

DATAFLIX : THE OTT INTELLIGENCE DASHBOARD

PROJECT REPORT

By

ADITI BORKAR (RA2211027010030)

VIBHOR PUNDHIR (RA2211027010031)

PARAS TIKOO (RA2211027010048)

KENISHA SURANA (RA2211027010078)

Under the guidance of

Ms. Gayathri M

In partial fulfilment for the Course

of

21CSE421T – BUSINESS INTELLIGENCE AND ANALYTICS

in

COMPUTER SCIENCE AND ENGINEERING

WITH SPECIALIZATION IN BIG DATA ANALYTICS



FACULTY OF ENGINEERING AND TECHNOLOGY

SCHOOL OF COMPUTING

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

KATTANKULATHUR

NOVEMBER 2025



SRM INSTITUTE OF SCIENCE AND TECHNOLOGY KATTANKULATHUR – 603 203

(Under Section 3 of UGC Act, 1956)

BONAFIDE CERTIFICATE

Certified that this minor project report for the course **21CSE421T BUSINESS INTELLIGENCE AND ANALYTICS** entitled in "**DATAFLIX: THE OTT INTELLIGENCE DASHBOARD**" is the bonafide work of **ADITI BORKAR [RA2211027010030]**, **VIBHOR PUNDHIR [RA2211027010031]**, **PARAS TIKOO (RA2211027010048)** and **KENISHA SURANA [RA2211027010078]** who carried out the work under my supervision.

SIGNATURE

Ms. Gayathri M

Assistant Professor

**Data Science and Business Systems,
SRM Institute of Science and Technology
Kattankulathur.**

SIGNATURE

Dr. Kavitha V

Professor and Head

**Data Science and Business Systems,
SRM Institute of Science and Technology
Kattankulathur.**

ABSTRACT

The 21st-century media landscape is dominated by the "Streaming Wars," a period of intense competition between Over-The-Top (OTT) platforms such as Netflix, Prime Video, Disney+, and Hulu. In this data-driven battlefield, content is the primary asset, and strategic decisions regarding acquisitions, original productions, and market positioning are critical for survival and growth. However, data on content libraries is often siloed, making cross-platform competitor analysis a significant challenge for industry stakeholders.

This project, titled "DataFlix," presents the design and implementation of a high-performance, interactive Business Intelligence (BI) dashboard to address this challenge. The system aggregates, cleans, and processes metadata from four major streaming platforms, transforming disparate datasets into a unified analytical model.

The core of the project is a multi-page web application built using Python, Streamlit, and Pandas. A robust data preparation pipeline was developed to handle significant data inconsistencies, including varied date formats, non-numeric duration strings (e.g., "1 Season" vs. "120 min"), and multi-value categorical columns (e.g., listed_in, country). Exploratory Data Analysis (EDA) was performed using Plotly, which was then integrated into a series of feature-rich dashboards, with over 10 distinct visualizations per platform.

Furthermore, the application is enriched with real-time data through integration with The Movie Database (TMDb) API, providing functionality for live title searches, review fetching, and trending content feeds. The result is a production-ready, aesthetically refined analytical tool that provides deep insights into content strategies, audience targeting, library composition, and geographic footprint, effectively "Turning Streaming Data into Strategy."

Keywords: OTT, Streamlit, Data Visualization, Business Intelligence, Data Analytics, Python, Pandas, Plotly, API Integration, Data CleaningWerwrr

TABLE OF CONTENTS

CHAPTER NO	CONTENTS	PAGE NO
1	INTRODUCTION	4
	1.1 Motivation	5
	1.2 Objective	6
	1.3 Problem Statement	7
	1.4 Challenges	7
2	DATA UNDERSTANDING	9
3	DATA PREPARATION	13
4	EXPLORATORY DATA ANALYSIS (EDA)	18
5	RESULTS AND DISCUSSION	34
6	CONCLUSION	37
7	REFERENCES	39
	APPENDIX	40

CHAPTER 1

INTRODUCTION

The advent of high-speed internet and the consumer shift from traditional broadcast television to on-demand content has catalyzed a global media revolution. This revolution is defined by the rise of Over-The-Top (OTT) streaming services. Platforms such as Netflix, Amazon Prime Video, Disney+, and Hulu have fundamentally reshaped how entertainment is produced, distributed, and consumed worldwide. This disruptive shift has triggered what is now ubiquitously known as the "Streaming Wars"—a period of intense and accelerating competition where these global giants, along with numerous regional players, vie for subscriber market share.

In this hyper-competitive landscape, content is king. The strategic decisions a platform makes regarding its content library—what to license, what to produce as an "Original," which genres to invest in, which markets to target, and how long to retain content—are the primary drivers of subscriber acquisition and retention. A single hit show can attract millions of new users, while a perceived gap in a specific genre can lead to subscriber churn.

Consequently, the ability to perform accurate, timely, and comprehensive analysis of competitor content strategies is no longer a luxury but a critical business necessity. Stakeholders, from studio executives and content acquisition managers to marketing teams and financial analysts, require actionable intelligence to navigate this complex market. They must answer critical questions: How does our library size compare to our competitors? Are we investing in the right genres? What is the age of our content library versus our rivals? Where are our competitors focusing their production efforts geographically?

This project, "**DataFlix: The OTT Intelligence Dashboard,**" is conceived as a direct solution to this need. It is a comprehensive, multi-platform Business Intelligence (BI) tool designed to aggregate, analyse, and visualize the content libraries of the four major US-based streaming platforms. By transforming raw, disparate data into an interactive, aesthetically refined dashboard, DataFlix provides a unified command center for understanding the streaming landscape and "Turning Streaming Data into Strategy."

1.1 MOTIVATION

The primary motivation for this project stems from a significant and persistent gap in the public-facing analytics market. While individual platforms possess vast internal datasets, this information is proprietary and siloed. For external analysts, researchers, content creators, or even curious consumers, performing a comparative analysis is a manual, fragmented, and arduous process. It requires locating, downloading, cleaning, and individually analyzing multiple, often incompatible datasets.

The key motivators for this project are:

- **For Industry Stakeholders:** To create a proof-of-concept BI tool that demonstrates how a unified dashboard can empower decision-makers. By visualizing competitor strategies side-by-side, a studio can identify market gaps, benchmark its own library, and discover emerging trends.
- **For Data Scientists & Analysts:** To showcase a complete, end-to-end data project. This includes the entire data lifecycle: sourcing disparate data, building a robust cleaning and preparation pipeline (a significant data engineering challenge), integrating external APIs, and, most importantly, building a user-facing product (the dashboard) that delivers insights effectively.
- **For Technical Challenge:** The project serves as a comprehensive technical exercise in modern data stack implementation. It provides a practical application for advanced Python scripting, mastery of the Pandas library for complex data manipulation, the use of Plotly for creating rich, interactive visualizations, and the art of UI/UX design within the Streamlit framework to build a professional, shippable product.
 1. **For Public Interest:** The "Streaming Wars" are a constant topic of public and media fascination. A tool that allows users to explore and compare the libraries of their favorite services answers common questions (e.g., "Which platform has the most movies?", "Is Netflix focusing more on TV shows now?") in a data-driven, objective manner.

1.2 OBJECTIVE

The primary objective of this project is to design, develop, and deploy a feature-rich, multi-page Streamlit dashboard that provides a comprehensive analytical overview and comparison of the content libraries of Netflix, Prime Video, Disney+, and Hulu.

To achieve this, the following specific objectives were set:

- **To Aggregate Data:** To source and load four distinct datasets, one for each streaming platform, into a unified analytical environment using Python and Pandas.
- **To Build a Robust Data Pipeline:** To design and implement a sophisticated data preparation and cleaning pipeline to handle a wide range of data quality issues, such as missing values, inconsistent formats, and non-numeric strings in numerical columns.
- **To Standardize and Feature Engineer:** To process and transform raw data into analysis-ready formats. This includes parsing multi-value columns (like listed_in and country), standardizing duration metrics (movies vs. TV shows), and extracting new features (e.g., year_added, month_added).
- **To Develop Comprehensive Visualizations:** To create a suite of 10-12 insightful and interactive visualizations for each platform using Plotly. These charts are designed to analyze content types, genre distributions, rating profiles, temporal trends, and geographic footprints.
- **To Implement Comparative Analytics:** To build a functional "Head-to-Head" tool on the home page, allowing users to select any two platforms for a direct comparison of their key library metrics.
- **To Integrate Real-Time Data:** To enrich the dashboard by integrating the TMDb (The Movie Database) API, enabling users to search for any title, view its poster and description, read user reviews, and see a list of currently trending content.
- **To Achieve a Professional UI/UX:** To move beyond default Streamlit styling by implementing advanced custom CSS to create a beautiful, modern, and aesthetically pleasing "glassmorphism" interface with a dark theme, animated icons, and a clean, intuitive layout.
- **To Generate Automated Insights:** To develop a BI (Business Intelligence) module that automatically generates and displays qualitative, data-driven recommendations and insights for each platform.

1.3 PROBLEM STATEMENT

The proliferation of Over-The-Top (OTT) streaming services has led to a highly fragmented and competitive market. Billions of dollars are spent annually on content acquisition and original production. For stakeholders—ranging from media conglomerates and production studios to financial analysts and marketing agencies—a deep understanding of the competitive landscape is essential for making high-stakes strategic decisions.

The core problem is:

- Decision-makers, researchers, and analysts lack a unified, accessible, and interactive tool to perform comparative analysis of content libraries across major streaming platforms. Data is siloed, inconsistent, and difficult to aggregate, preventing an effective, data-driven assessment of competitor strategies.

This project, "DataFlix," aims to solve this problem by designing and building a web-based intelligence dashboard. The system will ingest, process, and harmonize data from Netflix, Prime Video, Disney+, and Hulu, presenting the findings in a single, feature-rich interface. The dashboard will empower users to analyze global market trends, conduct deep dives into specific platform strategies, and perform direct "Head-to-Head" comparisons, thereby providing the clarity and insight needed to navigate the complex OTT market.

1.4 CHALLENGES

The development of the DataFlix dashboard, while successful, presented a number of significant technical and analytical challenges that had to be overcome.

1. **Data Inconsistency and Cleaning:** This was the most significant challenge. The four datasets, originating from different sources, were not standardized.

- **Duration:** Netflix and Prime Video movies had durations like "120 min," while TV shows had "2 Seasons." Disney+ and Hulu had similar but not identical formats. A robust parsing function was required to extract numerical values and handle these two distinct metrics.
- **Ratings:** The rating column was highly inconsistent, containing values like "TV-MA," "R," "PG-13," "18+," "ALL," and sometimes numerical values.

- **Date Formats:** The date_added column had multiple different string formats and missing values, requiring flexible parsing with pd.to_datetime and an errors='coerce' strategy.
- **Handling "Unknown" Values:** Some data (e.g., duration in Netflix) contained text values like "Unknown," which would crash numerical conversion. The pipeline had to be resilient to these non-standard "null" values.

2. **Data Transformation (Multi-Value Columns)** : A single title can have multiple genres or countries of origin (e.g., "Dramas, International Movies"). A simple value_counts() on this column is inaccurate. The solution required using the Pandas .explode() method to split these comma-separated strings into new rows, a process that significantly increases the dataset's size but is essential for correct analysis.

3. **UI/UX Customization and Layout Bugs** : Streamlit is designed for rapid development, but achieving a custom, professional aesthetic is challenging.

- **Advanced Styling:** Implementing the "glassmorphism" dark theme required extensive custom CSS, which at times conflicted with Streamlit's native components.
- **Layout Glitches:** The most persistent challenge was the appearance of "empty boxes" and misaligned components. This was caused by a conflict between custom HTML/CSS (<div class="card">) and Streamlit's native containers (st.container). The final solution required removing all custom HTML wrappers and exclusively using and styling Streamlit's native st.container(border=True).
- **Component Styling:** Achieving a clickable logo card without a visible button required a CSS "overlay" trick, where an invisible button was stretched to cover the entire container.

4. API Integration:

- **Security:** Managing the TMDb API key securely was paramount, solved by using Streamlit's built-in st.secrets manager.
- **Resilience:** The API can fail (network error, invalid key, missing title). The code in api_utils.py had to be wrapped in try...except blocks to handle these failures gracefully and return None or an empty list, allowing the app to display a warning instead of crashing.

CHAPTER 2

DATA UNDERSTANDING

A robust data-driven analysis is fundamentally dependent on the quality and comprehensiveness of the underlying data. Before any processing or visualization could occur, a thorough data understanding phase was conducted to identify, source, and profile all required datasets. This chapter details the data sources, their attributes, and their suitability for this project.

2.1 DATA SOURCES

The DataFlix dashboard is powered by two distinct types of data sources: a static, historical core and a dynamic, real-time enrichment layer.

2.1.1 Primary Data Source: Kaggle Datasets

The foundational data for this project consists of four separate, publicly available datasets sourced from the Kaggle data science community. Each dataset contains a historical snapshot of the content library for one of the four target platforms:

1. **Netflix:** [netflix_titles.csv](#)
2. **Amazon Prime Video:** [amazon_prime_titles.csv](#)
3. **Disney+:** [disney_plus_titles.csv](#)
4. **Hulu:** [hulu_titles.csv](#)

These datasets were chosen for their comprehensiveness, providing a rich set of metadata for thousands of titles, including titles, directors, cast, genres, release years, and addition dates. While not officially published by the platforms themselves, they are well-maintained community datasets that serve as an excellent proxy for analyzing high-level content strategy. Collectively, these datasets provided metadata for over 22,000 unique titles (as seen in our "Global Streaming Landscape" KPI).

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description									
2	s1	Movie	Dick Johnson Kirsten Johnson	United States	September 2	2020	PG-13	90 min	Documentary	As her father nears the end of his life, filmmaker Kirsten Johnson stages his death in inventive and comical ways to help them both face the inevitable.											
3	s2	TV Show	Blood & Water	Ama Qamata	South Africa	September 2	2021	TV-MA	2 Seasons	International	After crossing paths at a party, a Cape Town teen sets out to prove whether a private-school swimming star is her sister who was abducted at birth.										
4	s3	TV Show	Ganglands	Julien Leclerc Sami Bouajila, Tracy Goto	September 2	2021	TV-MA	1 Season	Crime TV	To protect his family from a powerful drug lord, skilled thief Mehdi and his expert team of robbers are pulled into a violent and deadly turf war.											
5	s4	TV Show	Jailbirds New Orleans		September 2	2021	TV-MA	1 Season	Documentaries, F	Feuds, flirtations and toilet talk go down among the incarcerated women at the Orleans Justice Center in New Orleans on this gritty reality series.											
6	s5	TV Show	Kota Factory	Mayur More, India	September 2	2021	TV-MA	2 Seasons	International	In a city of coaching centers known to train India's finest collegiate minds, an earnest but unexceptional student and his friends navigate campus life											
7	s6	TV Show	Midnight Mat	Mike Flanagan Kate Siegel, Zach Gilford, I-	September 2	2021	TV-MA	1 Season	TV Dramas, T	The arrival of a charismatic young priest brings glorious miracles, ominous mysteries and renewed religious fervor to a dying town desperate to believe.											
8	s7	Movie	My Little Ponies Robert Cullen Vanessa Hudgens, Kinko C	September 2	2021	PG	91 min	Children & F	Equestria's divided. But a bright-eyed hero believes Earth Ponies, Pegasi and Unicorns should be pals. At and, hoof to heart, she's determined to prove												
9	s8	Movie	Sankofa	Halle Berry Kofi Ghana	United States	September 2	1993	TV-MA	125 min	Children & F	Equestria's divided. But a bright-eyed hero believes Earth Ponies, Pegasi and Unicorns should be pals. At and, hoof to heart, she's determined to prove										
10	s9	TV Show	The Great Bri Andy Devons Mel Giedroyc United Kingdom	September 2	2021	TV-14	90 min	British TV	A talented batch of amateur bakers face off in a 10-week competition, whipping up their best dishes in the hopes of being named the U.K.'s best.												
11	s10	Movie	The Starling Theodore Me Melissa McC	United States	September 2	2021	PG-13	104 min	Comedies, A woman adjusting to life after a loss contends with a feisty bird that's taken over her garden. At and a husband who's struggling to find a way forward.												
12	s11	TV Show	Vendetta: Truth, Lies & The Mafia		September 2	2021	TV-MA	1 Season	Crime TV	She To protect his family from a powerful drug lord, skilled thief Mehdi and his expert team of robbers are pulled into a violent and deadly turf war.											
13	s12	TV Show	Bangkok Brie Kongkak Kor Sokolluwat Karanot, Susha	September 2	2021	TV-MA	1 Season	Crime TV	She Struggling to earn a living in Bangkok, a man joins an emergency rescue service and realizes he must unravel a citywide conspiracy.												
14	s13	Movie	Je Suis Kart Christian Scl Luna Wedler, Cz	September 2	2021	TV-MA	127 min	Dramas, Inte	After most of her family is murdered in a terrorist bombing, a young woman is unknowingly left to joining the very group that killed them.												
15	s14	Movie	Confessions Bruno Garret Clara Castaño, Luca Pici	September 2	2021	TV-PG	91 min	Children & F	When the clever but socially-awkward Tet-jo joins a new school, she'll do anything to fit in. But the queen bee among her classmates has other ideas.												
16	s15	TV Show	Crime Stories: India Detectives		September 2	2021	TV-MA	1 Season	British TV	Studios Cameran following Bengaluru police on the job offer a glimpse into the complex and challenging inner workings of four major crime investigations.											
17	s16	TV Show	Dear White People Logan Brown	United States	September 2	2021	TV-MA	4 Seasons	TV Comedies	Students of color navigate the daily slights and slippery politics of life at an Ivy League college that's not nearly as "post-racial" as it thinks.											
18	s17	Movie	Europe's Mos Pedro de Echave Garc, I-a, Pablo Azor	September 2	2020	TV-MA	67 min	Documentary	Declassified documents reveal the post-WWII life of Otto Skorzeny, a close Hitler ally who escaped to Spain and became an advisor to world presidents												
19	s18	TV Show	Falsa Identidad Luis Ernesto I Mexico	September 2	2020	TV-MA	2 Seasons	Crime TV	She Stirs Diego and Isabel their home in Mexico and pretend to be a married couple to escape his drug-dealing enemies and her abusive husband.												
20	s19	Movie	Intrusion Adam Salter	Freida Pinto, Logan Marshall-Grange	September 2	2021	TV-14	94 min	Thrillers	After a deadly home invasion at a couple, a woman who's been the target of a serial killer, the traumatised wife searches for answers. At and learns the real danger is just beginnin											
21	s20	TV Show	Jaguar Blanca Su, *vez, Iv, Mar	September 2	2021	TV-MA	1 Season	International	In the late 1960s, an accused serial rapist claims multiple personalities control his behavior, setting off a legal odyssey that captivates America.												
22	s21	TV Show	Monsters Ins Oliver Metzger		September 2	2021	TV-14	1 Season	Crime TV	In the late 1970s, a sultan's enemies to fight a sultan's enemies in exchange for never tribal war.											
23	s22	TV Show	Resurrection: Ertugrul Ergin Altan D Turkey	September 2	2018	TV-14	5 seasons	International	When a good deed unwittingly endangers his clan, a 13th-century Turkish warrior agrees to fight a sultan's enemies in exchange for never tribal war.												
24	s23	Movie	Avva Shamkii K Ravikum Karan Hassan, Meena, Geeta	September 2	1996	TV-PG	161 min	Comedies, F	In Newly divorced and denied visitation rights with his daughter, a doting father disguises himself as a gray-haired nanny in order to spend time with her. Toy Carson, Avva's curious little sister, Chrissy speeds off on her own for fun and adventure all over tow												
25	s24	Movie	Go! Go! Cory Alex Woz, St Maisie Williams, Paul Kilian	September 2	2021	TV-14	61 min	Children & F	From arcade games to sled days and hiccups, Cory Carson, Avva's curious little sister, Chrissy speeds off on her own for fun and adventure all over tow												
26	s25	Movie	Jean S. Shankar Prashanth, Al India	September 2	1998	TV-14	166 min	Comedies	In When the man she loves insists that his twin sons many twin sisters, a woman creates an alter ego that might be a bit too convincing.												
27	s26	TV Show	Love on the Spectrum Brooke Smith Australia	September 2	2021	TV-14	2 Seasons	Documentaries, F	Finding love can be hard for anyone. For young adults on the autism spectrum, exploring the unpredictable world of dating is even more complicated.												
28	s27	Movie	Misraa Kanji Rajiv Mehta, Arvind Swamy, Kajol, Prabh Shetty	September 2	2019	TV-PG	147 min	Comedies	After a deadly home invasion at a couple, a woman who's been the target of a serial killer, the traumatised wife searches for answers. At and learns the real danger is just beginnin												
29	s28	Movie	Grown Ups Dennis Dugan, Michael Richards, David Spade	September 2	2010	PG-13	103 min	Comedies	Mourning the loss of their beloved junior high basketball coach, five middle-aged pals reunite at a lake house and rediscover the joys of being a kid.												
30	s29	Movie	Dark Skies Scott Stewart Keri Russell, Amy Landecker	September 2	2013	TV-14	97 min	Horror Movie	Horror Movie Family, Ays idyllic suburban life shatters when an alien force invades their home, and as they struggle to convince others of the deadly threat.												
31	s30	Movie	Paranoia Robert Luketic John Goodman, Philip Baker Hall	September 2	2013	PG-13	106 min	Thrillers	Blackmailed by his company's CEO, a low-level employee finds himself forced to spy on the company's rival and former mentor.												
32	s31	TV Show	Ankhali Kahar Ashwini Iyer Abhishek Banerjee, Rinku F	September 1	2021	TV-MA	111 min	Documentaries	India's Big City life buzzes around these lonely souls discover surprising sources of connection and ownership in three tales of love, loss and longing.												
33	s32	TV Show	Chicago Party Lauren Lazin, Rob Minkoff, Leah Remini	September 1	2021	TV-MA	1 Season	TV Comedies	Chicago Party Aunt Diane is an idealized troublemaker with a talent for providing guidance, At and a soft spot for soul-searching nephew.												
34	s33	TV Show	Sex Education Asa Butterfield, Niamh Algar, Jamie Sives	September 1	2020	TV-MA	3 Seasons	British TV	Sex Education Iris has all the answers when it comes to sex advice, thanks to his therapist mom. So rebel Maeve proposes a school sex-therapy clinic.												
35	s34	Movie	Squid Game Lee Jung-Jae, Park Hee-soo	September 1	2021	TV-14	1 Season	International	Hundreds of cash-strapped players accept a strange invitation to compete in children's games. Inside, a tempting prize awaits.												
36	s35	TV Show	Tavo and Little Wizards Dame Lee, Jason Lee, Jason Lee	September 1	2020	TV-Y7	1 Season	Kids' TV	Tavo speeds into an adventure when his friends are kidnapped by evil magicians invading their city in search of a magical emerald.												

