

# REPORT

## E-Commerce Customer Churn Prediction

---



### **Problem Statement**

Customer churn, or attrition, poses a significant challenge for businesses across various sectors, including telecommunications, e-commerce, and SaaS companies. Monitoring and analyzing the departure of customers, as well as identifying those who remain loyal, is of paramount importance. Understanding the factors driving these decisions is essential. By delving into customer behavior and preferences, businesses can make more informed decisions, optimize their services, and reduce churn rates. This proactive approach not only safeguards revenue but also enhances overall profitability. Therefore, the ability to track, analyze, and act upon customer behavior is critical in the ongoing effort to retain and

---

satisfy customers, ultimately ensuring the long-term success and sustainability of any customer-centric business.

## **GOAL**

- Create a predictive model to effectively pinpoint customers who may be inclined to leave the company, a process known as churn prediction. By examining the given variables, this model can identify at-risk customers, enabling the company to proactively take measures to retain them and mitigate churn.
- Prior to model building, conduct a comprehensive exploratory analysis of the provided customer data. This step involves a deep dive into the data to uncover patterns, trends, and customer characteristics. Through this analysis, the company can gain valuable insights into customer behavior, preferences, and attributes. This understanding not only aids in identifying the factors influencing churn but also informs future decision-making. It equips the company with the knowledge needed to enhance customer satisfaction and loyalty, ultimately reducing churn and bolstering its bottom line.

## **Churn Reasons**

Churn in business operations can be categorized into two primary types: voluntary churn and involuntary churn. Voluntary churn occurs when customers make a conscious decision to discontinue their subscription or usage of a service. This might happen because they found a better offer elsewhere or because they had a negative experience with the service. On the other hand, involuntary churn happens when customers are forced to leave the platform due to external factors, such as payment failures resulting from issues like maxed-out credit cards.

While a certain level of churn is expected in any business, high churn rates can have a substantial negative impact on a company's bottom line. It also serves as a reflection of the overall perception of the company among its customer base, which is a critical factor in its long-term growth and sustainability. Detecting significant churn might seem intuitive, but it remains a guessing game without access to key data points.

---

To effectively address churn and its impact, it's essential to delve into the various factors influencing it. A comprehensive understanding of these factors provides actionable insights for businesses to make informed decisions and take steps to reduce churn. By analyzing and addressing both voluntary and involuntary churn, companies can work towards improving customer satisfaction, retention, and, consequently, their long-term success.

## **EXECUTION**

- To begin the process of working with the provided customer data effectively, the initial step involves a thorough examination of the dataset. This step is vital for understanding the structure of the variables and the quality of the data. The key aspects to be addressed include data quality, the presence of missing values, and potential errors.
- Data quality checks are crucial to ensuring that the data is accurate and reliable. This involves identifying inconsistencies, inaccuracies, or anomalies in the dataset. Missing values, if present, need to be handled appropriately. Depending on the significance of missing data, imputation techniques may be required.
- The subsequent phase is exploratory data analysis (EDA), which aims to unveil insights into the data's distribution and relationships between variables. Detecting outliers or anomalies is part of this process, as they can skew the analysis. Furthermore, identifying correlations and patterns between variables provides a deeper understanding of the dataset.
- The next step is data pre-processing. This includes cleaning the data by addressing missing values and converting variables to appropriate data types. Data quality issues uncovered during EDA are resolved, ensuring that the dataset is ready for subsequent analyses or modeling.
- Visualization is a powerful tool during EDA, enabling the conversion of data into comprehensible graphs and charts. Visual representations offer valuable insights into customer behavior and characteristics, making it easier to grasp trends and tendencies.

Overall, this systematic approach, encompassing data overview, EDA, and pre-processing, is essential for extracting meaningful insights and ensuring the dataset's readiness for more

---

advanced analyses or predictive modeling. It lays the foundation for informed decision-making and insights into customer behavior.

## ANALYSIS

Analysis of the numeric features within the customer dataset provides valuable insights into factors that impact churn, offering guidance for proactive customer retention strategies.

**1. Tenure** Longer customer tenure is linked to a reduced likelihood of churn. This is intuitive, as it suggests that customers who have been with the company for an extended period are more likely to be satisfied with the service.

**2. City Tier** Interestingly, the churn rate appears to be relatively consistent across different city tiers. This indicates that the city tier is not a strong predictive factor for churn. It implies that customer satisfaction and loyalty can transcend geographical boundaries.

**3. Warehouse-to-Home** Shorter distances between the warehouse and the customer's home result in lower churn rates. Faster deliveries are likely to enhance customer satisfaction, making this an actionable insight for logistics and operations teams.

**4. Time spent on the app** The data suggests a positive correlation between spending more time on the app and reduced churn. This indicates that active app engagement is a positive sign, possibly reflecting the convenience and value the app offers to customers.

**5. Number of Devices Registered** Customers with multiple registered devices exhibit lower churn rates. Offering convenience and accessibility across devices can contribute to customer retention.

**6. Satisfaction Score** As expected, higher satisfaction scores strongly correlate with lower churn. Customer satisfaction is a critical factor in retaining customers.

**7. Number of Addresses** While the trend is subtle, there is a slight decrease in churn as the number of addresses associated with a customer increases. This could indicate that customers with more addresses are more loyal or have a stronger connection with the brand.

