Comparative Analysis of Age Restrictions and Rotten Tomatoes Scores

Netflix vs. Disney+

Submitted by:

Saber Sojudi Abdee Fard

Date:

14 January 2025

Table of Contents

I. INTRODUCTION

Streaming services like Netflix, Disney+, Prime Video, and Hulu have completely transformed the way films and TV content enters people's lives. While Netflix has a very diverse viewership, portraying the content from family friendly to maturity dramas, Disney+ is very much compromised as a children's entertainment platform because of the Disney, Pixar, and Marvel franchises. However, this leaves us with two interesting questions: Are Movies on Disney+ Actually Meant for Young Audience When Compared to Netflix? And if so, would this variety in content lead to a greater Rotten Tomatoes score for Netflix?

Despite the persistent view that Disney+ caters mainly to children, several blockbuster Marvel and Star Wars films featuring intense action or darker themes also populate its catalog. This potentially challenges the assumption that all Disney+ content is child-oriented. Meanwhile, Netflix's extensive range includes everything from original teenage rom-coms to gritty thrillers, prompting consumers to see it as a "something-for-everyone" platform. At first glance, it might appear that Netflix's wide genre coverage and adult-oriented features would lead to a higher collective Rotten Tomatoes score overall, especially if critics favor complex, mature-themed films. Yet, the Marvel Cinematic Universe titles some of which enjoy high critical acclaim live on Disney+ and could balance or even surpass Netflix's ratings. Consequently, two central questions arise:

- Are movies on Disney+ truly aimed at younger audiences compared to Netflix's offerings?
- Does Netflix hold an edge in Rotten Tomatoes scores, or could Disney+ match or exceed Netflix's average critic ratings?

To address these questions, this report utilizes a publicly available dataset from Kaggle that consolidates key information about movies across four major streaming platforms. Specifically, I focus on Disney+ and Netflix, examining two main attributes: (1) the age classification of each film, and (2) its Rotten Tomatoes rating. Both of these metrics serve as crucial indicators of a platform's content orientation: a lower age classification suggests accessibility for families and younger audiences, while higher Rotten Tomatoes scores point to strong critical reception. By scrutinizing these measures, we can glean whether Disney+ indeed revolves around child-friendly content and if Netflix's eclectic library garners higher critical favor.

This report is presented in the following structure. Section 2 concerns data acquisition, cleaning, and also the employed statistical methods, with the standard reference citations to the textbooks. The section 3 is descriptive analysis that gives some of the key performance measures and other visual presentations. Section 4 provides the selected hypothesis tests and statistical results. Last but not the least, Section 5 summarizes the findings, limitations, and future work with regards to expanding this study.

II. DETAILED DESCRIPTION OF THE PROBLEM

Every platform performs a specialized curation of films as diverse as the streaming services market today. it can be said with some amount of certainty that Disney+ is seen as somewhat more family-friendly than Netflix. Therefore, our main intention with this report is a systematic comparison of the two platforms for actually giving more children-oriented films differentiated by age restriction and the extent to which higher-rated films would be available on Netflix as compared to Disney+ measured using scores obtained from Rotten Tomatoes as one of the measures of quality.

These questions deal directly with the general motivation sketch above:

A. Age Restriction Comparison

Are the age ratings for movies on Disney+ systematically lower than those on Netflix? This question examines whether Disney+ does show a bias toward a younger audience than does Netflix.

B. Rotten Tomatoes Score Comparison

More likely are films on Netflix to bear lower or higher Rotten Tomatoes scores compared to films available on Disney+? This will assess whether rating by critics tends to favor fragment of the catalog in one of the platforms.

These questions fall well within practical aspects: families wishing to purchase a streaming service can, thereby, choose a closely associated one that offers some content for children, while filmaholics would more likely gravitate toward platforms well-known to host critically acclaimed films. The plan is to use Descriptive Statistics and Formal Hypothesis Tests to Determine whether any Observed Differences are Significant.

