
ENTERTAINMENT / OTT STREAMING INDUSTRY

Strategic Content Investment Analysis for a Global OTT Platform

Mayank Sharma

2401010270

Hrishi Dongre

2401010191

Aman Soni

2401010062

Himanshu Pandey

2401020095

Aarya Srivastav

2401010008

Sarvjeet Yadav

2401010424

SECTOR CONTEXT



On-Demand Growth

The OTT streaming industry is driven by on-demand consumption, global accessibility, and high competition.



Content Differentiation

Content quality, diversity, and freshness directly influence user engagement and platform differentiation.



Cost Strategy

Since content acquisition involves high costs, strategic decisions must be supported by structured data analysis.

KEY STAKEHOLDER

Content Strategy & Investment Team

Responsible for portfolio mix & regional acquisition

PROBLEM STATEMENT



What insights from historical content distribution—type, genre, country, rating, and growth trends—can guide future content investment and expansion strategy for a global OTT platform?

PRIMARY OBJECTIVE

To support strategic content acquisition, production, and regional expansion decisions using historical metadata analysis.

Data Engineering: Source to Sink

DATA SOURCE & VOLUME

Netflix Movies & TV Shows

Kaggle Dataset Archive

8,808

TOTAL ROWS

2008–2021

TIME PERIOD

Based on Date Added

12

RAW COLS



21

FINAL COLS

Post-Feature Engineering

MAJOR DATA ISSUES ADDRESSED



Missing Values

Imputed or handled gaps in Director, Cast, Country, and Rating fields.



Text Formatting

Standardized inconsistent spacing, mixed case, and delimiter formats.



Mixed Duration Units

Separated mixed formats (e.g., "90 min" vs "2 Seasons") into distinct numeric value and unit columns.

KEY ANALYTICAL COLUMNS SELECTED

● STANDARD ● ENGINEERED

type

release_year

date_added

year_added

primary_genre

primary_country

rating

Rating_Group

duration_value

duration_unit

Content_Lag



Execution Note: All data cleaning, transformation, and preparation were performed natively in Google Sheets.

Metadata-Based KPIs



Total Titles
Catalog Volume



% Movies vs TV Shows
Format Split



Year-wise Content Growth
year_added



Genre Distribution
primary_genre



Country Contribution
primary_country



Rating Group Distribution
Target Audience



Duration / Season Patterns
Content Length



Content Lag
year_added - release_year

Why These KPIs Matter

- 01** Evaluate content mix and portfolio balance between short-form (Movies) and long-form (TV) formats.
- 02** Identify risks associated with over-concentration in specific genres or geographical regions.
- 03** Assess audience targeting strategies through maturity rating segmentation (Kids vs. Adults).
- 04** Measure the pace of catalog expansion and the "freshness" of content acquisition (Lag Analysis).
- 05** Support data-driven investment planning by highlighting under-served opportunities.

Key Insights from Metadata Analysis



Movie Dominance

Movies dominate the catalog, forming approximately **70% of total titles**, indicating a historically movie-heavy portfolio strategy.



Strategic Shift to TV

TV Shows exhibit **stronger growth post-2016**, highlighting an increasing strategic emphasis on serialized content for retention.



Genre Concentration Risk

Drama and Comedy dominate the genre mix, creating potential over-concentration risks while niche genres remain under-served.



Geographic Imbalance

The **United States and India** are the largest content contributors, showing a significant geographic imbalance in the global library.



Mature Audience Targeting

Adult-oriented ratings form the largest audience segment, reflecting a shift towards mature demographic targeting over family content.



Limited Series Preference

Most TV Shows have only **1-2 seasons**, indicating a preference for limited-series formats to manage production risk.

ANALYSIS TYPE

Country Expansion Gap Analysis

WHAT WAS DONE

Content distribution was examined using `primary_country` and `year_added` to systematically compare dominant markets against emerging regions.

Important Clarification:

This analysis is descriptive and exploratory based on historical metadata, not predictive modeling of future performance.

NEW UNDERSTANDING GAINED



US Market Saturation Signals

The United States shows signs of content saturation with slowing growth rates in recent years, suggesting a need to pivot acquisition strategies elsewhere.



International TV Growth

International TV Shows have grown significantly faster than domestic content post-2018, validating the global "local-for-local" content strategy.



Emerging Region Potential

Several high-population emerging regions remain significantly underrepresented in the catalog relative to their market size, indicating clear expansion potential.

