

Anime Stats

Process Book

CS 5630 - Visualization for Data Science
December 4, 2022

Dylan Hansen
Ervin Chhour
Khris Thammavong

1. Overview and Motivation

Anime Stat aims to display many different statistics of trends in anime over time. More specifically, the changes in trends of anime over the course of time with a focus on the trends of genres. We'll compare the number of shows for each genre along with finding the cumulative amount of genres over time.

We chose this topic because everyone in our group watches a ton of anime and we thought it would be interesting to research and do a visualization on various statistics on the topic. Whether it be if we want to see how a favorite genre of ours compares to other genres or looking at if there is any correlation between genres and seasons, seeing if there are any genres that air more often in a certain season, we feel like visualizing this data will be exciting and interesting for fans of anime.

2. Questions

What are the most popular genres of anime per season? How much anime of each genre is currently out? What are the most popular genres overall? What are the most highly rated genres? Out of the most popular anime, what is the rating distribution among those? How popular are the most highly rated anime? For all of the anime of a certain popularity level or of a certain rating, how many of those are older vs newer?

The overall benefit to answering these questions is to gain a better understanding of how anime trends have evolved over time and how the anime community views those trends both in terms of how many people consume those trends and how good they are.

3. Data

We have managed to collect our data using a python script from the API of MyAnimeList.net which is one of the most commonly used websites for viewing information about anime. For each anime it stores its genre, season released, popularity, and average user rating, all of which is necessary to our project. If we need more data, we could also use the API of Anilist, another website that stores the same information about anime. Using a second website would give us a wider data set and would lead to more accurate results.

<https://myanimelist.net/>

<https://myanimelist.net/apiconfig/references/api/v2>

<https://anilist.co/>

<https://anilist.gitbook.io/anilist-apiv2-docs/>

4. Data Processing

During our project proposal, we didn't expect to do much data cleanup since we were thinking that we'd get the data in the right form after some minor data processing. However when we were actually collecting the data, we had to take the metadata of each anime and for each season we count how many genres are in that season. It was a bit of work collecting just that specific data, and more than we expected to have to be doing. Besides that, the main bit of processing we had to do was grouping all the data that we receive from the API into an easier to manage data structure.

The data we are mainly concerned with a particular anime is its genre, rating, popularity, and release date. The lists of anime along with the needed quantities of each anime can be obtained very easily via the API.

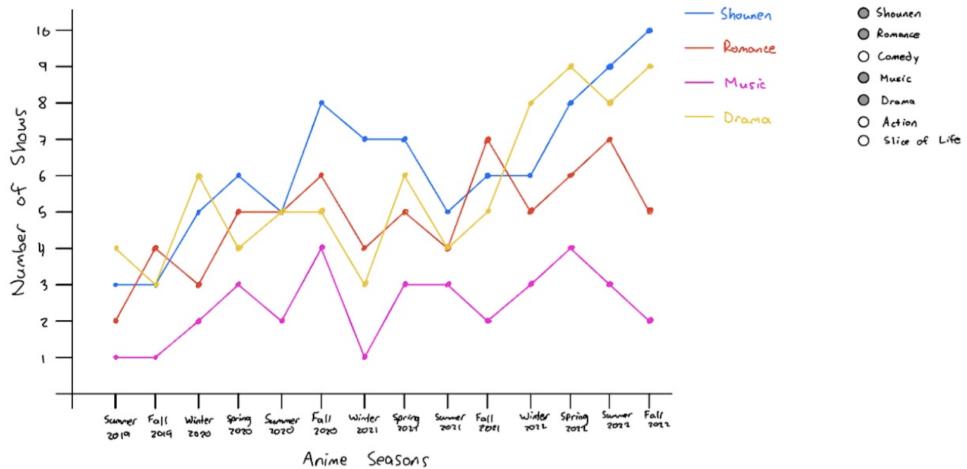
5. Design Evolution

Our initial visualizations were drawn out on a digital white board and we just planned out the basic features of how our graphs would interact with other things such as filters. We didn't put too much thought into our design layout until we had to do the peer feedback assignment in class. After seeing how much thought and effort the team we were giving feedback to put into their visualizations, we realized that we had to put more thought into our layouts. We also noticed that they used a visualization tool to draw out their sketches. So for our new designs we used Figma to draw them out cleanly as it was a visualization tool that we had experience using before.

The main difference between our old and new designs is the way we drew out our interactions for our graphs in one area. During our peer feedback assignment we got feedback from Brendan, one of the other team members reviewing our project, and he gave us the idea to have one interactive area for all our graphs. This cleaned up the layout of our visualizations and made it easier to see where you could change stuff to interact with our graphs.

A **line graph** comparing different genres of anime and seeing how many were released per season. Includes some way to filter out/see specific genres.

Proposal Design:



Milestone Design: We came up with two different genre picking designs. One is just a list of all the available genres and the other would let the user search through them.

