

# Assignment 1 – Data Exploration

## Analysis Approach

To explore factors contributing to subscriber churn at StreamFlow, a structured analysis was conducted, starting with exploratory techniques, progressing to feature engineering and in-depth churn analysis.

**Initial Exploratory Data Analysis:** Descriptive statistics were calculated to understand structure of the dataset. Data types, particularly date fields, were formatted correctly, and missing values were checked to ensure data completeness.

**In-Depth Analysis:** Data was grouped to analyse churn rate based on location, reason for churn, genre preferences, and subscription type. Visualizations like correlation heatmaps, boxplots and scatterplots were used to uncover trends, correlation and outliers.

**Feature Engineering:** New features, 'loyalty levels', 'engagement levels' and 'satisfaction levels' were engineered, to further engineer another feature, 'customer types' more granular analysis. This segmentation helped develop personalized campaigns and retention strategies.

- Loyalty levels: Calculated using subscription start and churn dates, divided into quartiles.
- Engagement levels: Based on monthly hours watched, divided into quartiles.
- Satisfaction levels: Determined by user ratings, categorized into high (4-5), moderate (3), and low (1-2).

## Key Metrics

- Churn Rate (%): High churn rate is a sign of user dissatisfaction, impacting revenue negatively. Understanding churn under different parameters and while applying various filters helps pinpoint factors and areas of churn. Understanding such patterns will then support formulating retention strategies.
- Subscription Length: Reflects user loyalty and service value perception, essential for revenue stability and growth. It also helps in understanding what keeps the user engaged over time, and identify factors of retention in both short and long term. It can help develop retention strategies like loyalty programs, or rewards programs for loyal customers.
- Customer Life-Time Value: Higher CLV signifies long-term customer retention and profitability. CLV provides an insight into user's potential revenue contributions over time, spending patterns and behavior. It is imperative to enhance customer satisfaction, ensure continuous value addition and offer premium features when aiming to increase CLV.

## Findings and Insights

**Demographic Analysis:** Highest churn rates exist in East region due to better service offerings from competitors and price sensitivity. Central and North regions experience moderate to high churn, driven by competitive offerings and personal reasons, suggesting that enhancing service quality and personalized support can improve retention. The South and West regions show lower churn rates, but addressing content dissatisfaction in the West and personal reasons in the South can further enhance loyalty. Age-related analysis identified significant churn in the 26-35 and 36-45 age groups, especially in the East and Central regions. This is again driven by sensitivity to service quality and price. The 18-25 age group shows lower churn rates, indicating a stable segment, particularly in the South and West. The 56-65 age group in the North needs personalized support to reduce churn.

**Customer Type Analysis:** High churn rates are shown by customers classified as 'At-Risk Engaged' (4.7%) and 'At Risk Moderate' (3.8%), needing service quality improvements and personalized retention efforts. 'Critical' (3.1%) and 'Engaged Risk' (3.4%) customers also exhibit high churn due to low satisfaction and loyalty. Moderate users, such as Potential Stars and Steady Stars, present growth opportunities and can be converted into high-value customers. Critical users, like Critical and Disengaged Critics, require immediate attention to improve satisfaction. 'Super Stars' (12.3%) and 'Steady Stars' (6.3%) show high engagement, satisfaction, and loyalty, suggesting that rewarding these customers can maintain their retention. 'Loyal but Unhappy' (4.6%) and 'Frustrated' (4.8%) customers, despite being engaged and loyal, are at risk of future churn if their dissatisfaction is not addressed. High engagement significantly reduces churn, with only 11.2% of highly engaged customers churning. High satisfaction correlates with lower churn (8.1%), while low satisfaction increases churn to 17.0%. High loyalty almost eliminates churn, with 0% churned, whereas low loyalty correlates with a 3.3% churn rate. Boosting engagement, satisfaction, and loyalty through quality improvements, personalized support, and loyalty programs is crucial.