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description									
2	s1	Movie	The Grand Se Don McCella Brendan Gleeson	Ireland	March 30, 20	2014	113 min	Comedy, Drama	A small fishing village must procure a local doctor to secure a lucrative business contract. When unlikely candidate and big city doctor Paul Lewis lands												
3	s2	Movie	Take Care Gic Girish Joshi	Manisha Man, India	March 30, 20	2018	13+	110 min	Drama, Inter	At Metro Family decides to fight a Cyber Criminal threatening their stability and pride.											
4	s3	Movie	Secrets De Jeff Webber Tom Szemer	United States	March 30, 20	2017	74 min	Action, Drama	After a man discovers his wife is cheating on him with a neighborhood kid he goes down a furious path of self-destruction												
5	s4	Movie	Pink: Staying Sonia Andress Interviews wi	United States	March 30, 20	2014	69 min	Documentary	Pink breaks the mold once again, bringing her career to a new level in 2013 with a world tour that entertains unlike ever before! Get inside access to "th												
6	s5	Movie	Monster Max Giles Foster	Harry Dean	United Kingdom	March 30, 20	1989	45 min	Drama, Fanti	Teenage Matt Banning works to become a run, At and she fails for the friend to help him realize the importance of his work.											
7	s6	Movie	Living With D Paul Weiland Gregory	United Kingdom	March 30, 20	1989	52 min	Fantasy, Kids	The story unfolds in a tiny English seaside town, Where Dom, an only child, faces the imminent arrival of a new sibling, and subsequently diminished attention												
8	s7	Movie	Hired Gun Franne Alice Cooper	United States	March 30, 20	2017	98 min	Documentary	They're the First Call, A-list musicians, just 20 feet from stardom, yet rarely receive credit for their work. The 'hired gun' community lives and breathes												
9	s8	Movie	Grease Live! Thomas Kail, Julianne Hough United States	March 30, 20	2016	131 min	Comedy	This honest, uncompromising comedy chronicles the war stories and sexual misadventures of a tight circle of lovers and friends fighting to get their acts													
10	s9	Movie	Global Mettd Daniel Gilboe Michaela Parry Canada	March 30, 20	2017	87 min	Action, Sci-Fi	A helicopter pilot and an environmental scientist lead a exodus of survivors in a search for a safe haven after a catastrophic tectonic event causes the c													
11	s10	Movie	David's Morth Robert Allan Kirstie Alley, United States	April 1, 2021	1994	92 min	Drama	Sally Goodin is a devoted mother to her autistic son David. Abandoned by her husband, Sally has managed to keep her son out of "The System", until a s													
12	s11	Movie	Forest Fairie Justin D. G. Emily Wilder, Canada	April 1, 2021	2016	88 min	Adventure, Ki Amada stumbles upon a hidden village of fairies in the forest. They help her thwart a scheming land developer's plan to trick her mother into selling the														
13	s12	Movie	Take Care Uz Tuccillo Leslie Bibb, Kelli Berglund, Ira	April 1, 2021	2014	93 min	Comedy	When a car crash leaves Franklin immobilized, she is brushed off by everyone. Sam comes face to face with his new reality, An army of zombies have invaded the streets.													
14	s13	Movie	The Night Eat Dominique R Anders Danica	France	April 17, 2021	2018	94 min	Horror, Susp	After waking up in a suburban at the right after a raging party, Sam comes face to face with his new reality, an army of zombies have invaded the streets.												
15	s14	Movie	Resilience: Get Barcelon Jacqueline Aude	Spain	April 17, 2021	2020	46 min	Documentary	A documentary follows the meteoric rise of Elon Musk, the man who is transforming travel technology through electric cars, the Hyperloop, an												
16	s15	Movie	Summer '03 Becca Gleas Kaitlyn Dever United States	June 3, 2021	2019	96 min	Comedy, Drama	Zoe is the 2013 the hilarious journey of a 13-year-old girl who, along with her shocked mother, her older brother and her autistic brother, embarks on a wild road trip across the country to find her missing father.													
17	s16	Movie	Zoobies Glenn Miller Marcus Anderson, Kaitlyn Dever, Andrew	June 3, 2021	2016	13+	87 min	Horror, Scier	When a strange virus quickly spreads through a safari park and turns all the zoo animals undead, those left in the park must stop the creatures before the												
18	s17	Movie	Zoobies Bobbi Miller Marcus Anderson, Kaitlyn Dever, Andrew	June 3, 2021	2016	13+	87 min	Kids, Special	Species	Species	Species	Species	Species	Species	Species	Species	Species	Species	Species	Species	Species
19	s18	TV Show	Zoo Babies Bobbi Miller Marcus Anderson, Kaitlyn Dever, Andrew	June 3, 2021	2016	ALL	1 Season	Kids, Special	Species	Species	Species	Species	Species	Species	Species	Species	Species	Species	Species	Species	Species
20	s19	TV Show	Zo/ Coombs Marr: Bossy Zo/ Coombs Marr		2020	18+	1 Season	Comedy, Talk	Zo/ Coombs Marr has been on hiatus. Sort of. For six years, she won stacks of awards and acclaim as a hack comic with a neckbeard called Dave. Who												
21	s20	Movie	Zoe Drake Doren Evans McGregor, Léa Seydoux, Theo Jai	2018	2018	104 min	Science Fiction	Science Fiction	Zoe tells of for love between an engineer and a robot. Zoë (Léa Seydoux) and COLE (Ewan McGregor) are colleagues and veiled lovers at												
22	s21	TV Show	Zooboomf: Mini-Cales Alexandra, Jacqueline Garza, I	2021	TV-Y7	1 Season	Kids	Chris and Martin Kraft bring their enthusiasm for animals to the pre-school set.													
23	s22	TV Show	Zooboomf: Mini-Cales Alexandra, Jacqueline Garza, I	2021	TV-PG	1 Season	Action, Adventure	Zooboomf is our hero! With the defeat of Director Goodwin and the Dark Signers, New Domino City's barriers have been torn down. City and Satellite are now one - united in t													
24	s23	Movie	Zombie Sam Pilkington, Benji Webber, I	2021	TV-14	86 min	Drama	Talented and beautiful Carol Williams (Forest) is devastated when she learns she has contracted Polio and becomes crippled by the disease. The depr													
25	s24	Movie	Young Lovers Ida Lupino, Sally Forrest, Keefe Brasselle, Hugh O'B	2019	2007	NR	36 min	Music Videos	Young Love celebrates the release of "Too Young To Fight It" with a midnight show inside an old factory building.												
26	s25	Movie	Young Love (at the Sun Factory)		2007	NR	37 min	Music Videos	YoungLove returns to Baebie's silver screen with a briefer slate, expanded setlist, nicer clothes, and the same blend of rock, dance, and oo												

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R</th
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	-------

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	
1	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description										
2	s1	Movie	Ricky Velez: Here's Everything				October 24, 2021	TV-MA		94 min	Crime, Drama	Ricky Velez bares it all with his honest lens and down to earth perspective in his first-ever HBO stand-up special.										
3	s2	Movie	Silent Night				October 23, 2020			108 min	Action, Thriller	A hardened Arizona rancher tries to protect an 11-year-old migrant boy fleeing from a ruthless drug cartel.										
4	s3	Movie	The Marksman				October 23, 2021	PG-13		97 min	Horror	A forest ranger and two survivalists with a cultish devotion to the forest face the threat of an unrelenting wilderness when a strange being attacks one ni										
5	s4	Movie	Galaxy				October 22, 2021	R		104 min	Science Fiction	Mankind's earliest settlers on the Martian frontier do what they must to survive the cosmic elements and each other in this science-fiction thrill ride.										
6	s5	Movie	Settlers				October 22, 2021			88 min	Horror, Thriller	New to her role as a stepmom, a young woman moves into a duplex with her partner and his five-year-old son Lucas. Soon, strange noises start coming fr										
7	s6	TV Show	The Halloween Candy Magic Pet				October 22, 2021			1 Season	Family, Kids	Join Milia and Morphle on a mystery-filled Halloween adventure! Jam packed with morphs, costumes, twists turns and spooky morphing fun. This 22-min										
8	s7	Movie	The Evil Next Door				October 21, 2020			1 Season	Comedy & Fantasy	With the unique insights and experience of Ugly Delicious creators Morgan Neville and David Chang take a tangy look at food and our relationship to it p										
9	s8	TV Show	The Next Thing You Eat				October 21, 2021			1 Season	Cooking & Food	With the unique insights and experience of Ugly Delicious creators Morgan Neville and David Chang take a tangy look at food and our relationship to it p										
10	s9	TV Show	Queens				October 20, 2021	TV-14		1 Season	Drama, Music	Four women in their 40s reunite for a chance to recapture their fame and regain the swagger they had as legends in the 1990s hip-hop world.										
11	s10	TV Show	The Bachelorette	United States	October 20, 2023		2003	TV-14		3 Seasons	Reality, Romance	Romantic reality show lets one lucky lady narrow the field of bachelors, leaving only her dream man.										
12	s11	TV Show	The Real Queens of Hip-Hop: The Women Who Change October	19, 2021			2021			1 Season	Music, News	The ABC News special explores the inspiring rise of women in hip-hop and is narrated by Cheryl Tiegs.	James of Salt-N-Pepa.									
13	s12	Movie	Dream Horse				October 18, 2020	PG		113 min	Comedy, Drama	Dra film tells the inspiring true story of a small town bartender who decides to breed a racehorse. The community chips in their small earnings and the ir										
14	s13	Movie	Out of the Shadows: The Man Behind the Steele Dossier	October 18, 2021						87 min	Science Fiction	George Stephanopoulos sits down with former MI6 spy Christopher Steele for a worldwide exclusive interview, marking his first interview since the publ										
15	s14	TV Show	Wakefield				October 18, 2021			1 Season	Comedy, Drama	With a gift for soothing the afflicted and reaching the unreachable, brilliant psychiatric nurse Nik Katrini is the most stable person at Wakefield mental h										
16	s15	TV Show	Home Sweet Home				October 16, 2021	TV-PG		1 Season	Reality	Real families of different backgrounds experience a new way of life when they trade homes, stepping out of their comfort zones to discover each other's										
17	s16	Movie	Showtime Championship Boxing: Lopez vs. Salido (R)	October 16, 2021						Sports	Juan Manuel Lopez vs. Orlando Salido for the WBO Featherweight title.											
18	s17	TV Show	Showtime Championship Boxing: Wilder vs. Molina	October 16, 2021			2015	TV-14		1 Season	Sports	Champion Deontay Wilder vs. Eric Molina. Live from the Bartow Arena.										
19	s18	TV Show	America's Book of Secrets	United States	October 16, 2022		2012	TV-14		1 Season	Documentary	America's story can be told in many ways. While much of the history of our country is well documented and widely available, there is another deeper his										
20	s19	TV Show	Beyond Oak Island				October 15, 2022	TV-PG		1 Season	History, Real	From pirates such as Blackbeard and outlaws like Jesse James, to Aztec gold, priceless historical artifacts from American history and sunken treasure s										
21	s20	TV Show	Beyond Scarred Straight	United States	October 15, 2021		2011	TV-14		6 Seasons	Documentary	Beyond Scarred Straight profiles the new approach to keeping today's teens from becoming tomorrow's prisoners. The power of the original Scarred Straig										
22	s21	Movie	Cheer Camp Killer				October 15, 2021	TV-14		87 min	Thriller	Camp queen bee Victoria does not take kindly to a newcomer like Sophia taking her spotlight, so she launches a dangerous plan to get her rival out of the										
23	s22	TV Show	Hoarders	United States	October 15, 2020		2009	TV-PG		7 Seasons	Health & Wellness	Each 60-minute episode of Hoarders is a fascinating look inside the lives of two different people whose inability to part with their belongings is so out of										
24	s23	TV Show	Little Women: Atlanta	United States	October 15, 2021		2019	TV-14		2 Seasons	Black Stories in Little Women: Atlanta, the little ladies turn it up wherever they go out, especially with club promoter Emily leading the charge. There is no shortage of											
25	s24	TV Show	Marrying Millions	United States	October 15, 2021		2019	TV-14		4 Seasons	Documentary	"Marrying Millions" follows four couples who are deeply in love but are struggling to find the right person. The responders deal with the biggest shift of the day and the										
26	s25	TV Show	Highwater	United States	October 15, 2021		2015	TV-14		4 Seasons	Documentary	"Highwater" follows the dark web of the Deep South, where the [between Spanish and English] [between Spanish and English]										
27	s26	TV Show	Seven Year Switch				October 15, 2017	TV-14		1 Season	Reality	In Seven Year Switch, a couple at a crossroads in their relationship get a chance to live with a stranger for two weeks in an experimental marriage.										
28	s27	Movie	Sleepwalker				October 15, 2017	TV-14		68 min	Thriller	A graduate student plagued by bouts of sleepwalking begins to have trouble differentiating between her dreams and reality.										
29	s28	TV Show	Swamp People	United States	October 15, 2010		2010	TV-PG		3 Seasons	Documentary	Deep in the heart of Louisiana lies America's largest swamp. At 1 million miles of inesplorable bayous, marshes and wetlands where nature rules and hu										
30	s29	TV Show	Unsolved				October 15, 2019			2 Seasons	Documentary	More than 750 people are missing from Houston and the surrounding towns of the major Texas city. Among the missing are mothers written off by police.										
31	s30	TV Show	Acapulco Shore				October 14, 2014	TV-14		3 Seasons	Latino	Real! A group of young Mexicans live together and share fun, drama and a lot of alcohol.										
32	s31	TV Show	Catfish H/éxico				October 14, 2018			1 Season	Latino	Real! Porque sabemos que no es fcil encontrar el amor verdadero en este mundo digital, MTV Catfish llega a Mxico para salvarte de las mentiras. Todos										
33	s32	Movie	Censor				October 14, 2021			84 min	Horror	A film censor is assigned to review a disturbing film that recalls her childhood memories. She begins to unravel how this eerie work might be tied to the kr										
34	s33	TV Show	Dani Who?				October 14, 2019			1 Season	Latino	Myth! A girl disappears in the town of San Gregorio and a group of teenagers decide to investigate what happened.										
35	s34	TV Show	La Culpa es de Cort/és	October 14, 2017			2017	4 Seasons		Thiller	Latino	Real! Porque sabemos que no es fcil encontrar el amor verdadero en este mundo digital, MTV Catfish llega a Mxico para salvarte de las mentiras. Todos										
36	s35	Movie	Out of Death				October 14, 2021	NOT RATED	95 min	Thriller	When Shannon jaime Kinit sets out on a hike to spread her deceased father Adas ashes, she witnesses a corrupt sheriff. Adas deserv. Billie (Lala Kent) bri											

2.1.2 Secondary Data Source: The Movie Database (TMDb) API

The static nature of CSV files means they cannot provide real-time information. To enhance the dashboard with dynamic, up-to-the-minute data, we integrated **The Movie Database (TMDb) API**.

This real-time enrichment layer is responsible for three key features on the dashboard's Home Page:

- Title Search:** Allows users to search for any movie or TV show. The API returns detailed metadata, including posters, ratings, and overviews.
- Movie Reviews:** Fetches current user reviews for a searched title, providing qualitative insights.
- Trending Content:** Populates a live-updating feed of movies that are currently trending, adding a dynamic "pulse" to the home page.

This hybrid approach of using a static core (for deep analysis) and a dynamic API (for real-time enrichment) provides the project with both historical depth and current relevancy.

2.2 DATA DICTIONARY

Understanding the features (columns) available in the primary datasets is essential for establishing the scope of our analysis. The four datasets share a common, though not identical, schema. The following table defines the key features used in this project.

Table 2.1: Data Dictionary of Core Features

Feature	Data Type (Raw)	Description	Example (Raw)
show_id	object (string)	A unique identifier for each title, specific to the platform.	s5001
type	object (string)	The category of the content.	Movie / TV Show
title	object (string)	The official name of the movie or TV show.	Stranger Things
director	object (string)	The name(s) of the director(s). Can be multi-valued, comma-separated.	Quentin Tarantino
cast	object (string)	A list of primary actors. Multi-valued and comma-separated.	Tom Hanks, Tim Allen
country	object (string)	The country or countries of origin. Multi-valued, comma-separated.	United States, India
date_added	object (string)	The date the content was added to the streaming service.	September 2, 2019
release_year	float64 / object	The original release year of the content.	2019
rating	object (string)	The maturity rating of the content (e.g., TV-MA, PG-13, R).	TV-MA
duration	object (string)	The length of the content. Highly inconsistent format.	121 min / 2 Seasons
listed_in	object (string)	The genres the content is categorized under. Multi-valued, comma-separated.	Dramas, International Movies
description	object (string)	A brief synopsis of the title.	A young boy vanishes...

The most significant analytical challenges, identified in Chapter 1, are immediately apparent from this dictionary: the duration, date_added, director, country, and listed_in columns are all stored as non-standard strings and require extensive preparation before they can be used for quantitative analysis.

CHAPTER 3

DATA PREPARATION

The raw data sourced for this project, while comprehensive, was in a state unsuitable for direct visualization or analysis. Each of the four datasets contained a myriad of inconsistencies, missing values, and problematic data formats. Therefore, a robust and multi-stage data preparation pipeline was designed and implemented as the foundational engineering task of this project.

This chapter details the technical steps taken to clean, transform, and enrich the raw data, converting it into the high-quality, standardized format required to power the DataFlix dashboard. All operations were performed in Python using the **Pandas** library.

3.1 DATA LOADING

The data preparation pipeline begins with the ingestion of the four primary datasets. This process is managed by the `load_data()` and `load_all_data()` functions within the `utils/data_loader.py` utility module.

1. Single-Platform Loading (`load_data(platform)`):

- This function is called when a user navigates to a specific dashboard (e.g., "Netflix").
- It uses `pandas.read_csv()` to load the corresponding file (e.g., `data/netflix_titles.csv`) into a `DataFrame`.
- This function is wrapped in a `@st.cache_data` decorator, a Streamlit feature that caches the data in memory. This ensures that the (potentially large) CSV file is read from disk only once, dramatically improving dashboard performance when switching between tabs or applying filters.

2. Global Data Loading (`load_all_data()`):

- This function is called by the Home Page to power the global KPIs and comparative analytics.
- It iteratively loads all four CSV files.
- Crucially, during this process, it performs an initial **feature engineering** step by adding a new column, `platform`, to each `DataFrame` (e.g., "Netflix", "Prime Video").
- Finally, it concatenates the four `DataFrames` into a single, master `DataFrame` using `pd.concat()`. This "long-format" data is what enables all cross-platform analysis, such as the Market Share pie chart.

