Netflix User Behaviour Data Analysis Report

1. Dataset Description

1.1 Source: Kaggle Netflix users dataset (users.csv) – 10,300 records.

1.2 Columns:

Column Name	Data Type	Description
first_name	String	User's first name
last_name	String	User's last name
age	Integer	User's age
gender	String	Male/Female
country	String	User's country
state_province	String	State or province
city	String	City
subscription_plan	String	Type of subscription (Basic, Standard, Premium+)
subscription_start_date	Date	Date subscription started
is_active	Boolean	Whether the user is currently active
monthly_spend	Float	Amount spent per month
primary_device	String	Device used (Laptop/Desktop)
household_size	Integer	Number of people in the household
created_at	String	Account creation timestamp

1.3 Data Quality:

- Missing values exist in some columns (e.g., age, monthly_spend).
- Duplicate users were detected in a few name/email entries.
- Outliers found in age (under 10, over 90) and monthly_spend (very high values beyond Premium+).
- Overall dataset is large, diverse, and representative across multiple countries.

---

2. Operations Performed

2.1 Data Cleaning & Exploration

- Checked and handled missing/null values.
- Removed/flagged duplicate entries.
- Summarized numerical columns: mean, median, standard deviation, interquartile range (IQR).
- Identified outliers in age and monthly_spend.

2.2 Descriptive Analytics

- Gender distribution (pie chart).
- Subscription plan distribution (bar chart).
- Device usage distribution (pie chart).
- Country distribution of users (bar chart).
- Age distribution (histogram, boxplot).
- Monthly spend distribution (histogram, density plot).

2.3 Relationship Analysis

- Age vs. Monthly Spend (scatter plot).
- Subscription Plan vs. Monthly Spend (boxplot).
- Gender vs. Device usage (stacked bar).
- Household size vs. Monthly Spend (line/boxplot).
- Country vs. Subscription Plan distribution (stacked bar).

---

3. Key Insights

3.1 Demographics

- Age distribution clusters between 20–40 years → dominant young adult audience.
- Average user age: ~32 years.
- Very few users above 60, indicating limited senior adoption.
- Household size typically ranges between 1–4, aligning with family subscriptions.

3.2 Subscription Trends

- **Standard plan** has the largest share of users, followed by **Premium+**, then **Basic**.
- Premium+ adoption is higher among larger households (3+ members).
- Active users make up ~82% of the dataset, showing strong engagement.

3.3 Spending Insights

- Monthly spend range: $5 – $100, with most clustering between $10–40.
- Average monthly spend: ~$27.
- Outliers exist with spend >$80 (possibly errors, bundles, or corporate accounts).
- Higher household sizes correlate with higher spending.

3.4 Device Usage

- **Laptop and Mobile** are the dominant devices, with TV usage significant in family households.
- Desktops are least common among users.
- Younger users prefer Mobile, while older users lean towards TV and Laptop.

3.5 Geographic Spread

- USA leads in user base, followed by UK, India, and Canada.
- Regional variations exist:
    - USA/UK – higher Premium+ adoption.
    - India – higher Standard plan adoption.
- State-level variation shows concentration in urban/province capitals.

---

4. Recommendations

4.1 Targeted Engagement

- Launch senior-friendly plans to attract users above 55.
- Create family-oriented bundles for households with 4+ members.

4.2 Pricing & Plan Strategy

- Investigate outlier spends >$80 – check if genuine or billing errors.
- Expand Standard plan features since it attracts the majority of users.
- Introduce localized pricing for emerging markets like India.

4.3 Device Optimization

- Enhance mobile streaming quality since mobile usage dominates younger users.
- Improve Smart TV integration for family households.
- Optimize desktop/web experience for office-time users.

4.4 Geographic Expansion

- Invest in local content for high-user countries (India, USA).
- Encourage growth in underrepresented markets by offering regional promotions.

4.5 Future Analytics Opportunities

- Build churn prediction models using age, spend, and activity.
- Segment users by household size, spending, and device type for marketing campaigns.
- Track subscription upgrades/downgrades over time to optimize pricing strategies.