

# INTRODUCTION

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data. It transforms complex data sets into visual formats that are easier to interpret, allowing for quicker decision-making and more effective communication of insights.

## Key Concepts in Data Visualization:

1. **Charts and Graphs:** Common types include bar charts, line graphs, pie charts, scatter plots, and histograms. Each type serves different purposes depending on the nature of the data and the insights you want to extract.
2. **Dashboard:** A dashboard is a collection of visualizations displayed together, often in real-time, providing an overview of key metrics and trends.
3. **Data Storytelling:** This involves using visualizations to narrate the story behind the data, making it easier for audiences to understand and engage with the insights.
4. **Interactivity:** Modern data visualizations often include interactive elements like filters and drill-downs, allowing users to explore data from different angles.

Tableau is a leading data visualization tool that empowers users to transform complex data sets into interactive, easy-to-understand visualizations and dashboards. With its intuitive drag-and-drop interface, Tableau allows users to create a wide range of charts and graphs without needing advanced technical skills. It connects to various data sources, including spreadsheets, databases, and cloud services, enabling real-time data analysis and integration. Tableau's interactive dashboards offer dynamic features like filtering and drill-downs, making it possible to explore data from multiple perspectives. Its capabilities in advanced analytics, coupled with robust sharing options through Tableau Server and Tableau Online, make it a powerful solution for data-driven decision-making and effective communication of insights.

## **ABOUT THE DATASET**

### **1.History of Summer Olympics.xlsx (from 1896 to 2020)**

1. Year = the year of the Olympics.
2. Host Country = the country hosting the Olympics.
3. Country = the countries participating in the Olympics.
4. Bronze = the number of bronze medals won.
5. Silver = the number of silver medals won.
6. Gold = the number of gold medals won.
7. Is Host = whether or not the country was the host of the Olympics.
8. Total Medals = total number of medals won.

### **2.all\_athlete\_games.csv**

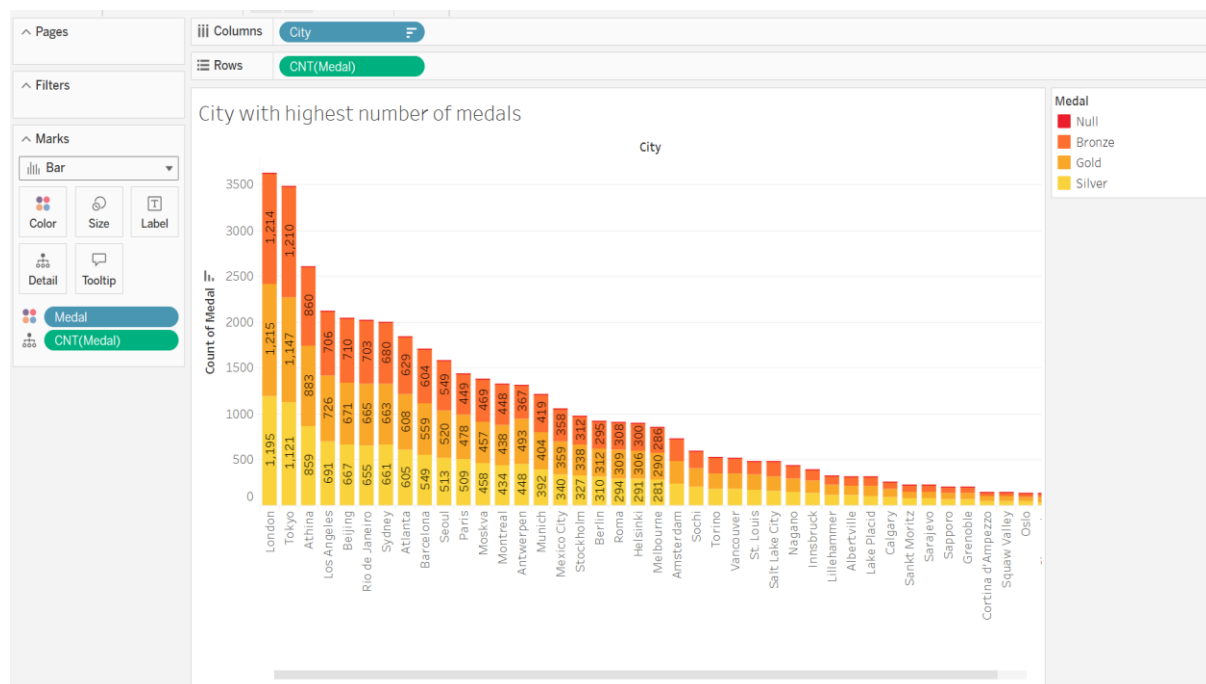
1. Entry ID = Index Number
- 2.Name = Name of athlete
- 3.Gender = Gender of Athlete
- 4.Age = Age of athlete
- 5.Team = The team to which athlete belong
6. NOC = NOC Representative of Country to which athlete belong
7. Year = Year in which athlete participated
8. Season = Olympics Season in which athlete participated
- 9.City = The city which hosted Olympics in that year
10. Sport = The sport in which athlete participated
11. Event = The event in that sport
12. Medal = Medal won by athlete (if any)

# PROBLEM STATEMENT:

“To identify the top-performing athletes and countries based on medal counts, examine the most prevalent sports, and analyse the gender distribution of medallists and the effect that the host country have in the medals won at the Olympics. Additionally, we seek to explore patterns in participation by evaluating the number of participants from each city and year, as well as the age distribution of athletes. The goal is to uncover meaningful patterns and trends that can inform strategic decisions for future Olympic participation and planning.”

