Netflix Data Exploration & Insights

A Comprehensive Analysis of Movies and TV shows

By – Rishav Sinha

Project Overview

Objective: Explore Netflix's catalog to uncover key insights related to genre of movies most releases, ratings received by the movie or TV show, year-wise number of movies released, popularity etc.

Approach: Data Cleaning, Data manipulation, and Visualization

Key Focus Area: Content distribution by Genre, Ratings and Year.
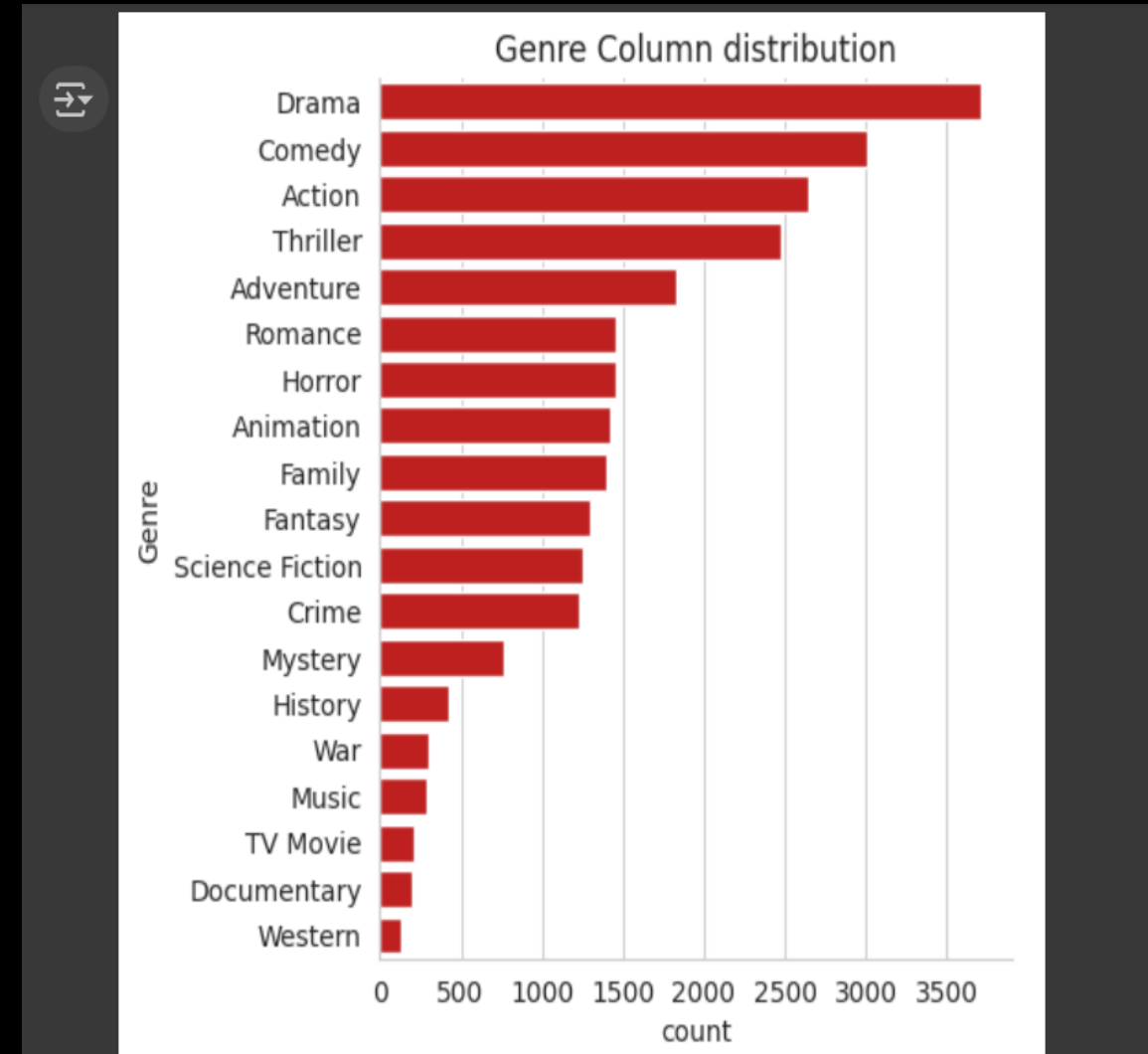
Dataset Information

- **Dataset**: Netflix Movie and TV shows.

