Business Requirements Document (BRD)

Project Title: Netflix Content Data Analysis

Prepared By: Abhishek Darbare
Date: 05/1/2026

---

1. Business Overview.

Netflix is one of the leading global video streaming platforms, offering a wide range of movies and TV shows across multiple genres, languages, and countries. With millions of subscribers worldwide, Netflix relies heavily on data-driven insights to optimise its content strategy, enhance user engagement, and expand its global presence. This project analyses the Netflix content dataset to uncover patterns in content distribution, growth trends, genre popularity, and regional contributions. The analysis aims to help stakeholders understand how Netflix's content library has evolved and how it aligns with audience preferences and market expansion strategies.

---

2. Objectives and Goals

The key objectives of this analysis are:

- To analyze the distribution of Movies and TV Shows on Netflix.
- To study how Netflix's content library has grown over the years.
- To identify the most dominant genres on the platform.
- To determine which countries contribute the most content.
- To analyze movie duration patterns and TV show season trends.
- To understand how fresh the content is when added to Netflix.
- To extract meaningful business insights that can support content strategy and regional expansion decisions.

---

3. Domain Knowledge (Retail Analytics)

Domain: Media Analytics / Streaming Platform Analytics

This domain focuses on:

- Media Analytics / Streaming Platform Analytics
- Optimising content acquisition strategies
- Improving user engagement and retention
- Supporting content recommendation systems
- Guiding regional and genre-based content investment

Media Analytics/Streaming Analytics is widely used by video streaming platforms, production studios, OTT platforms, digital content creators, media distribution companies, and entertainment networks to understand audience behaviour, optimise content strategies, and improve viewer engagement.

Key industry applications include:

- Content recommendation engines
- Subscriber growth analysis
- Market expansion planning
- Performance evaluation of original vs licensed content
- Genre and regional content optimization

---

4. Data Description

The dataset includes the following key attributes:

Column Name	Description
show_id	Unique ID for each title
type	Movie or TV Show
Title	Title of the content
Director	Director of the content
Cast	Main cast members
Country	Country of production
date_added	Date when content was added to Netflix
release_year	Year the content was released
rating	Content rating (TV-MA,PG,etc.)
duration	Movie duration in minutes or number of seasons for TV shows
listed_in	Genre(s)
description	Brief description of the content

5. Functional Requirements

a. Data Preprocessing

i. Load the Netflix dataset and inspect its structure.
ii. Check and handle missing values in relevant columns.
iii. Convert the date_added column into datetime format.
iv. Clean and standardize text columns like country, genre, and cast.
v. Extract numeric values from the duration column.
vi. Create new derived features such as year and month added.

b. Exploratory Data Analysis (EDA)

i. Content Type Distribution

1.Analyze the ratio of Movies vs TV Shows**.**

ii. Content Growth Over Time

1. Study how the number of titles added changes over the years.
2. Compare growth patterns of Movies and TV Shows.

iii. Genre Analysis

1. Identify the most popular genres on Netflix.
2. Understand audience preferences through genre dominance.

iv. Country-wise Analysis

1. Identify top content-producing countries.
2. Study regional contributions to Netflix's content library.

v. Duration and Season Analysis

1.Analyze typical movie duration ranges.
2. Analyze the distribution of TV show seasons.

vi. Content Freshness

1. Calculate how old the content is when it is added to Netflix.
2.Understand Netflix's focus on new vs old content.

---

6. Tools and Technologies

The analysis will be performed using the following tools:

- **Programming Language:** Python
- **Libraries:** Pandas, NumPy, Matplotlib, Seaborn
- **Development Environment:** Jupyter Notebook

7. Deliverables

The final deliverables for this project will include:

- A complete EDA report with visualizations and interpretations.
- A cleaned and feature-engineered Netflix dataset.Data-driven recommendations based on findings.
- Business insights and strategic recommendations based on analysis.
- A GitHub repository containing:
  • Jupyter Notebook
  • Dataset
  • README documentation