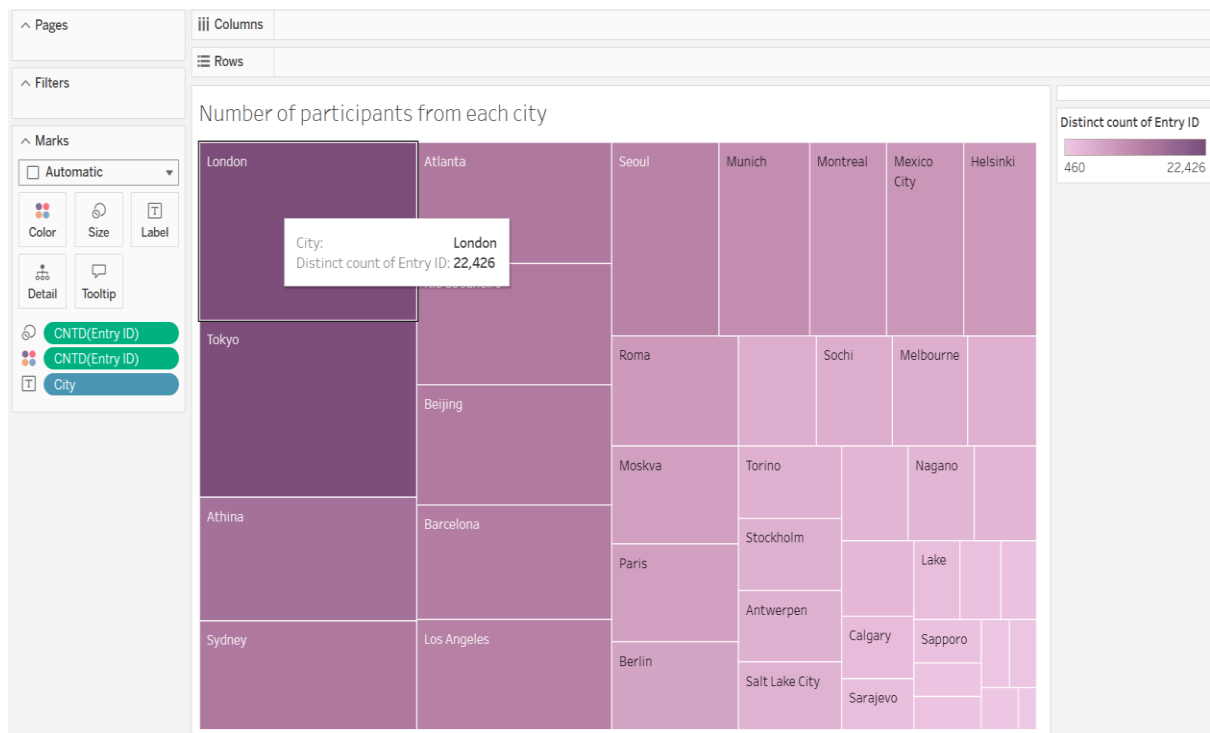
## ANALYSING THE PROBLEM – PART 1

### 1.To find the city with highest number of medals



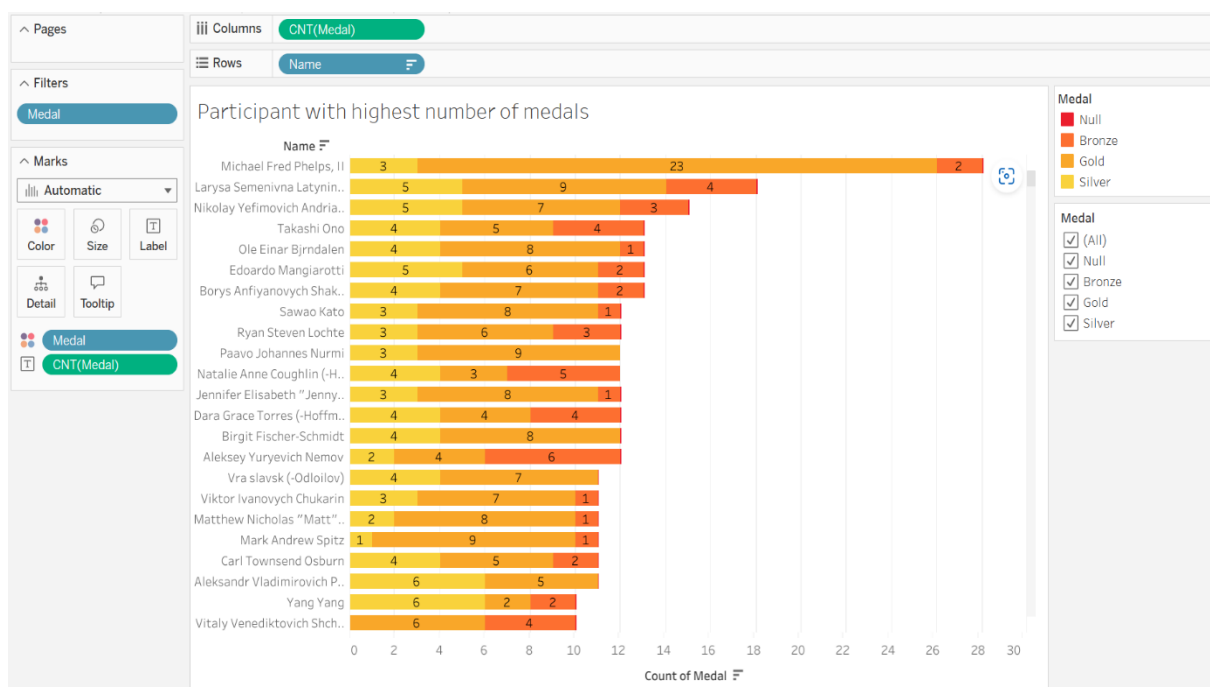
Inference: London is the city with the highest number of medals (1896 – 2020). The number of gold medals is 1,215, Silver is 1,195 and Bronze is 1,214 with a grand total of 3,624 medals

## 2.To find the total number of participants from each city from 1896 to 2020



Inference: Here, again we find that London is the city with highest number of participants across the years which is 22,426. This explains the high number of medals won. Second highest is Tokyo.

