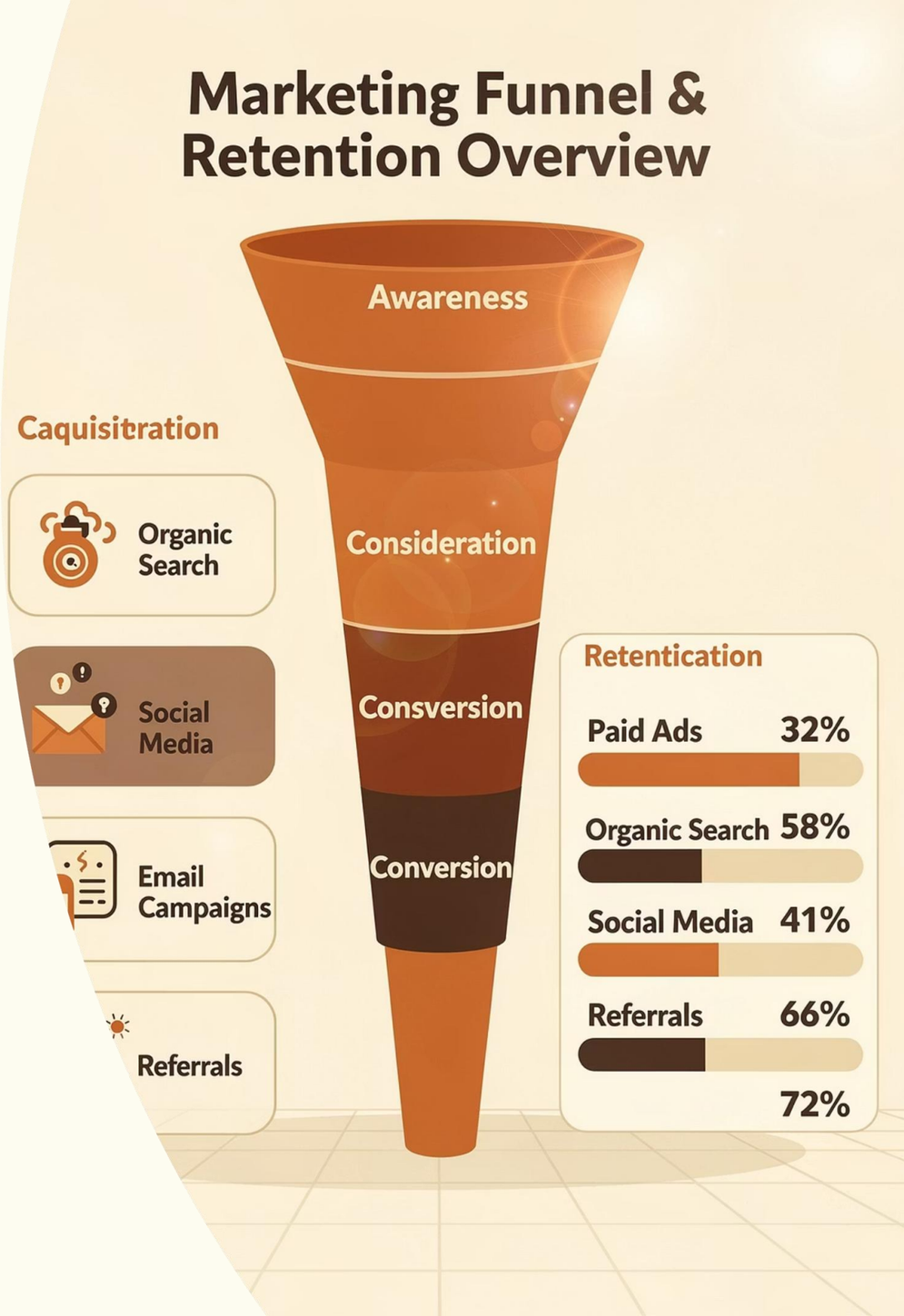
Paid Ads Churn

High volume, low retention

10%

Referral Churn

Lower volume, strong retention



The "Holiday Shopper" Effect

Acquisition Seasonality Analysis

Customers acquired in different months exhibit distinct retention patterns. A notable "January Hangover" effect shows that customers acquired in December (often impulse-driven holiday shoppers) experience significantly higher churn rates compared to those acquired in July (typically more deliberate, high-intent shoppers).

