Project Proposal

Basic Info

- Project title: Paris 2024 Olympic Summer Games
- Team members:
    - TingtingZhou(u1432080, u1432080@umail.edu)
    - Christina Le (u1408077, u1408087@utah.edu)
    - Andy Tran (u1399300, u1399300@utah.edu)
- Project repo: https://github.com/paupaudragon/dataVisualization

Background and Motivation

Our team, made up of one Data Science major and two Computer Science majors, is always looking for ways to stay up to date with the latest trends in both our fields and the world. This summer, with so much 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 been about bringing people together from around the world, and that's what inspired us to choose it for our data visualization project.

What makes the Olympics special to us is not just the excitement of the big events but 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 that 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.

Project Objectives

Our primary objective is to explore how athletes from smaller or less represented countries perform and participate in the Olympic Games. Specifically, we aim to answer the following questions:

- How does the representation of athletes from smaller countries compare to larger countries in terms of participation and medal counts?
- What trends can we observe in the geographical distribution of Olympic medals over time?
- Are there any underdog athletes or countries that have performed surprisingly well compared to historical performance or expectations?

We also want to learn how to effectively visualize large datasets in a way that tells a story, and highlights diversity and inclusiveness. By accomplishing this, our project will:

- Promote awareness of the global nature of the Olympics.
- Provide casual sports fans and data enthusiasts with insights into trends and stories they might not have noticed.
- Showcase our ability to work with real-time data and present it in visually appealing, user-friendly ways.

Ultimately, our visualization will emphasize the unity and competitive spirit that define the Olympic Games, and encourage viewers to appreciate the broader narratives behind the competition.

Data:

https://www.kaggle.com/datasets/piterfm/olympic-games-medals-19862018

https://www.kaggle.com/datasets/piterfm/paris-2024-olympic-summer-games/data

Data Processing

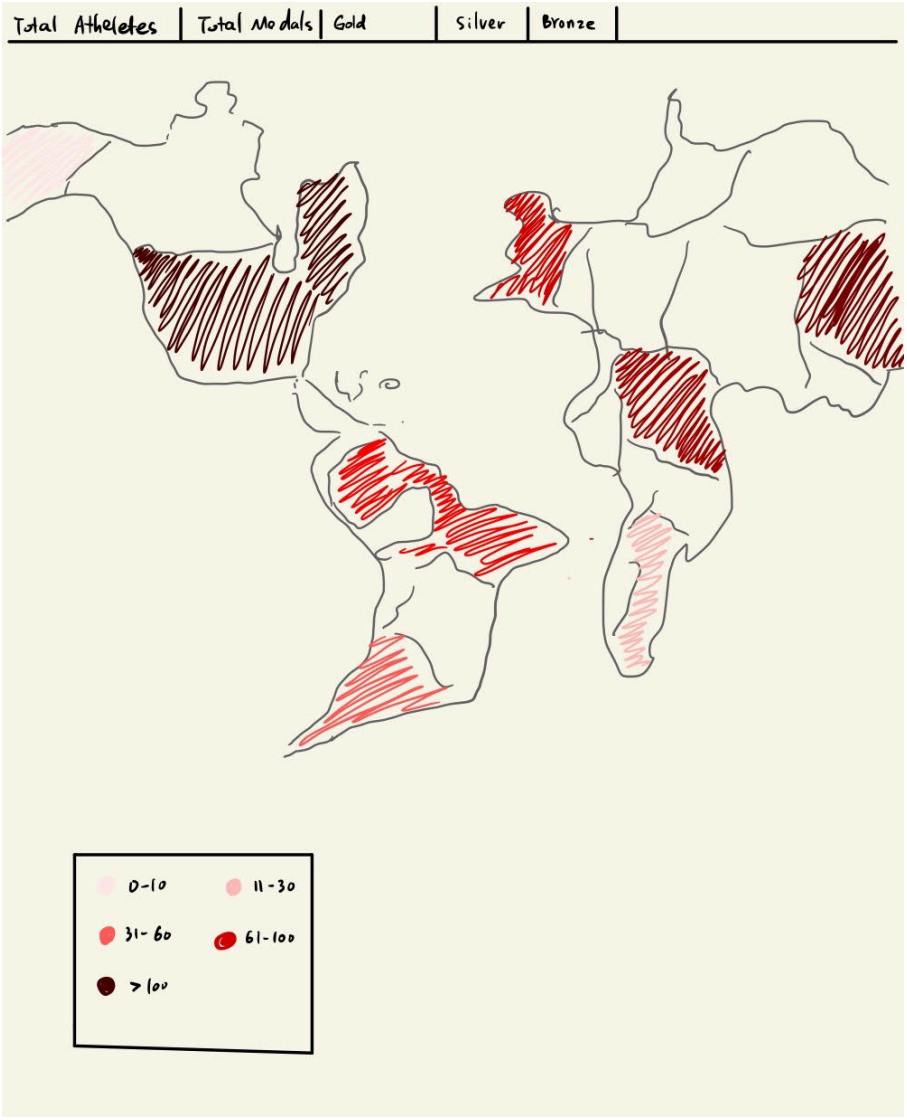
The data cleaning process will be light as we would just have to remove data categories and data that do not pertain to our questions/scope, like the winter olympics since we are solely focusing on summer olympics. We plan to derive any data about an athlete, where they come from, and how many metals they received. It is likely that we may deal with missing data depending on whether a country has participated prior to the 2024 olympics. There may also be some missing data between sets that would require cleaning dependent on what is missing/differences in naming.