## 3.To find the participant with highest number of medals



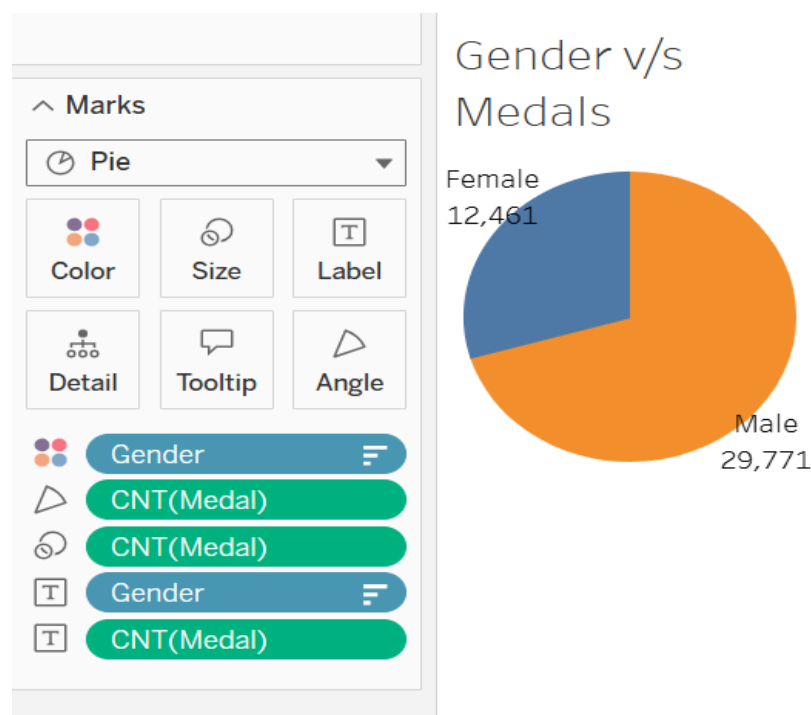
Inference: Olympic athlete with highest number of medals is Michael Fred (male category). Phelps, a swimmer, has accumulated a total of **28 Olympic medals** over his career. His medal tally includes:

- 23 Gold Medals, 3 Silver Medals, 2 Bronze Medals

The second highest, also the top female athlete is Larysa Semenivna Latynina. Latynina, a gymnast, has won a total of **18 Olympic medals** throughout her career:

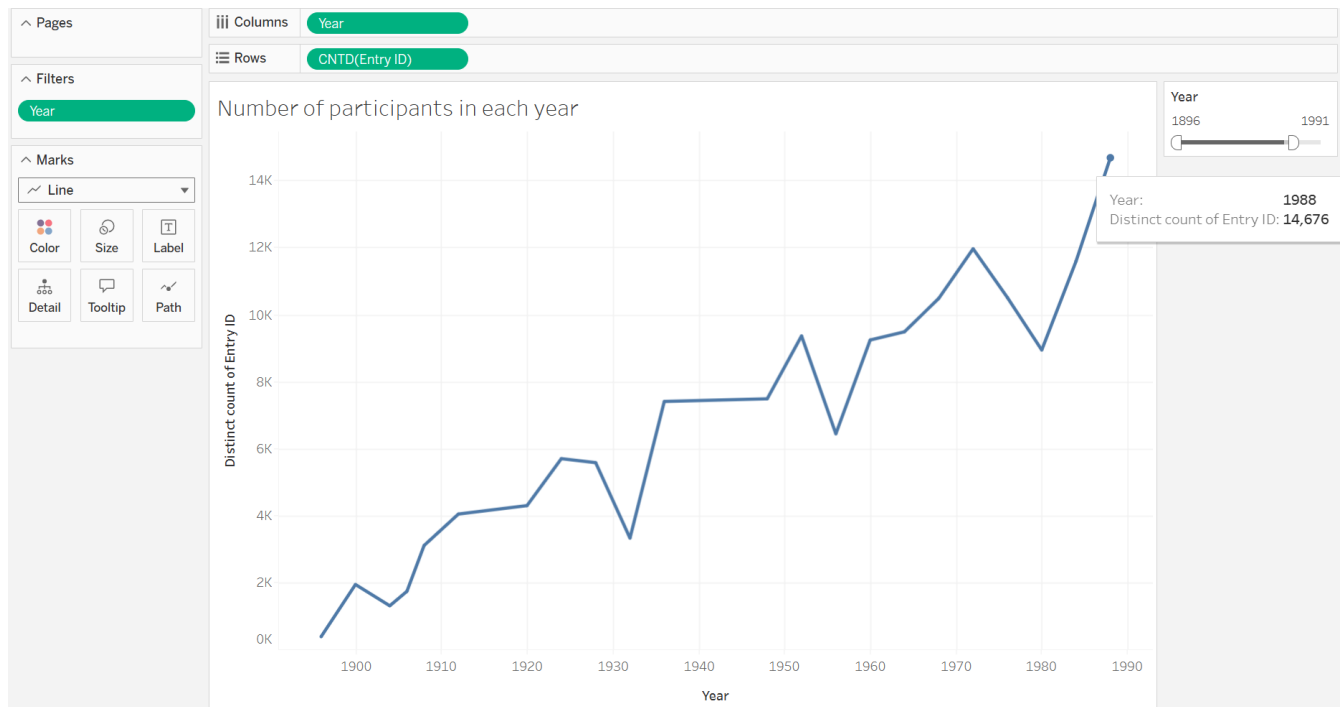
- 9 Gold Medals, 5 Silver Medals, 4 Bronze Medals

#### 4. Female v/s Male ratio distribution



Inference: We see that a total of 29,771 male players and 12,461 female players have participated in Olympics over the years. Initially, male athletes vastly outnumbered female athletes. This imbalance was stark, with female participation often less than 10% of the total. Female participation has grown substantially, particularly since the 1980s. By the 2010s and 2020, women made up nearly half of the total participants, reflecting broader societal changes towards gender equality in sports.

