Age Ratings and Critics' Scores: A Streaming Platforms Analysis

NAME: SHARVARI KODGULE

TABLE OF CONTENTS

INTRODUCTION

The proliferation of streaming platforms has revolutionized the way audiences consume movies, offering a diverse range of genres and styles tailored to various demographics. Among these platforms, Disney+ and Netflix stand out as major players, each with distinct brand identities and audience bases. Disney+ has established a reputation for catering primarily to family-friendly content, enriched by its vast collection of Disney classics, Pixar animations, and franchises like Marvel and Star Wars. Conversely, Netflix is widely regarded as a hub for a broader and more mature audience, featuring a mix of original productions, international titles, and genres that span across thrillers, documentaries, dramas, and more.

This report explores whether these perceptions align with the actual content available on these platforms. Specifically, the analysis addresses two key questions:

1. Is the age restriction for movies on Disney+ lower than for movies on Netflix?
2. Are the movies on Netflix generally rated higher on Rotten Tomatoes than those on Disney+?

To answer these questions, a dataset containing movie information across four platforms Netflix, Hulu, Prime Video, and Disney+—is analyzed. The dataset includes details such as movie titles, release years, age restrictions, Rotten Tomatoes scores, and platform availability. This report will employ descriptive and inferential statistical methods to explore and test these hypotheses, providing data-driven insights into the differences between Disney+ and Netflix.

Through this analysis, we aim to challenge preconceived notions about these platforms and offer a nuanced understanding of their movie offerings. The findings could have implications for audience segmentation, content strategy, and platform positioning in the competitive streaming landscape.

DATA DESCRIPTION

The dataset provides information on movies available across four major streaming platforms—Netflix, Hulu, Prime Video, and Disney+. The dataset contains 11 columns with 9,515 entries. Here's a brief description of the columns:

It contains the following columns:

- ID: Unique identifier for each movie.
- Title: The name of the movie.
- Year: The release year of the movie.
- Age: Age restriction for the movie, such as "7+" or "18+".
- Rotten Tomatoes: Critics' score as a percentage, presented as a string (e.g., "98/100").
- Netflix, Hulu, Prime Video, Disney+: Binary indicators (1 for available, 0 for not available) showing the platforms where the movie is accessible.
- Type: Indicator for content type, which is focused on movies in this dataset.

A preliminary inspection reveals missing values in the Age and Rotten Tomatoes columns. These gaps could impact the analysis and will be addressed during preprocessing.

I. Approach:

1. Data Preprocessing:

- Dropping the ‘unnamed’ column as it serves no use.
- Addressing missing data in Age and Rotten Tomatoes using appropriate imputation or exclusion strategies to minimize bias.

2. Descriptive Analysis

- Summarizing the distribution of movies by age restrictions across Disney+ and Netflix to understand platform-specific audience targeting.
- Calculating statistical measures (mean, median, and standard deviation) for Rotten Tomatoes scores to assess content quality.
- Visualizing data through histograms, bar charts, and box plots for a comprehensive understanding.

3. **Inferential Analysis**

   - Performing hypothesis testing to answer the following:
     a. Are age restrictions for movies on Disney+ lower than those on Netflix?
     b. Is there a significant difference in Rotten Tomatoes scores between movies on Disney+ and Netflix?
   - Using non-parametric tests if the data distribution does not meet the assumptions for parametric tests.

II. Tools Used:

1. **Python Libraries:**

   - Pandas for data cleaning and manipulation.
   - Matplotlib and Seaborn for creating insightful visualizations.
   - SciPy and Statsmodels for conducting statistical hypothesis tests.

2. **Visualization Techniques:**

   - Histograms and box plots to compare distributions.
   - Bar charts to illustrate platform-specific content trends.

3. **Statistical Tests:**

   - Mann-Whitney U test or independent t-test for age restriction comparison.
   - Independent t-test for analyzing Rotten Tomatoes score differences if data assumptions are met.

By following these approaches, the report aims to draw meaningful insights about platform-specific audience focus and content quality.

METHODOLOGY

1. Distribution Plots:

Distribution plots are essential for visually examining the distribution of sample data by comparing the empirical data with theoretical values expected from a specific distribution. In Python, the distplot() function from the Seaborn library provides various options for visualizing univariate or bivariate data distributions. These plots help in understanding the underlying structure of the data and identifying any deviations from expected patterns.

