**Module 6: Critical Thinking Assignment**

Sophia Balentine

Colorado State University Global

MIS581 - Capstone - Business Intelligence and Data Analytics

Dr. Osama Morad

July 20, 2025

**Introduction**

In the highly competitive world of digital streaming services, companies constantly seek strategies to increase viewership and profitability while retaining subscribers. The massive growth in streaming consumption over the past decade has intensified the need for platforms to understand what drives audience engagement and revenue. This Capstone Project uses quantitative research to investigate key aspects like the impact of content library volume, contract term length, and the influence of viewer ratings and online popularity on engagement and profit. Using predictive analytics, this project will help provide data driven recommendations for decision makers within the streaming industry.

The results of this Capstone Project aim to provide actionable insights for streaming platforms that face increasing competition and evolving audience expectations. Understanding which types of content, contracts, and viewer feedback mechanisms drive profitability and sustained engagement will help platforms optimize their content strategies and improve their competitive positioning in the market. The platforms need a boost, as Alex Weprin, a journalist with the Hollywood Reporter, noted in his 2023 article:

“Netflix now finds itself atop the subscription streaming heap, ending 2022 with 231 million paid subscribers and $5.6 billion in profits. […] The rest of the industry, meanwhile, is looking at 2024. That’s when Disney, NBCU, Paramount and WBD all say they expect — or at least hope — to swing to a profit in their streaming businesses.” (Weprin, 2023).

While Netflix is very profitable, almost all other streaming platforms have not even gotten out of the red. For them to continue to produce and hopefully one day see success, it is important to determine the factors that have helped profitable companies like Netflix get there.

I will gather references and conduct analysis to determine which hypotheses to support/reject and answers to my research questions. I have found multiple sources that will help me find the answers I am looking for. I have gathered articles that report streaming platform profits, statistics that detail viewer and critic ratings, and more to help with this process.

**Objectives and Organizational Benefits**

The following are my objectives for the Capstone Project that I would like to achieve by completing this analysis:

1. Analyze data that includes observations from at least three major streaming platforms to identify the top content traits linked to higher user engagement and profits.
2. Compare different approaches to content library structures to determine the most cost effective and high profit strategies.
3. Analyze the relationship between viewer ratings and platform subscribership using data from major streaming services, if that relationship exists.
4. Explore if the length of term for subscription contracts has an effect on the amount of customers who churn (join and then quickly leave a subscription contract).

The following are potential organizational benefits from the completion of my analysis and this project:

1. Improved content investment decisions: By analyzing what traits are linked to higher or lower profitability and viewership, the business can decide to devote more or less funds to certain projects or licensing fees.
2. Enhance the user experience: By analyzing and better understanding the influence that user ratings/popularity metrics have on audiences, the business can make themselves more competitive and retain subscribers.
3. Predictive modeling: The business can utilize predictive models built based on my analysis that will forecast things like viewer behavior or content success. These forecasts can give the company a competitive advantage and increased profits.
4. Potential cost saving: By analyzing the popularity and engagement of different features, the company can potentially cut funding to less successful projects and save money.
5. Competitive advantage: The market is flooded with choices for streaming platform subscriptions. The entertainment industry has moved farther away from cable television and the focus has been placed on on-demand, at your fingertips media. With so many choices for consumers, it is important that platforms strive for an advantage over their counterparts. Multiple parts of my analysis aim to provide that advantage whether through better content curation based on engagement rates, more effective subscription modeling, or whatever else.

**Overview of Study**

My study seeks to understand the key drivers behind profitability and sustained engagement across digital streaming platforms. As competition increases and traditional broadcasting fades, platforms must identify what content strategies lead to subscriber retention, increased engagement, and ultimately higher profit margins.

This research focuses on three core areas: the amount of content a platform offers, the effectiveness of different subscription models, and the impact of user feedback and critical ratings on platform success.

Through quantitative analysis of multiple datasets, including content libraries, viewer ratings, and subscription statistics this study aims to find what makes a streaming strategy successful. The study also leverages real world trends and examples. These include Netflix's market dominance and the surprising success of platforms like Amazon Prime Video, whose main business is not focused on streaming. This will help me to reject or accept hypotheses and contextualize my results.

Ultimately, this study will offer strategic recommendations for decision makers in the industry who are seeking evidence-based practices to optimize their content portfolio, attract and retain subscribers, and increase overall profitability.