C. Data Source and Collection

The dataset we got from Kaggle includes four major streaming companies (Netflix, Hulu, Prime Video, and Disney+) that have pooled all their movies, into one database. The essential characteristics that matter to our study include:

- Title (name of the movie)
- Year (release year)
- Age Restriction, such as G, PG, PG-13, R
- Rotten Tomatoes Score (a numeric measure of critical reception)
- Platform Availability, (binary indicators for Netflix, Hulu, Prime Video, and Disney+).

D. Variables and Scale Levels

And the most important variables for our specific analysis include:

- Age restrictions: ordinal regarding the recommended age group: G < PG < PG-13 < R. While it is a text label, this ordinality should be distinguished as evolving from more child-friendly to adult-oriented.
- Rotten Tomatoes Score: Continuous (0-100). It measures aggregated critical reviews.
- Platform Indicators (Netflix, Disney+): Binary variables indicating whether the film is accessible on each of the services (1 = Accessible, 0 = Not Accessible).

E. Handling Data with Absences and Irrelevances

A quick scan suggests that some entries in our data might have no Rotten Tomatoes scores simply because no information was available. Since we are comparing the Age restrictions and Rotten Tomatoes scores, such movies with these two variables missing should be addressed carefully, probably by excluding them from the analysis or using possible imputation strategies, depending on the degree of missingness.

By focusing on these two areas of age orientation and critic-based quality, we will be able to provide the data-based insight into whether Disney+ is really more youth-oriented and to show that Netflix has an edge on critical content.

III. METHODS

I describe here the more sophisticated statistical techniques and visual tools used in the investigation of research questions about age restrictions and Rotten Tomatoes scores between films on Disney+ and those on Netflix. I begin with a short summary of the data handling steps involved and then list statistical and graphical techniques. Each method is referenced by standard textbooks to be given a proper theoretical justification.

A. Preprocessing and Data Handling

Before the application of any descriptive or inferential methods, the dataset followed these steps:

1) Type Conversion

Converting Rotten tomatoes score form originally as a string, like "85/100," into a numerical format 85.0, in that way, it could be mathematically processed.

2) Missing Data Treatment

Rows with missing Age or Rotten Tomatoes Scores fields are erased by python(Foundation 2023) using pandas(McKinney 2010) library. This listwise deletion has been chosen purely for simplicity. On the other hand, although some kind of other imputation methods uses mean/median imputation or multiple imputations to impute missingness, rows were left out as the fraction of missing had been relatively low in this case.(Han 2011)

3) Ordinal Mapping of Age Restrictions

Even though the age restriction labels are ordinal (all, 7 +, 13 +, 16 +, 18+), these were mapped to integer values to be calculable. The Mapping Empowers Numerical Operations along Rank based Statistical Tests.

4) Sub setting

Two subset datasets were created:

- Netflix subset: movies for which Netflix = 1.
- Disney+ subset: movies for which Disney+ = 1.

B. Descriptive Statistical Techniques

All the basic statistics like mean, variance and standard deviation are well-known types and, as such, not explicated here. But i used few advanced descriptive techniques:

1) Graphical techniques

a) Histogram

Histograms are a graphical representation of data by varying heights of bars. It divides the range into bins and shows the number (or frequency) of observation concentrated in each bin.

A histogram represents percents by area. It consists of a set of blocks. The area of each block represents the percentage of cases in the corresponding class interval. With the density scale, the height of each block equals the percentage of cases in the corresponding class interval, divided by the length of that interval. With the density scale, area comes out in percent, and the total area is 100%. The area under the histogram between two values gives the percentage of cases falling in that interval. A variable is a characteristic of the subjects in a study. It can be either qualitative or quantitative. a quantitative variable can be either discrete or continuous. A confounding factor is sometimes controlled for by cross-tabulation.(Freedman 2007) Formally, for a numeric variable X, a histogram plot:

$$f(x_i) = \frac{n_i}{N.h}$$

where n_i is the count of observations falling in the i-th bin, N is the total number of observations and h is the bin width. This provides a non-parametric estimate of the underlying distribution of X.

b) Box-and-Whisker Plot (Box Plot)

It summarizes the distribution of a dataset in terms of its five-number summary: minimum, first quartile (Q_1), median (Q_2), third quartile (Q_3), and maximum. Outliers are usually marked with points beyond these whiskers. They are particularly valuable in making comparisons across several groups such as Netflix versus Disney+.(Tukey 1977)