---

**8. Complain** The data shows that a higher number of complaints is associated with higher churn, although the relationship is not extremely strong. While complaints are not the sole factor, they do have a negative impact on customer satisfaction.

**9. Order amount hike from last year** Customers who increased their spending from the previous year are less likely to churn. This insight suggests the importance of retaining big spenders by offering incentives and personalized experiences.

**10. Coupon used** The use of coupons is linked to lower churn rates. Coupons can act as loyalty incentives, encouraging customers to stay with the company.

**11. Order count** Higher-order counts are associated with lower churn. Frequent usage of the service builds habits and loyalty, making these customers less likely to leave.

**12. Day since last order** The longer the time since the last order, the higher the churn rate. This underlines the importance of recency in predicting churn, indicating that targeting customers who haven't engaged recently might be a priority for retention efforts.

In summary, this analysis provides actionable insights for businesses to develop customer-centric strategies aimed at reducing churn. Factors like customer satisfaction, app engagement, delivery speed, and personalized incentives play significant roles in retaining customers and improving their long-term loyalty.

## INSIGHTS

The exploratory data analysis (EDA) provides a rich understanding of various factors impacting customer behavior and churn within the dataset.

**1. Gender Influence** There isn't a significant gender-based difference in average orders. However, males have a higher churn rate, suggesting the need to cater to their interests to reduce churn.

**2. Marital Status** Married customers make up the largest segment. Strategies should consider products appealing to both singles and married customers, as singles are more prone to churning.

---

**3. City Tier** While City Tier 2 has the highest tenure rate and City Tier 3 has the highest order average, City Tier alone doesn't appear to be a strong churn predictor.

**4. Satisfaction Score** Customers with lower satisfaction scores tend to spend less time on the app. However, there isn't a strong correlation between satisfaction score and app usage duration.

**5. Address Density** There's a negative correlation between city tier and number of addresses, indicating that address density varies by city type. Larger cities have fewer addresses on average.

**6. Complaints** There is a weakly negative relationship between complaints and the days since the last order. However, this doesn't strongly affect churn, as 68% of complainers don't churn.

**7. Device Preference** Mobile phone users are more likely to churn, suggesting potential issues with the mobile app's user experience.

**8. Warehouse Locations** Warehouse-to-home distances are consistent across city tiers, indicating that warehouses are strategically placed in lower tiers to serve customers.

**9. Product Categories** Laptops, accessories, and mobile phones are preferred across city tiers, with different payment methods based on city tier.

**10. Coupon Usage** More coupons are used to result in lower churn rates, indicating the effectiveness of incentives in retaining customers.

**11. Satisfaction and Order Count** The satisfaction score doesn't significantly impact the order count.

**12. Cashback and Churn** There is no significant relationship between cashback and order count, but there is a positive relationship with the churn rate.

**13. Complaints and Churn** Complaining doesn't strongly affect churn, as most complainers stay.

---

In summary, this EDA highlights various customer behavior factors, including gender, marital status, satisfaction, address density, and complaint impact. While some factors, like coupon usage, offer actionable insights for retention strategies, others, like satisfaction scores, are less predictive of churn. Understanding these relationships allows the company to focus on areas with the greatest potential for reducing churn and improving overall customer satisfaction and loyalty.

## Recommendations and Conclusion

Based on the insights derived from the exploratory data analysis (EDA), here are some key recommendations for the company to reduce customer churn and enhance overall customer satisfaction:

- 1. Target Male Customers** Since males exhibit a higher churn rate, it's essential to consider introducing products and services that cater to their specific interests. Conduct market research or surveys to better understand their preferences.
- 2. Address Single's Needs** Given that singles are more likely to churn, the company should pay attention to product offerings that resonate with both single and married customers. This approach ensures a broader appeal and helps retain a diverse customer base.
- 3. Product Pricing Strategy** In CityTier 2, where medium-tier customers are prevalent, it's crucial to offer products with pricing that reflects their purchasing power. Avoid extremely expensive or excessively cheap products, striking a balance that fosters trust and meets their needs effectively.
- 4. Enhance Customer Feedback Mechanisms** To supplement or even replace traditional satisfaction scores and complaints, consider introducing alternative feedback channels. A dedicated hotline or regular phone calls to collect feedback and resolve issues in real time can foster stronger customer relations.
- 5. Evaluate the App Experience** Investigate the mobile version of the app for potential user interface (UI) and user experience (UX) issues. Addressing any problems in the mobile app can significantly enhance customer satisfaction and reduce churn among mobile users.

---

**6. Target High-Value Customers** When customers have reached an order amount increase of 12%-15%, it's an opportune time to focus on capturing their attention with products that align with their preferences. Personalized offers and recommendations can be effective in this context.

**7. Cross-Selling Electronics** For customers who have recently purchased electronic goods, explore cross-selling opportunities by offering related electronic accessories such as keyboards, mice, headphones, and other complementary products. This strategy can boost the average order value and customer loyalty.

In conclusion, the data-driven insights gleaned from the EDA can guide the company in implementing targeted strategies to reduce churn and enhance customer satisfaction. By understanding the nuanced behavior of different customer segments and adapting product offerings and feedback mechanisms accordingly, the company can improve its customer retention efforts and overall business performance. These recommendations can serve as a foundation for a customer-centric approach that fosters lasting customer relationships and strengthens the company's competitive position.