

**Netflix Movies and TV Shows:**

**Visualizations, Recommendation, EDA**

**By**

Rawand Aziz

Trifa Khalid

Asan Tahsin

**Supervised by**

Dr. Miran Taha Abdullah

January, 2024

**Table of Contents………………………………………………………………….2**

**Introduction……………………………………………………………………….3**

**Problem Statement………………………………………………………………..4**

**Solution Method…………………………………………………………………..5**

**Implementation……………………………………………………………………6**

**Results Discussion………………………………………………………………..12**

**Project Conclusion……………………………………………………………….13**

**References………………………………………………………………………...14**

**Introduction**

This data set was created to list all shows available on Netflix streaming and analyze the data to find interesting facts. This data was acquired in July 2022 containing data available in the United States.

The ambitious work titled "Netflix Movies and TV Shows: Visualizations, Recommendation, EDA" embarks on an all-encompassing exploration and accurate analysis of the expansive dataset provided by Netflix. This project, which places a keen focus on visualizations, recommendation systems, and exploratory data analysis (EDA), is chosen strategically in light of the growing influence of streaming platforms in contemporary society. The uncountable potential benefits it promises for daily life, coupled with the exponential surge in content on platforms like Netflix, underscore the pressing need to understand user preferences, refine recommendation strategies, and extract profound insights from the vast content landscape.

As we navigate the problem of streaming platforms, the sheer diversity and availability of content present an intricate challenge for users in discovering material that resonates with their preferences. This project aims to address this challenge head-on, recognizing the profound impact it has on user satisfaction and engagement. Thus, a meticulous examination and enhancement of the recommendation system become not just an endeavor but a necessity, aligning with the evolving dynamics of user expectations and the ever-expanding content repertoire.

**Problem Statement**

The core challenge at the heart of this project lies in the intricacies of content recommendation on Netflix. The voluminous library and diverse user base contribute to the complexity of curating personalized recommendations effectively. This challenge is not merely a technical hurdle but holds profound significance as it directly influences the user experience. In an era where streaming platforms play a crucial role in entertainment consumption, the need for an optimized and efficient recommendation system is more critical than ever. The interplay between user preferences, content attributes, and the evolving content landscape creates a dynamic problem that demands a thoughtful and innovative solution.

**Solution Method**

Navigating the multifaceted landscape of content recommendations requires a methodical and sophisticated approach. The method employed in this project orchestrates a harmonious blend of data visualizations, recommendation algorithms, and exploratory data analysis.

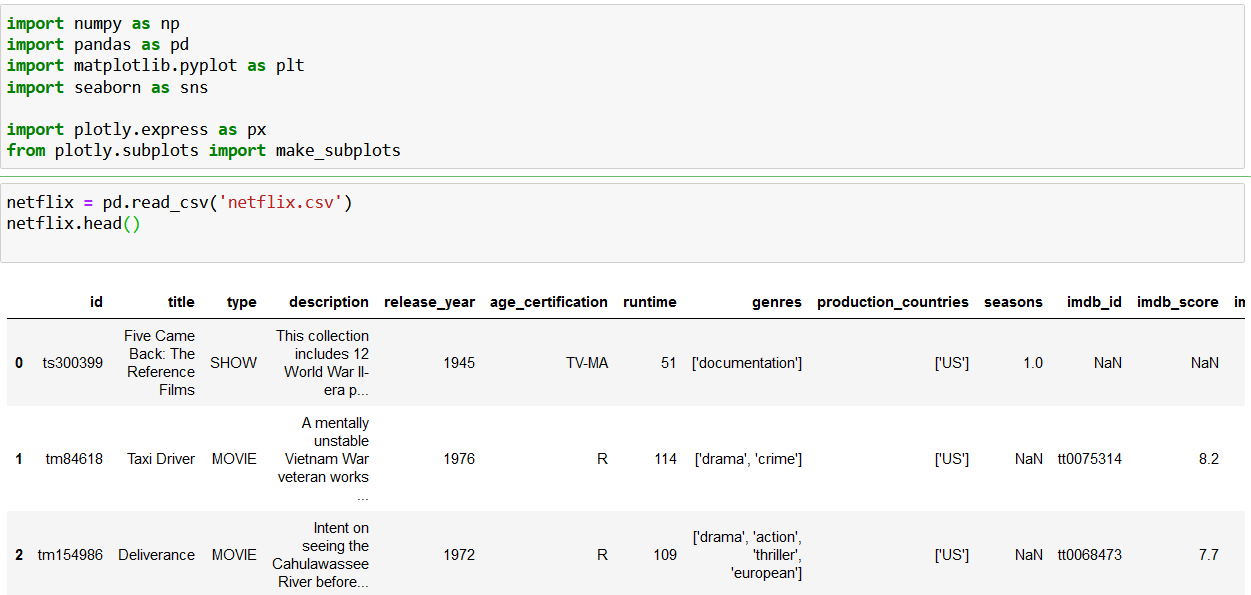
* Developing a content-based recommender system using the genres and/or descriptions.
* Identifying the main content available on the streaming.
* Network analysis on the cast of the titles.
* Exploratory data analysis to find interesting insights.

