Data Visualization
VizMoiCa

Process Book

Antoine Métivier (IN)
Julien Mettler (IN-CS)
Etienne Monnin (IN)

Our path

Since all of us are ***anime* enjoyers** at different degrees, we knew from the beginning that we wanted to base our project on this topic.

Finding the data and a story to tell

Since *anime* are a very popular form of entertainment nowadays, we were sure to find relevant and exhaustive datasets on this topic, so we started by looking on **Kaggle**. Our idea was to first explore existing datasets, and only then, once we got a clear view about the information they stored, we came up with a theme and proposed some ideas of visualizations that align with it. Since the first dataset contains information about users of **MyAnimeList** (MAL) and we already knew the global aspect of this social network, we decided to focus our main axis on the **worldwide influence of animes**, with MAL serving as a microcosm of this phenomenon.

Splitting the work

We decided to tell our story along **four main axes/visualizations**, and each of us assigned itself one of them and **sketched our ideas**, searching information on the lectures and on the web on how we could use **D3.js** features to implement them.

Preprocessing

Then came the preprocessing of the data, which essentially consisted of applying numerous operations to our datasets, **computing statistics** that we wanted to show on our graphs, **merging** some of our datasets together and **cleaning** them. For the studios dataset, we had to **scrap MAL website** ourselves, as there was no information online that covered our specific needs.

Our path

Implementing

With all the data computed, we coded each on our own the visualizations, applying what we had learned about **HTML/CSS/JavaScript/D3.js** from examples in the lectures, exercise sessions and tutorials from the Web. We also used libraries such as **Bootstrap** and **fullPage.js.**

Due to time constraints, we made the decision to omit the additional idea of implementing a recommender system, which we had previously discussed in Milestone 2.

Getting all the work together

Once our visualizations were ready, we decided to gather all our work into a single website. This task proved more difficult that expected, as there were some conflicts in the names of our variables and in the styles we defined. For the layout of the website, we had agreed at the beginning of the project to make it as a **single scrollable page**, which required to put our graphs one after the other, yet some graphs were not initially meshing well as they were slightly overlapping. The library **fullPage.js** proved very useful to solve these problems, as we could divide easily our pages into well-separate **sections**, and **transition smoothly** between each visualization.

At this point, we were not entirely satisfied by the look of our website, as we all used different fonts, colors, so we agreed on a white background and on a font to use everywhere, to create a uniform **visual identity** for our website.

As a last step, we added further sections for the home page, the footer and the texts that introduce each visualization, to guide the user into using them and navigating between the axes of our story.

Visualization 1
Circle packing

Motivation

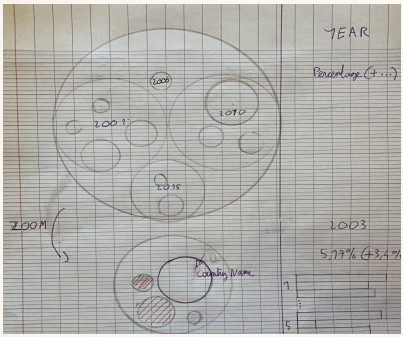
Since the beginning our objective for the first visualization has been to provide an **introduction** through the **history** of Anime. Our intention is to highlight its remarkable **growth** over time and to illustrate its importance as a **prominent medium** in today's cultural landscape.

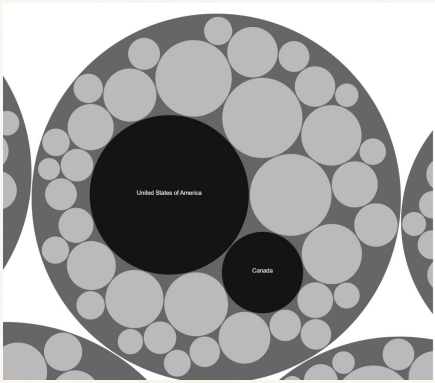
Challenges

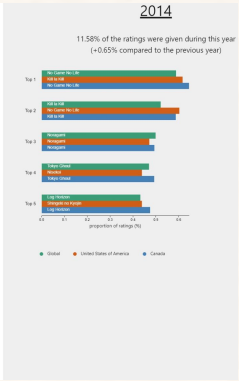
- Dealing with **lengthy names** of certain anime required finding a compromise.
- **Modifying the axis, legends and adjusting the width of each bars** when a circle is clicked was more complex than anticipated.
- Incorporating a **large amount of information into a single bar plot** presented a significant challenge.

Changes since Milestone 2

- **Anime names** are displayed within each respective **bar**.
- The **country name** is now shown on the **selected circle**.
- When no selection is made, the **number of ratings** and the **time period (From 2007 to 2018)** are now visible.







Visualization 2
Word cloud

Motivation

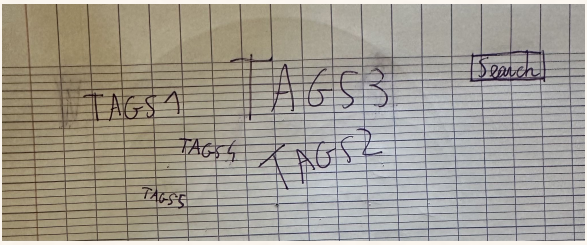
With the second visualization, our aim was to demonstrate that Anime have a **universal** appeal and that almost every character can be **relatable** to viewers from diverse backgrounds.

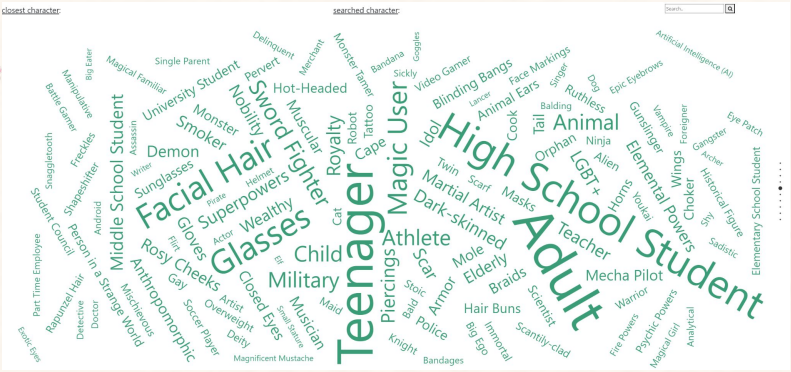
Challenges

- After milestone 2, we felt that the visualization lacked interactivity. It was then necessary for us to **introduce additional interactive elements** in a significant manner.
- Finding a way to **make the search bar practical**.

Changes since Milestone 2

We incorporated the functionality to **display the character with the highest Jaccard similarity** based on the selected traits, providing the user with the closest match.





Visualization 3
World Map

Motivation

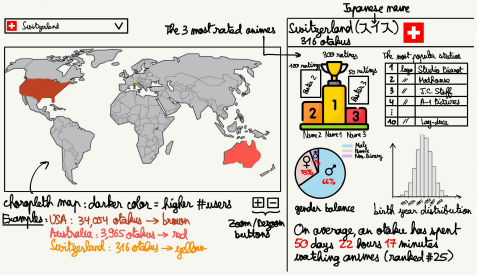
Our intention was to provide users with a **user-friendly** method to **compare otaku country trends** across multiple axes, reflecting specificities of their viewers, while also providing a clear understanding of how anime has become a **globally appreciated** form of entertainment.

Challenges

- The biggest challenge regarded the management of the **map dimensions** when zooming on a particular country, we encountered many display bugs, such as some parts of the map being cropped, some paths not being responsive to our clicks.
- As we were unable to find an existing dataset that compiled anime studios and their logos, we had to **manually scrape** the MAL pages of each studio ourselves.
- Managing the **logic of the country and studio tabs** proved tricky when it came to switch between the "country focus" mode (the user has selected or clicked on a country), and the "studio focus" mode (the user has selected a studio) and we had to handle many display bug. For instance, when the user selects a studio after having selected a country, the country tab should be cleared, and vice-versa.

Changes since Milestone 2

- Smallest countries can be hard to distinguish when we select one. We thus decided to **center and zoom on each country** whenever it is selected/clicked.
- As a result, we **removed the zoom and dezoom buttons**, as we thought that adding another zoom/dezoom feature would uselessly overload the user experience.
- We added a **reset button**, that sets the map to its initial state, i.e. disables the focus on a country or a studio, and acts also as a replacement for the dezoom button, as it resets the map back to its original size.
- We decided against including a legend in the gender pie chart since we assumed the male and female symbols are familiar to the users. Including a legend would have introduced redundant information.



Visualization 4
Genre Explorer

Motivation

We have developed this interactive graph with the aim of visualizing and exploring the **distribution of genres in anime**, offering users an immersive and informative experience. Our objective was to provide anime enthusiasts and fans with an effortless way to **discover trends and relationships among various genres** found in our extensive database, as well as to introduce them to **new anime with unfamiliar genres**. By utilizing this graph, users can enhance their understanding of **genre combination**s and their impact within the anime industry.

Challenges

- In the database, **the MAL links to the anime posters were not updated**, so I had to update them manually.
- I had trouble finding a **compromise on text size** because some anime have names that are too long.
- Anime usually have **genres that are interrelated**, handling this aspect thus posed a challenging task.
- In general, I found it difficult to manage the **size of svg and div files**, and how to resize them.

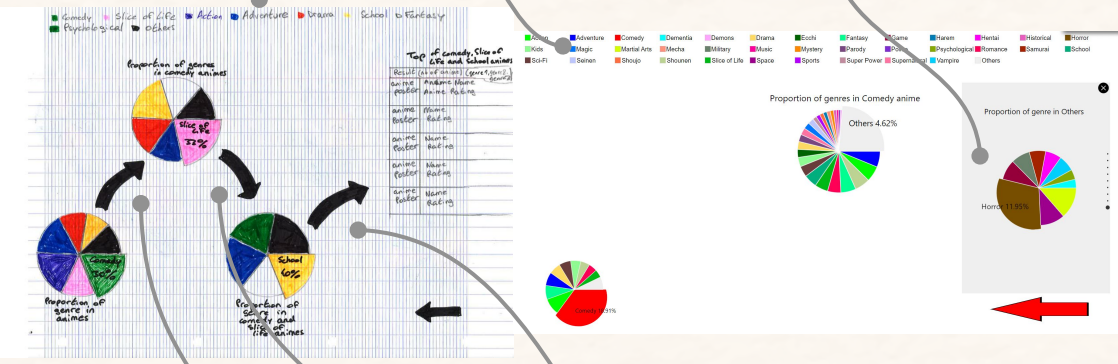
Inserted Ideas

Click on Legend

I added an enhanced feature to the graph legend, enabling users to click on the caption to **view the proportion of the corresponding genre in the current graph**. Double-clicking on the caption selects the genre. This improves navigation and allows users to explore and analyze specific genres of interest with ease.

Other proportion

I added a feature that lets users double-click on the **"Others" category** to identify **the ten least common genres within a specific anime genre**. This provides deeper insights into the genre distribution, highlighting less prevalent genres contributing to the "Others" category.



Arrows between subgraphs were omitted due to difficulties in creating visually appealing and appropriately sized arrows. We believe that including arrows would not significantly enhance the overall comprehensibility of the visualization. Considering design considerations, we concluded that the absence of arrows ensures a clear and easily understandable graph.

Dropped Ideas

Peer assessment

Throughout the entire project, our team consistently strived to collaborate and make all the significant decisions collectively. To facilitate this process, we diligently scheduled **two meetings every week**, where we engaged in thorough discussions and arrived at the most optimal choices. Here is a breakdown of the parts completed by each of us.

Etienne

I worked mainly on **visualizations 1 and 2**.

Preprocessing:

- Handling all data related to time and characters including the computation of all the relevant statistics.

Implementation:

- All components of visualizations 1 and 2

Julien

My work mainly focused on all the elements related to **visualization 3**.

Preprocessing:

- Geocoding the user locations, cleaning and merging the resulting dataset with the GeoJSON data.
- Computing the datasets of statistics for each country and studio.
- Scraping the studios data on MyAnimeList.

Implementation:

- Coding all the visualizations and interactions of the world map and its subgraphs for the countries and studios statistics.
- Adapting our website into a single scrollable page with fullPage.

Finally, I took care of the screencast.

Antoine

In general I took care of the part related to genre data and **visualization 4**.

Preprocessing:

- Computing the genre proportions in the anime dataset, identified the most common genre, and created a pie chart to visualize the genre distribution.
- Computing various statistics, e.g., the number of anime per genre.
- Updating the website by adding anime images.

Implementation:

- Developing visualizations and interactive features for the genre graph and its subgraphs, including the top anime section.
- Crafting a transition text introducing the genre graph.
- Incorporating a navigation tool for seamless full-page scrolling.