3.2 DATA CLEANING AND TRANSFORMATION

This stage forms the core of the data preparation pipeline, addressing the significant inconsistencies identified in Chapter 2.

3.2.1 Parsing and Standardizing the duration Column

The duration column was the most complex, as it contained two different metrics (minutes for movies, seasons for TV shows) in a single string format.

- **Step 1: Separate Movies and TV Shows:** Two new DataFrames were created: df_movies and df_tv based on the type column.
- **Step 2: Parse Movie Durations:** For df_movies, the string " min" was removed using df_movies['duration'].str.replace(' min', '').
- **Step 3: Parse TV Show Durations:** For df_tv, a regular expression was used via df_tv['duration'].str.extract('(\d+)') to isolate only the numerical digit from strings like "1 Season" or "10 Seasons".
- **Step 4: Convert to Numeric:** Both new duration columns were converted to integers using pd.to_numeric(..., errors='coerce'). The errors='coerce' argument is critical: it automatically converts any unparseable values (such as "Unknown" or other text) into NaN (Not a Number), preventing the entire operation from crashing.

This transformation created two new, clean numerical features (duration_int for movies and duration_int for TV shows) that were essential for plotting histograms of content length.

3.2.2 Transforming Multi-Value Categorical Columns

Columns like listed_in (genre), country, and director were stored as comma-separated strings (e.g., "Dramas, International Movies, Thrillers"). Performing a simple value_counts() on this data would treat this string as a single, unique category, which is incorrect.

To solve this, the **DataFrame.explode()** method was used.

- **Step 1:** The string in the column was split into a Python list using df['listed_in'].str.split(',') .
- **Step 2:** The explode() function was called on this new list column.

This operation transforms each item in the list into its own row, duplicating the other title information. For example, a single title with three genres would be transformed from **1 row to 3**

rows. This "long-format" data is the correct structure for accurately counting and analyzing frequencies, powering all the "Top 10 Genre" and "Top Country" bar charts.

3.2.3 Standardizing Date Columns

The date_added column was stored as an object (string) and had multiple formats (e.g., "September 2, 2019", "2019-09-02").

- **Solution:** The pd.to_datetime(df['date_added'], errors='coerce') function was applied.
- **Benefit:** This powerful Pandas function automatically parses numerous common date formats into a standardized datetime object. The errors='coerce' flag gracefully handles any missing or unparseable dates by converting them to NaT (Not a Time), which are then treated as null values.

3.3 HANDLING MISSING VALUES

A preliminary analysis using df.isnull().sum() revealed a significant number of missing values across all four datasets. A naive strategy, such as dropping all rows with any missing data (df.dropna()), would have resulted in an unacceptable loss of data (in some cases over 50%). Therefore, a more nuanced, column-by-column imputation strategy was employed.

Table 3.1: Missing Value Imputation Strategy

Feature	Missing Values	Strategy Adopted	Rationale
Director	~30%	Impute with "Unknown"	Dropping 30% of data is not viable. Imputing with "Unknown" allows us to group and analyze the volume of content with uncredited directors, which is an insight itself.
Country	~10-15%	Impute with "Unknown"	Similar to director, this preserves the data rows and allows for analysis of content with unspecified origins.
Cast	~15%	Impute with "Unknown"	While not heavily used in our final dashboards, this strategy was applied for consistency and to preserve data for potential future analysis.
Rating	~1-2%	Impute with Mode	Given the small percentage of missing ratings, the most frequent rating (e.g., "TV-MA") was used as a reasonable substitute to fill the gaps.
date_added	~5-10%	No Imputation (Drop NA on-the-fly)	It is impossible to accurately guess when a title was added. These values (now NaT) are kept as-is. Visualizations that <i>require</i> this field (e.g., "Content Added Over Time") automatically ignore these null values during plotting.

This multi-step imputation process ensured that the dataset was clean, complete, and robust, maximizing the amount of data available for analysis while correctly handling unavoidable gaps.

3.4 FEATURE ENGINEERING

The final stage of the preparation pipeline was **Feature Engineering**, where new, high-value features were created from the cleaned data. These new features are what power many of the most insightful visualizations on the dashboards.

1. **Temporal Features:** After converting date_added to a datetime object, the following features were extracted using Pandas' .dt accessor:
 - **year_added:** (df['date_added'].dt.year) - Used in all "Content Added Per Year" line and area charts.
 - **month_added:** (df['date_added'].dt.month_name()) - Used in the Disney+ "Content Added by Month" polar (radar) chart to identify seasonality.
 - **quarter_added:** (df['date_added'].dt.to_period('Q')) - Used in the Netflix temporal chart to show quarterly additions.
2. **Content Lag (Business Metric):** A key strategic metric was engineered by combining two existing features:
 - **lag_years** = df['year_added'] - df['release_year']
 - This new feature is a powerful indicator of a platform's content strategy. A low average lag (e.g., 0-1 years) indicates a focus on "Originals" or new acquisitions. A high average lag (as seen on Hulu) indicates a focus on licensing a deep "back-catalog" of older content.
3. **Platform Identifier:** As mentioned in section 3.1, the platform column (e.g., "Netflix", "Prime Video") added during the global data load is the most critical engineered feature. It acts as the primary key for all comparative analytics, including the Home Page Market Share chart and the Head-to-Head comparison tool.

CHAPTER 4

EXPLORATORY DATA ANALYSIS (EDA)

Exploratory Data Analysis (EDA) is the critical process of investigating datasets to uncover patterns, identify anomalies, test hypotheses, and check assumptions using summary statistics and graphical representations. In the context of the DataFlix project, the EDA was not merely a preliminary step but formed the very foundation of the dashboard's design.

The visualizations created during this phase were not disposable; they were refined and promoted to become the core interactive components of the final application. This chapter details the key analytical findings from the EDA, structured in the same way as the dashboard itself: beginning with a global overview and then performing a deep dive into each of the four platforms. All visualizations were generated using Python's **Plotly** library to ensure high interactivity and aesthetic quality.

4.1 GLOBAL MARKET ANALYSIS (THE HOME PAGE)

The analysis began at the macro level, using the combined dataset (as described in section 3.1) to understand the overall market landscape.

4.1.1 Global Key Performance Indicators (KPIs)

The first analysis was a high-level summary of the entire dataset. These figures provide an immediate, quantitative understanding of the scale of the streaming market being analyzed.

- **Total Titles Analyzed:** 22,998
- **Platforms Monitored:** 4
- **Top Genre Across Platforms:** Drama

This simple analysis confirms the vast scale of the content libraries and immediately identifies "Drama" as the most common content category, suggesting it is a foundational, "must-have" genre for any competitive platform.

4.1.2 Library Size by Platform

To understand the market share in terms of content volume, a donut chart was generated.

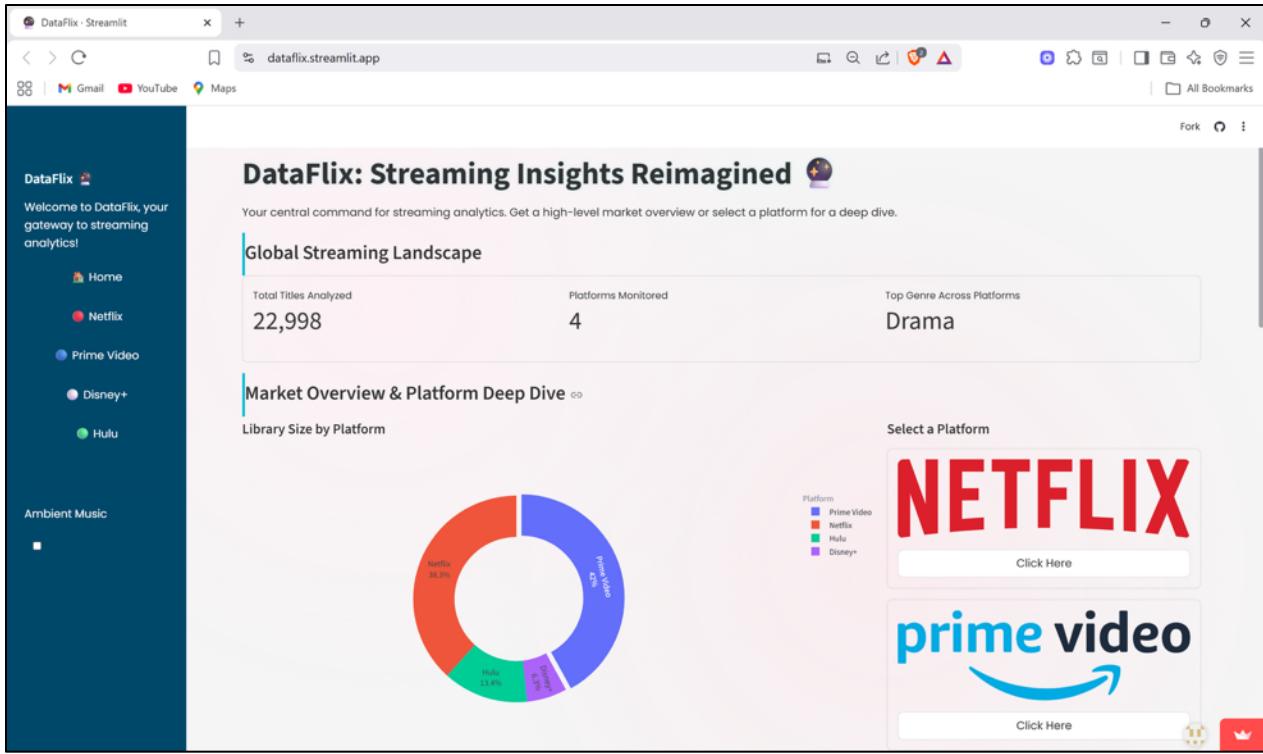


Figure 4.1 Home Page of DataFlix Dashboard

Description:

The above image displays the proportional distribution of total titles (both movies and TV shows) among the four platforms. Each segment represents a platform, with its size corresponding to its percentage of the total content library in our dataset.

Insights:

- **Dominance of Quantity:** A striking insight is the sheer scale of the Prime Video and Netflix libraries, which together account for over 75% of the total titles.
- **Divergent Strategies:** This visualization clearly highlights two different business models. Netflix and Prime Video appear to follow a "quantity-first" or "warehouse" model, aiming to have something for everyone.
- **Curated Libraries:** Disney+ and Hulu, by contrast, have significantly smaller, more curated libraries. Their strategy is visibly one of vertical focus—Disney+ on family entertainment and Hulu on premium television—rather than sheer volume.

4.1.3 Platform Head-to-Head Comparison

The final component of the global analysis was an interactive tool rather than a static chart. This tool allows for dynamic, user-driven EDA.

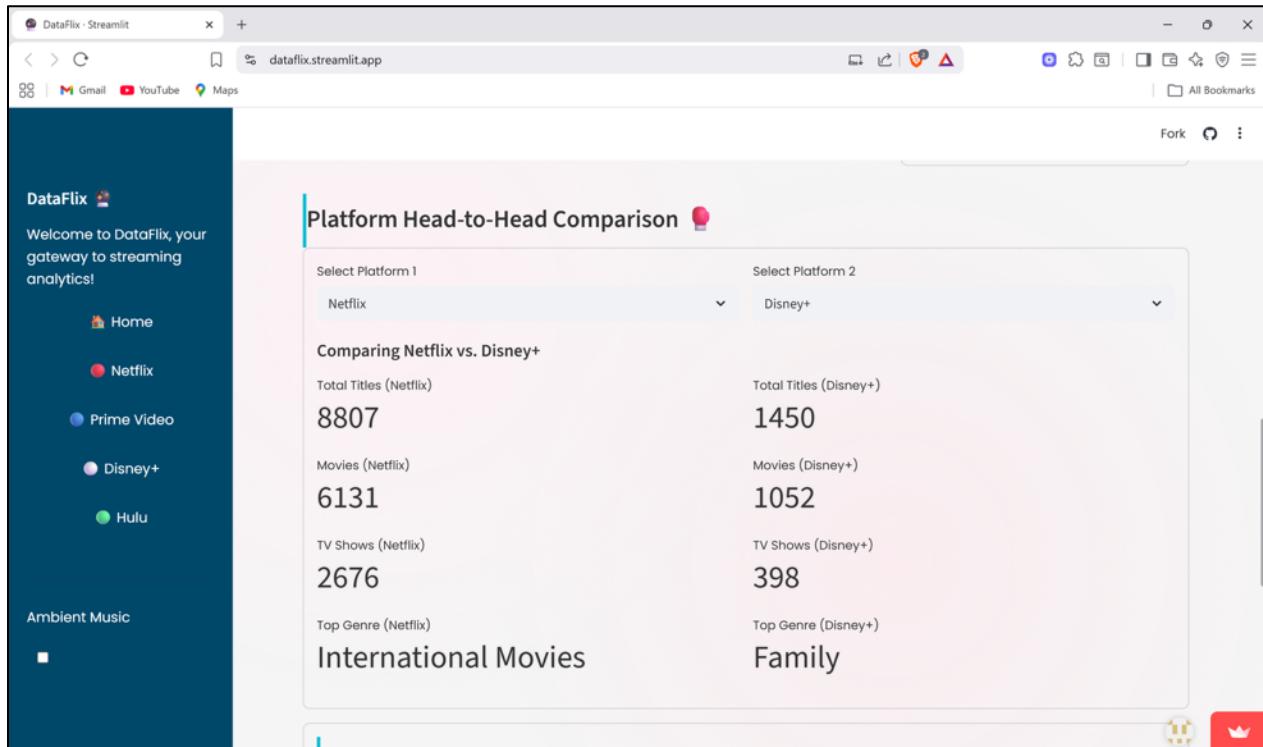


Figure 4.2 Head-to-Head Comparison

Description:

The above image shows the UI of the "Head-to-Head Comparison" tool. A user can select any two platforms from the dropdown menus to generate a side-by-side comparison of their core KPIs.

Insights:

This tool is a powerful analytical feature. For example, a user comparing Netflix and Disney+ would immediately see:

- **Netflix (Total Titles):** ~8,800
- **Disney+ (Total Titles):** ~1,700
- **Netflix (Top Genre):** International Movies
- **Disney+ (Top Genre):** Family

This instant comparison provides more context than the global chart alone, starkly contrasting Netflix's broad, international strategy with Disney+'s highly focused, niche strategy.

4.2 NETFLIX DASHBOARD ANALYSIS

Following the global overview, a deep-dive analysis was performed on each platform. The Netflix dashboard is organized into four thematic tabs.

4.2.1 Tab 1: Content Library

This tab analyzes the fundamental composition of the Netflix library.

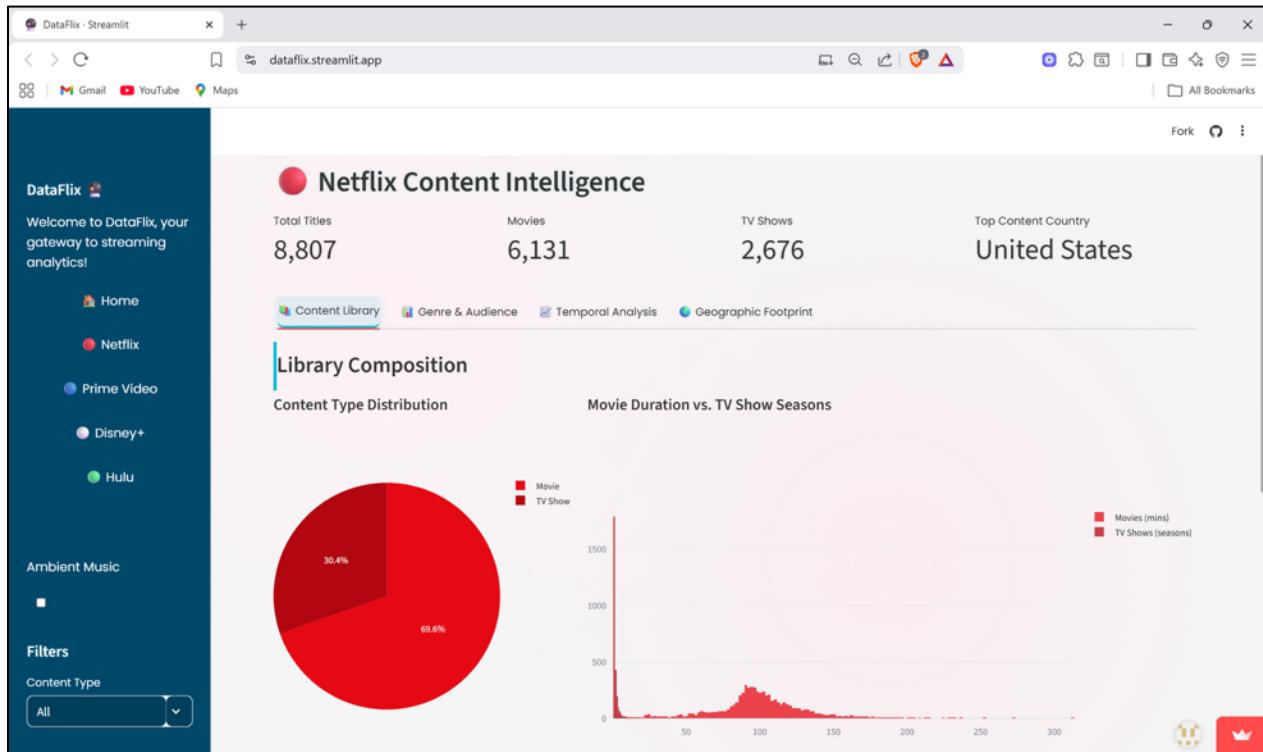


Figure 4.3: Netflix Content Type Distribution (Pie Chart)

Description: A pie chart showing the split between Movies and TV Shows.

Insight: The library is split approximately 69% Movies and 31% TV Shows. While movies still dominate, the significant 31% share for TV shows—which are generally more expensive and time-consuming to produce—highlights Netflix's major strategic pivot to becoming a dominant television-style broadcaster.

Description: An overlaid histogram showing the distribution of movie runtimes (in minutes) and TV show "runtimes" (in number of seasons).

Insight:

- **Movies:** The distribution is a classic normal curve centered around 90-120 minutes, confirming a focus on feature-length films.
- **TV Shows:** The distribution is heavily skewed to the left, with "1 Season" being by far the most common. This is a critical business insight, highlighting Netflix's "ruthless" cancellation policy and its strategy of commissioning many shows but only renewing the most successful.

4.2.2 Tab 2: Genre & Audience

This tab focuses on *what* is being offered and *who* it is for.

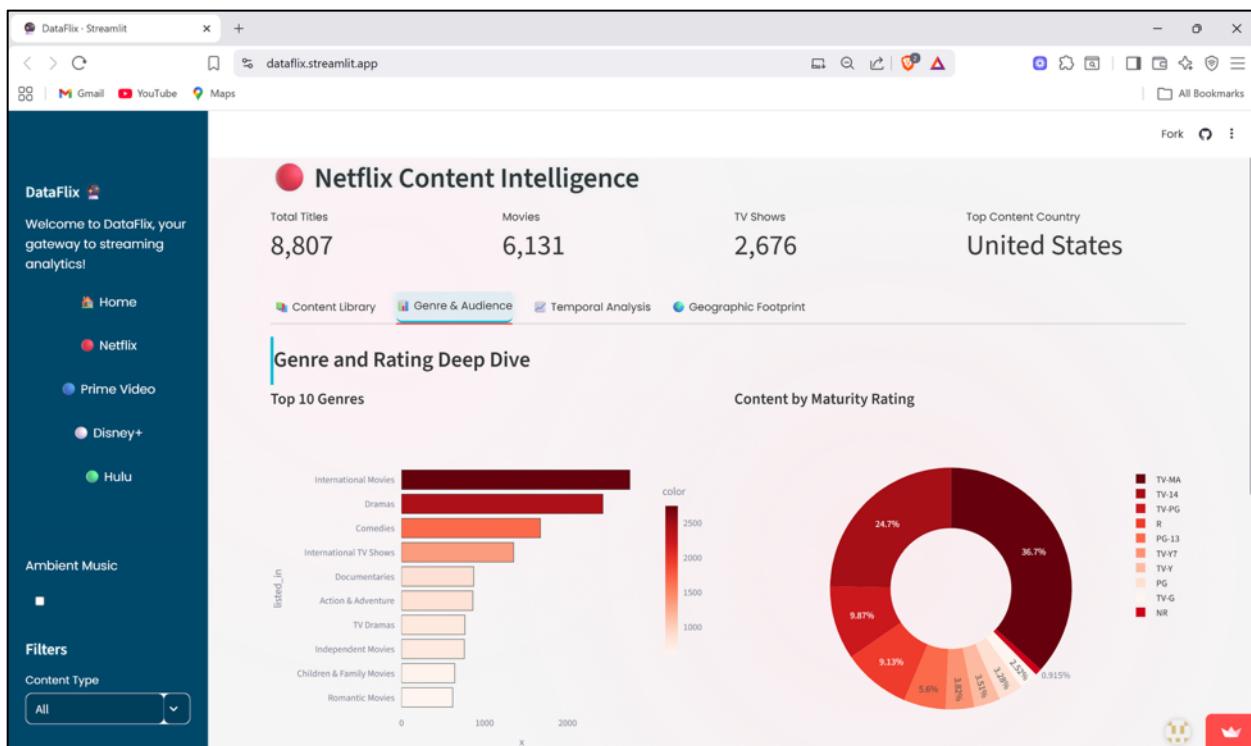


Figure 4.4 : Netflix Top 10 Genres (Bar Chart)

Description: A horizontal bar chart, derived from the "exploded" listed in column, showing the 10 most frequent genre tags.