The filled circle encodes the median, a measure of the center, or location, of the distribution. The upper and lower ends of the box are the upper and lower quartiles. The distance between these two values, which is the interquartile range, is a measure of the spread of the distribution. The middle 50% or so of the data lie between the lower and upper quartiles. If the interquartile range is small, the middle data are tightly packed around the median. If the interquartile range is large, the middle data spread out far from the median. The relative distances of the upper Data and lower quartiles from the median give information about the shape of the distribution of the data. If one distance is much bigger than the other, the distribution is skewed.(Cleveland 1993)

c) Line chart (for CDF)

For some variables, I studied its empirical cumulative distribution function (CDF). Let $x_1,...,x_n$ be the observed values of a random sample $X_1,...,X_n$. For each number x ($-\infty < x < \infty$), define the value $F_n(x)$ as the proportion of observed values in the sample that are less than or equal to x. In other words, if exactly k of the observed values in the sample are less than or equal to x, then $F_n(x) = k/n$.(DeGroot 2012) The empirical CDF at a point x is defined as

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1\{X_i < x\}$$

A line chart of $F_n(x)$ is thus intended to show how quickly the distribution accumulates the probability mass within a certain threshold.

2) Statistical Methods

The tests sought to determine whether there were statistically significant differences between Netflix and Disney+ in the two key variables: Age and Rotten Tomatoes Score.

a) Normality Checkup

Before the final verdict is going to be given on whether to go for a parametric test or a non-parametric one, it is pertinent to see how each group's (Netflix and Disney+) data look when assessed for normality.

- Shapiro-Wilk Test.

This is the definition of the Shapiro-Wilk test statistic W:

$$W = \frac{(\sum_i a_i x_i)^2}{\sum_i (x_i - \overline{x})^2}$$

where x_i is the i-th order statistic of the sample, $\overline{x}$ is the sample mean, and ai are constants derived from the covariance matrix of the order statistics. Less than a p-value (α=0.05) means the data deviate significantly from the normality.(Shapiro and Wilk 1965)

b) Parametric Test: Two Sample t-test

For Rotten Tomatoes scores, if we assume (or approximate) normality or rely on the Central Limit Theorem for sufficiently large sample sizes, we may use a two-sample t-test. Specifically:

- Null Hypothesis (H_0): μNetflix = μDisney
- Alternative Hypothesis (H_1): μNetflix ≠ μDisney

Welch's t-statistic is:

$$t = \frac{\overline{X_1 - X_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where $\overline{X_1}$ and $\overline{X_2}$ are sample means of the two groups, s_1^2 and s_2^2 are sample variances, and n_1 and n_2 are sample sizes. This version of t-test does not require equal variances, hence making it a robust choice. If p-value goes below chosen significance level (α=0.05), we reject H_0.(DeGroot 2012) *c)* Non-Parametric Tests

Since Age is ordinal by nature and may breach the premise of normality, hence I have gone for non-parametric methods which do not assume the normality reference.

- Mann-Whitney U Test

The Mann-Whitney U test (otherwise known as Wilcoxon rank-sum test) compares equally different distributions of two documents. The null hypothesis states that both samples come from the same distribution.

The Mann−Whitney U-test is used to compare two unrelated, or independent, samples. The two samples are combined and rank ordered together. The strategy is to determine if the values from the two samples are randomly mixed in the rank ordering or if they are clustered at opposite ends when combined. A random rank ordered would mean that the two samples are not different, while a cluster of one sample's values would indicate a difference between them. (Corder 2014) For two groups of observations X={x_1,…,x_m} and Y={y_1,…,y_n}, we rank all observations together and U is based on sums of these ranks. One side alternative can be tested by the specified direction: $\text{Age}_{\text{Disney+}} < \text{Age}_{\text{Netflix}}$

- Kolmogorov-Smirnov test

We can use the Kolmogorov–Smirnov two sample test to analyze two different data samples for independence. Our data must meet two assumptions. Observations $X_1, \ldots, X_m$ are a random sample from a continuous population 1, where the X-values are mutually independent and identically distributed. Likewise, observations $Y_1, \ldots, Y_n$ are a random sample from a continuous population 2, where the Y-values are mutually independent and identically distributed. Also, the two samples are independent.(Corder 2014)

