

My Film-Watching Activity Report

Introduction

Films have always been a part of my life. While the Star Wars saga and The Lord of the Rings played crucial roles in shaping my identity, the rest of the movies I watched in my youth merely served as pastimes, a reason to go out with my family, or background noise while I built Lego sets. Everything changed during the COVID pandemic. I had a lot of time on my hands to explore new hobbies. Cinema was the one interest that stuck, and my appreciation for the art form has been growing exponentially since.

As someone who loves statistics and documenting information, I naturally wanted to visualize my viewing activity. Letterboxd is a platform that does this well, but I wanted to see if I could create dashboards by myself, with visualizations that were not available on the app. The two main questions that I wanted to ask were:

1. How did my selection of films change over time?
2. How much more active have I been with watching films?

I designed dashboards to answer these questions, creating general visualizations based on four years' worth of data, and also plots for each of the four years. In this report, I will cover where I got the data from. I will then interpret the visualizations before answering my two motivating questions directly.

Data

The data that I used for this project came from two places: Letterboxd and TMDB. I have been logging every film that I have watched since March 2022. I started doing this on IMDb, but then exported my data and switched to Letterboxd in September 2023. When I download my data from Letterboxd, they provide numerous sheets. Using R, I joined a few of them together and selected the statistics that I considered useful.

For TMDB, I used a Python script to use the platform's API to create a dataset. This dataset consisted of statistics for the films found in my Letterboxd log. Once I obtained this, I joined my TMDB dataset and my various Letterboxd datasets together in R.

After combining the data, I had to perform *a lot* of data cleaning and wrangling to put the data in a format that was usable. There were numerous errors in the data that I had to fix on a case-by-case basis. Figuring out how to make use of the multiple sheets and narrow my manufactured data frames to non-overwhelming ones was a hassle, but rewarding.

Explaining and Interpreting Visualizations

1. Genres

The first plot to interpret is *Figure 1.0*, which is a pie chart displaying how many films I've seen from each genre. It's important to note that most films belong to multiple genres, so they are represented in more than one of the wedges. For instance, *Back to the Future* is a science fiction, fantasy, and comedy film; therefore, it is represented in the counts for each of those three wedges. I also want to note that the "Other" wedge consists of war stories, westerns, and TV movies. The counts for those three genres were insignificant, so they were grouped together for convenience.