2. Outliers:

An outlier is an observation that significantly deviates from the other data points in the dataset, either being much larger or smaller. Outliers can arise due to natural variability in the data or errors during data collection or recording. Analyzing these anomalous points, often called outlier analysis, is crucial for identifying and addressing potential issues. Visualization methods, such as box plots, are commonly used to detect outliers, while imputation or exclusion methods are used for handling them.

3. Hypothesis Testing:

A hypothesis is an assumption or idea, specifically a statistical claim about an unknown population parameter. For example, a judge assumes a person is innocent and verifies this by reviewing evidence and hearing testimony before reaching a verdict.

Hypothesis testing is a statistical method that is used to make a statistical decision using experimental data. Hypothesis testing is basically an assumption that we make about a population parameter. It evaluates two mutually exclusive statements about a population to determine which statement is best supported by the sample data.

Key Parameters of Hypothesis Testing:

1. **H_0(Null hypothesis):**
   This represents the default assumption, stating that there is no effect or relationship between the groups being analyzed. For example, H_0 might claim that the production of a company is equal to 50 units per day.

2. **H_a(Alternative hypothesis):**
   The alternative hypothesis challenges the null hypothesis, suggesting a deviation. For instance, H_1 could assert that the production is not equal to 50 units per day.

3. **Types of Tests:**
   - **One-tailed Test:** Evaluates the data for a specific direction (e.g., greater than or less than).
   - **Two-tailed Test:** Examines the possibility of deviation in either direction, focusing on whether variances or means differ without specifying the direction.

4. **Level of Significance (α):**
   This indicates the threshold for accepting or rejecting H_0, often set at 0.05 (5%), meaning the results should be 95% reliable.

5. **p-value:**
   The p-value represents the probability of observing the results under the assumption that H_0 is true. If the p-value is below the significance level (e.g., $p < 0.05$), the null hypothesis is rejected in favor of the alternative.

Shapiro-Wilk Test

The Shapiro-Wilk test, which is denoted by W, is a test of whether a data set has been drawn from an underlying normal distribution. The null and alternative hypotheses being tested are

H_0: The data have been drawn from a normal distribution

H_a: The data are drawn from a non-normal distribution

The W statistic is calculated as follows:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Where,

- x_i are the ordered sample values ($x_{(1)}$ is the smallest)
- a_i are coefficients calculated based on the means, variances, and covariances of order statistics from a normal distribution.

- $\bar{x}$ is the mean of samples.

Applications:

This test is effective for small sample sizes and is widely regarded for its sensitivity in detecting deviations from normality.

Decision Criteria :

If p<α (e.g., 0.05), reject H_0 , indicating the data does not follow a normal distribution.

Levene's Test

The Levene's Test, used to assess the equality of variances across groups, is a statistical test for homogeneity of variances. It tests the null hypothesis that the variances in each group are equal. The null and alternative hypotheses being tested are:

H_0: The variances are equal across the groups.

H_a: At least one group's variance is different from the others.

The Levene statistic is calculated as follows:

$$W = \frac{(N-k)}{(k-1)} \cdot \frac{\sum_{i=1}^{k} n_i (\bar{Z}_\iota - \bar{Z})^2}{\sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Z_{ij} - \bar{Z}_\iota\right)^2}$$

where:

- k is the number of groups,
- N is the total number of observations,
- n_i is the number of observations in group i,
- $\bar{Z}_\iota$ is the mean of the absolute deviations from the central measure of group i,
- $\bar{Z}$ is the overall mean of the absolute deviations.

Applications:

Levene's test is useful in comparing the variances of multiple groups, especially before running an analysis of variance (ANOVA), which assumes equal variances.

Decision Criteria:

- If p<α (commonly α=0.05), reject H_0, indicating that the variances are significantly different across groups.
- If p>αp, fail to reject H_0, suggesting that the variances are not significantly different.

Mann-Whitney U Test

The Mann-Whitney U test, also known as the Wilcoxon rank-sum test, is a non-parametric test used to determine whether there is a difference between two independent groups on a continuous or ordinal variable. It is particularly useful when the assumptions of a t-test (such as normality of the data) are not met. The null and alternative hypotheses being tested are:

H_0: The distributions of the two groups are identical (no difference).

H_a: The distributions of the two groups are different.

The test statistic, U, is calculated as follows:

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - \sum R_1$$

where:

- n1 and n2 are the sample sizes of the two groups,
- R1 is the sum of the ranks for the group with sample size n1,
- The ranks are assigned to all the data points from both groups combined, and each data point is given a rank based on its value.