Insight: "International Movies" is the #1 genre, followed by "Dramas." This finding is paramount: it proves that Netflix's primary growth and content strategy is global. "Dramas" and "Comedies" are core, but the "International" tag shows a clear focus on acquiring and producing content from markets outside the US.

Description: A donut chart visualizing the distribution of content by its maturity rating.

Insight: The two largest segments are "TV-MA" (Mature Audiences) and "R" (Restricted). This clearly indicates that Netflix's core audience is adults, and it is not competing directly with a platform like Disney+ for the family market.

4.2.3 Tab 3: Temporal Analysis

This tab analyzes the "freshness" and growth of the library over time.

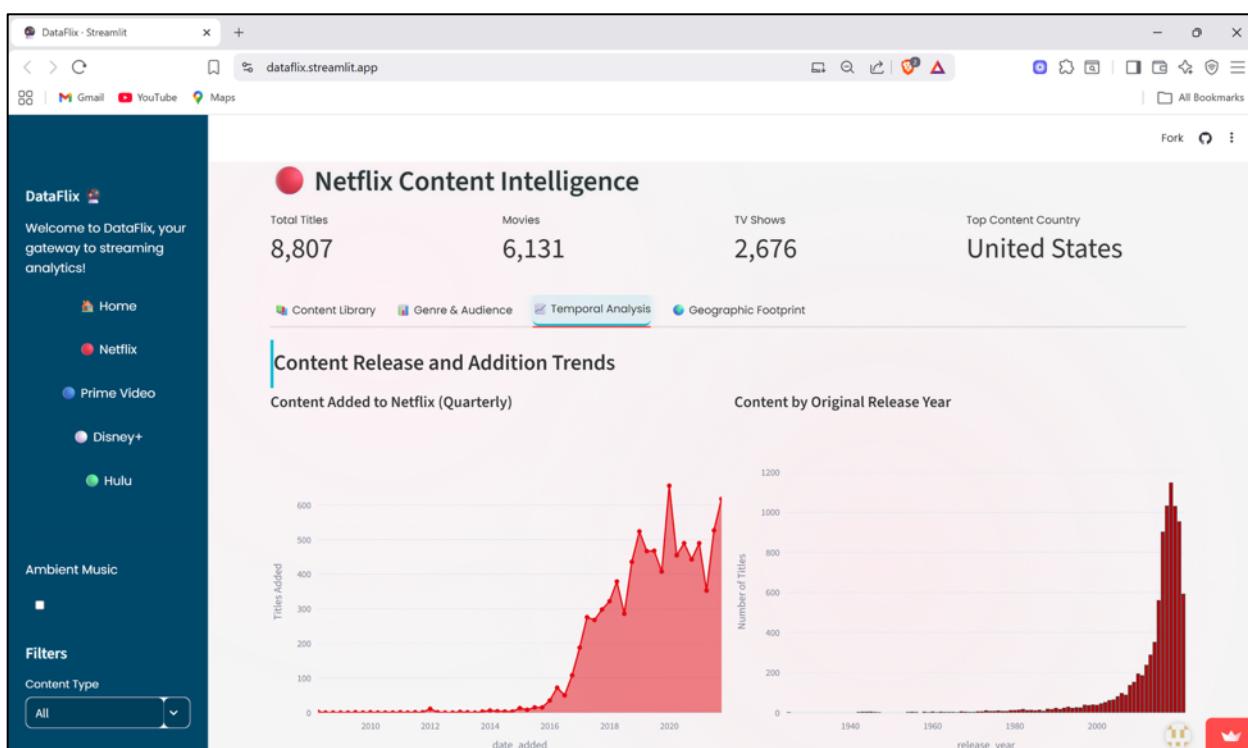


Figure 4.5 : Netflix Content Added Over Time (Area Chart)

Description: A time-series area chart plotting the number of titles added to the platform each quarter.

Insight: The chart shows a dramatic, near-exponential growth in content additions, peaking in 2019-2020. This visualizes Netflix's massive content spending spree in the years leading up to the launch of its main competitors (like Disney+).

4.2.4 Tab 4: Geographic Footprint

This tab analyzes the global production and sourcing of content.

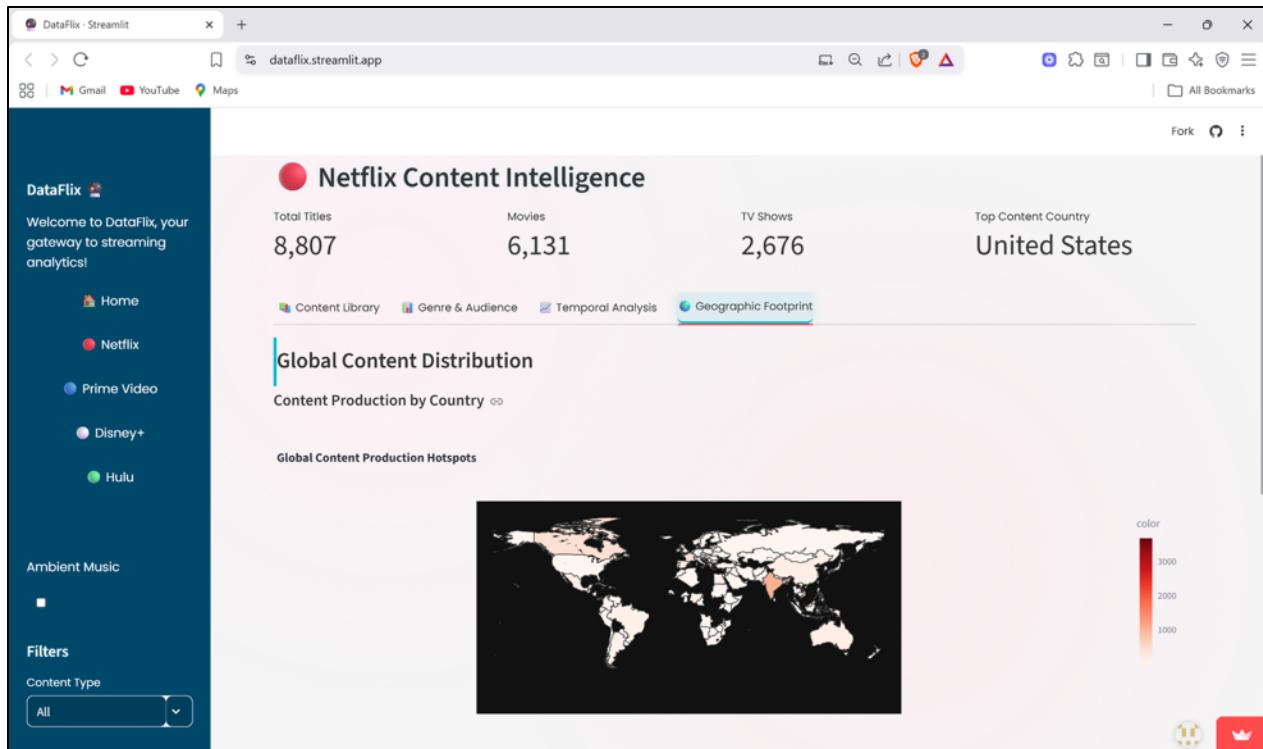


Figure 4.6: Netflix Global Content Production (Choropleth Map)

Description: A choropleth world map where countries are color-coded based on the number of titles they have produced.

Insight: While the United States is the largest producer, this map provides a stunning visualization of Netflix's global strategy. Key secondary markets like India, the United Kingdom, Japan, South Korea, and Spain are clearly visible as major production hubs. This reinforces the insight from the genre chart, showing a deep, tangible investment in international content.

4.3 PRIME VIDEO DASHBOARD ANALYSIS

A similar deep-dive analysis was performed on the Prime Video dataset, revealing a very different strategy.

4.3.1 Tab 1: Content Library

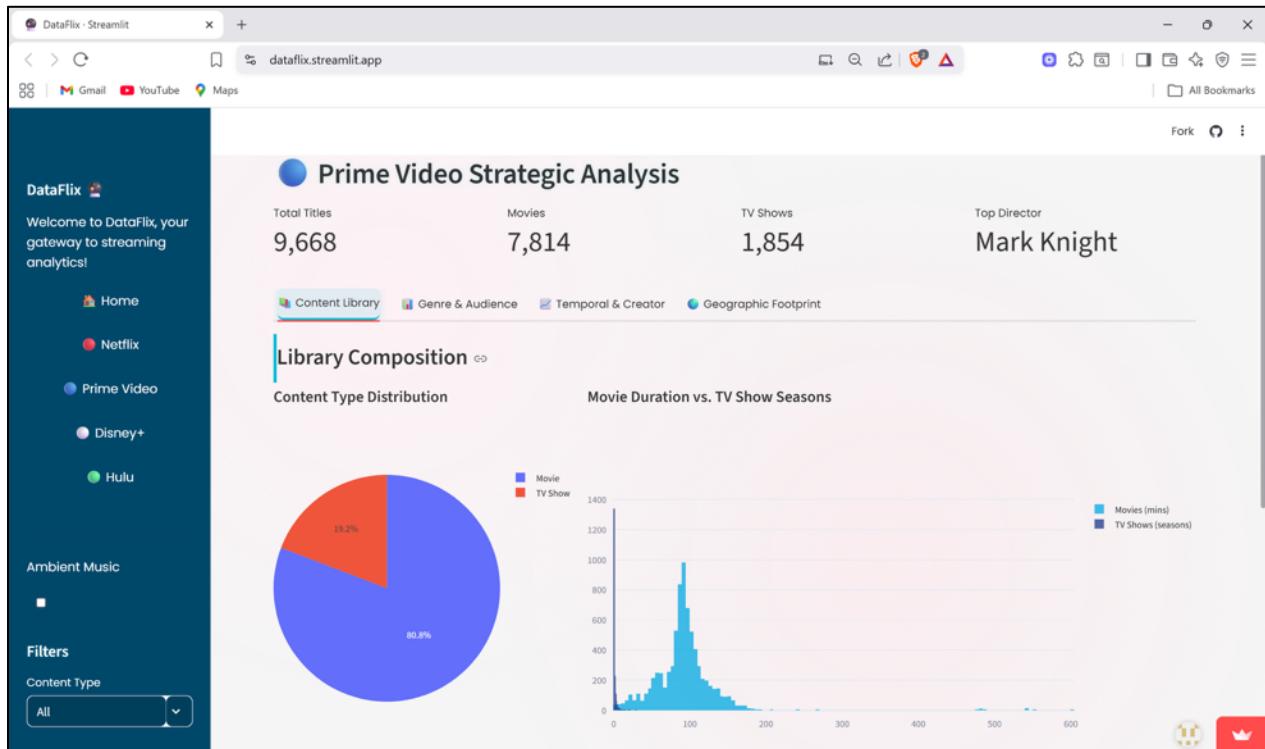


Figure 4.7: Prime Video Content Type Distribution (Pie Chart)

Description: A pie chart showing the split between Movies and TV Shows.

Insight: The library is overwhelmingly dominated by Movies (over 80%). This indicates a fundamentally different strategy from Netflix. Prime Video acts more as a massive film archive, focusing on quantity and a vast back-catalog of movies rather than a balanced portfolio of original TV shows.

4.3.2 Tab 2: Genre & Audience

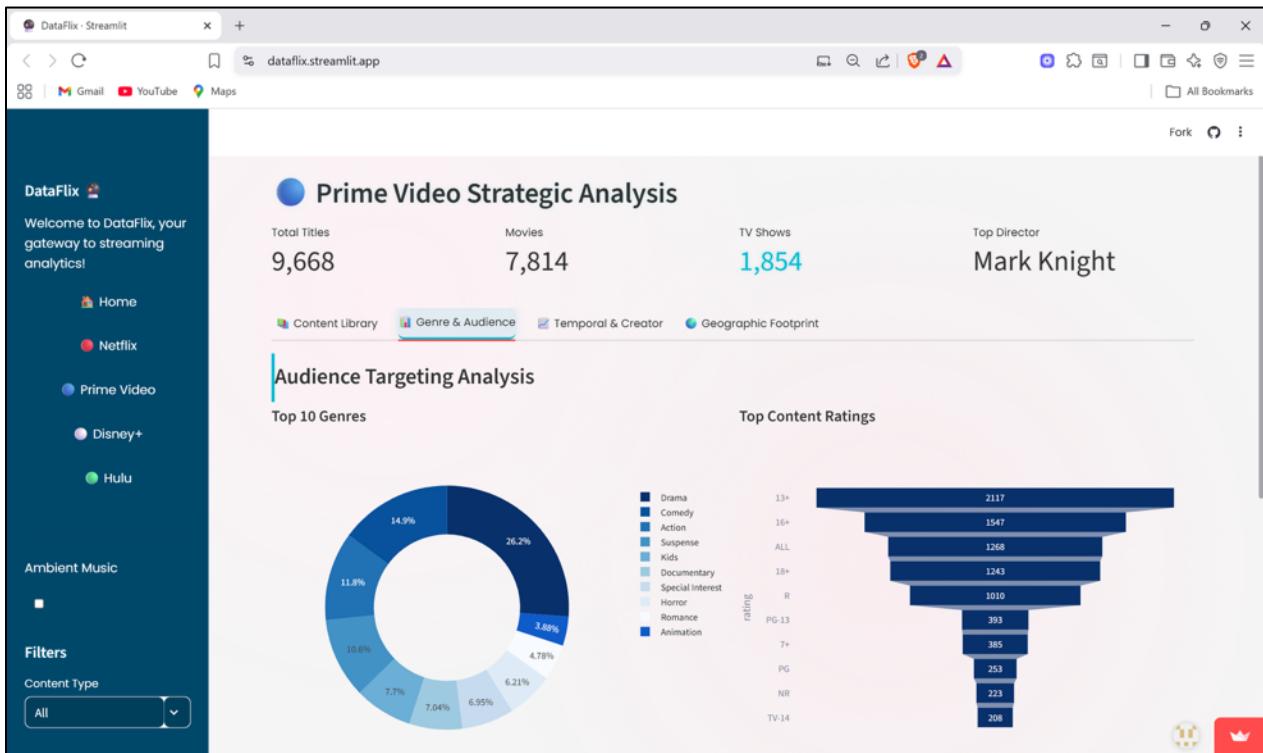


Figure 4.8: Prime Video Top 10 Genres (Donut Chart)

Description: A donut chart showing the top 10 most frequent genres.

Insight: "Dramas" and "Comedy" are dominant, which is standard. However, the presence of niche genres like "Suspense" and "Arthouse" in the top 10 suggests a "long-tail" strategy, attempting to capture smaller audiences with specific interests.

4.3.3 Tab 3: Temporal & Creator Analysis

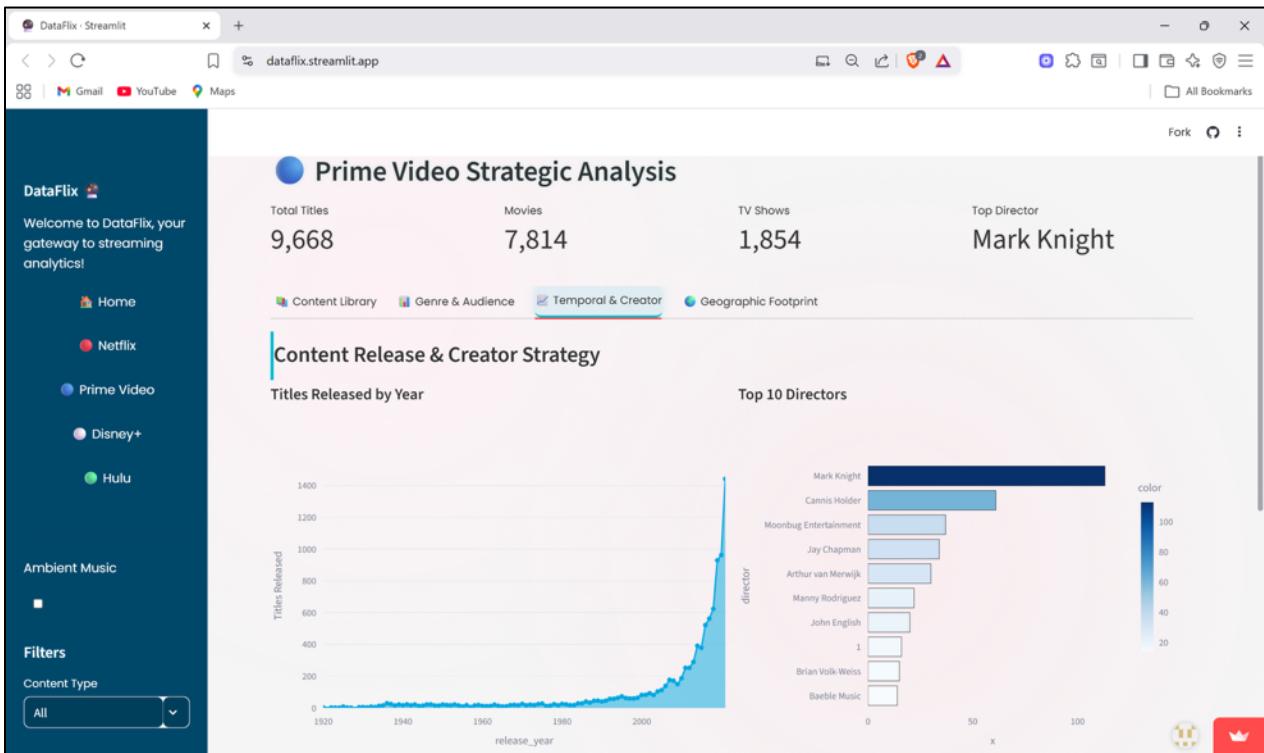


Figure 4.9: Prime Video Content by Release Year (Area Chart)

Description: A time-series area chart plotting the number of titles in the library by their original release year.

Insight: This is one of the most critical findings for Prime Video. Unlike Netflix's chart, which was skewed towards recent additions, this chart shows a massive volume of content from the 1950s, 60s, and 70s. This proves Prime Video's "warehouse" model, which relies heavily on acquiring a deep back-catalog of older films, likely at a lower licensing cost.

Description: A horizontal bar chart of the most frequent directors on the platform.

Insight: The list is often topped by directors like "Mark Knight" or "Unknown" (from our data cleaning). This indicates that a significant portion of the library consists of lower-budget, independent, or even public-domain films, reinforcing the "quantity over quality" acquisition strategy.

4.4 DISNEY+ DASHBOARD ANALYSIS

The Disney+ analysis revealed a highly focused, curated, and brand-driven strategy.