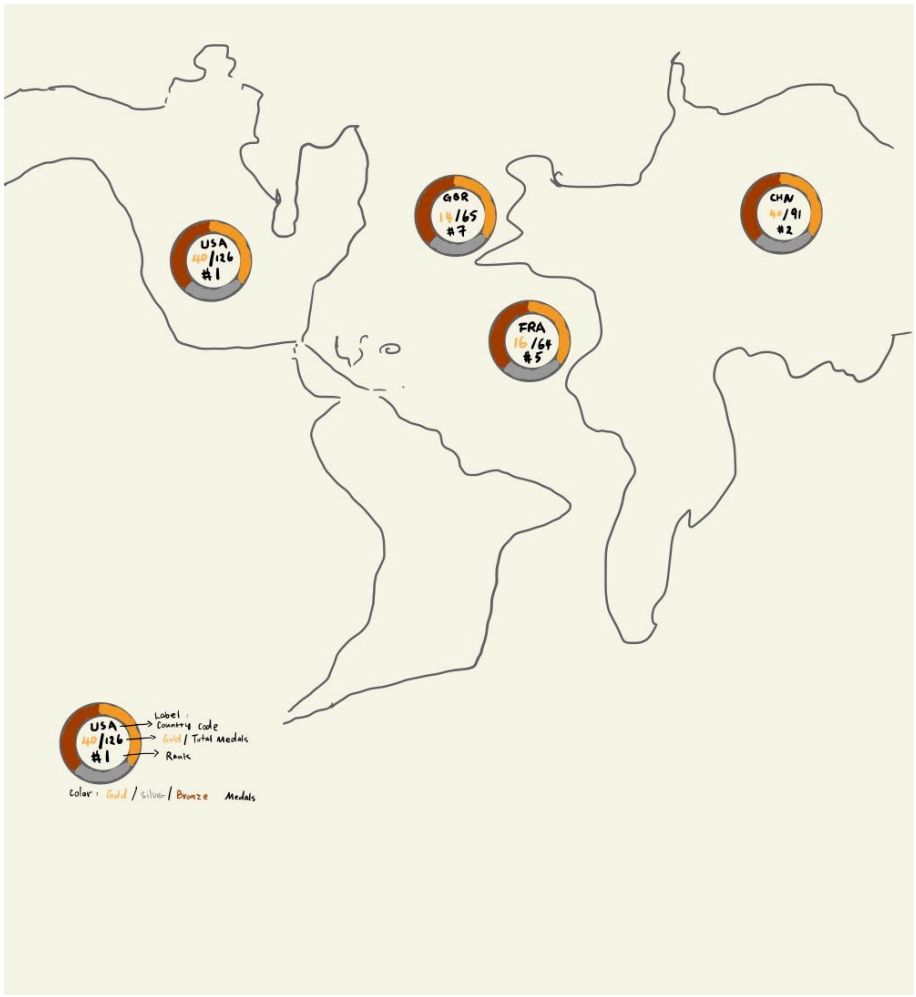
Visualization Design

- 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.

