

Interactive Visualization of Olympic games

Process Book

Kiranmayee.D

Sai Vamshi.D

Table of Contents:

[Background and Motivation](#)

[Questions:](#)

[Data](#)

[Data Preprocessing:](#)

[Design Evolution:](#)

[Design 1:](#)

[Year slider:](#)

[Scatterplot](#)

[Marks and Channels](#)

[Table Chart:](#)

[Marks and channels:](#)

[Gender Inequality](#)

[Categorical Data:](#)

[Aggregated Data](#)

[Bar Graph](#)

[Design 2:](#)

[Tree map](#)

[Line chart and Box plot](#)

[Design 3:](#)

[Histogram](#)

[Word Map:](#)

[Final Design Implementation:](#)

[Scatter plot and Table chart](#)

Team Information

Sai Vamshi D	U1266122
Kiranmayee D	U1265299

Github repo:

<https://github.com/saivamshidobbali/120YearsOfOlympics>

Background and Motivation

The Olympic games are the world's foremost sports competition that first took place in 1896 in Athens, Greece. They are held every four years, with Summer and Winter Games alternating by occurring every 2 years apart. Thousands of athletes from all over the world participate in the events. It is one of the greatest events where the whole world come together and cheer for their idols. The games are considered not just between the athletes, but between the nations. Athletes are representatives of their nation, showing the world that their country can compete in the world.

It is important for people to know the significance of the Olympic games and their countries position in the performance. Even though mainstream media, social media promotes and telecasts news regarding Olympic games, it only happens during the time of the events. Also, there are no tools that represent the history of the games over the years which people can access whenever needed. Therefore, we propose an interactive visualization tool for the Summer and Winter Olympics that presents, the number of medals won, the country which won, details of the athletes, categories of sports throughout the years.

Questions:

The aim of the project is to :

1. Effectively design a visualization system where users can interact and gain information on the Olympic games.
2. User should be able to select the year to see the results of the games pertaining to that year.
3. Is the country's GDP and the number of medals won correlated?
4. How is the participation of women athletes changing from history?
5. Comparison between Gold, Silver and Bronze medals won by each country?

Data

This is a historical dataset on the modern Olympic Games, including all the Games from Athens 1896 to Rio 2016. Note that the Winter and Summer Games were held in the same year up until 1992. After that, they staggered them such that Winter Games occur on a four-year cycle starting with 1994, then Summer in 1996, then Winter in 1998, and so on. A common mistake people make when analyzing this data is to assume that the Summer and Winter Games have always been staggered.

Dataset:

<https://www.kaggle.com/heesoo37/120-years-of-olympic-history-athletes-and-results>

Data Preprocessing:

Need to drop unused columns and limit the number of countries represented to 20 using python. Need to filter out noisy data and remove the duplicates. Data regarding countries' GDP need to be gathered and combined with Olympics dataset.

Data files used are:

athlete_events_modified.csv

gdppc.csv

truncated_latest.csv

Used the csv data from the above files and converted into json using d3.nest() method.

Design Evolution:

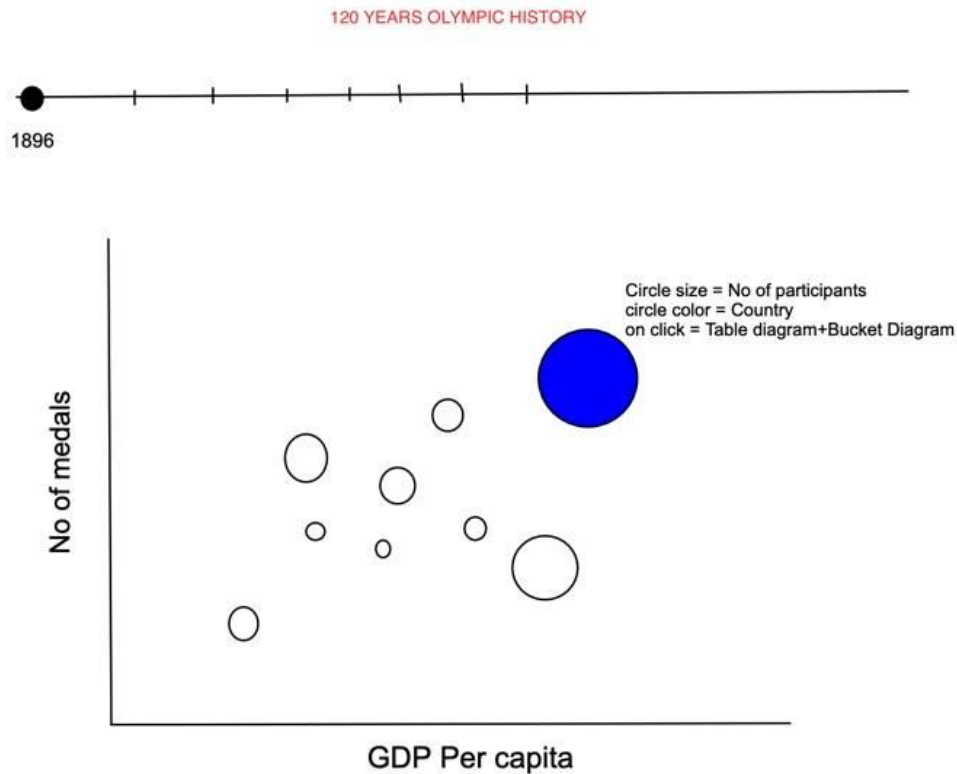
Design 1:

Year slider:

It synchronizes all the views in this design with a particular year. (i.e 1896 - 2016). It also shows information related to the location of the Olympics in that particular year.

Scatterplot

Scatterplot shows a comparison between the number of medals won by a country and its GDP per capita. How is an increase in GDP resulted in an increase in participation from that country and also an increase in the number of medals won by that country over a period of 120 years.



Marks and Channels













- Color represents the country name.
- Size of the circle represents the total number of participants in the Olympics from that particular country
- The position of the circle encodes the number of medals won by the country and its GDP per capita.
- Year slider fetches information for that particular year.
- By clicking a particular circle that circle gets highlighted, the rest of the circles are diluted and also a drop-down listing information related to each and every sport played by that country is seen in Medals Table.

Table Chart:

This table showcases information on the number of medals (gold, silver, bronze) won by each country in each sport. All the columns can be sorted.

This table is linked to year slider, it fetches information of a particular year.

Initially only aggregated values of medals won by a country are shown as bars, on clicking a specific country, we can see a drop-down listing information related to each and every sport played by that country and also respective country is highlighted in scatterplot.

COUNTRY	GOLD	SILVER	BRONZE	TOTAL
USA				
<i>Swimming</i>				
<i>Gymnastics</i>				
<i>Shooting</i>				
<i>Weight lifting</i>				
CHINA				
<i>Swimming</i>				
<i>Gymnastics</i>				
<i>Shooting</i>				
<i>Weight lifting</i>				

Marks and channels:

Marks = Horizontal Bars

Channels = Colors, Length of the bar.

Gender Inequality

Each of the circles represent a particular athlete, hover on the circle to get athlete information. These athletes are colored based on their gender and segregated into different sports categories they represent.

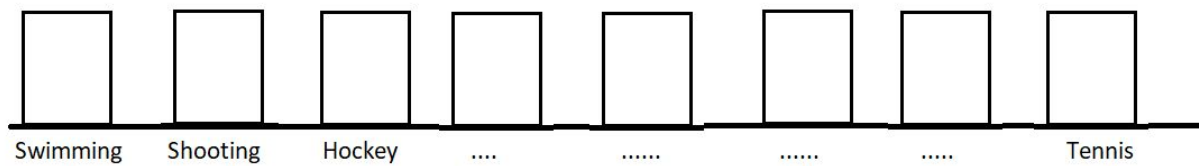
There is a toggle switch provided at the top, which when pressed all these athletes are sorted into three different bins based on their gender. This shows which shows the inequality between different genders.

These bins are synchronized with the year slider, showing data related to a particular year.

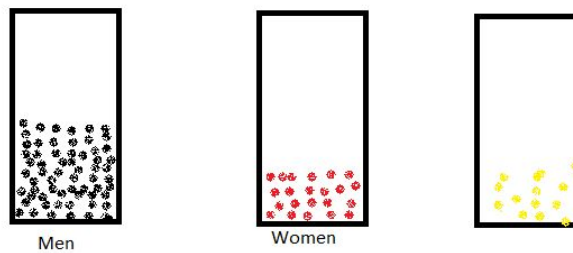
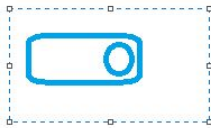
Categorical Data:



Aggregated Data



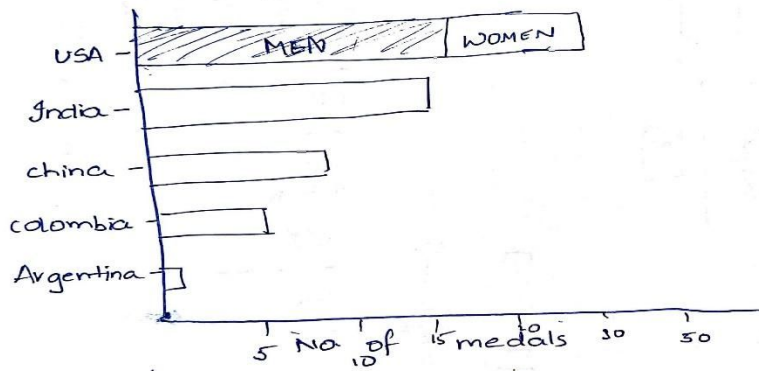
Gender Inequality



Bar Graph

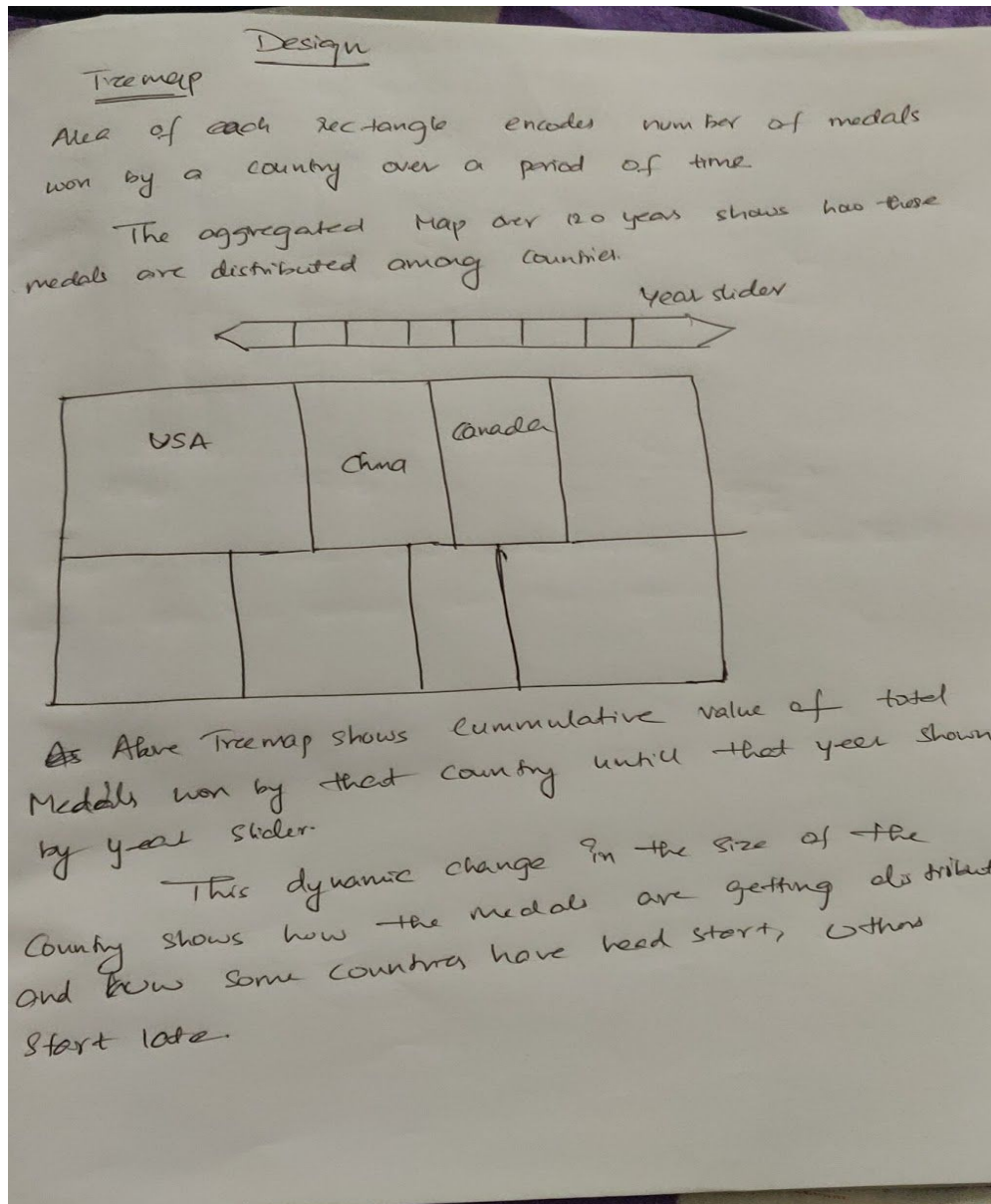
This bar graph shows, ranking of different countries based on the number of medals won by them until that particular year. It is also synchronized by year slider with other views showing data until a particular year. Y-axis represents the country name and X-axis represents the number of medals won. It also encodes male and female medal winner in the bar graph using color, basically it is a stacked bar graph.

