EDA Executive Summary

Anna Wagman

Data Science Final

Contents

Introduction:

My EDA compares and contrasts the titles availale on Hulu, Disney+, and AmazonPrime. I chose to examine data about streaming services because I have always loved and watched a lot of movies. I personally subscribe to all three websites, along with a few other streaming services such as Netflix and HBOmax. Throughout my EDA, I extract specific variables such as genre, type, and duration. In order to interpret these variables through metrics andd visual graphics, I had to clean and recode a lot of the variables and inputs. My first two visualizations highlight the variable type, which has two possible values: "Movie" or "TV Show". If one wanted to be economical and optimize their options, my first visualization depicts the total amount of available titles per service. As shown in my pie chart, AmazonPrime has significantly more options than Hulu and Disney+. However, my most interesting finding was from the variable duration. In order to evaluate "duration", I had to individually parse the duration based on type. Specifically, for type = "Movie", my findings were very surprising.
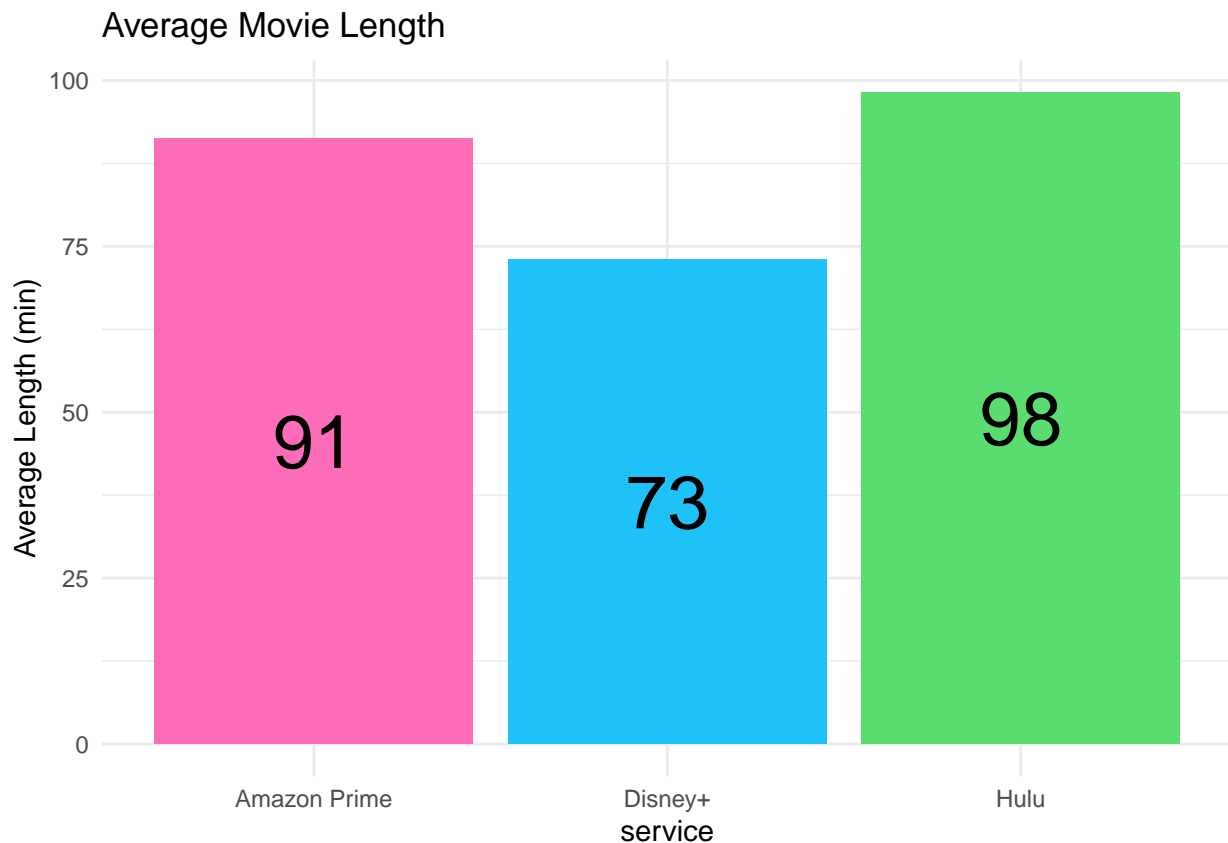
Data Sources:

https://www.kaggle.com/shivamb/hulu-movies-and-tv-shows/version/1. https://www.kaggle.com/shivamb/amazon-prime-movies-and-tv-shows https://www.kaggle.com/shivamb/disney-movies-and-tv-shows

Highlight 1:

Average Movie duration per service in minutes

```
Movie_av_plot
```
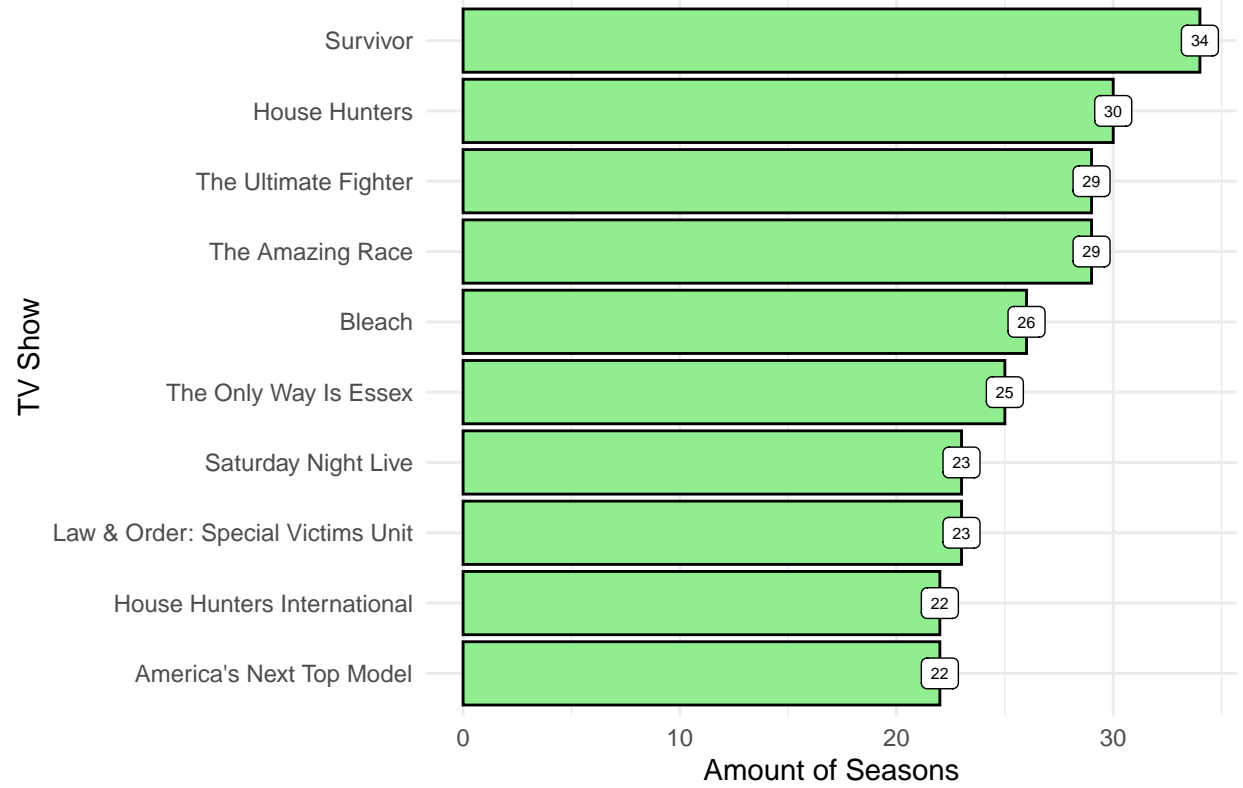
Average Movie Length

In this bar chart, I highlight the average (mean) duration of movies in minutes. I find it super interesting that movies on Disney+ are on average significantly shorter than on Hulu and AmazonPrime. Movies on Disney+ average at 1 hour and 13 minutes which is pretty short for a movie. AmazonPrime and Hulu have similar movie length averages of 91 and 98 respectively, which is about 1.5 hours. 1 hour and a half is what I predicted the average movie length would be. I guess it makes sense that Disney+ is mostly family / animated movies which tend to be shorter than dramas and horror films ..etc

