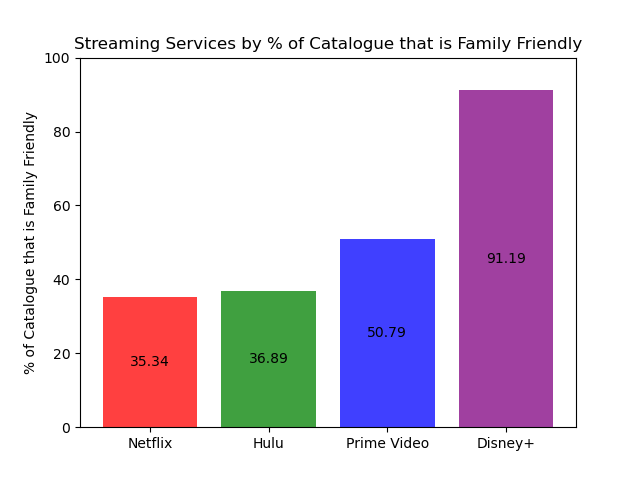
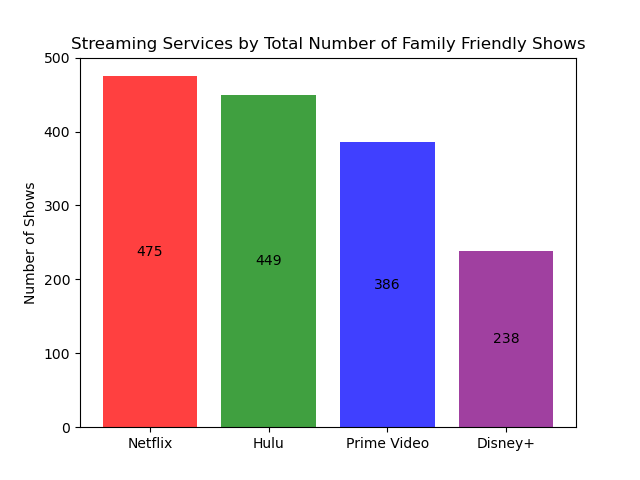
Project 1 Write-Up

Our group set out to answer questions about streaming services based on a dataset of the TV show offerings of 4 different streaming platforms: Netflix, Hulu, Amazon Prime and Disney+. We came up with three questions that could help us identify which streaming service is best and also when the best shows have come out.

Our first question was which streaming service is the best for families. In order to answer this, we took the csv file and created a dataframe to hold the data and dropped any entries that were missing an age value. We then created dictionaries to hold counters and looped through the dataframe, counting the total number of shows from each streaming service and how many shows are for ages “all” or “7+”. Once we had that information, we calculated the percentage of shows that are family friendly per streaming service. From our analysis, we found that Disney+ has the highest percentage of shows on their platform that are family friendly, with 91.19% of their content being family friendly. This was vastly greater than any other service, with Amazon Prime being the closest at only 50.79%. The data implies that Disney+ could be trying to position themselves as a family friendly streaming service by having a majority of their content fit that label.

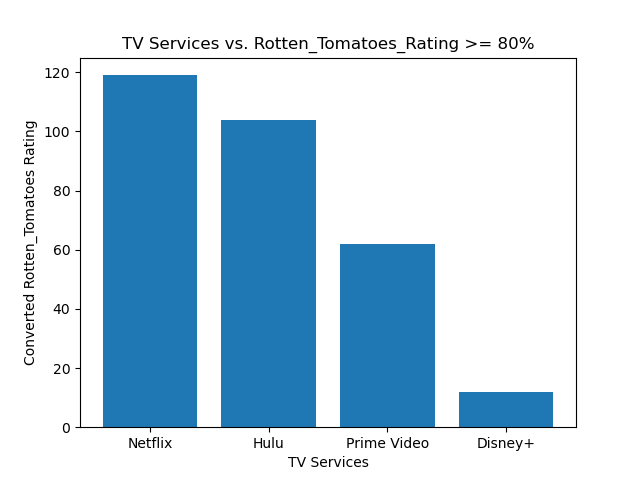
To some, this would be enough to answer the question of which streaming service is the best for families, but digging a little deeper revealed that this question may have multiple answers. Netflix, which only had 35.34% of their content rated as family friendly, actually had the highest total number of family friendly shows with 475, followed closely by Hulu with 449. Disney+ has the least total family friendly shows with 238. Therefore, one could reasonably say that Netflix is the best streaming service for families because it has the most family friendly content. 

