IST 652:
The Final Project

Produced By:

Jonathan Durbin & Patricia Meadows

The daring adventure of two students looking to explore Movie and TV Show offerings from online platforms that have taken too much of their time.

POPCORN

Background

What we are working with:

6 Datasets with two CSV files each.

One CSV file contained information on the TV Shows and Movies of each platform. The IMDB and TMBD scores were included in these files.

  *TMDb is an open-source platform for rating movies

  *IMDb is a publicly accessible platform owned by Amazon

The other CSV file contained information from the credits of each platform's offerings.

Prepping the Data

Cleaning:

Step 1: Download the CSV files to our local computers

Step 2: Iterate through each raw file reformatting and removing any rows that cannot be saved, adding the platform for the file, and rewriting them to a cleaned folder on our local computer.

Step 3: Create one file for each of the two types of CSVs available; combining all 6 platform's data.
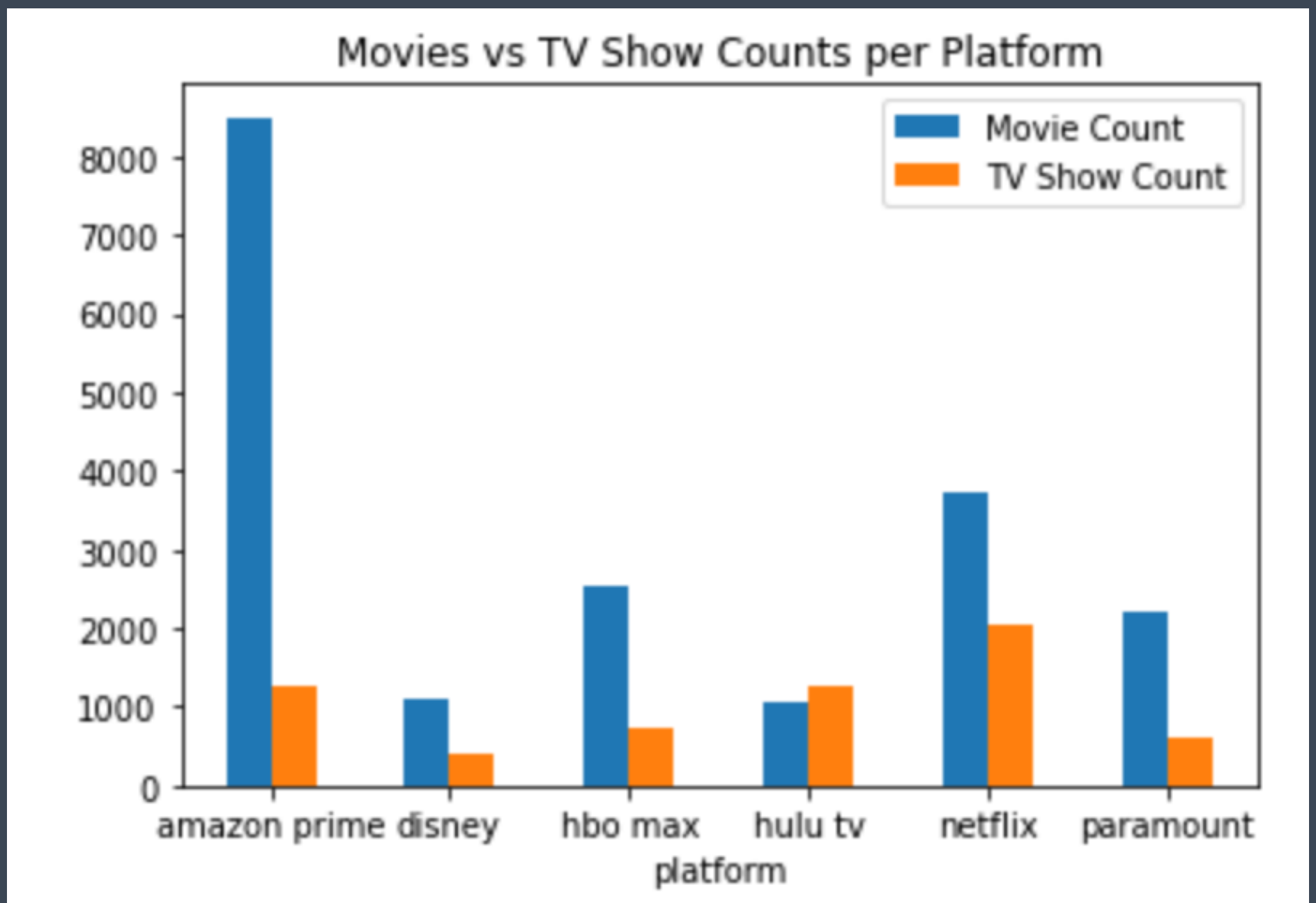
Final Files:

Credits File: This is a DataFrame of 366,429 Actors and Directors.

