stc TV Data Analysis

By: Nuha Alanezi

April 2025

User Behavior Analysis

Top 10 Watched Programs by Total Watch Time (Mar 14, 2017 – Apr 30, 2018)

Legend:
- The Boss Baby
- The Amazing pider-Man
- The Expendables
- Moana
- Trolls
- Bean
- The murfs
- Hotel Transylvania
- Cloudy With a Chance of Meatballs
- The Man With The Iron Fists

Pie chart values: 18.7%, 12.4%, 12.4%, 10.8%, 10.1%, 8.97%, 8.46%, 6.91%, 5.98%, 5.42%

Tooltip: program_name=The Boss Baby / Total watch time in houres=2,961.350833 / program_class=MOVIE

1. The Boss Baby leads with 18.7% of the total viewing time among the top 10, making it the most popular program.

2. The Amazing Spider-Man and The Expendables follow closely with 12.4% each. Moana, Trolls, and The Smurfs also appear prominently, highlighting a strong preference for Animated content.

3. The chart shows a dominance of Movies (especially animated and action genres), suggesting viewers are highly engaged with visually rich, family-friendly content.

User Behavior Analysis

Viewing Behavior by Program Type (Movies vs Series/Episodes)

Top Chart: Total Duration Spent by Program Class71.1% of the total watch time was spent on Series/Episodes.28.9% of the total watch time was spent on Movies.

Users tend to spend more time on Series, possibly due to binge-watching behavior and the longer total duration of episodes.

Bottom Chart: Total Users Watching by Program Class74.4% of users watched Movies.25.6% of users watched Series/Episodes.
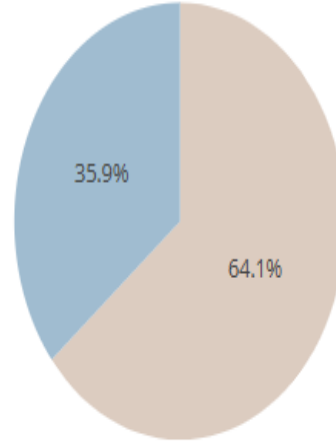
While more users watch Movies, the average time spent per user is higher for Series.

This suggests that: Movies attract more casual viewers, while Series engage users for longer periods.

User Behavior Analysis

Movie Viewers by Quality Preference (HD vs SD)

No of Users who Watched MOVIE Vs Program Quality

- HD
- SD

35.9%

64.1%

- **64.1% of users watched movies in HD quality.**

- **35.9% of users watched movies in SD quality.**

User Behavior Analysis

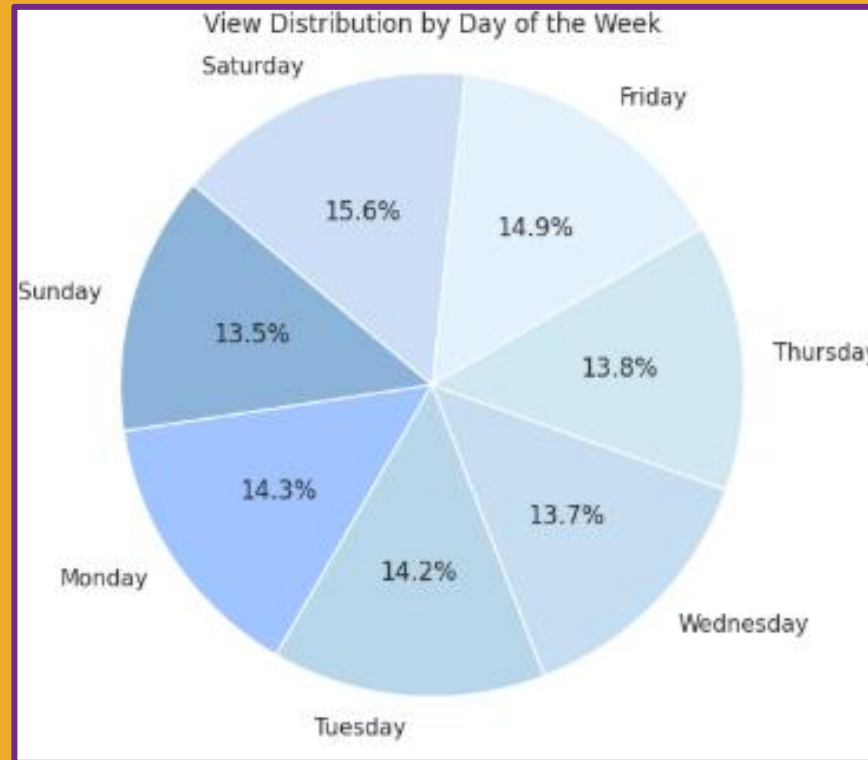Series/Episodes Viewers by Quality Preference (HD vs SD)

