DATA ANALYTICS REPORT

EXPLORATORY DATA ANALYSIS OF NETFLIX USER SEGMENTATION & BEHAVIOR EDA

User Behaviour, Purchase Patterns & Lifetime Value

NETFLIX

TOOLS USED

Python, Pandas, Matplotlib, Seaborn & Plotly

PRESENTED BY

Tejas Jadhav

GITHUB

TABLE OF CONTENTS

PROBLEM STATEMENT & OBJECTIVE

Problem Statement:

Netflix serves a diverse global audience with varying engagement levels, subscription choices, and content preferences, making it critical to identify the key drivers of sustained watch time, premium adoption, and long-term user value.

Objective:

To explore the Netflix user dataset to uncover patterns in user behavior, engagement intensity, subscription adoption, and content preferences that can inform data-driven decisions around engagement growth, premium monetization, and long-term user value.

EDA WORKFLOW

For this analysis, a structured workflow was followed, involving data collection, understanding, cleaning, exploration, and summarization of insights to allow a clear understanding of the dataset and its trends.

01 Data Collection

- Gathered the Netflix Dataset from Kaggle

02 Data Understanding & Anomaly Detection

- Looked at data distributions
- Found missing values, outliers and unusual patterns

03 Data Cleaning & Treatment

- Fixed missing or incorrect values
- Standardized formats for consistency

04 Exploratory Analysis

- Univariate Analysis: e.g Distribution of Age, Users by Subscription, Top 10 Genres etc
- **Bivariate:** Subscription Type vs Average Watch Time, Age Group vs Favorite Genre etc
- Multivariate Analysis: Watch Time Hours by Age & Subscription Type

05 Insights & Reporting

- Summarized patterns and trends
- Highlighted key findings for developers and businesses

KEY QUESTIONS EXPLORED

- How does geographic distribution influence user scale versus engagement depth across markets for Netflix?
- Does subscription tier meaningfully impact watch time, or is engagement largely independent of pricing level?
- Which age segments contribute most to total viewing hours and sustained platform engagement?
- To what extent is watch time concentrated among a small group of highly engaged users versus the broader user base?
- How evenly are content preferences distributed across genres, and does any genre act as a primary engagement driver?
- Are users upgrading subscription tiers in response to higher content consumption, or due to feature-based value perception?
- Which user segments represent the strongest opportunities for targeted upsell and retention strategies?
- How does engagement vary over time, and are there identifiable periods of peak and plateaued usage?

DATA OVERVIEW

The dataset provides insights into Netflix user behavior, covering demographics, subscription types, content consumption patterns, and recency of platform usage. It captures key engagement signals such as watch time and genre preferences, which are essential for understanding viewing intensity, subscription value, and long-term user engagement.

Data Source: Kaggle

Dataset Size

25,000
Records

8
Features

Purchase Diversity

3
Subscription Type

10
Countries

7
Genre

DATA OVERVIEW

Below is a detailed description of the feature set:

Dataset Features	Type	Feature Description
User_ID	Numerical (Discrete)	Unique identifier assigned to each Netflix user
Name	String	User name or anonymized identifier
Age	Numerical (Discrete)	Age of the user in years
Country	Categorical	Country from which the user accesses Netflix
Subscription_Type	Categorical	Subscription plan type (Basic, Standard, Premium)
Watch_Time_Hours	Numerical (Continuous)	Total number of hours spent watching content on the platform
Favorite_Genre	Categorical	User's most frequently watched content genre
Last_Login	Date / Time	Timestamp of the user's most recent platform activity

DATA QUALITY CHALLENGES & ANOMALIES

Few inconsistencies were found in the dataset, which could have affected the analysis if left unaddressed.

DATA ANOMALIES

- The Last_Login column is stored as a string and needs conversion to a datetime format to enable time-based engagement analysis.
- Subscription_Type, Country, and Favorite_Genre are categorical in nature and should be explicitly stored as categorical variables for efficient segmentation and aggregation.

DATA CLEANING & TREATMENT

Inconsistencies were addressed, and key features were cleaned and standardized for analysis.

DATA CLEANING SUMMARY

- Categorical fields such as Country, Subscription_Type, and Favorite_Genre were converted to appropriate categorical data types,
- Last_Login was standardized to a datetime format to enable accurate time-based engagement analysis

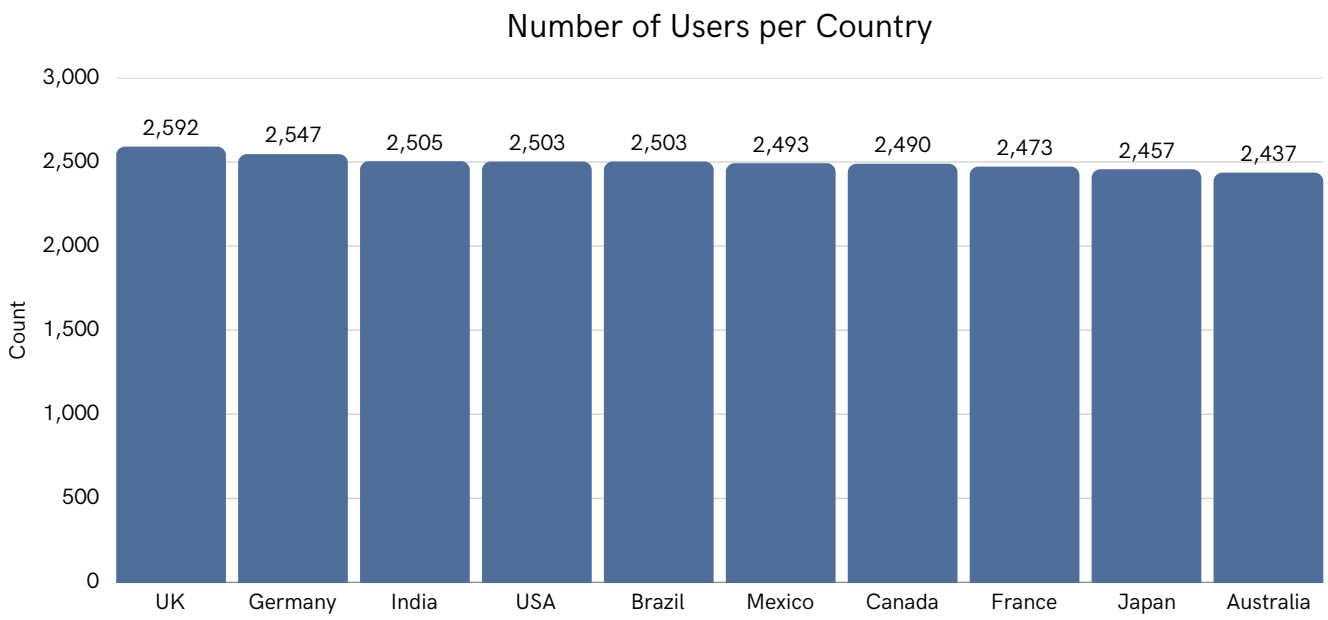
INSIGHTS

Customer Profile & Engagement

User distribution is evenly balanced across countries, with no single market dominating the platform

44.47%

User Traffic is from the United States



Key observations

- User counts are tightly clustered across all countries, showing minimal variation in market size.
- No country stands out as a clear leader or laggard in total user representation.