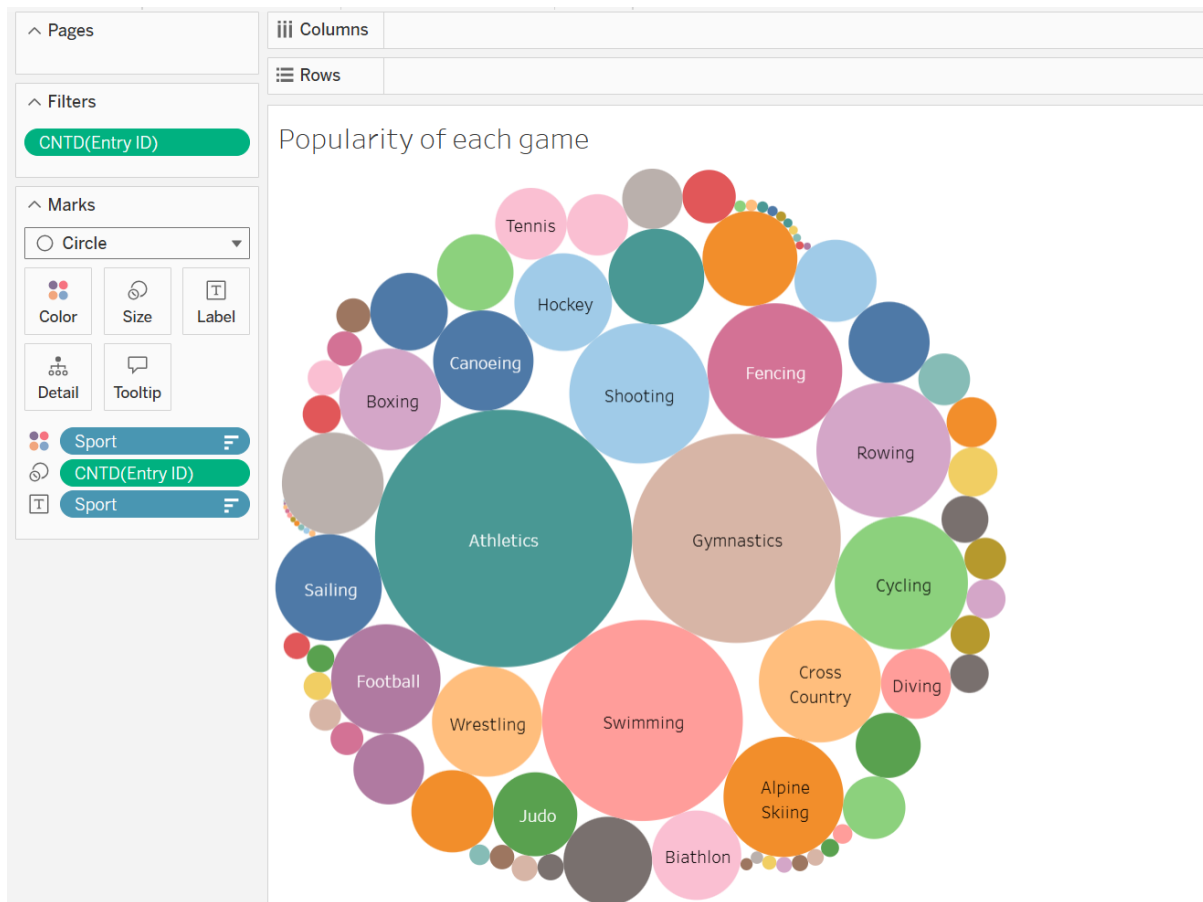
## 5.The trend of number of participants over the years



Inference: The number of participants has grown from a few hundred in the early Games to over 11,000 in recent Olympics. The Olympics have expanded to include a more diverse range of countries and athletes, reflecting the global nature of the Games. The trend highlights the growing inclusivity and scale of the Olympics, reflecting broader societal changes and the increasing global interest in sports.

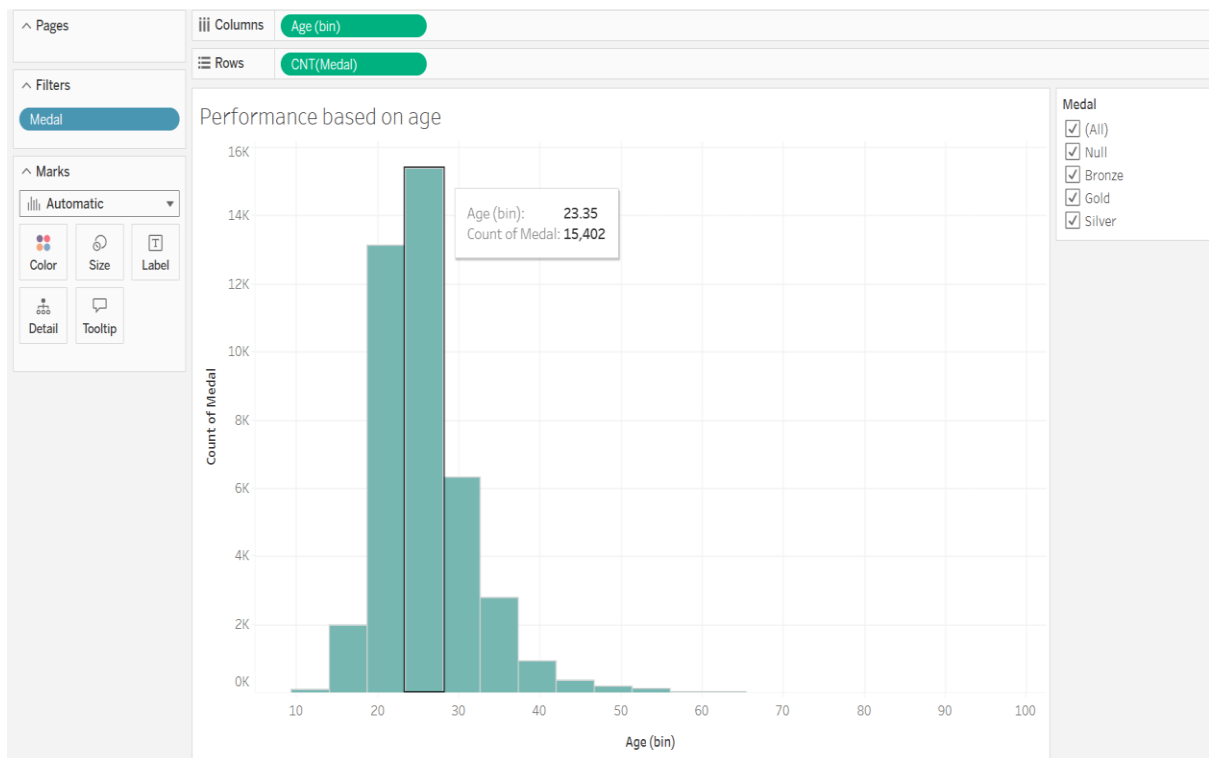
## 6. To find the most popular game of all time

Inference: Determining the "most popular game of all time" can depend on the criteria used, such as global reach, cultural impact, or sales figures. Athletics, Gymnastics, Swimming, Shooting, Fencing, Rowing are among the most popular games of all time.