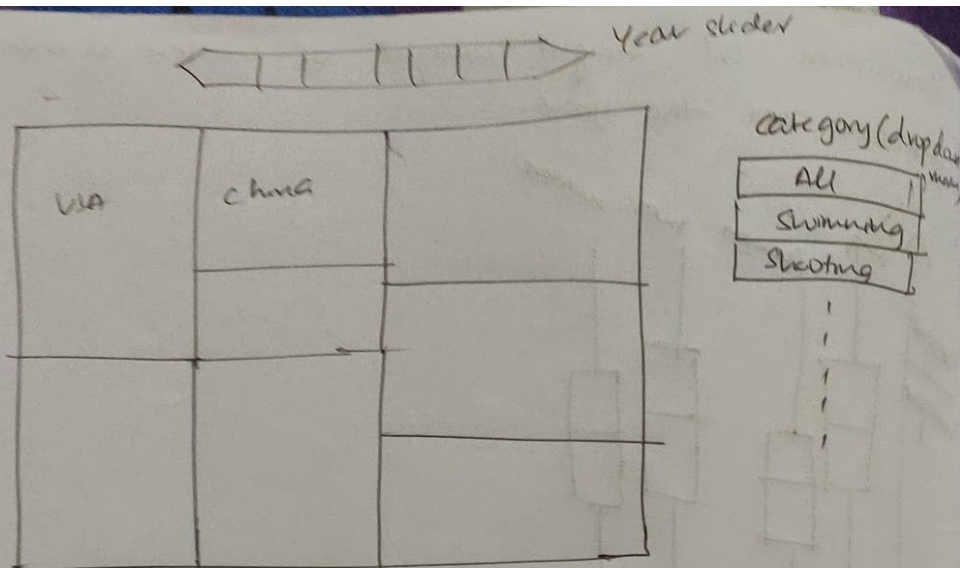
TOP 10 COUNTRIES RANKING TILL YEAR



Design 2:

Tree map





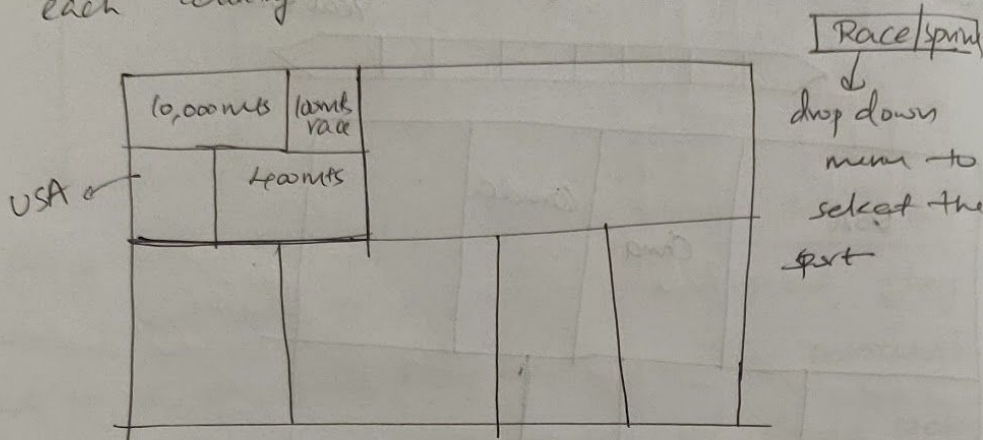
Along with Aggregated tree map that shows all the medals won by a country over a period, they will be a dropdown menu to choose one particular sports category, which when selected shows statistics for that one ~~the~~ sport alone.

Tree map represents Medals won By top 15 countries only.

Mark: ~~A~~ Rectangular shape

Channel: Area of Rectangle

treemap further gives us information about subcategory of each sport, like mixed double Men's singles, women's singles in tennis. We can get information about No. of medals won by each country in these sub-categories.



Color — encodes country

Area — Medals won by each country

Handwritten notes and diagrams illustrating the relationship between the number of trials and the probability of success.

Top Diagram: A line graph showing the probability of success (Y-axis) versus the number of trials (X-axis). The curve starts at 0 and increases, leveling off towards 1.0 as the number of trials increases.