The design process unfolds systematically, commencing with the collection and preprocessing of the voluminous Netflix dataset. Subsequently, the project delves into the intricate world of data visualization, unraveling patterns and nuances in content distribution across genres. The crux of the methodology lies in the development of a recommendation system that intricately weaves together user preferences and content attributes. Leveraging advanced machine learning algorithms, the system adopts a dual-pronged strategy, seamlessly integrating collaborative filtering and content-based filtering to fortify the accuracy and relevance of recommendations.

**Implementation**

The tangible manifestation of the project comes to life through a meticulous implementation strategy, providing an immersive showcase of the user interface and operational functionality. This section transcends mere textual elucidation, opting for a visual feast that captures the essence of the project's intricacies. Screenshots and code snippets interplay to offer a comprehensive insight into the data preprocessing pipeline, the visual narrative crafted through various visualizations, and the intricacies of the recommendation generation process. A deliberate emphasis is placed on the incorporation of user feedback and preferences into the recommendation system, highlighting the inherently user-centric nature of the project.

**Objectives**:

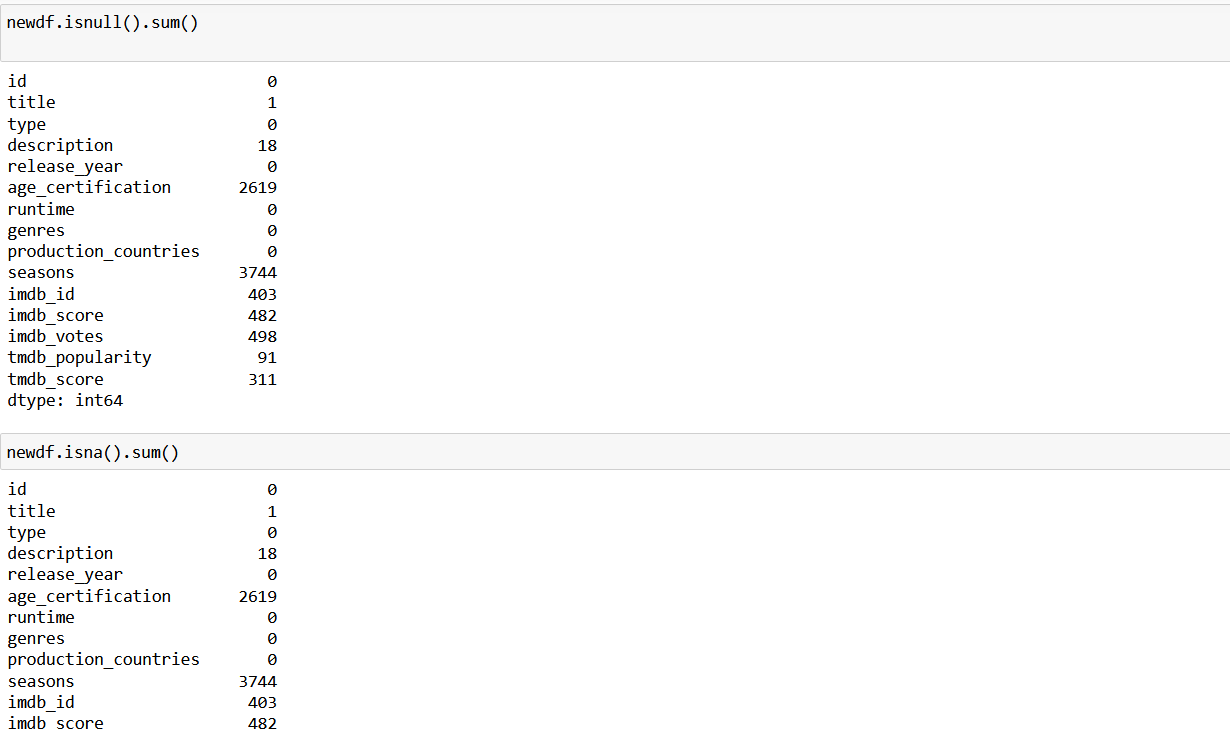
1. **Loading and exploring data**
2. **Cleaning Data:**