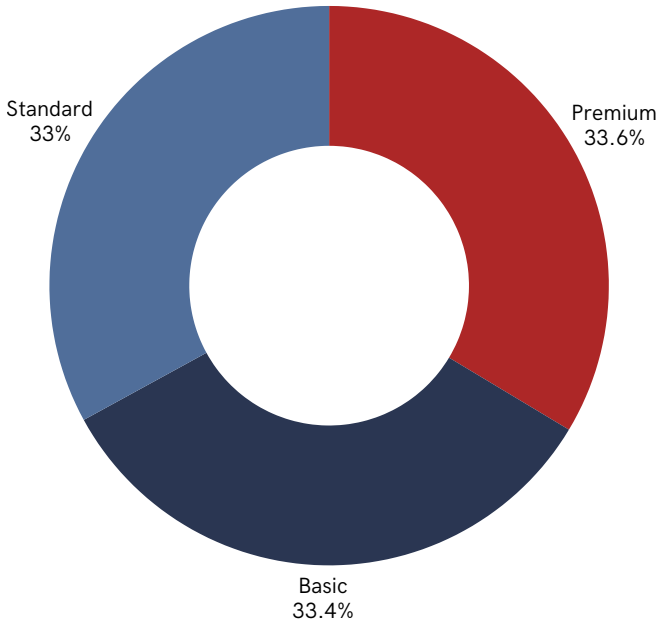
Business Insights

- Netflix benefits from a well-diversified global user base, reducing reliance on any single geography.
- With user scale largely maxed out across regions, future gains are more likely to come from higher engagement and premium tier adoption rather than user growth.

Subscription adoption is evenly distributed, indicating no single pricing tier dominates user preference

58.20%

of User Sessions come from Desktops



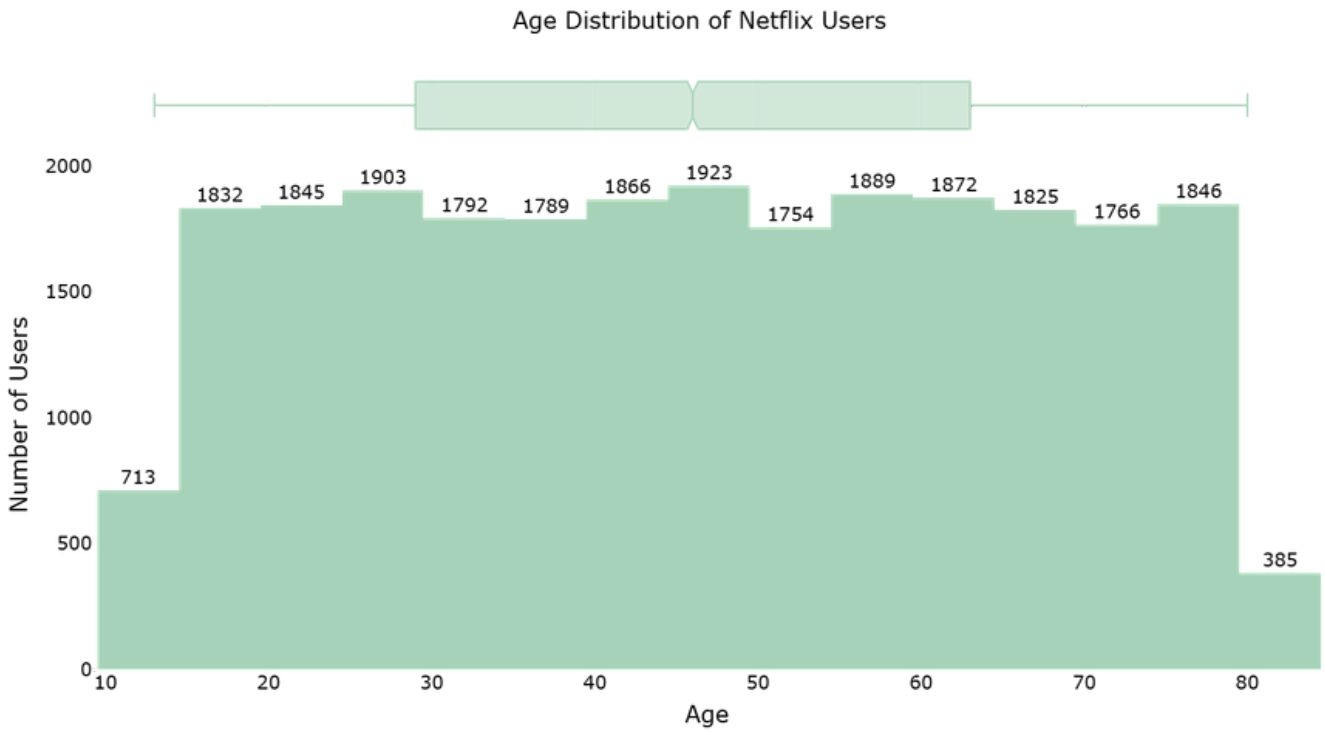
Key observations

- Premium, Basic, and Standard plans have nearly identical user counts, with only marginal differences between tiers.
- No subscription type shows a clear adoption advantage over the others.

Business Insights

- Netflix has achieved a well-balanced subscription mix, reducing dependency on any single pricing tier.
- Revenue growth will likely depend more on driving upgrades and increasing engagement per user than shifting users across plans.

Netflix engagement spans a broad age range, with strongest representation among working-age users