This left us with two correct answers to our question of which streaming service is best for families. Disney+ is the best streaming service if you are looking for a mostly family-friendly streaming service, while Netflix is the best streaming service if you are looking for the platform with the most family friendly content.

The second question was about which service has the most highly rated shows based on the given dataset, this giving us the opportunity for better decision which service is more suitable. IMDb and Rotten Tomatoes rating services are in used to help us identify the answer.

The bar graphs below allow for visualization of the result after data analyzing for each rating service:

* Rotten Tomatoes: Netflix had the highest TV shows with total count of 119 with rating >= 80% compare to others



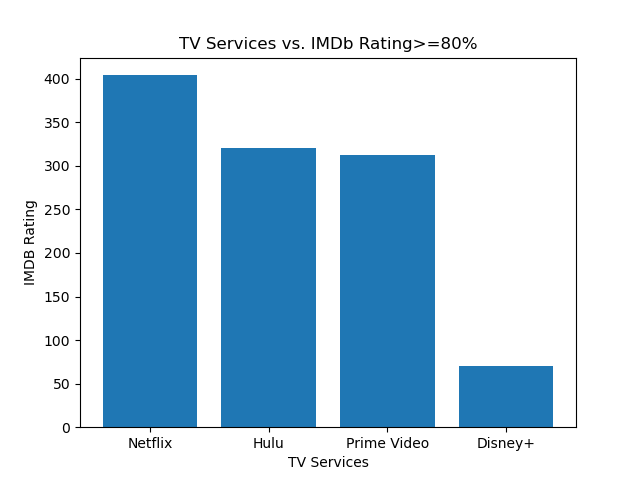
104

62

12

119

IMDB: Netflix had the highest TV shows with total count of 404 with the rating >= 80% compare to others



320

312

70

404

On the results of 2 rating services, Netflix has more highly rated TV show choices for us.

For our third question regarding the identification of the most highly rated shows by year, the rating services become an important differentiator. Not only are the Internet Movie Database [IMDb] and Rotten Tomatoes unique in scales used; popularity in the marketplace; and years each covers, they lead to different answers to this seemly straightforward question, “Which year had the most highly rated shows?”. There is a limitation in the dataset, as it only goes through 2022.

The line graphs below allow for visualization of a peak year and dataframes can summarize the years and counts.

* IMDb best year is 2019, as noted below, where we have 85 shows ranked greater than 8.0 /10.0 (or 80%):

A close up of numbers

Description automatically generatedA graph showing the number of years

Description automatically generated

* Conversely, for Rotten Tomatoes, there were 23 shows ranked greater than 80 / 100 (or 80%) in 2015. These results are shown below:

A close up of a text

Description automatically generatedA graph showing the growth of tomatoes

Description automatically generated

The discrepancy in the years identified as having the highest counts means that we should further investigate any kind of equivalence between the ratings systems. So, if we look at a single year, 2019, we find that there is only a modest association between IMDb ratings and Rotten Tomatoes ratings. This year is one of two the team could evaluate given it is when IMDb had the most highly rated shows. The R-value is 0.49, indicating that a high score in IMDb ratings does not necessarily indicate a high score in Rotten Tomatoes. For example, the scatterplot does have points where high IMDb scores for Television shows where Rotten Tomatoes scores are low. There does not seem to be an equivalence between the two rating systems:

A diagram of a number of blue dots with a red line

Description automatically generated

If the range is expanded to cover the years between 2015 (Rotten Tomatoes’ year of most highly rated shows) through 2019 (IMDb’s year of most highly rated shows), there is much change in R-value. In fact, it decreases slightly to 0.47. This means with even more data included in the analysis, there isn’t any increase in equivalence. Ultimately, this seems to indicate that the year of most highly rated shows is predicated on which ratings system is your favorite.

A graph with blue dots and red lines

Description automatically generated

Ultimately, the year with the most highly rated shows is predicated on which ratings system is the user’s favorite.