**Research Questions and Hypotheses**

To guide this investigation, I developed the following research questions:

1. Do platforms that offer more content choices pull in more subscribers?
2. Do viewer ratings or online popularity metrics on originally produced content significantly influence number of subscribers?
3. Does the type of subscription contract a consumer signs influence how long they stay a subscriber?

These questions aim to address both the operational and strategic challenges faced by streaming services that are investing heavily in content production and licensing. Understanding these factors can help platforms refine their content acquisition strategies, content library decisions, and utilize all tools at their disposal.

Based on the research questions, I have formulated the following hypotheses:

1. Hypothesis 1
   1. H₀: There is no significant relationship between the amount of content a platform offers in their library and sustained subscriber retention.
   2. Hₐ: There is a significant relationship between the amount of content a platform offers in their library and sustained subscriber retention.
   3. There are differing opinions across the entertainment industry on what strategy is best as Marah Eakin suggested in a Wired article: “The act of discovering content on a platform, is almost impossible.” (Eakin, 2024)

This ongoing debate within the industry highlights the importance of understanding library organization strategies. Are platforms risking ‘lost’ content for the sake of having more options? By examining historical data from multiple platforms, this hypothesis will test whether bigger is actually better.

1. Hypothesis 2
   1. H₀: Viewer and critic ratings (such as Rotten Tomato scores or Letterboxd reviews) are not significantly correlated with high subscriber numbers.
   2. Hₐ: Viewer and critic ratings (such as Rotten Tomato scores or Letterboxd reviews) are significantly correlated with high subscriber numbers.
   3. Viewer ratings and critical acclaim are often leveraged by platforms as promotional tools. Business Insider, the wide known publication, publishes lists of the most popular weekly streaming television shows in the United States. In one of these posts from April of 2021, they analyzed the demand for the TV show and compared it to the Rotten Tomatoes critic rating. It was very interesting to see that one of the most in demand shows, The Handmaid’s Tale on Hulu, at 26.1 times more in demand than the average show, only had a 56% score on Rotten Tomatoes (The Business Insider, 2021). This would lead me to believe that ratings have less bearing on popularity than we would think. This project will assess whether these popularity indicators significantly drive actual engagement and revenue, or if they merely serve as superficial marketing tools/numbers.
2. Hypothesis 3
   1. H₀: Shorter contract lengths are not significantly correlated with customer churn (i.e. customers leaving a streaming service quickly after signing up).
   2. Hₐ: Shorter contract lengths are significantly correlated with customer churn.
   3. Contracts come in all different kinds of shapes and sizes. They can range from month to month all the way to multiple year lengths. By analyzing the dataset I have found on customer churn from a simulated telecommunications company, I will be able to discern whether the contract that customers choose has a bearing on how long they stay. Answering this question will provide the business’ decision makers with strategies that would increase profits moving forward. For example, based on an article from World Screen, a publication that covers the international entertainment industry:

“January [2024] garnered nine of the ten highest daily streaming levels ever, with Saturday, January 13 marking the most-streamed day in history, totaling 40.8 billion minutes streamed–driven by Peacock’s coverage of the first exclusively streamed NFL playoff game.” (Brzoznowski, 2024).

This suggests that offering short term contracts for large events may be extra profitable for streaming platforms, especially if they increase the contract prices for these special showings. Overall, usually the longer a subscriber stays with the company, the more revenue they are producing.

**Literature Review**

My process for writing the literature review has focused on gathering a wide range of credible sources to support my topic and problem statements. I have primarily been using the CSU Global Library to search for scholarly articles and industry reports. I use keywords to search like “Streaming Services,” “Customer Churn,” “Streaming Profits,” or “Streaming Services Original Content.” One strategy that has been especially helpful is looking through the bibliographies of the articles/journals that I find. This often leads me to additional relevant sources I might not have come across through keyword searches alone, and it helps me understand how researchers have built on each other’s work overtime.

One challenge I’ve faced is that some articles focus heavily on streaming platforms that are not used in the United States or rely on outdated data, which does not always directly address the factors leading to higher profits today. Because of this, I have had to piece together information from different perspectives and sources to build a clear connection to my topic and make sure my literature review stays relevant. To narrow my focus, I also sometimes include specific streaming platforms in my searches, like Netflix, Hulu, or Disney+, to find articles that discuss their business’ trends, content strategies, and subscriber behaviors in more detail. Overall, staying organized, taking careful notes, and using the reference lists of strong sources have really helped strengthen my literature review and ensure that I am covering my topic from multiple angles.

