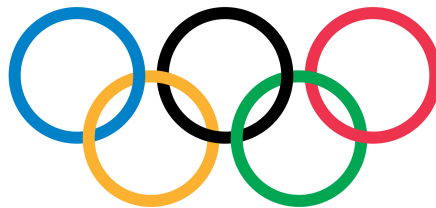




## Ring of Time : An Olympic Legacy



Louis Duval  
Tanguy Decléty  
Léa Blinière

Professors: Vuillon Laurent Gilles Marie

Data Visualization - COM - 490

Spring 2024

École Polytechnique Fédérale de Lausanne

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Presentation of the Project and Its General Objective . . . . .	1
<b>2</b>	<b>Data</b>	<b>1</b>
2.1	Data Kaggle . . . . .	1
2.2	Data Scrapping . . . . .	1
2.3	Additional Data . . . . .	2
<b>3</b>	<b>Design and Development Process</b>	<b>2</b>
3.1	Challenges . . . . .	2
3.2	Design Decisions . . . . .	3
<b>4</b>	<b>Final Visualization</b>	<b>3</b>
4.1	Software and Libraries Used for Visualization . . . . .	3
4.2	Organization of the Website . . . . .	3
4.3	Description of the Final Visualizations . . . . .	4
<b>5</b>	<b>Peer Assessment</b>	<b>6</b>
5.1	Louis Duval . . . . .	6
5.2	Tanguy Decléty . . . . .	6
5.3	Léa Blinière . . . . .	6
<b>6</b>	<b>Conclusion</b>	<b>7</b>
<b>7</b>	<b>Appendix</b>	<b>8</b>
7.1	Interactive Map . . . . .	8
7.2	Historical Timeline and Dynamic Table of the Olympics . . . . .	9
7.3	Gender Representation Over the Years . . . . .	10
7.4	Evolution of Disciplines Over the Years . . . . .	10
7.5	Medals Distribution Over the Years . . . . .	11



# 1 Introduction

## 1.1 Presentation of the Project and Its General Objective

The idea for our project was born from the media enthusiasm surrounding the 2024 Olympic Games. We realized that, amidst this influx of information about the upcoming Games, few people were truly aware of their history, creation, and evolution over time. Our online research revealed that searching for "history of the Olympics" mainly leads to the Wikipedia page [1], which, although comprehensive, remains very traditional with a straightforward presentation and often cumbersome tables. The official Olympic Games website [2] also dedicates very little space to their history.

Our objective is to make the history and evolution of the Olympic Games accessible to everyone through engaging and visually appealing animations. This project proved to be vast, covering several important aspects. We explored the distribution of host cities worldwide with an interactive map, traced the evolution of Olympic venues with a dynamic timeline, and showcased the increase in the number of disciplines, from 10 to nearly 49 as we approach the Paris 2024 Games. Additionally, we highlighted the evolution of women's participation in sports competitions and analyzed sports performances over time.

These elements have been integrated to create an interactive, aesthetically pleasing, and user-friendly website, aiming to inform a broad audience that the Olympic Games are much more than just a sports competition. They are a showcase of our world in many aspects.

## 2 Data

### 2.1 Data Kaggle

**Olympic Data [3]:** This dataset was obtained from Kaggle, a leading platform in the machine learning and data analysis community. It compiles comprehensive information on the Olympic Games, boasting a perfect "usability" score of 100%. This score indicates high data quality in terms of completeness, credibility, and compatibility, underscoring the dataset's reliability. Its extensive download count within the data science community reflects its value for analysis and visualization efforts in our Olympic Games study. The dataset covers modern editions of the Summer and Winter Olympics from 1896 to 2016, excluding data related to the Tokyo 2021 Games and the Paralympic Games. It includes detailed metadata for each participant, such as unique identifiers (ID), demographic details (age, height, weight), team affiliations, National Olympic Committee (NOC), Games season (summer or winter), host city, sports, specific events, and, where applicable, medal achievements.

### 2.2 Data Scrapping

To supplement our primary dataset, we scraped additional data from the internet:

**French Wikipedia Olympic Games page [4]:** We extracted and translated tables from the French Wikipedia page into English. These tables had the advantage of containing the correspondence of the host city to its country and continent, along with the year, which was missing in our primary dataset. This allowed us to develop the Olympic Games timeline and its associated dynamic table, showing the evolution of the number of Games by city, country, and continent for both the Summer and Winter

Olympics.

**Olympedia performance data :** Inspired by Joseph Cheng’s data scraping method on Olympedia [https://github.com/josephwccheng/olympedia\\_web\\_scraping/tree/main](https://github.com/josephwccheng/olympedia_web_scraping/tree/main), we were able to retrieve many performance metrics for Olympic Games events. These kind of information like time taken in a track event or points made during an gymnastic event were missing from our original dataset. Many more information were taken as Olympedia is an extensively detailed website made by dedicated Olympic historians and statisticians, known as MADmen, but were filtered out for storage and execution purposes. A notebook, named `datascrap.ipynb`, shows how it the scrapping was done while another notebook `datapreproc.ipynb` describes the cleaning and filtering of our different datasets. The cleaning and filtering is far from being perfect due to the multiple different layouts and titled columns present on the website. The dataset (without the performance metrics) was also available on Kaggle, holding a usability score of 10.00 [?].