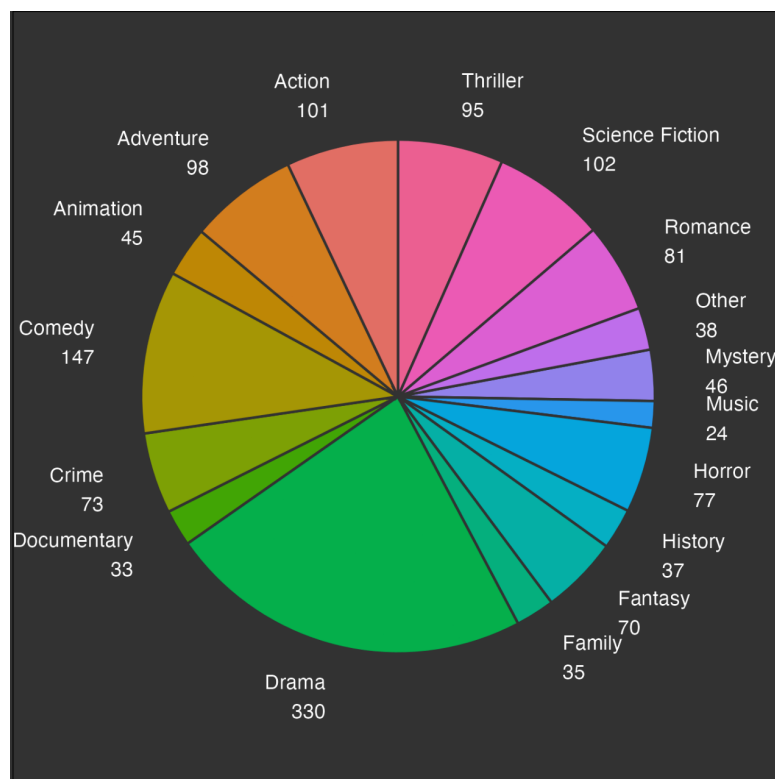


Figure 1.0 - Representing the genres of the films I've seen

It's clear to see that I watched more dramas than anything else - a whopping 54%, which isn't surprising considering that it's quite a versatile genre. I'm surprised to see that 24.1% of the films were comedies. I don't seek out funny movies, yet it's my second most-watched genre. I don't think every film that has a joke should be considered a comedy, but perhaps TMDB assigns the comedy genre label generously. I'm also surprised to see that thrillers made up just 15.5% of the movies I watched. I thought it would be a much more popular genre when investigating my activity.

2. When I Watch Movies

I used comparative charts to see if *when* I watched a movie changed over the years. This would most simply be seen across the days of the week (*Figure 2.0*). I also mapped the films that I saw across the weeks (*Figure 2.1*) and months (*Figure 2.2*) of the year to see if my activity could be mapped to events that occurred in my life.

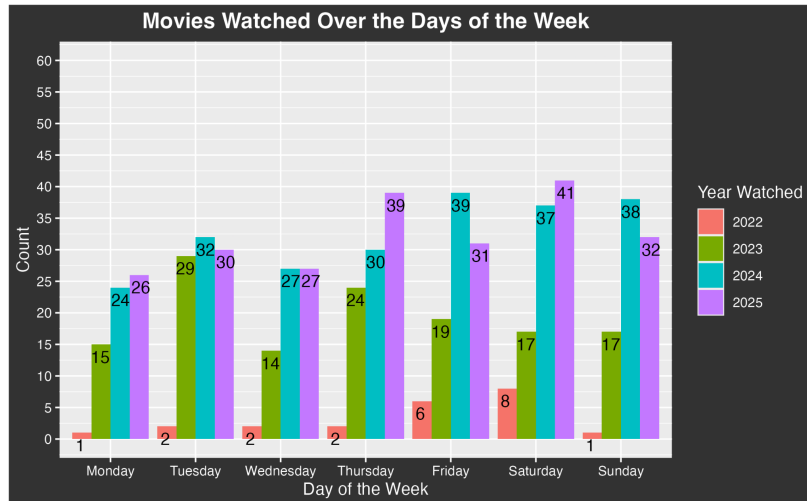


Figure 2.0 - How frequently I watched films over the days of the week

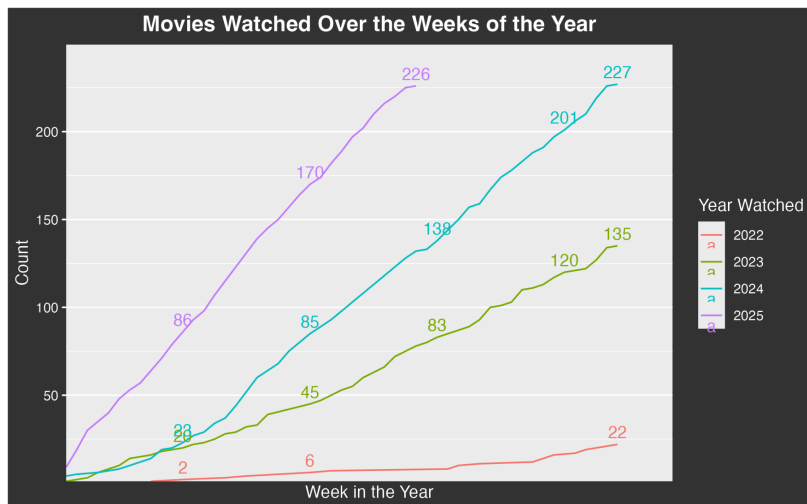


Figure 2.1 - How frequently I watched films over the weeks of the year. It better represents my rate of film-watching over the year