Charts
Genres

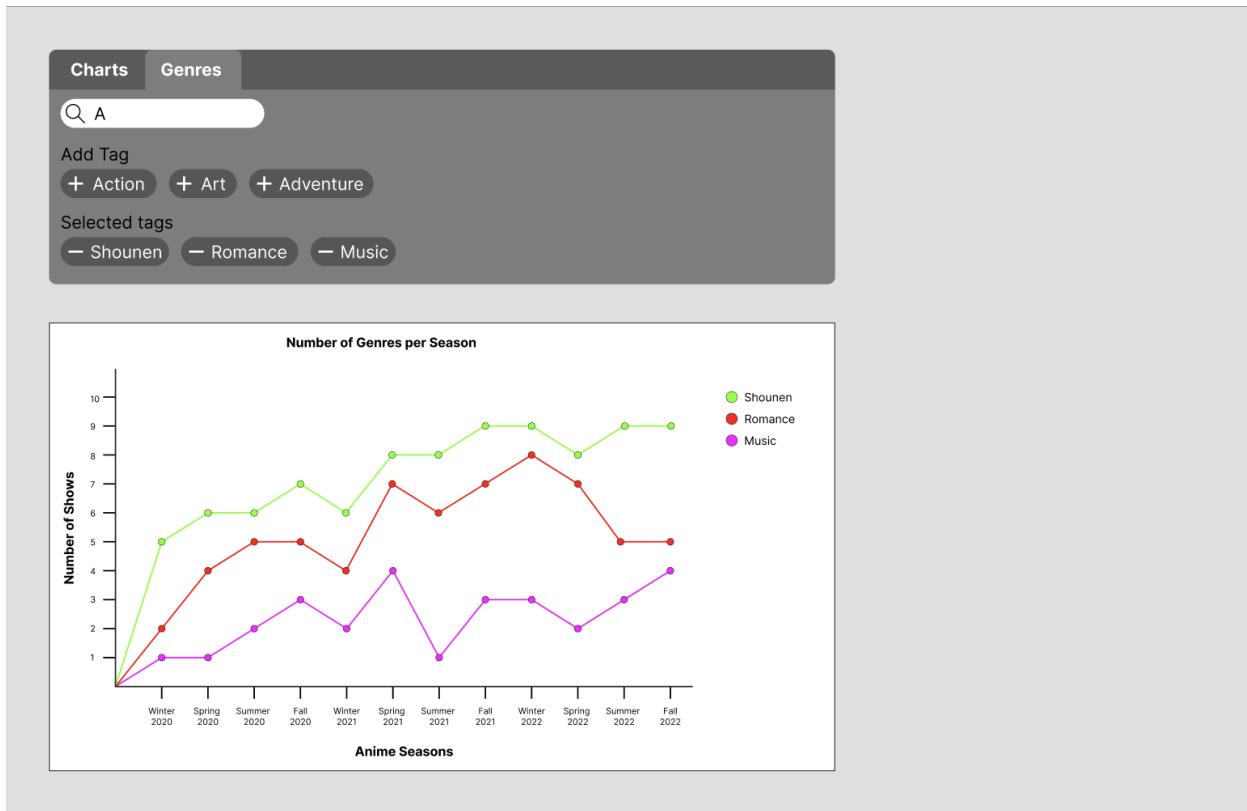
Shounen Seinen Slice of Life Cooking Seinen Slice of Life Cooking
 Fantasy Shoujo Action Music Shoujo Action Christmas
 Mystery Romance Art Cats Cars Art Cats
 Drama Adventure Sports Dogs Adventure Sports Dogs

Number of Genres per Season

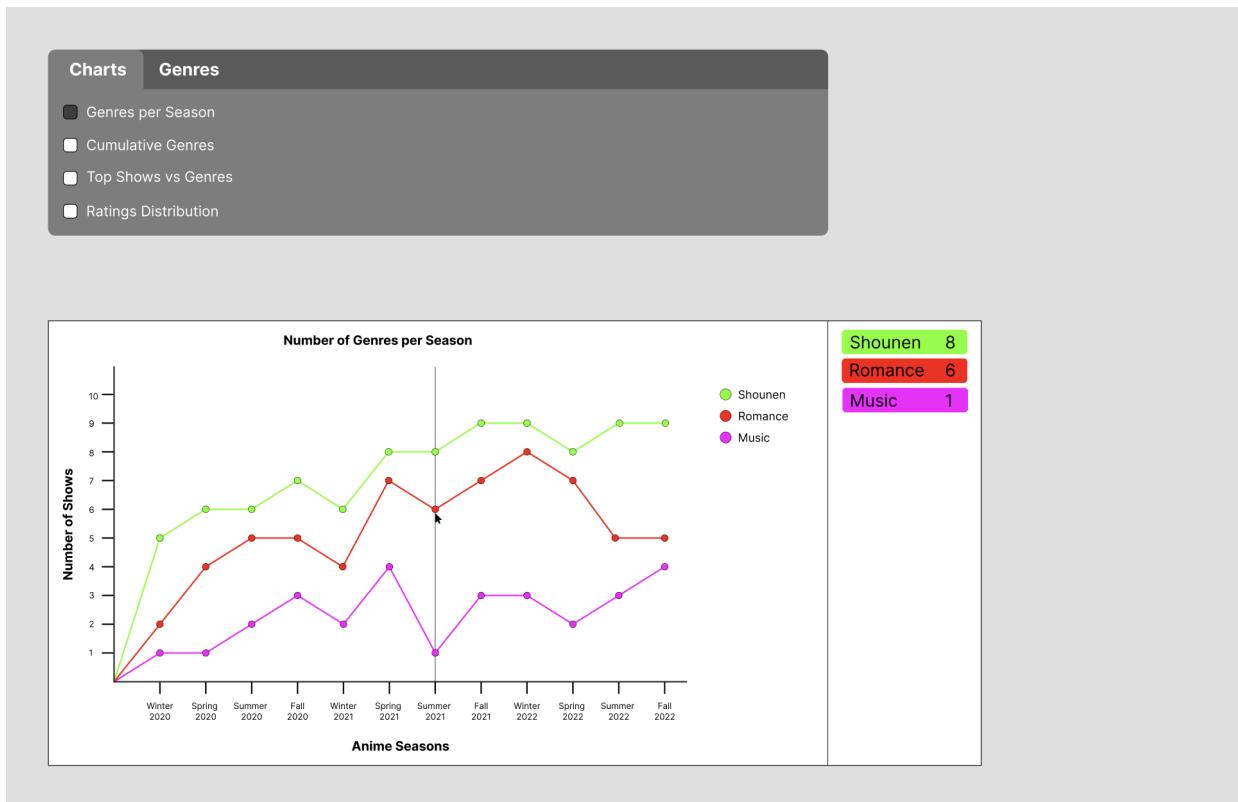
The graph tracks the following genres over time:

- Shounen**: Green line with circles, starting at 5 in Winter 2020 and ending at 9 in Fall 2022.
- Romance**: Red line with circles, starting at 2 in Winter 2020 and ending at 5 in Fall 2022.
- Music**: Magenta line with circles, starting at 1 in Winter 2020 and ending at 4 in Fall 2022.