December Acquisitions

Characterized by holiday shopping surges, these customers often present a higher churn risk.

1

2

January Trends

The post-holiday "hangover effect" leads to elevated churn rates for recently acquired users.

3

July Acquisitions

Typically deliberate and high-intent, these customers demonstrate strong, sustained retention.

Strategic Action: Refine Customer Lifetime Value (CLV) models to reflect these seasonal variations. Avoid allocating identical advertising budgets for "December users" and "July users" if the former statistically yields half the long-term value.

PRICING STRATEGY

Price Sensitivity Impact

The Analysis

Our analysis involved segmenting customers into `Discount_User` and `Full_Price_User` groups, allowing us to compare their retention rates and accurately assess the long-term impact of promotional pricing strategies.

If customer retention is predominantly observed among those who initially engaged with discounts, it indicates an acquisition of "bargain hunters" rather than genuinely loyal customers.

Strategic Insight

Elevated churn rates among discount users signal an unsustainable pricing strategy. This necessitates a shift towards emphasizing core product value over aggressive discounting.

Critical Finding: A reliance on discounts for customer acquisition is often indicative of a weak product-market fit and leads to unsustainable customer acquisition costs.

Operational Friction Analysis

Delivery Time Impact on Customer Retention

Customer churn extends beyond product performance, often stemming from operational inefficiencies. This analysis segments customers by delivery speed (e.g., 0-3 days vs. 7+ days) to rigorously correlate logistics performance with observed churn rates.

Optimal Delivery (0-3 Days)

Exhibits the baseline churn rate, suggesting high satisfaction with timely service.

Extended Delivery (4-6 Days)

Demonstrates a noticeable increase in churn risk, indicating customer impatience.

Delayed Delivery (7+ Days)

Correlates with a doubling of the baseline churn rate, highlighting a critical operational bottleneck.

The stark disparity in churn rates—where customers experiencing deliveries exceeding seven days churn at twice the rate of those receiving items within two days—underscores that the primary driver of attrition lies in logistics, not inherent product quality.

Strategic Recommendation: Develop a comprehensive business case to justify investment in expedited shipping solutions or the establishment of new warehouse locations.

VIP ANALYSIS

Strategic Impact of High-Value Customer Churn

Recognizing that not all customer churn carries the same weight, we analyzed the precise revenue impact of departing users within our most valuable customer segments. This approach allows us to prioritize retention efforts where they yield the greatest strategic benefit.

Loss of One VIP Customer

Represents an approximate $1,000 revenue impact, highlighting the significant value of each individual high-tier client.

Loss of Ten New Customers

Equates to a combined $100 revenue impact, demonstrating the differential financial consequence compared to high-value churn.

> Strategic Imperative: A sharp increase in VIP customer churn demands immediate, highly personalized "white glove" retention campaigns. This targeted approach is significantly more effective than broad, automated outreach for mitigating substantial revenue loss.

Predicting At-Risk Customers

Leveraging RF Analysis: Recency & Frequency

By analyzing customer Recency (days since last login) and Frequency (total purchases), we classify active customers into distinct risk zones. This methodology acts as a proactive tool, enabling intervention before churn fully materializes.

High-Risk VIPs

Frequent purchasers (>5 times) exhibiting recent inactivity (60+ days since last visit).

- Initiate targeted win-back campaigns.
- Deploy personalized incentive programs.
- Mandate direct, personalized outreach.

Drifting Customers

Customers with infrequent purchase patterns and recent inactivity.

- Implement automated re-engagement email sequences.
- Provide attractive discount offers (e.g., 10% off).
- Deliver tailored product recommendation updates.

It is crucial to deploy a "Win-Back" email campaign (e.g., "We miss you, here is 10% off") targeting High-Risk VIPs *prior* to their official churn.

BEHAVIORAL PATTERNS

Behavioral Cooling Detection

Basket Size Decline Analysis

By comparing the quantity of items in a user's first order against their last order for customers who ultimately churned, we uncovered a crucial pattern. Customers rarely disappear abruptly; instead, a gradual "cooling down" period typically precedes full disengagement.