## 7. To find the most effective age group for participating in the Olympics

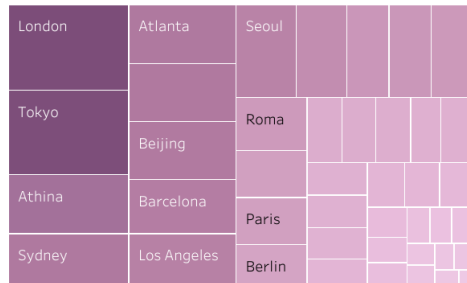
Inference: When considering the most effective age group for Olympic athletes across all sports collectively, we find that age group of 20 – 30 yields more number of medals. Athletes in their 20s and early 30s are often at their physical peak, possessing the strength, speed, endurance, and agility required for high-level competition. This age range allows athletes to harness peak performance while also benefiting from years of training and experience. This range encompasses the optimal combination of physical capabilities, experience, and strategic understanding necessary for peak performance at the Olympic level.



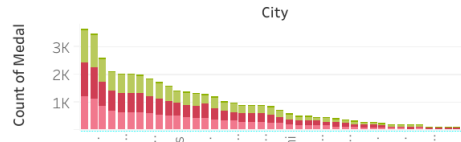


# DASHBOARD 1

Number of participants from each city

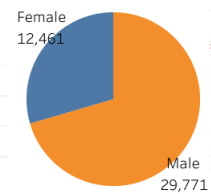


City with highest number of medals

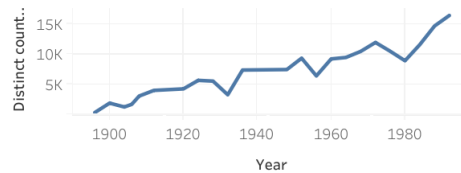


Medal  
Null  
Bronze  
Gold  
Silver

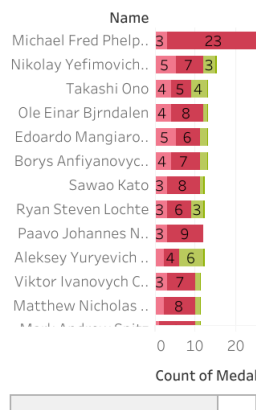
Gender v/s Medals



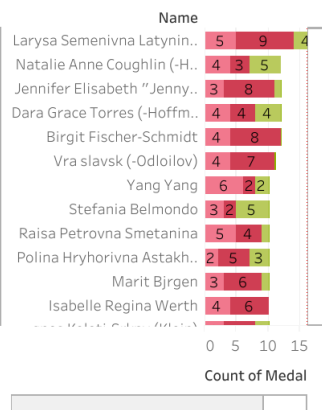
Number of participants each year



Best male performer



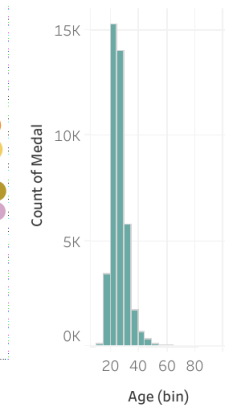
Best female performer



Popularity of each game



Performance based on age

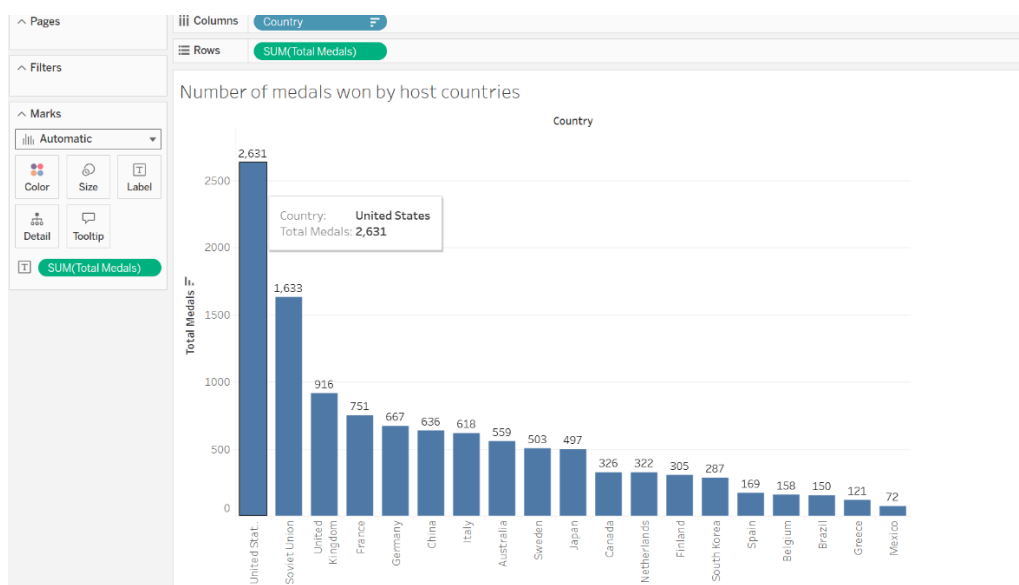
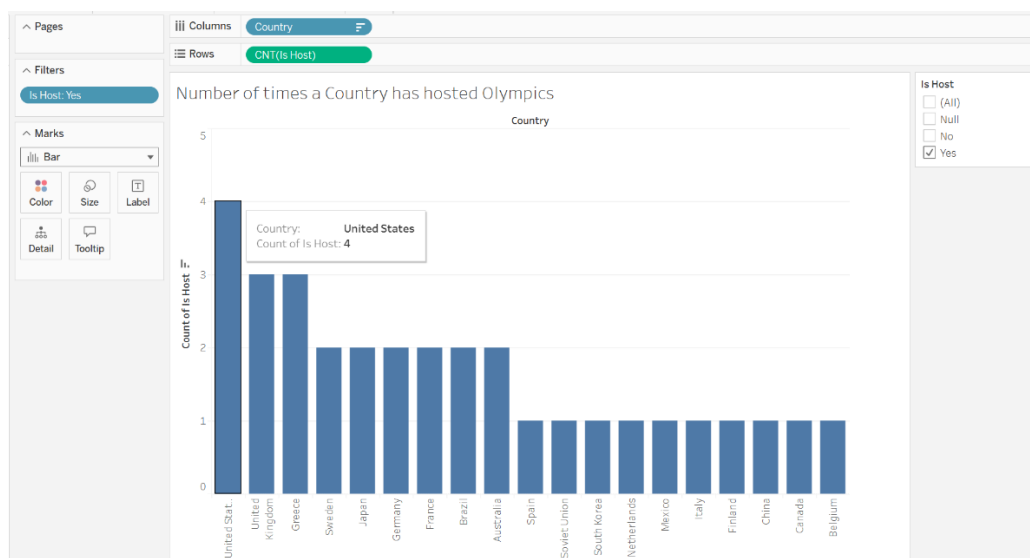


This dashboard gives a comprehensive view of all the graphs analysed earlier and gives useful insights like identifying which countries consistently perform well and have the highest medal counts, suggesting strong national sports programs and effective athlete training, how the number of medals or the performance of countries and athletes has evolved over the years, potentially highlighting improvements or declines in certain sports or nations, determines the most successful age groups, which could indicate the peak performance age for various sports or events, analyse the gender distribution among athletes and medallists to understand trends in gender equality and participation in different sports, recognize which sports are most popular or competitive. These insights can help stakeholders, such as sports organizations, policymakers, and fans, understand patterns in Olympic success and make informed decisions about training, policy, and future events.

