**Module 2: Critical Thinking Assignment**

Sophia Balentine

Colorado State University Global

MIS581 - Capstone - Business Intelligence and Data Analytics

Dr. Osama Morad

June 22, 2025

**Table of Contents**

|  |  |
| --- | --- |
| Page 3 | Dataset Descriptions and Dictionaries |
| Page 11 | Tools/Techniques for Analysis |
| Page 12 | Rationale for Dataset Choice |
| Page 14 | References |

**Dataset Descriptions and Dictionaries**

I have found nine datasets so far that I am utilizing for this project. I imported all of the datasets into SAS Studio and ran them to ascertain the variable lists and data dictionaries. The following were my results:

1. User Churn
   1. Description: This is a simulated dataset that observes the customers of a fictitious telecommunications company and what factors contribute to churn (quickly joining and leaving their contract).
   2. Variables List:

A screenshot of a table

AI-generated content may be incorrect.

* 1. Data Dictionary:

A screenshot of a computer

AI-generated content may be incorrect.

1. Rotten Tomato Scores
   1. Description: This dataset includes the ratings from both audience and critics pulled from the popular website Rotten Tomatoes for all original television shows created by streaming platforms.
   2. Variables List:

A table with numbers and letters

AI-generated content may be incorrect.

* 1. Data Dictionary:

A screenshot of a computer

AI-generated content may be incorrect.

1. Streaming Services
   1. Description: This is a list of all current (as of the creation of the dataset) streaming platforms including number of subscribers, areas served, and other important information to discern popularity amongst them.
   2. Variables List:

A table with numbers and letters

AI-generated content may be incorrect.

* 1. Data Dictionary:

A screenshot of a computer

AI-generated content may be incorrect.

1. Netflix Stock Price
   1. This is a dataset containing Netflix’s, the publicly traded streaming giant, stock prices from 2002 to 2024.
   2. Variables List:

A screenshot of a computer

AI-generated content may be incorrect.

* 1. Data Dictionary:

A screenshot of a computer

AI-generated content may be incorrect.

1. Amazon Prime Video TV Shows with Ratings
   1. Description: This is a listing of all original TV productions from Amazon Prime Video, their IMDb rating, and the demographics of viewers.
   2. Variables List:

A screenshot of a table

AI-generated content may be incorrect.

* 1. Data Dictionary:

A screenshot of a computer

AI-generated content may be incorrect.

1. Amazon Prime Video Content
   1. Description: A list of all content available on Amazon Prime Video (at the time the dataset was created).
   2. Variables List:

A screenshot of a table

AI-generated content may be incorrect.

* 1. Data Dictionary:

A screenshot of a computer

AI-generated content may be incorrect.

1. Disney+ Titles
   1. Description: A list of all content available on Disney+ (at the time the dataset was created).
   2. Variables List:

A table with numbers and letters

AI-generated content may be incorrect.

* 1. Data Dictionary:

A screenshot of a computer

AI-generated content may be incorrect.

1. Hulu Titles
   1. Description: A list of all content available on Hulu (at the time the dataset was created).
   2. Variables List:

A screenshot of a table

AI-generated content may be incorrect.

* 1. Data Dictionary:

A screenshot of a computer

AI-generated content may be incorrect.

**Tools/Techniques for Analysis**

For my Capstone Project, I will be using SAS Studio, Microsoft Excel, and Tableau to combine, clean, and analyze the multiple datasets I have collected related to streaming services. By leveraging SAS Studio’s advanced data manipulation and statistical analysis tools, I will merge various data sources, perform comprehensive data cleaning to ensure accuracy and consistency, and run detailed analyses to uncover meaningful trends and relationships. I will also use Microsoft Excel for preliminary exploration, creating pivot tables, checking for data anomalies, and summarizing key metrics in an easily accessible format. Finally, I will utilize Tableau to design dynamic, interactive visualizations and dashboards that clearly communicate findings and make it easier for stakeholders to interpret the results. Together, these tools will help me generate valuable insights and develop actionable recommendations to increase viewership and maximize profits for streaming platforms.

**Rationale**

The nine 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 original content is as popular as licensced product. 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 of the library certain companies are devoting to original content, licensed content, or different genres. If a certain streaming service is more successful and has only 20% comedic TV shows and a heavier focus in dramatic movies, 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. This also applies to the Netflix Stock Prices dataset, which introduces a financial perspective, enabling us to explore correlations between content strategies, audience reception, and a company’s market performance.

By combining customer data, content ratings, platform comparisons, and financial metrics, I can generate a multidimensional analysis that highlights the drivers behind viewership trends and profitability. 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.

**References**

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/?utm\_source=chatgpt.com

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).