K-S test compares the cumulative distribution functions of two samples:

$$D = \sup\ |F_n(x) - F_m(x)|$$

F_n and F_m are the empirical CDFs of the two groups. Hence, a significant result implies that they differ in overall shape.

d) Effect Size: Cohen's d

Cohen's d was calculated to quantify the size of differences in Rotten Tomatoes scores. For two independent samples,(Virtanen 2020, Foundation 2023)

$$d = \frac{\overline{X_1 - X_2}}{s}$$

where

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Provision is made ford = .10 (.10) .80 (.20) 1.40. Conventional definitions of ES have been offered above, as follows: small: d= 0.20, medium: d= 0.50, large: d = 0.80.(Cohen 1998)

IV. Evaluation

A. Descriptive Analysis

1) Age Restriction

To find out whether Disney+ does actually provide children-friendly content more than Netflix, I then looked into Age for both platforms. The next table displays the average, median, and standard deviation though age is ordinal, I still retain all these statistics as an indicative snapshot:

Platform	Count	Mean (Years)	Median (Years)	Std. Dev. (Years)
Netflix	1,898	13.544	16	5.560
Disney+	725	4.102	0	4.589

TABLE I. THE AVERAGE, MEDIAN, AND STANDARD DEVIATION OF TWO PLATFORMS AGE RESTRICTION

Observations:

- The Disney+ subset appears to have less stringent average age-restrictions, indicating not so much of adult content (e.g., 0, 7, 13).
- Netflix spans a wider range which includes more of the adult-rated entries.

a) Box and Whisker Plot Age

Figure 1 and 2(below) illustrates a box and whisker age restriction comparison. The box gives us the interquartile range (IQR), the median is the straight line across the box, and the whiskers extend to 1.5×IQR from either quartile. Values beyond this range are identified as outliers with different marks.(Tukey 1977)

Fig. 1. Box and whiskers of Netflix movies Age restriction

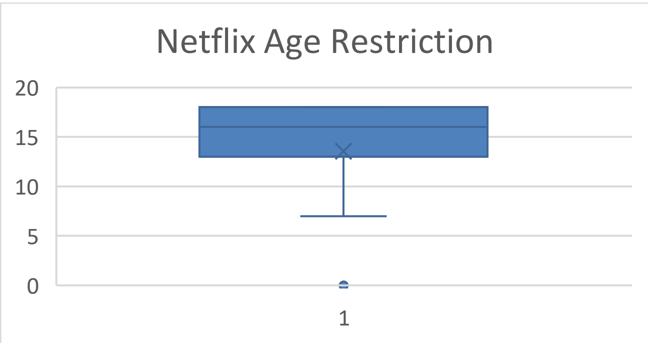
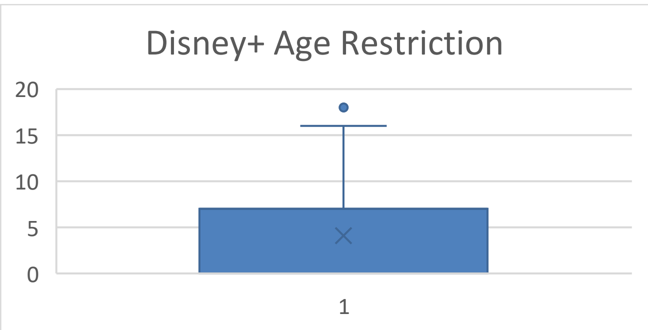


Fig. 2. Box and whiskers of Disney+ movies Age restriction



Interpretation:

- The Disney+ box seems to consolidate towards lower age ratings (0-7).
- Netflix seems to carry a higher mean age rating, implying that it has relatively more mature content.