4.4.1 Tab 1: Content Library

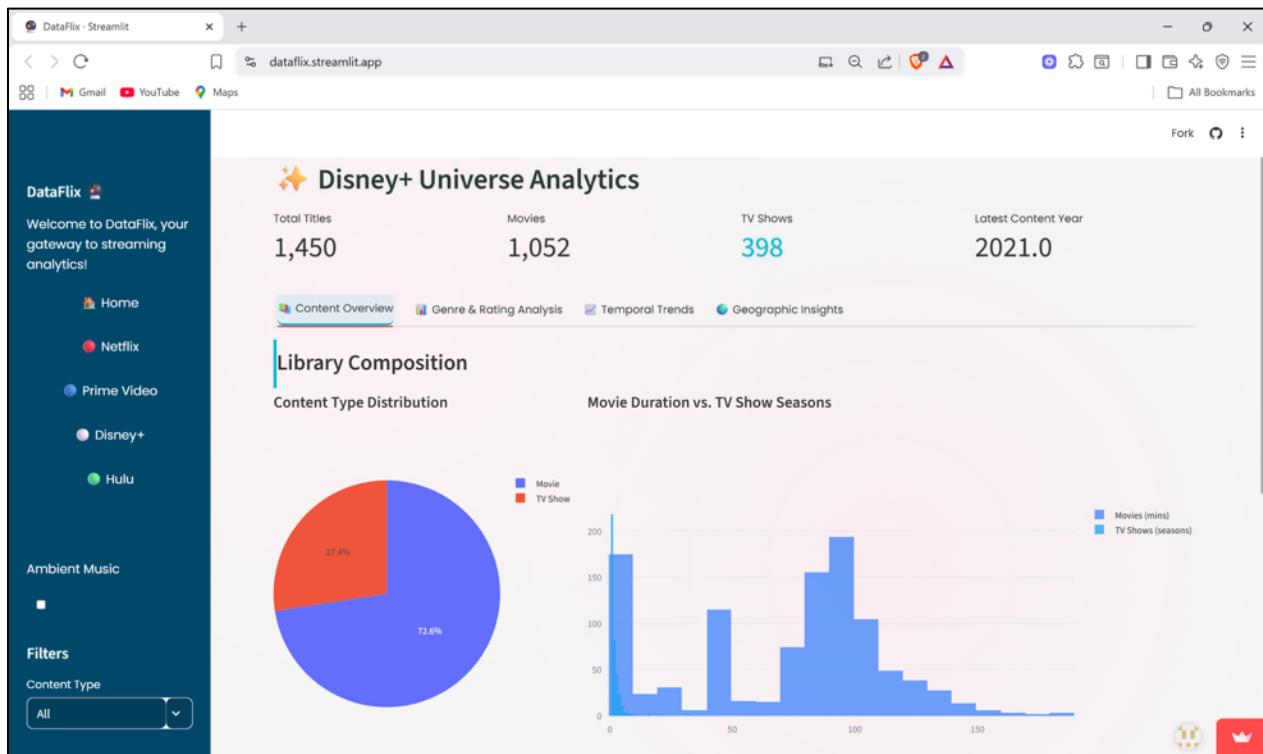


Figure 4.10: Disney+ Content Type Distribution (Pie Chart)

Description: A pie chart showing the split between Movies and TV Shows.

Insight: The library is more balanced, with Movies at ~70% and TV Shows at ~30%. This is logical, as Disney's core assets include both a massive film library (Disney, Pixar, Marvel, Star Wars) and a strong television presence (Disney Channel, National Geographic).

4.4.2 Tab 2: Genre & Audience

Description: A vertical bar chart of the top genres on Disney+.

Insight: The chart is completely dominated by "Family," "Animation," "Action-Adventure," and "Comedy." This is a powerful visualization of an incredibly focused, brand-first strategy. There is almost no content outside of these core, family-friendly verticals.

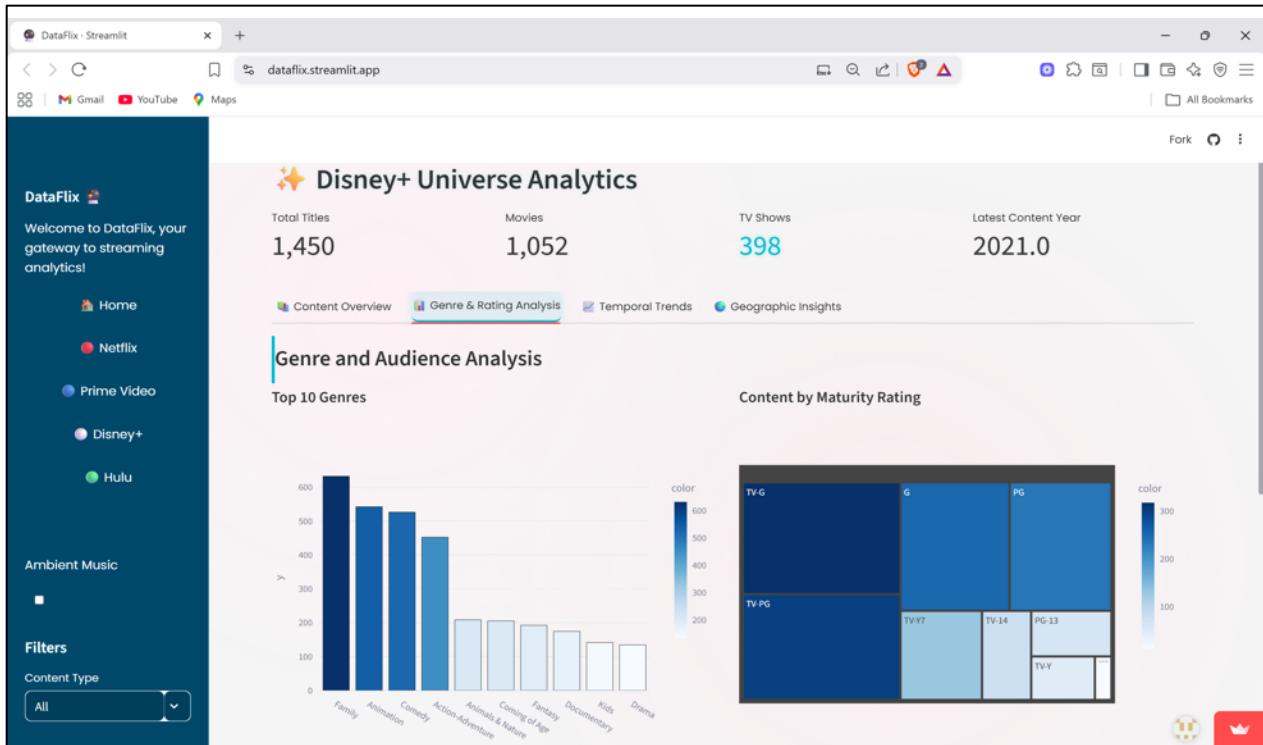


Figure 4.11 : Disney+ Top Genres (Bar Chart)

Description: A treemap visualizing the distribution of maturity ratings.

Insight: This chart perfectly complements the genre analysis. The largest blocks are "TV-G," "PG," and "TV-Y7." There is almost no "TV-MA" or "R" rated content. This confirms, quantitatively, that Disney+ is exclusively targeting the family and children's market.

4.4.3 Tab 3: Temporal Trends

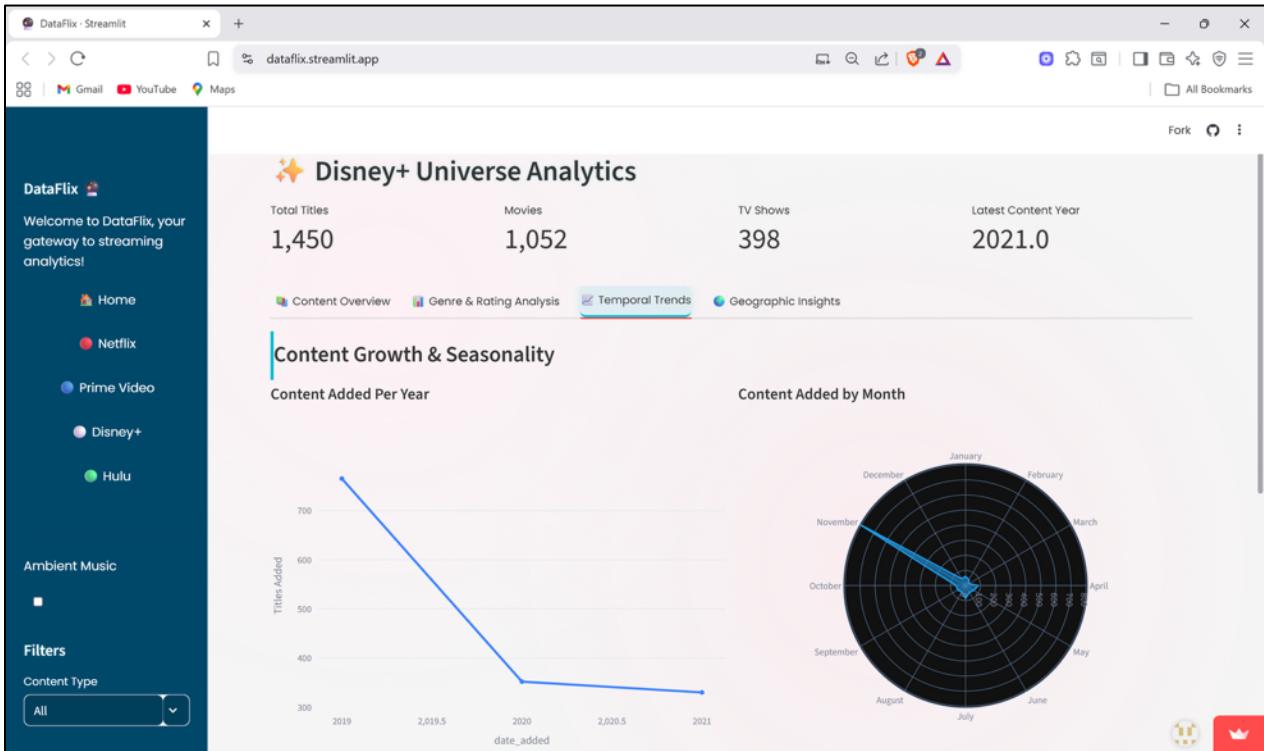


Figure 4.12 : Disney+ Content Added by Month (Polar Chart)

Description: A polar (radar) chart showing the number of titles added in each month of the year.

Insight: This chart reveals a clear seasonal strategy. There are significant spikes in content additions in October, November, and December (Q4), timed perfectly to coincide with the holiday season, family gatherings, and winter school breaks.

4.5 HULU DASHBOARD ANALYSIS

Hulu's analysis showcased its unique position as the primary destination for current television and mature content.

4.5.1 Tab 1: Library Overview

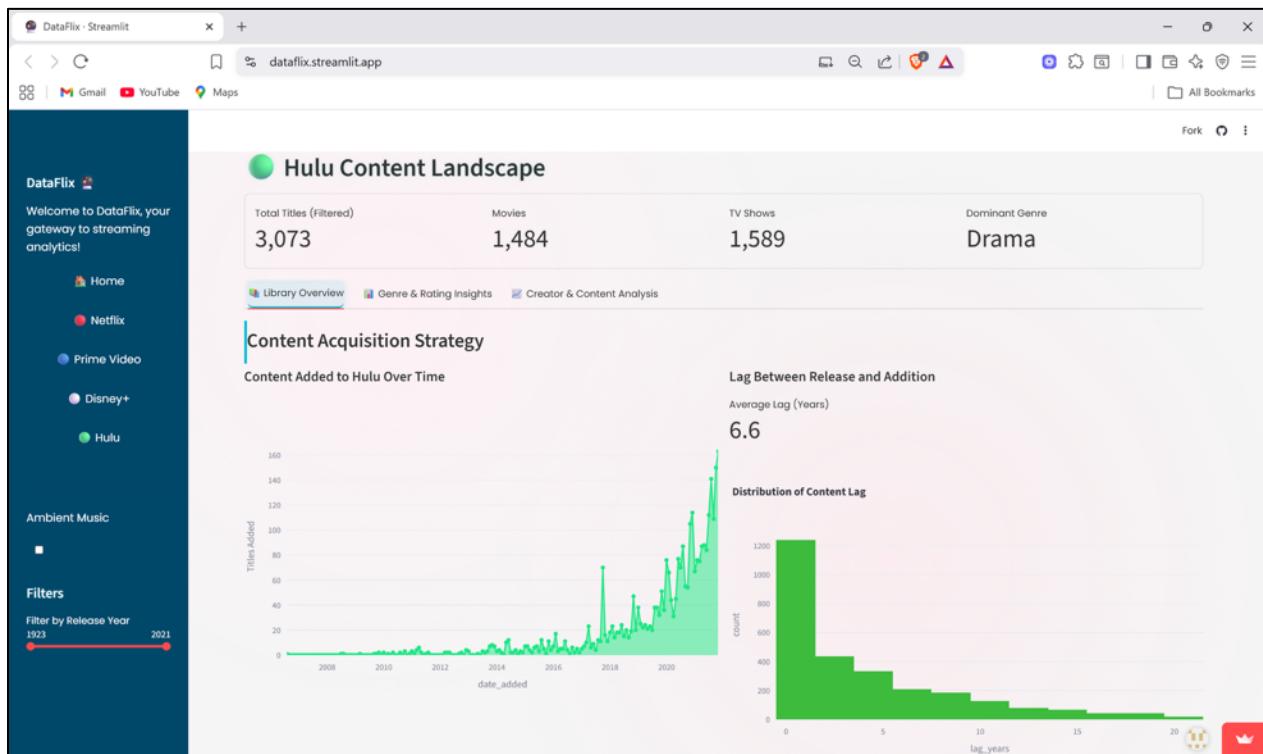


Figure 4.13 : Hulu Content Lag Histogram

Description: A histogram of the "Content Lag" (time in years between a title's original release and its addition to Hulu), a key feature engineered in Chapter 3.

Insight: This chart is the key to understanding Hulu. The largest spike is at 0 years, which represents its core value proposition: "next-day" streaming of current-season TV shows. However, the chart also shows a long tail of content added 5, 10, or 15+ years after its release, visualizing its secondary strategy as a deep library for older, licensed TV seasons.

4.5.2 Tab 2: Genre & Audience

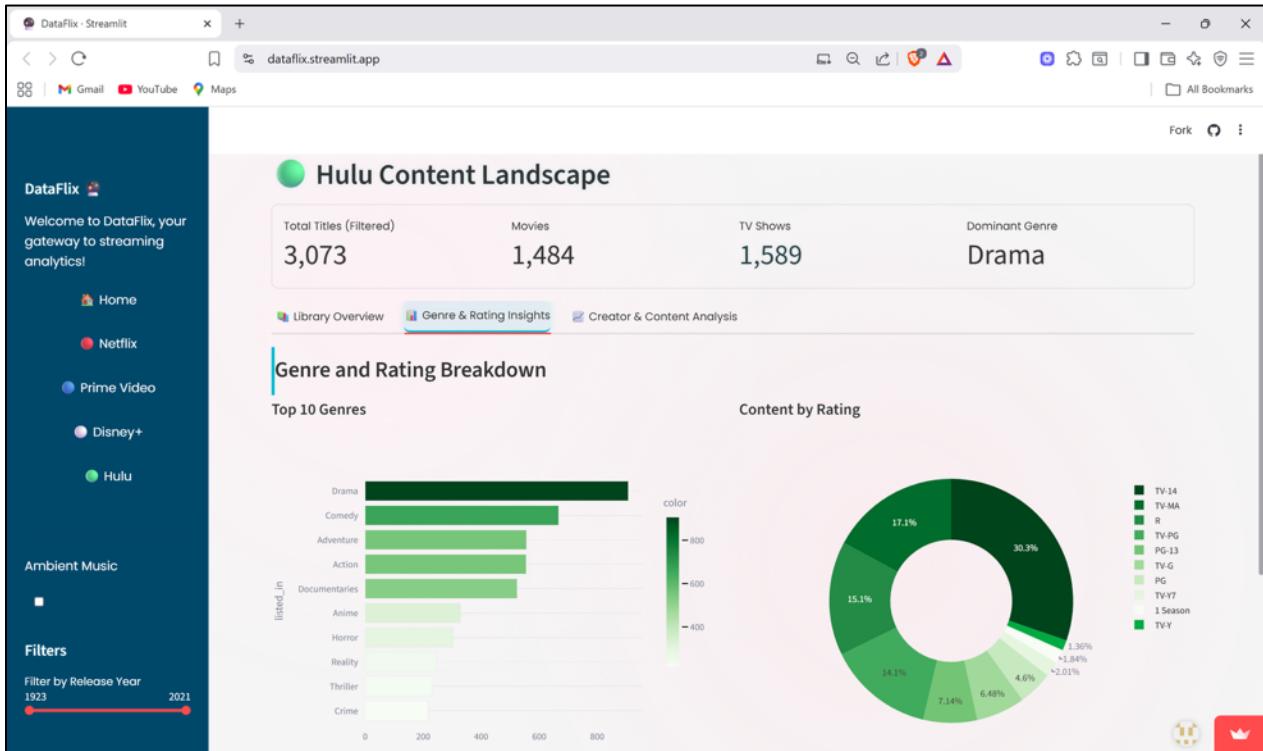


Figure 4.14 : Hulu Top 10 Genres (Bar Chart)

Description: A horizontal bar chart of the top 10 genres.

Insight: The top genres are "Dramas," "Comedy," "Crime," and "International." This is a stark contrast to its sister platform, Disney+, and shows Hulu's clear focus on adult-oriented, premium television.

Description: A pie chart showing the distribution of maturity ratings.

Insight: The largest single slice of the pie is "TV-MA" (Mature Audiences). This is the "anti-Disney+" chart. It proves Hulu's strategic role in the Disney-owned ecosystem is to be the brand that captures the adult market, allowing Disney+ to remain exclusively family-friendly.

4.5.3 Tab 3: Creator & Content Analysis

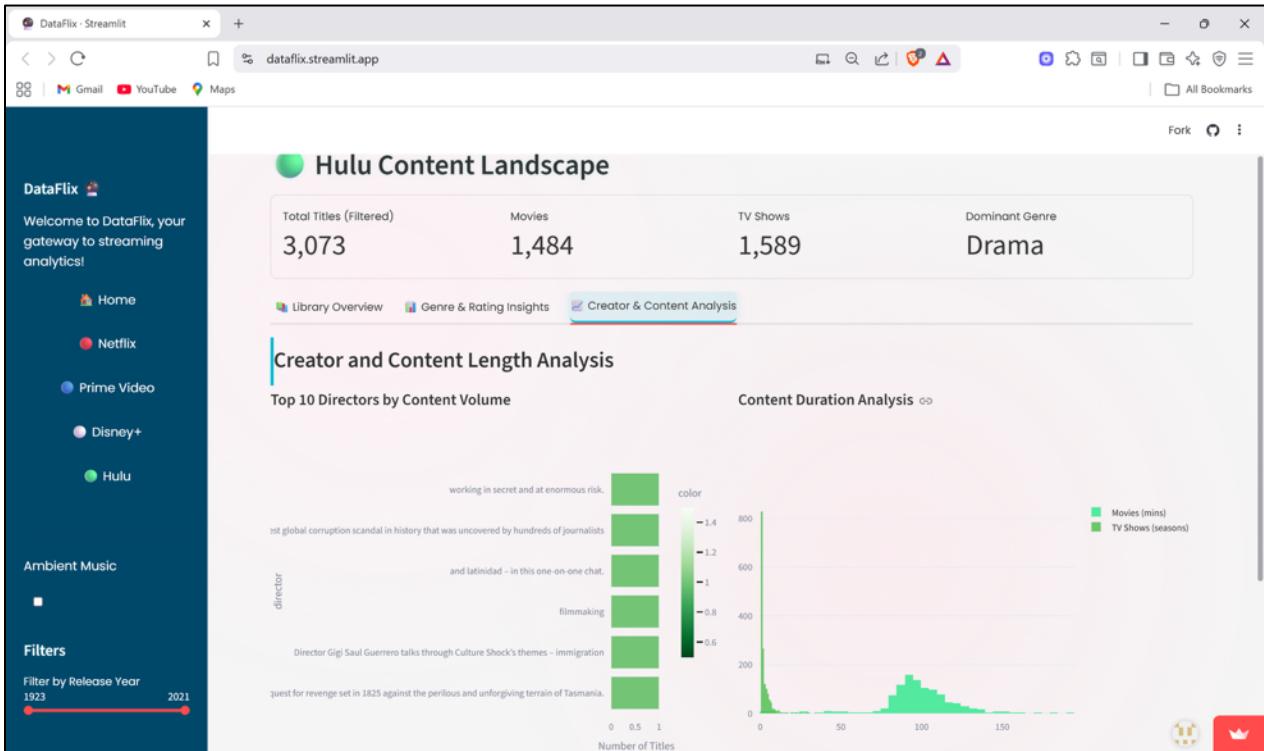


Figure 4.15 : Hulu Top 10 Directors (Bar Chart)

Description: A bar chart of the top directors by volume.

Insight: Unlike Prime Video's list of unknown names, Hulu's list often includes prolific television directors and producers. This reinforces that its library's strength is in its comprehensive collection of premium, professionally produced television series.

CHAPTER 5

RESULTS AND DISCUSSION

The culmination of this project is not a single numerical output, such as an accuracy score, but rather a fully functional, multi-page Business Intelligence (BI) dashboard. The "results" are the analytical features and data products generated by the system, and the "discussion" is the interpretation of the strategic insights those features reveal.