No of Users who Watched SERIES/EPISODES vs Program Quality

SD
HD

44.4%   55.6%

- **55.6% of users watched in SD quality.**

- **44.4% of users watched in HD quality.**
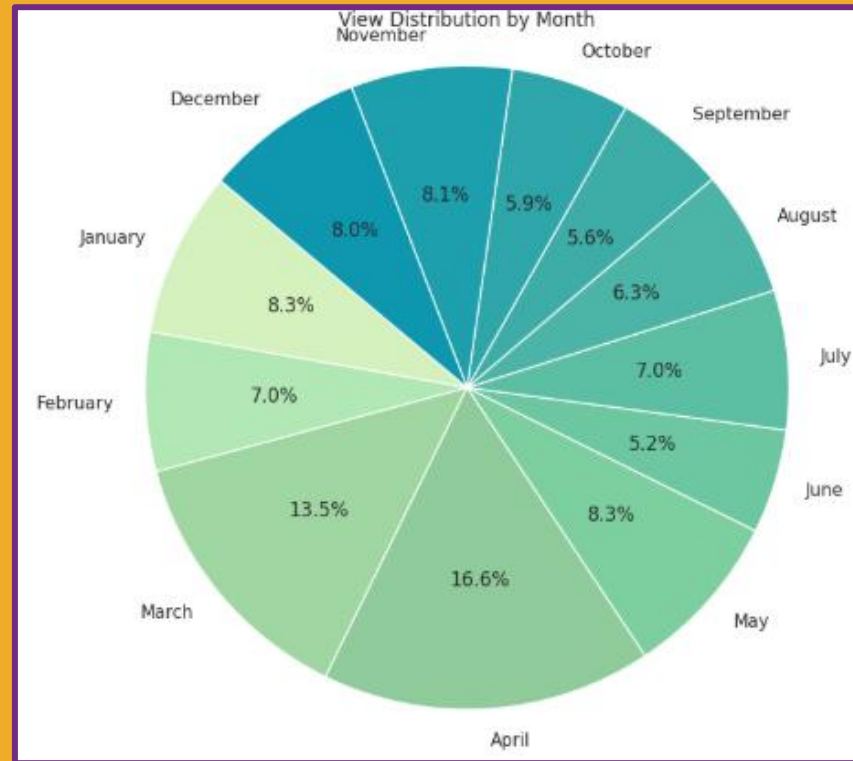
User Behavior Analysis

User Viewing Patterns by Weekday

- **Saturday (15.6%) and Friday (14.9%) have the highest viewership, indicating peak engagement during weekends.**