**Research Design - Methodology**

The multiple datasets I selected provide a comprehensive foundation for analyzing the key factors that influence streaming service viewership, customer churn, content performance, and overall profitability.

The first dataset, focused on User Churn, allows me to examine the demographics, behaviors, and contract details that contribute to customers leaving. This is critical information for developing retention strategies.

Datasets like Rotten Tomato Scores and the Amazon Prime Video TV Shows with Ratings are most helpful when trying to discern if ratings have any bearing on platform success. By looking at the average scores across different platforms, we can see who is implementing the most profitable strategies.

The content lists from different platforms like Disney+ and Hulu give insight to just how much these companies are offering to their viewers. If a certain streaming service is more successful and has only a quarter of the content available as their less profitable competitor, then we can try to analyze the correlation between their success and content choices. This gives insight into how reception to content impacts subscriber satisfaction and platform loyalty.

Additionally, the Streaming Services dataset offers an overview of competing platforms’ reach and subscriber counts, which is essential for benchmarking market position and identifying competitive advantages. One area of data I would appreciate is internal data. Data from inside some of these companies about their content suggestion algorithm’s success, their profits, recent data etc. But, much of this data is proprietary and not open to the public, so I have been making do with what is available.

By combining customer data, content ratings, and platform comparisons, I can generate a multidimensional analysis that highlights the drivers behind viewership trends and increased subscribership. Together, these datasets support my goal of developing actionable recommendations that streaming platforms can use to retain customers, produce desirable content, and strengthen their market presence.

**Research Design – Method**

To test these hypotheses, a quantitative research approach using predictive analytics will be employed. The methodology will include descriptive statistics, correlation analyses, regression models, and hypothesis testing through t-tests or ANOVA where applicable.

The research will utilize secondary data gathered from sources like Kaggle such as viewership trends, content library lists, and ratings from established platforms like Rotten Tomatoes. I will also utilize simulated data that observes a fictitious telecom company’s contracts and user pool.

I will use the following tools and models:

* SAS Studio: SAS Studio will serve as the primary tool for data cleaning, transformation, and conducting advanced statistical tests, such as multiple linear regression and hypothesis testing.
* Microsoft Excel: Excel will be used for preliminary data organization, summary statistics, and simple visualizations that assist in the initial exploration of trends and anomalies.
* Tableau: Tableau will help create interactive dashboards and visualizations to effectively communicate findings to stakeholders, allowing for clear presentation of trends, correlations, and model outputs.

For example, the use of a tool like correlation coefficients will determine the strength and direction of relationships between viewer ratings and actual viewership figures, while using regression models will help test whether variations in profit margins can be explained by content type or release strategy.

**Research Design – Limitations**

While this research aims to offer comprehensive insights, several limitations should be acknowledged:

1. Data Availability and Scope: Much of the analysis relies on secondary datasets. These may not reflect the most recent platform strategies or include proprietary data only available internally to streaming companies.
2. Sampling Bias: Some datasets may overrepresent U.S. based platforms or demographics, limiting the applicability of findings to international markets.
3. Correlation vs. Causation: While the research uses statistical tools like regression to infer relationships, the results may show correlation rather than causation. Viewer engagement may be influenced by factors not captured in the available data.
4. Platform Variability: Each streaming service operates under different business models (some with ads, some subscriber only, etc.), which could affect profitability in ways that are difficult to normalize across datasets.
5. Timeframe Gaps: Some datasets reflect a particular time window (like COVID-19 lockdown periods for example) and may skew behavioral trends not representative of normal user behavior.

These limitations will be carefully considered during the analysis, and findings will be interpreted with caution and clearly contextualized.

**Research Design – Ethical Considerations**

Given the sensitive nature of data in the entertainment industry, strict ethical standards will be followed throughout this project. These include:

* Privacy Protection: This project will not involve the collection or use of any personal or confidential user data. All data will be aggregated from publicly available sources.
* Data Source Credibility: Only reputable and verifiable sources will be used to ensure the integrity and accuracy of the analysis.
* Transparency: Research methods, statistical models, and data sources will be fully documented and disclosed to allow for reproducibility and verification.
* Bias Mitigation: Special attention will be given to identify and address any potential biases in the data. For instance, data that overrepresents specific regions or demographics could skew findings. This will be controlled for during analysis when possible.

By adhering to these ethical considerations, the research will maintain high standards of integrity and respect for privacy, aligning with the broader ethical principles that guide data driven decision making.