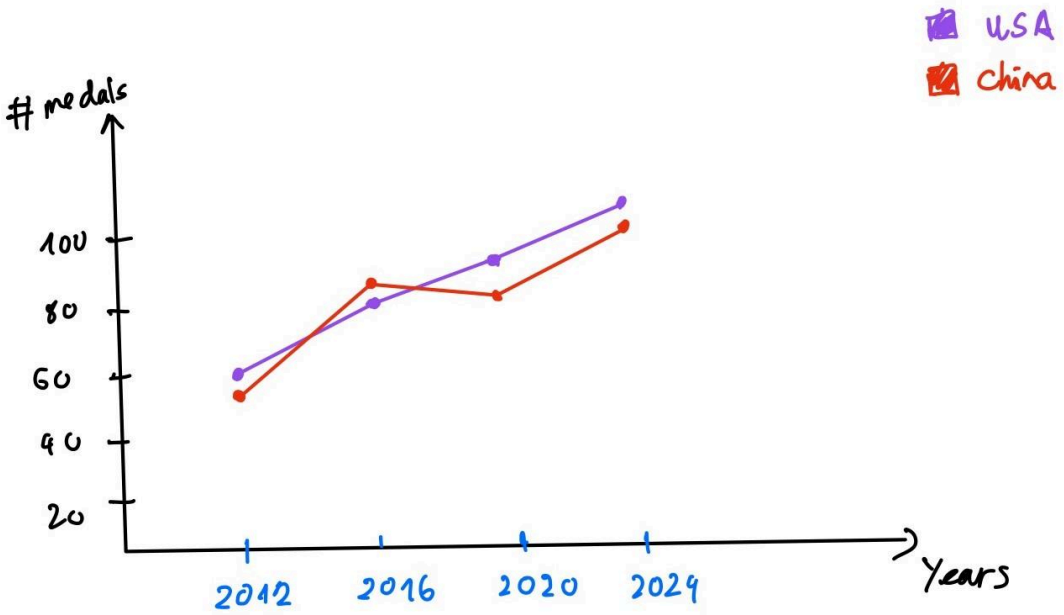


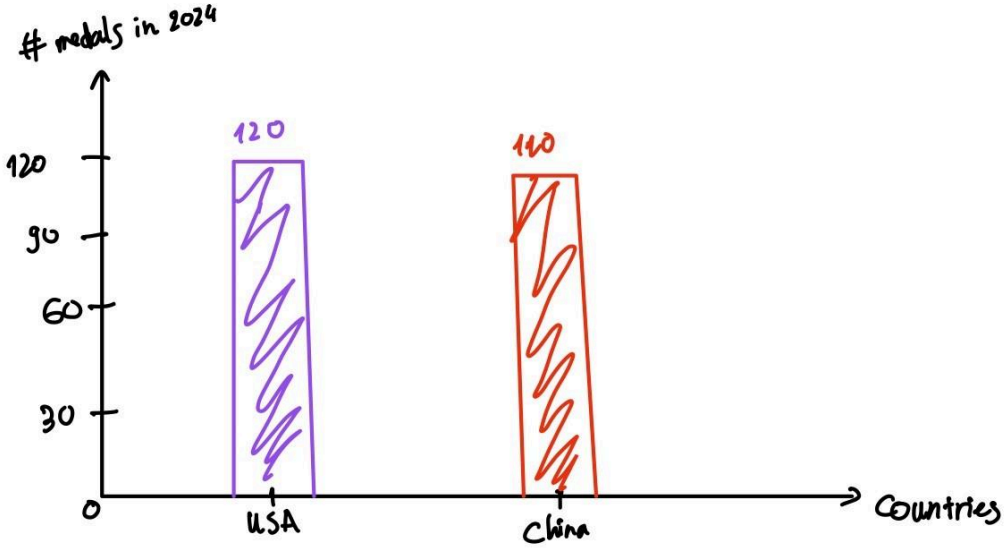


- Timeline and Medal Count Visualization

  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.