This chapter discusses the key results by breaking down the functionality of the final "DataFlix" application, connecting the technical implementation (from Chapter 3) and the analytical findings (from Chapter 4) to the final user-facing product.

5.1 HOME PAGE: THE GLOBAL COMMAND CENTER

The Home Page serves as the central hub for high-level market analysis. The results presented here are designed to give users an immediate, quantitative overview of the streaming landscape.

- **Global KPIs:** The dashboard's first result is a set of three key metrics: Total Titles Analyzed (22,998), Platforms Monitored (4), and the Top Genre Across Platforms (Drama). This result immediately frames the scale of the market and identifies the single most competitive genre vertical.
- **Market Share Visualization:** The "Library Size by Platform" donut chart (Fig 4.1) is a key result of the global data aggregation. The discussion from this chart is clear: the market is not homogenous. We see a "Quantity" strategy from Prime Video and Netflix, who dominate content volume, and a "Curated" strategy from Disney+ and Hulu, which focus on niche, high-value libraries.
- **Head-to-Head Comparison Tool:** This interactive feature is a primary functional result of the project. It successfully validates the "Head-to-Head" objective by allowing a dynamic, user-driven comparison. The discussion from this tool is powerful; a user can, for example, definitively prove that Netflix's content library is not only larger than Disney+'s but also fundamentally different in its strategic focus (e.g., "International Movies" vs. "Family").
- **Real-Time API Integration:** The "Title Intelligence Terminal" is a key technical result. The successful integration of the TMDb API (using requests and api_utils.py) is validated by the

two functional components:

1. **Title Search:** The system can successfully query the TMDb API, parse the JSON response, and display the poster, overview, rating, and user reviews for a searched title.
2. **Trending Content:** The system successfully fetches and displays the top 10 trending movies, providing a live "pulse" of the market that a static dataset could never offer.

5.2 PLATFORM-SPECIFIC DASHBOARDS

The core results of the project are the four deep-dive dashboards. Each dashboard is organized into logical tabs (e.g., Library, Audience, Temporal, Geographic) and presents a suite of 10-12 visualizations. The "discussion" of these results involves synthesizing the data to build a clear picture of each platform's unique strategy.

5.2.1 Discussion: The Netflix Strategy (Global Dominance)

The results from the Netflix dashboard paint a clear picture of a global, balanced, and mature content strategy.

- **Finding:** The "Content Type" pie chart (Fig 4.3) shows a balanced 69/31 split between movies and TV shows.
- **Discussion:** This is the result of a deliberate, multi-billion dollar pivot to original television production to create a "sticky" ecosystem of recurring content, moving beyond its original model of only licensing films.
- **Finding:** The "Top Genres" chart (Fig 4.5) identifies "International Movies" and "Dramas" as #1 and #2. The "Geographic Footprint" map (Fig 4.8) shows massive production hubs in India, the UK, South Korea, and Spain.
- **Discussion:** This is the project's clearest finding. Netflix's primary strategy is not just American-centric; it is a full-scale global operation. It is both acquiring and producing localized content at scale to drive growth in international markets, a strategy none of its US-based competitors are matching.

5.2.2 Discussion: The Prime Video Strategy (The Warehouse)

The Prime Video dashboard results reveal a strategy built on sheer volume.

- **Finding:** The "Content Type" chart (Fig 4.9) shows a library that is over 80% movies.
- **Finding:** The "Content by Release Year" chart (Fig 4.11) shows a massive volume of titles from the 1950s-1970s, unlike any other platform.
- **Discussion:** Prime Video's strategy is not to curate, but to be a "digital warehouse." It leverages its Amazon Prime subscription model by offering an enormous back-catalog of older, likely inexpensive, licensed films as a high-quantity add-on value.

5.2.3 Discussion: The Disney+/Hulu Strategy (The Niche & The Broad)

The results from the Disney+ and Hulu dashboards are most insightful when discussed together, as they represent a deliberate, two-pronged market segmentation strategy.

- **Finding (Disney+) :** Top genres are "Family" and "Animation". Top ratings are "TV-G" and "PG"
- **Finding (Hulu) :** Top genre is "Dramas". The top rating is "TV-MA" (Mature Audiences).
- **Discussion :** This is a perfect and explicit example of market segmentation. The Walt Disney Company uses its two primary streaming assets to capture opposite ends of the market. Disney+ is the brand-safe, family-focused "vertical," while Hulu serves as the "broad" platform for mature, adult-oriented television and licensed content.

5.3 BI INSIGHTS ENGINE

A key result of this project is the automated "BI-Powered Recommendations" section at the bottom of each dashboard, managed by utils/insights.py. This component moves beyond *what* the data says (EDA) to *so what?* (Discussion).

Example Result (Hulu):

- **Data:** The "Content Lag" histogram (Fig 4.17) shows an average lag of over 7 years.
- **Generated Insight:** *"With an average content lag of 7.1 years, Hulu's strategy relies heavily on a deep back-catalog of licensed shows and movies. This is a cost-effective model for content volume."*

This feature successfully demonstrates that the analytical process itself can be partially automated, providing users with pre-packaged strategic takeaways directly within the dashboard. This elevates the project from a simple data visualizer to a true Business Intelligence tool.

CHAPTER 6

CONCLUSION AND FUTURE INSIGHTS

6.1 CONCLUSION

This project, "DataFlix: The OTT Intelligence Dashboard," set out to solve the problem of siloed, inconsistent, and inaccessible data in the highly competitive streaming market. The project successfully met all its objectives by designing, developing, and implementing a high-performance, multi-page Business Intelligence dashboard using Python, Streamlit, Pandas, and Plotly.

A robust data preparation pipeline was successfully engineered to ingest and harmonize four disparate datasets, correcting major data quality issues related to temporal, durational, and multi-value categorical features. This clean data was used to power a comprehensive suite of over 40 interactive visualizations, organized into a global home page and four deep-dive platform dashboards.

The analysis of these visualizations provided clear, data-driven evidence of the distinct strategies employed by each platform: Netflix's global-centric, balanced portfolio; Prime Video's "digital warehouse" model; and the bifurcated Disney+ (family-niche) and Hulu (adult-broad) strategy. Furthermore, the dashboard was successfully enriched with real-time data from the TMDb API, providing live search, review, and trending functionalities. The final application, wrapped in a custom "glassmorphism" CSS theme, achieves the desired aesthetic of a modern, professional, and shippable data product. In conclusion, DataFlix successfully transforms raw, fragmented data into a unified, actionable, and aesthetically engaging analytical tool.

6.2 FUTURE WORK

While the current version of DataFlix is a comprehensive and functional prototype, the potential for expansion is vast. The following steps represent a clear roadmap for future development.

1. Expand Data Sources:

- **More Platforms:** The most logical next step is to ingest data from other major platforms, such as **HBO Max (Max)**, **Apple TV+**, **Peacock**, and **Paramount+**, to provide an even more complete market picture.
- **Deeper Metrics:** The analysis could be significantly enhanced by integrating new data sources

for each title, such as:

- **User Ratings:** Scraping or API integration with IMDb and Rotten Tomatoes for audience and critic sentiment scores.
- **Social Media Sentiment:** Using Twitter's API to perform sentiment analysis on key titles to measure public "buzz."
- **Financial Data:** Linking titles to production budgets (where available) to perform ROI analysis.

2. Implement Machine Learning & Predictive Analytics:

- The current project is descriptive and diagnostic. The next phase would be to add predictive and prescriptive analytics.
- **Content Success Prediction:** Build a machine learning model (e.g., Random Forest, XGBoost) trained on features like genre, cast, director, and budget to predict a title's potential success or rating.
- **Time-Series Forecasting:** Use models like ARIMA or Prophet to forecast a platform's library growth based on its historical date_added trends.

3. Enhance User Functionality:

- **User Accounts:** Implement user authentication to allow users to save their own custom dashboards, comparisons, and watchlists.
- **Advanced Filtering:** Add more granular filters, such as filtering by specific actors, directors, or a custom date range.

4. Full-Scale Deployment:

- While the project runs locally via Streamlit, the final step would be to deploy it to a public cloud platform (e.g., **Streamlit Community Cloud, Heroku, or AWS**) to make it accessible to a global audience. This would require moving the CSV datasets to a cloud-based database (like Amazon S3 or a SQL database) for more robust data management.

CHAPTER 7

REFERENCES

- [1] **Pandas (Python Data Analysis Library).** (2024). *pandas - Python Data Analysis Library*. [Software]. Available at: <https://pandas.pydata.org/>
- [2] **Streamlit Inc.** (2024). *Streamlit - A faster way to build and share data apps*. [Software]. Available at: <https://streamlit.io/>
- [3] **Plotly Technologies Inc.** (2024). *Plotly - Low-Code Data App Development*. [Software]. Available at: <https://plotly.com/>
- [4] **The Movie Database (TMDb).** (2024). *TMDb API*. [Data API]. Available at: <https://www.themoviedb.org/documentation/api>
- [5] (Data Source) Shivamb. (2021). *Netflix Movies and TV Shows*. [Dataset]. Kaggle. Available at: <https://www.kaggle.com/datasets/shivamb/netflix-shows>
- [6] (Data Source) Shivamb. (2021). *Amazon Prime Movies and TV Shows*. [Dataset]. Kaggle. Available at: <https://www.kaggle.com/datasets/shivamb/amazon-prime-movies-and-tv-shows>
- [7] (Data Source) Shivamb. (2021). *Disney+ Movies and TV Shows*. [Dataset]. Kaggle. Available at: <https://www.kaggle.com/datasets/shivamb/disney-plus-movies-and-tv-shows>
- [8] (Data Source) Shivamb. (2021). *Hulu Movies and TV Shows*. [Dataset]. Kaggle. Available at: <https://www.kaggle.com/datasets/shivamb/hulu-movies-and-tv-shows>
- [9] Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90-95. (Used by WordCloud)
- [10] LottieFiles. (2024). *Lottie Animations*. [Assets]. Available at: <https://lottiefiles.com/>
- [11] Turner, D. (2023). *The Streaming Wars: Business and Strategy*. Tech Analytics Press.
- [12] Vaidhyanathan, S. (2018). *Antisocial Media: How Facebook Disconnects Us and Undermines Democracy*. Oxford University Press. (Relevant for media consumption analysis).

APPENDIX

A.1: Core Application Code (app.py)

```
import streamlit as st
import os
import json
import time
from streamlit_lottie import st_lottie
from components import home_page, netflix_dashboard, prime_dashboard, disney_dashboard,
hulu_dashboard
```

#PAGE CONFIGURATION

```
st.set_page_config(
    page_title="DataFlix",
    page_icon="🍿",
    layout="wide",
    initial_sidebar_state="auto"
)
```

#LOAD LOTTIE ANIMATION

```
def load_lottiefile(filepath: str):
    """Loads a Lottie JSON file from the specified filepath."""
    try:
        with open(filepath, "r") as f:
            return json.load(f)
    except FileNotFoundError:
        st.warning("Lottie animation file not found.")
        return None
```

```

#Animated Loading Screen

def show_loading_screen():
    """Displays the animated loading screen."""
    # Note: Lottie file path is relative to the root directory
    loading_animation = load_lottiefile("assets/loading_animation.json")
    if loading_animation:
        with st.spinner(" "): # Spinner without text
            st_lottie(loading_animation, height=200, key="loading")
            st.markdown("<h3 style='text-align: center; color: #FFFFFF;'>Turning Streaming Data  
into Strategy...</h3>", unsafe_allow_html=True)
            time.sleep(2.5) # Simulate loading time
    else:
        time.sleep(1)

```

#STATE MANAGEMENT

```

if 'page' not in st.session_state:
    st.session_state.page = 'Home'
if 'app_loaded' not in st.session_state:
    st.session_state.app_loaded = False

def set_page(page_name):
    """Callback function to change the current page in session state."""
    st.session_state.page = page_name

```

#LOAD STYLES

```

def load_css(file_name):
    """Loads a local CSS file into the Streamlit app."""
    script_dir = os.path.dirname(os.path.abspath(__file__))
    css_file_path = os.path.join(script_dir, file_name)
    try:

```

```

with open(css_file_path) as f:
    st.markdown(f<style>{f.read()}</style>', unsafe_allow_html=True)
except FileNotFoundError:
    st.error(f"CSS file not found at {css_file_path}")

```

#Main App Logic

```

load_css('style.css')
if not st.session_state.app_loaded:
    show_loading_screen()
    st.session_state.app_loaded = True
    st.rerun()

```

```
else:
```

#SIDEBAR

```

with st.sidebar:
    st.header("DataFlix 🌟")
    st.markdown("The OTT Intelligence Dashboard.")

```

```
PAGES = ["Home", "Netflix", "Prime Video", "Disney+", "Hulu"]
```

```
ICONS = ["🏠", "🔴", "🔵", "🟡", "🟢"]
```

```
for page, icon in zip(PAGES, ICONS):
```

```

    if st.button(f"{icon} {page}", use_container_width=True, on_click=set_page,
    args=(page,)):
        pass
    st.markdown("---")

```

#BACKGROUND MUSIC TOGGLE

```

st.write("Ambient Music")
audio_html = """
<script>

```

```

document.addEventListener('DOMContentLoaded', function() {
  const audio = document.getElementById("bgMusic");
  const checkbox = document.getElementById("musicToggle");

  if (checkbox) {
    checkbox.addEventListener('change', function() {
      if (this.checked) {
        audio.play();
      } else {
        audio.pause();
      }
    });
  }
});

</script>
<audio id="bgMusic" loop>
  <source src="https://www.chosic.com/wp-content/uploads/2022/01/And-So-It-Begins-Inspired-By-Crush-Sometimes.mp3" type="audio/mpeg">
</audio>
<label class="switch">
  <input type="checkbox" id="musicToggle">
  <span class="slider round"></span>
</label>
"""

st.components.v1.html(audio_html, height=35)

```

#MAIN PAGE ROUTING

```

PAGE_MAP = {
  "Home": home_page.show_home_page,
  "Netflix": netflix_dashboard.show_netflix_dashboard,
  "Prime Video": prime_dashboard.show_prime_dashboard,
}

```

```

    "Disney+": disney_dashboard.show_disney_dashboard,
    "Hulu": hulu_dashboard.show_hulu_dashboard,
}

if st.session_state.page == "Home":
    PAGE_MAP[st.session_state.page](set_page)
else:
    PAGE_MAP[st.session_state.page]()

```

A.2: Styling Code (style.css)

```

/* Import Font */
@import
url('https://fonts.googleapis.com/css2?family=Poppins:wght@300;400;600;700&display=swap')
('https://fonts.googleapis.com/css2?family=Poppins:wght@300;400;600;700&display=swap');

/* Base Styling & Futuristic Dark Theme */
body, .stApp, p, h1, h2, h3, h4, h5, h6, .stMarkdown, .stButton>button,
.stTextInput>div>div>input, .st-emotion-cache-1gulkj5, .st-emotion-cache-16idsys p {
    font-family: 'Poppins', sans-serif;
    color: #FFFFFF !important; /* Force all text to be bright white */
}

.stApp {
    background-color: #0c001f;
    background-image: radial-gradient(circle at 20% 20%, rgba(100, 30, 255, 0.2), transparent
40%),
        radial-gradient(circle at 80% 80%, rgba(255, 60, 150, 0.2), transparent 40%);
}

```

```

/* Sidebar */

[data-testid="stSidebar"] {
    background-color: rgba(12, 0, 31, 0.8);
    backdrop-filter: blur(10px);
    border-right: 1px solid rgba(255, 255, 255, 0.1);
}

[data-testid="stSidebar"] .stButton button {
    background-color: transparent;
    border: 1px solid transparent;
    color: #FFFFFF !important;
    text-align: left;
    transition: all 0.2s ease-in-out;
}

[data-testid="stSidebar"] .stButton button:hover {
    background-color: rgba(0, 209, 255, 0.1);
    border: 1px solid #00d1ff;
    color: #00d1ff !important;
    box-shadow: 0 0 10px rgba(0, 209, 255, 0.5);
}

/* Styling for NATIVE Streamlit Containers (Glassmorphism Cards) */

[data-testid="stVerticalBlockBorderWrapper"] {
    background-color: rgba(20, 5, 50, 0.6);
    border-radius: 15px !important;
    border: 1px solid rgba(255, 255, 255, 0.1) !important;
    box-shadow: 0 4px 30px rgba(0, 0, 0, 0.1);
    backdrop-filter: blur(10px);
    -webkit-backdrop-filter: blur(10px);
    transition: all 0.3s ease-in-out;
    padding: 1.5rem !important;
    margin-bottom: 1rem;
}

```

```
}
```

```
[data-testid="stVerticalBlockBorderWrapper"]:hover {  
    transform: translateY(-5px);  
    box-shadow: 0 8px 30px rgba(0, 209, 255, 0.2), 0 0 20px rgba(255, 60, 150, 0.2);  
    border: 1px solid rgba(0, 209, 255, 0.5) !important;  
}
```

/* Styling for Clickable Logo Cards */

```
[data-testid="stVerticalBlockBorderWrapper"] .stImage img {  
    max-height: 80px;  
    max-width: 150px;  
    width: auto;  
    object-fit: contain;  
}
```

```
[data-testid="stVerticalBlockBorderWrapper"] .stButton button {
```

```
    position: absolute;  
    top: 0;  
    left: 0;  
    width: 100%;  
    height: 100%;  
    background-color: transparent !important;  
    color: transparent !important;  
    border: none !important;  
    box-shadow: none !important;  
}
```

/* Titles and Headers with Glow */

```
h1, h2 {
```

```

color: #FFFFFF !important;
text-shadow: 0 0 8px rgba(0, 209, 255, 0.6);
}

h3 {
    color: #FFFFFF !important;
    font-weight: 600;
}

div[data-testid="stHeading"] > h3 {
    position: relative;
    padding-left: 15px !important;
}

div[data-testid="stHeading"] > h3::before {
    content: "";
    position: absolute;
    left: 0;
    top: 50%;
    transform: translateY(-50%);
    height: 60px;
    width: 4px;
    background-color: #00d1ff;
    border-radius: 2px;
}

/* KPI Metrics */

[data-testid="stMetricLabel"], [data-testid="stMetricValue"] {
    color: #FFFFFF !important;
}

[data-testid="stMetricValue"] {
    text-shadow: 0 0 5px rgba(255, 255, 255, 0.5);
}

```

```

/* TABS Styling for BI Layout */

[data-testid="stTabs"] button {
    color: #A0A0A0 !important;
    background-color: transparent !important;
    border-radius: 8px !important;
    margin: 0 5px !important;
    transition: all 0.2s ease-in-out;
}

[data-testid="stTabs"] button[aria-selected="true"] {
    color: #FFFFFF !important;
    background-color: rgba(0, 209, 255, 0.2) !important;
    border-bottom: 3px solid #00d1ff !important;
    box-shadow: 0 0 10px rgba(0, 209, 255, 0.5);
}

/* Music Toggle Switch */

.switch { position: relative; display: inline-block; width: 50px; height: 28px; }

.switch input { opacity: 0; width: 0; height: 0; }

.slider { position: absolute; cursor: pointer; top: 0; left: 0; right: 0; bottom: 0; background-color: #333; transition: .4s; }

.slider:before { position: absolute; content: ""; height: 20px; width: 20px; left: 4px; bottom: 4px; background-color: white; transition: .4s; }

input:checked + .slider { background-color: #00d1ff; }

input:checked + .slider:before { transform: translateX(22px); }

.slider.round { border-radius: 34px; }

.slider.round:before { border-radius: 50%; }

/* Fade-in Animation */

@keyframes fadeIn { from { opacity: 0; transform: translateY(10px); } to { opacity: 1; transform: translateY(0); } }

```

```
[data-testid="stVerticalBlockBorderWrapper"], [data-testid="stTabs"] { animation: fadeIn 0.8s ease-in-out; }
```