- **Entries** : 9826 rows , 9 columns.

- **Key Columns**: Release date, Title, Popularity, Vote Count, Vote_Average, Genre.

- **Missing Data Handling** : No missing or null values in the dataset.

Data Cleaning & Transformation

- **Conversions** : Converted 'Release_date' column to datetime format and then drop day and time to keep year only.

- **Dropping** : Dropped unwanted columns like 'Overview', 'Original_Language', 'Poster_URL'.

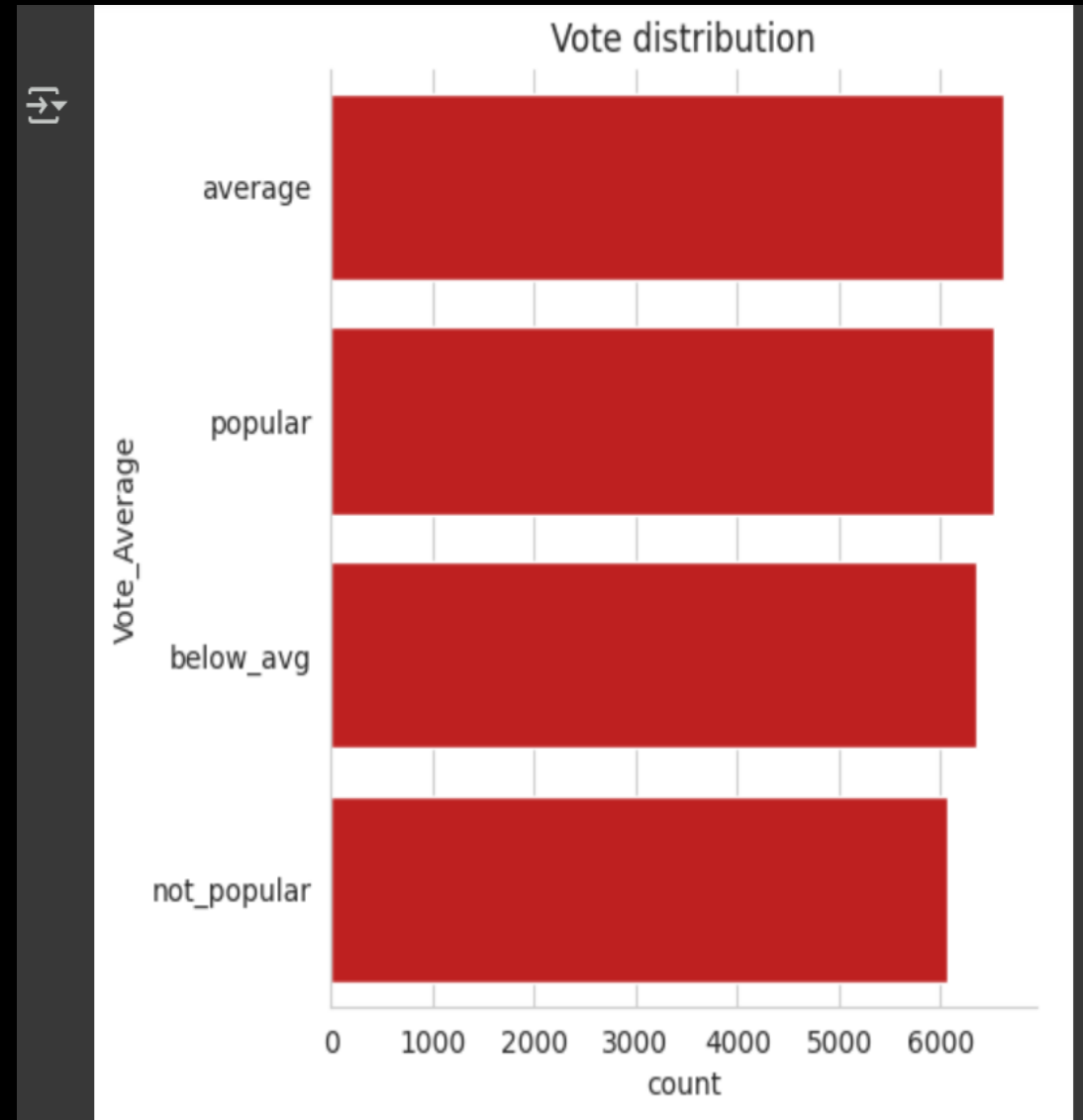- **Manipulation** : Exploded multiple genre into single row