Dashboard Walkthrough

EXECUTIVE VIEW

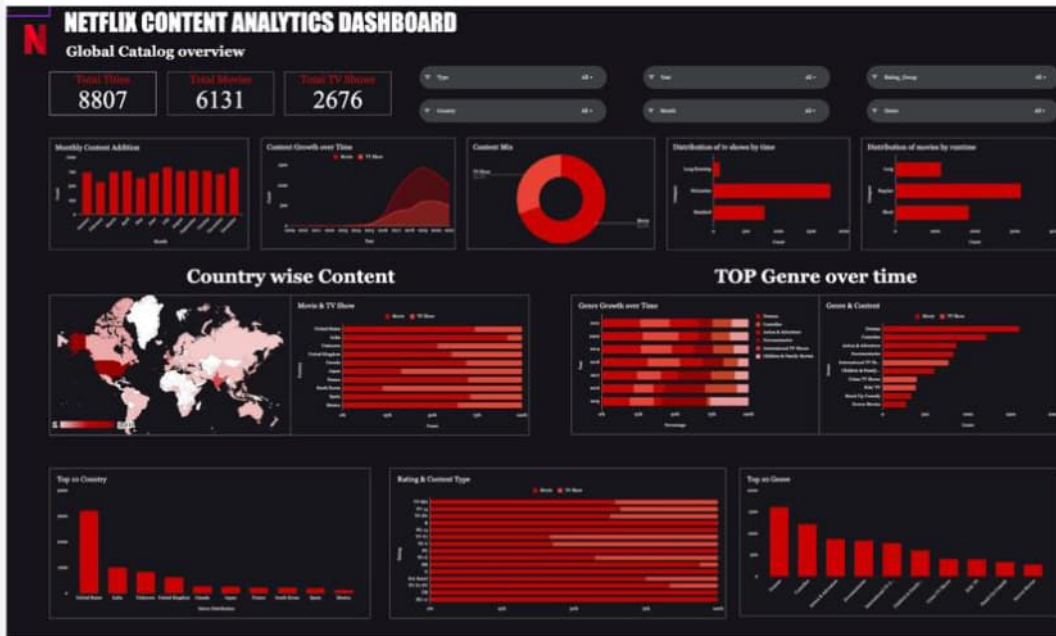
- ✓ Total Titles & Catalog Size
- ✓ Movies vs TV Shows Split
- ✓ Year-wise Content Additions
- ✓ Top Genres and Countries

OPERATIONAL VIEW

- ✓ Drill-downs by Country & Genre
- ✓ Duration & Season Patterns
- ✓ Content Lag & Trend Exploration

DECISION USAGE

Enables leadership to monitor portfolio balance and analysts to identify diversification and expansion opportunities using interactive filters.



Recommendations

01



Increase Investment in TV Shows

Shift content budget allocation towards serialized TV content to drive longer engagement periods and improve subscriber retention rates compared to standalone movies.

🔗 LINKED TO: GROWTH TRENDS & FORMAT SHIFT

02



Expand Regional Production

Mitigate US-market saturation risk by establishing production hubs in high-growth emerging regions (e.g., India, Brazil) to support a "local-for-local" strategy.

🔗 LINKED TO: GEOGRAPHIC IMBALANCE

03



Anchor & Diversify Genres

Maintain Drama and Comedy as core retention anchors while allocating 15-20% of budget to test niche, underrepresented genres for audience expansion.

🔗 LINKED TO: GENRE CONCENTRATION

04



Diversify Rating Categories

Develop specific content slates for Family (TV-Y/G) and Teen (TV-14) segments to balance the portfolio's heavy skew towards Mature (TV-MA) audiences.

🔗 LINKED TO: RATING DISTRIBUTION

05



Prioritize Limited Series

Continue prioritizing 1-2 season limited series formats to test new IP with managed financial risk before committing to long-term multi-season renewals.

🔗 LINKED TO: SEASON PATTERNS

So What? Why Approve This?



Strategic Value

- ✓ Improved clarity on **content mix** and diversification needs for long-term portfolio health.
- ✓ Better alignment of **content investment** with existing catalog structure and identified gaps.
- ✓ Reduced risk from **genre and regional over-concentration** through data-backed diversification strategies.



Operational Value

- ⚡ Faster, **data-driven decision-making** enabled by the new automated KPI dashboards.
- ⚡ More efficient allocation of **content planning efforts** by focusing resources on high-impact areas.

Limitations & Next Steps



Project Limitations



No Financial or Viewership Data

Analysis lacks direct revenue figures, subscriber churn rates, or hours viewed, limiting financial ROI calculation.



Metadata-Only Scope

Insights are derived solely from structured catalog attributes (genre, duration, country) rather than content quality metrics.



Historical Data Cutoff

Dataset coverage ends in mid-2021, meaning recent strategic shifts post-2021 are not reflected in this study.



Future Next Steps



Integrate Engagement Metrics

Incorporate third-party viewership estimates (e.g., Nielsen, FlixPatrol) to correlate catalog depth with actual consumption.



Add Financial Dimensions

Layer in estimated production costs and regional revenue data to transform descriptive analysis into ROI modeling.



Develop Predictive Models

Move from descriptive analytics to predictive machine learning models to forecast future content success probabilities.