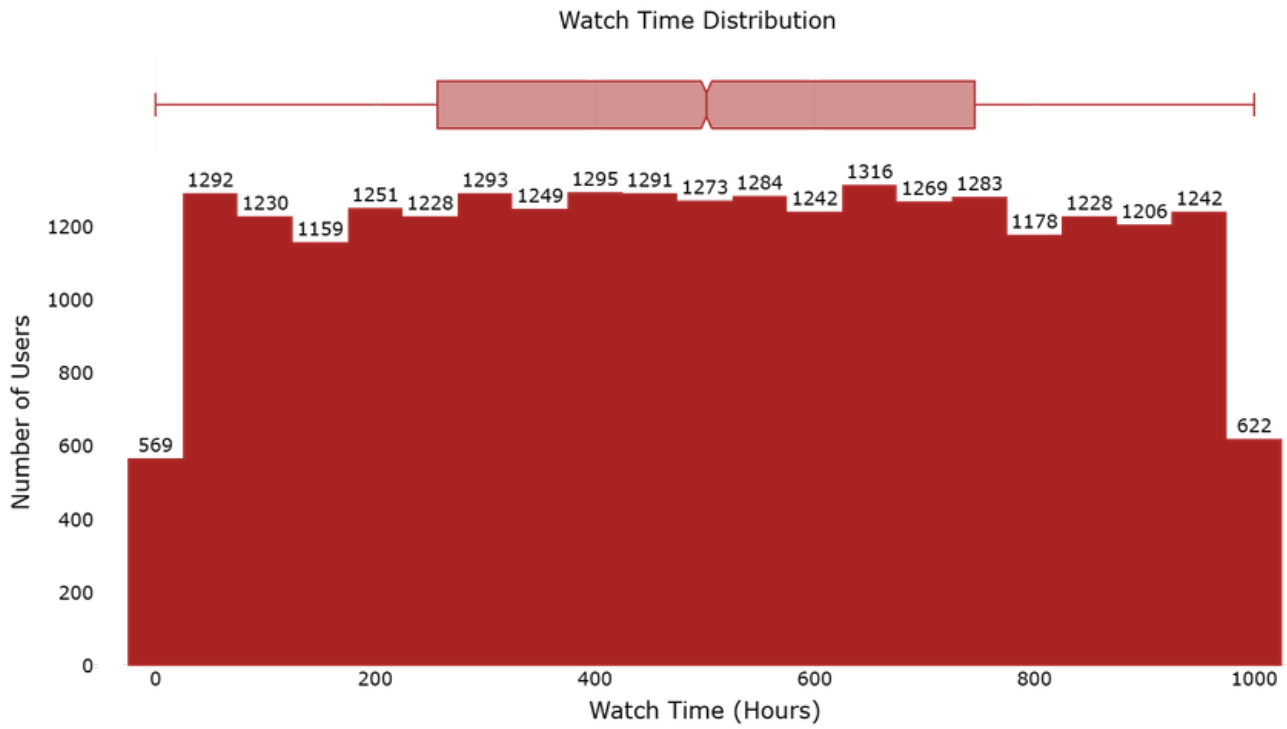
Key observations

- Users are distributed across a wide age spectrum, with higher concentration between early adulthood and late middle age.
- Very young and very old age groups form a noticeably smaller share of the user base.

Business Insights

- Netflix primarily serves a broad working-age audience, supporting diversified content investments across genres.
- Growth opportunities exist in age-tailored content and experiences to deepen engagement at the younger and older ends of the spectrum.

Watch time is widely dispersed, indicating the coexistence of casual viewers and highly engaged power users



Key observations

- Watch time spans a very broad range, with users distributed across low, medium, and extremely high viewing hours.
- The absence of a sharp peak suggests no single "typical" consumption level dominates the platform.

Business Insights

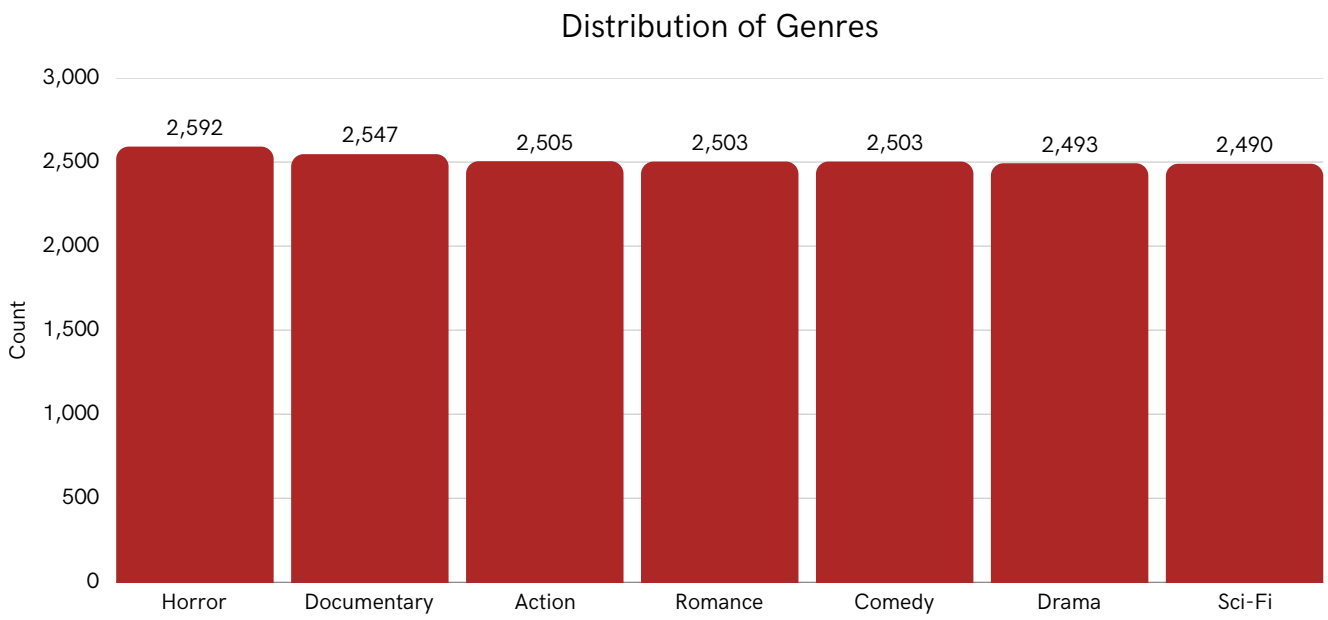
- Netflix serves both light and heavy users, with a subset likely driving a disproportionate share of total viewing hours.
- Deepening engagement among mid-tier viewers could yield meaningful gains, while protecting high-usage users remains critical.

Content Preference & Consumption

User preferences are evenly spread across genres, with no single content category dominating demand

38.85%

Of the Interactions is driven by Apparel



Key observations

- The top genres show very similar user counts, indicating broadly balanced content consumption.
- Even traditionally niche genres maintain comparable popularity to mainstream categories.