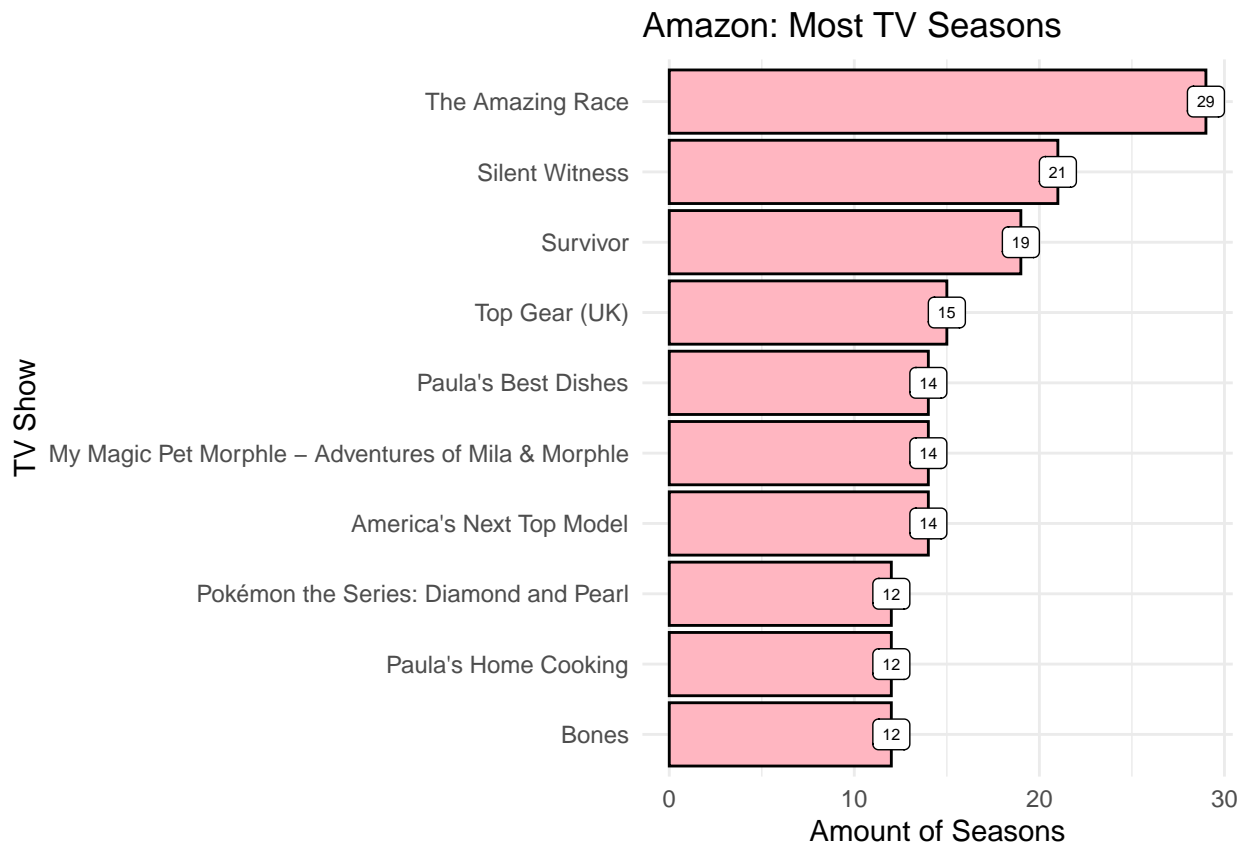## Highlight 2: Longest running TV Shows per service

TV Shows with the most seasons on Hulu, Disney+, and AmazonPrime

```
hulu_seasons
```

Hulu: Most TV Seasons

TV Show	Amount of Seasons
Survivor	34
House Hunters	30
The Ultimate Fighter	29
The Amazing Race	29
Bleach	26
The Only Way Is Essex	25
Saturday Night Live	23
Law & Order: Special Victims Unit	23
House Hunters International	22
America's Next Top Model	22

```
amazon_seasons_plot
```

Amazon: Most TV Seasons

In these two bar charts, I calculated the 10 shows with the most seasons on Hulu and AmazonPrime. What I love about this plot is that I unintentially extracted TV Shows that are available on both services. As you can see above, both Hulu and AmazonPrime have the TV Shows "Survivor" and "America's Next Top Model", which both have a ton of seasons. However, they have each have a different amount of seasons available. "Survivor" has 34 seasons on Hulu and 19 seasons on AmazonPrime. In my fourth Visualization, I found how many overlapping titles the services have which each other, i.e. titles available on both. However, I returned the number n = total overlap and not the specific titles. So this visualization helped highlight a few of the titles available on Hulu and AmazonPrime

Conclusion:

I learned that AmazonPrime, Hulu, and Disney+ are far more different than I had expected. I did not expect AmazonPrime to have so many more titles than Hulu and Disney+. AmazonPrime has over 9000 titles, while Hulu and Disney+ have about 3000 and 1500 respectively. Additionally, I did not expect the distribution of Movies vs TV Shows to be so uneven on each service. The percentage of Movies on AmazonPrime and Disney+ are both over 70%. However, Hulu is almost 50/50 Movies and TV Shows. In conclusion, if you are looking to be economical and subscribe to only 1 service, AmazonPrime has by far the most options. However, if you are specifically interested in watching TV Shows over Movies, I would recommend subscribing to Hulu.