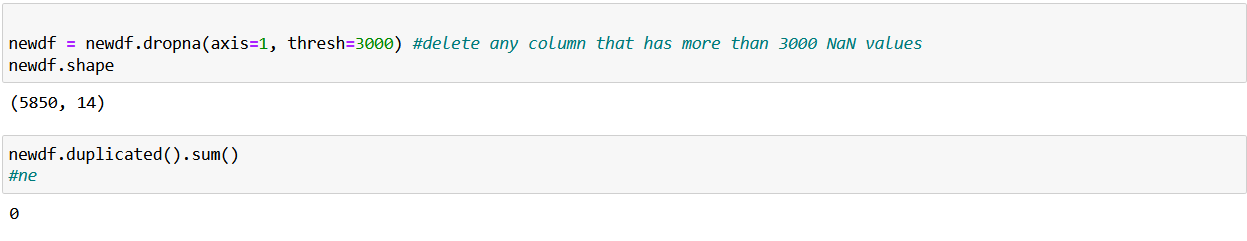
- Find Null/Nan values

- Find Duplicated values

- Fill Missing values

- Drop unnecessary columns with many missing values





1. **Visualization:**

Figure 1: Show Movie VS Show types

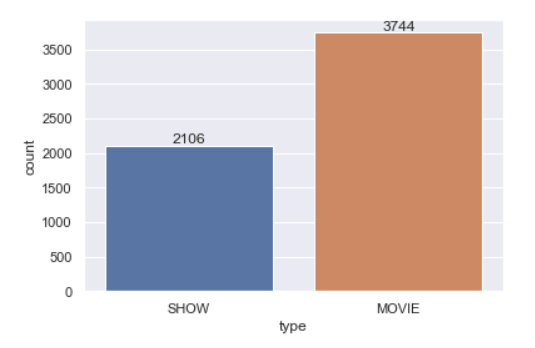


Figure 2: Show number of movies & shows in each year

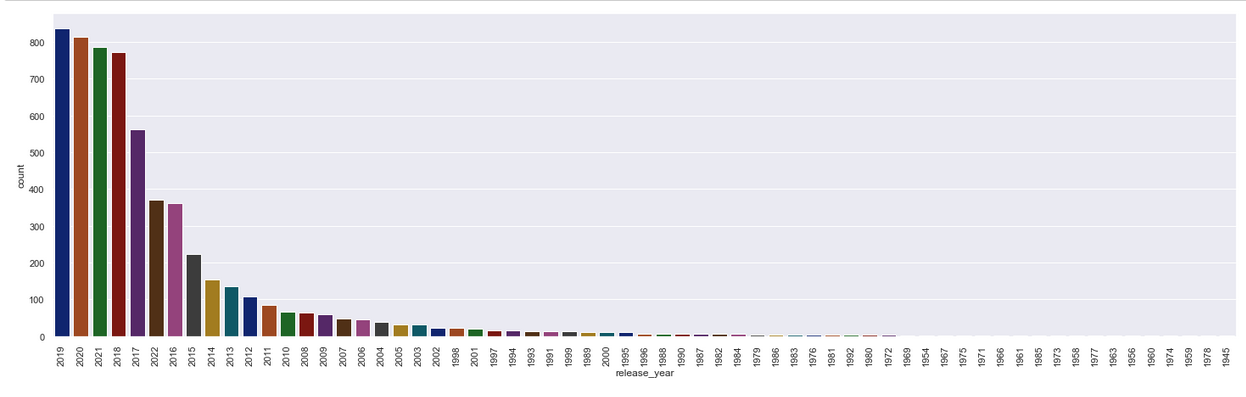


Figure 3: Showing number of movies/shows in top 10 country

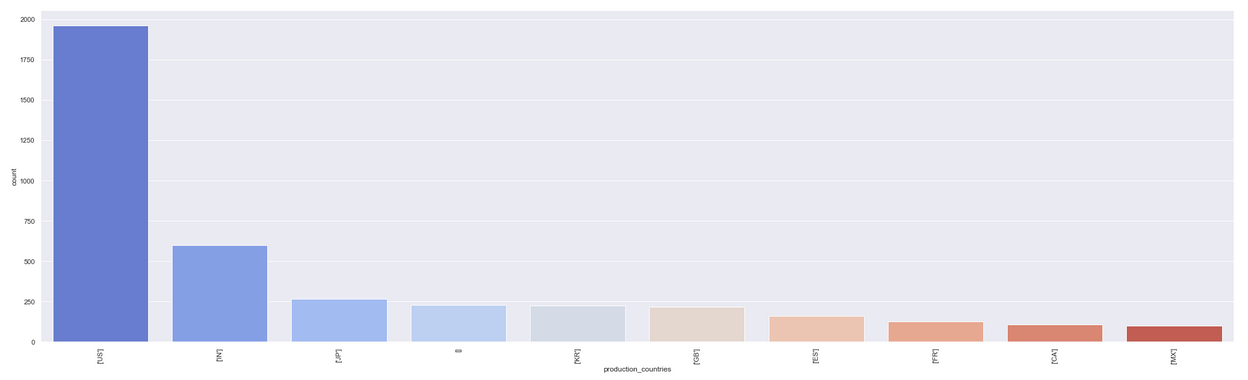


Figure 4: a histogram plot for MOVIES with drawing MEAN and MEDIAN

Mean: dashed red line Median: yellow line

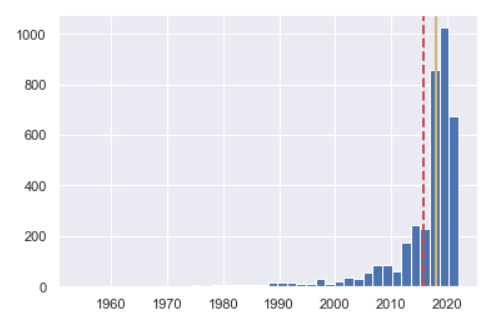


Figure 4: Score of each Movies/Shows based on their duration

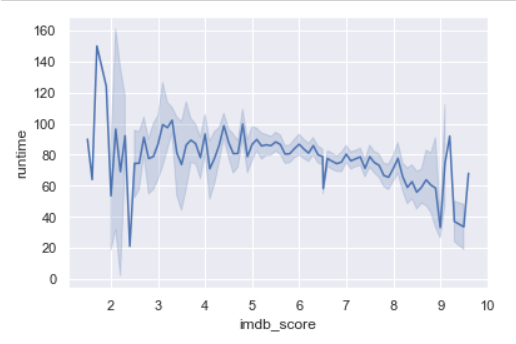


Figure 5: Showing IMDB score by release year of the movie/show

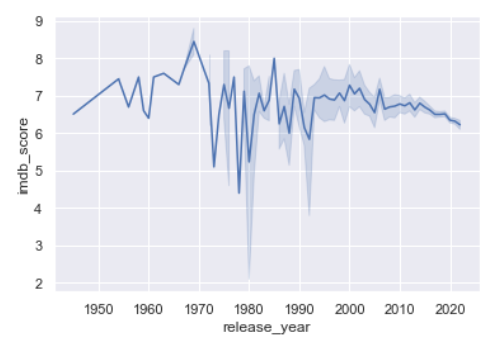
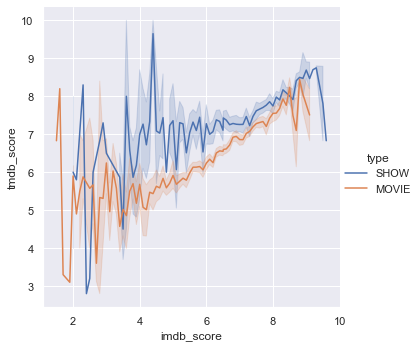


Figure 6: IMDB vs TMDB score comparison



**Results Discussion**

The crescendo of the project unfolds in the presentation and subsequent discussion of the results attained. This section is an artful amalgamation of data insights, strategically adorned with an array of graphs, charts, and visualizations. Each element is meticulously placed to illuminate the key findings extracted from the troves of data. The exploration delves