

Netflix Movies Dataset Analysis

Dataset Overview

- Dataset Size: The dataset contains 9,827 entries with 9 columns initially.
- Columns: The columns include `Release_Date`, `Title`, `Overview`, `Popularity`, `Vote_Count`, `Vote_Average`, `Original_Language`, `Genre`, and `Poster_Url`.
- Data Cleaning:
 - Converted `Release_Date` to the year format.
 - Dropped unnecessary columns: `Overview`, `Original_Language`, and `Poster_Url`.
 - Checked for duplicates: No duplicates were found.

Genre Distribution

- Code Steps:
 - Split the `Genre` column into individual genres.
 - Exploded the list of genres into separate rows.
 - Converted `Genre` to a categorical data type.
 - Created a count plot to visualize the distribution of genres.
- **Visualization:** A horizontal bar chart showing the count of movies per genre.
- **Insights:**
 - The most common genres are Drama, Comedy, and Action.
 - Thriller and Adventure genres also have significant representation.
 - This indicates a strong viewer preference for these genres.

Votes Distribution

- Code Steps:
 - Categorized the `Vote_Average` column into four categories: 'Not Popular', 'Below Avg', 'Average', and 'Popular'.
 - Created a count plot to visualize the distribution of these categories.
- **Visualization:** A horizontal bar chart showing the count of movies in each vote category.

- Insights:

- Most movies fall into the 'Average' category.
- Fewer movies are categorized as 'Popular' or 'Not Popular'.
- This suggests a balanced distribution of movie ratings.

Movies Released Per Year

- Code Steps:

- Created a histogram to show the number of movies released each year.

- Visualization: A histogram with years on the x-axis and the number of movies on the y-axis.

- Insights:

- The year 2021 had the highest number of movie releases.
- This peak could be due to increased production or availability of movies in that year.

Most and Least Popular Movies

- Code Steps:

- Identified the most and least popular movies based on the `Popularity` column.

- Insights:

- Most Popular Movie: "Spider-Man: No Way Home" (Action, Adventure, Science Fiction).
- Least Popular Movies: "The United States vs. Billie Holiday" (Music, Drama, History) and "Threads" (War, Drama, Science Fiction).

Top 10 Most Popular Movies

- Code Steps:

- Sorted the dataset by `Popularity` in descending order.
- Extracted the top 10 movies.

- Insights:

- The top 10 most popular movies include:

1. Spider-Man: No Way Home
2. The Batman
3. No Exit
4. Encanto

- These movies span various genres, indicating that popularity is not confined to a single genre.

Vote Average Distribution

- Code Steps:

- Created a box plot to show the distribution of vote averages.

- **Visualization:** A box plot showing the spread of `Vote_Average` values.

- Insights:

- The median vote average is around the 'Average' category.
- There are some outliers indicating exceptionally high or low ratings.

Popularity Trend Over Years

- Code Steps:

- Grouped the data by `Release_Date` and calculated the mean `Popularity` for each year.
- Created a line plot to visualize the trend.

- **Visualization:** A line plot with years on the x-axis and average popularity on the y-axis.