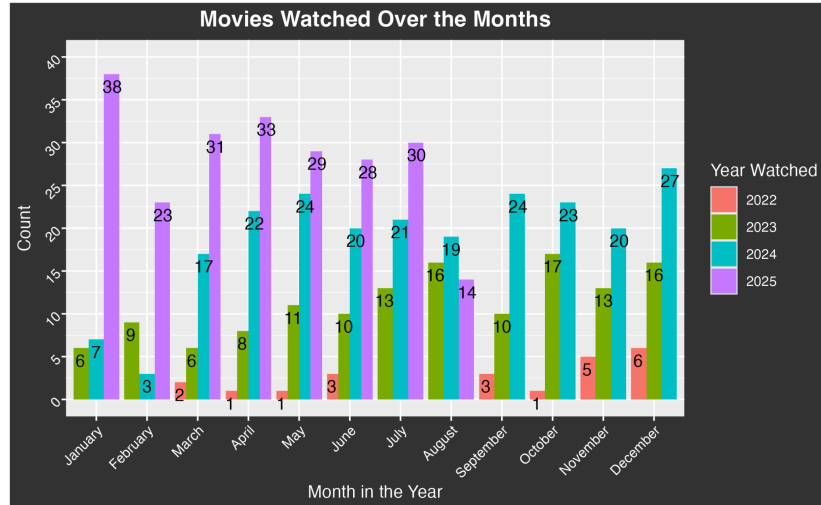


Figure 2.2 - How frequently I watched films over the months in the year

Before getting into the year-by-year plot-by-plot breakdown, the most obvious change is that the number of films that I watch has been increasing substantially yearly. Given that I'm making this three-quarters of the way through 2025, I'm predicting that this year's total will be higher than ever. I think that I'm approaching a ceiling, though.

In 2022, it is clear to see that I watched films most frequently on Fridays and Saturdays. This is expected, given that it's the end of the school week. In 2023, my activity was more consistent throughout the week. This makes sense because I wasn't in school or working until the end of August - I was available most days. I took a film history class that Fall, where we watched a film every Tuesday, which is why that bar is much higher than the rest. My activity over 2024 and 2025 was pretty similar in distribution, which is expected.

Figure 2.1's weekly breakdown doesn't imply much. One thing that I can see is that the rate at which I watch films has increased year by year and has remained consistent throughout that year. There aren't any significant plateaus that can be aligned with events in my life. This surprised me a little, but there truly weren't prolonged periods where I was forced to put films to the side.

It's surprising to see (from Figure 2.2) that I watched fewer films in the first half of 2023 than I did in the latter half. I did not have school or work in the first half of the year, and I thought that I watched quite a few movies during that time. In March of 2024, despite a difficult course load, I strongly committed to watching more movies (following a challenging February). My numbers took a considerable leap in that month. My rate was consistent through the rest of the semester, over the summer, and up to the end of the year.

December 2024 and more so in January of 2025, my numbers were larger than normal because I took two 25-hour flights, in which I binged movies. By this time, my love for cinema had grown considerably to the point that I watched a movie every possible chance I could get. In August, I worked more than I ever have and was simultaneously doing Summer courses that required a

lot of my time, so I had to go easy on the film-watching.