# ANALYSING THE PROBLEM – PART 2

While watching the Olympics we often come across the question “What effect does the host country have in the medals won at the Olympics?”. The advantage of host country in any sporting activity is well known, as the participants will have familiarity of the field, and also there is a great support from the home crowd. Host countries are expected to win 3 times the medals that they were winning while playing as away. We can answer this question with evidence by analysing the Olympic host dataset using charts and graphs that give a clear and deeper understanding. We do this analysis in 2 parts:

## 1.The number of times a country has hosted Olympics and the number the medals they have won over the years

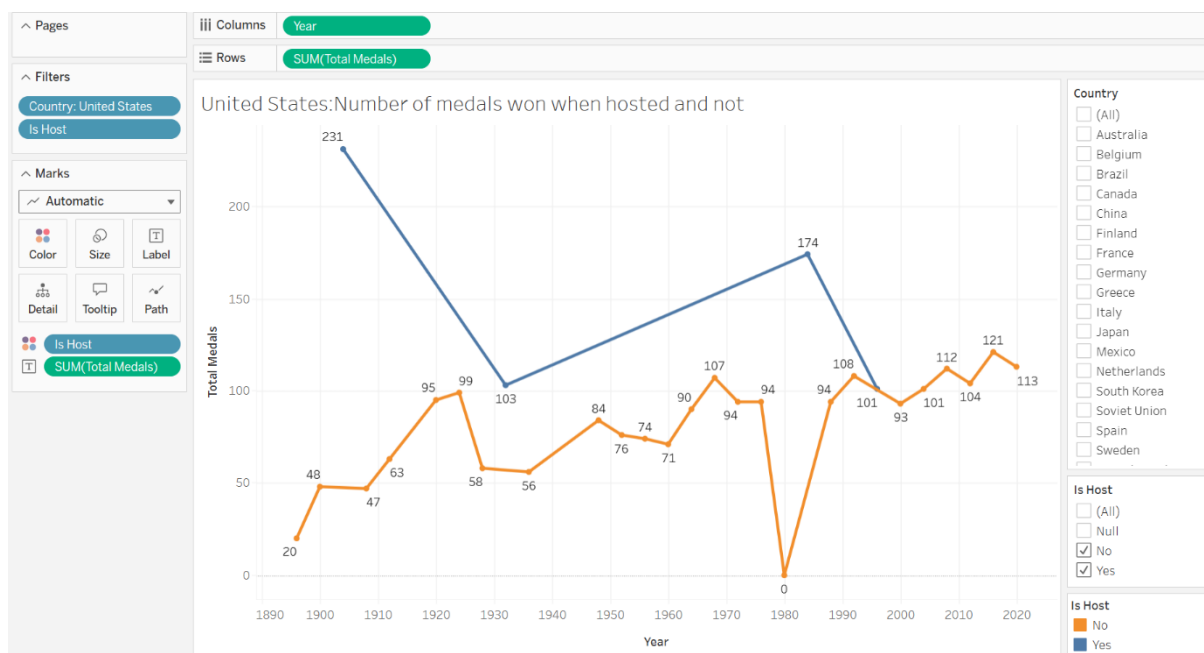


The first graph shows that the USA has hosted it the most, i.e. 8 times followed by France, which had hosted it 5 times. The second graph shows the total medal winnings of countries that hosted Olympics, and shows that USA has won the most medals, and then it is Germany, followed by France. This hints that the countries which have hosted more Olympics have won more medals.

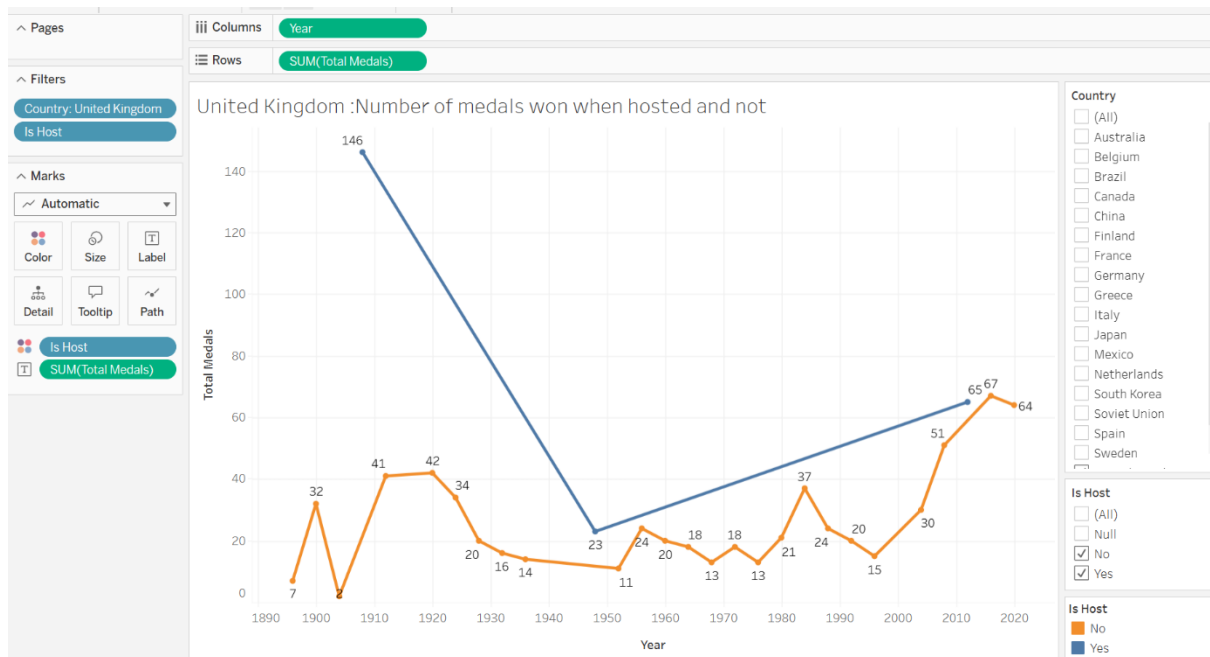
## 2.The variation of number of medals when a country is the host