## 2.3 Additional Data

**Interactive Map Development:** To develop the interactive map on our main page, we extracted data from an open-access world map demonstration [5]. This included the coordinates of each path, allowing us to trace our SVG map as interactively as possible. Additionally, we extracted the exact longitude and latitude for each Olympic host city and its corresponding position on the map, ensuring accurate placement.

**Historical Timeline Development :** For creating a visually appealing historical timeline, we extracted the ISO codes for each country. This allowed us to utilize an API [6] to display the corresponding country flags, enhancing the visual aspect of our timeline.

**Country Coordinates Data:** To accurately visualize the gender representation map, we utilized the ‘countries.geojson’ file. This file provided the precise geographical coordinates for each country, enabling us to create a detailed and accurate world map. By integrating this data, we ensured that each country’s representation on the map corresponds correctly to its real-world location, thus enhancing the reliability and visual quality of our interactive map.

## 3 Design and Development Process

### 3.1 Challenges

**Geopolitical Changes Over Time :** The modern Olympic Games span over 100 years, during which the world map has undergone numerous changes. Some countries have changed names, ceased to exist, or have been divided or unified. This posed a significant challenge, especially when comparing the performance of countries in the Olympics or representing them on the map.

**Variety of Topics to Cover :** The Olympic Games impact society in many different ways. We aimed to address as many of these aspects as possible, such as the inclusion of women and the internationalization of the Games. However, we decided not to cover the Paralympic Games despite their importance in promoting inclusivity in sports. This decision was made due to time constraints and the vast scope of the analysis.

**Map Design :** Designing the map of the homepage was a significant challenge because many host cities are concentrated in certain areas, particularly the Winter Olympics around the Alps. To address

this, we implemented a zoom feature by region on the map. This was a major technical challenge since the map is entirely designed using SVG paths and coordinates.

**Performance Metrics & World Records :** Comparing and retrieving compelling metrics was also a hindrance. Indeed the data was stored in many formats and different columns, was sometimes missing or not really indicative. That is why certain choices were made to mainly represent time metrics aswell as points while many events may have missing or confusing values. Also getting the metrics for world records was deemed too challenging to do and maybe a bit out of topic. This led to us to replace this idea by representing new Olympic Records in track and swimming events instead.

**Animating simulation :** We also had the objective to display a few different animations for athletics or swimming events unfortunately dealing with the dataset and implementing such visualization was pretty difficult and time consuming. We only opted for a one kind of "Race" visualization.

## 3.2 Design Decisions

**Flag Representation :**As mentioned, geopolitical changes over time presented a real challenge, particularly in representing flags. The API we used for flags only includes the current countries' flags. Therefore, we decided to represent the host cities with the flag of the current country where the host city is now located.

**Geopolitical Impact :** The Olympics have often been used as a platform to address global issues. Initially, we intended to use our timeline to represent some of these geopolitical events to highlight their societal impact. However, the sheer number of events to represent was overwhelming. Selecting only a few would have been subjective, and representing all of them would have significantly complicated our design. Thus, we decided to represent only the events that directly impacted the organization of the Olympics, such as cancellations or postponements.

# 4 Final Visualization

## 4.1 Software and Libraries Used for Visualization

To achieve our project's visual and interactive goals, we employed a variety of tools and libraries:

**JavaScript (D3.js and SVG):** JavaScript, along with the D3.js library and SVG, was crucial for developing the interactive map, the historical timeline, and the dynamic tables. D3.js allowed us to create data-driven visualizations with precision and flexibility, while SVG provided the necessary scalability and crisp visuals.

**Python:** We used Python extensively for data scraping and manipulation. Libraries such as Beautiful Soup facilitated the extraction of data from web pages, while Pandas and Numpy were indispensable for handling and processing our data efficiently. For the [Olympedia.org](https://olympedia.org) data-scraping other libraries were used such as aiohttp for asynchronous HTTP requests and asyncio for writing concurrent code using the `async/await` syntax.

## 4.2 Organization of the Website

Our website is divided into 4 distinct sections:

**Main Page** : Homepage of our site, this is where you'll find the interactive map of the host cities, described in detail below.

**Olympics**: This section provides an overview of the **medal distribution** 8 for every edition for the top 10 countries ranked according to the number of Gold medals then the number of Silver medals if equal in gold medals, and finally the number of Bronze medals to decide in case of a draw in gold and silver medals.

Just under this distribution a **podium visualization** ?? is displayed where you can select the Game's edition, disciplines and specific events to discover the medalists. The source of the data is also added under the visualization if you want to check the event in more details. In case of group events, most of the time, the group corresponding country is displayed.

Finally two pretty similar **performance simulation** are available for track and swimming final events ???. You can select what specific final event you want to see and a simple animation will present the corresponding event. As a little bonus if a Olympic Record has been beaten you will be able to see it ??.

This page allows you to simply and quickly find out performance metrics and rankings of any event you'd like.

**History**: This section provides a comprehensive overview of the Olympic Games history through various visualizations. It begins with a chronological timeline showcasing the evolution of host countries, supported by a dynamic table that counts the number of Olympics held by continent and country. Following this, a gender representation map illustrates the progress towards gender equality in Olympic participation, providing insights into the advancements and challenges faced by different countries. The evolution of disciplines is depicted through a tree map, revealing changes in the number of athletes participating in each sport. Lastly, the medals distribution graph highlights the dominance and emerging competitors in the Summer Olympics over time. Together, these visualizations offer a detailed and engaging exploration of the historical aspects of the Olympics.

**About Us**: This page gives a brief description of the project behind the site and introduces the team members.

## 4.3 Description of the Final Visualizations

**Interactive Map of Host Cities** This interactive map serves as the entry point to our website, immediately highlighting the global aspect of the Olympic Games. On this world map, you can observe the locations of host cities for the modern Olympics. The Summer and Winter Olympics are distinguished by color coding (orange for Summer and blue for Winter), which is consistently used throughout the site.

The map, developed in SVG, is highly interactive. When you hover over a region, such as North America 1, it changes color and its name is displayed. A simple click zooms into that region, displaying only that part of the world. Similarly, in the zoomed-in view, hovering over a country reveals its name and changes its color 2. A "back" button appears below the map in zoom mode to return to the global view.

In any mode (zoomed or not), hovering over a host city displays essential information such as its name, the country, and the year(s) it hosted the Olympics 3. This map provides a straightforward and effective way to observe the distribution of host cities around the world, highlighting the concentration in North America and Europe (particularly in the Alps for the Winter Olympics due to obvious geographical reasons).

**Historical Timeline of the Olympics** The history page of the Olympics opens with a chronological timeline showing the evolution of host countries over time. This timeline is accompanied by a dynamic table that displays the number of Olympic Games held per continent and country for both the Summer and Winter Olympics 4.

A slider bar allows users to move through different time periods, synchronized with a red marker indicating the current position on the timeline. To maintain consistency and clarity, the same color coding is used throughout: Summer Olympics are represented by an orange bar, while Winter Olympics are denoted by a blue bar, matching the colors in the table below.

When hovering over a flag on the timeline, detailed information is displayed, including the host city, host country, year, and the number of times the city has hosted the Olympics. Some bars are dashed to indicate that those Olympics were canceled (notably during wars), and a note accompanies these entries to explain the reason for the cancellation. Other events, such as the Tokyo 2021 Olympics, include notes explaining why they were postponed.

This timeline provides a simple and interesting visualization, allowing users to observe the evolution of the Olympics over time. It highlights the later introduction of the Winter Olympics and the alternating schedule between the Summer and Winter Games. The accompanying table complements this visualization by showing the count of host cities by continent and country over time.

**Gender Representation Over the Years** The gender representation map provides a visual representation of the ratio of female to male athletes from different countries in the Summer Olympics over the years. Utilizing a color scale, the map highlights the progress towards gender equality in Olympic participation, with darker colors indicating a balanced representation. This visualization is crucial for understanding the advancements and challenges in achieving gender parity in sports. By tracking these changes, we can identify which countries have made significant progress and which are lagging, thereby highlighting the global disparities in gender representation. *This map serves as a powerful tool to analyze the effectiveness of policies aimed at promoting gender equality in sports worldwide.*

**Evolution of Disciplines Over the Years** The evolution of disciplines is depicted through a tree map, which reveals the number of athletes participating in each sport for a specific year. This visualization helps in understanding how the popularity of different sports has changed over time. Larger blocks represent sports with higher participation, offering a visual representation of their prominence in the Olympic Games. Observing these trends allows us to see the introduction and growth of new sports, as well as the enduring popularity of traditional events. It also highlights shifts in athlete interest and investment in various sports, providing a broader picture of how the Olympic program has evolved to reflect changing athletic priorities and global interests.

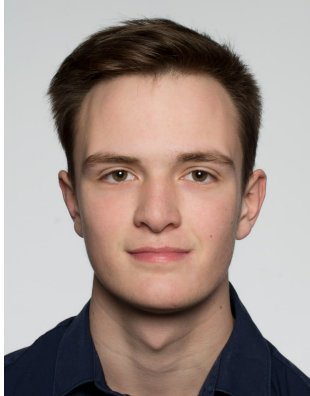
**Medals Distribution Over the Years** The medals distribution graph offers a comprehensive view of how medals have been awarded across different countries in the Summer Olympics from 1896 to 2016. Each colored area in the plot represents the cumulative number of medals won by a nation or group of countries, clearly indicating the dominance of certain nations over time. This visualization is particularly insightful as it shows not only historical powerhouses but also emerging competitors in the Olympic Games. By studying these trends, we can better understand the changing dynamics of international sports competition and recognize periods of significant achievement for various countries.



## 5 Peer Assessment

In this section, we detail the contributions of each team member to the project.

### Team Members



Louis Duval



Tanguy Decléty



Léa Blinière

#### 5.1 Louis Duval

Responsible of the Olympic page including the data-scraping and processing of the metrics data.

#### 5.2 Tanguy Decléty

#### 5.3 Léa Blinière

Contributions include the development of the main page (particularly the interactive map), the About Us page, the historical timeline and its associated dynamic table in the History page, as well as the data collection and preprocessing related to these components.

# 6 Conclusion

## 7 Appendix

### 7.1 Interactive Map

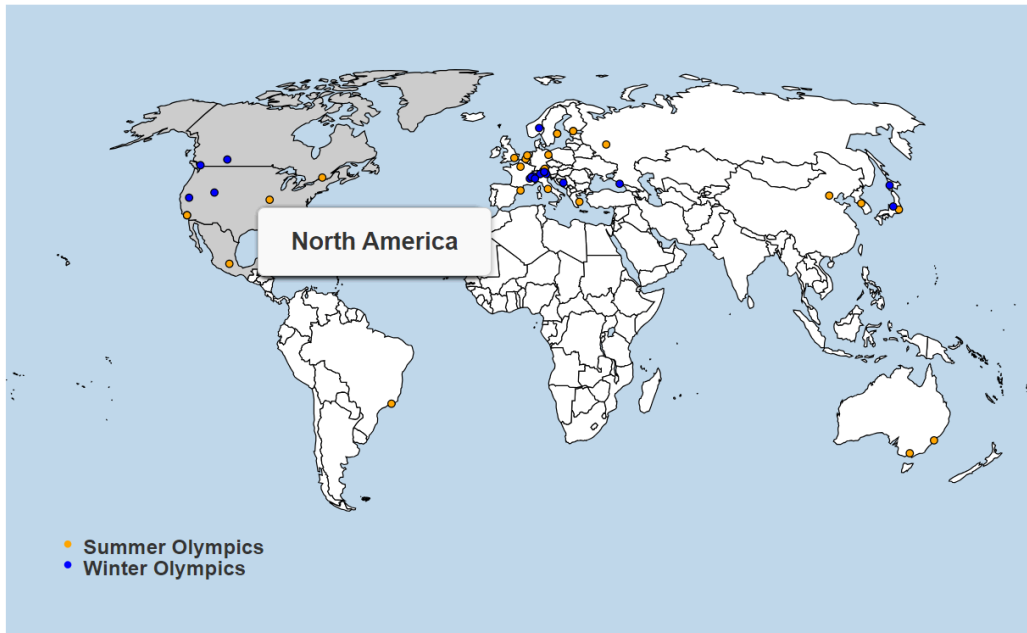


Figure 1: Map showing the interaction with a region.

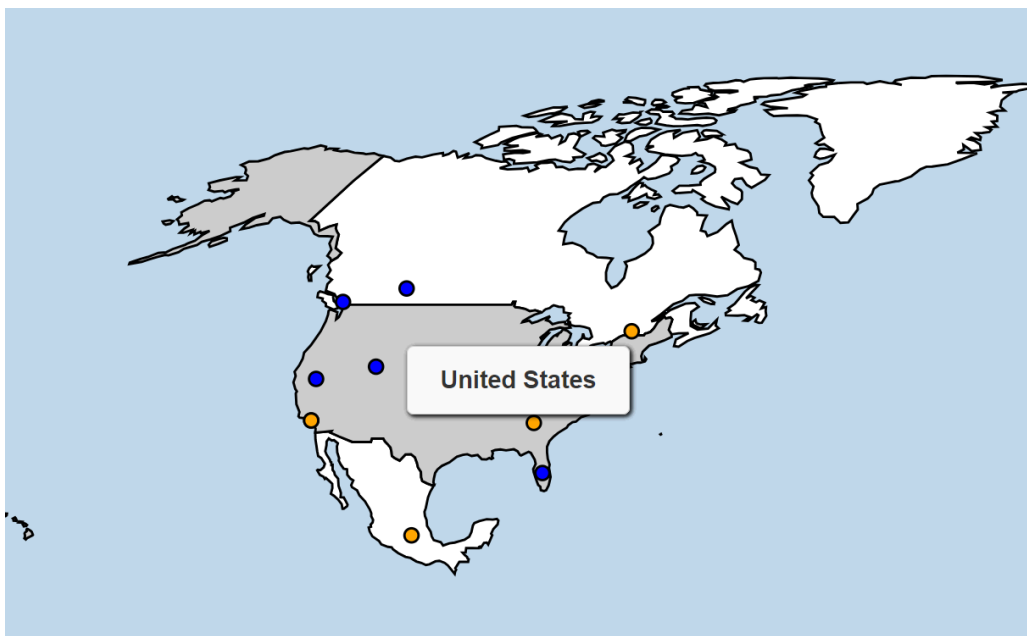


Figure 2: Map showing interaction with a country.

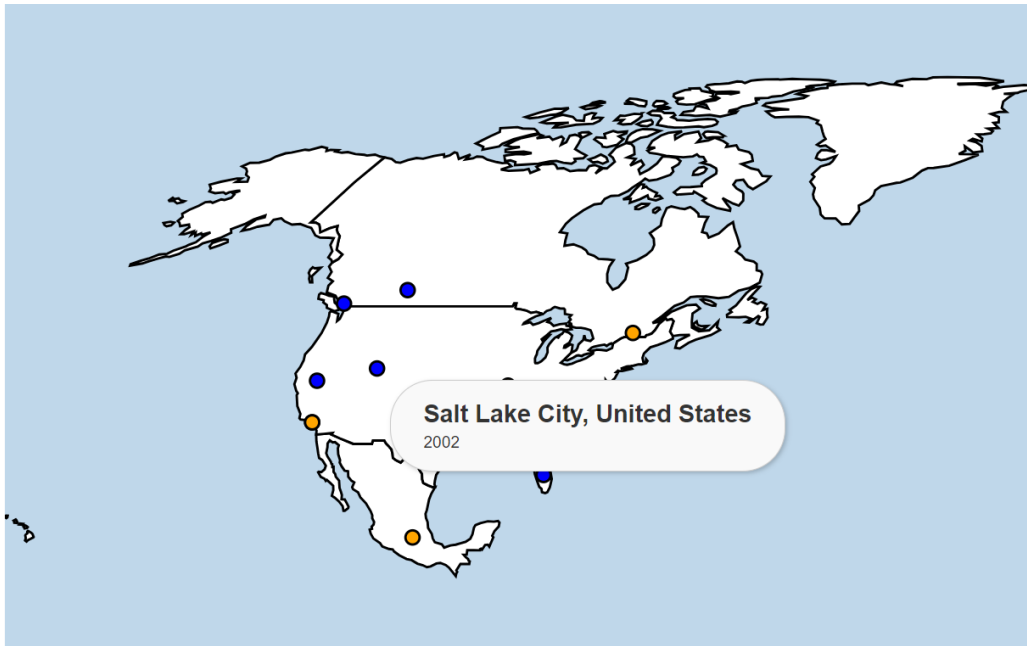


Figure 3: Detailed view of a host city.

## 7.2 Historical Timeline and Dynamic Table of the Olympics

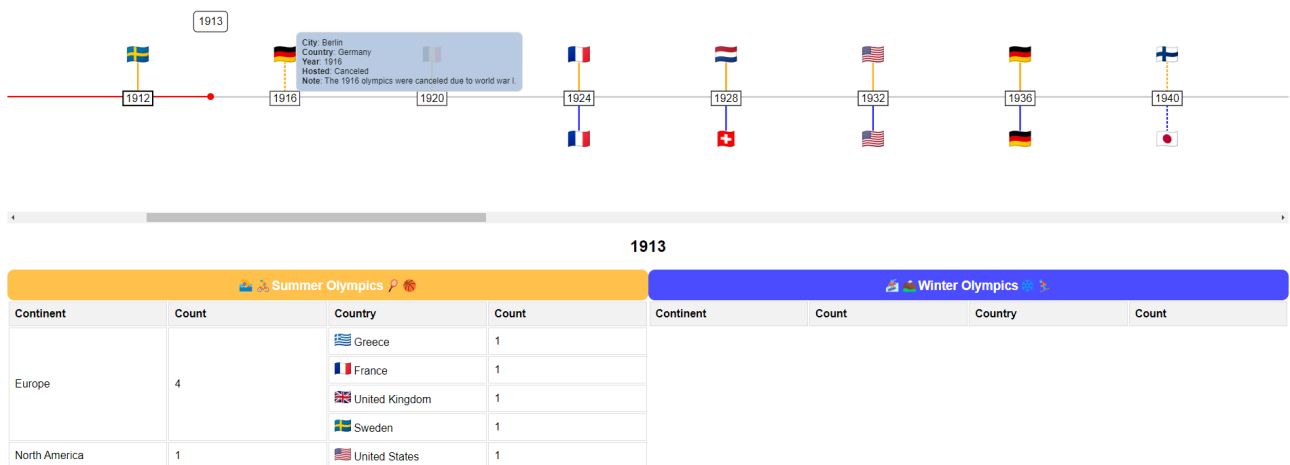


Figure 4: Historical timeline of the Olympic Games, showing host cities and dynamic table of counts by continent and country.

7.3 Gender Representation Over the Years

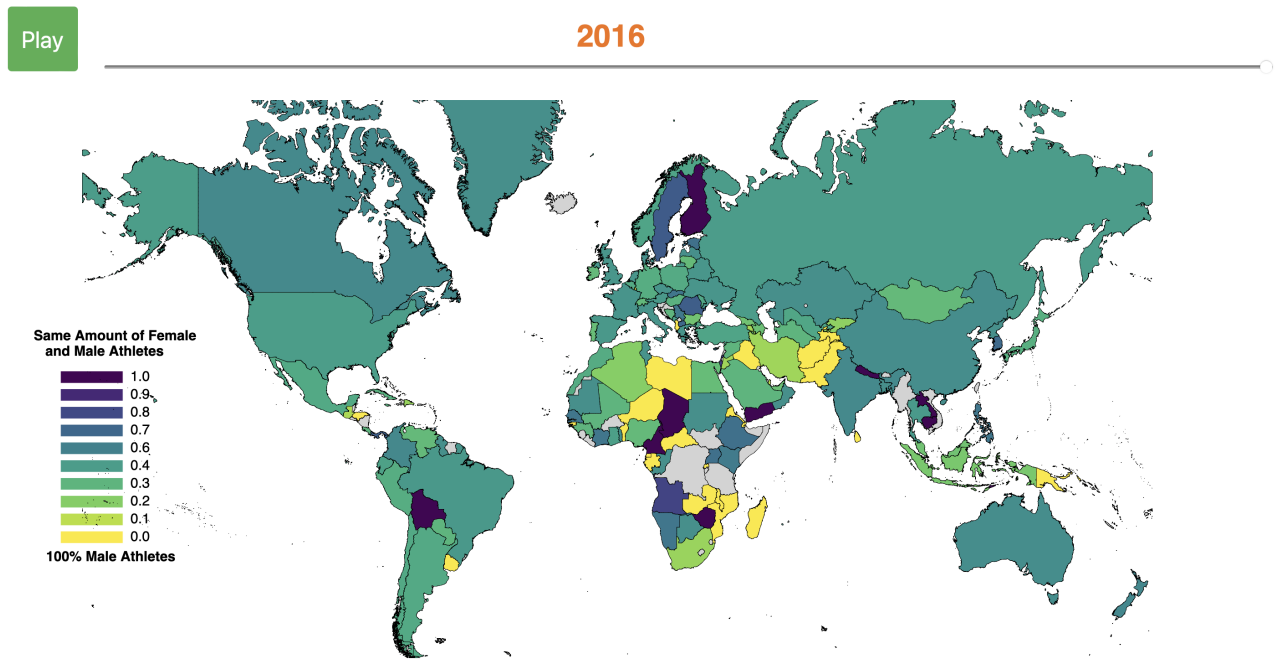


Figure 5: Gender representation map of the Olympic Games, illustrating the ratio of female to male athletes by country over the years.

7.4 Evolution of Disciplines Over the Years

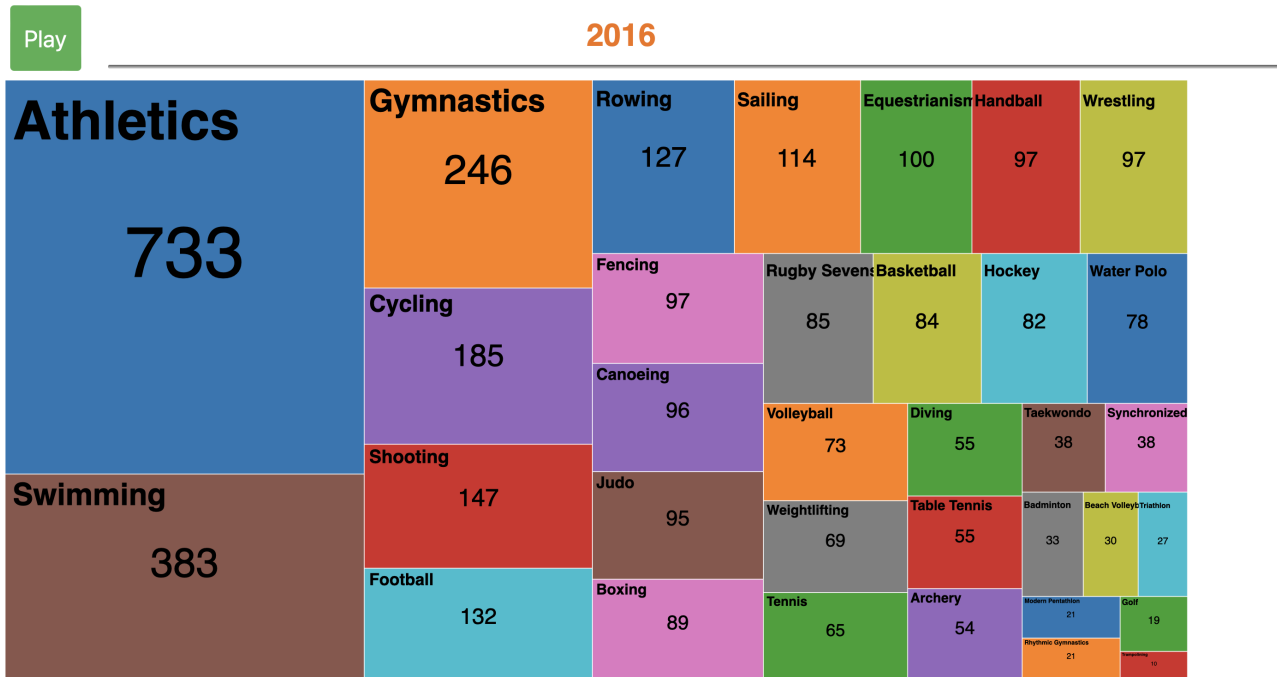


Figure 6: Tree map showing the evolution of disciplines in the Olympic Games, highlighting the number of athletes participating in each sport over the years.

## 7.5 Medals Distribution Over the Years

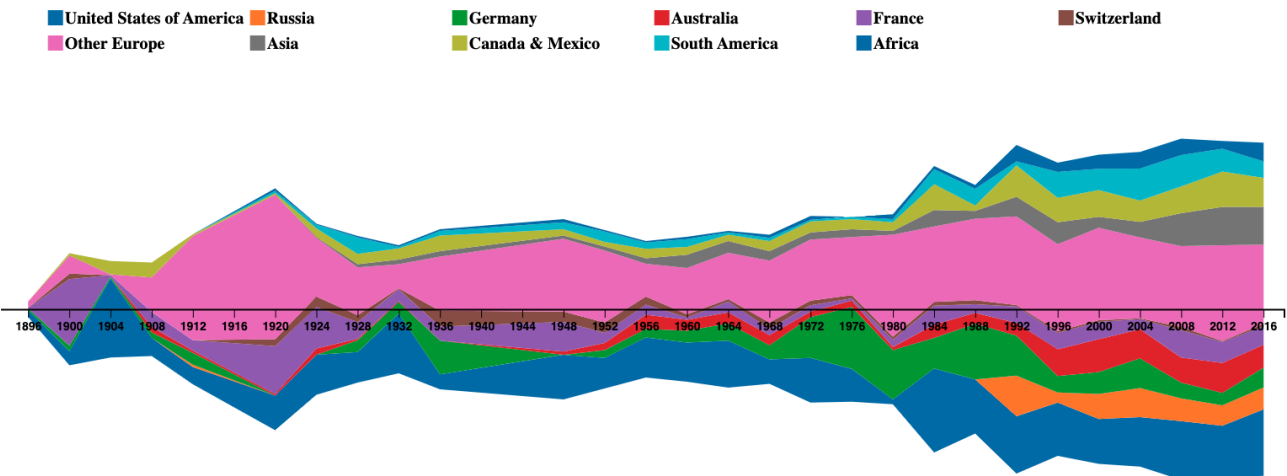


Figure 7: Area plot displaying the distribution of medals among countries in the Summer Olympics from 1896 to 2016.

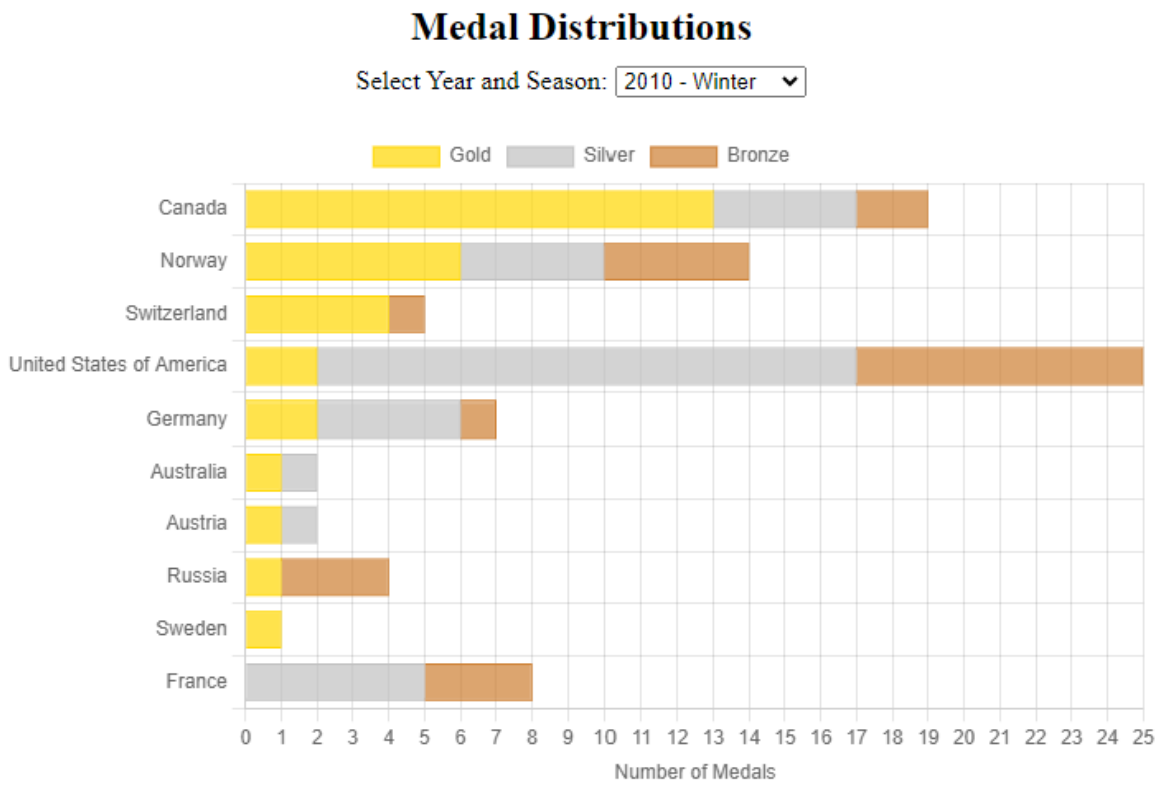


Figure 8: Country medal distribution by Olympic Game Edition

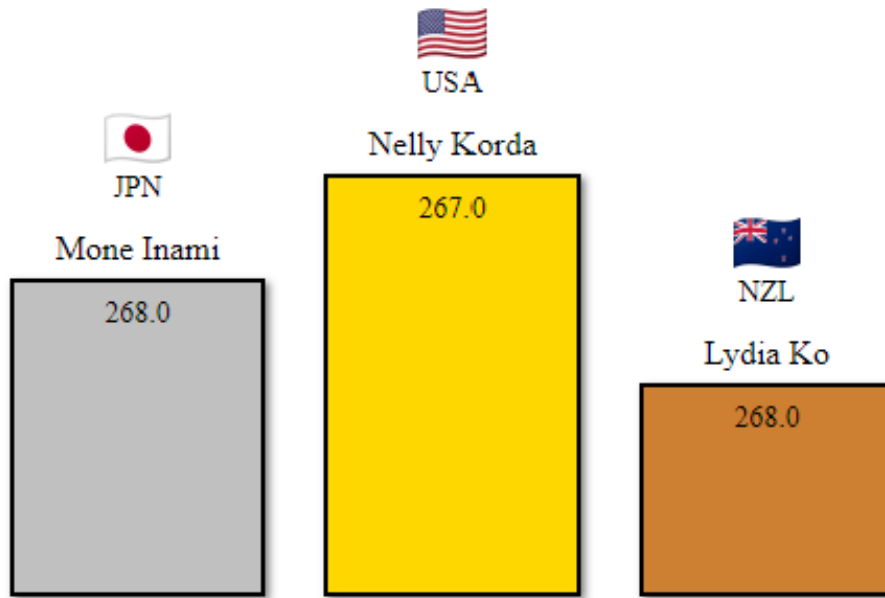
## Olympic Podium Visualization

2020 Summer Olympics ▼

Golf ▼

Individual, Women ▼

### 2020 Summer Olympics Golf - Individual, Women



[Source](#)

Figure 9: Event specific podium visualization

## Finals Swimming Events - Simulation

Choose an edition: 2000 Summer Olympics Choose an event: 100 metres Butterfly, Men Load Event Play Speed Up Speed Down Speed: x1.00

Anatoly Polyakov
Thomas Rupprath
Michael Klim
Ian Crocker
Takashi Yamamoto
Mike Mintenko
Lars Frölander
Geoff Huegill

[Source](#)

Previous Olympic Record: Denis Pankratov (1996 Summer Olympics) - 52.27s

Figure 10: Swimming event specific animation

## Finals Track Events - Simulation

Choose an edition: 2008 Summer Olympics Choose an event: 100 metres, Men Load Event Play Speed Up Speed Down Speed: x1.00

		Rank: 4	9.93s	Churandy Martina
		Rank: 7	10.01s	Marc Burns
New Record!		Rank: 1	9.69s	Usain Bolt
		Rank: 2	9.89s	Richard Thompson
		Rank: 6	9.97s	Michael Frater
		Rank: 8	10.03s	Darvis Patton
		Rank: 3	9.91s	Walter Dix
		Rank: 5	9.95s	Asafa Powell

[Source](#)

Previous Olympic Record: Donovan Bailey (1996 Summer Olympics) - 9.84s

Figure 11: Track event specific animation with Olympic Record



## References

- [1] Wikipedia, “Olympic games.” [https://en.wikipedia.org/wiki/Olympic\\_Games](https://en.wikipedia.org/wiki/Olympic_Games). [Online; accessed May 30, 2024].
- [2] Olympics.com, “The history of the olympic games.” <https://olympics.com/en/news/the-history-of-the-olympic-games>. [Online; accessed May 30, 2024].
- [3] Bhanu Pratap Biswas, “Olympic data.” <https://www.kaggle.com/datasets/bhanupratapbiswas/olympic-data>. [Online; accessed May 30, 2024].
- [4] Wikipédia, “Jeux olympiques.” [https://fr.wikipedia.org/wiki/Jeux\\_olympiques](https://fr.wikipedia.org/wiki/Jeux_olympiques). [En ligne ; consulté le 30 mai 2024].
- [5] Simplemaps, “Customizable world map.” <https://simplemaps.com/custom/world/qQCmlXaE>. [En ligne ; consulté le 30 mai 2024].
- [6] FlagCDN, “Flagcdn - country flags.” <https://flagcdn.com/>. [En ligne ; consulté le 30 mai 2024].