Titles File: This is a DataFrame of 25,430 Movie and TV Offerings from all the providers.
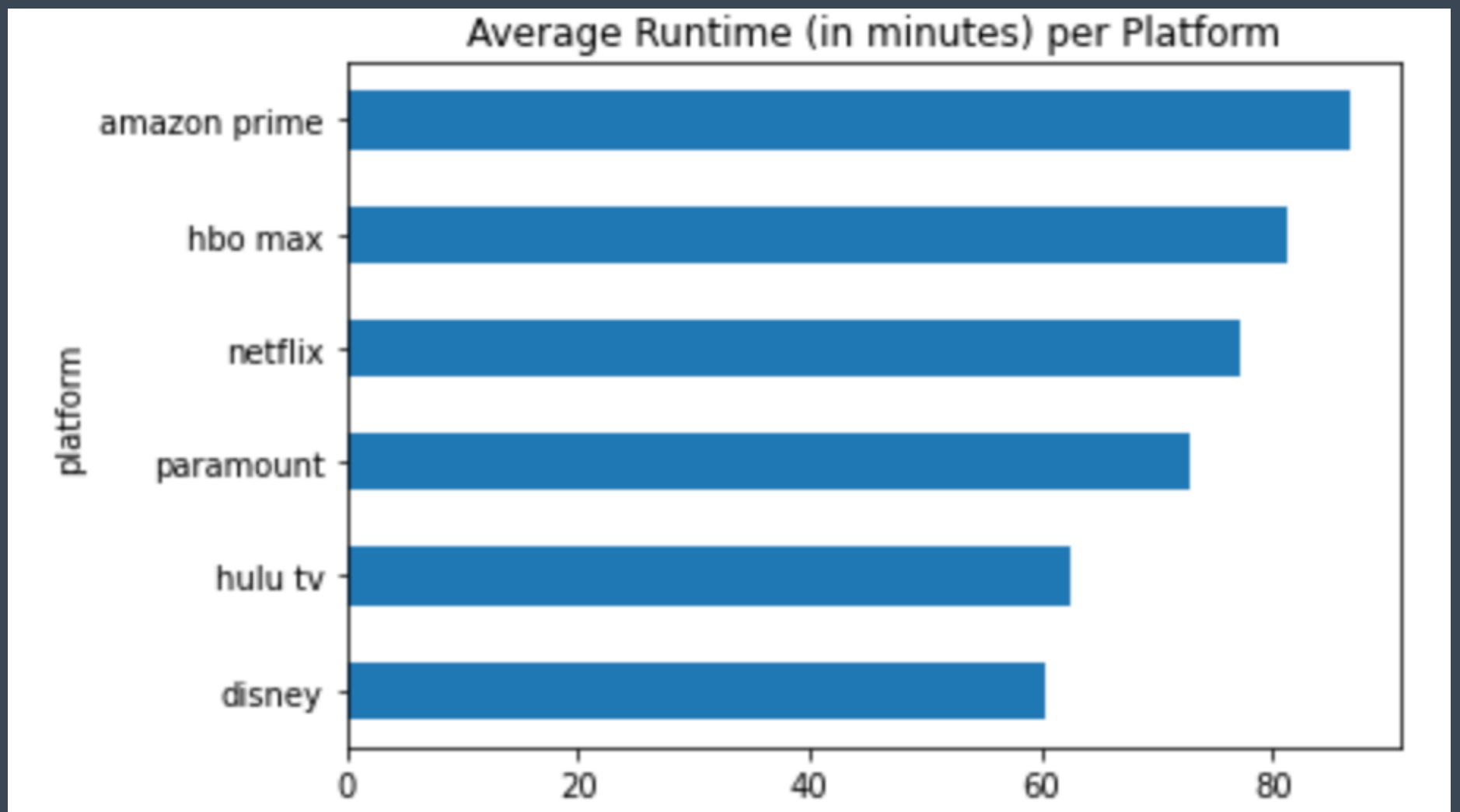
Movies or TV Shows?

Movies vs TV Show Counts per Platform

Netflix looks to have a great 2:1 ratio of Movies to TV Shows.

How do you find any movie on Amazon??

How Much Time Do You Have?