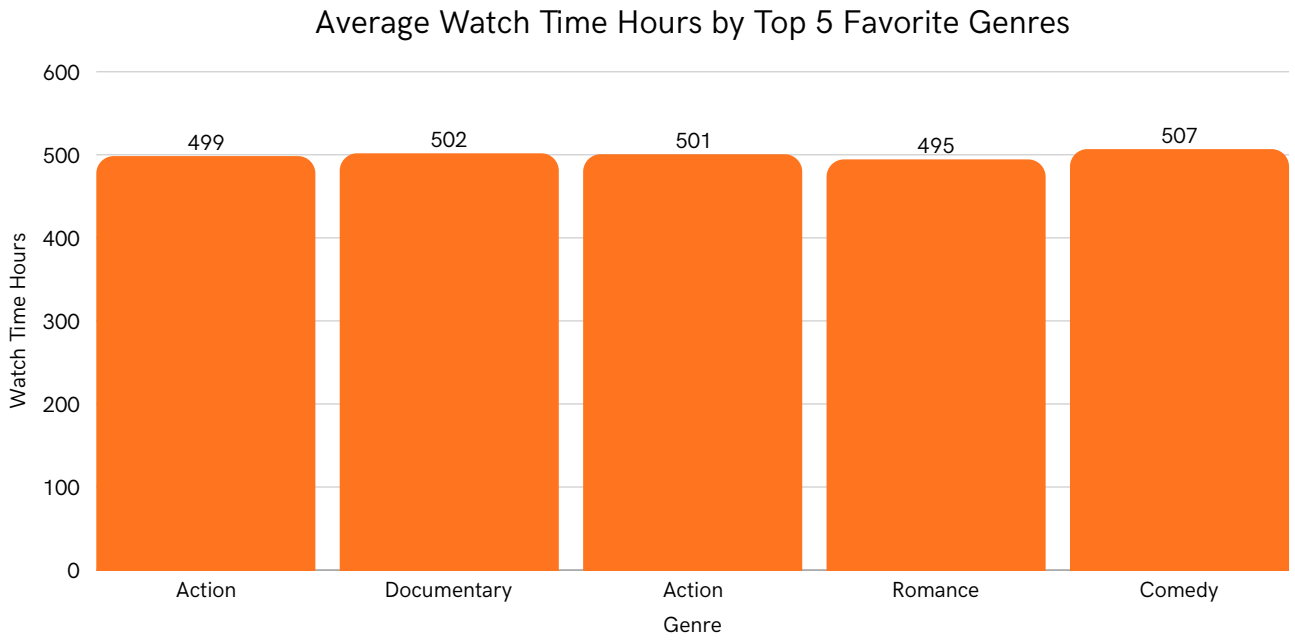
Business Insights

- Netflix benefits from a diversified content portfolio rather than reliance on a single genre.
- Content investment should remain balanced across genres, with personalization driving discovery rather than over-concentration in any one category.

Watch Time Remains Consistent Across Top Genres, Indicating Broad-Based Content Engagement

76%

Of Users search for Google related products



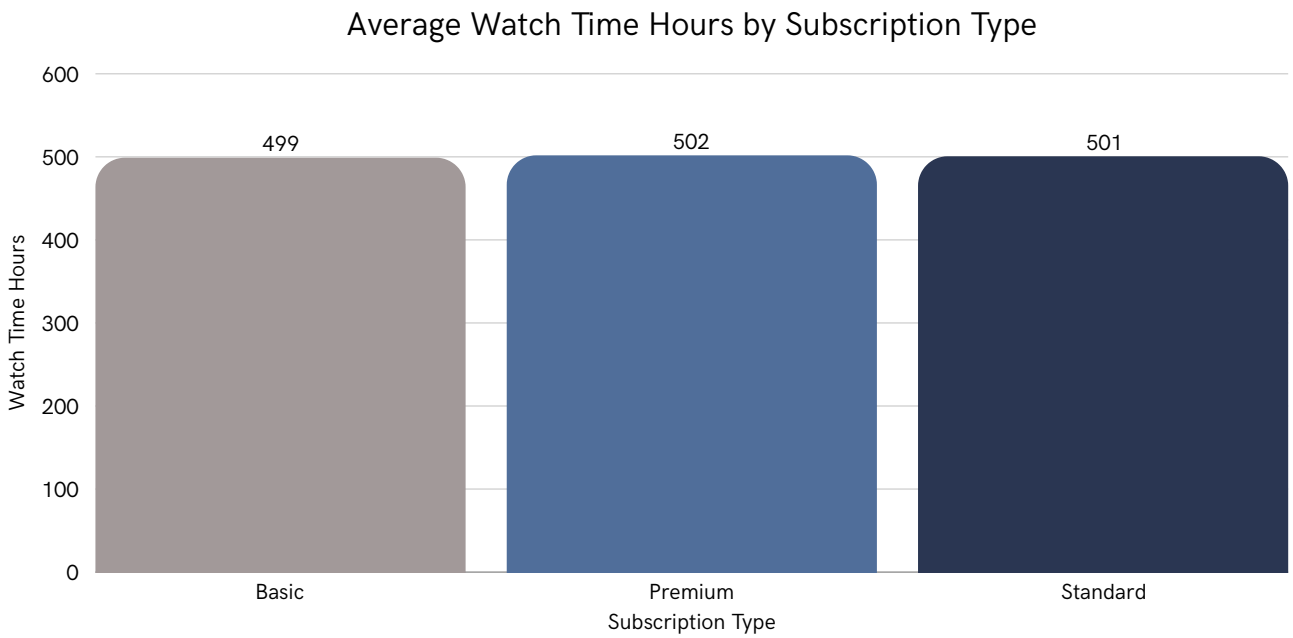
Key observations

- Average watch time across Action, Comedy, Documentary, Horror, and Romance is nearly identical (~495–507 hours).
- No single genre significantly outperforms others in terms of engagement depth.

Business Insights

- Engagement appears platform-driven rather than genre-dependent, suggesting strong overall content stickiness.
- Investment strategy should prioritize content quality and volume across genres rather than over-concentrating on one category.

Engagement Levels Are Nearly Identical Across Subscription Tiers, Limiting Behavioral Differentiation



Key observations

- Average watch time is very similar across Basic (~503 hrs), Premium (~501 hrs), and Standard (~497 hrs) plans.
- Higher-tier subscriptions do not show materially higher engagement compared to lower-tier plans.