Genre Distribution

- **Visualization** : Bar chart showing Genre wise movie counts from Highest to Lowest.

- **Insights** : Drama is the most frequent genre (~14%). Comedy, Action & Thriller are also popular.

Vote Distribution

- **Visualization** : Bar chart showing vote distribution category-wise.

- **Insights** : ~25.5% movies are in popular category. Drama dominates with ~18.5% of the popular movies.
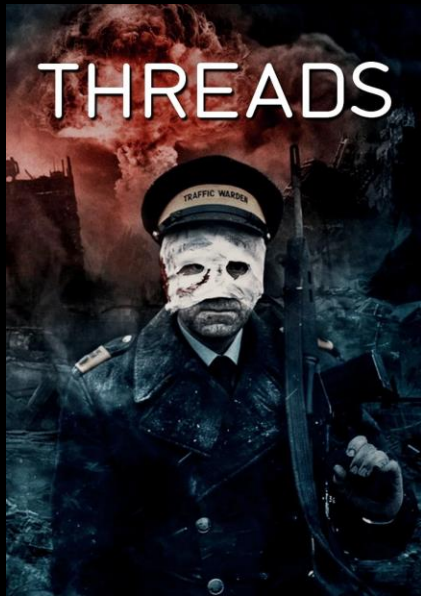
Highest Popularity Movie