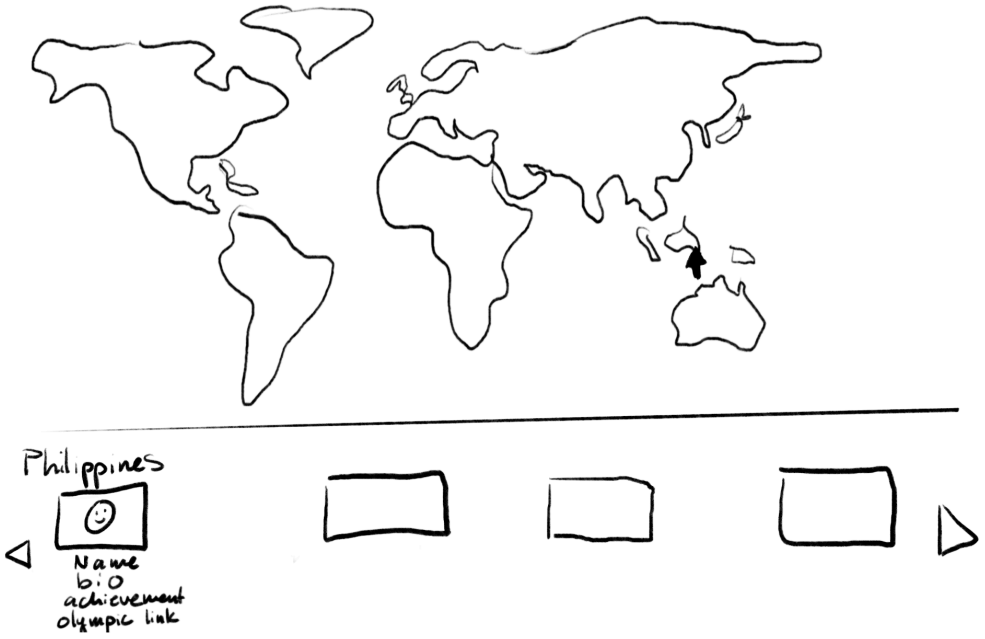

USA
China
medals
100
80
60
40
20
2012
2016
2020
2024
Years


medals in 2024
120
90
60
30
0
120
110
USA
China
Countries

- Athlete Profile Highlights

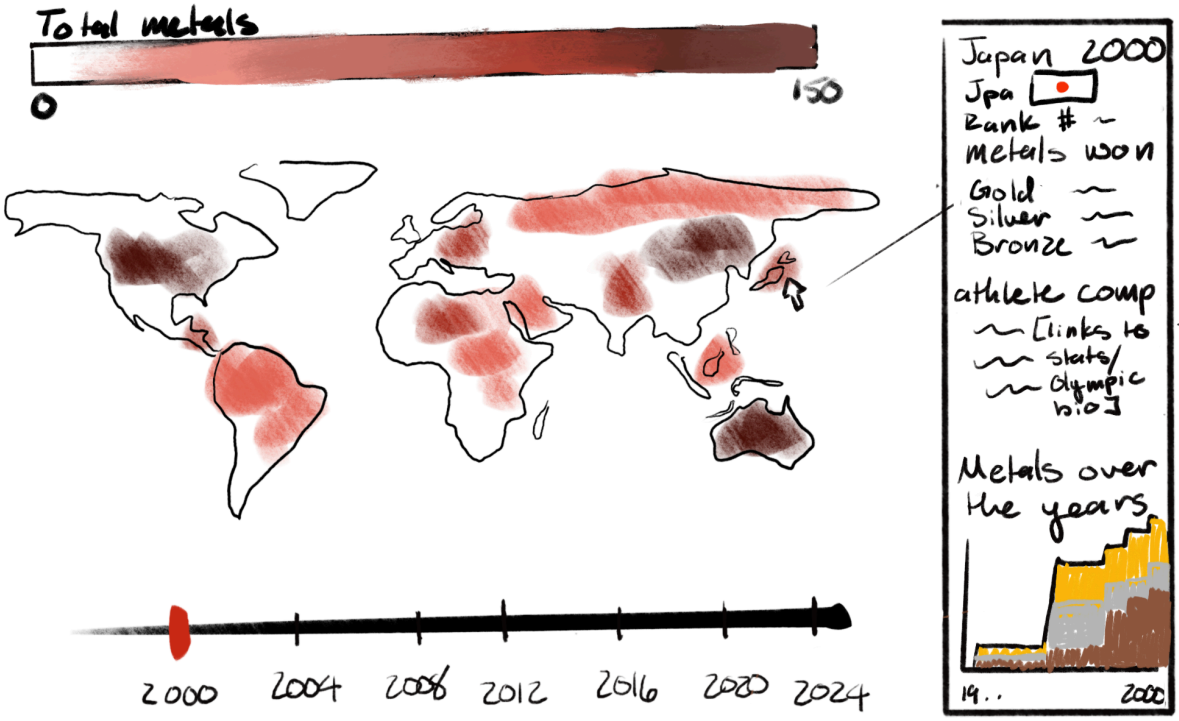
  An interactive dashboard where users can select athletes from smaller or underrepresented countries and see individual profiles that include stats, bios, and achievements.

  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 metals 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 that given year. The interactions between these features provides 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 users and data enthusiasts.



Must-Have Features

- Interactive geographical map showing country participation and medal counts.
- Timeline visualizing the historical performance of countries in terms of medals.
- Athlete profile view for detailed statistics on athletes from smaller countries.
- Responsive design for optimal viewing on different devices.

Optional Features

Integration of real-time data updates from the ongoing Olympics.

- Advanced filtering options for users to focus on specific sports, regions, or demographics.
- Social sharing feature allowing users to share visualizations on social media.
- Customizable visual themes (e.g., dark mode).
- Wider olympic timeline (beyond 16 years)
- Inclusion of winter and paralympics

Project Schedule

Week 5: Learn necessary technologies for the project

Week 6: Finish The Line/Pie/Chart simple visualizations

Week 7: Prepare for Midterm

Week 8: Fall Break

Week 9: Implement multiple views

Week 10: Finish Geo Graph

Week 11: Filtering and aggregation

Week 12: Extra features

Note: this schedule is subjected to changes. It is aggressive now because we aim to start early so we can have time to improve and prepare for final exams.