**Documentation for Netflix Content Analysis**

**1. Objective of the Analysis:** This report provides an in-depth analysis of Netflix’s content library, aiming to uncover patterns and trends in content distribution, growth, and geographical influence. By evaluating Netflix's vast library of movies and TV shows, we aim to answer critical questions about the platform's content strategy and user engagement.

**2. Data Overview:** The data used in this analysis covers content added to Netflix from 2008 to 2021. It includes information about:

* **Content Type:** Differentiating between Movies and TV Shows.
* **Content Categories:** Genres such as Dramas, Comedies, Documentaries, International Movies, and more.
* **Country of Production:** Identifying the origin countries of the content.
* **Content Added Date:** Month and year when the content was added.
* **TMDB Ratings:** Ratings sourced from TMDB for evaluating content quality.
* **Duration:** The runtime of movies and TV shows, categorized as short, medium, or long.

**3. Methodology:** The data was processed through various techniques such as:

* **Data Cleaning:** Identifying and handling missing or incomplete data (e.g., missing genres or unknown countries).
* **Aggregation:** Grouping data by content type, categories, release years, and countries.
* **Trend Analysis:** Using time series analysis to understand the volume of content added over the years and detecting seasonal patterns.
* **Geospatial Analysis:** Mapping the geographic distribution of Netflix content using tools like **Folium** for visualizing country-specific trends.

**4. Key Findings:**

* **Content Type Distribution:** There is a clear trend where **Movies** dominate over **TV Shows**, especially in recent years.
* **Most Popular Categories:** Genres like **International Movies**, **Dramas**, and **Comedies** top the list in terms of volume.
* **Geographic Distribution:** Countries such as the **United States**, **India**, and **United Kingdom** lead the content production, while countries in Europe and Asia also contribute significantly.
* **Trend Over Time:** There has been a dramatic increase in the number of new titles added to Netflix since 2016, with 2019 and 2020 showing the largest growth.

**5. Seasonal Patterns:**

* Content is added throughout the year, but certain months, such as **July**, **December**, and **September**, have seen a higher volume of new additions.
* Peak content additions are observed during the **holiday seasons** and **summer months**, suggesting strategic content planning around global events.

**6. TMDB Ratings Analysis:**

After integrating ratings from the TMDB API, we focused on the top 200 movies from TMDB and identified which of them are available on Netflix. The results show that Netflix's content generally consists of highly rated titles. Among these, Schindler's List stood out as the highest-rated movie.

The TMDB ratings reveal that Netflix includes several critically acclaimed titles in its library, particularly in the Action and Drama genres, with consistently high ratings, averaging around 8.0 or more.

**7. Limitations:**

* The dataset may contain missing or incomplete records, especially for genres or countries that are not always listed clearly.
* This analysis relies on data available until 2021 and may not reflect more recent additions to Netflix’s catalog.

**8. Conclusion:** This documentation presents an overall picture of Netflix's content strategy and provides insights into how its content library has evolved. By leveraging geographic, genre, and temporal data, we have a deeper understanding of Netflix’s global footprint and its content-driven engagement model.