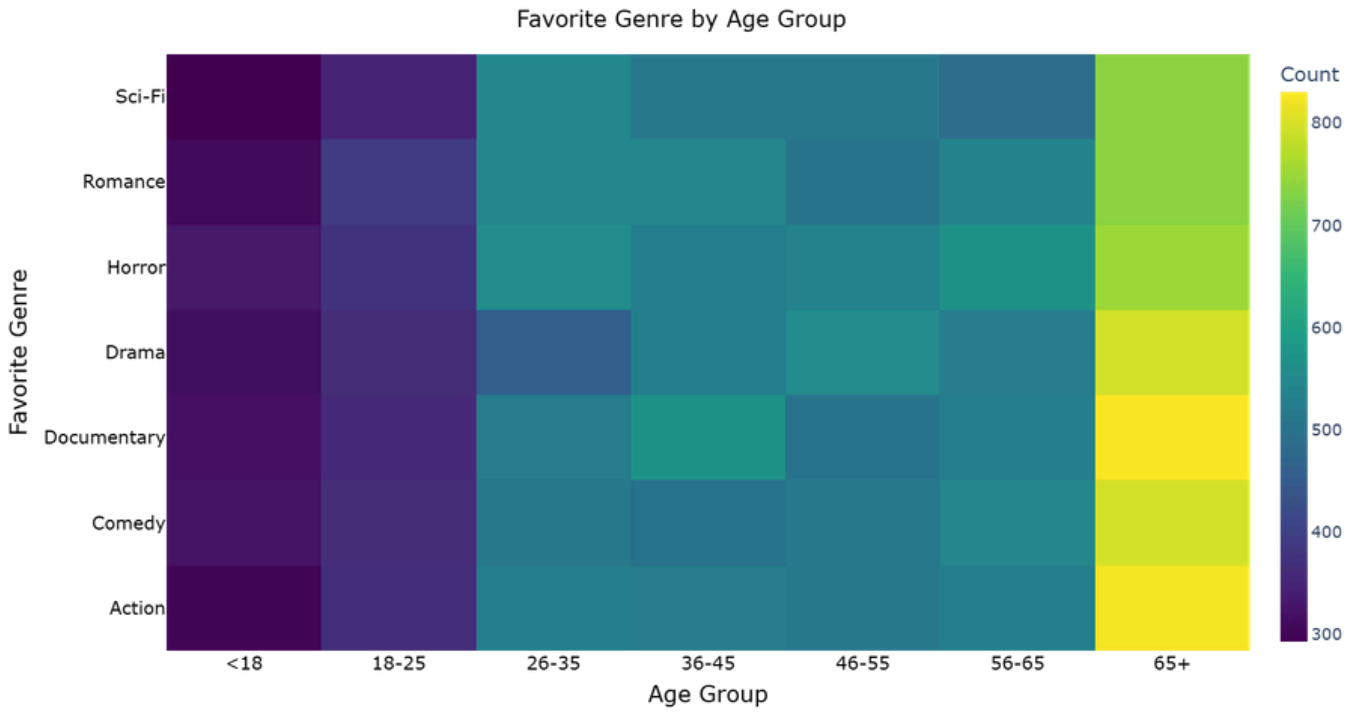
Business Insights

- Users are not upgrading for increased consumption, suggesting value perception may be driven by features (quality, screens) rather than usage.
- Upsell strategies should emphasize premium benefits beyond content volume to justify tier differentiation.

Content Preferences Shift with Age, with Older Users Showing Significantly Higher Genre Engagement

$22

Typical Price (Median)



Key observations

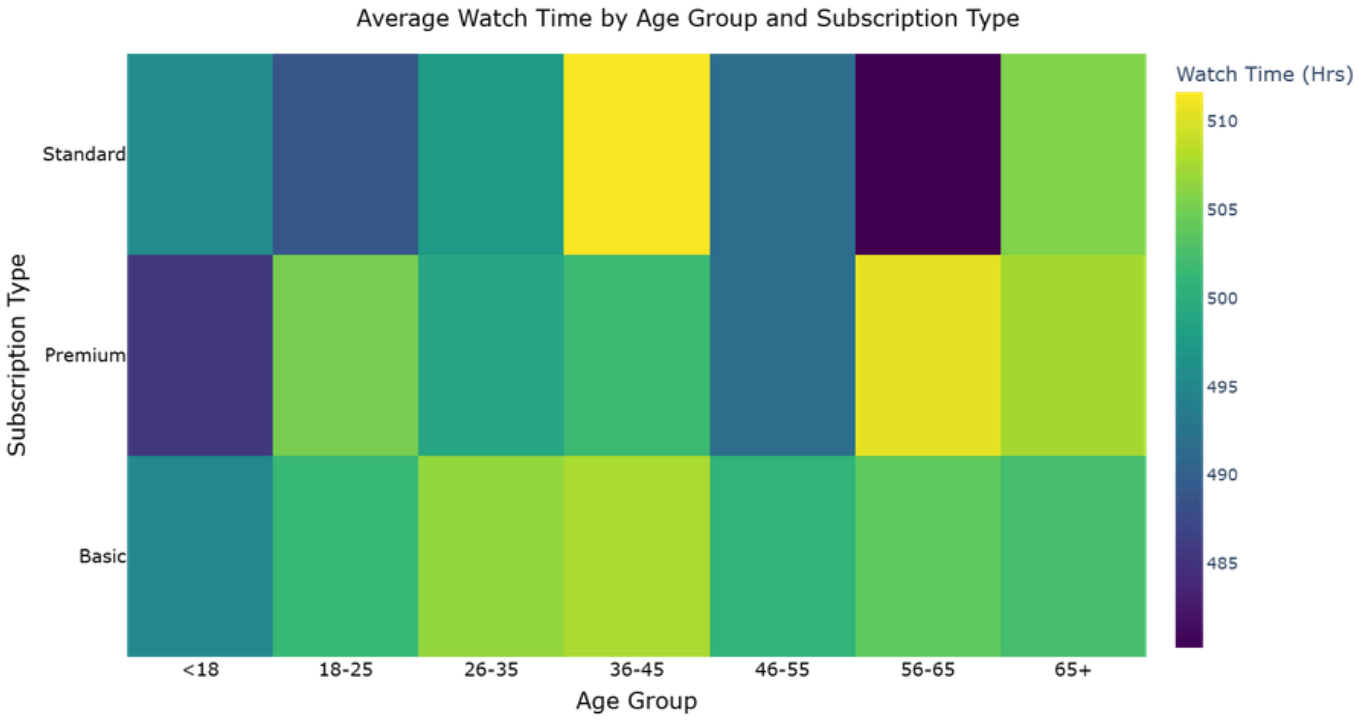
- Users aged 65+ show the highest engagement across nearly all genres, significantly exceeding younger age groups.
- Younger segments (<25) display comparatively lower genre counts, with gradual increases across middle age groups.

Business Insights

- Older users represent a highly engaged and potentially high-retention segment, warranting tailored content and marketing strategies.
- Growth opportunities may lie in increasing engagement among younger audiences through targeted, trend-driven content.