Bottom Diagram: A bar chart showing the probability of success (Y-axis) versus the number of trials (X-axis). The bars represent the probability of success for different numbers of trials (1, 2, 3, 4, 5, 6, 7, 8, 9, 10). The probability increases with the number of trials, leveling off towards 1.0.

Text:

Handwritten notes in Hindi and English:

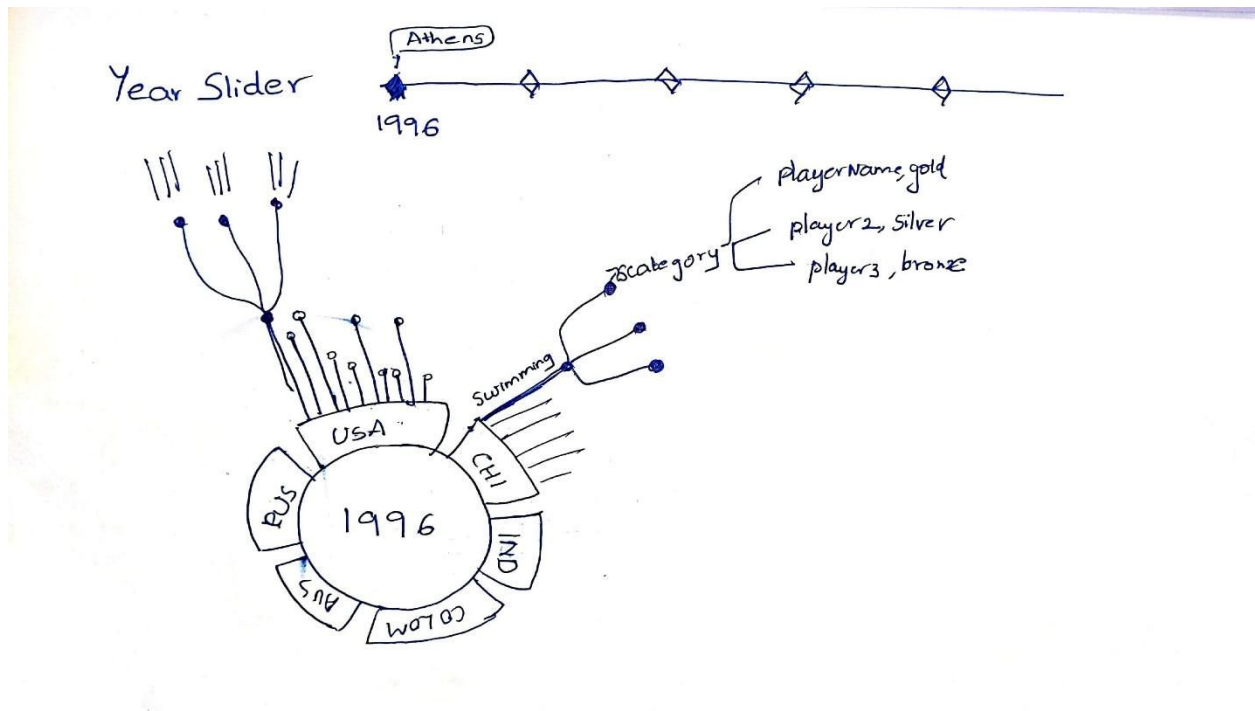
- Handwritten in Hindi: "प्रयोगों की संख्या बढ़ाई जाये" (Increase the number of trials).
- Handwritten in English: "The probability of success increases with the number of trials." (This text is written vertically along the right side of the bottom diagram).
- Handwritten in Hindi: "प्रयोगों की संख्या बढ़ाई जाये" (Increase the number of trials).
- Handwritten in English: "The probability of success increases with the number of trials." (This text is written vertically along the right side of the bottom diagram).

Above Line chart shows difference between no. of male and female Athletes over the years.

This actually shows gender gap between male vs female over the years.

Design 3:

Histogram



The above histogram is synchronized to year by year slider, wherein each segment represents a country and bars on top of the country represents medals won in each category, for example, swimming, shooting, tennis, etc.

When clicked on these bars it expands to represent sub-categories like 100mts Race, 400 mts race in running which further expands to list players who won the medals.

Word Map:

It shows aggregated values of medals won by the top 20 players in the world over the last 120 years.

Player with the highest number of medals is represented by the largest font, player with the lowest number of medals with the smallest font.