- **Viewership is relatively consistent on weekdays, ranging from 13.5% to 14.3%.Sunday has the lowest share (13.5%).**
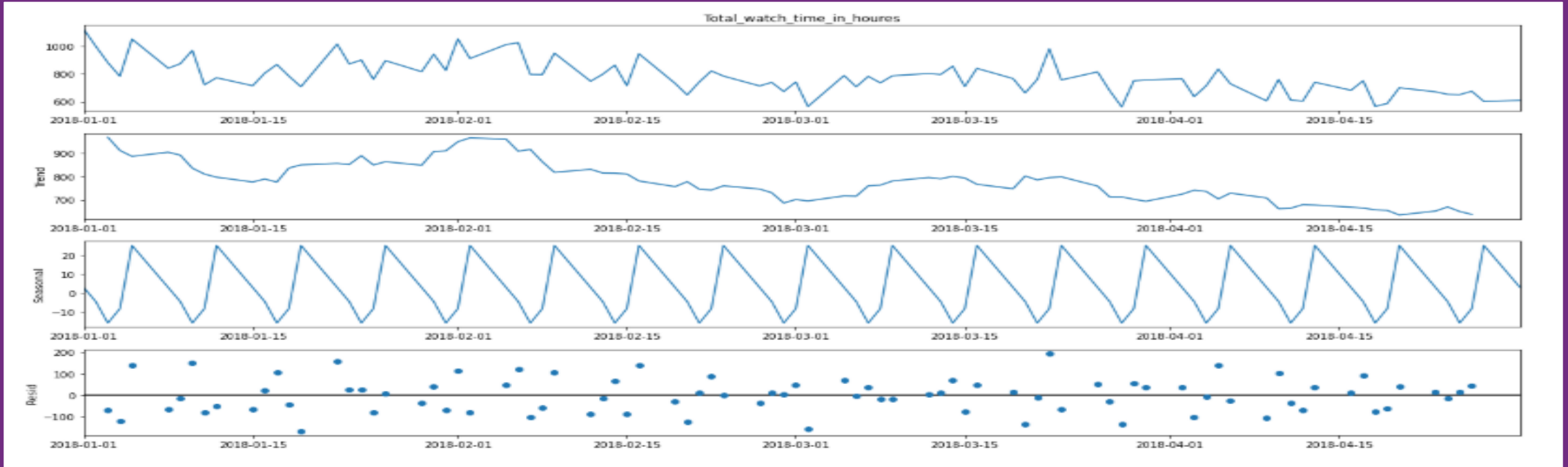
User Behavior Analysis

Monthly Viewership Trends
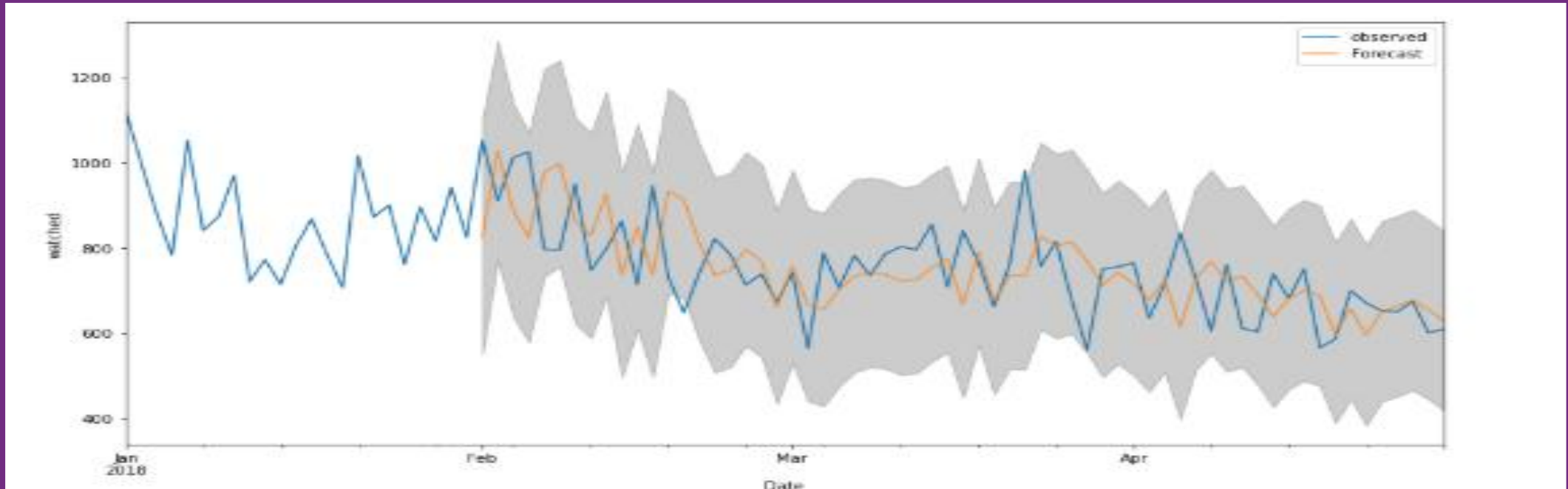

View Distribution by Month

- **April (16.6%) and March (13.5%) have the highest viewership, indicating a significant spike in spring.**

- **May and January are tied with 8.3%, showing strong start-of-year and post-spring engagement.**

- **June (5.2%), October (5.9%), and September (5.6%) have the lowest activity.**

Watch Time Forecasting – Time Series Decomposition

- The trend component shows a gradual decline in overall watch time during the later months.