Engagement Value & Premium Users

Premium and Standard Subscribers Aged 18-45 Drive the Highest Average Watch Time



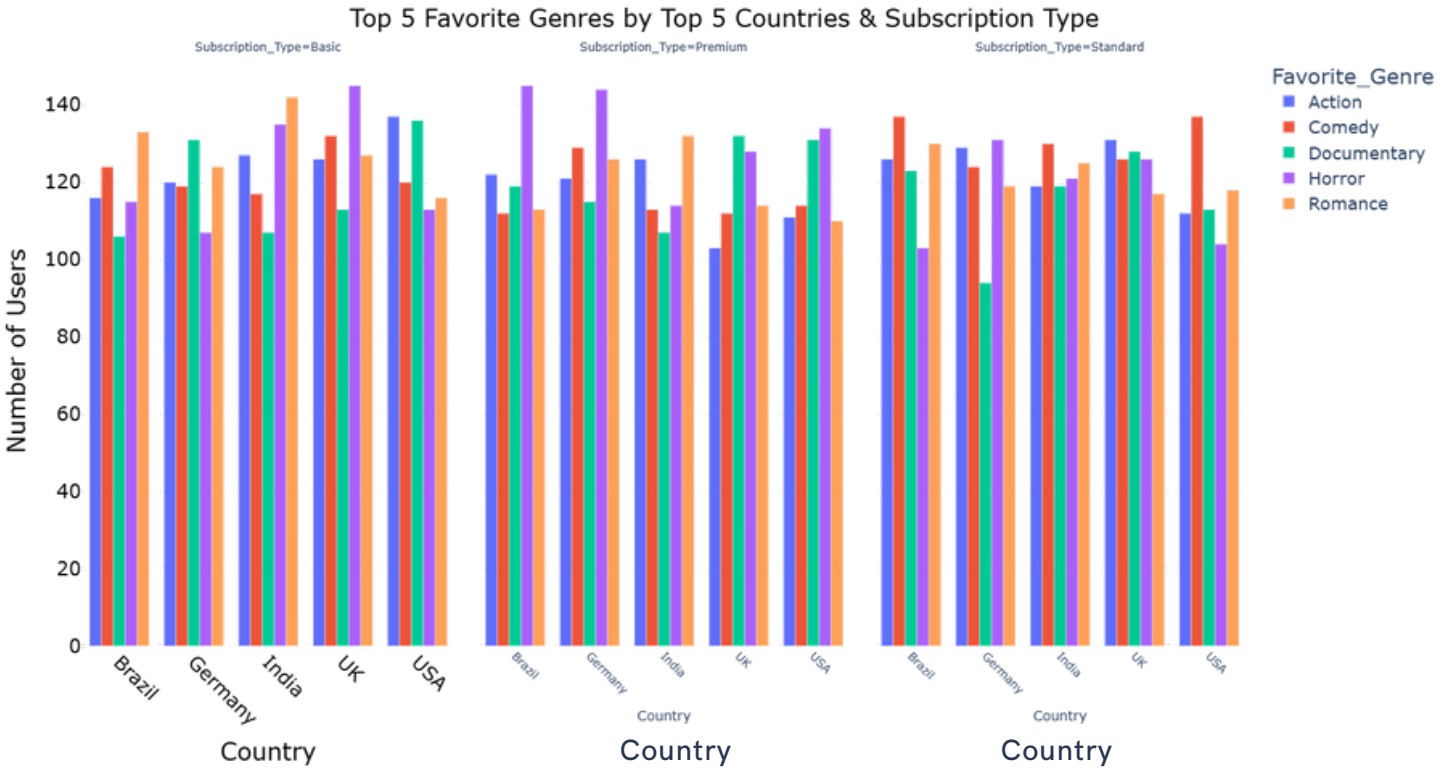
Key observations

- Users aged 18-45 on Premium and Standard plans watch significantly more hours than other age groups, showing that mid-age higher-tier subscribers are the most engaged.
- Basic subscribers have steadier but lower watch times across all ages, indicating underutilization of content despite access.

Business Insights

- Focus retention, personalized recommendations, and upsell campaigns on 18-45 Premium and Standard users to maximize engagement and subscription value.
- Encourage Basic users in this age range to upgrade by promoting exclusive content and features that boost watch time.

Horror Leads Among Premium Subscribers, Comedy Dominates Basic and Standard Across Countries



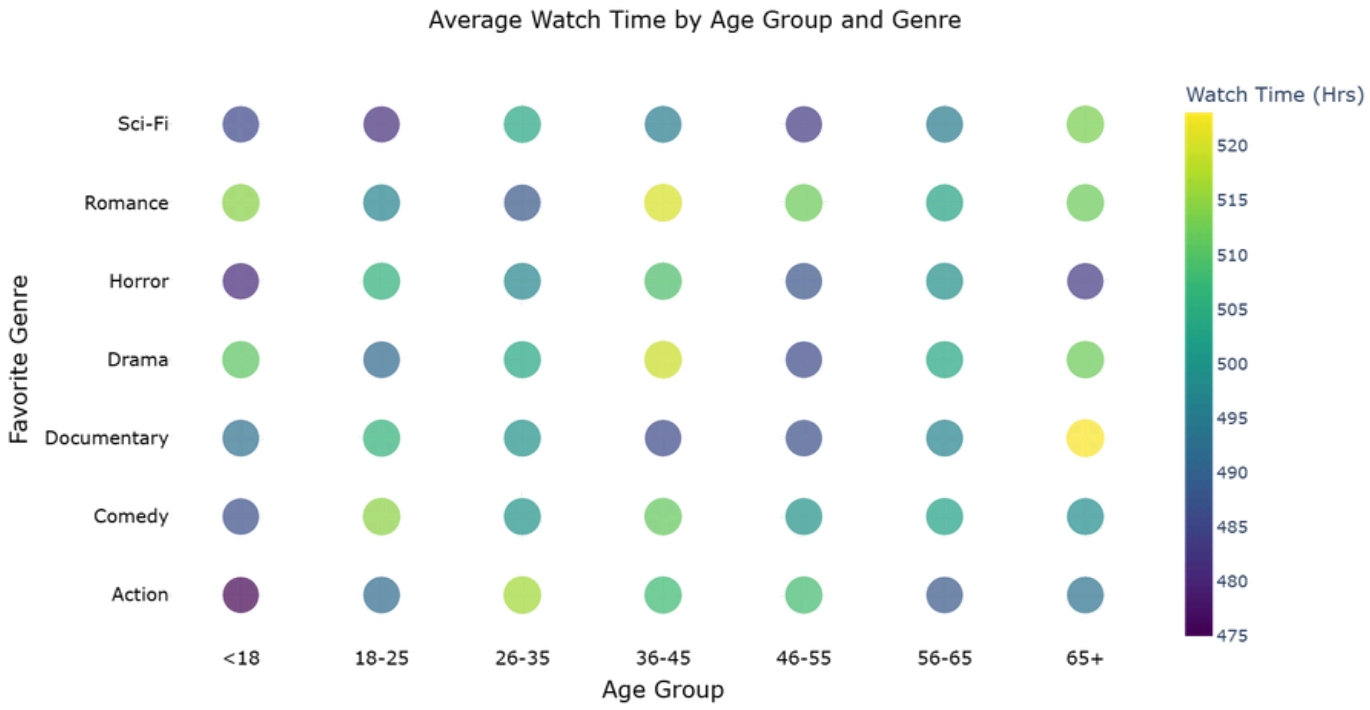
Key observations

- Horror is the top genre for Premium users across most countries, reflecting a strong engagement with high-intensity content among higher-paying subscribers.
- Comedy dominates preferences for Basic and Standard users, with genre popularity varying significantly by country, highlighting regional differences.

Business Insights

- Prioritize content acquisition and marketing campaigns toward Horror for Premium tiers to reinforce engagement and retention.
- Leverage regional and subscription-based genre trends to localize content strategies and improve relevance for diverse markets.

