MBAN 6110: DATA EXPLORATION WITH STREAMFLOW

MBAN 6110 – Section T

Christine Tang

Student ID: 214-151-401

Instructor: **Delina Ivanova**

May 28, 2024

Executive Summary

The analytical approach consisted of <u>six</u> steps: (1) Load data and view descriptive statistics, (2) Identify variables and validating logic, (3) Analyze and treat missing data, (4) Analyze and treat outliers, (5) Conduct data cleaning, feature-engineering, and encoding, and (6) Utilize variables for statistical analysis. Key metrics to investigate churn and potential reasons for downgrading plan ("Churned", "Monthly_Hours_Watched", "Average_Session_Length", "Subscription_Duration_days", "Lifetime_Value", "Encoded_Subscription_Type_Basic") were created and evaluated with statistical tests. Important insights were discovered, regarding (1) Device experiences negatively affecting Mobile users, (2) Content Catalog opportunities to explore with genre and location, and (3) Exploring price and promotional activities that may have happened in the past. Recommendations focus on improving device experience, marketing opportunities targeting different types of viewers, and examine past activities and potential partnerships considerations.

The Analytical Approach to StreamFlow Data

For the missing 754 rows of data from "Date_of_Churn" and "Reason_for_Churn", the nature of the missing data was missing not at random (MNAR), as there is a systematic, logical pattern. The missing rows represented existing subscribers, thus, the treatment was to not modify the missing rows data, to avoid introducing bias or affecting data structure. The remaining rows were analyzed separately to understand churn. Regarding outliers, two continuous variables ("Monthly_Hours_Watched" and "Average_Session_Length") had outliers identified and the treatment was to keep outliers, to avoid introducing bias or affecting the data structure. Winsorization transformation attempts were completed, however, outliers were left in the data. Pandas Date-time variables were created to calculate subscription durations and encode "year" and "quarter, creating groupings for "Age", "Monthly_Hours_Watched", "Average_Session_Length". Categorical variables were converted using one-hot encoding for statistical testing. Various charts were utilized (i.e. boxplot, histograms, bar charts) were utilized to look at distributions and relationships with other variables. Statistical testing (i.e. Chi-Squared, T-Test, ANOVA and Pearson-R) was completed to look at significance relationships.

Key Metrics

To understand churn, the following metrics were utilized: "Churned", "Monthly_Hours_Watched", "Average_Session_Length", "Subscription_Duration_days", and "Lifetime_Value". This understanding informs StreamFlow about controllable factors that affect retention efforts. "Subscription_Duration_days" is used to determine factors that influence how long a user to stay or churn. "Lifetime_Value" is utilized to create a revenue estimation, understanding the value (or lost revenue opportunity) of specific groups, setting targets for retention. A secondary concern was downgrading to basic plan; one column ("Encoded_Subscription_Type_Basic") was analyzed with other metrics to understand what factors influence users to select the basic plan versus other plans.

Summary of Insights and Findings

The analysis derived three important findings. Firstly, when comparing Devices Used vs. Churned, there is a statistically significant relationship that indicates the proportion of "Mobile" users who churn is

higher than other groups, and "PC" users have the lower churn. "Tablet" users had the highest mean vs. non-Tablet users for subscription duration and lifetime value. Secondly, there is opportunity for StreamFlow to analyze content catalog differences for location and genre. The analysis revealed that, for churned users, "documentary" fans were more likely to rate "Content Dissatisfaction" as a reason. "East" users and "Comedy" fans have a lower proportion of having basic plans vs. other regions. Finally, there are implications based on pricing and promotions that affect Churn rates. It was analyzed that mean subscription duration for StreamFlow users is ~2.1 years, aligning with the finding that there is a statistically significant relationship between Churn Year 2024 and churn reason for "price increase". "Above-average" viewers (~1.9-3.0 hours/session) had a higher proportion of 2022 subscription. Finally, there is a relationship between payment type and the viewer type, with "Above-average" viewers using PayPal for payment.

Recommendations for StreamFlow

From "Device Experience" findings, StreamFlow should conduct user-experience (UX/UI) studies to understand why Mobile users churn at a higher proportion versus other device type. Recent examples include Apple TV+, with noticeable button issues affecting user experience (Diaz, 2023). PC users had lower churn; aspects such as search, or video issues should be examined for mobile users. From "Content Catalog" findings, StreamFlow should identify and analyze regional differences for content offerings and consider marketing opportunities to highlight these differences to improve churn. For non-East users, there could be marketing promotions to highlight the better content in StreamFlow's services. Some additional market research should be conducted for East users, to understand why these users typically do not subscribe to basic plans vs other plans. It should be investigated if the catalog for "documentary" content was recently reduced, contributing to churn. Content acquisition efforts should be implemented to bring documentary content, and marketing campaigns to highlight newer content. From "Pricing and Promotions" findings, StreamFlow should track user who joined with discounts, and consider implementing efforts to communicate StreamFlow's value. Finally, there's corporate partnership opportunities to encourage plan upgrades and reduce churn. For example, "Above-Average" viewers should be incentivized to upgrade due to the benefits and reduce churn. These efforts will help reduce churn and switching users to other plans.

Business Impact Estimation

Mobile experience improvements will improve customer satisfaction and encourage positive word-of-mouth for StreamFlow. Since the average churn of SVOD service is ~37% in the United States (Stoll, 2023). However, StreamFlow's churn rate is 24.6%, which is lower. Therefore, a target of 20% would be applied. The "lifetime value" is $2351/year; achieving at least 20% less churn for this group will generate $470/year. Marketing efforts for non-East users to upgrade plans is valued at $2340/year, so achieving 20% less churn will generate $468/year. Finally, partnering with other organizations will increase brand awareness and credibility of StreamFlow to all users, and offering bundled services with financial services like PayPal will likely discourage users from leaving. An example marketing campaign targeting "Above-average" users who use PayPal has a lifetime value of $2702 USD/year from preventing churn, with 20% less churn target of $540.

Works Cited

Diaz, Jesus. *From Netflix to HBO, the Terrible Design of Streaming Is ...*, Fast Company, 23
    May 2023, www.fastcompany.com/90882683/netflix-hbo-terrible-design-streaming.

Stoll, Julia. "SVOD Service Churn Worldwide by Generation 2022." *Statista*, Statista, 15 May
    2023, www.statista.com/statistics/1306198/churn-paid-streaming-worldwide-generation/.