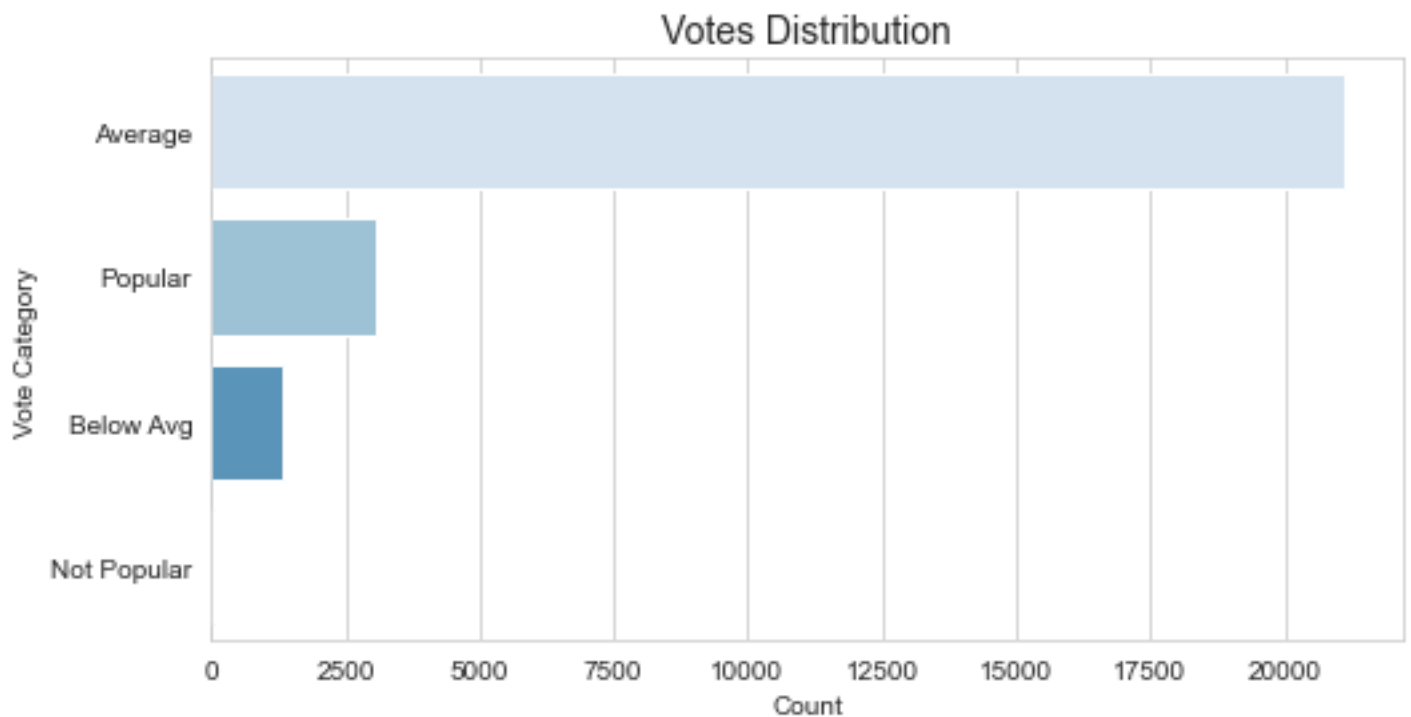
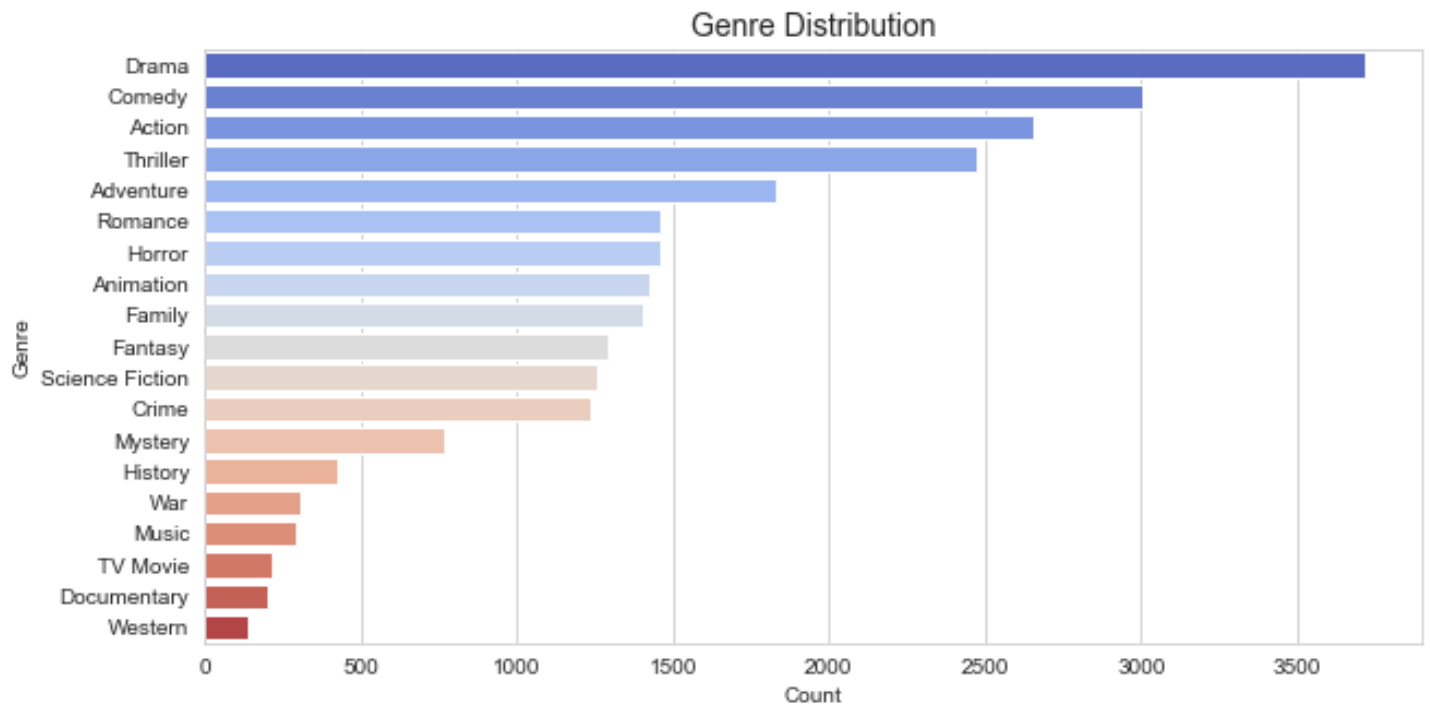
- Insights:

- The popularity trend shows fluctuations over the years.
- Certain years, such as 2021, show higher average popularity, likely due to the release of blockbuster movies.

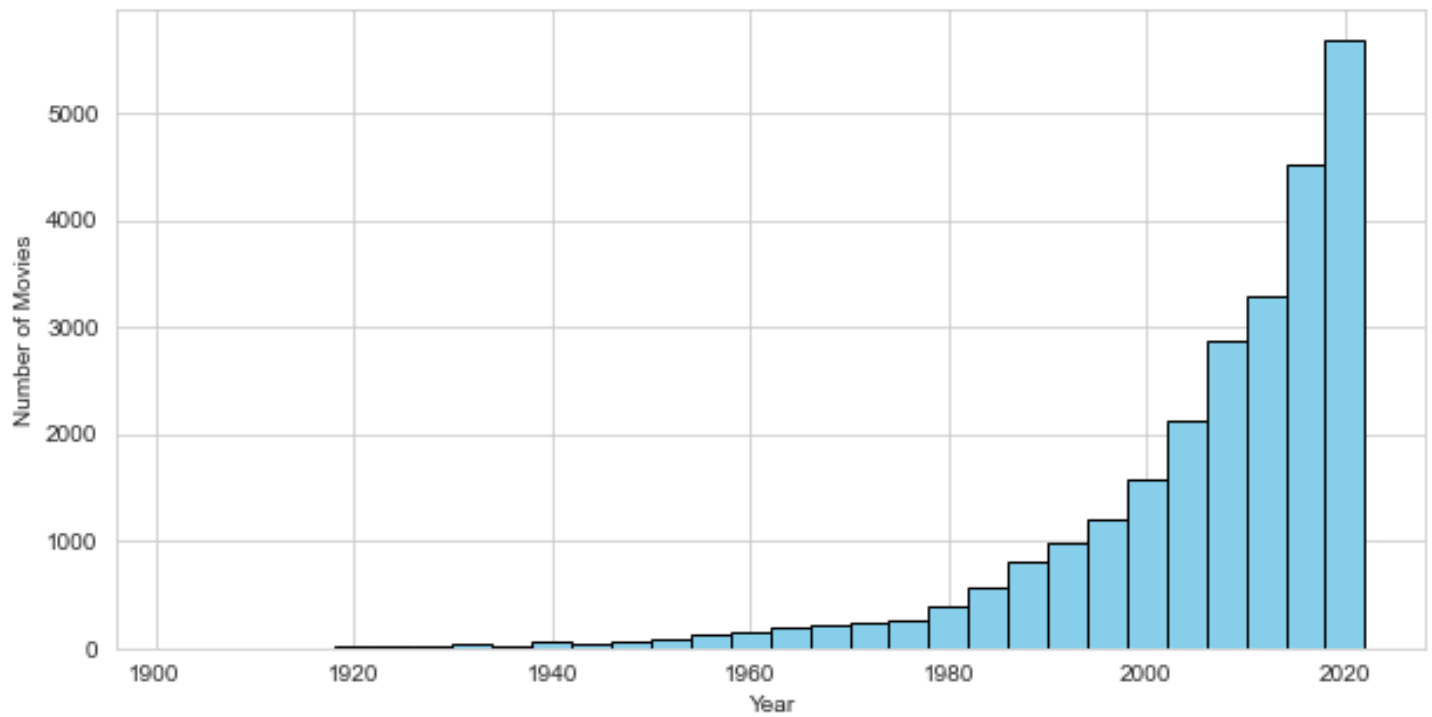
Conclusion

- Genre Preferences: Drama, Comedy, and Action are the most prevalent genres, reflecting viewer preferences.
- Popularity and Ratings: While most movies have average ratings, a few stand out as highly popular or unpopular.