Romance and Drama Attract Highest Watch Time Among Middle-Aged and Senior Viewers



Key observations

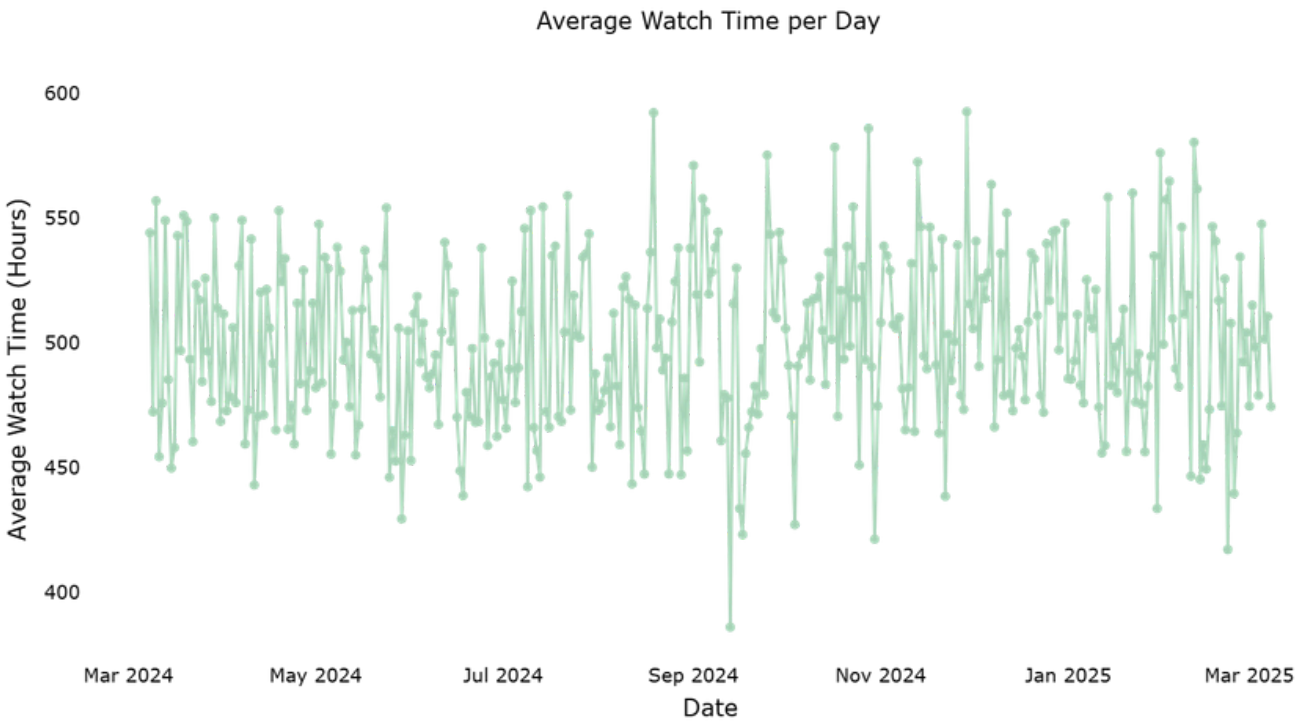
- Watch time peaks for Romance and Drama in ages 36-45 and 65+, showing these genres resonate strongly with middle-aged and senior audiences.
- Younger viewers prefer Comedy and Action, but their engagement in these genres is moderate compared to older groups in Romance and Drama.

Business Insights

- Invest in Romance and Drama content and promotions to engage middle-aged and senior subscribers, maximizing their platform usage.
- Use age-based genre preferences to personalize recommendations and marketing campaigns, boosting retention and satisfaction.

Temporal Trends & Activity Patterns

Daily watch time remains stable throughout the year, with short-term volatility but no sustained growth or decline trend



Key observations

- Daily average watch time fluctuates within a consistent range (~450–550 hours), showing no long-term upward or downward trend.
- Occasional sharp spikes and dips occur, but they are temporary and quickly return to baseline levels.

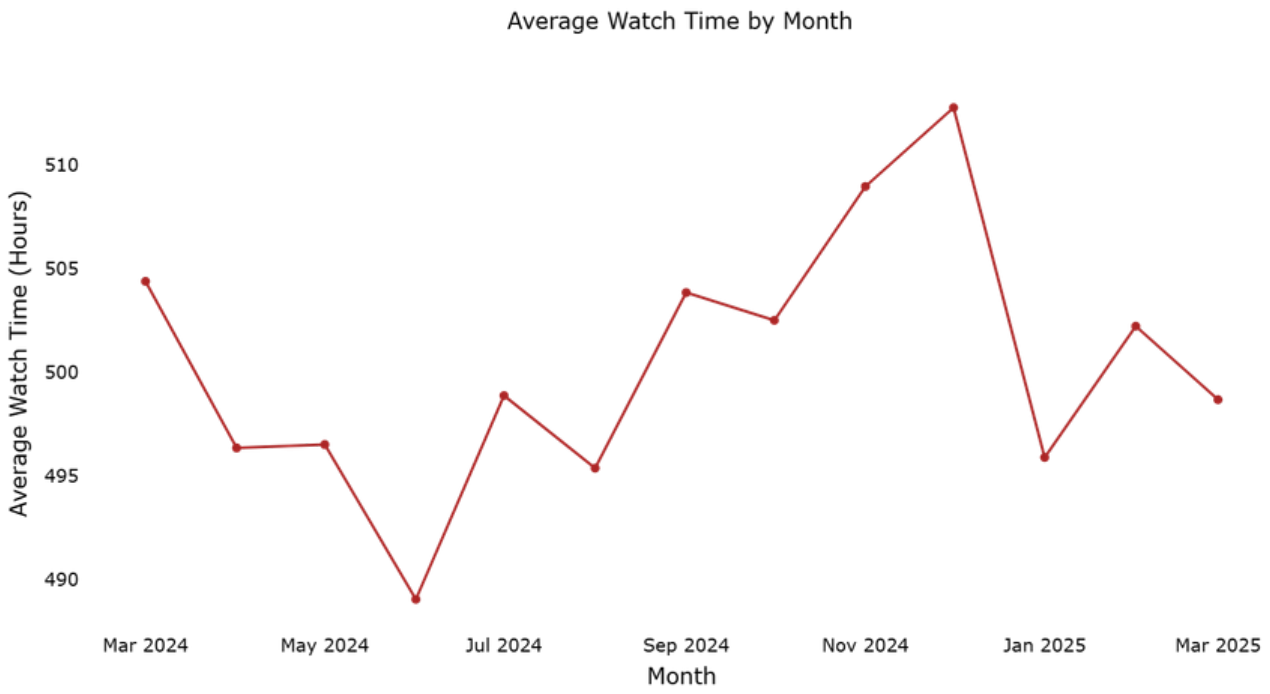
Business Insights

- Engagement is steady but plateaued, indicating strong retention but limited organic growth in viewing intensity.
- Identifying drivers behind spike days (e.g., content releases or campaigns) can help replicate high-engagement moments strategically.