2) Rotten Tomatoes score

We now turn to compare the distributions of Rotten Tomatoes scores across Disney+ and Netflix. Key descriptive statistics have been summarized below:

Platform	Count	Mean (Score)	Median (Score)	Std. Dev. (Score)
Netflix	1,898	60.499	60	12.808
Disney+	725	61.769	61	12.982

TABLE II. THE AVERAGE, MEDIAN, AND STANDARD DEVIATION OF TWO PLATFORMS ROTTEN TOMATOES SCORE

Observations:

- With the mean, median and standard deviations looking fairly similar across platforms, suggesting differences are not enormous.

a) Histograms for Rotten Tomatoes Score

Figures 3 and 4 show the histograms for Rotten Tomatoes scores for Netflix and Disney+, respectively. Each histogram uses bins of equal width 1 across the 0-100 range.

Fig. 3. Histogram of Netflix Rotten Tomatoes Score

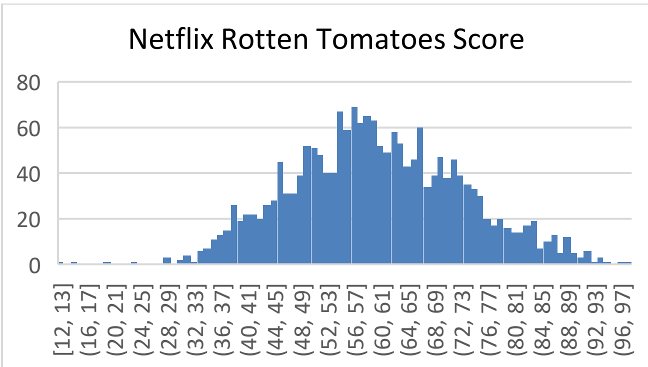
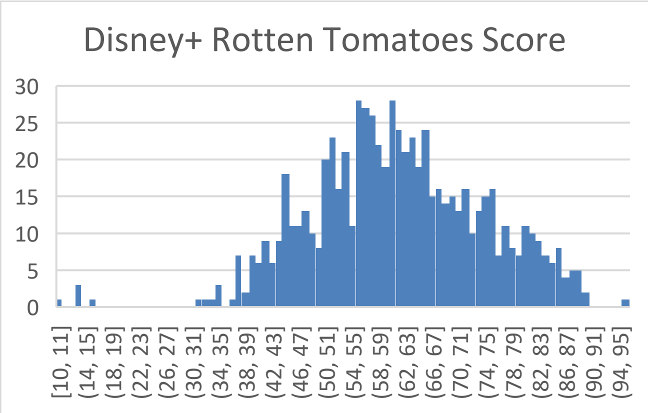


Fig. 4. Histogram of Disney+ Rotten Tomatoes Score



Interpretation:

Both distributions appear roughly bell-shaped, though Netflix's distribution might be wider due to having more movie titles.

b) Box-and-Whisker Plot for Rotten Tomatoes

The two distributions have been compared under one box-and-whisker chart in Figure 5 and 6. Their central tendencies appear to be similar.

Fig. 5. Box and whiskers of Netflix Rotten Tomatoes Score

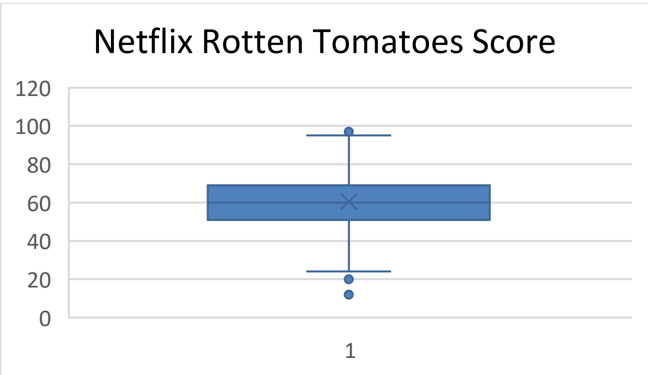
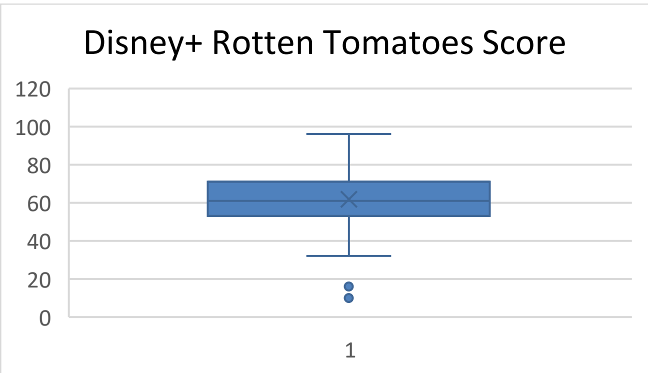