into the distribution of genres, user ratings, and the ephemeral nature of content popularity. Yet, the discourse transcends mere numerical metrics, extending into the qualitative realm. The evaluation of the recommendation system's efficacy becomes a focal point, utilizing precision, recall, and user satisfaction metrics to gauge its impact on the user experience. Moreover, the insights harvested from the exploratory data analysis are dissected and laid bare, enriching the understanding of user behavior and content preferences on the Netflix platform.

Advantages of Netflix Movies and TV Shows Visualizations, Recommendation, EDA:

* Personalized Recommendations: Netflix uses advanced algorithms to provide highly personalized content suggestions, enhancing user satisfaction.
* Enhanced User Engagement: Visualizations and data analysis help Netflix understand user behavior, leading to improved platform engagement.
* Content Discovery: Data-driven recommendation systems promote diverse content discovery beyond mainstream options.
* Efficient Content Production: Netflix analyzes data to inform content production decisions, increasing the likelihood of creating popular shows and movies.
* Optimized User Interface: Visualizations and analytics inform Netflix's user interface optimization, creating an intuitive and user-friendly experience.
* Dynamic Content Library: Netflix's recommendation algorithms adapt to changing viewer preferences, ensuring a dynamic and relevant content library.

**Project Conclusion**

In the tapestry of conclusion, this project stands as a testament to accomplished milestones in unraveling the intricacies associated with content recommendations on Netflix. The robust implementation of an advanced recommendation system, coupled with visually compelling data insights, contributes significantly to an elevated and personalized user experience. As we stand at the crossroads of achievement, this project is not merely a culmination but a launchpad for future endeavors. The horizon beckons with prospects for refining the existing model, incorporating real-time user feedback mechanisms, exploring the frontiers of advanced machine learning models, and extending the analysis to encompass a broader spectrum of streaming platforms. This project, therefore, not only serves as a milestone but as a catalyst for future endeavors aimed at making entertainment consumption more than just personalized, but immersive, interactive, and utterly enjoyable for users across the global digital landscape.

# **References:**

# <https://www.w3schools.com/>

# <https://www.kaggle.com/>