**Findings**

*Hypothesis 1: Content Library Volume versus Subscriber Numbers*

To test whether the amount of content in a streaming platform’s library significantly correlates with sustained subscriber retention, I used datasets that measured both total content offerings and subscriber metrics across the streaming platforms Amazon Prime, Netflix, Hulu, Apple TV+, HBO Max, and Disney+. The first thing I did was find the total number of subscribers that each of these six platforms had. The following are my findings:

A screen shot of a graph

AI-generated content may be incorrect.

Next, I found the total amount of content each of the six platforms offers in their content libraries. The following are my findings:

A screen shot of a graph

AI-generated content may be incorrect.

After finding these important numbers, I performed a correlation analysis between the number of titles offered to viewers and the total subscriber count. The following is the correlation analysis table:

A screenshot of a computer screen

AI-generated content may be incorrect.

Netflix, which had the most extensive content library, also had the highest subscriber count of 231 million. Apple TV+ on the other hand, has the least expansive library and the lowest amount of subscribers with only 20 million, a tenth of what Netflix supports. These two alone would support my hypothesis that platforms that offer more content are more popular. Hulu is something of an outlier though, and disproves this theory. Hulu has the second highest amount of content offered to viewers with 3,090 pieces of content to choose from but, sports the second lowest number of subscribers with only 45.3 million.

The correlation analysis further supports what the data initially shows. When I tested the correlation between the number of subscribers and the volume of a platform’s library, I found the Pearson correlation coefficient was 0.62407 and the P-Value was 0.1854. This shows that there is a strong positive correlation between subscribers and content but, that correlation is not statistically significant. This is most likely due to the low number of observations I had to work with. This is a good example of correlation does not equal causation. Just because the two are correlated it does not mean that the amount of content offered is causing high or low subscribers. With this, the null hypothesis is accepted. There is no significant relationship between the amount of content a platform offers in their library and sustained subscriber retention.

*Hypothesis 2: Viewer Ratings Effects on Subscribership*

The second hypothesis examined whether audience scores significantly influence subscriber numbers. I used Rotten Tomatoes scores for original content across the six platforms used in the first analysis which includes Amazon Prime, Netflix, Hulu, Apple TV+, HBO Max, and Disney+. To prepare the dataset with the viewer scores, the first thing I had to do was make the data usable. I started by converting the two columns that hold the audience and critic scores from text strings (percentages) to a decimal. This was the output, a dataset that I could use for analysis:

A screenshot of a data

AI-generated content may be incorrect.

Next, I generated histograms to visualize the distribution of viewer ratings across each of the platforms side-by-side:

A screenshot of a graph

AI-generated content may be incorrect.

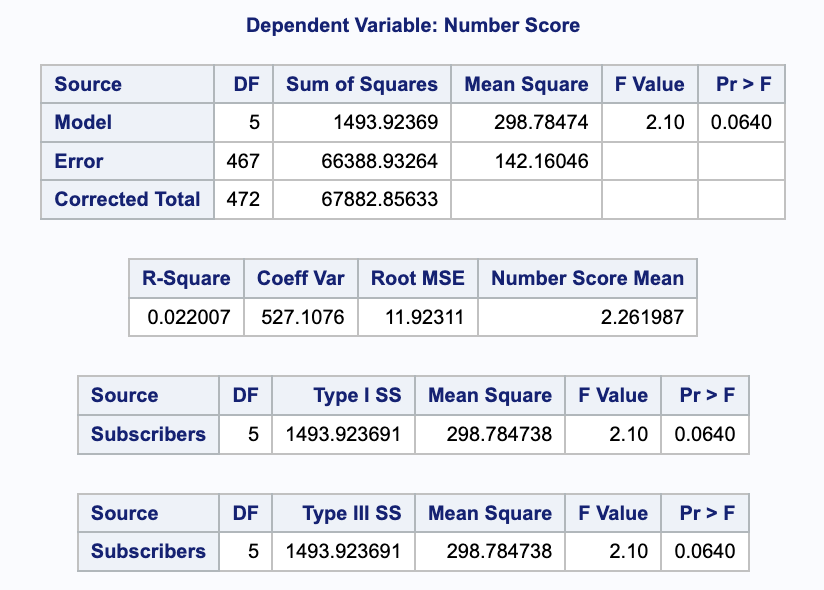
I also generated a line-graph of the findings from the histogram to better understand the layout of the data:

A graph of different colored lines

AI-generated content may be incorrect.