Average Runtime (in minutes) per Platform
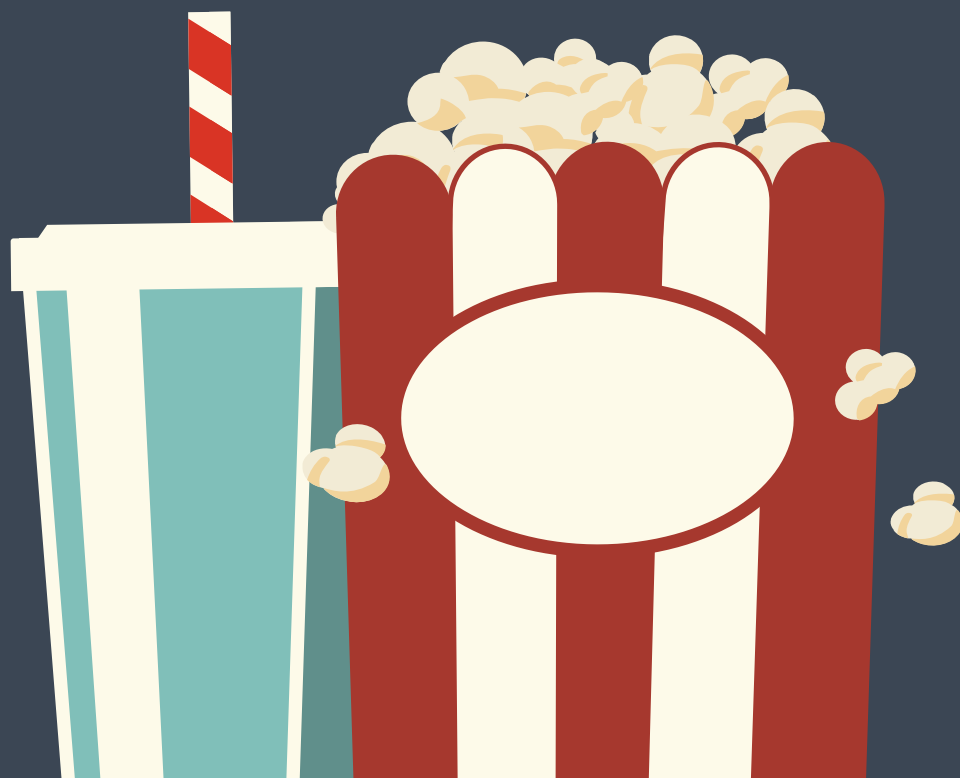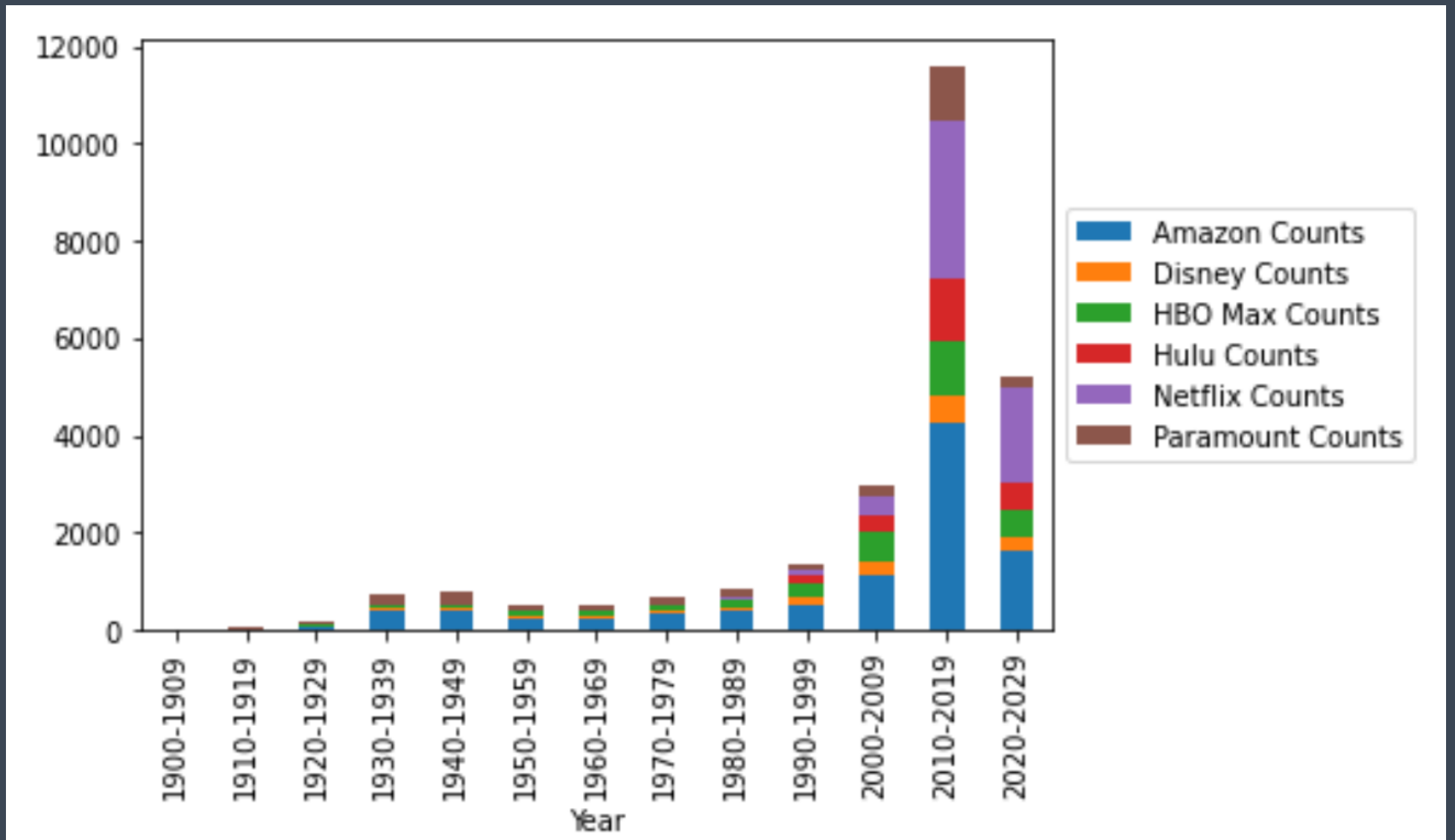
Disney tailors towards family and children content and Hulu also is known for quick watching so the shorter average runtimes make sense.