Final Design Implementation:

1) Scatter plot and Table chart

Choosing scatter plot from 1st design, as it can easily incorporate GDP data and number of participants. Also choosing table chart from 1st design as it acts as a support for the scatter plot as it gives more information like medals won by a country in specific sports category.

- In Scatter plot, X-axis, Y-axis , circle size parameters can be chosen.
- Play button is provided to observe how each country progressed in all the years.
- Color is organized for countries according to their. continent
- When a circle(particular) country is hovered, respective country is shown in table with much more details, by drilling down into number of medals own by the country in different sports category that year.

Team	Gold	Silver	Bronze	Total
Great Britain	22	12	20	54
Soviet Union	66	54	67	187
East Germany	31	64	23	118
Hungary	25	4	24	53
France	22	14	30	66
India	0	1	0	1
Italy	22	29	14	65
Australia	60	54	69	183
North Korea	0	3	1	4
Jamaica	0	3	20	23
Canada	4	23	4	31
Pakistan				
Japan	5	8	31	44
United States	130	51	61	242
China	39	17	23	79
Portugal	0	2	0	2
Qatar	0	1	0	1

Team	Gold	Silver	Bronze	Total
Great Britain	22	20	12	54
Sailing	3	4	0	7
Modern Pentathlon	1	0	1	2
Badminton	0	0	2	2
Rowing	13	4	0	17
Athletics	2	2	2	6
Canoeing	0	1	1	2
Equestrianism	0	4	0	4
Cycling	1	3	6	10
Shooting	1	1	0	2
Boxing	1	0	0	1
Judo	0	1	0	1

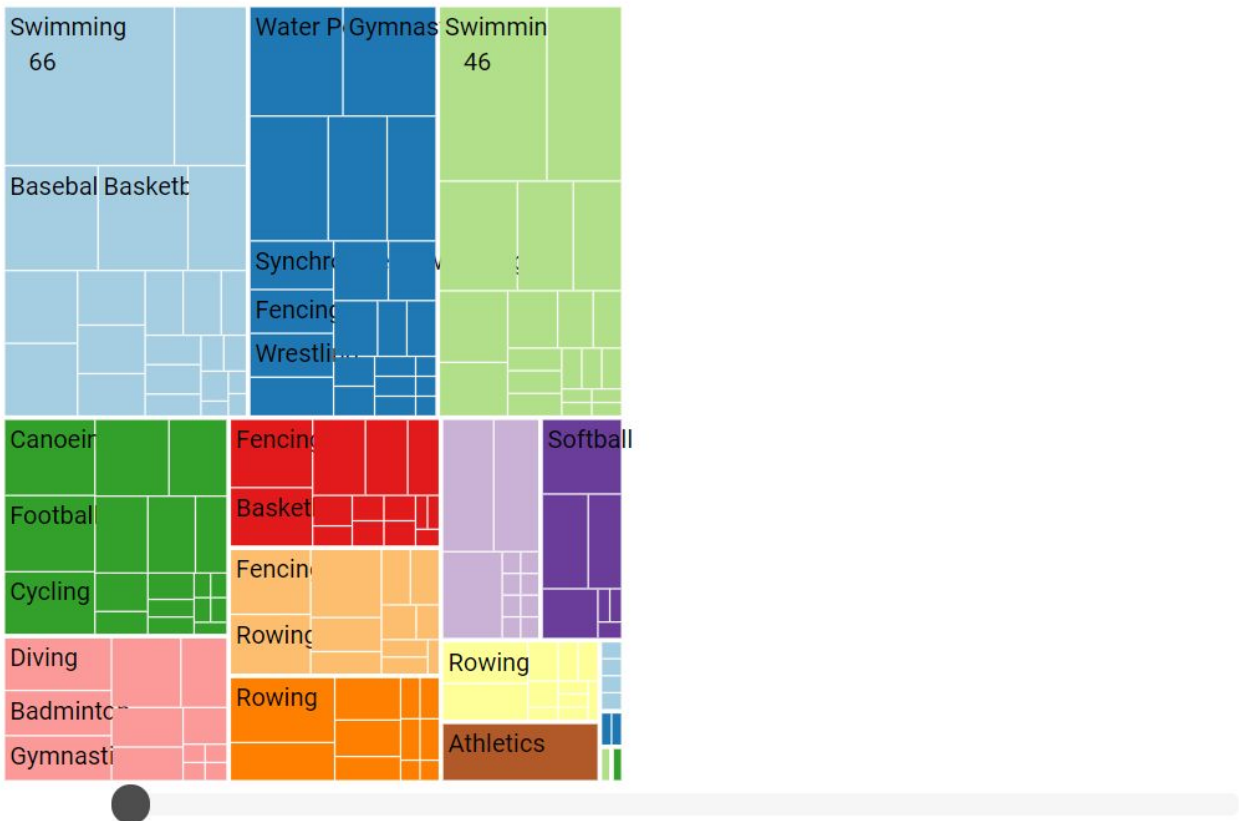
Team	Gold	Silver	Bronze	Total
United States	130	61	51	242
Soviet Union	66	67	54	187
Australia	60	69	54	183
China	39	23	17	79
East Germany	31	23	64	118
Hungary	25	24	4	53
France	22	30	14	66
Italy	22	14	29	65
Great Britain	22	20	12	54
Japan	5	31	8	44
Canada	4	4	23	31
Jamaica	0	20	3	23
North Korea	0	1	3	4
Portugal	0	0	2	2
India	0	0	1	1
Qatar	0	0	1	1