3. Writing Reviews

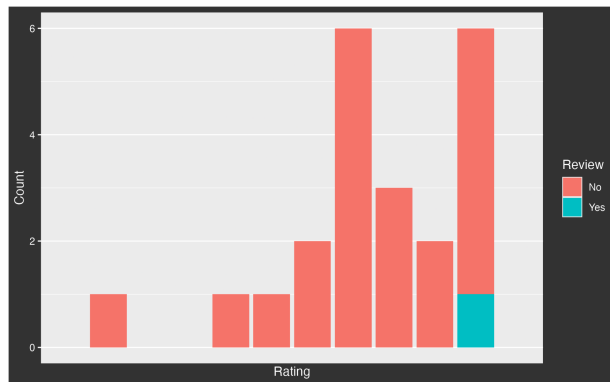


Figure 3.0

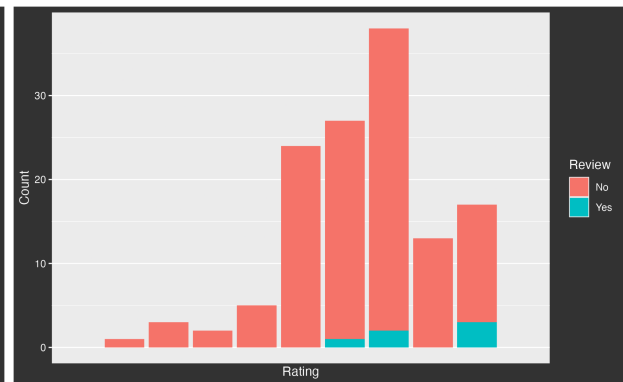


Figure 3.1

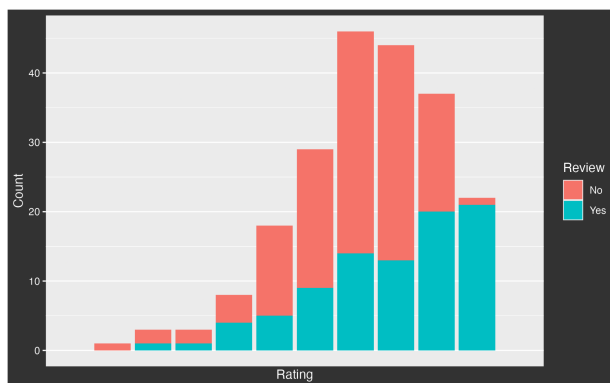


Figure 3.2

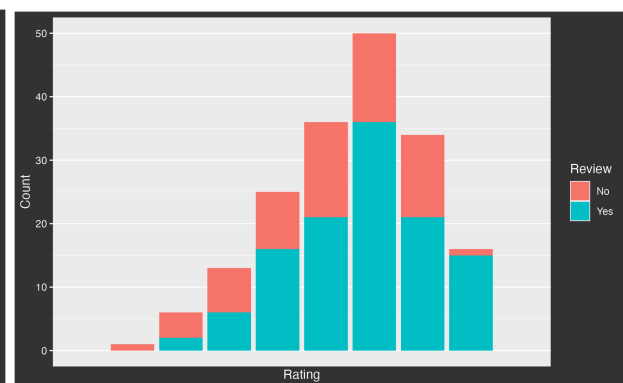


Figure 3.3

Figure 3.0-3.3 showcases the rating distribution for films that I've reviewed

Every year, as my film literacy improves and as I find more to say about what I watch, I leave more reviews. Out of the films I saw in 2022, only one film was reviewed (4.5%) - even then, I retroactively reviewed it in a different year. In 2023, I reviewed 6 films (4.4% of my yearly count). In 2024, that number jumped up to 40.5% and leaped up to 62.4% in 2025. This is even more outstanding when considering my increased volume of movies watched.

In 2023, I barely reviewed any films; the ones that I reviewed were movies that I gave a higher rating to. In 2024, that changed. Even though I had more reviews for higher-rated films, I attribute that statistic to me watching more 'good' movies. I reviewed almost every film that I assigned a perfect rating to. I also reviewed half of the 4.5/5-rated movies. Looking at the lower end of the distribution, I reviewed approximately half of the films that I rated 2 stars and below. Perhaps I was better able to articulate why I disliked those films? Or maybe I became more passionate about my opinions.

Finally, we move on to 2025. I am very pleased to see this distribution because the ratings for my reviewed films share a similar distribution to the ratings overall. I reviewed about 60% of the higher rated films and approximately 40% of the lower rated films. It seems that I'm still less likely to review a low-rated film, but they're solid percentages across the board.

4. Diversity in Film Languages and Production Countries

I almost exclusively watched English-language movies throughout 2022 and 2023. I watched more foreign films in 2024, but far fewer than I thought - only around 10% (estimating from the plot). My foray into foreign films doubled in 2025, with over 20% of those films being non-English-speaking.