Applications:

The Mann-Whitney U test is commonly used in situations where the data are not normally distributed or when sample sizes are small. It is widely applied in various fields, including medical research, psychology, and social sciences, to compare two independent groups.

Decision Criteria:

- If p<α (commonly α=0.05), H_0 reject , indicating a significant difference between the two groups.
- If p>α, fail to reject , H_0 indicating no significant difference between the two groups.

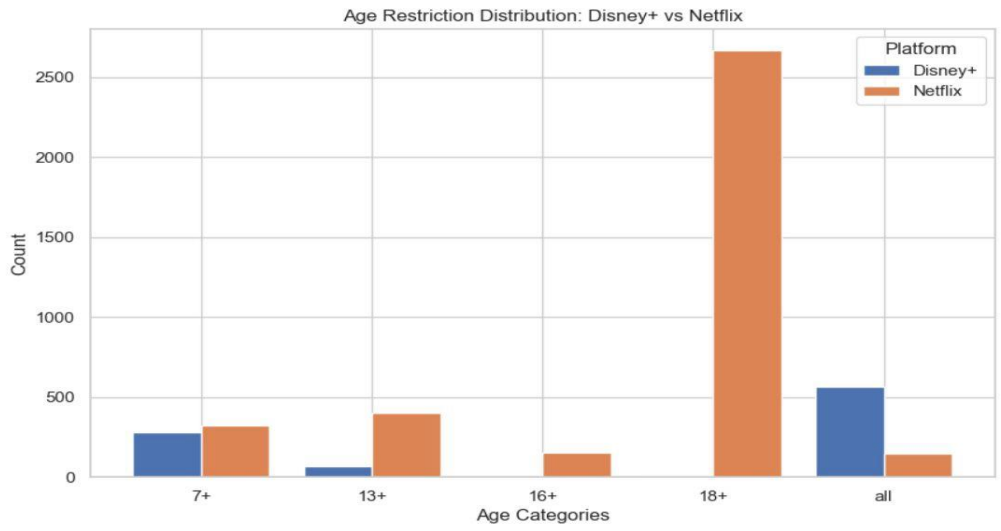
EVALUATION

This section evaluates the analysis by describing the application of descriptive and inferential statistical methods. The evaluation includes the rationale for the tests performed, their results, and the interpretation of findings.

I. Descriptive Analysis:

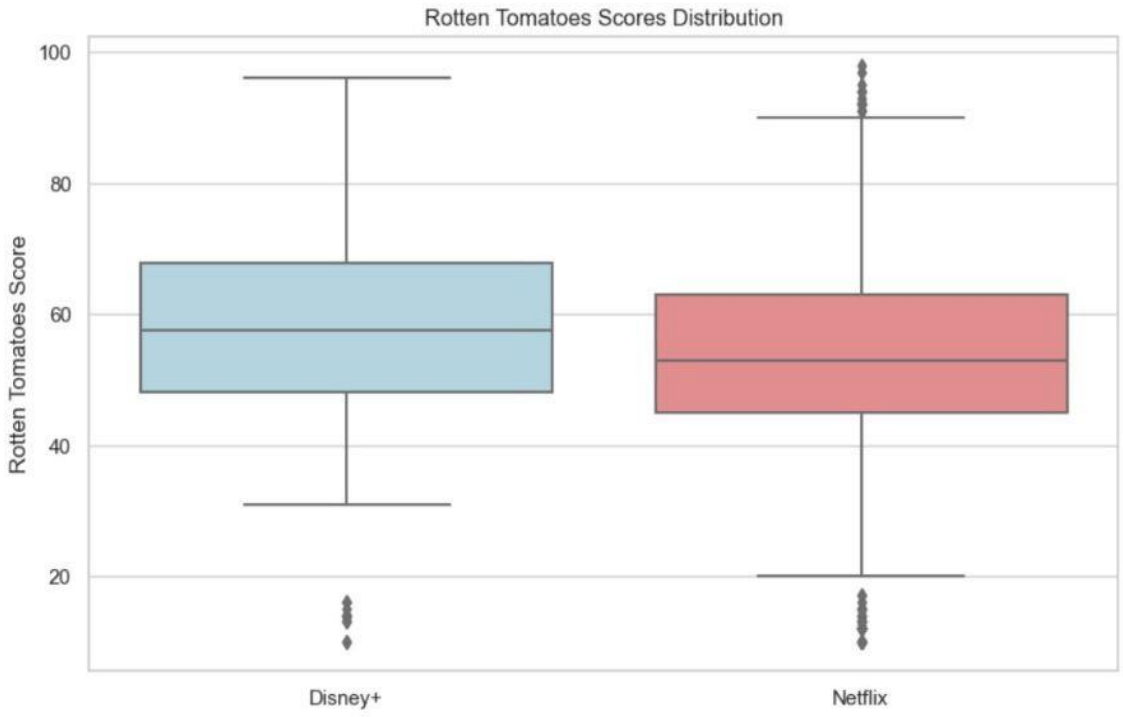
1. Age Restrictions:

- The distribution of age restrictions for movies on Netflix and Disney+ was visualized using a bar chart.
- Summary statistics, including the mean, median, mode, and frequency counts, provided insights into the platforms' audience focus.
- Findings:
  - Disney+ predominantly featured movies rated "7+" and "All," indicating a strong focus on family-friendly content.
  - Netflix, on the other hand, included a significant proportion of movies rated "18+," reflecting a broader audience demographic that includes adult viewers.
  - The mean age restriction for Disney+ was lower (1.2) compared to Netflix (2.8), further supporting the family-friendly reputation of Disney+.



2. Rotten Tomatoes Scores:

- Box plots were used to compare the Rotten Tomatoes score distributions between Netflix and Disney+.
- Descriptive statistics, including the mean, median, standard deviation, and interquartile range (IQR), summarized the central tendency and variability.
- Findings:
  - Disney+ had a slightly higher median score (57.5) than Netflix (53).
  - The interquartile range (IQR) for Disney+ (48–67.75) was narrower than that of Netflix (45–63), indicating less variability in movie quality on Disney+.
  - While the mean scores were close, Disney+ exhibited fewer low outliers compared to Netflix.



II. Inferential Analysis:

1. Normality Check (Shapiro-Wilk Test):

- Before selecting statistical tests, a Shapiro-Wilk test was conducted to assess whether the data followed a normal distribution.
- **Hypotheses:**
  - H_0: The data is normally distributed.
  - H_a: The data is not normally distributed.
- **Results:**
  - For both the Age and Rotten Tomatoes columns, the p-values were significantly lower than 0.05, indicating that the data was not normally distributed.
- **Rationale for Using Shapiro-Wilk Test:**
  - The Shapiro-Wilk test is a robust method for assessing normality in small to medium-sized datasets, ensuring that appropriate parametric or non-parametric tests are chosen.

2. Variance Equality Check (Levene's Test):

- Levene's test was used to determine whether the variances of age restrictions and Rotten Tomatoes scores were equal across the two platforms.
- **Hypotheses:**
  - H_0: Variances are equal.
  - H_a: Variances are not equal.
- **Results:**
  - Age restrictions: The p-value was less than 0.05, indicating unequal variances.
  - Rotten Tomatoes scores: The p-value was greater than 0.05, indicating equal variances.
- **Rationale for Using Levene's Test:**
  - This test is widely used to validate assumptions about variance, which is critical for choosing the appropriate statistical test.

3. **Age Restrictions (Mann-Whitney U Test):**

- The Mann-Whitney U test was used to compare age restrictions between Netflix and Disney+.
- **Hypotheses:**
  - H_0: The age restrictions for movies on Disney+ and Netflix are the same.
  - H_a: The age restrictions for movies on Disney+ are lower than those on Netflix.
- **Results:**
  - U-statistic: 2,137,331.0.
  - p-value: 1.0.
  - Since the p-value was greater than 0.05, we failed to reject H_0.
- **Rationale for Using Mann-Whitney U Test:**
  - As the data violated normality assumptions (confirmed by the Shapiro-Wilk test), the Mann-Whitney U test was the most suitable non-parametric alternative to the t-test for comparing distributions.

4. **Rotten Tomatoes Scores (Mann-Whitney U Test):**

- The Mann-Whitney U test was also used to compare Rotten Tomatoes scores between Disney+ and Netflix.
- **Hypotheses:**
  - H_0: Rotten Tomatoes scores for Disney+ and Netflix are the same.
  - H_a: Rotten Tomatoes scores differ between Disney+ and Netflix.
- **Results:**
  - U-statistic: 1,984,865.0.
  - p-value:3.32×10^{-15}.
  - Since the p-value was less than 0.05, H_0 was rejected, indicating a significant difference in scores.
- **Rationale for Using Mann-Whitney U Test:**
  - Despite equal variances (as per Levene's test), the non-normal distribution of scores necessitated using the Mann-Whitney U test instead of the independent t-test.