As you can see from the charts, Amazon Prime Video has the largest amount of content rated highly, over 80%. All of the streaming services are skewed to the left which tells us that viewers tend to rate higher rather than lower.

Next, using the same subscriber numbers for each platform that were found in the previous analysis, I performed a one-way ANOVA analysis between average platform ratings and subscriber numbers:



This is a low F-Value, at 2.10, indicating that the differences in average ratings between amounts of subscribers are only slightly greater than the variation within the groups. It's not a strong signal. The P-Value on the other hand, 0.0640, is just above the threshold of 0.05 for significance. These are the results from the Levene and Welch’s method ANOVA tests:

A screenshot of a test results

AI-generated content may be incorrect.

As you can see, they have P-Values of 0.0478 and 0.03 respectively. They also have the same low F-Value. This could tell us that though the correlation is very low, it may not be due to random chance. That the relationship between number of subscribers and ratings, though small, is statistically significant.

With the combination of all of this data I think it is safe to tentatively accept the null hypothesis. Viewer and critic ratings (such as Rotten Tomato scores or Letterboxd reviews) are not significantly correlated with high subscriber numbers. This lines up with what we saw in the article above from entertainment sources, that some high-performing shows (like The Handmaid’s Tale on Hulu) had poor critic ratings but exceptional viewership, suggesting that critical acclaim and popularity do not always align. It may be more accurate to guess that a few highly engaging shows drive platform growth, and other factors like cultural trends or social media hype play bigger roles.

*Hypothesis 3: Contract Length and Subscriber Churn*

To explore whether contract length impacts user churn, I used a simulated dataset from a telecom service model. The first thing I did was filter the data to include only customers who subscribed to streaming services and later churned (i.e., canceled their subscription quickly after beginning it). After I had cleaned the data, I created a bar chart of the frequency of types of churn users’ contract terms; these included month-to-month, one year, or two years. The following are those results:

A graph of a customer

AI-generated content may be incorrect.

As you can see above, the vast majority of churn occurred among the month-to-month users. One-year and two-year contracts had significantly lower churn rates. To confirm the statistical significance of length of contract, I did a correlation analysis using the variables tenure (how long a customer has been with the company) and contract term (month-to-month, one year, or two years). The following are the results:

A screenshot of a computer screen

AI-generated content may be incorrect.

As you can see above, the F Value is 278.08 and the P-Value is <0.0001. The F value tells us that they are strongly correlated. It is also positive, so as one increases, so does the other. This means, the longer-term contract, the longer someone is likely to stay on as a user. The P-Value confirms that contract length and churn are significantly associated (p < .0001), indicating that shorter contract types are strongly correlated with increased churn.

To further support this conclusion, I utilized one of the other variables available in the dataset which was monthly cost, or how much a user is paying monthly, regardless of contract term. I ran a correlation analysis with a scatter plot testing the correlation between monthly cost and tenure. The following are my results:

A screen shot of a screen

AI-generated content may be incorrect.

As you can see above, there is somewhat of a curve that trends upwards and skews left. The Pearson correlation coefficient tells us that the relationship between tenure and monthly cost is nowhere near as strong as that between tenure and contract term. This is really interesting, and may suggest that people do not care as much about the cost of their subscription but other features like advertisements, content available, etc.

Based on all of this analysis, we can accept the alternative hypothesis and reject the null. Shorter contract lengths are significantly correlated with customer churn.

**Recommendations**

*Hypothesis 1: Content Library Volume versus Subscriber Numbers, Null Hypothesis Accepted*

Since we rejected the alternative hypothesis and accepted the null hypothesis, we can conclude that the data suggests content volume alone is not a strong predictor of subscriber retention.

This indicates that library strategy rather than raw volume is more strongly tied to long-term success. The hypothesis was not strongly supported by a significant statistical relationship, but further multivariate testing may reveal interactions between other important variables.

My recommendation is that streaming platforms analyze the success of their recommendation algorithms and the platform’s overall “findability”. This analysis has led us to believe that a giant like Netflix cannot significantly or directly contribute their success to having the largest library. If a streaming platform like Netflix were to further analyze how well their like/dislike model works and how widely it is being used, they would be able to gather further insight into what can sustain viewership and high subscriber numbers. Making content easy to find and making suggestions that viewers like may be the ultimate key to success.