Fig. 6. Box and whiskers of Disney+ Rotten Tomatoes Score



B. Inferential Study

A bit more closely following this descriptive preamble, I conduct statistical hypothesis tests to find whether any of the tested differences show up as reliably significant.

1) Age Restriction Hypothesis Test

a) Rationale for the Test:

- Variable Type: Age is mapped from categorical labels (e.g., “7+”) to ordinal integers (e.g., 7).
- Non-Normality: Age restrictions are not naturally continuous or normally distributed.
- Selected Method: The Mann-Whitney U Test was chosen because it is a non-parametric method that compares the distribution of ranks between two independent groups

b) The Mann-Whitney U Test(Virtanen 2020):

- The age values of Netflix and Disney+ and combined and sorted in ascending order.
- The number of ranks for Disney+ (R_D) and for Netflix (R_N) has been summed.
- $U_D = R_D - \frac{n_d(n_d+1)}{2}$ and $U_N = R_N - \frac{n_N(n_N+1)}{2}$ calculated.
- The smaller U is used as the test statistic (some software use the larger but p-values are consistent either way).
- p-value needs to be computed (exact or approximate) based on the distribution of U or normal approximation. For Large Samples the distribution of U approaches a normal distribution, and the mean and variance of U are given by:
- Mean: $\mu_U = \frac{n_1 n_2}{2}$
- Variance: $\sigma_U^2 = \frac{n_1 n_2(n_1+n_2+1)}{12}$

- Z-score: $Z = \frac{U-\mu_U}{\sigma_U}$ Excel(Corporation 2023)
  Formula: =NORM.DIST(Z,0,1,TRUE)*2

Z-score can then be used to find the p-value using the standard normal distribution.

Sum of Netflix Ranks	3011850.5
Sum of Disney+ Ranks	431709
$U_{Netflix}$	1209699.5
$U_{Disney+}$	168534
Mean of U	688025
Standard deviation of U	17346.3618
z-score	-29.9481244
p-value	4.654E-197

TABLE III. THE MANN-WHITNEY U TEST CALCULATED VALUES

Python result:(Foundation 2023) (Virtanen 2020)MannwhitneyuResult(statistic=np.float64(166350.5), pvalue=np.float64(2.656927667628983e-212))

- Null Hypothesis (H_0): Distribution of age restrictions is equal for Disney+ and Netflix.
- Alternative (H_1): Age restrictions on Disney+ will generally be lower (more child-oriented) than those of Netflix.

Result: The extremely small p-value suggests that we can reject H_0 at the 5% significance level.

2) Rotten Tomatoes Score Hypothesis Tests

a) Rationale for the Test:

- Using the Shapiro-Wilk Test by python,(Virtanen 2020, Foundation 2023) both Netflix and Disney+ scores produced:
  Netflix scores Statistic: 0.9960128052099871,
  P-value: 6.57418175404767e-05
  Disney+ scores Statistic: 0.9874354943028479,
  P-value: 7.0263151868374596e-06
  A large p-value (> 0.05) indicates no strong evidence to reject normality. Combined with the bell-shaped histograms, this justifies a parametric approach.

b) T-Test:

- Null Hypothesis (H_0): There is no difference in Rotten Tomatoes score between Netflix and Disney+.
- Alternative (H_1): There is a significant difference in Rotten Tomatoes scores between Netflix and Disney+.

Using ttest_ind(...) from Python(Virtanen 2020, Foundation 2023), with equal_var=False (Welch's t-test):

T-Statistic: 2.25, P-Value: 0.0247

According to that output, the value of p-value was less than 0.05 indicating difference in scores statistically significant.

c) Additional Tests

To confirm the results of t-test and to remedy the distributional and sample size imbalance raised, some non-parametric or alternative tests were conducted as follows:

- Mann-Whitney U Test:(Virtanen 2020)

Null: The distributions of Rotten Tomatoes scores do not differ.