Firstly, looking at the USA's medals won, we can see that the number of medals won by it has always increased when it hosted the Olympics. In some cases, the host country prevails in successive Olympics after they host, such as Japan. Greece (2004), United Kingdom (2012) and United States (1996), have won between 10–20 % more medals when they were the hosts. Its only Brazil (2016) which failed to score more when they had hosted it (20 per cent less than when UK Olympics in 2012), whereas Greece managed to bag only 20 percent medals in 2008 after the Greece Olympics.

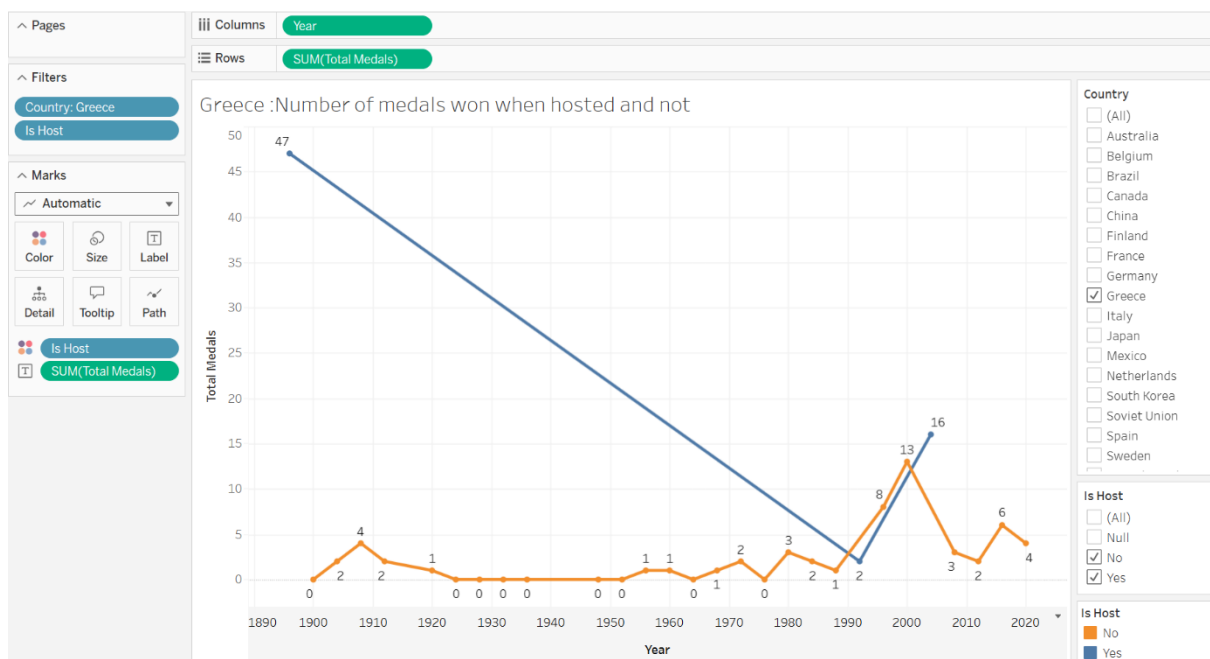
- USA



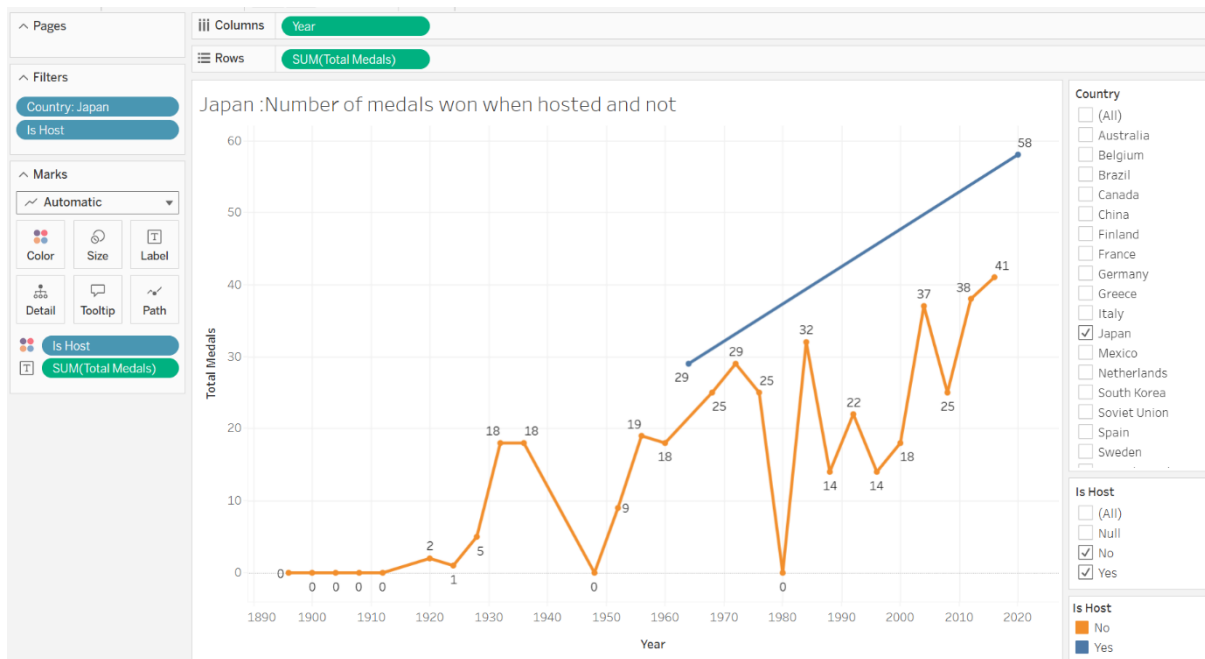
- United Kingdom



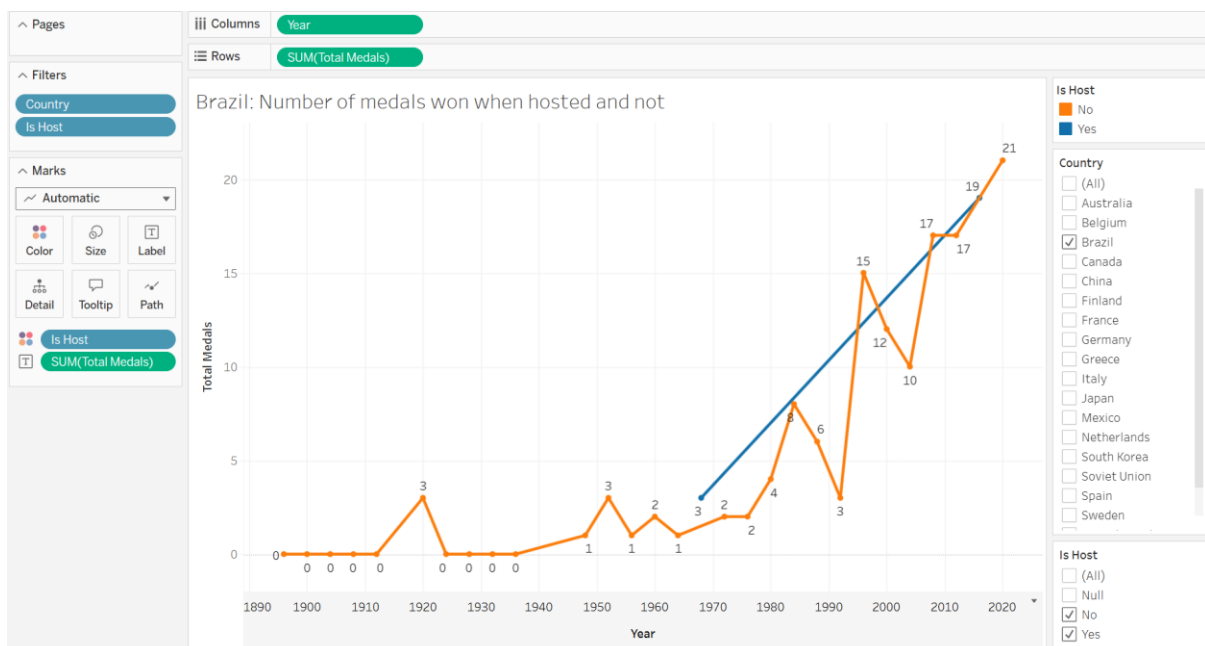
- Greece



- Japan



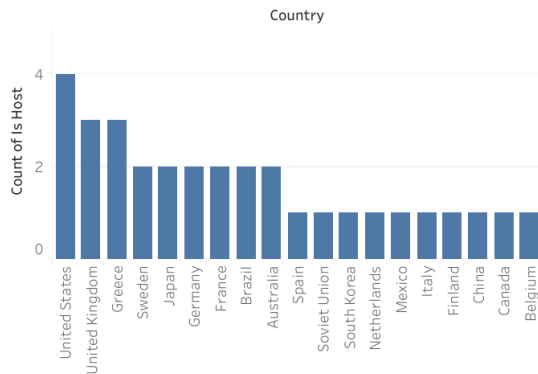
- Brazil



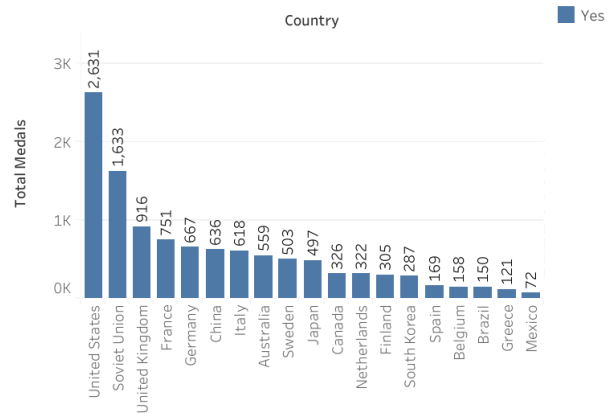
It is clear that the host countries have always a better chance of winning medals in the Olympics; they can win at least 10–20 percent more medals. The reasons why the host countries win more medals are due to home field advantage, crowd support and easier qualification