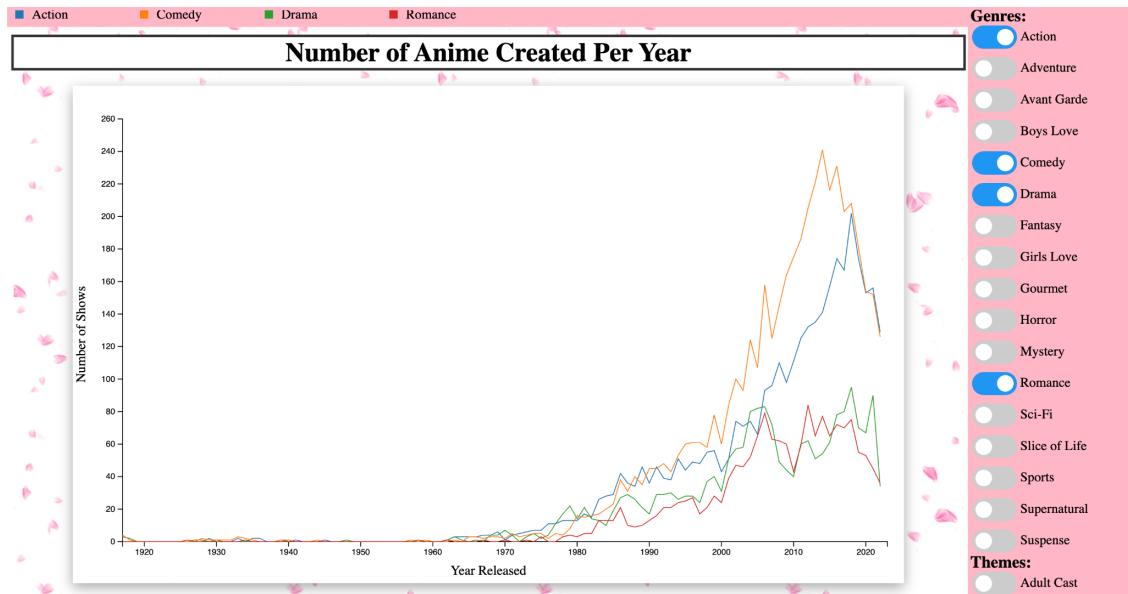
Season	Shounen	Romance	Music
Winter 2020	5	2	1
Spring 2020	6	4	1
Summer 2020	6	5	2
Fall 2020	7	5	3
Winter 2021	6	4	2
Spring 2021	8	7	4
Summer 2021	8	6	1
Fall 2021	9	7	3
Winter 2022	9	8	3
Spring 2022	8	7	2
Summer 2022	9	5	3
Fall 2022	9	5	4



Interactive Element:

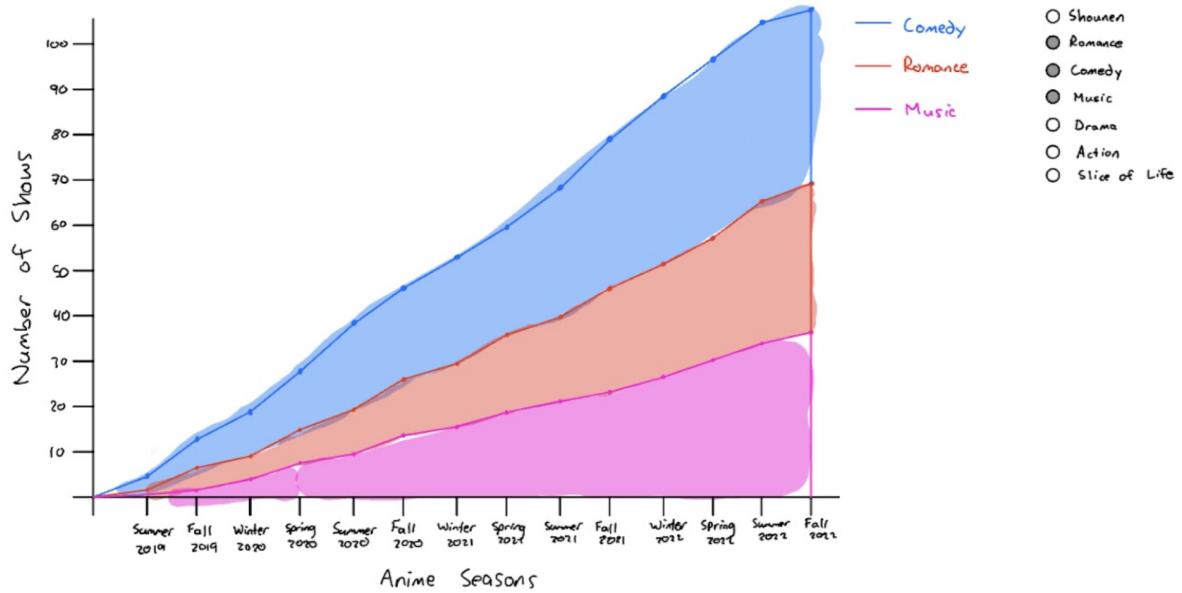


Final Project Submission Design:

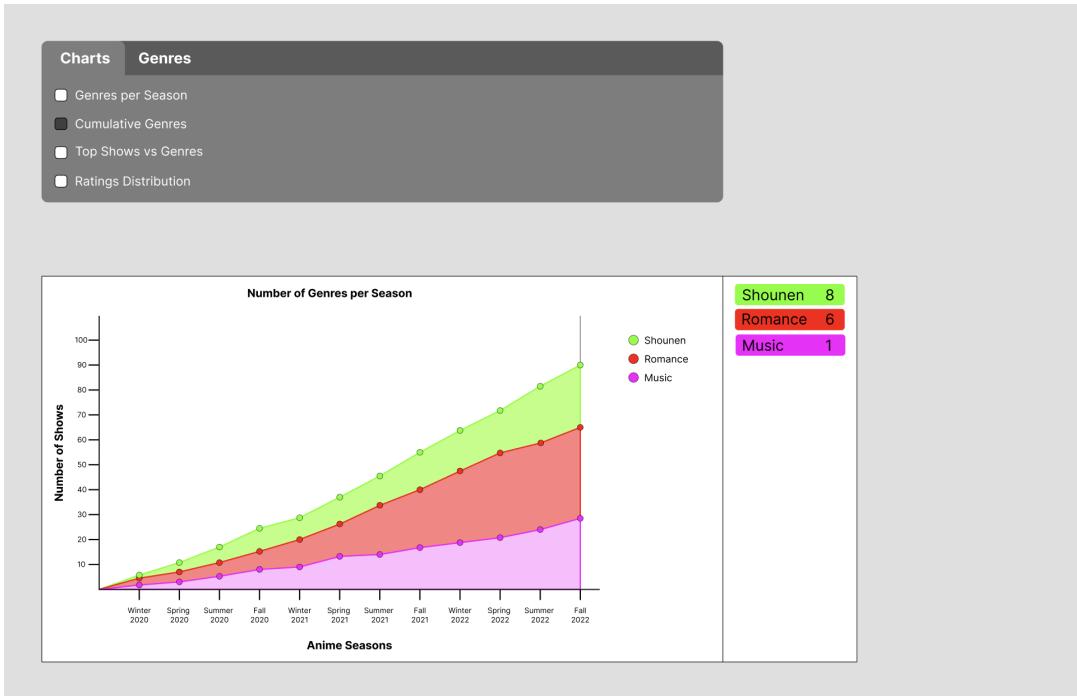


An **area graph** comparing different genres of anime and seeing how many were released per season *cumulatively*. Includes some way to filter out/see specific genres.