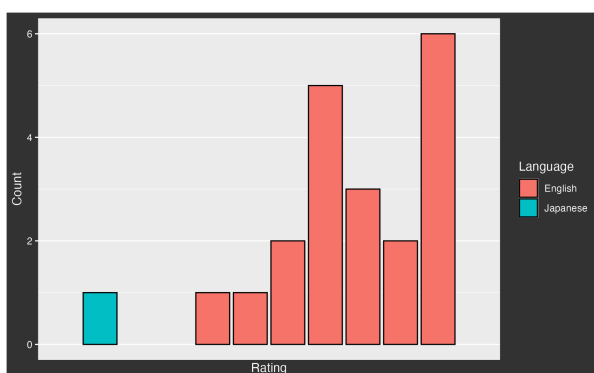


Figure 4.0

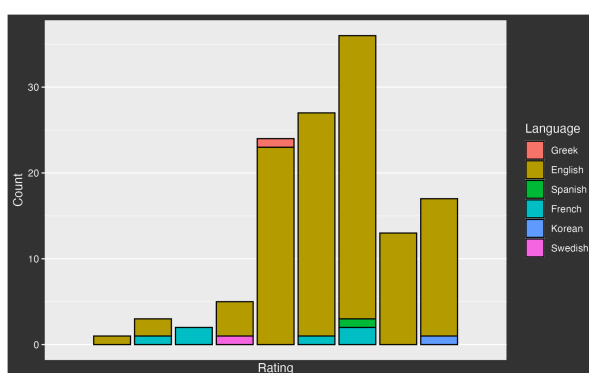


Figure 4.1

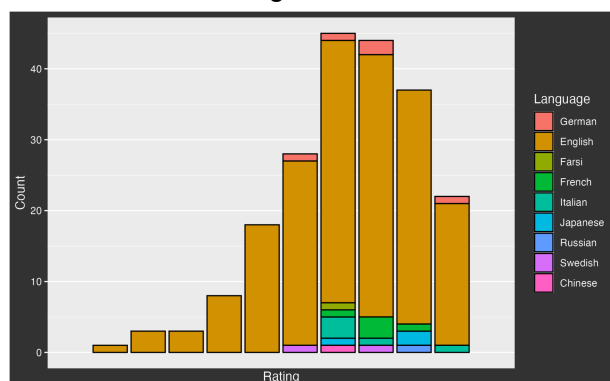


Figure 4.2

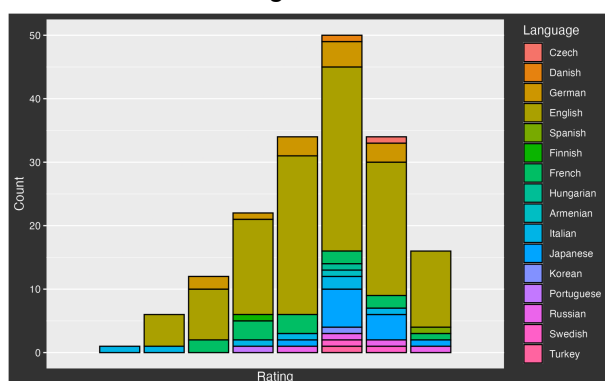


Figure 4.3

Figures 3.0-3.3 display my rating distribution of films, color-coded by spoken language