Result: p <0.05, in agreement with a difference.

- Kolmogorov-Smirnov (K-S) Test:(Virtanen 2020)

Null: The two samples come from the same continuous distribution.

This was the only test which returned p≥0.05, hence suggesting the absence of any significant difference in the overall shape of the distributions. This method uses CDF of the data, Figures 7 and 8 shows Line chart of CDF of two platforms which visually justifies the result of this test. Also, Cohen's d: 0.0985 This indicates a very small effect size, meaning the difference in means is negligible in practical terms.

Fig. 7. Line chart of CDF of Netflix Rotten Tomatoes Score

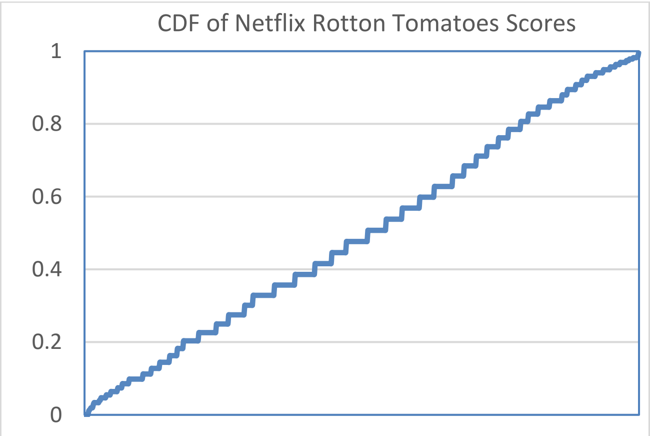
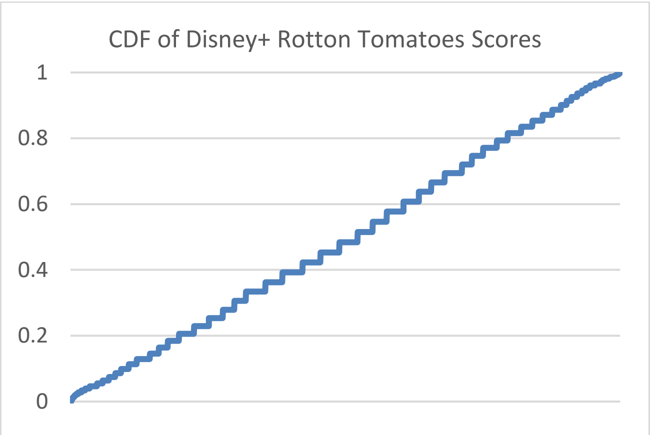


Fig. 8. Line chart of CDF of Disney+ Rotten Tomatoes Score



C. Presentation of Findings

1) By Statistical Importance Alone:

While multiple methods converge on significance (except for the K-S test in the Rotten Tomatoes comparison), we have not yet addressed effect sizes or real-world implications.

2) Sample Size Effects:

Netflix is around 1898 while Disney+ has only about 725 entries, and huge sample sizes have small p values, regardless of modest absolute differences between groups.

3) Potentially Distributional Factors:

Even though the Shapiro-Wilk test gave a suggestion of approximate normality, the absence of significance in the K-S test may be taken to show similarity in the overall distribution of the scores (broadly bell-shaped).

The findings from the analysis reveal that indeed, it seems that Disney+ concentrates on lower-age rating categories, while Netflix traps a much wider range of material for that older audience. The difference in Rotten Tomatoes scores, however, while it seemed small from graphical exploration, most hypothesis tests threw up statistically significant results except K-S at α=0.05.

V. SUMMARY

This project aimed at making comparisons between the two biggest streaming platforms: Disney+ and Netflix in two aspects: (1) age restrictions, which would indicate whether Disney+ genuinely serves a younger audience; (2) Rotten Tomatoes scores which would clarify if movies from Netflix generally receive higher critic ratings than from Disney+. Data used was from a Kaggle dataset containing titles for different online platforms, although focusing mainly on the subsets Disney+ and Netflix.