- A clear weekly seasonality pattern is visible (peaks and dips every 7 days), possibly reflecting weekend vs.weekday behavior.

- Residuals are minor, indicating that most of the data is well-explained by the trend and seasonality.
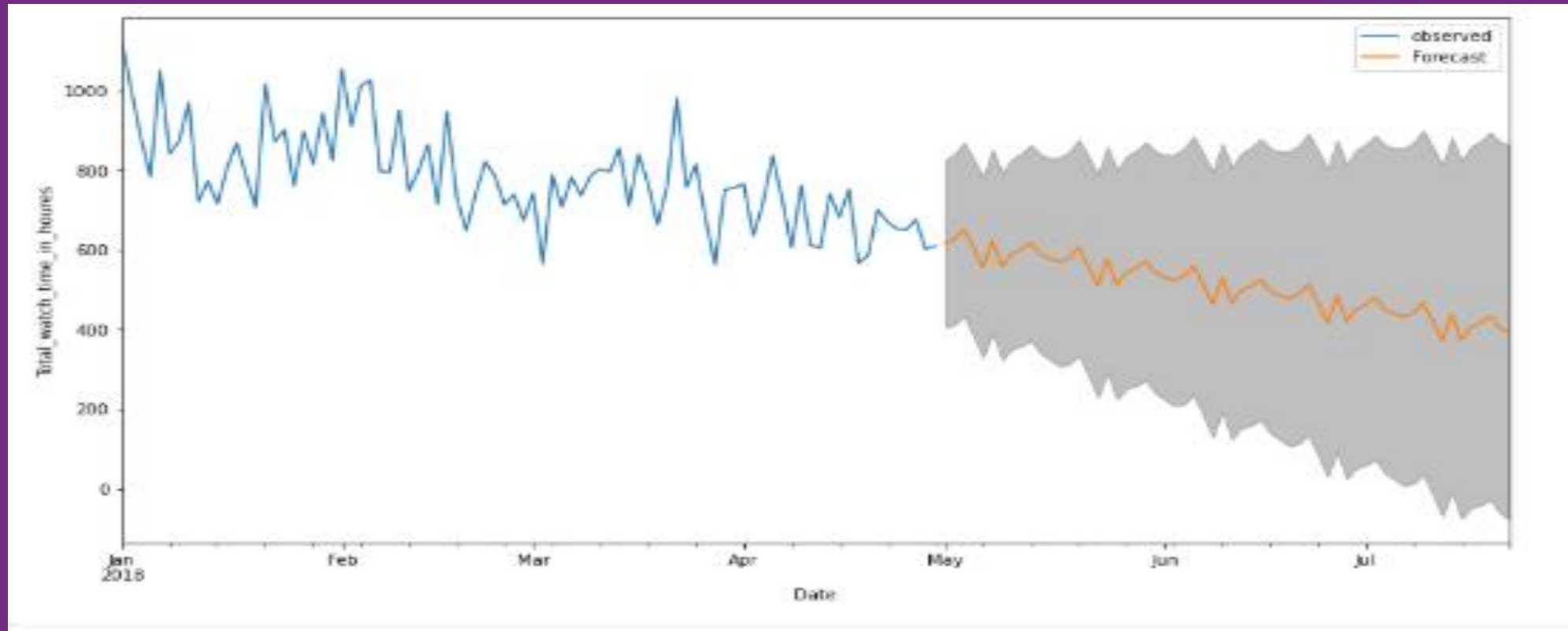
Watch Time Forecast vs Actual (SARIMAX Model)

- The forecast closely follows the actual trend in the short term (early February), but the uncertainty grows in later months.

- There is a visible drop in both actual and predicted values across the months, suggesting declining user engagement.

- The model is stable and captures trend + seasonality, but precision decreases as we forecast further into the future.