A significant negative percentage change in basket size (e.g., a -30% reduction) strongly indicates diminishing interest and provides an invaluable early warning signal for potential churn.

30%

Average Basket Decline

Observed prior to churn

Strategic Action: When analytics reveal a shrinking basket size, initiate a targeted "Volume Discount" (e.g., "Buy 5 items, get 20% off") precisely when a VIP customer's average basket size declines by 10%.

Share of Wallet Analysis

Cross-Category Revenue Patterns

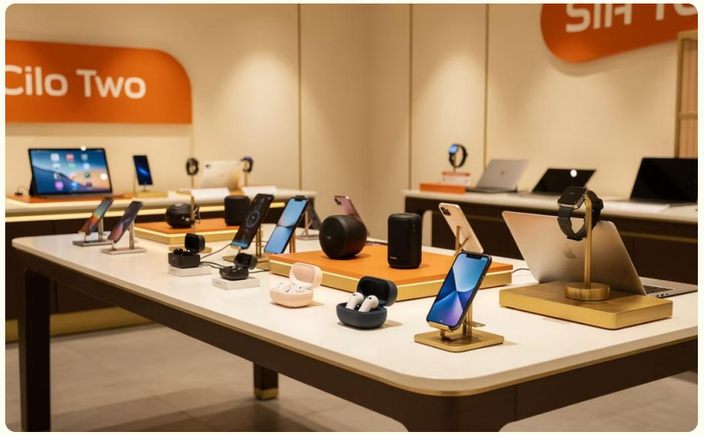
This analysis compares the revenue share across various product categories for both active and churned users. It highlights a critical pattern: customers often do not cease all purchasing activity abruptly, but rather reduce engagement or abandon specific product categories prior to full churn.

Home Decor

Exhibits a disproportionate churn signal, indicating potential underlying issues with product quality, design, or value proposition. Immediate investigation and corrective action are warranted.

Apparel

Demonstrates stable customer retention. Continued investment and optimization of the current strategy are recommended to sustain this performance.



Electronics

Identified as a rapidly growing category, presenting significant opportunities for strategic expansion and increased market penetration. Focus on innovation and aggressive marketing to capitalize on this trend.

Key Insight: A strong correlation exists between disproportionate spending on the "Home Decor" category by churned users and their subsequent departure. This pattern strongly suggests critical underlying issues within that category, such as product quality deficiencies or uncompetitive pricing strategies.

The Critical Window of Customer Churn

Uncovering Retention Patterns Through Tenure Analysis

By grouping churned customers according to their tenure (lifespan), we can pinpoint critical "drop-off" periods. Analyzing _when_ customers disengage offers crucial insights into _why_ they might be leaving.

1

First Week

High churn rates often indicate issues with the initial onboarding experience.

2

Months 1-3

This early engagement phase is crucial for assessing product-market fit and user adoption.

3

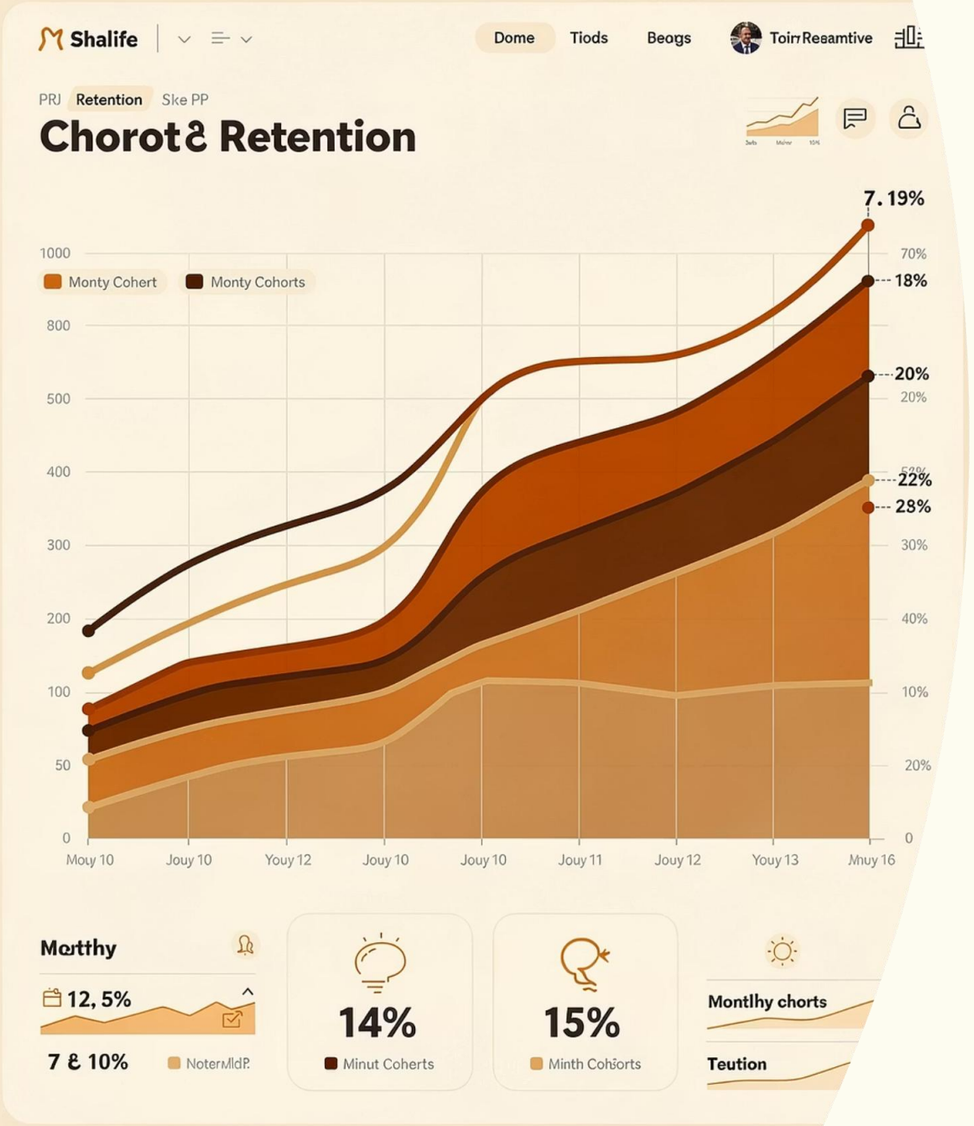
Months 6+

Churn at this stage typically points to diminishing long-term engagement or evolving customer needs.

- Addressing Early Churn: Optimize the onboarding experience, personalize welcome communications, and enhance initial product education.

- Mitigating Late Churn: Develop strategies to foster sustained engagement, such as implementing loyalty programs and offering exclusive, value-driven content.

Strategic Recommendation: Implement automated communications and targeted offers, timed to deploy proactively _just before_ identified drop-off points, thereby extending the customer lifecycle.



Cohort Analysis

The Retention Curve Over Time

This analysis meticulously tracks the retention rates of various monthly signup cohorts, providing crucial validation for product enhancements and measuring the enduring impact of strategic initiatives.

The Insight

Demonstrably, if the 'March Cohort' exhibits superior retention compared to the 'January Cohort,' it unequivocally confirms that features introduced in March have effectively enhanced product stickiness.

The Action

Leverage this insight to rigorously measure the long-term impact of product evolutions, thereby transcending the limitations of immediate sales metrics.

January Cohort

Establishes the foundational retention benchmark.

March Cohort

Reflects enhanced retention post-significant feature deployment.

CONCLUSION

Key Findings & Impact

This initiative successfully converted raw transactional data into actionable business intelligence. Through advanced Python-based data cleaning and precise SQL segmentation, we pinpointed critical churn drivers and developed effective proactive retention strategies.

Delivery Logistics

Delivery time is a significant churn driver; customers receiving items in 7+ days exhibit double the churn rate.

Seasonal Behavior

December holiday shoppers demonstrate significantly higher churn rates compared to typical July purchasers.

Proactive Intervention

A prioritized list of "At-Risk VIP" customers was delivered to the marketing team for immediate, targeted engagement.

Strategic Transformation: Our approach shifted from reactive measures to proactive retention, enabling us to identify at-risk customers before churn and implement targeted intervention campaigns.

Report Prepared by

Anshu Raj
ar.anshu2611@gmail.com