Data Analysis Report: Streaming Service User Behavior

Introduction

This analysis aims to provide insights into user behavior, content consumption, and the overall performance of a streaming service, focusing on several factors such as genre popularity, age demographics, subscription types, and device usage. The findings will be instrumental in guiding marketing strategies for the streaming service.

Dataset Overview

The dataset consists of 6214 entries with the following columns:

- **User_ID**: Unique identifier for the user.
- **Session_ID**: Unique identifier for the user's session.
- **Device_ID**: Identifier for the user's device.
- **Video_ID**: Identifier for the video content being viewed.
- **Duration_Watched (minutes)**: Time spent watching the video (in minutes).
- **Genre**: Genre of the video content (e.g., Action, Comedy, Drama).
- **Country**: Country of the user.
- **Age**: Age of the user.
- **Gender**: Gender of the user (e.g., Male, Female).
- **Subscription_Status**: User's subscription status (Free, Premium).
- **Ratings**: User's rating for the content (1 to 5 scale).
- **Languages**: Language of the content being viewed.
- **Device_Type**: Type of device used (e.g., Smartphone, Tablet).
- **Location**: Location where the interaction event occurred.
- **Playback_Quality**: Quality of video playback (e.g., HD, SD, 4K).
- **Interaction_Events**: Count of interaction events during the session (e.g., clicks, likes).

Data Exploration and Summary

Dataset Structure

- The dataset contains 17 columns, including 10 categorical variables and 6 numerical variables.
- There are no missing values in the dataset.

Key Observations

1. **Genre Popularity**: The distribution of genres is as follows:
2. Documentary: 1065
3. Thriller: 1050
4. Sci-Fi: 1041
5. Comedy: 1033
6. Action: 1030

7. Drama: 995

8. **User Demographics**:

9. **Age**: The average age of users is approximately 39.67 years, with a standard deviation of around 17.57 years.
10. **Gender**: The dataset reflects a fairly balanced gender distribution; 3148 Female (50.6%) vs. 3066 Male (49.4%).

11. **Subscription Status**: The user base is evenly split between Free (3111) and Premium (3103) subscriptions.

12. **Device Usage**:

13. The predominant device types used are:

    - Smart TV: 1306
    - Desktop: 1270
    - Smartphone: 1233
    - Tablet: 1217
    - Laptop: 1188

14. **Playback Quality**:

15. The distribution of playback quality is fairly even among 4K, HD, and SD resolutions.

16. **User Engagement**:

17. Average duration watched by genre indicates that users tend to spend the most time on "Action" content (43.52 minutes) followed closely by "Comedy" (42.91 minutes).
18. Ratings by genre reveal that "Action" has the highest average rating (3.04) while "Documentary" has the lowest (2.93).

Grouped Analysis

1. **Duration Watched by Age**: Average duration watched varies by age, suggesting different engagement levels at different ages.
2. **Ratings by Genre and Subscription Status**: Premium users tend to rate content slightly higher in most genres compared to Free users.
3. **Interaction Events by Device**: Smart TVs and Desktops show higher interaction events compared to other devices, suggesting that these platforms may warrant more focus in marketing strategies.

Correlation Analysis

A correlation analysis indicates minimal relationships among variables, with slight correlations between age and duration watched, and ratings and interaction events.

Recommendations for Marketing Strategies

1. **Focus on Popular Genres**: Prioritize marketing campaigns for genres such as Action and Comedy, especially targeting Premium subscribers who show longer engagement times.
2. **Tailored Campaigns for Younger Audiences**: Create content and promotions aimed at younger demographics who exhibit higher engagement with specific genres.
3. **Optimize User Experience on Common Devices**: Enhance streaming quality and user experience on Smart TVs and Desktops, as these are the most used devices.
4. **Regional Marketing Strategies**: Develop marketing initiatives based on genre preferences and user interactions in specific countries.
5. **Highlight Top-Rated Content**: Use high user ratings in promotional materials to attract new users and retain existing subscribers.

Conclusion

The analysis provides actionable insights into user behavior and preferences for a streaming service. By leveraging these insights, the company can enhance its marketing strategies to better engage existing users and attract new ones, ultimately leading to increased user satisfaction and retention.