Treemap and Line Chart

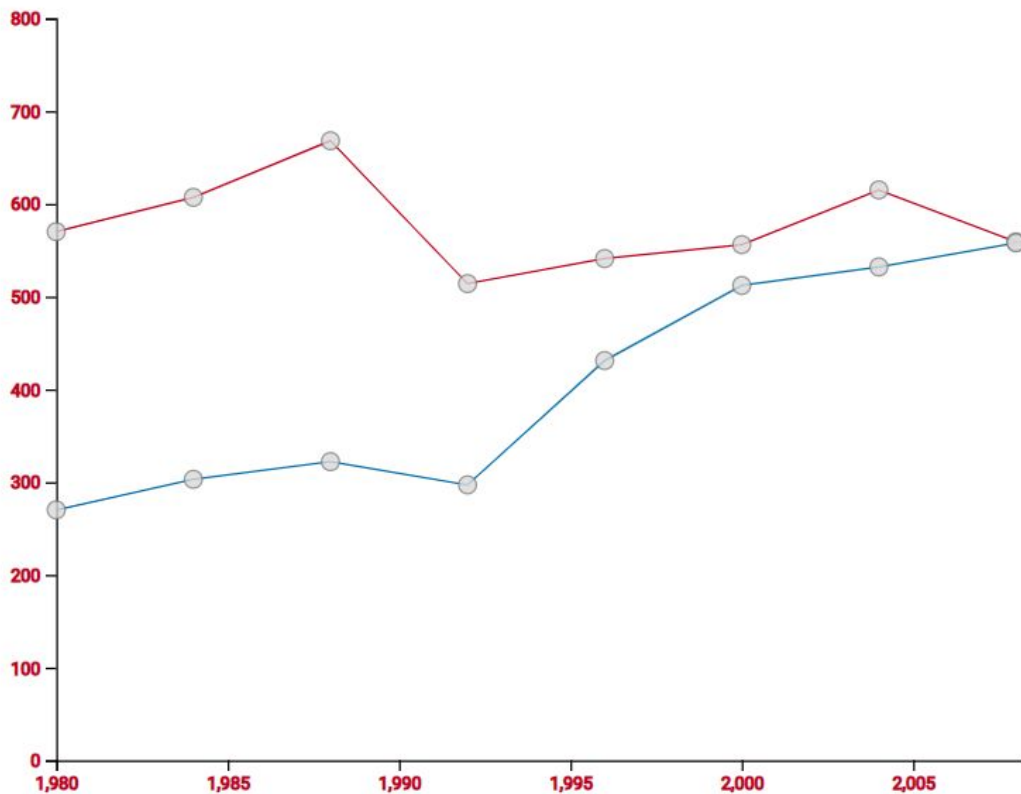
- Treemap shows distribution of medals among different countries
- If there is enough space, name of the sport and medals won by a particular country in that sport is displayed if not it is blank.
- Hover option is provided with tool tip to view country name, medals won and sports category.

- Treemap is synchronised with year slider to observe how each country progressed in each sports category over the years.