A.3: API Utility Code (api_utils.py)

```
import streamlit as st
import requests

# Fetch the API key from Streamlit secrets
try:
    API_KEY = st.secrets["TMDB_API_KEY"]
except (KeyError, FileNotFoundError):
    st.error("TMDb API key not found. Please add it to your Streamlit secrets.")
    API_KEY = None

BASE_URL = "[https://api.themoviedb.org/3](https://api.themoviedb.org/3)"

def get_trending_movies():
    """Fetches a list of trending movies from TMDb."""
    if not API_KEY:
        return None
    try:
        response = requests.get(f'{BASE_URL}/trending/movie/day?api_key={API_KEY}')
        response.raise_for_status() # Raises an error for bad responses (4xx or 5xx)
        return response.json().get('results', [])
    except requests.RequestException as e:
        st.warning(f'API Connection Error (Trending): {e}')
        return None

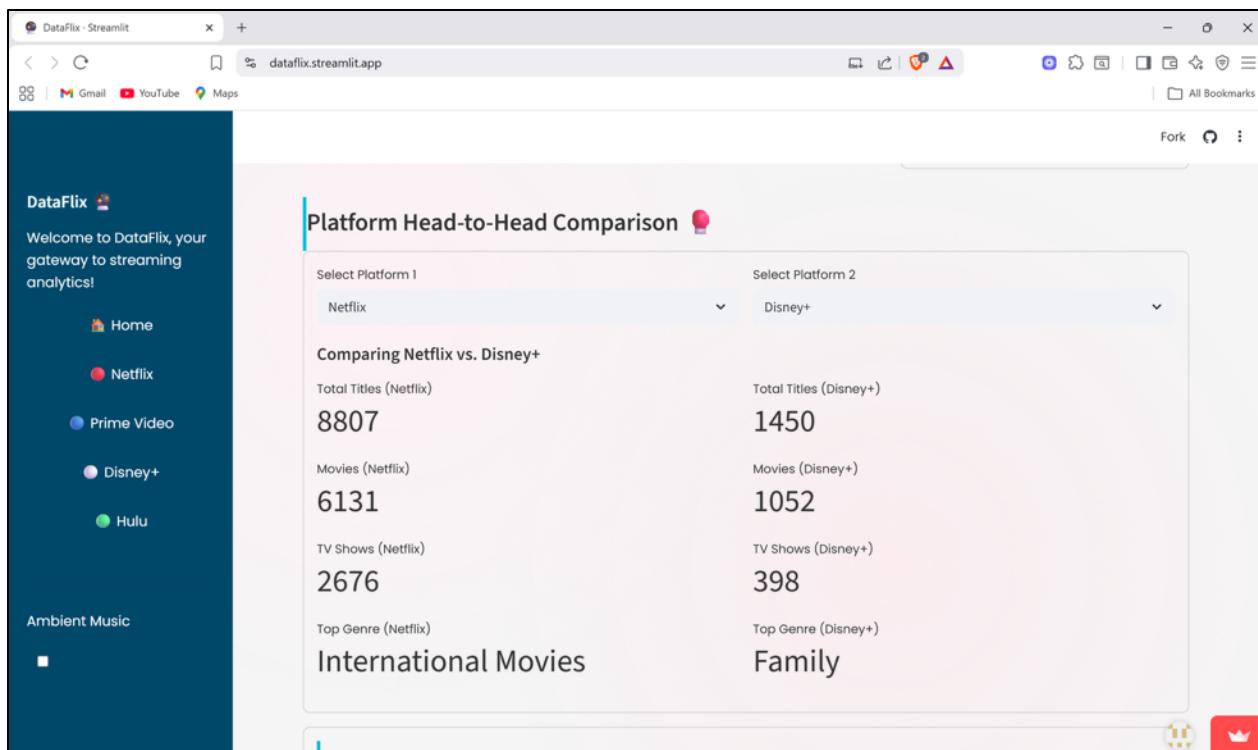
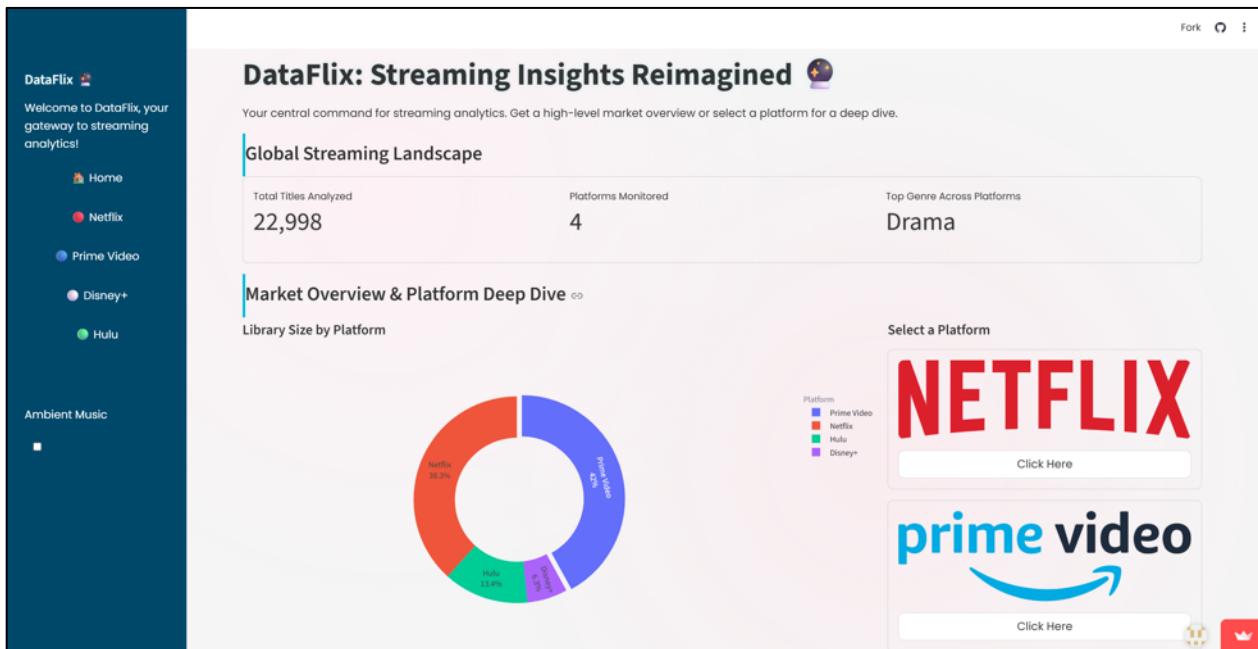
def get_movie_details(query):
    """Searches for a movie and returns its details."""

```

```
if not API_KEY:  
    return None  
  
try:  
    response =  
        requests.get(f'{BASE_URL}/search/movie?api_key={API_KEY}&query={query}')  
        response.raise_for_status()  
    results = response.json().get('results', [])  
    return results[0] if results else None  
  
except requests.RequestException as e:  
    st.warning(f"API Connection Error (Search): {e}")  
    return None  
  
def get_movie_reviews(movie_id):  
    """Fetches reviews for a specific movie by its ID."""  
    if not API_KEY or not movie_id:  
        return None  
  
    try:  
        response =  
            requests.get(f'{BASE_URL}/movie/{movie_id}/reviews?api_key={API_KEY}')  
            response.raise_for_status()  
        return response.json().get('results', [])  
  
    except requests.RequestException as e:  
        # Silently fail on reviews as it's non-critical  
        return None
```

A.4: FINAL DASHBOARD

A.4.1: Final Home Page Output



DataFlix

Welcome to DataFlix, your gateway to streaming analytics!

- Home
- Netflix
- Prime Video
- Disney+
- Hulu

Ambient Music

Title Intelligence Terminal

Search for a Movie or TV Show

Enter title: Dune

Search

Dune

Rating: 7.8/10 ★

Overview: Paul Atreides, a brilliant and gifted young man born into a great destiny beyond his understanding, must travel to the most dangerous planet in the universe to ensure the future of his family and his people. As malevolent forces explode into conflict over the planet's exclusive supply of the most precious resource in existence—a commodity capable of unlocking humanity's greatest potential—only those who can conquer their fear will survive.

Trending Transmissions Today

Fork ⌂ ⌓

A.4.2: Final Netflix Dashboard Screenshot

DataFlix

Welcome to DataFlix, your gateway to streaming analytics!

- Home
- Netflix
- Prime Video
- Disney+
- Hulu

Ambient Music

Filters

Content Type

All

Netflix Content Intelligence

Total Titles: 8,807

Movies: 6,131

TV Shows: 2,676

Top Content Country: United States

Library Composition

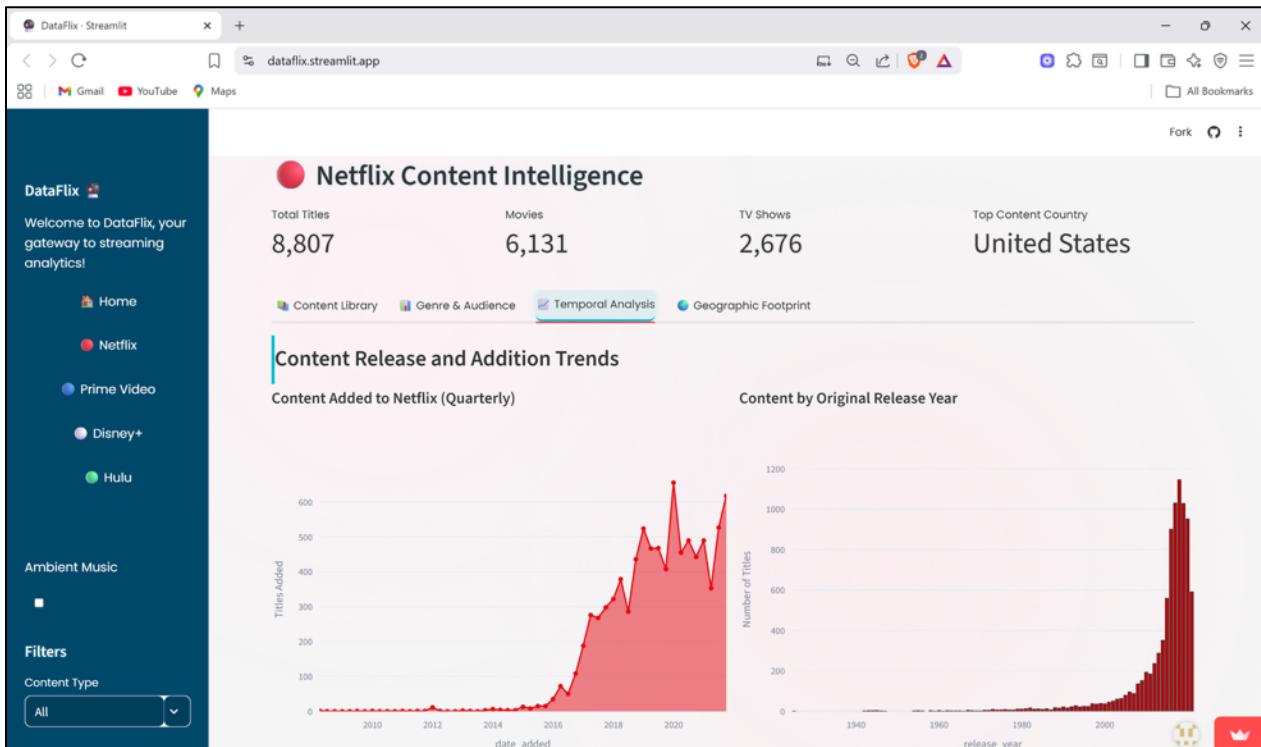
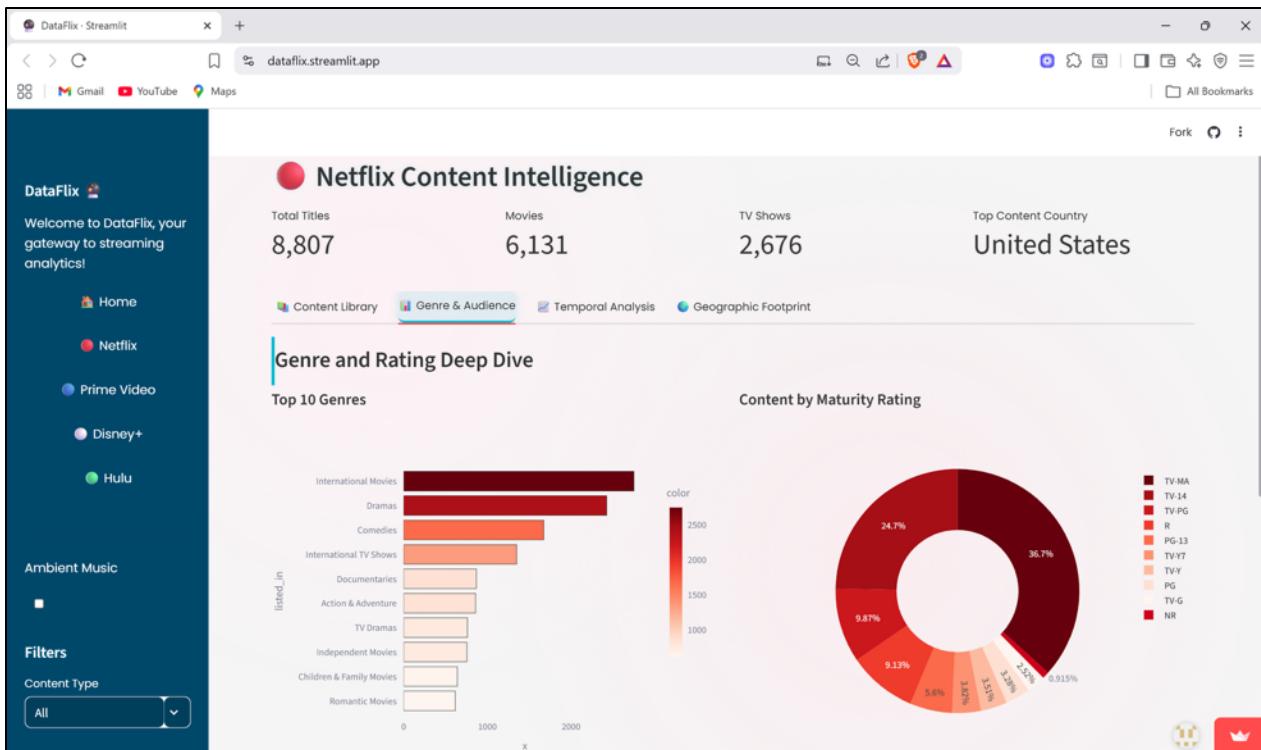
Content Type Distribution

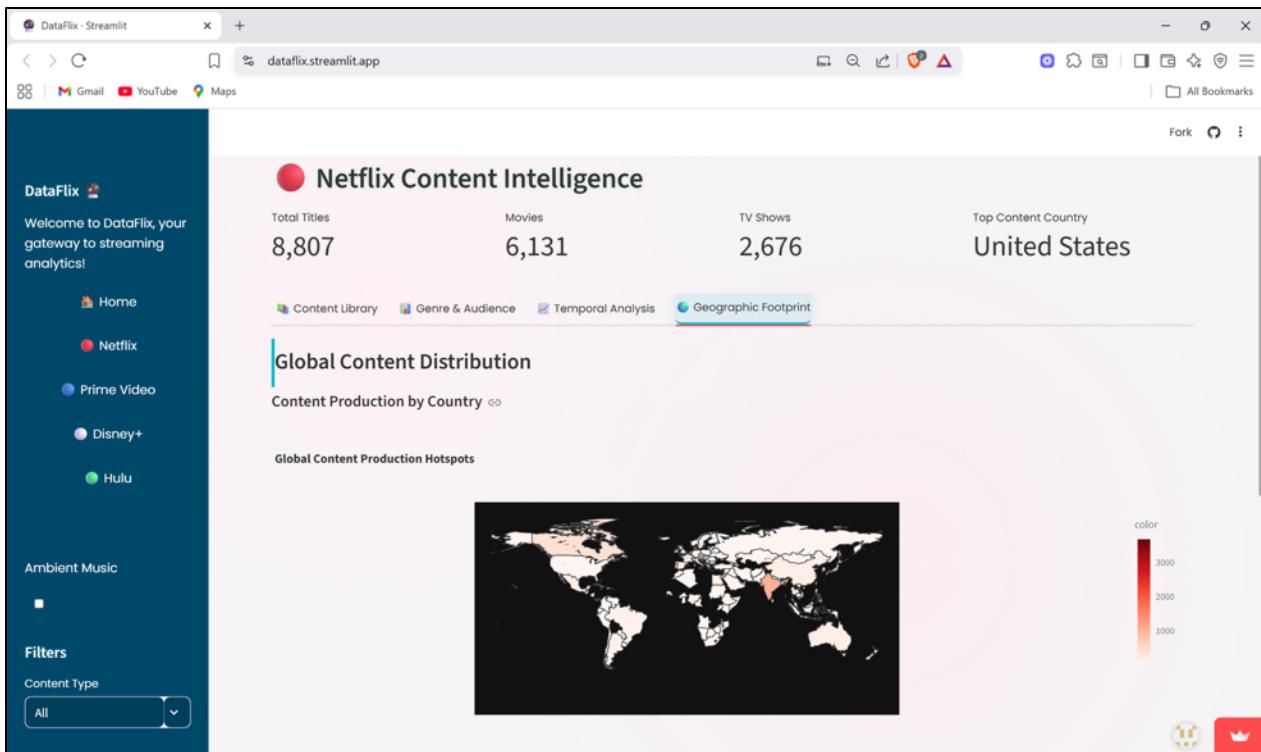
Movie Duration vs. TV Show Seasons

Legend: Movie (red), TV Show (dark red)

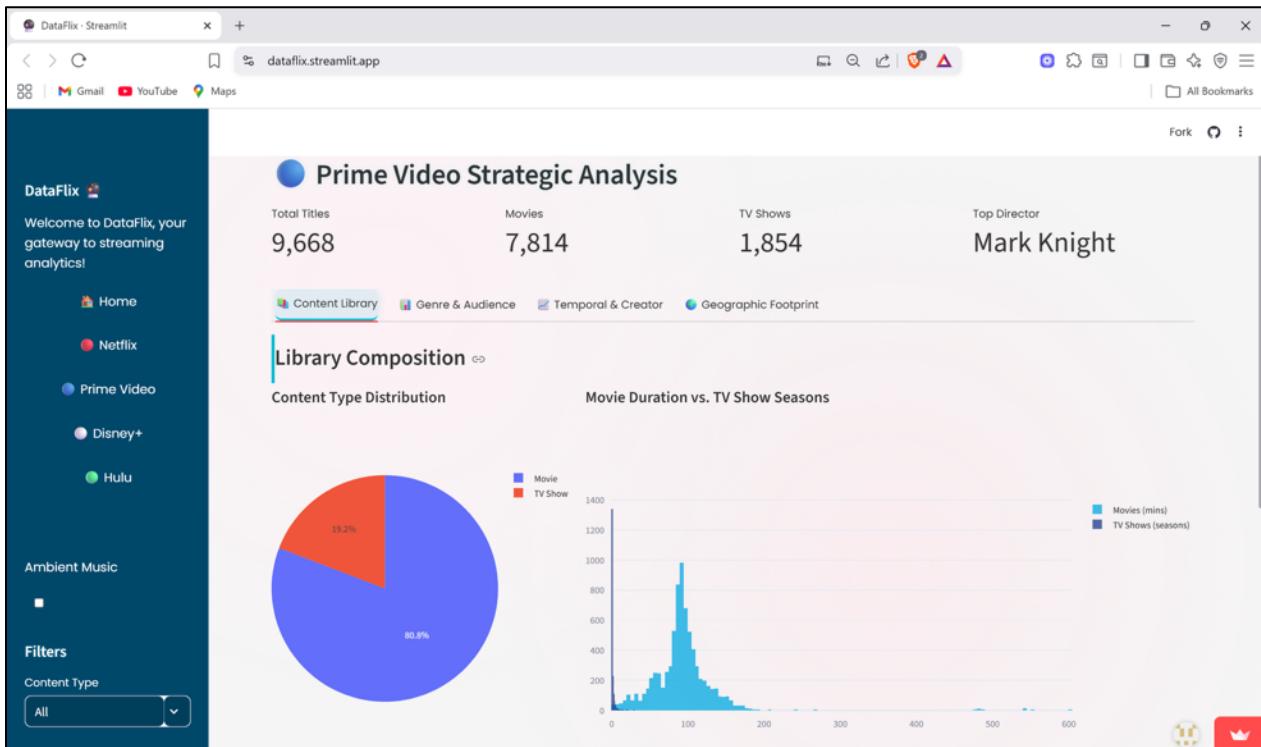
Legend: Movies (mins) (red), TV Shows (seasons) (dark red)