Watch Time Forecast vs Actual (SARIMAX Model)

Watch Time Forecast – Next 2 Months (May to July 2018)

- **A noticeable downward trend is expected in total watch time across the coming two months.**

- **The model predicts a steady decline from ~600 hours/day down to nearly 400 hours/day, with possible dips even lower.**

- **The confidence interval widens into June–July, suggesting higher uncertainty the further out we forecast.**

Top 5 Recommendations Using Machine Learning

- **Recommendation Engine – Nearest Neighbors (Cosine Similarity)**

- **Objective: Suggest similar content for users who watched Moana.**

- **Model Details:**

Model used: NearestNeighbors from sklearn.

Similarity metric: Cosine Distance (measures how close user preferences are).

Dataset: Sparse matrix of users vs. watched titles.

- **Top 5 Recommendations for "Moana":**

	Title	Distance
0	Trolls	0.427642
1	Surf's Up : WaveMania	0.470576
2	The Mermaid Princess	0.506638
3	The Boss Baby	0.551443
4	The Jetsons & WWE: Robo-WrestleMania!	0.561058

Lower distance = more similar
All recommendations are animated or family-friendly, perfectly aligned with *Moana's* genre and tone.

Top 5 Recommendations Using Machine Learning

	Title	Distance
0	Trolls	0.427642
1	Surf's Up : WaveMania	0.470576
2	The Mermaid Princess	0.506638
3	The Boss Baby	0.551443
4	The Jetsons & WWE: Robo-WrestleMania!	0.561058

Lower distance = more similar
All recommendations are animated or family-friendly, perfectly aligned with *Moana's* genre and tone.

The system effectively captures user behavior and genre-based similarities.

It can be personalized per user to enhance retention and time spent on the platform.

"The Boss Baby" appears again, reinforcing its popularity and relevance.

Top 5 Recommendations Using Machine Learning

Top 5 Highest Rated Genres:

1. Romance – 2.9355
2. Animation – 2.6132
3. Drama – 2.6433
4. Comedy – 2.4190
5. Sci-Fi – 2.8879

```
program_genre
Action                                  2.358474
Adventure                               2.125523
Animation                               2.613295
Biography                               1.821320
Comedy                                  2.418921
Crime                                   2.284653
Documentary                             2.011512
Drama                                   2.643298
Family                                  2.677141
Horror                                  2.340451
NOT_DEFINED_IN_UMS                      1.886439
Romance                                 2.935547
SERIES_NOT_ADDED_UNDER_ANY_GENRE        2.481481
Sci-Fi                                  2.885779
Thriller                                2.316255
Wrestling                               1.125000
Name: rating, dtype: float64
```

Romance receives the highest average rating, suggesting strong viewer satisfaction.

Animation and Drama follow closely, which aligns with the top watched content trends.

Wrestling has the lowest rating (1.15), indicating either niche appeal or viewer dissatisfaction.

Future Work: Recommendations to Improve the Service

- EXPAND AND ENHANCE THE RECOMMENDATION SYSTEM USING HYBRID MODELS (USER-BASED + CONTENT-BASED).

- INTEGRATE MORE PERSONALIZED ANALYTICS BASED ON USER DEMOGRAPHICS AND PREFERENCES.

- OFFER MORE FLEXIBLE VIDEO QUALITY OPTIONS: HD, FHD, AND 4K TO SATISFY QUALITY-SENSITIVE USERS.

- INTRODUCE SMART NOTIFICATIONS BASED ON VIEWING PATTERNS (E.G., WEEKEND BINGE ALERTS).

- OPTIMIZE CONTENT RELEASE SCHEDULING USING FORECASTED HIGH-TRAFFIC PERIODS.

- CONTINUOUSLY TEST AND ANALYZE UNDERPERFORMING GENRES (LIKE WRESTLING) FOR IMPROVEMENT OR REMOVAL.

- IMPLEMENT AI-DRIVEN CONTENT TAGGING TO IMPROVE SEARCH, FILTERING, AND DISCOVERY EXPERIENCES.

Contact us

Thank you for your time!

For questions, collaboration, or feedback, feel free to reach out:

LinkedIn: https://www.linkedin.com/in/nuha-alanezi-805a77212/

GitHub: https://github.com/naz50/Data-analysis-at-STC

Special Thanks To: