OLYMPIC MEDALS BY COUNTRY

FINAL PROJECT

Andy Tran
Christina Le
Tingting Zhou

Table Of Contents

Project Proposal

Process Book

Project Proposal

Background and Motivation

Our team, one Data Science major and two Computer Science majors is always looking for ways to stay up to date with the latest trends in our fields and the world. This summer, with so many ongoing global political contentions, the Paris 2024 Olympics felt like a refreshing opportunity to focus on unity and the athletic feats of our peers. The Olympics has always brought people from around the world together, which inspired us to choose it for our data visualization project.

What makes the Olympics special to us is the excitement of the big events and the chance to learn about smaller countries and their athletes—people who don't normally get the spotlight but are still part of this incredible global event. We want to highlight diversity and inclusiveness in our project by showing the stories behind the data.

From a technical standpoint, the Olympics provides a rich and dynamically updated dataset, which gives us a great opportunity to show off our skills in both data analysis and system design. By working with the latest data, we hope to create visualizations that are engaging, insightful, and accessible to everyone, whether they're casual sports fans or data enthusiasts.

Ultimately, this project is about more than just data—it's about using technology to tell the story of the Olympics and what it stands for: unity, diversity, and the spirit of competition. We hope that by showcasing the accomplished athletes from all over the world, our project will help others see the Games not just as a competition but as a celebration of people coming together through difficult times.

Related Work

Our Olympic medals data visualization is largely inspired by the way television broadcasts present live Olympic events, particularly the swimming competitions. The visualization techniques used during the races—such as real-time comparisons of swimmer positions and times—were influential in how we approached displaying dynamic competition data. We wanted to capture that same sense of progression and competition in our visualization of Olympic medal counts.

In addition to this, we conducted further research into effective visual storytelling for the Olympics and came across the blog post by Flourish on visualizing the Olympics (Flourish Blog). This resource offered valuable insights into different approaches to presenting sports data, which helped shape our thinking on how to present Olympic medal information in an engaging and informative way. These influences guided us to create a clear and interactive visualization that highlights key trends and allows users to explore the data in depth.

Initial Design Concepts

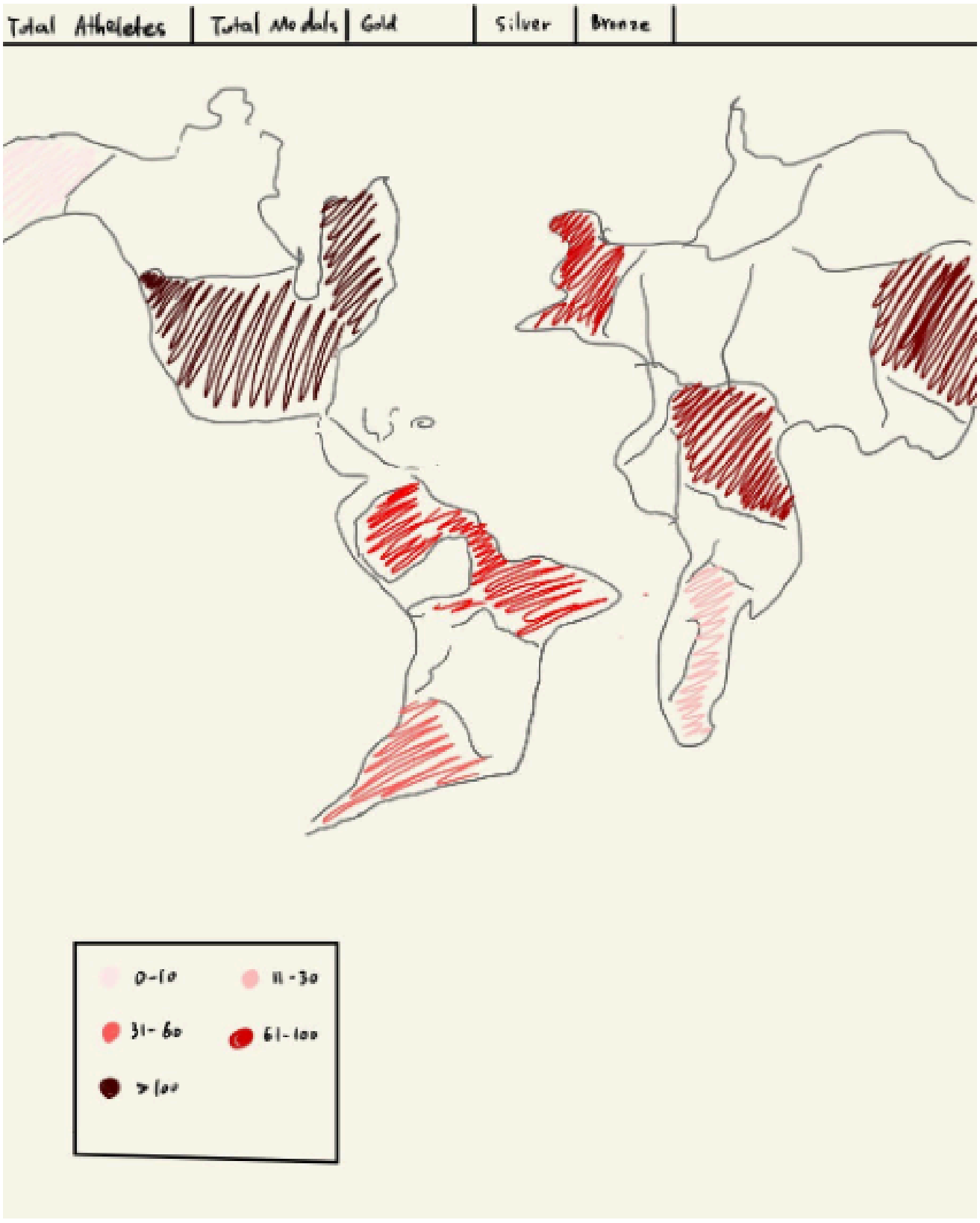
Geographical heat map
A world map where countries are color-coded based on the number of athletes or medals won. Hovering over each country provides additional statistics like athlete count, medal breakdown, and historical performance.

Justification: This design effectively highlights geographical distribution and lets users easily compare performance between countries. The visual encoding of color helps users quickly grasp the scale of participation and performance.

Timeline and Medal count
A timeline that displays medals won by different countries over time, with each country represented by a line graph or bar chart. This allows users to track the rise and fall of Olympic success for countries across multiple years.

Justification: Line graphs and bar charts are well-suited for showing trends over time, making it easy to spot performance changes. Time is visually encoded along the x-axis, and medal counts along the y-axis, allowing users to see patterns in performance.

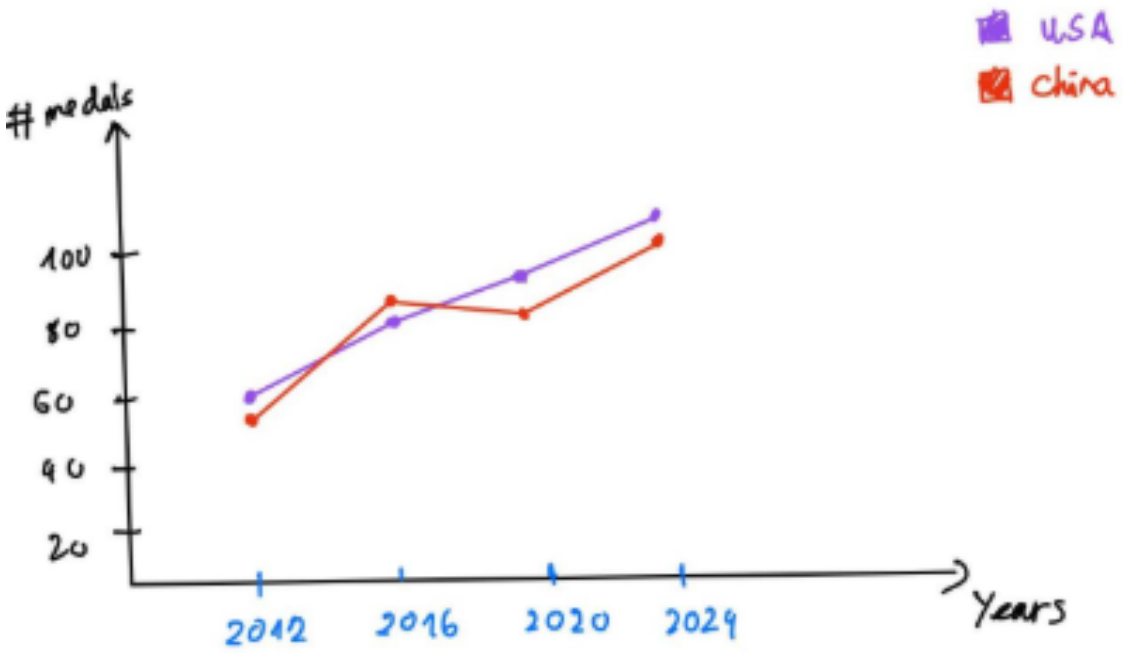
Heat Map Sketch



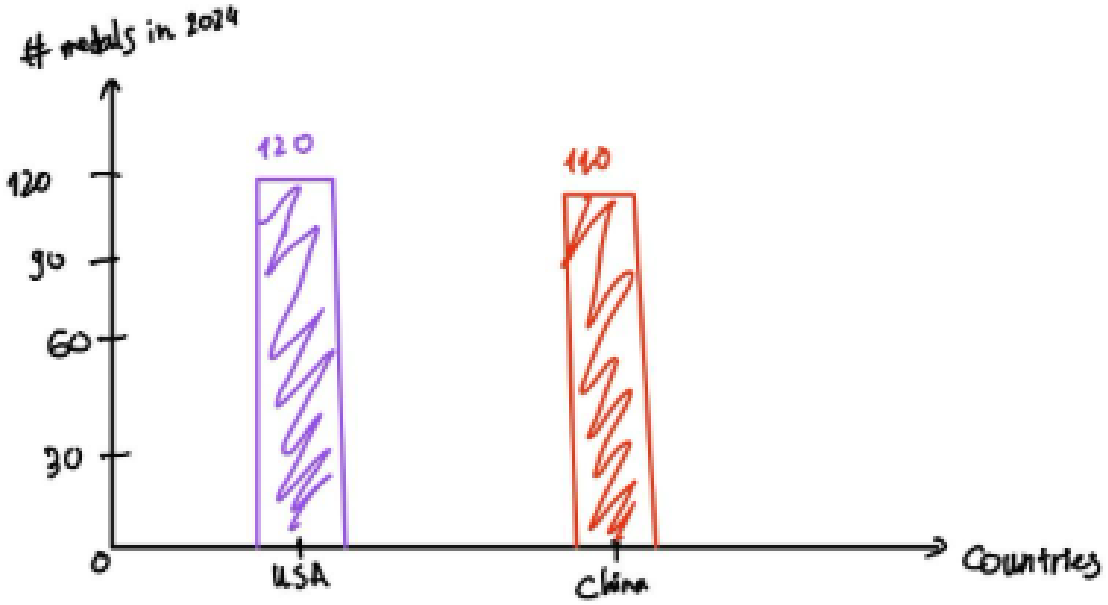
Hover Map Sketch

Project Proposal

Timeline Sketch



Medal Count Sketch



Initial Design Concepts

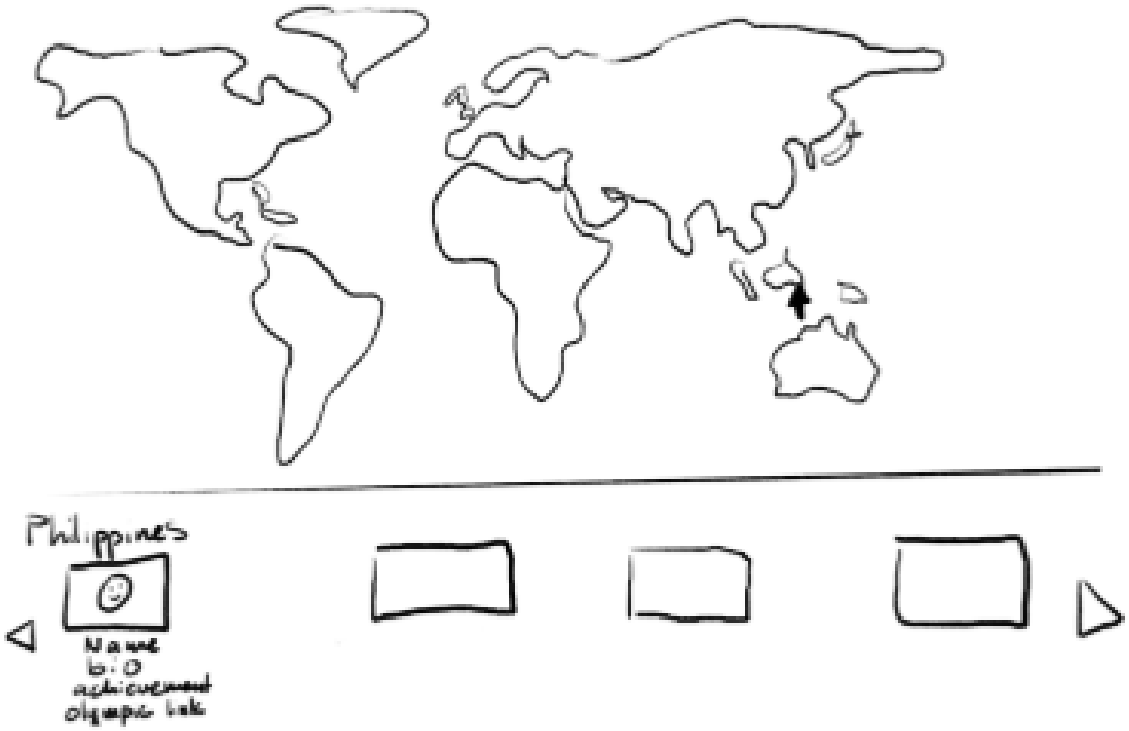
Athlete Profile Highlights
An interactive dashboard where users can select athletes from smaller or underrepresented countries and see individual profiles that include stats, bios, and achievements.

Justification: This focuses on the personal stories of athletes, providing an opportunity to highlight diversity and inclusivity. It gives depth to the data by allowing users to engage with the human aspect behind the numbers

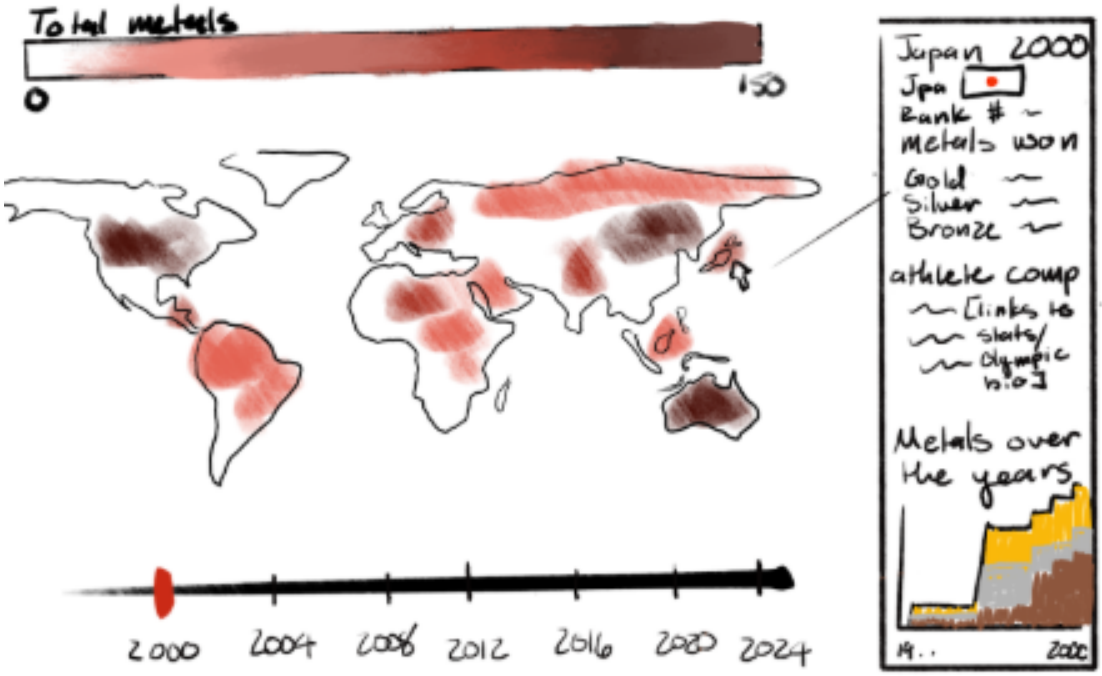
Final Design
The final design will include an interactive geographical heat map denoting medals won with a timeline slider that changes the heat map depending on the year of the slider. Upon hovering, countries will list the medal count/athletes participating in that given year. The interactions between these features provide a comprehensive narrative allowing users to explore general trends, but also a specific athlete's story which makes the data more engaging and insightful for both casual and data enthusiasts.

Project Proposal

Athlete Profile Sketch



Final Concept Sketch



Questions

Initial Aims

At the outset of our project, we aimed to answer several core questions related to Olympic medals data:

- Which countries have been the most successful in terms of total medal counts?
  - We wanted to provide a clear view of which nations have dominated the Olympics historically, across all sports, and over time.

- How have Olympic medal trends evolved?
  - We were interested in visualizing how different countries have risen or fallen in the medal rankings throughout Olympic history, highlighting any significant changes in dominance or performance over the years.

- What is the distribution of medals across different sports?
  - This question aimed to explore whether certain sports are consistently dominated by particular countries or regions.

Questions

Evolved Questions

As the project evolved, new questions emerged based on our exploration of the data:

- How do countries compare in terms of medals won relative to their population or GDP?
  - While initially focused purely on medal counts, we started considering how external factors like population size or economic strength might influence a country's Olympic success.

- Similarly, how does a country's assignment of athletes per sport affect medals won?
  - Some countries seem to dominate a sport and we wanted to see if the number of qualified athletes sent to an event influenced how many medals are won by a single country in a single event.

- Are there specific years or events where smaller or less dominant countries saw a sudden spike in medal wins?
  - This led us to look into historical or geopolitical factors that could have contributed to unusual shifts in the data.

Questions

Further Analysis

During our analysis, we also began to consider new questions, such as:

- How do host countries perform compared to their usual averages?
  - This question arose from observing that host nations often tend to perform better when the Games are held on home soil.

- What are the gender trends in medal distributions over time?
  - As gender equality in sports became more prominent, we explored how the ratio of male to female medalists has changed across various sports and Olympic games.

These questions not only guided our data visualization decisions but also helped us present a more comprehensive narrative around the Olympic Games and medal trends.

Data

Data source
Wikipedia of the Summer Olympics from 2008 to 2024. We didn't use any scraping or cleanup for the data acquisition and process.

We did try to find data for each athlete but it was hard to find data on athletes that were not well known, so we did not use any athlete data as it would defeat our initial intentions of including athletes in our visual.

E**xploratory Data Analysis**

- What visualizations did we use to initially look at we data? What insights did we gain? How did these insights inform our design?
  - To begin our analysis, we utilized several visualizations to explore the Olympic medals data and gain a deeper understanding of the underlying trends. Our primary focus was on displaying the total number of medals (gold, silver, bronze) won by all countries, and we used different approaches to visually represent this information.

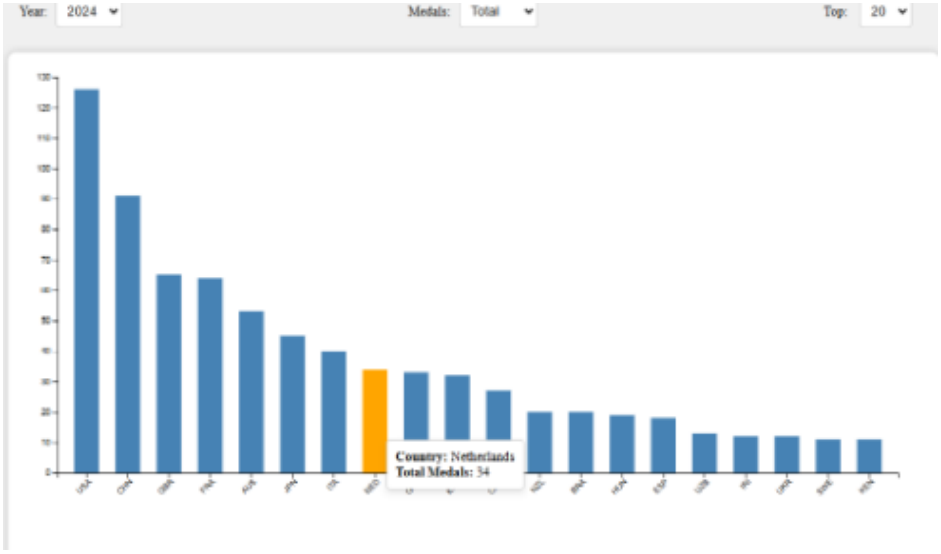
Implementation

Initial Visualizations

Bar Charts for Medal Distribution:

In our initial exploration, we utilized bar charts to compare total medal counts across countries, offering a clear view of the top-performing nations in Olympic history. By further breaking down the data into individual medal types—gold, silver, and bronze—we revealed nuanced insights into each country's strengths. This approach highlighted that some countries excel in winning certain types of medals, such as consistently earning more bronze while securing fewer golds, and uncovering intriguing patterns in medal distribution.

These bar charts proved particularly valuable for visualizing both overall and specific-year performances, enabling users to identify dominant nations in the Olympics. By adhering to basic design principles, such as proximity for grouping data and clear comparison, we ensured that relative performance differences among countries were immediately apparent.



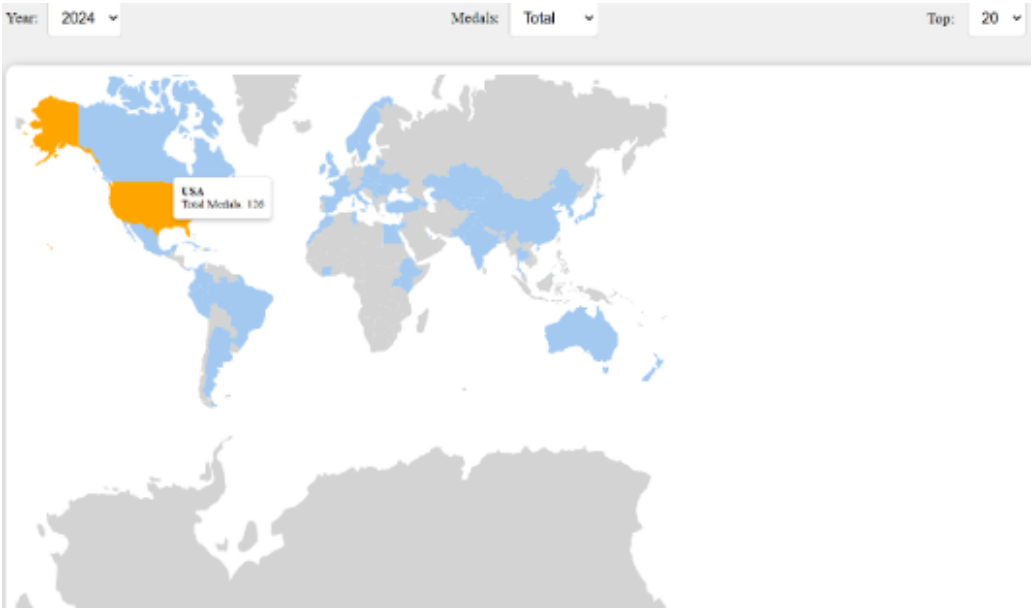
Implementation

Initial Visualizations

Heatmaps for Country Medals Over Time:

We also created heatmaps to visualize how countries' medal counts evolved over different Olympic Games. This time-series visualization revealed periods of dominance for specific nations, such as the United States and China, and how newer countries have emerged as strong competitors in more recent years. The heatmap made it easier to identify patterns in medal wins over time, particularly during certain eras.

To represent the geographical distribution of Olympic medals, we used geo maps that allow users to visualize which regions of the world excelled in each Olympic year. The countries are shaded based on their total medal counts, giving users a clear view of regional dominance. The map adheres to the principle of spatial organization, allowing users to connect the data with real-world geography intuitively.

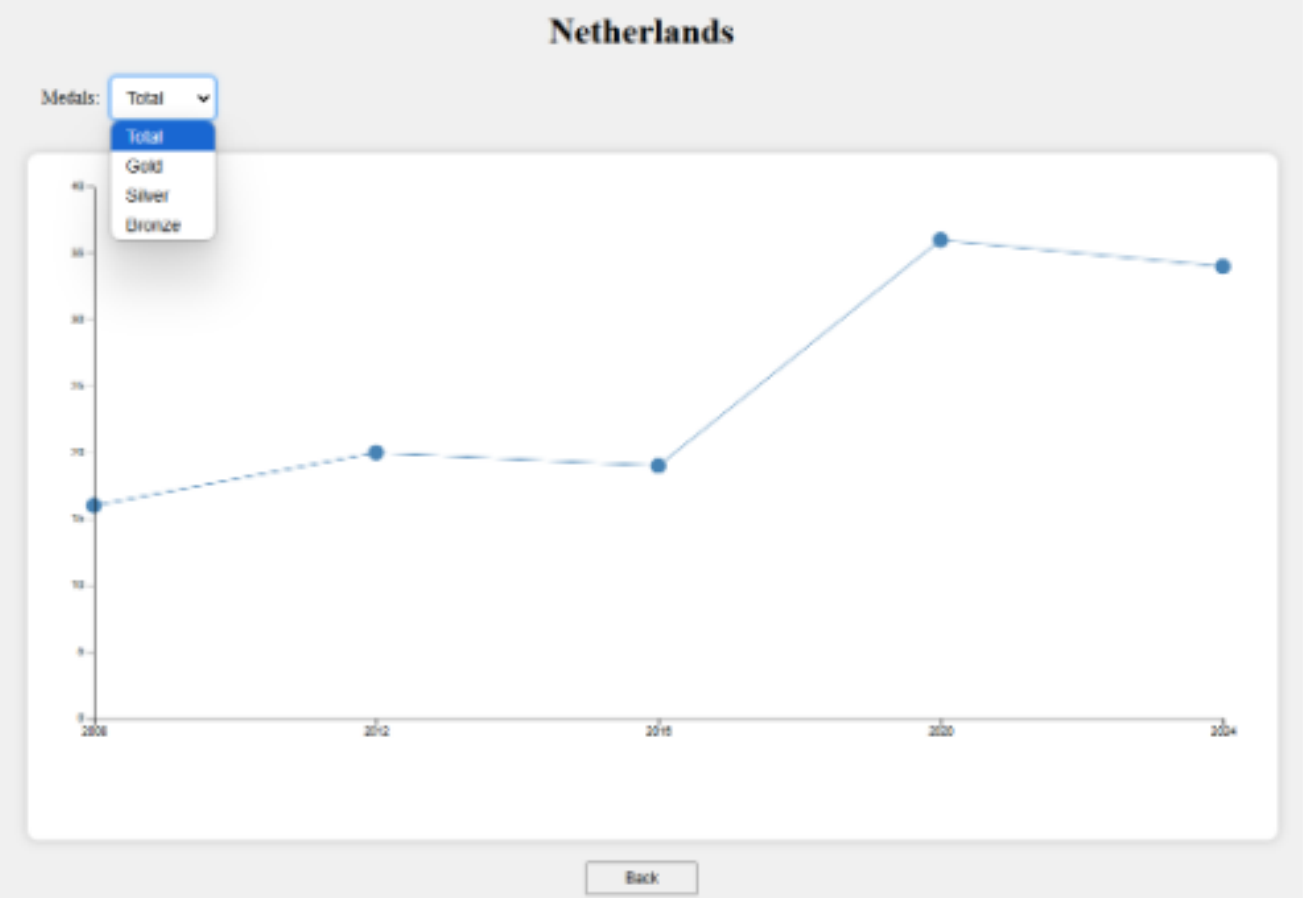


Implementation

Initial Visualizations

Line Charts for Medal Types:

We have a responsive line chart that appears when a user clicks on a country in the bar chart. This line chart shows the amount of the selected medals (Gold, Silver, Bronze, and Total) a country has won each year. This visualization helps users compare medal counts throughout the year. The chart is to be updated to include comparisons between two or more countries.



Implementation

Initial Visualizations

Insights Gained:

From these initial visualizations, we gained several important insights:

- Country Dominance: The bar charts and heat maps highlighted the dominance of countries like the United States, Russia, and China in terms of total medal counts, while also showing smaller countries' sporadic spikes in medal wins.

- Consistency vs. Specialization: The breakdown of medals into gold, silver, and bronze revealed that while some countries consistently rank at the top in all categories, others specialize more in certain medal types. For instance, some nations frequently win more bronze medals, suggesting a higher frequency of reaching finals but not securing the top spot.

Implementation

Design Evolution

Insight informed design:

These insights played a critical role in shaping our overall design decisions for the final visualization:

- Emphasis on Medal Breakdown: Based on our discovery of how different countries vary in their success across gold, silver, and bronze medals, we decided to incorporate a clear breakdown of medals by type in our final visualizations, making it easy for users to explore these differences.

- Interactive Elements for Time-Based Trends: The heatmap's ability to showcase changes over time inspired us to include interactive time-based visualizations that allow users to explore how countries have performed over different Olympic years. This dynamic aspect adds depth to the analysis and encourages further exploration.

- Comparative Visualization: We also recognized the value of side-by-side comparisons between countries in terms of their medal distributions, leading us to design visualizations that facilitate direct comparisons of overall performance and medal type breakdowns between nations.

Implementation

Design Evolution

Current Design:

At this point, the bar chart, line chart, and geo map visualizations are managed in three separate scripts. Both visualizations allow users to filter by year and medal type. For example:

- The bar chart script updates medal counts based on the selected year or medal type, helping users focus on specific periods or medal categories.

- The line chart script updates medal counts based on the selected country and medal type, helping users see trends between countries throughout the years 2008-2024.

- The geo map script reflects geographical medal distributions based on these same filters, offering insights into which countries excel in different Olympic years or for specific medals.

As we integrate the bar charts, geo maps, and line chart into a single interface, we will ensure that users can filter the data by year, medal type, and country across all charts simultaneously. This will streamline the user experience and improve interactivity, following the principle of consistency across all visualizations.

It is also important to note that ChatGPT was used during the debugging process and to simplify our code for all three scripts.

Implementation

Design Evolution

In the development of our Olympic medals data visualization, we explored different approaches to effectively represent the data and allow for interactive exploration. Initially, we focused on two main visualizations: bar charts, line charts, and geo maps, each of which enables users to filter the data by year and medal type (gold, silver, and bronze). These visualizations provided a foundation for understanding both the overall medal counts and the geographical distribution of Olympic success. We wanted to keep these simple visuals so that we could use color, shape, and size as our focus. We changed the color of our map to one that stood out more so that we could depict a strong range of colors to show the distribution of medals in our heat map. Shape was used in our line charts as the spikes showed increases/decreases in a country's winnings. Size is used in our bar chart to give a direct comparison between the two countries. We chose to keep the bar chart despite it being similar to the lines charts because we wanted the line charts to show trends between countries over time, and the bar chart looks only at a specific year between two countries.

Implementation

Deviations from the Proposal

Hover Map: We found it a bit redundant to have a pie chart of a county's medal-type distribution since we already had the heat map that can be updated depending on the user's choice. Additionally, if a user were to choose a medal type that wasn't the total the pie chart would only show the medal that was selected which would be useless.

Athlete Profile Highlights: When trying to find data on athletes we realized that a lot of athletes had missing data. So, we couldn't represent all athletes in a profile highlight. Only athletes that got medals or were popular had viable data, and so it felt like it defeated our initial purpose of highlighting not well-known athletes for their performances,

Timeline Slider: We opted for a drop-down menu instead since the line chart will always look at all years, and having a timeline slider felt cluttering next to our drop-downs.

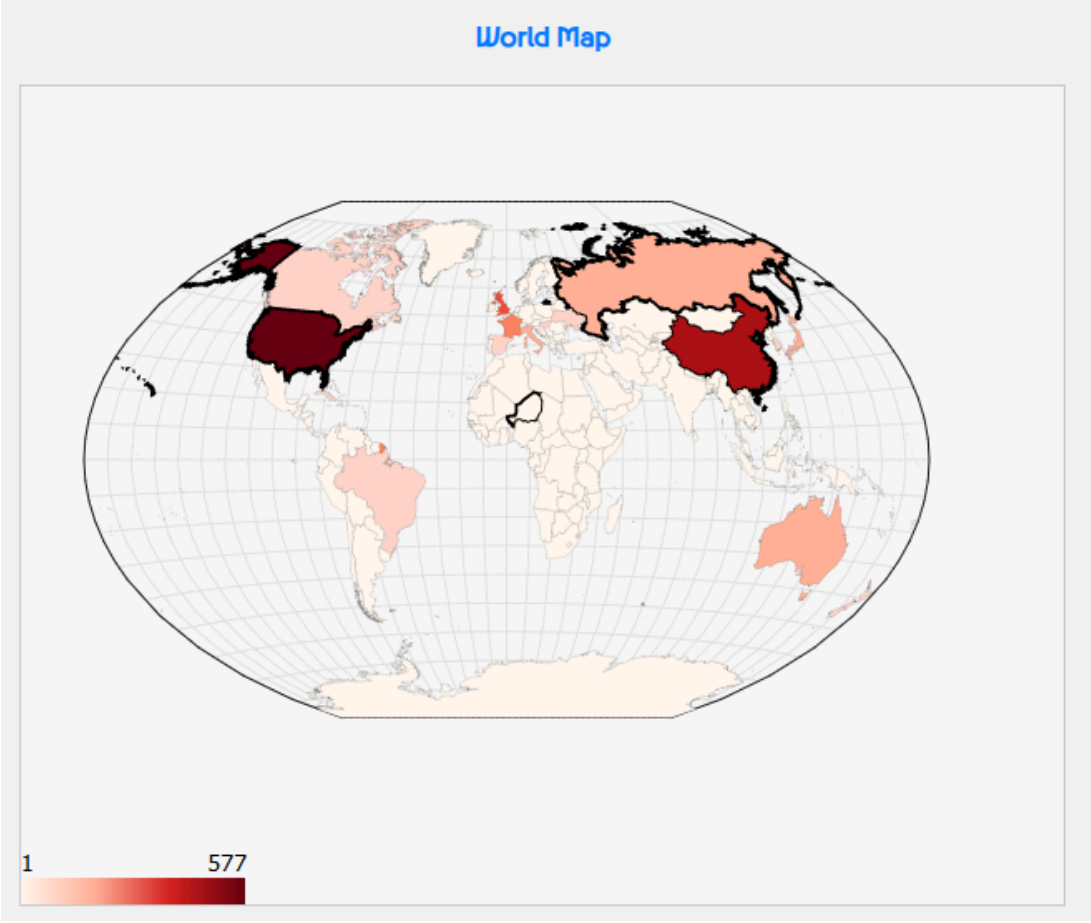
Final Concept Hover Highlights: We could not figure out a way to properly implement this, and, again, this felt redundant since we wanted to use our implemented line and bar charts.

Final Visualization

Map

The final design for our map remained similar to the initial implementation, but we changed the design of the map so that we could include a heat map. Before our map was only a single shade of blue indicating that there were medals but it was not clear the distribution between countries.

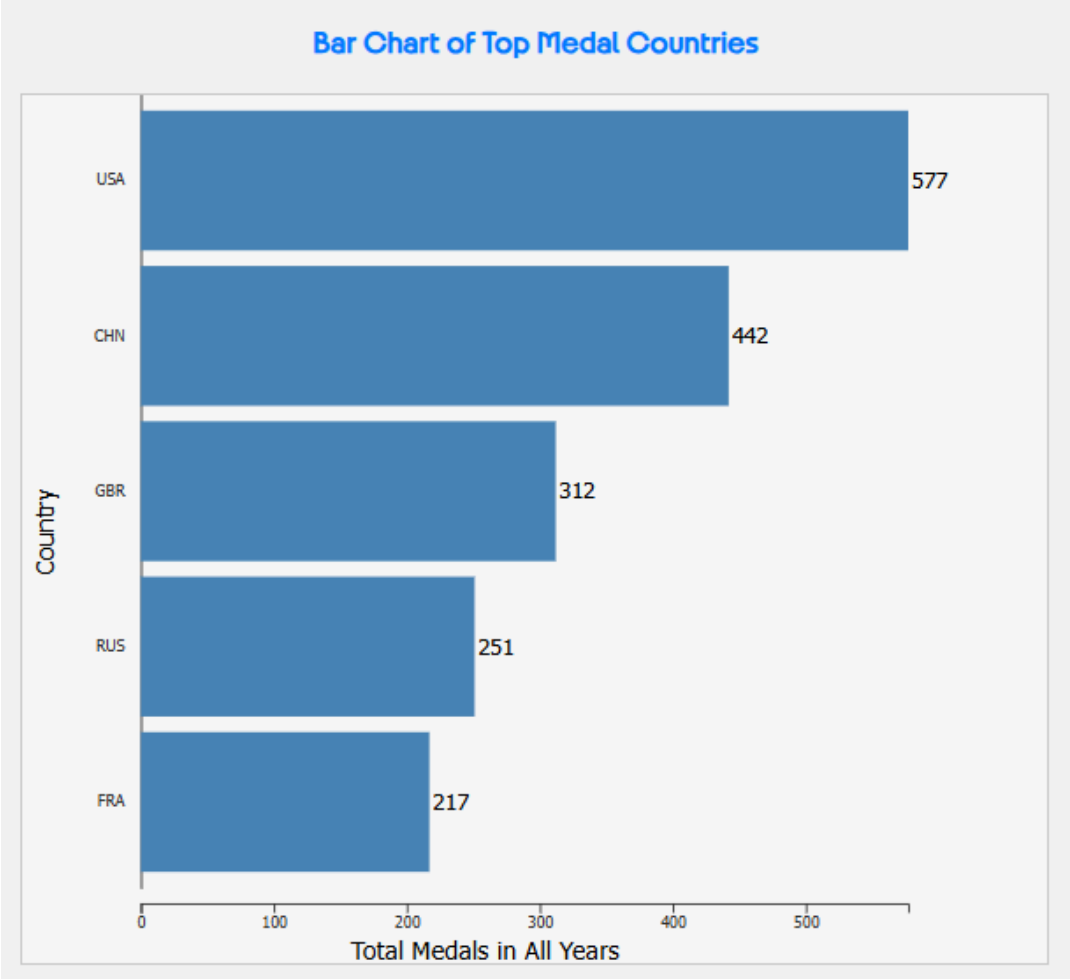
The picture below shows a user's click interaction with countries with a bold outline being the ones chosen. This interaction also updates the respective bar and line charts as well as highlights the country's name in our side list.



Final Visualization

Bar Chart

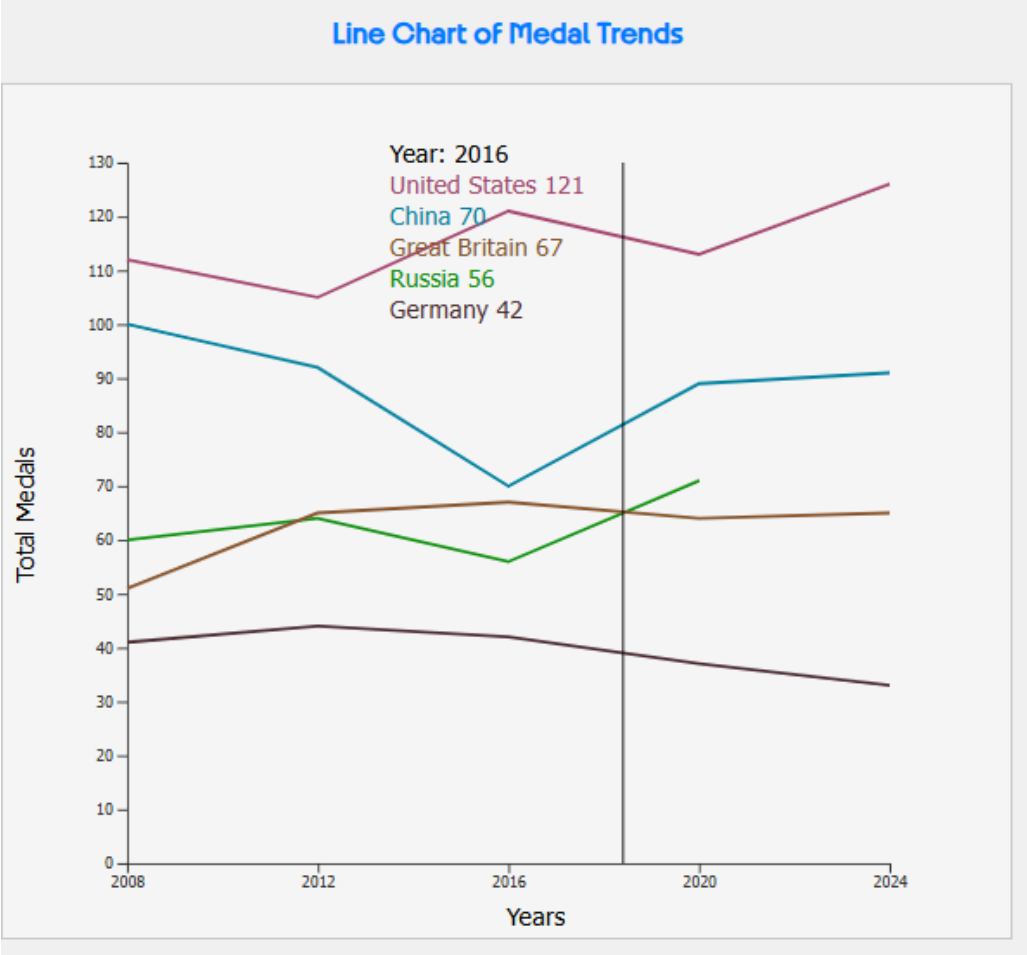
Our bar chart remained similar to our initial implementations as well. However, we made the bar chart more dynamic with animations and label changes as users choose different countries/medal types/years. On the initial launch, we changed the bar chart so that it only focused on the top five countries of all time so that our page didn't look too cluttered like it did in our initial implementation of the bar chart. The number at the end of the charts is the count for that country so that users don't have to guess based on the axis.



Final Visualization

Line Chart

Our final line chart can draw multiple countries. We implemented a mouse move element so that users can see the medal counts of the countries during a specific year. The line chart implementation is similar to the bar charts with it initially comparing the top five countries of all time, and the line chart updates based on the user's country and medal type selections. The time-series data adds a new dimension to our analysis, following the principle of continuity, as users can trace performance trajectories across years.



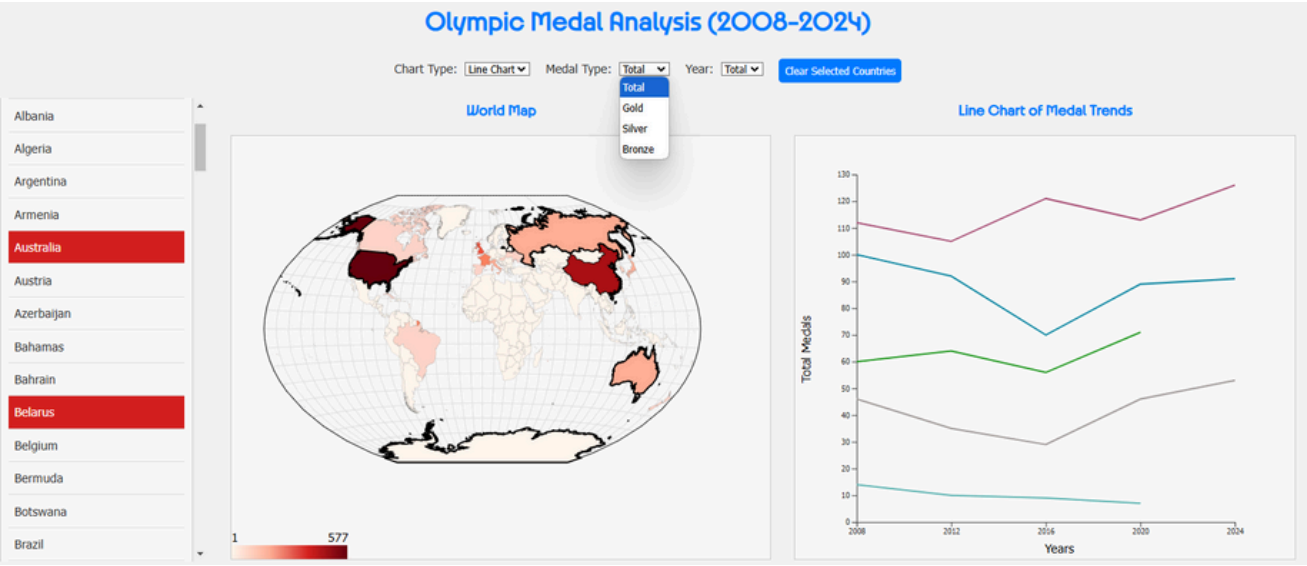
Final Visualization

User Interface

Drop Down: Users can choose what chart they want (line or bar) as well as the type of medal (Gold, Silver, Bronze, and Total) and year (2008-2024). This was the simplest and most user-friendly way to implement these selections.

Country Selection: Users can select a country by clicking on it on the map or the participants list to the left. We realized that selecting the country on the map may be hard for some users since a small country would require precise movements. However the participants list only includes countries that do have data because the line and bar chart only draw a country if it has data, so we found it meaningless to include it in our list. If we had included all existing countries it may clutter the list as well.

Clear Selection: This resets the visuals back to initial launch status.



Evaluation

We learned that most of the economic powerhouses/ largely well-known countries are the ones holding most of the medals. We were able to answer most of our questions through our map and line charts since those provided a visual distribution between countries and trends throughout the years. The questions we could not answer revolved around athletes since we were unable to find representative data.

Our visualizations work well for their intended purposes. Further improvements would be to have more labels on the map.

We would like to also try and find more representative data in the future so that we can integrate these charts:

- Scatter Plots for Medal Count vs. GDP or Population
  - To bring a deeper analytical perspective, this visualizes the relationship between a country's medal count and socio-economic factors such as GDP or population size. By using scatter plots, users can explore whether wealthier or more populous countries tend to win more medals, providing insights into the potential correlation between a nation's resources and its Olympic success.
- Stacked Area Charts for Medal Composition
  - A stacked area chart would show the composition of gold, silver, and bronze medals over time for each country. This will give users a more nuanced understanding of a country's medal makeup over different Olympic years, as well as its overall success trajectory.