Amazon Prime is far more movie based so a longer average runtime is expected.
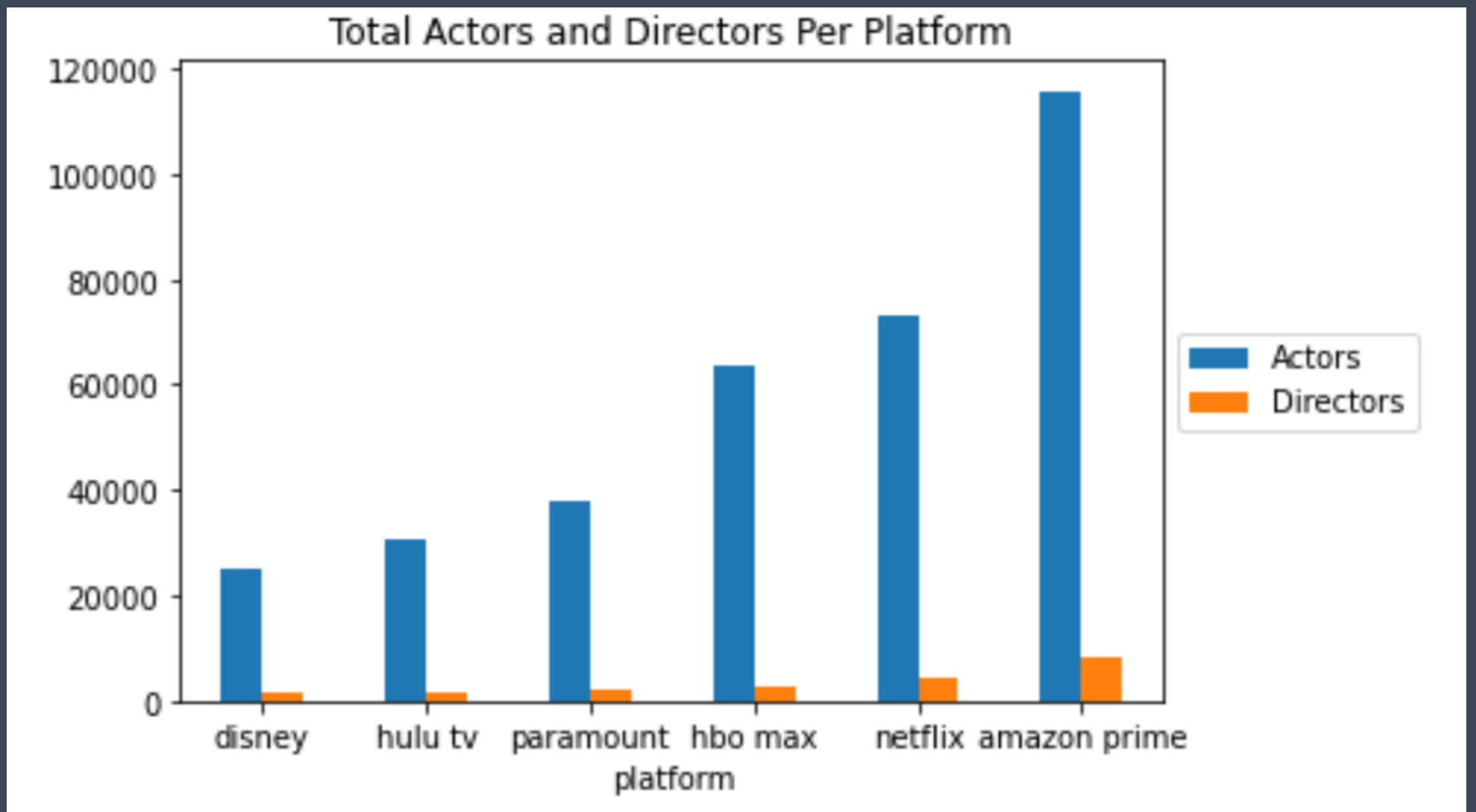
How Recent are the Offerings?

