

■ Netflix Dataset Analysis Report

1. Dataset Overview

- Total Records: 8807
- Total Columns: 12
- Columns: show_id, type, title, director, cast, country, date_added, release_year, rating, duration, listed_in, description

2. Data Quality Check

- Missing Values: director (29.9%), cast (9.3%), country (9.4%), date_added (0.1%), rating (0.04%), duration (0.03%)
- Duplicates: 0
- Missing values handled (director → 'Unknown', country → mode value)

3. Dataset Composition

- Movies: 6131
- TV Shows: 2676
- Movies dominate Netflix content (~70%).

4. Release Year Analysis

- Range: 1925 – 2021
- Most Releases: 2017–2019
- Content production rapidly grew in the last decade.

5. Country-Wise Content Distribution (Top 10)

1. United States – 3649
2. India – 972
3. United Kingdom – 419
4. Japan – 245
5. South Korea – 199
6. Canada – 181
7. Spain – 145
8. France – 124
9. Mexico – 110
10. Egypt – 106

6. Genres (Top 10)

Documentaries, Stand-Up Comedy, Dramas, International Movies, Comedies, Action & Adventure, Romantic Movies, Thrillers, Horror Movies, TV Dramas

7. Ratings Distribution

- Most Common Ratings: TV-MA, TV-14, TV-PG, R, PG-13
- Covers wide audience: Kids → Adults

8. Movie Duration Analysis

- Average duration: ~90 minutes
- Most movies: 80–120 minutes
- Few outliers (very short/long movies)

9. Trends Over Time

- Content added peaked around 2019
- TV Shows rose significantly after 2015

10. Correlation Insights

- release_year correlated with rating & type
- date_added strongly correlated with year_added

11. Key Takeaways

- Netflix is movie-heavy but TV Shows are rising
- US & India are biggest content producers
- Documentaries & Comedies dominate
- Content targets all age groups
- Most movies ~90 minutes