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

- **Genres** : **Action, Adventure, Sci-Fi.**

- **Most popular movie in the dataset.**
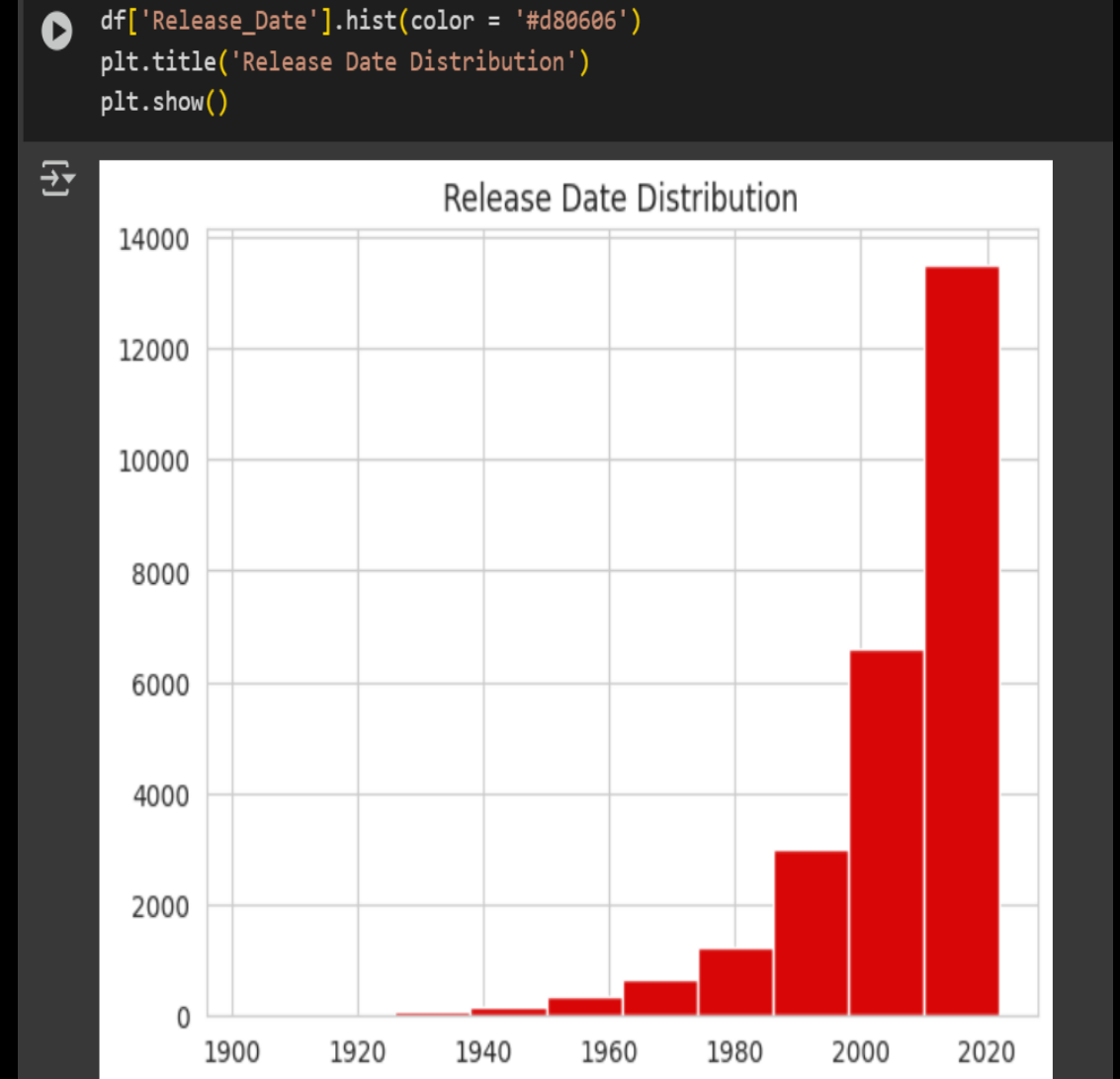
Lowest Popularity Movies

- **Movie** : The United States VS. Billie Holiday.

- **Genres** : Movie, Drama, History.

THE UNITED STATES VS. BILLIE HOLIDAY

THREADS

- **Movie** : Threads.

- **Genres** : Drama, Sci-Fi, War.

Yearly Release Trend

- **Visualization** : Bar chart showing year – wise number of movies released.

- **Insights** : Peak releases in 2020. Netflix rapidly increased it's movie release in recent years.

```
df['Release_Date'].hist(color = '#d80606')
plt.title('Release Date Distribution')
plt.show()
```

Key Insights

- **Drama** is Netflix's most frequent genre.

- **Spider-Man No Way Home** is most popular.

- **2020** saw the most releases.

- **Large portion** of the content is well received (Popular Category)

Conclusion

- The analysis highlights that Drama is the backbone of Netflix's content library, consistently being the most dominant genre across categories.

- Action and Adventure genres also play a key role in driving popularity, with big franchises like *Spider-Man: No Way Home* attracting massive audiences.

- The trend analysis shows that Netflix significantly ramped up production in 2020, possibly due to the pandemic-driven demand for online entertainment.

- A large portion of Netflix movies fall under the "Popular" and "Average" vote categories, showing that the platform maintains a balance of critically well-received and mainstream content.

- Movies with multiple genres (Drama + Thriller, Drama + Comedy) tend to appear more frequently, suggesting Netflix focuses on genre diversity to appeal to wider audiences.

- Overall, the findings reflect Netflix's strategy of blending quality storytelling (Drama) with commercial blockbusters (Action/Adventure) to dominate global streaming.

THANK YOU.