*Hypothesis 2: Viewer Ratings Effects on Subscribership, Null Hypothesis Accepted*

After analyzing the data available, the null hypothesis was accepted and the alternative was rejected. My recommendation is that companies utilize their internal data to do further analysis. I would suggest that platforms who have access to the proprietary data that I am missing use it to maybe gauge correlation between viewer ratings and profits rather than subscribership. There is more variation seen in profits and the data comes with a length of dates rather than one snapshot of time. The analysis proved possibly significant, so I think it is important to continue to explore ratings but just in different ways.

*Hypothesis 3: Contract Length and Subscriber Churn, Null Hypothesis Rejected*

After careful and thorough analysis, we were able to reject the null hypothesis and accept the alternative hypothesis. Shorter contract lengths are significantly correlated with customer churn.

Knowing this, I recommend that companies take numerous steps. The first is to increase the cost of short-term contracts or special sign-up events. For example, if the platform will be broadcasting a highly anticipated sports event live, it would be in their best interest to offer a one time subscription ticket. This could mean that the user gets access to the one live event and then that access ends. This is an extremely short-term contract but, the company can charge substantially more for it since it is a special, one time thing.

The second recommendation I would make is more internal advertising. This analysis has shown that individuals with long-term contracts are much less likely to be a churned customer. This would imply that it is in companies’ best interests to turn as many short-term contract users as possible into long-term. By advertising on their own platform, such as banners on their web players or in program ads, they are saving the cost of buying ad space elsewhere. Companies would benefit from incentivizing longer term contracts (maybe a dollar less per month if you choose a two-year term for example) and then internally advertising those incentives to try and switch those with month-to-month contracts.

I think companies would also benefit from conducting further analysis into the timing of customer churn. If they can identify certain patterns like say for example, they get significantly more customers churning during the winter months/holidays, then they can stem off that activity. They could do this by offering incentives for long-term contracts during the winter months or in the reverse, charging higher prices for short-term during the winter months.

**Conclusion**

This analysis explored three hypotheses related to streaming service performance metrics; content library volume, viewer ratings, and contract length. It analyzed their impact on subscriber behavior. Across the first two hypotheses, the null hypotheses were accepted, indicating that neither content volume nor viewer ratings alone are strong predictors of subscribership. These findings suggest that streaming platforms should shift their focus from quantity and average ratings to enhancing content discoverability and leveraging internal data for more nuanced insights, such as profitability trends over time.

The third hypothesis, however, yielded a statistically significant result: shorter contract lengths are closely associated with increased churn. This finding highlights a clear opportunity for platforms to stabilize their subscriber base through strategic pricing, exclusive access models, and internal promotion of long-term contracts. Additional research into churn timing may uncover seasonal patterns that can be addressed through targeted retention efforts.

Together, these findings suggest that subscriber behavior is shaped less by surface-level metrics and more by user experience and engagement strategies. For sustained growth, streaming companies must invest in intelligent content recommendation systems, flexible pricing models, and data-driven contract structures that encourage long-term customer loyalty.

**References**

Apple’s “For All Mankind” is one of top streaming shows in the US this week - and season 2 has a 100% Rotten Tomatoes critic score. (2021). The Business Insider (Blogs on Demand).

Brzoznowski, K. (2024, February 22). U.S. TV Viewing Hits Four-Year High. World Screen. https://worldscreen.com/u-s-tv-viewing-hits-four-year-high/

Buckley, T. (2024). Disney Targets $1 Billion in Streaming Profit in Fiscal 2025. Bloomberg.Com, N.PAG.

Eakin, M. (2024, June 27). Why Streaming Services Keep Screwing Up Binge-Watching. Wired. https://www.wired.com/story/streaming-services-keep-screwing-up-binge-watching/

Gregg, A., & Cho, K. K. (2025, January 22). Netflix hikes prices as its lead widens over other streaming services. The Washington Post.

Holloway, D. (2015, January 5). Worth watching, driven by demand: the need for more original content is fueling trends at networks, studios and streaming services. Broadcasting & Cable, 145(1).

Huston, C. (2022, November 2). The Streaming Service That’s Quietly Gaining Major Market Share: Paramount’s Pluto TV, which eschews original content in favor of leveraging CBS’ library programming, is on track to reach up to 120 million monthly users powering the free, ad-supported streamer. Hollywood Reporter, 428(32).

Prince, J., & Greenstein, S. (2018). Does Original Content Help Streaming Services Attract More Subscribers? Harvard Business Review Digital Articles, 2–4.

Weprin, A. (2023, March 1). The One Metric to Rule Them All: Streaming Profits. Hollywood Reporter, 429(8).