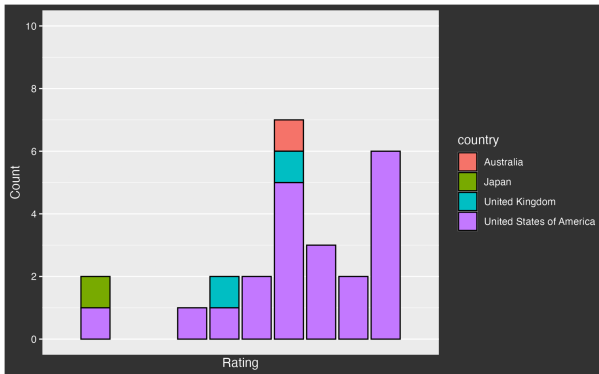


Figure 4.4

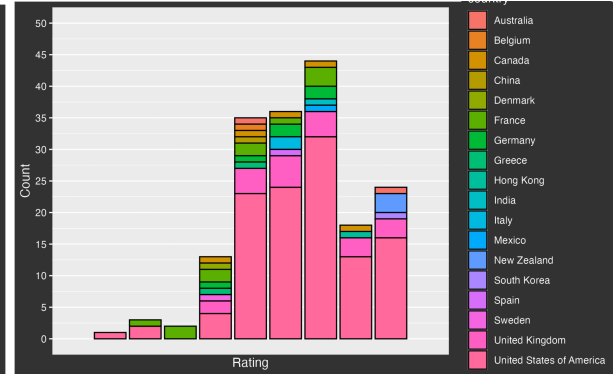


Figure 4.5

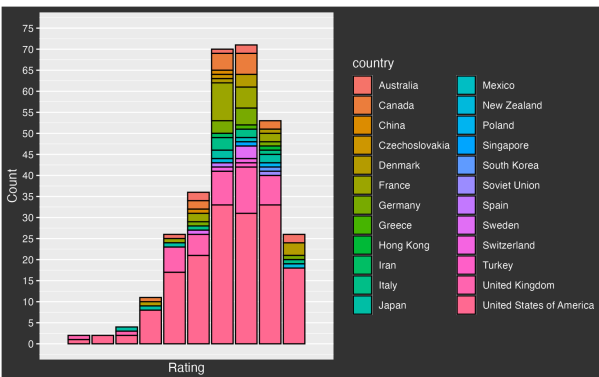


Figure 4.6

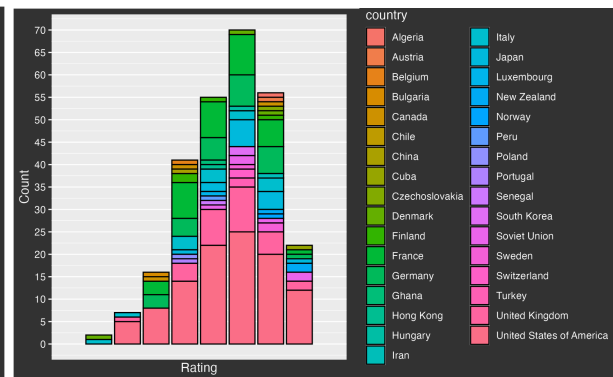


Figure 4.7

Figures 4.4-4.7 display my rating distribution, color-coded by production country

This increase in diversity is also reflected in the plots showing the countries where the films were produced. There are more entries in the plot legends and more variety in color across the different years. These plots are less interesting to me because there are many instances where multiple countries collaborate to get a film made. I think spoken languages (*Figures 4.0-4.3*) are better indicators.

Looking at how the variety relates to the distribution of ratings, it looks like most of the non-English-speaking movies that I watched in 2024 were 'good' movies (having a 4 star rating or above). In 2025, it looks like the distribution of non-English-speaking films better resembles a normal distribution, which is a good thing in my books. This speaks again to the increased diversity,

What isn't represented in these plots are silent films. I believe that I have been watching more of them, but these visualizations don't give me a perspective on that.

This increase in diversity is aligned with my intent to explore movies from all over the world. I have learned so much about different countries, their values, cinematic styles, and history.

Answering My Motivating Questions

1. How did my selection of films change over time?

The films that I've chosen to watch have been more diverse over the years, having different languages and being made in many more countries. In 2022, I only watched films from 2 different languages. That figure increased to 6 in 2023, then to 9 in 2024, then to 16 in 2025.

Most of the movies that I watch have been dramas and comedies, which are generally versatile genres. I explore fantasy, crime, mystery, and horror pretty evenly. However, regarding genres, I can not comment on how my selection has changed over time.

2. How much more active have I been with watching films?

It's very clear to see that I have been watching way more films than I ever have. Over the last four years, I have steamrolled my total from each previous year. My rate of film-watching has evidently increased and is pretty consistent over the months (despite varying amounts of commitment).

2025's total is continuing to increase, and I feel as though this is approaching the limit of what I am capable of (without sacrificing other necessities and responsibilities). With just over a year of college left before I hopefully kick off my career, I have to make the most of the time that I have!

To Improve and What to Explore Next Time

I think my visualizations have given me an insightful look into my film-watching activity. I noticed patterns that I wasn't aware of before. That being said, there is a lot more than I could do and would like to do.

As far as visualizations go, there are many that I would want to include next time so that I can answer more questions. These include:

- How my interest in silent films has changed over the years.
- Genre breakdown year-by-year to see if I've changed my preferences or if I explore more types of films.
- How often I rewatch films.
- The runtimes of the films that I watch year by year to see if they reflect the amount of free time I have.

A major change that I want to work on moving forward is making it an interactive dashboard. It would be a lot more convenient to use drop-downs to toggle between the different years' statistics. Using a platform like Tableau would also be a more professional display. It's also more convenient than doing manual graphic design work.