**Customer Life-Time Value Analysis:** The analysis reveals significant variation in Customer Lifetime Value (CLV), with a mean of \$465.61 and a high standard deviation of \$332.91. Approximately 50% of customers have a CLV below \$427.08, highlighting the need for targeted strategies.

**Standard Tier Scrutiny:** In the 'Standard' subscription tier, the East region shows the highest churn due to dissatisfaction with content and price increases. The 26-35 and 36-45 age groups exhibit the highest churn rates, particularly among the 'At Risk Moderate' and 'At-Risk Engaged' customer types. For the 18-25 age group, churn is relatively low across all customer types. The 56-65 age group in the North shows significant churn, indicating a need for personalized support.

## Recommendations to the Business

1. **Conduct Competitive Analysis:** StreamFlow must conduct a thorough competitive analysis specific to each region to understand service offerings and customer feedback. Feedback should be incorporated to adjust service offerings and pricing strategies.
2. **Personalized Engagement Programs for Younger Age Groups:** Younger age groups are critical, StreamFlow should implement targeted engagement programs and risk mitigation strategies for the 26-35 and 36-45 age groups, particularly in the East and Central regions, to address their sensitivity to service quality and price.
3. **Support for Older Age Groups:** Personalized support programs should be developed for the 56-65 age group in the North. Easy-to-access customer service channels and tailored assistance are some of the initiatives that can help improve retention rates.
4. **Address Dissatisfaction Among Engaged and Loyal Customers:** Implement a dedicated support team and develop personalized action plans for 'Loyal but Unhappy' and 'Frustrated' customers. Enhance communication channels and schedule regular follow-ups to ensure prompt and effective resolution of their issues. Implement targeted strategies to increase the engagement and satisfaction of customers below the median CLV of \$427.08, aiming to boost their lifetime value.
5. **Reward Loyal/High-Value Customers:** StreamFlow must recognize and reward 'Super Stars' and 'Steady Stars' with loyalty programs and exclusive benefits. Maintaining their retention by ensuring these programs are attractive and provide real value, will prove to be critical for the company's long term profitability and CLV.
6. **Content and Service Improvement:** Regularly update and improve content, with particular focus on popular genres or genres with higher dissatisfaction like Sci-Fi and Action, to stay competitive and meet user expectations. Moreover, improve platform experience for Mobile users, as data shows that by device usage, most users and hence most churn comes from to Mobile phone users, with the most popular reason being 'Found better service'.

## Business Impact Estimate

**Qualitative Benefit:** Improved customer satisfaction and loyalty will enhance the brand reputation and foster customer advocacy. Personalized support and targeted engagement programs will strengthen relationships, making customers feel valued and more likely to remain with the service. Addressing dissatisfaction will turn negative experiences into positive ones, further enhancing brand perception and trust, all of which

**Quantitative Benefit:** With a conservative approach, we aim to reduce the churn rate by 7% during the next year.

- **Churn Reduction:** By reducing churn by 7%, StreamFlow is expected to retain an additional 70 customers, leading to an increase in monthly revenue by approximately \$1,049.30 and an annual revenue increase of approximately \$12,591.60.
- **Customer Lifetime Value (CLV) Increase:** With a 7% increase in CLV, the new average CLV will rise from \$465.61 to approximately \$498.21. This results in an additional total CLV value of approximately \$32,592.85 for all subscribers.
- **Overall Financial Benefit:** The combined effect of reduced churn and increased CLV translates into substantial financial benefits, enhancing both immediate revenue and long-term customer value, ultimately driving business growth and sustainability.

The recommendations are supported by data from customer churn analysis and insights. Reducing churn directly correlates to increased revenue and profitability due to lower customer acquisition costs and higher CLV. Enhanced customer satisfaction and loyalty also contribute to word-of-mouth referrals, reducing marketing expenses. These economic benefits reflect a deep understanding of the business's financial dynamics and the significant positive impact of improved customer retention strategies. By leveraging the insights and implementing targeted retention strategies, StreamFlow can expect substantial qualitative and quantitative benefits, ultimately driving business growth and sustainability.