User engagement peaks in Q4, highlighting strong seasonal uplift in watch time

$18.7M

Total revenue over selected period



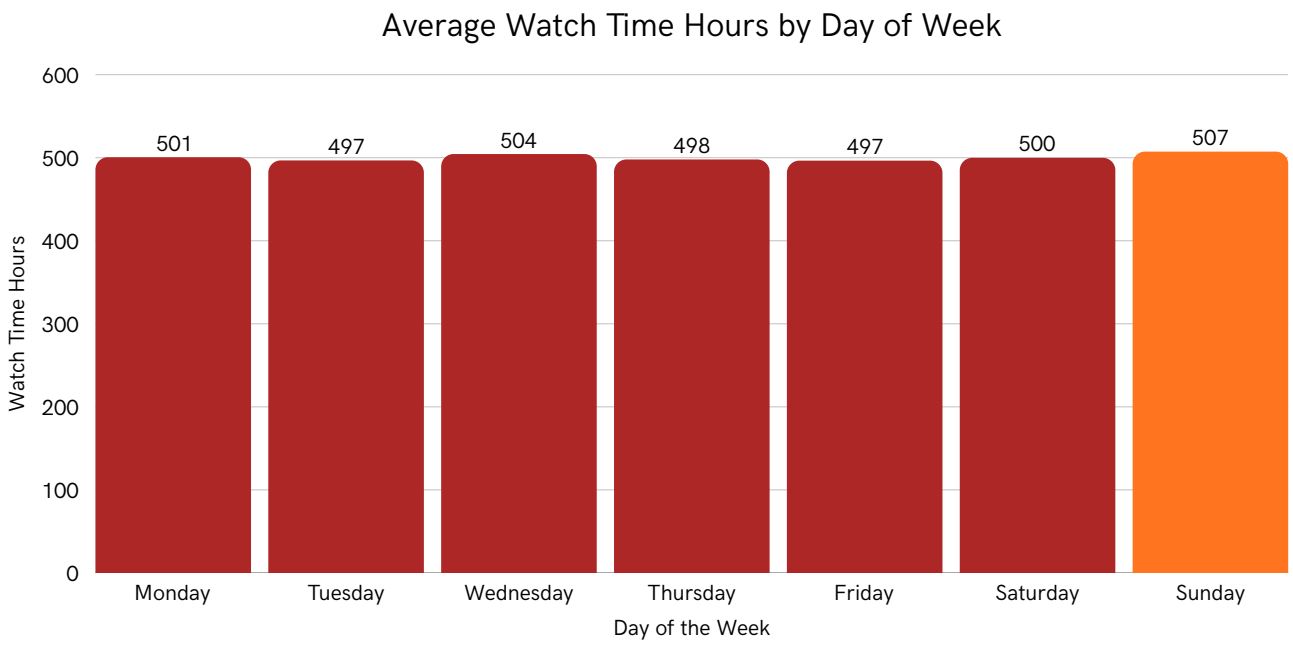
Key observations

- Watch time peaks in November–December, with December recording the highest monthly average.
- A noticeable dip occurs mid-year (around June) and again post-holiday in January.

Business Insights

- Q4 is a high-impact window for flagship releases, premium promotions, and acquisition campaigns.
- Mid-year and post-holiday dips may require targeted engagement strategies to prevent seasonal slowdowns or churn.

User engagement remains consistent throughout the week, with a slight weekend uplift



Key observations

- Watch time remains highly consistent across all weekdays, with minimal variance (~10-hour spread).
- Sunday records the highest engagement, while Tuesday/Friday are marginally lower.

Business Insights

- Netflix usage is habit-based rather than weekend-dependent, reflecting strong platform stickiness.
- Sunday presents a strategic opportunity for major content drops to capitalize on peak engagement.

BUSINESS / DEVELOPER TAKEAWAYS

- **Global User Scale Is Mature; Engagement Is the Core Growth Engine:** With a well-diversified user base across regions, Netflix growth will be driven by higher watch time and premium adoption rather than geographic expansion.
- **Balanced Subscription Mix Limits Structural Upside:** Even tier distribution reduces pricing dependency but shifts revenue focus toward upgrades and per-user engagement depth.
- **Working-Age Users Anchor Platform Value:** The 18-45 segment forms the backbone of engagement and monetization, while younger and older cohorts represent targeted growth opportunities.
- **Viewing Hours Are Highly Concentrated:** A small group of highly engaged users likely contributes a disproportionate share of total watch time, making their retention critical.
- **Platform Stickiness Outweighs Genre Dependence:** Consistent engagement across genres indicates that overall content quality and volume matter more than prioritizing individual categories.

BUSINESS / DEVELOPER TAKEAWAYS

- **Premium Value Is Feature-Led, Not Consumption-Led:** Higher tiers do not correlate with higher usage, signaling that upgrades are driven by experience features rather than content volume.

- **Genre Preferences Enable Precision Targeting**: Horror reinforces Premium engagement, while Romance and Drama sustain usage among older audiences, supporting age- and tier-based personalization.

- **Older Users Represent a High-Retention Segment:** Strong engagement among senior users presents opportunities for tailored content and marketing to maximize lifetime value.

- **Engagement Stability Signals Platform Maturity:** Plateaued watch time reflects strong retention but limited organic growth, requiring engineered engagement spikes through releases and campaigns.

- **Seasonality and Habitual Usage Create Predictable Leverage Points:** Q4 and Sundays offer the highest-impact windows for flagship launches, while mid-year and post-holiday periods need targeted re-engagement strategies.

CHALLENGES & OPPORTUNITIES

Limitations

- The dataset captures user behavior only within the Netflix platform and does not account for content discovery or influences outside the platform environment.
- User engagement metrics reflect activity within the observed timeframe and may underrepresent long-term viewing behavior or infrequent users.
- The Last_Login timestamp provides a snapshot of recent activity, limiting the ability to infer sustained behavioral changes or long-term engagement trends beyond the available data window.

Future Work / Opportunities:

- Incorporate user ratings, reviews, or feedback signals (if available) to better understand content satisfaction and its relationship with sustained engagement.
- Extend time-based analysis using longer observation windows to study how viewing intensity, subscription behavior, and genre preferences evolve over time.
- Perform deeper segmentation between light, mid-tier, and heavy viewers to identify behaviors associated with higher long-term engagement and subscription value.
- Enrich the dataset with content discovery and recommendation exposure metrics to assess their impact on watch time, genre exploration, and subscription upgrades for Netflix.

CONCLUSION

This analysis of the Netflix user dataset provides insights into user engagement patterns, subscription behavior, and content preferences across age groups, geographies, and genres. The findings highlight how engagement depth, subscription mix, and content diversity shape overall platform performance, while future work could extend this understanding through long term time-based analysis and deeper segmentation of viewing intensity and subscription upgrades for Netflix.

This analysis was conducted using Python, Pandas, Matplotlib, Seaborn and Plotly.

Author:

Tejas Jadhav

GitHub: **@tejas-jadhav**

LinkedIn: **Tejas Jadhav**