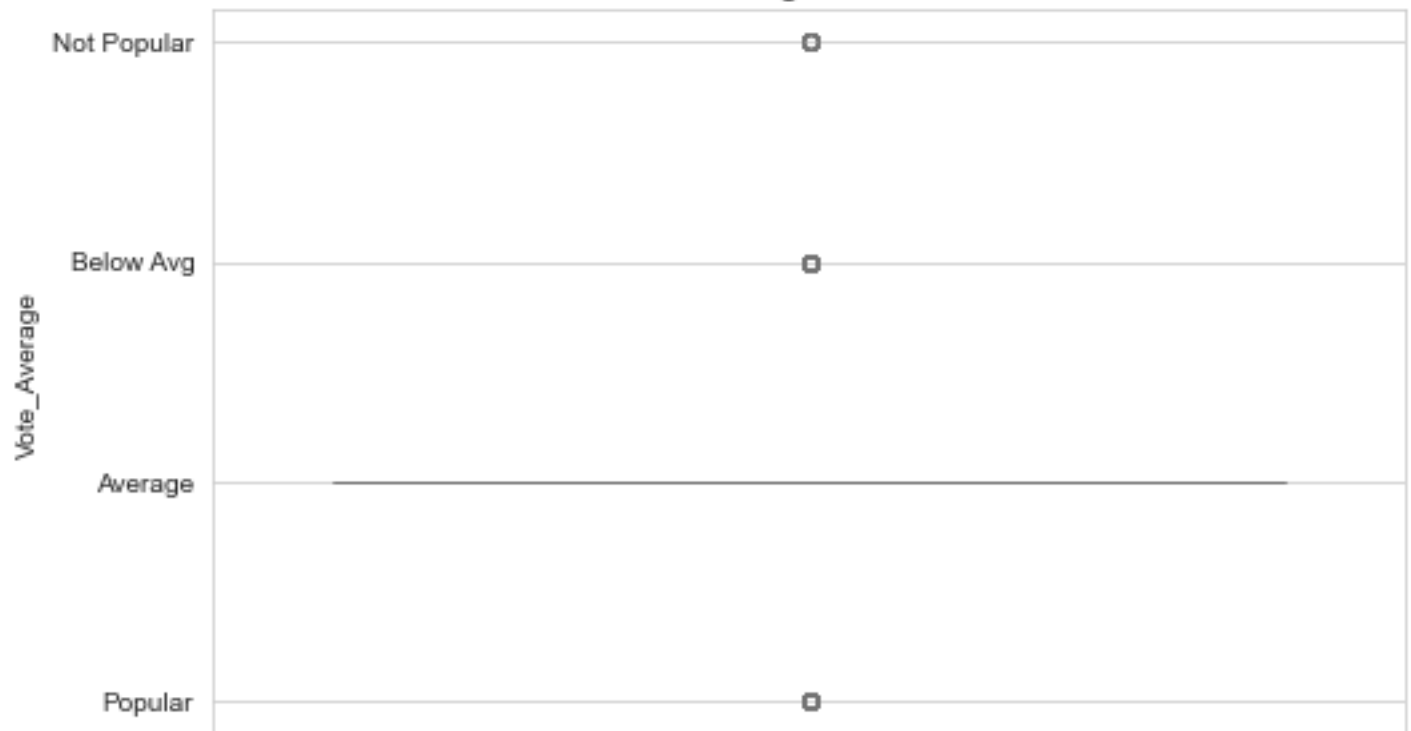
- Production Trends: The year 2021 marked a peak in movie releases, which could be due to various factors such as market demand or production capabilities.
- Trends Over Time: The popularity of movies has varied over the years, with certain years showing higher average popularity, likely due to the release of highly anticipated films.



Movies Released Per Year



Vote Average Distribution



Popularity Trend Over Years

