Netflix Movie Data Analysis - EDA Insights

This document summarizes the key insights and data preprocessing steps derived from the Exploratory Data Analysis (EDA) performed on the Netflix movie dataset. The focus was on cleaning the data, transforming variables, and understanding movie trends.

🔧 Data Preprocessing Highlights

• Date Columns: Converted from object type to datetime format for accurate temporal analysis.  
• Duplicates: Verified that there were no duplicate entries in the dataset.  
• Irrelevant Columns Dropped: Removed columns such as 'Overview', 'Language', and 'Poster URL' for cleaner data.  
• Vote\_Average Categorization: Transformed into categories:

Popular

Non-popular

Average

Below Average

• Genre Column Cleaning:

Whitespace removed

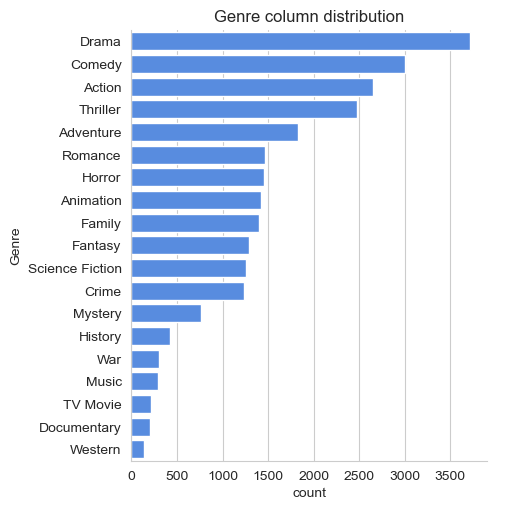
Genres split into list format for multi-label classification

Converted into categorical variable

🎯 Key Insights

* **🎬 Most Frequent Genre**

**Answer:** Drama



* **🌟 Movie with Highest Popularity**

**Movie:** *Spider-Man: No Way Home*

**Genre:** Action, Adventure

* **📉 Movie with Lowest Popularity**

**Movie:** *The United States vs. Billie Holiday*

**Genre:** Music, Drama, History

📆 Year with Most Movie Releases

**Year:** *2020*

Year with Most films :