Olympic medals distribution

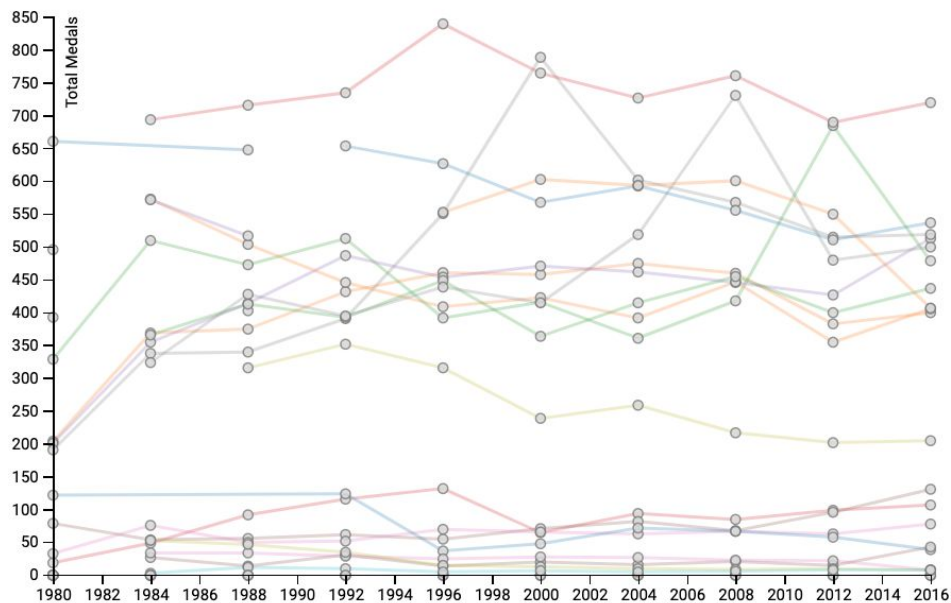


Line Chart(gender gap):



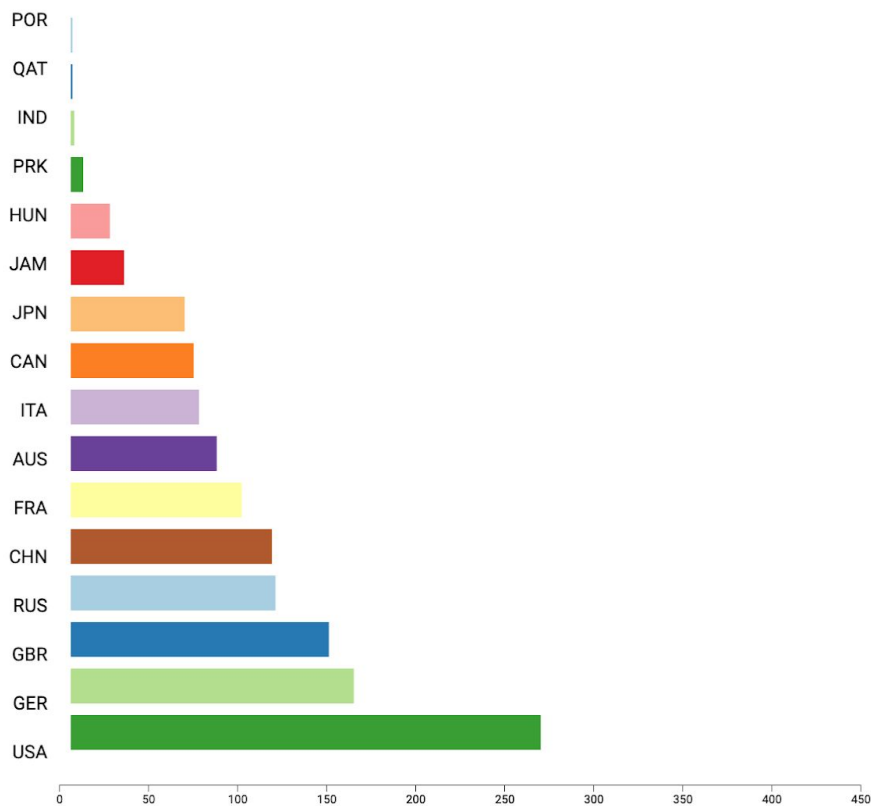
- This Line chart shows number of female and male participants over the years all over the world.
- Red line represents male participants, blue line represents female participants
- Hover option is provided to get exact value at each node.
- A drill down option is given to drill down into each country and observe gender gap in that country.

Over the years comparison of medals won by each country



- This Line graph shows the total number of medals won by all the countries through the years 1980 to 2016.
- When you hover over a line, the name of the country is displayed on the graph and the line is highlighted.
- When you hover over the circles, the count of medals won by that country in that year is presented.
- This graph helps in understanding which country started participating in which year and how they improved or decreased the number of medals won over the years.

Bar graph showing ranking of different countries based on number of medals won:



- This bar graph shows all the countries performance in number of medals won untill that year.
- A year slider is provided to move across all the years.
- The bars are presented in sorted in ascending order so that they can be compared throughout years.
- Transition is provided for the bars in the graph.