For the most part, all the platforms offer the same distribution of the age of their offerings.

Of note: the maximum release year in the data is 2022.

I Need Variety!

Total Actors and Directors Per Platform

Amazon offers far more variety in their Movies and TV Shows.

I Need Variety! (Continued)

Genre Counts for all platforms:

```
drama          : 11608
comedy         : 8867
thriller       : 5011
action         : 4833
romance        : 4241
documentation  : 3925
crime          : 3433
family         : 2975
scifi          : 2580
animation      : 2506
fantasy        : 2495
horror         : 2193
european       : 1804
music          : 1180
history        : 1090
western        : 855
reality        : 799
war            : 773
sport          : 693
no genre       : 423
```
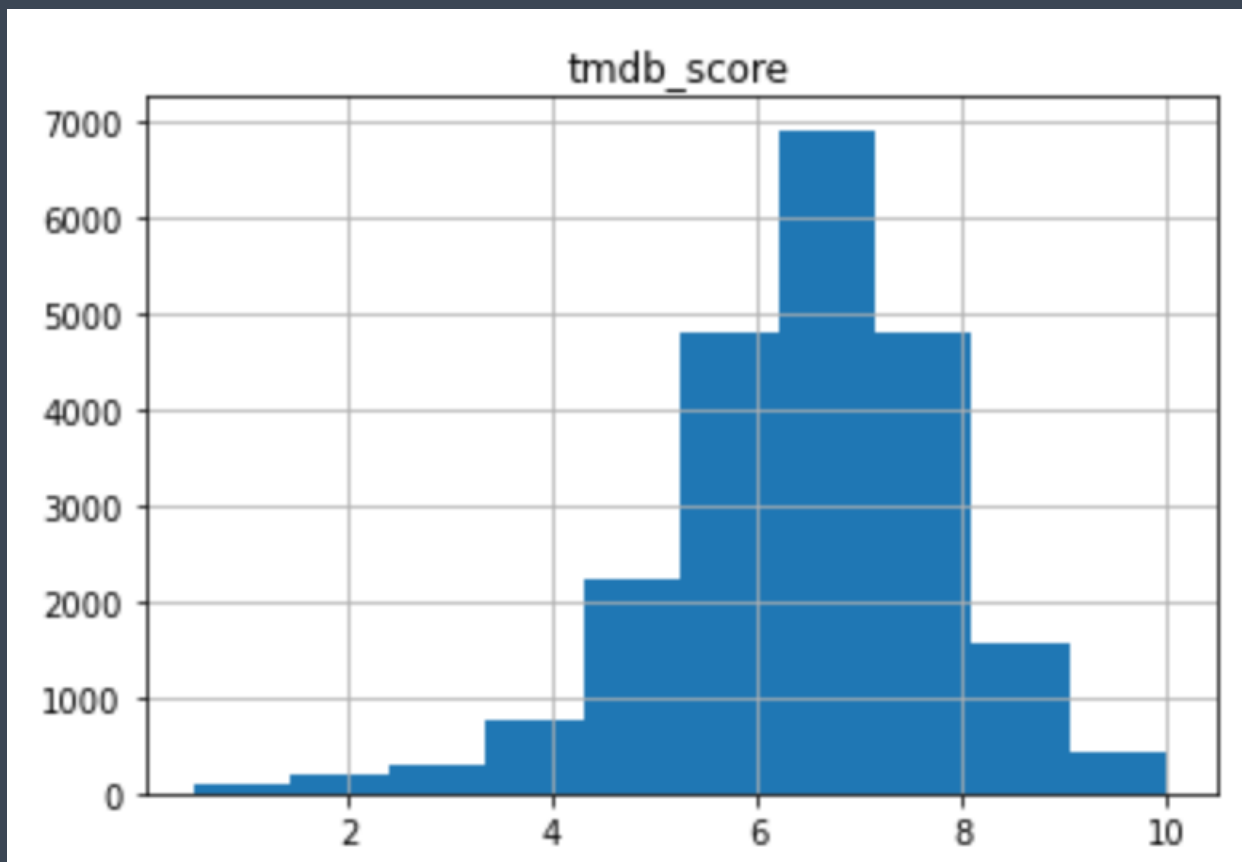
Across Platforms the most common genre will be in the drama category.
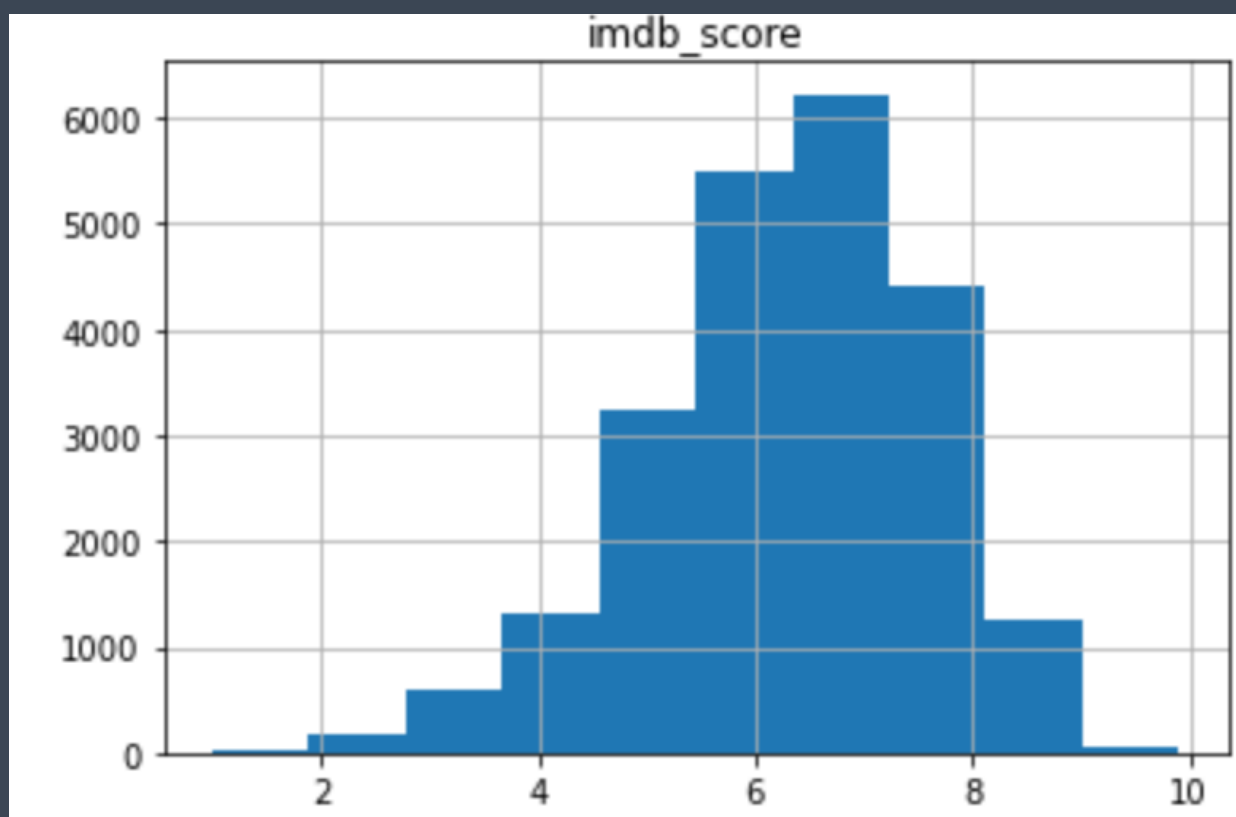
How Can You Tell What is Good?... Reviewers!
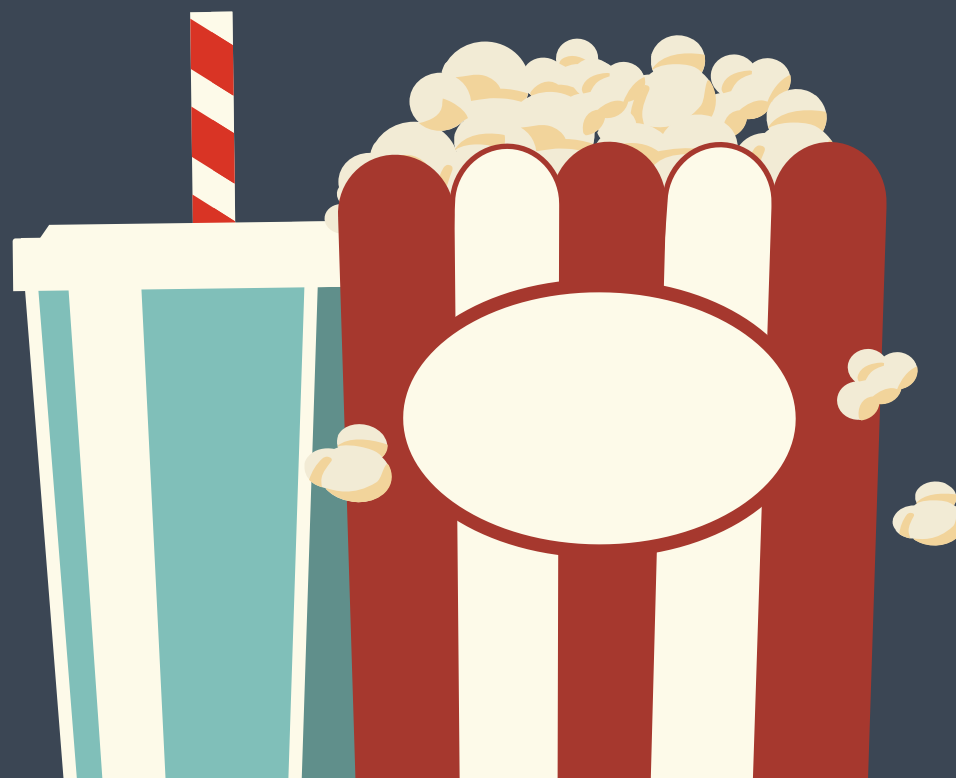
What to expect from IMDb and TMDb reviews.