Proposal Design:



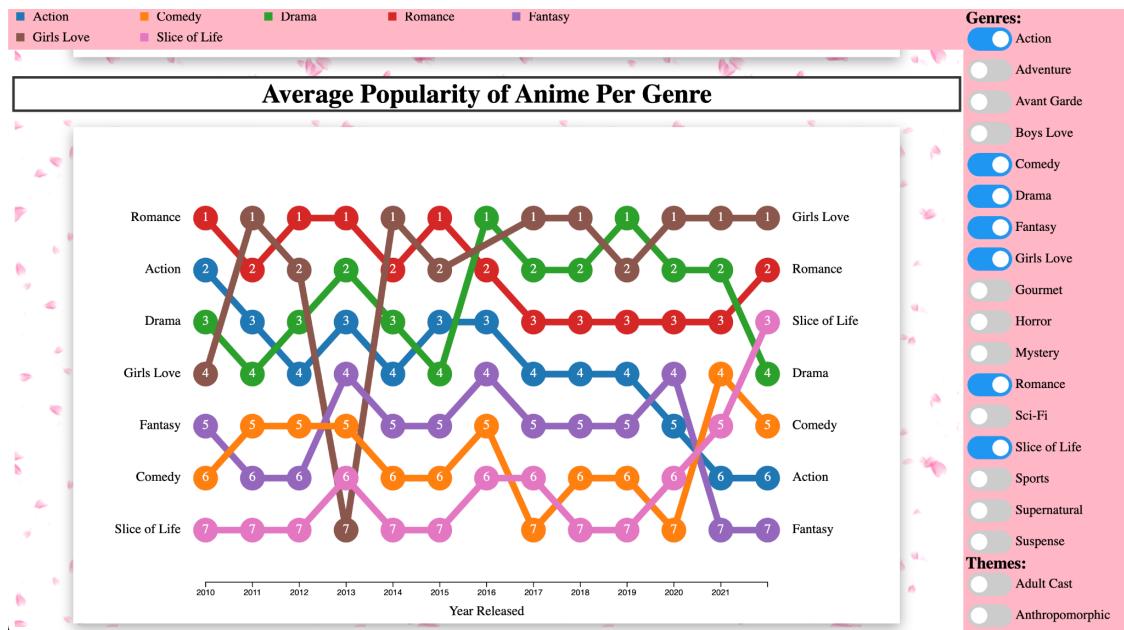
Milestone Design: We also added a separate tab with toggles that lets the user choose which graphs will display on the screen.



Final Project Submission Design: We thought about this graph in terms of what our visualizations show and figured that this graph didn't really add much in terms of questions answered, so we instead opted for a different visualization entirely.

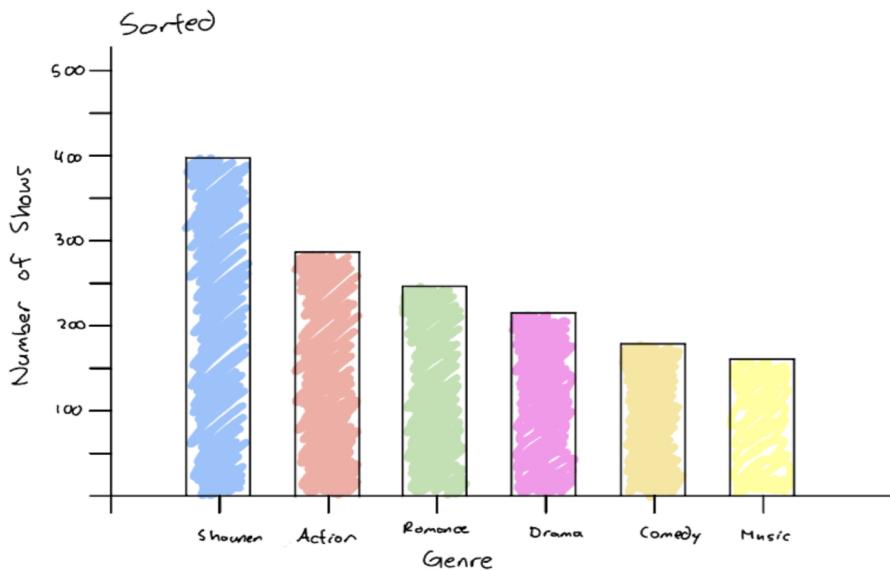
A **bump chart** that ranks genres based on the average popularity of the anime released in that year.

Bump Chart:



A **bar graph** showing the distribution of genres between the top 500 or so most popular anime. This will give a good visual as to what the most popular genres of anime are. Includes a toggle to show the top anime by rating rather than popularity.

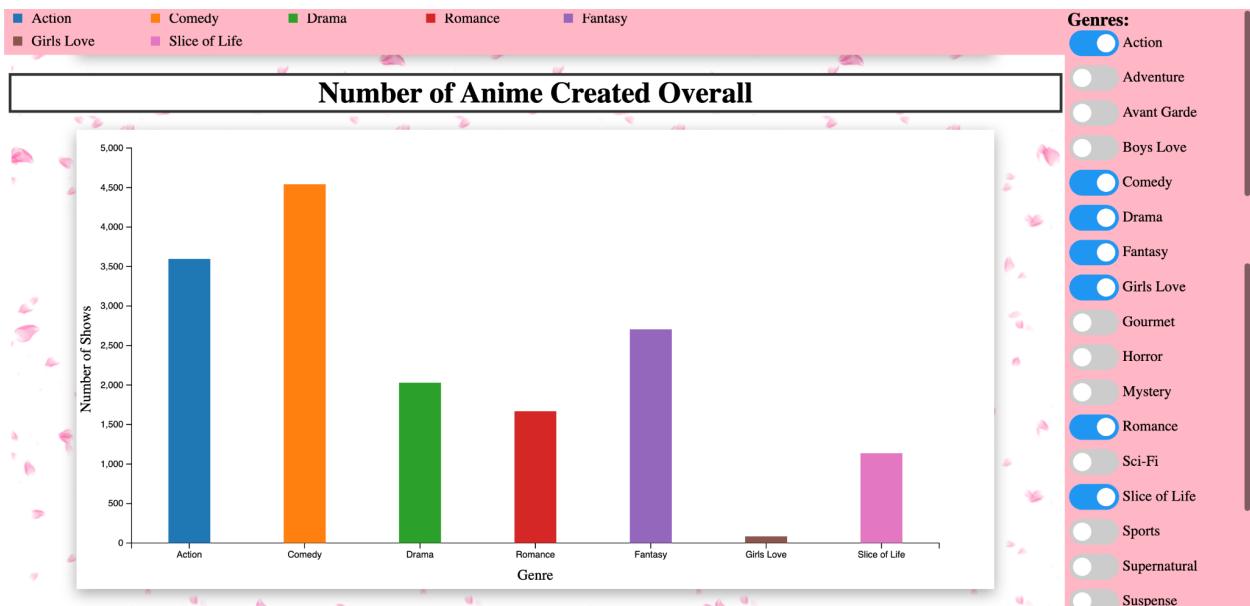
Proposal Design:



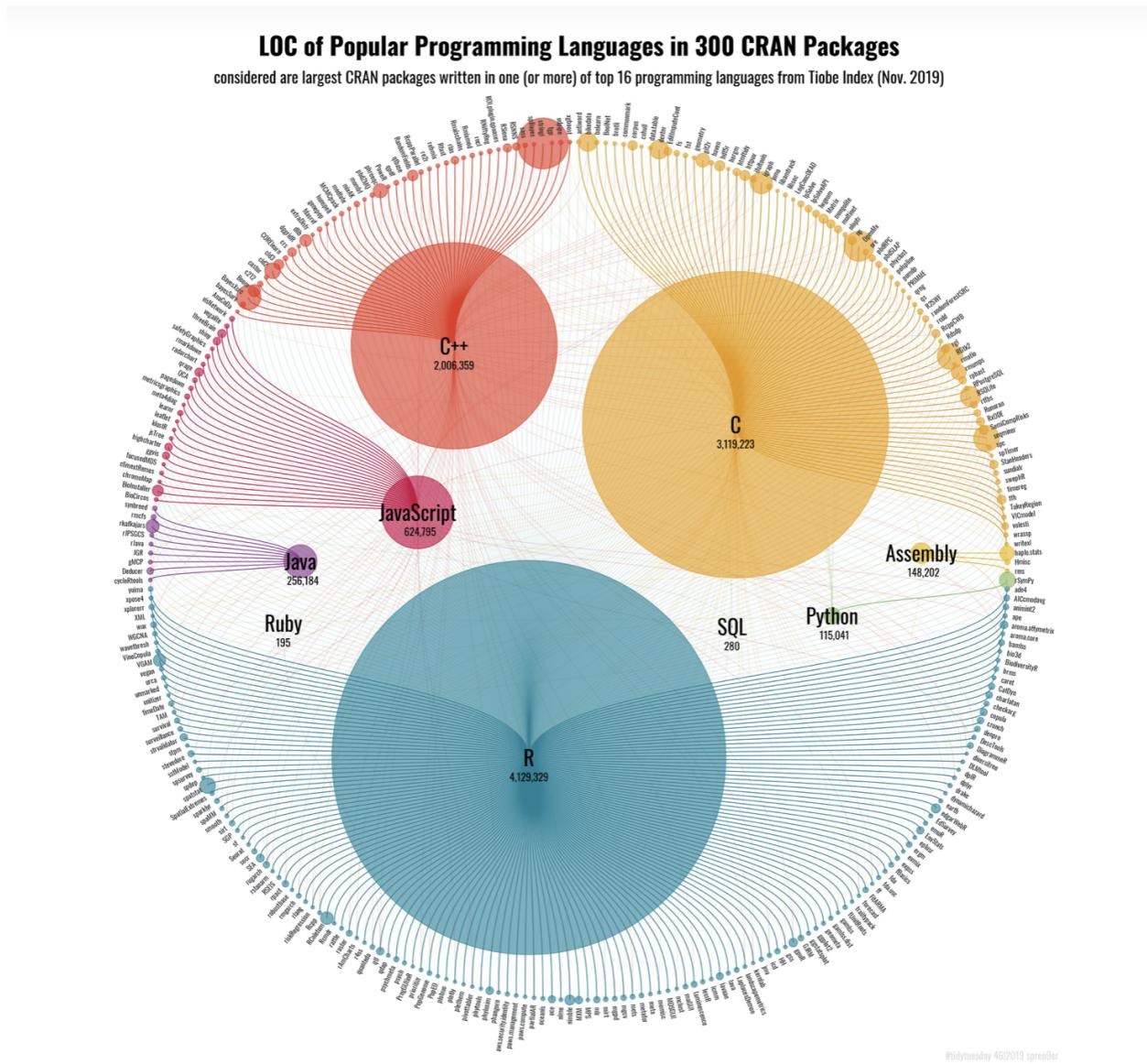
Milestone Design:



Final Project Submission Design:

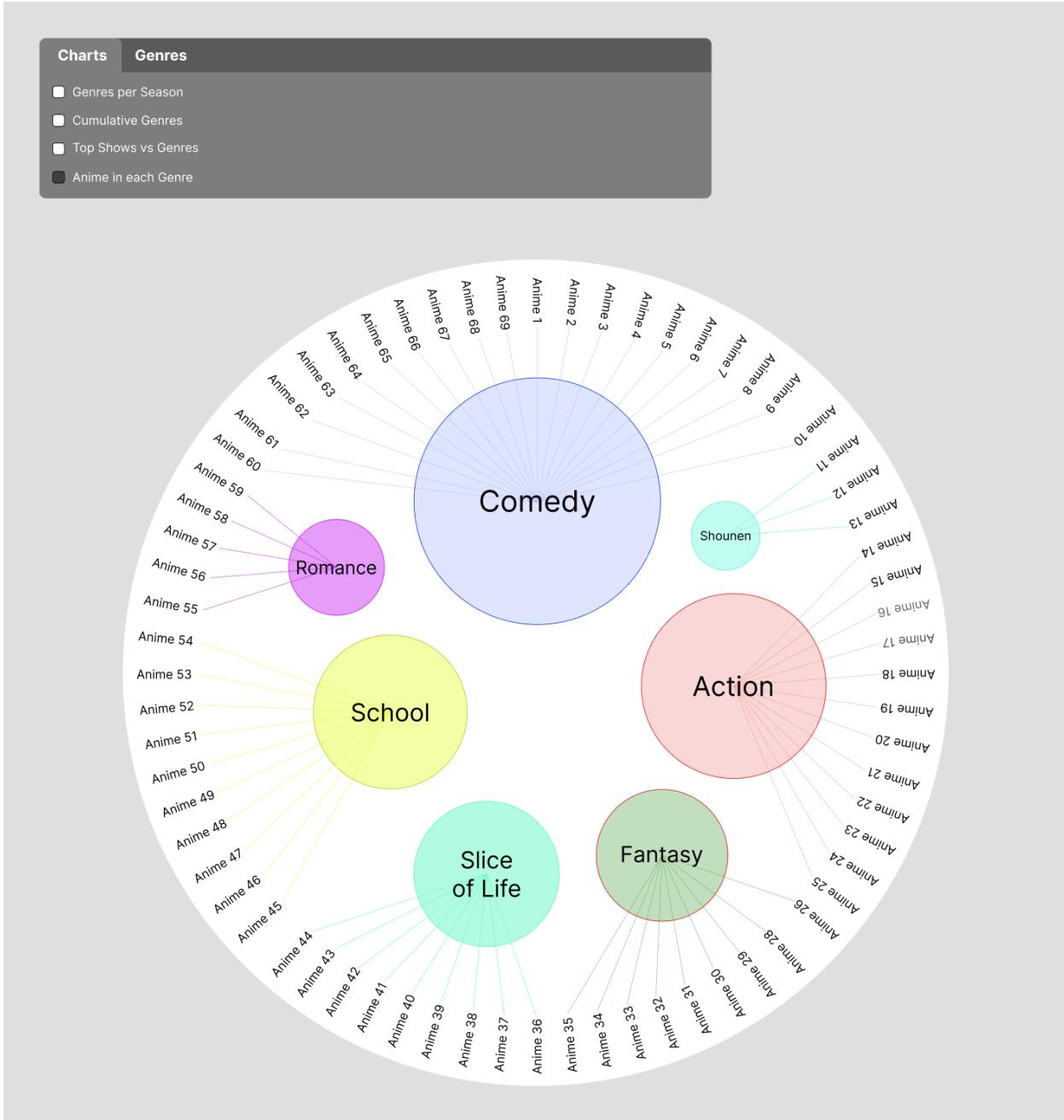


We found this graph design that we thought was really creative and could be implemented into our visualizations. For example, the inner circles could be genres and the lines connecting from the outermost circle could be the name of shows in that genre. This seems like it would be very difficult to implement so it might be something we look into later on if we have time.



<https://careerfoundry.com/en/blog/data-analytics/data-visualization-examples/>

Below is a very quickly drawn out sketch of what that might look like for our dataset:



In the end, we did not get to implement this chart but given more time we think it would still be something worth looking at for displaying our data.

6. Implementation

For our Final Project Submission, we ended up having 3 visualizations: A line chart, a bar chart, and a bump chart. Each chart has their own individual purpose for answering different questions about anime.

We have a legend at the top that keeps the color scheme consistent across all charts and also stays at the top of the screen when you scroll down. To the side we have interactive buttons that lets you filter the data that gets displayed. These selections stay consistent across all of our charts. Visualizations of our charts can be seen in the following pages.

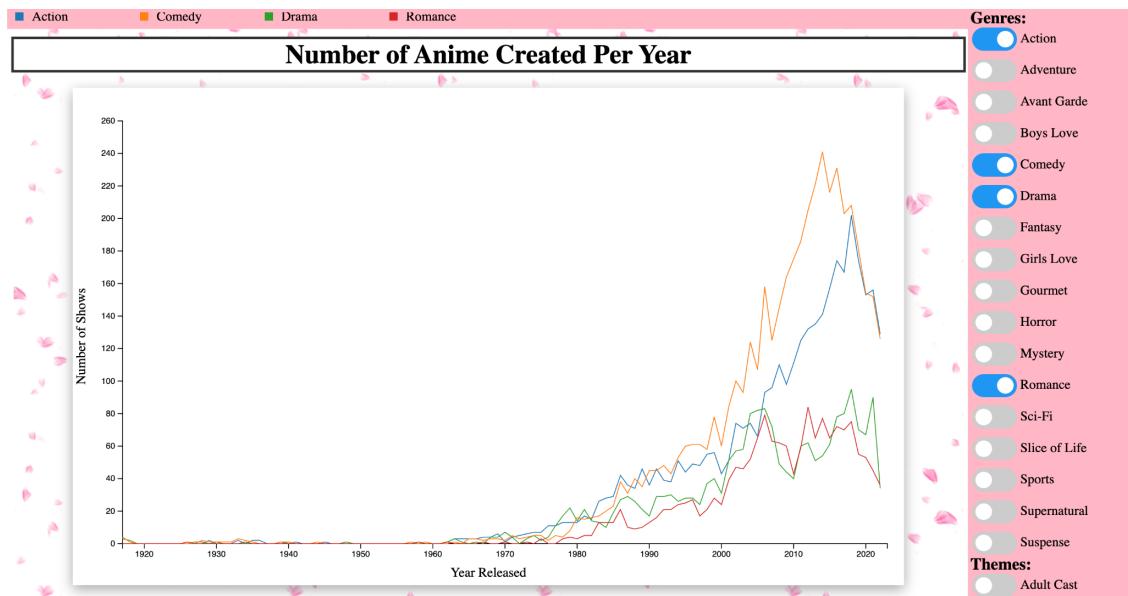
Our line chart tells us how many shows were created for a selected genre over the years and lets us see if there might be any correlation of shows created for genres over time. Our line chart's interactive element pops up a bar with the exact number of shows for each genre for the year hovered over.

We would have liked to have instead created a brush over this graph and tied its date selection to the bump chart we made, but we didn't have enough time to implement that feature.

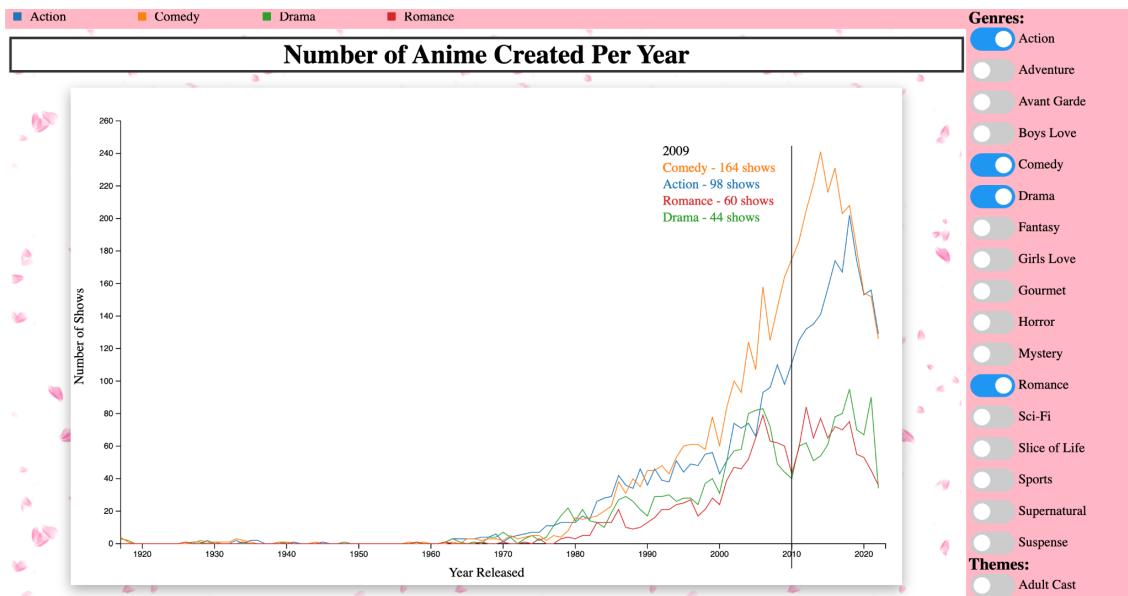
Our bar chart tells us how many shows there are overall for each genre selected. This lets us see which genre has the most shows ever created and what the overall catalog of anime would be filled with. Our bar chart's interactive element lets you hover over a bar and it will show the exact total number of shows for that genre at the top of the chart.

Our bump chart lets us see the average popularity of a genre over the past decade. This allows us to visualize how a certain genre might have become more or less popular over time and lets us see what people have started liking to watch more or less of. Our interactive element for the bump chart lets us hover over a circle/line for a genre which will then lower the opacity of all the other genre lines to make it easier to visualize the popularity of that genre over time.

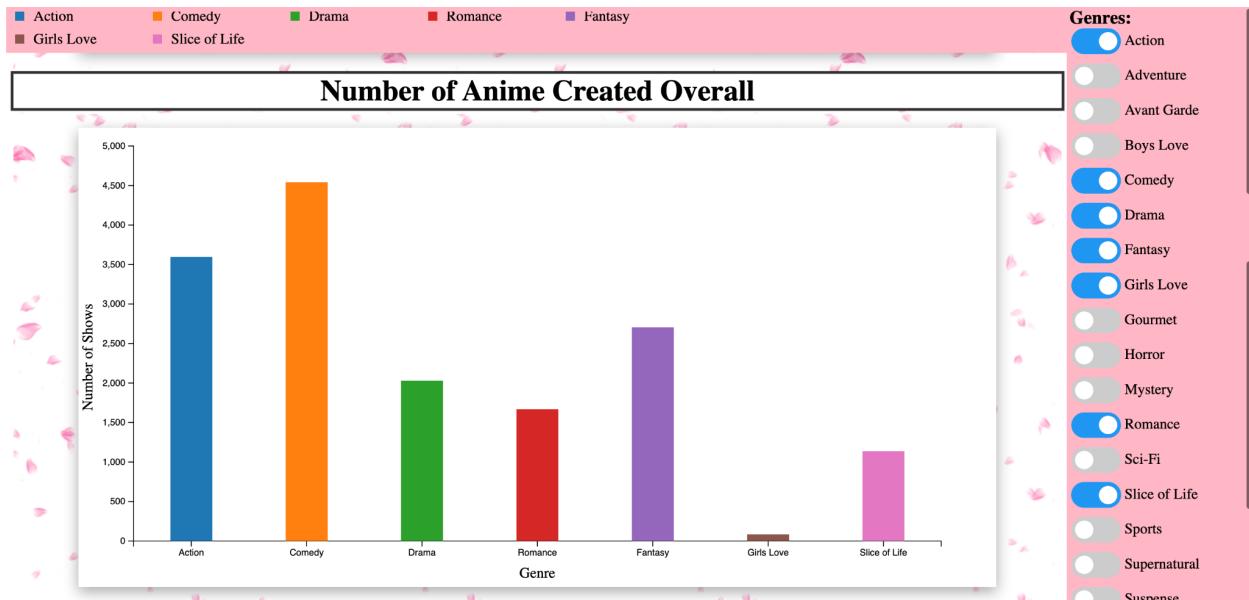
Line Chart:



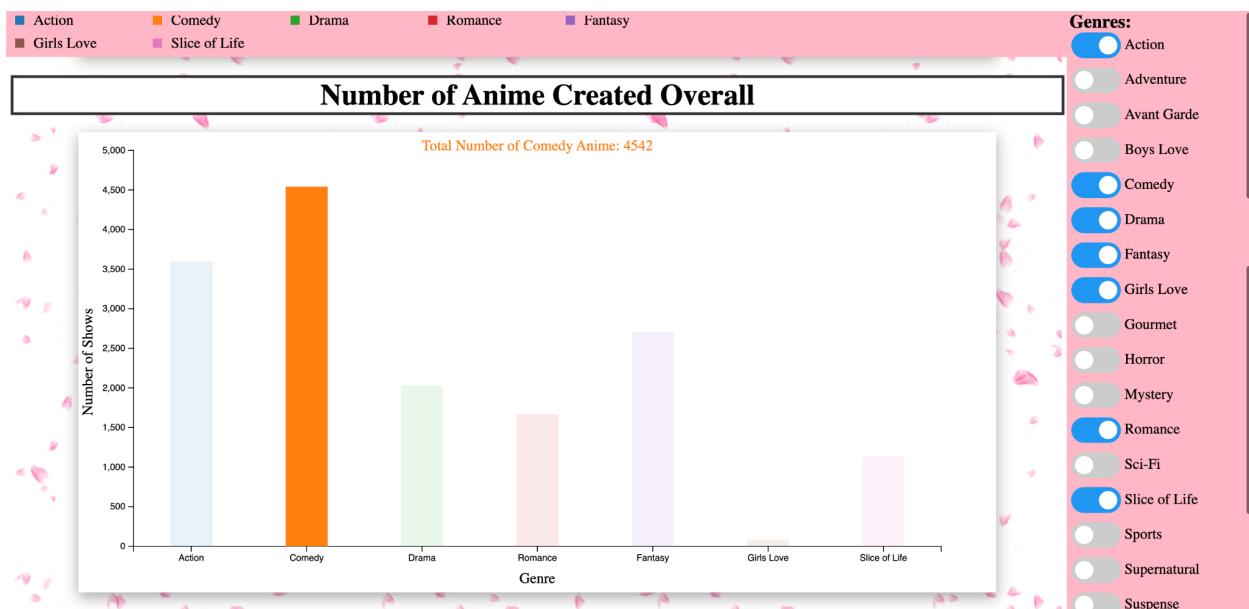
Interactive Element:



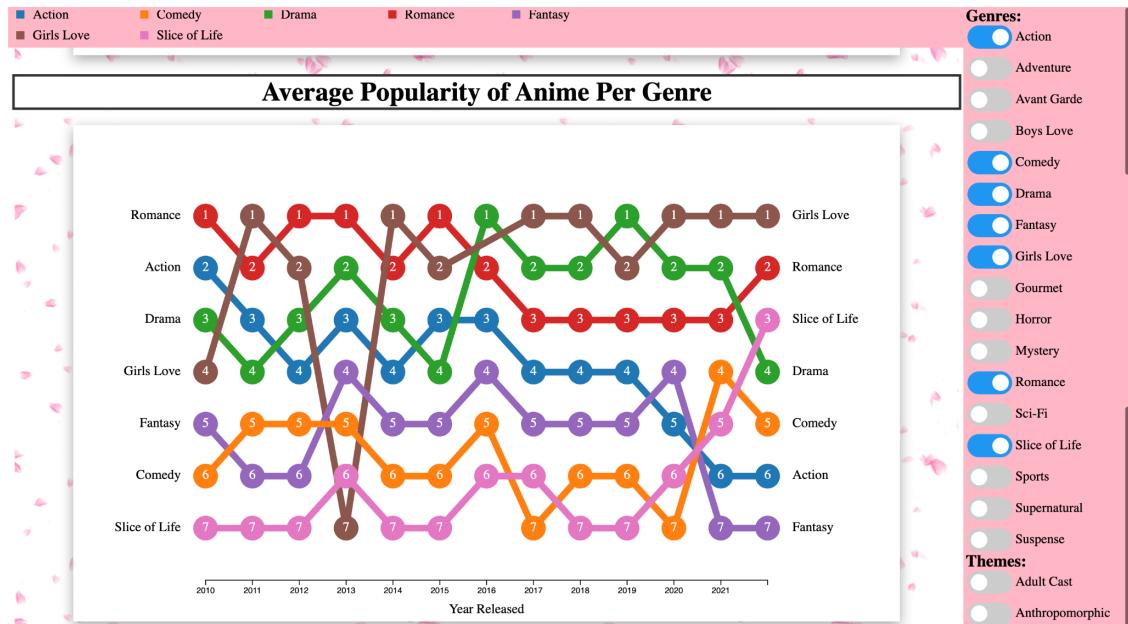
Bar Chart:



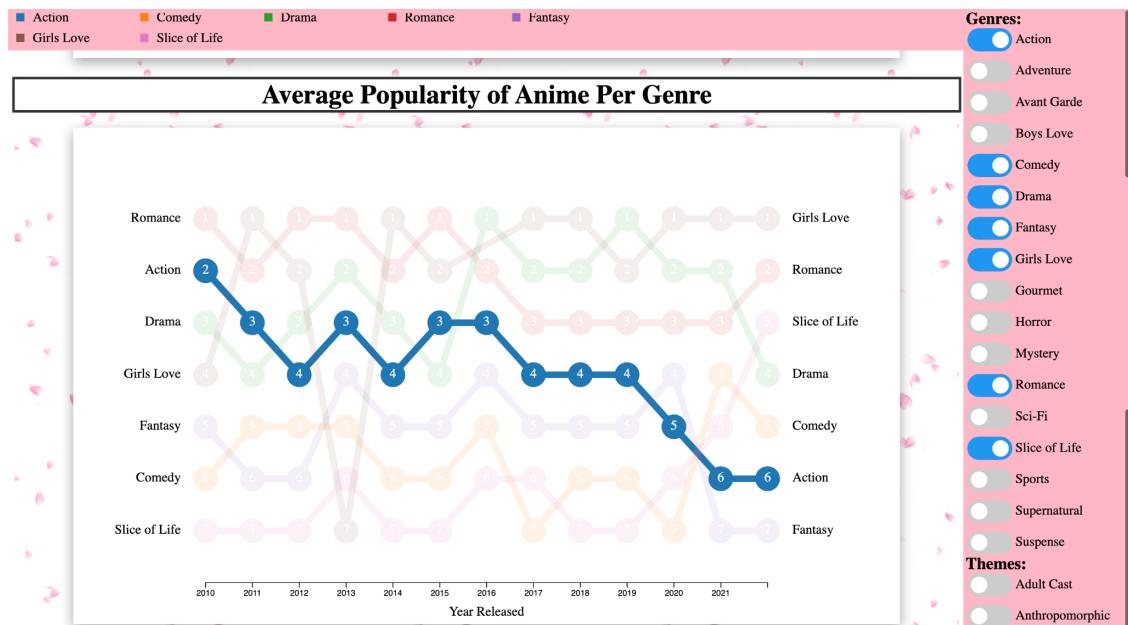
Interactive Element:



Bump Chart:



Interactive Element:



7. Evaluation

One thing we learned is how it can be hard to capture the ranking of the popularity of a genre. We decided that we would use an average of the popularity metric given by MAL, but it's not the best for all purposes, especially with some of the larger genres being affected by outliers more. If we were to come back and improve something, the way of ranking popularity would be one of the top things we would improve first.

Some other things we would like to improve include adding more visualizations and adding more interaction between visualizations. As previously mentioned with our line graph, we would've liked to connect that to our bump graph but couldn't add the functionality within our timespan.