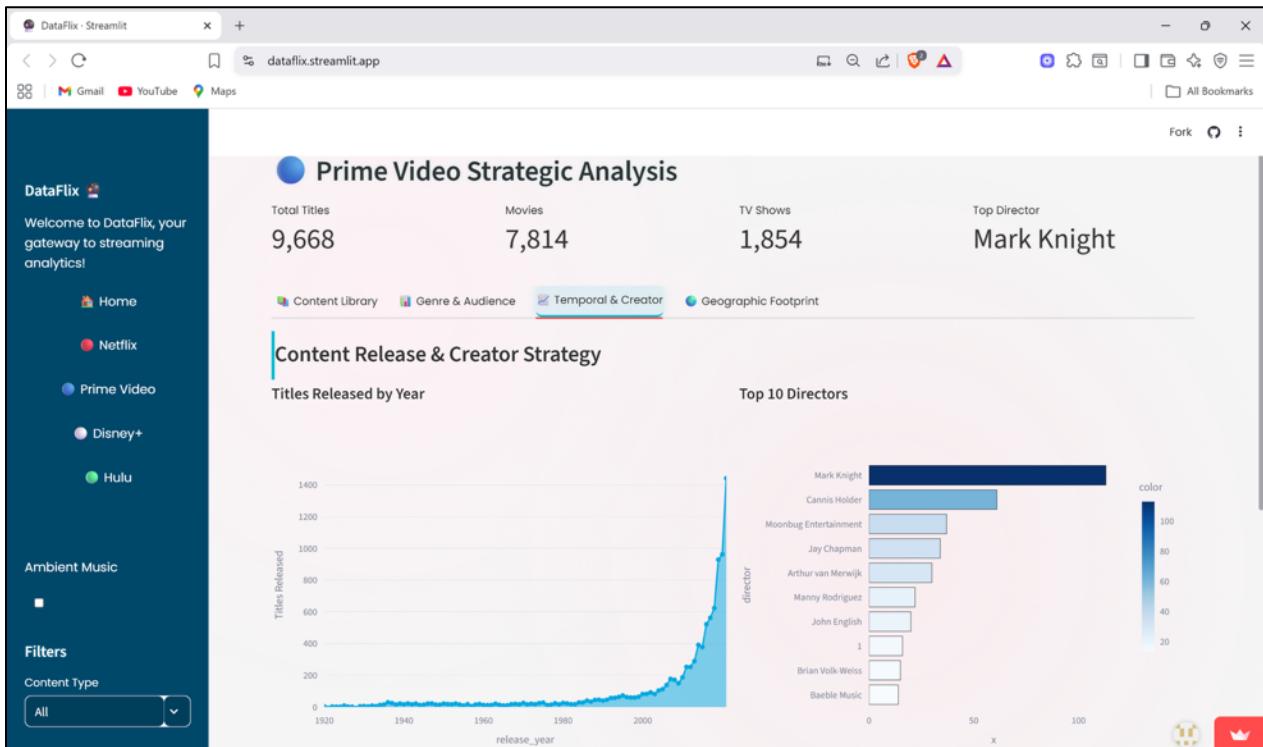
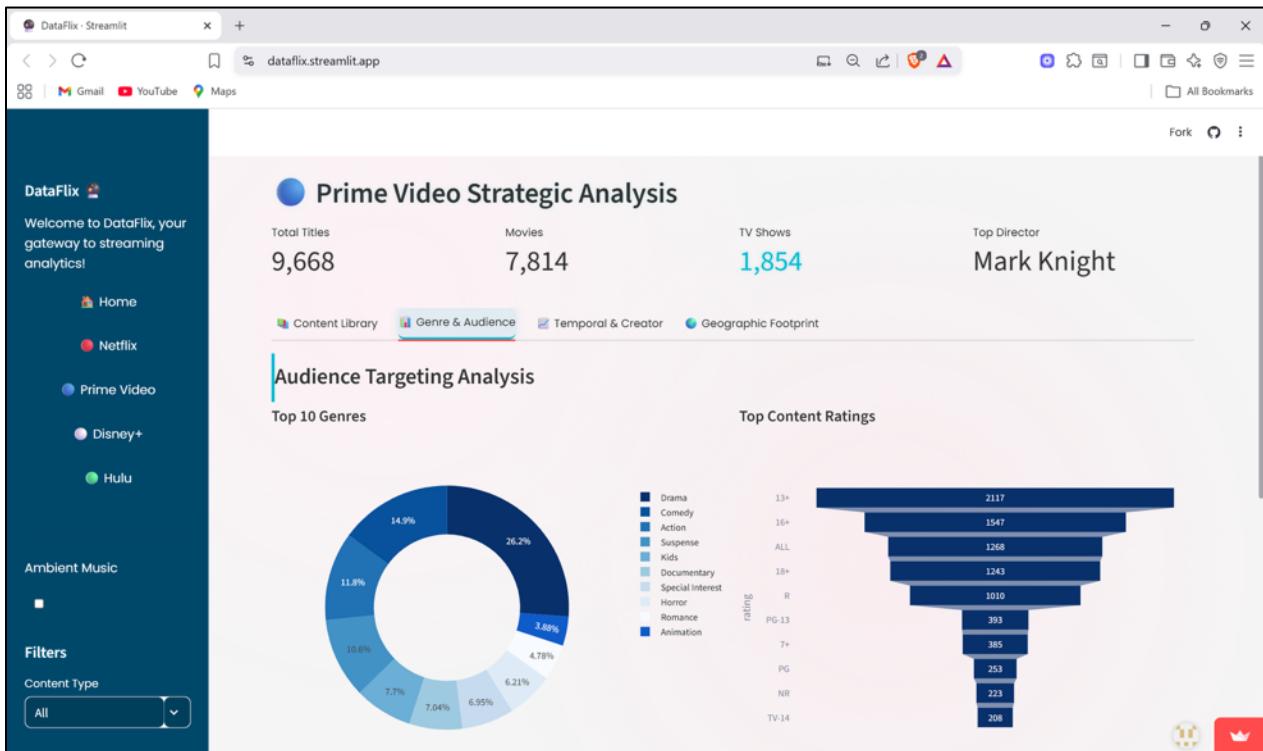
Fork ⌂ ⌓

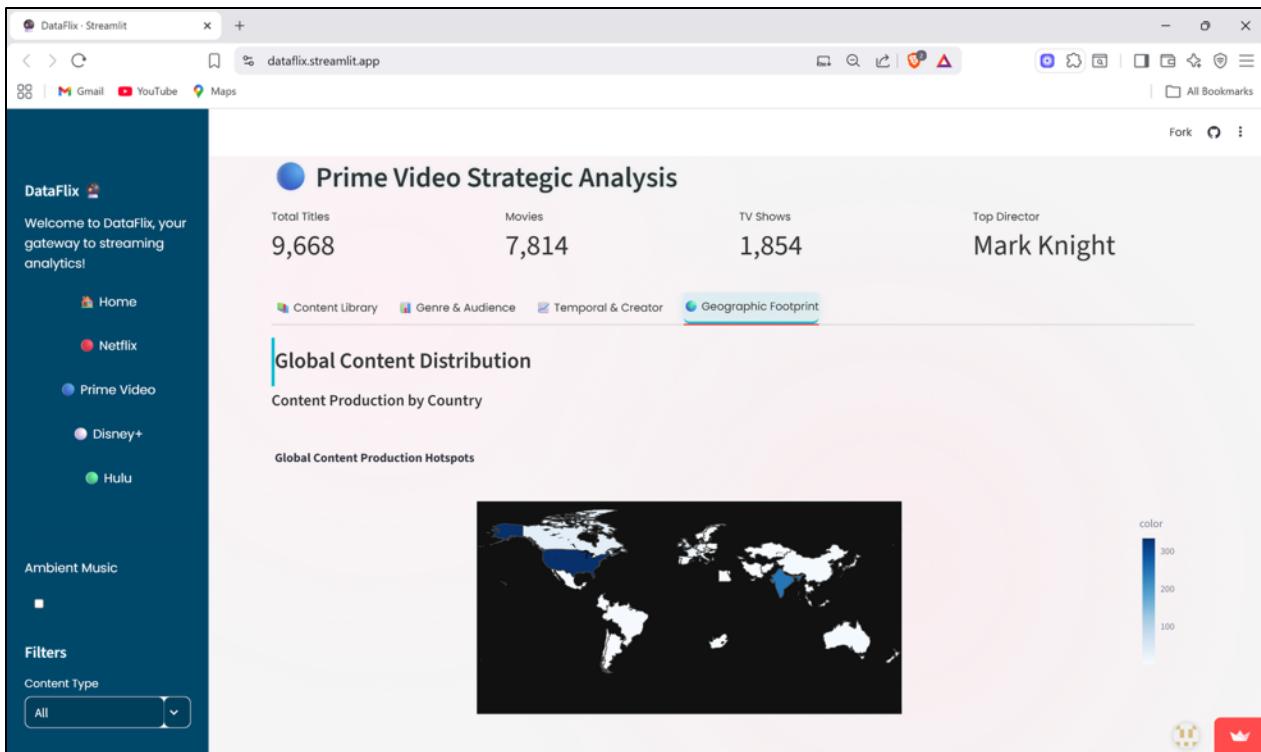




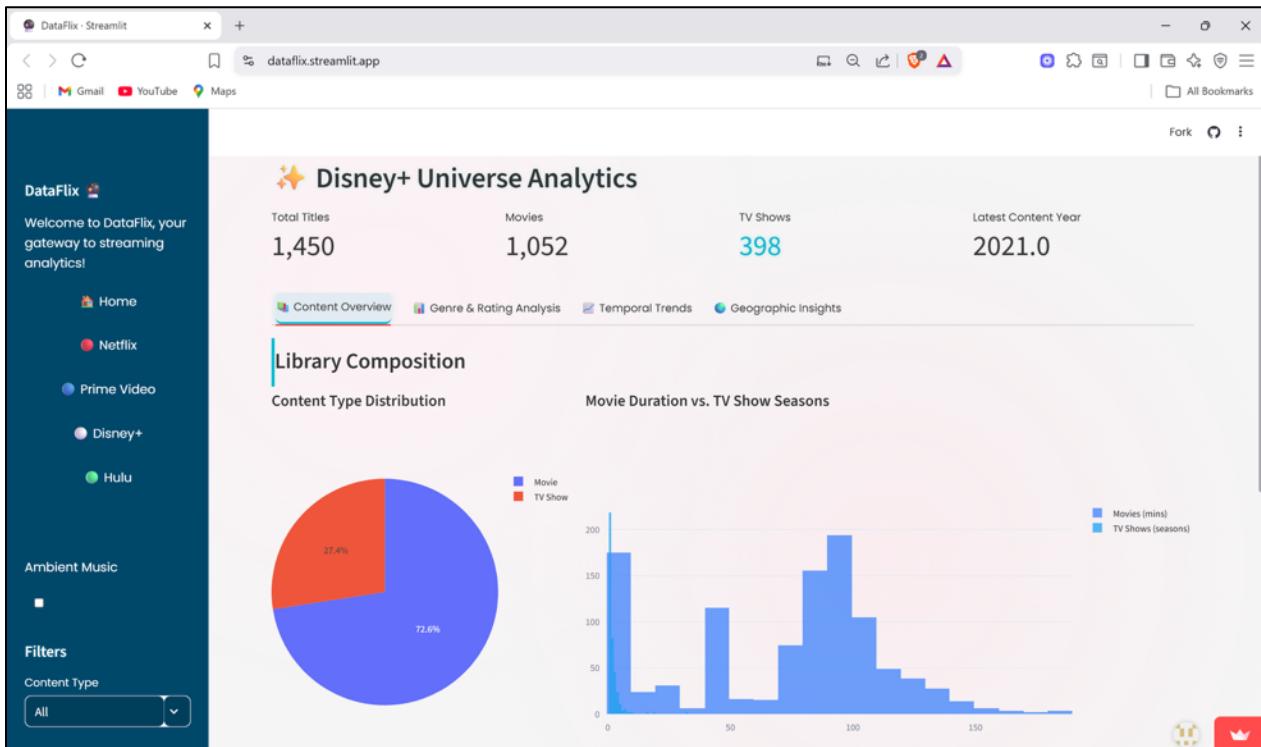
A.4.3: Final Prime Video Dashboard Screenshot

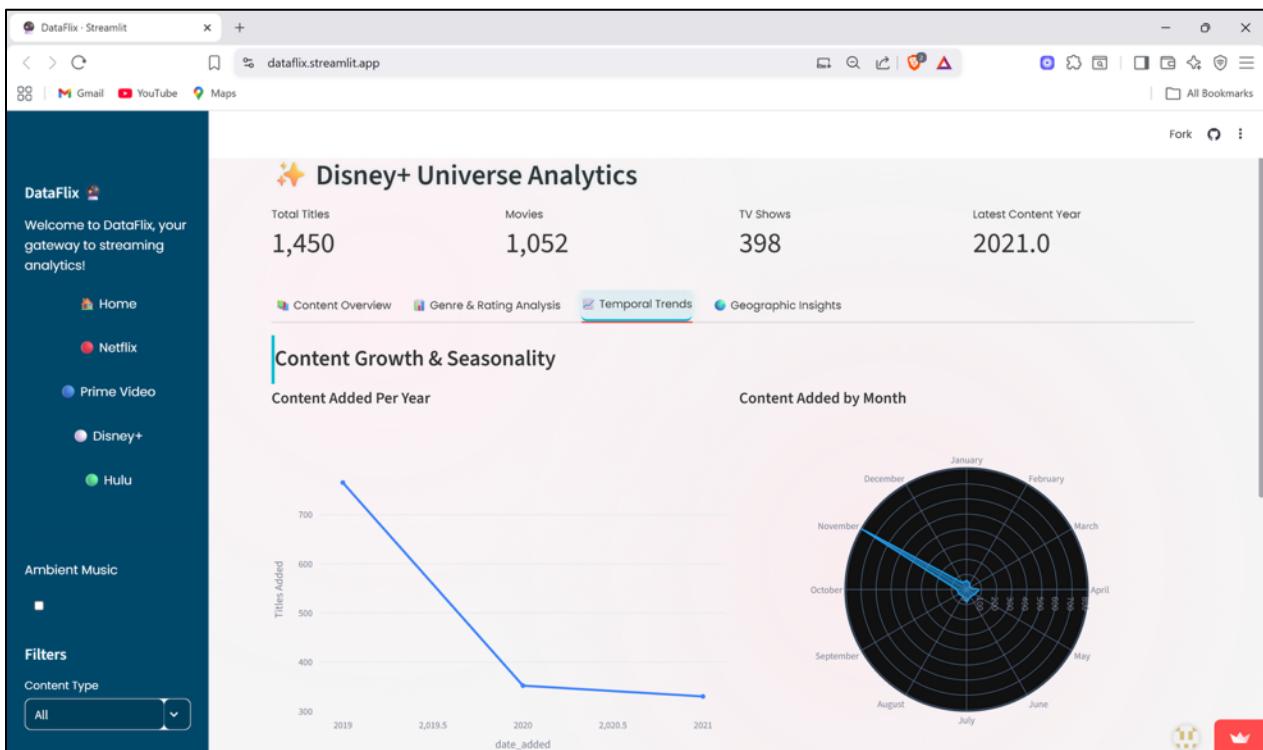
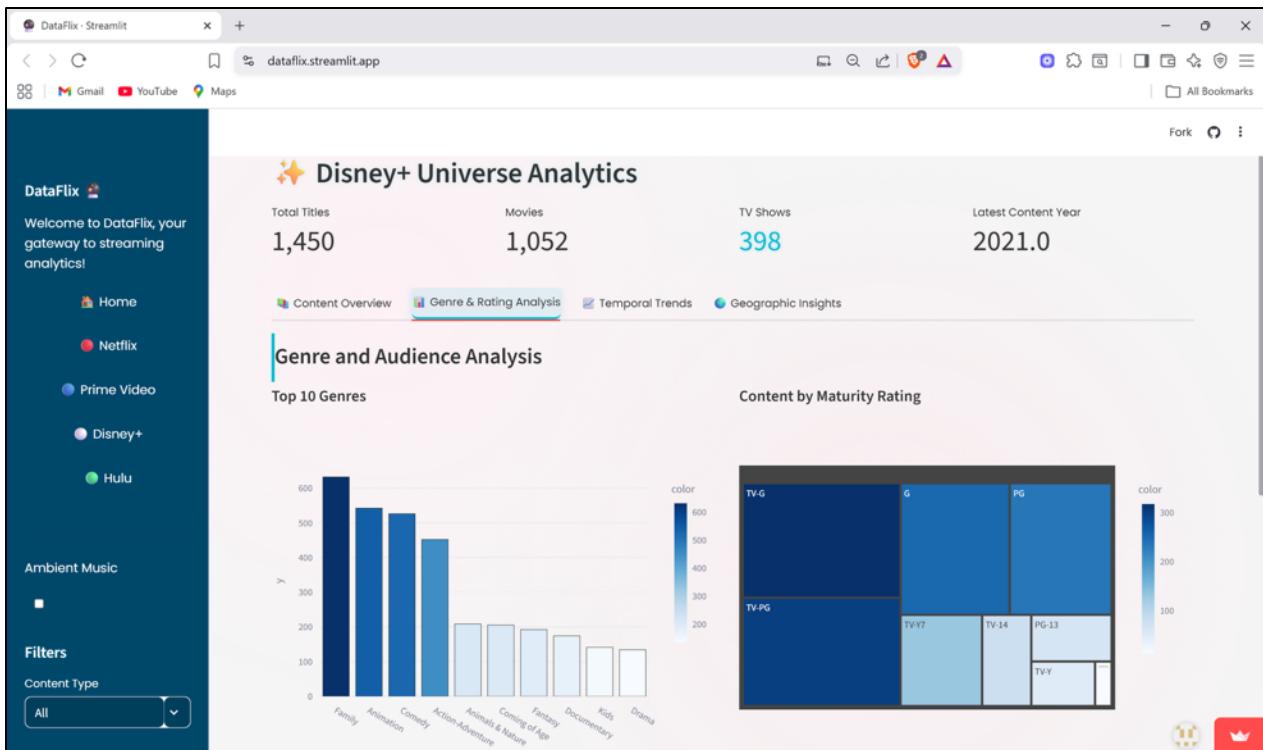


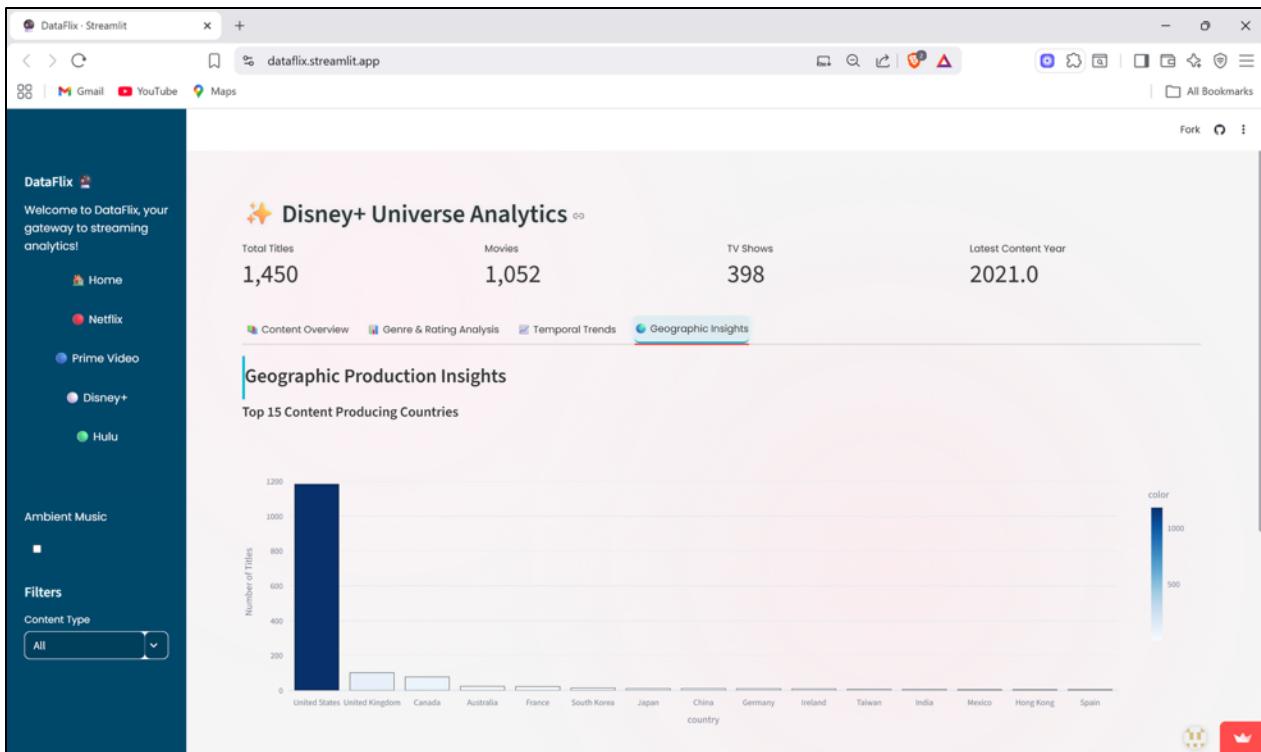




A.4: Final Disney+ Dashboard Screenshot







A.5: Final Hulu Dashboard Screenshot