TMDb

More reviews.

Wider Span.

IMDb

Tighter Grouping.

Will It Be Any Good?
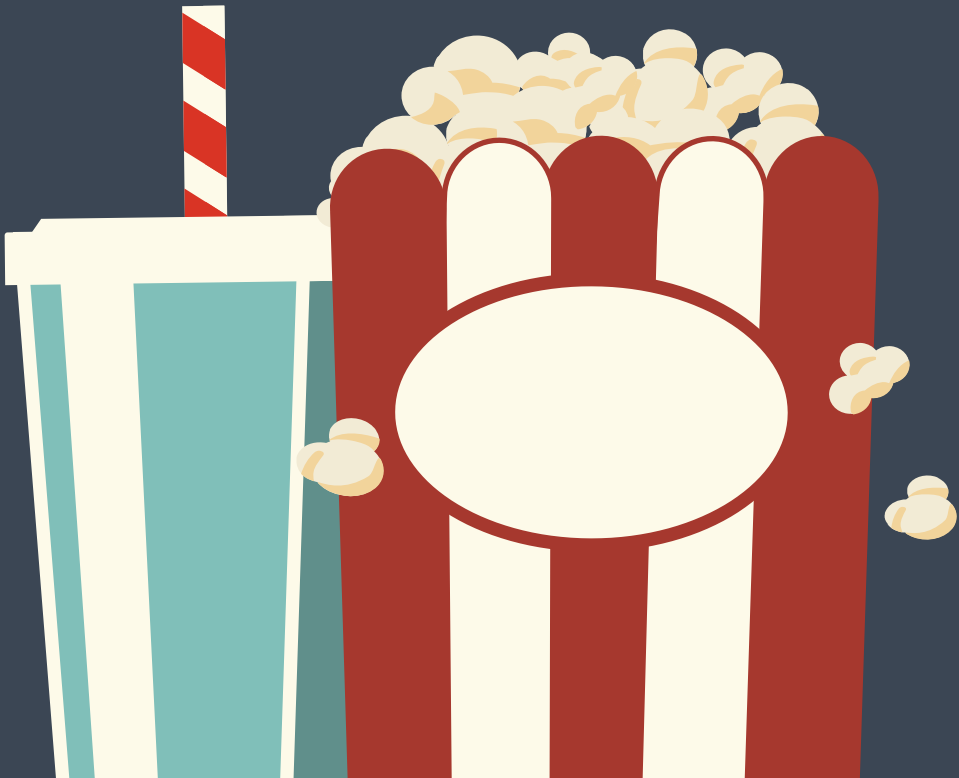
Average Ratings Per Genre across all platforms:

TMDb

animation	: 7.17
reality	: 7.15
documentation	: 6.93
history	: 6.91
family	: 6.82
fantasy	: 6.79
scifi	: 6.65
war	: 6.65
sport	: 6.63
music	: 6.55
drama	: 6.52
european	: 6.52
crime	: 6.49
comedy	: 6.46
action	: 6.43
romance	: 6.36
thriller	: 6.21
horror	: 5.78
western	: 5.66

IMDb

history	: 7.90
music	: 7.10
european	: 7.10
sport	: 7.00
fantasy	: 6.90
comedy	: 6.90
family	: 6.90
animation	: 6.90
scifi	: 6.10
romance	: 6.00
crime	: 5.10
drama	: 5.10
thriller	: 5.10
horror	: 4.80
reality	: 4.60
western	: 4.40
action	: 4.40
war	: 4.40
documentation	: 4.10

There is a difference in the popularity of genres per rating site.