III. Interpretation of Results:

1. Age Restrictions:

- The hypothesis test revealed no statistically significant difference in age restrictions between Disney+ and Netflix.
- However, the descriptive analysis supported the general perception that Disney+ focuses on family-friendly content, while Netflix offers a wider range of age-restricted content.

2. Rotten Tomatoes Scores:

- The hypothesis test showed a statistically significant difference in Rotten Tomatoes scores, with Disney+ having a higher median and fewer low outliers.
- This suggests that Disney+ maintains a more consistent level of movie quality compared to Netflix.

3. Summary Table:

Test Type	Variable	U-Statistic	p-value	Conclusion
Shapiro-Wilk Test	Age Restrictions	N/A	9.43×10^{-40}	Not normal
Levene's Test	Age Restrictions	N/A	7.27×10^{-9}	Unequal variance
Mann-Whitney U Test	Age Restrictions	2,137,331.0	1.0	No significant difference in restrictions.
Mann-Whitney U Test	Rotten Tomatoes	1,984,865.0	3.32×10^{-15}	Significant difference in scores.

This evaluation provides a comprehensive understanding of the statistical methods and their relevance, ensuring the validity and robustness of the conclusions.

SUMMARY

This analysis aimed to explore and compare the age restrictions and Rotten Tomatoes scores of movies available on Disney+ and Netflix. The first step in our analysis involved data preprocessing to clean the dataset. Missing values were identified in both the Age and Rotten Tomatoes columns. The Age column was also ordinally encoded to facilitate numerical analysis, and the dataset was further cleaned to ensure consistency and reliability.

For the descriptive analysis, we compared the distribution of age restrictions for movies on both platforms. The results showed that Disney+ primarily offers content that targets younger audiences, with a significant proportion of movies categorized as "7+" or "All," while Netflix had a substantial amount of "18+" content, catering to a broader, more mature audience. The Rotten Tomatoes scores for both platforms were analyzed to understand the quality of content available on each. Descriptive statistics revealed that Disney+ movies had a higher median Rotten Tomatoes score (57.5) compared to Netflix's median score of 53, with Disney+ showing less variability in the scores. Both platforms had similar mean scores, suggesting that while the overall quality might not differ drastically, Disney+ exhibited a slightly more consistent level of quality in terms of critics' assessments.

Hypothesis testing was then employed to test whether there were significant differences in age restrictions and Rotten Tomatoes scores between the two platforms. The Mann-Whitney U test, a non-parametric test, was used to compare the age restrictions for Disney+ and Netflix. The results indicated that there was no significant difference in age restrictions, meaning that despite Netflix's broader focus on adult content, both platforms offered a range of age-restricted movies. In contrast, when comparing Rotten Tomatoes scores, the Mann-Whitney U test revealed a statistically significant difference, with Disney+ movies tending to have higher scores, indicating a slight edge in quality according to critics.

In conclusion, while the analysis confirmed that Disney+ is often seen as a family-oriented platform, it did not show significantly lower age restrictions compared to Netflix. However, the Rotten Tomatoes score analysis did highlight that Disney+ generally offers movies that critics rate slightly higher than those on Netflix. This analysis provides a deeper understanding of the differences in content offerings on these two major streaming platforms and challenges some common perceptions about their target audiences and content quality.

BIBLIOGRAPHY

1. GeeksforGeeks . Understanding Hypothesis Testing. https://www.geeksforgeeks.org/understanding-hypothesis-testing/

2. PyData . Seaborn.distplot. https://seaborn.pydata.org/generated/seaborn.displot.html#seaborn.displot

3. Towards Data Science (2018). How to Compare Two or More Distributions. https://towardsdatascience.com/how-to-compare-two-or-more-distributions-9b06ee4d30bf

4. Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52(3-4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591

5. Quantifying Health. (2021, July 5). How to Report the Shapiro-Wilk Test. https://quantifyinghealth.com/report-shapiro-wilk-test/

6. Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. The Annals of Mathematical Statistics, 18(1), 50–60 https://doi.org/10.1214/aoms/1177730491

7. Statology. (2020, September 7). How to Report a Mann-Whitney U Test (With Example). https://www.statology.org/how-to-report-mann-whitney-u-test/

8. DATAtab. (n.d.). Levene Test • Simply explained. https://datatab.net/tutorial/levene-test

9. Seaborn.distplot - PyData https://seaborn.pydata.org/generated/seaborn.displot.html#seaborn.displot