criteria, hence more participation. However, it's not a guaranteed outcome for every host nation, and many factors contribute to the overall success of athletes at the Olympics.

# DASHBOARD 2

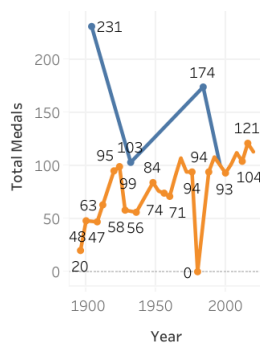
Number of times a Country has hosted Olympics



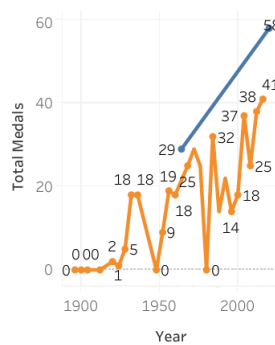
Number of medals won by host countries



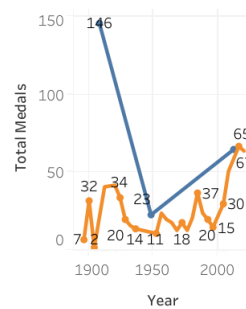
United States: Number of medals won when hosted and not



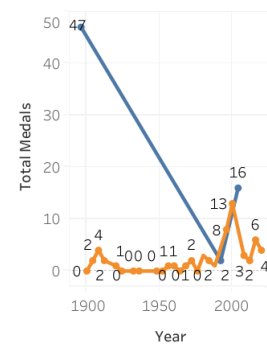
Japan: Number of medals won when hosted and not



United Kingdom: Number of medals won when hosted and not



Greece: Number of medals won when hosted and not



This dashboard is focused on showing various examples of how a host country has higher chances of winning more medals at Olympics. This dashboard shows graphs of USA, UK, Japan and Greece where the blue line indicates the hike of medals when these countries hosted and the orange line shows trend of medal count when they were not the hosts. It is very clear that host countries have a higher advantage than other participating countries.