Which Platform Has Good Offerings?

Average Rating per Platform

Legend: avg_IMDb, avg_TMDb

X-axis (platform): amazon prime, disney, hbo max, hulu tv, netflix, paramount
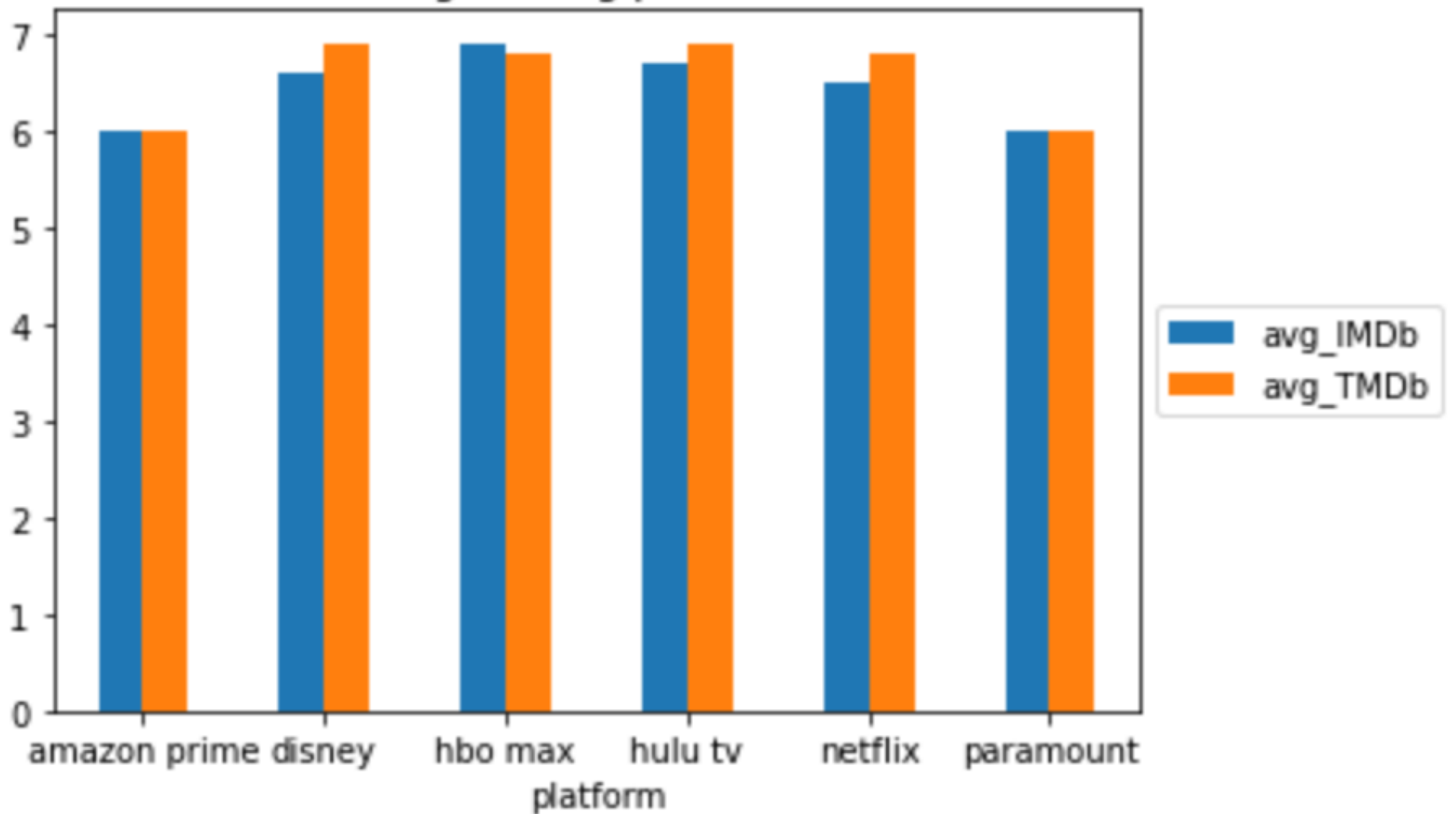
Nearly equal ratings for Amazon and Paramount.

Disney, Hulu, and Netflix are more highly rated on TMDb.

HBO Max is more highly rated on IMDb.

What Else?

Looking Beyond:

* Use a sentiment analyzer to review the Descriptions for the offerings to see how the top Movies and TV Shows differ per platform.

* Find more data on the monetary aspect of the platforms to see if it aligns with the IMDb and TMDb findings.

* Research how so many offerings have been made after only 2 years of the 2020s.

That's all Folks!