A. Age Restrictions

A strong nonparametric Mann Whitney U Test suggested that Disney+ movies statistically had lower age classifications than Netflix ($p < 0.05$). Descriptively, Disney+ exhibitions tilted more toward the categories "0+" and "7+," while Netflix had a wider range, also including content such as "18+." Hence, this data-deduced finding goes hand-in-hand with the popular interpretation that Disney+ is "child-friendly".

B. Rotten Tomatoes Scores

Multiple tests (including t-test, Mann-Whitney U Test) suggest that, in fact, differences between the Rotten Tomatoes scores in Netflix and Disney+ are statistically significant. Interestingly, all evidence, in terms of histogram and boxplot inspection, reveals these two distributions to be very much similar but tests p-values results smaller than 0.05 likely influenced by the large sample size in Netflix's catalog compared to Disney. However, The Kolmogorov-Smirnov (K-S) Test did not throw up significant differences but suggested that overall, the nature of distributions may be quite similar.

C. Real-World Interpretation and Discussion

The analysis now echoes consumer thinking that Disney+ generally offers more family-friendly films. This could also be a message to parents or guardians of kids looking for safer viewing options. On the flip side, there is definitely more content for feeding a wider audience in Netflix, which includes even more adult-oriented content preference. The marginally statistically significant difference along with Rotten Tomatoes would just owe to some factor-such as the nature of Disney franchises comparing to Netflix Originals, the mix of classics vs. contemporary hits, or simply the sample size contrast.

But, indeed, the practical size of the difference as opposed to its bare statistical significance deserves attention. In large datasets, even the most minor differences in means or medians give rise to p-values. This implies that the difference may be statistically confirmed yet hold real-world meaning that is small for the average viewer.

D. New Questions and Areas for Future Work

1) In terms of genre analysis:

one could break down the data by genre type (for example, action, comedy, family) to see if such differences are present when age ratings or critic scores are compared across categories.

2) Year of Release:

A time-based study could explore whether platforms pursued recency of title, hence impacting their Rotten Tomatoes scores.

3) User Ratings:

Inserting opinions from users who have other audience ratings or those from a site like IMDB would provide an alternative viewpoint of quality assessment.

4) Content Maturity Trends:

Auditing as to how the catalog of the portal's changes (for example, if the Disney+ increases its proportion of adult movies) could lead us to see the shift in age restriction distribution over time.

Thus, it would seem that Disney+ does cater to a younger audience in comparison with Netflix, and the Rotten Tomatoes analysis does find this, but the real effects of this modest difference remain uncertain. By looking at other genres and historical trends, one can help objective the reality of the changes happening in the world of streaming.

REFERENCES

Cleveland, W. S. (1993). Visualizing Data. Summit, NJ, Hobart Press.

Cohen, J. (1998). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ, Lawrence Erlbaum Associates.

Corder, G. W., & Foreman, D. I. (2014). Nonparametric Statistics: A Step-by-Step Approach (2nd ed.). Hoboken, NJ, Wiley.

Corporation, M. (2023). Microsoft Excel. Redmond, WA, Microsoft.

DeGroot, M. H., & Schervish, M. J. (2012). Probability and Statistics (4th ed.). Boston, MA, Pearson.

Foundation, P. S. (2023). "Math Module Documentation." from https://docs.python.org/3/library/math.html.

Foundation, P. S. (2023). Python Language Reference, version 3.x.

Freedman, D., Pisani, R., & Purves, R. (2007). Statistics (4th ed.). New York, NY, W. W. Norton & Company.

Han, J., Kamber, M., & Pei, J (2011). Data Mining: Concepts and Techniques (3rd ed.). Waltham, MA, Morgan Kaufmann.

McKinney, W. (2010). Data Structures for Statistical Computing in Python: 56-61.

Shapiro, S. S. and M. B. Wilk (1965). "An Analysis of Variance Test for Normality (Complete Samples)." Biometrika **52**: 591-611.

Tukey, J. W. (1977). Exploratory Data Analysis. Reading, MA, Addison-Wesley.

Virtanen, P., et al. (2020). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python." Nature Methods **17**(3): 261-272.