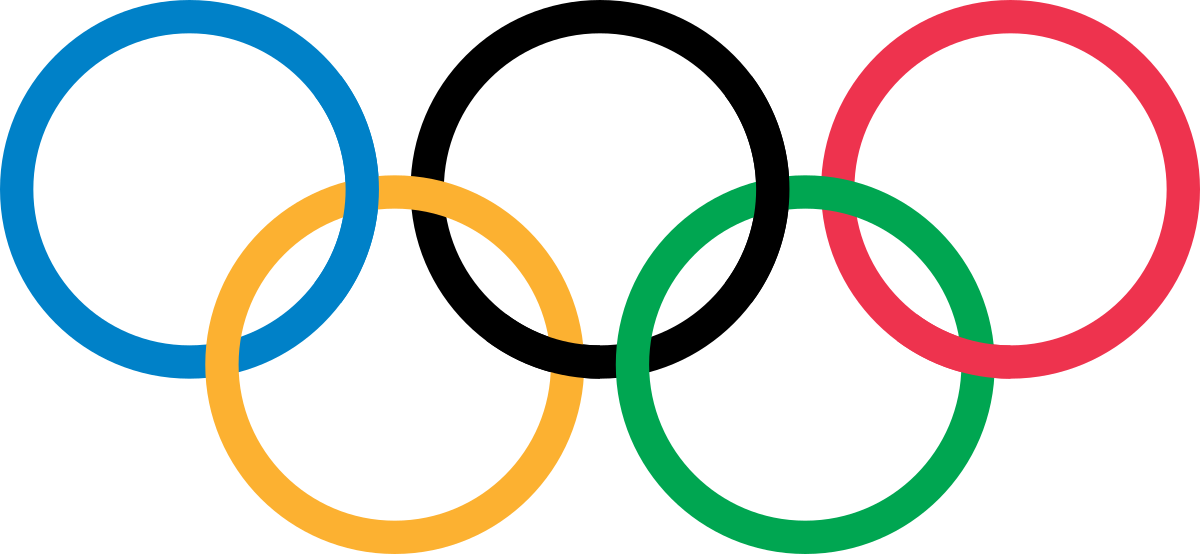
Capstone Project – Sports Analysis



Ajay Priyakar M

Ajay Priyakar M

**Overview**

The objective of this analysis is to gain valuable insights into the historical data of the Olympic Games by exploring various dimensions, including games, sports, events, participants, medals, and regional representation. The scope encompasses data from multiple editions of the Olympics, aiming to identify trends, patterns, and significant insights from different aspects of the Games. The primary goal is to provide a detailed overview of the Olympics' history and evolution, and the analysis will generate recommendations to enhance future Olympics' organization, sports selection, and representation of diverse regions. The final deliverable will be a comprehensive report and presentation showcasing the findings and historical significance of the Olympic Games

**The Process**

1. Data Acquisition from GitHub:

Obtain the requisite dataset from a designated GitHub repository, containing essential information on Sports Analysis, encompassing various sports and the performance of the participants across the regions and information about hosted Cities of the Olympics.

1. Data Cleaning and Transformation:

Data cleaning involves identifying and handling missing values, removing duplicates, and addressing outliers. Transformation includes feature encoding categorical variables to prepare data for analysis and modelling.

1. Connection with Tools:

Establish connections between the dataset and various analytical tools. Interface the dataset with Power BI and Excel facilitating seamless data integration and processing.

1. Problem Statement Solution in Power BI:

Utilize Power BI to delve into the specified problem statements. Employ its robust features for data visualization, exploration, and analysis, effectively deriving insights and solutions.

1. Exploratory Data Analysis (EDA):

Perform exploratory data analysis using either Excel or SQL Workbench, depending on the complexity of the analysis. Extract meaningful patterns, relationships, and trends from the data to inform subsequent decision-making.

1. Creation of Visual and Insightful PowerPoint:

Develop a compelling Power BI presentation by designing interactive and informative visualizations, leveraging diverse data sources. Craft concise narratives that uncover meaningful insights and trends, allowing stakeholders to explore and comprehend data-driven stories effortlessly.

1. Detailed Presentation:

Engaging in these strategic steps ensures the creation of a compelling presentation that not only showcases your data analysis prowess but also empowers your audience with actionable insights for informed decision-making in this sports analysis.

**Objective**

The aim of this analysis is to extract valuable insights from historical Olympic Games data by investigating various dimensions, encompassing games, sports, events, participants, medals, and regional representation. The primary goal is to offer a comprehensive overview of the evolution of the Olympic Games, identifying trends, patterns, and significant findings. Moreover, the analysis seeks to formulate recommendations to optimize future Olympics in terms of organizational aspects, sports diversity, and inclusive regional representation.

Solutions:

The analysis will culminate in a comprehensive report and a dynamic presentation. The report will present a detailed account of historical insights, trends, and patterns uncovered from the Olympic Games data. The presentation will encapsulate these findings through visually engaging slides, featuring interactive visualizations, impactful charts, and maps. The final deliverables will highlight not only the historical significance of the Olympic Games but also provide valuable recommendations to guide the future organization and enhancement of this global sporting event.

**Data Dictionary:**

**Table: Sports**

* Sport ID: Unique Identifier for each Sports
* Sports Name: Name of each Sports

**Table: Events**

* Event ID: Unique identifier for each Events
* Event Name: Name of each event
* Sport ID: Foreign Key referencing the Sport ID field from the Sports Table

**Table: Medal**

* Medal ID: Unique identifier for each Medals
* Medal Name: Name of each Medals

**Table: Games**

* Game ID: Unique identifier for each Games
* Game Name: Name of each Games
* Game Year: The Game conducted year
* Season: The Game conducted Season

**Table: Person**

* Person ID: Unique identifier for each Person
* Full Name: Full name for each person
* Gender: Gender of the person
* Height: Height of each person
* Weight: Weight of each person

**Table: Games City**

* City ID: Foreign Key referencing the City ID field from the City Table
* Game ID: Foreign Key referencing the Game ID field from the Games Table

**Table: City**

* City ID: Unique identifier for each City
* City Name: Name for each City

**Table: Person Region**

* Person ID: Foreign Key referencing the Person ID field from the Person Table
* Region ID: Foreign Key referencing the Region ID field from the NOC Region

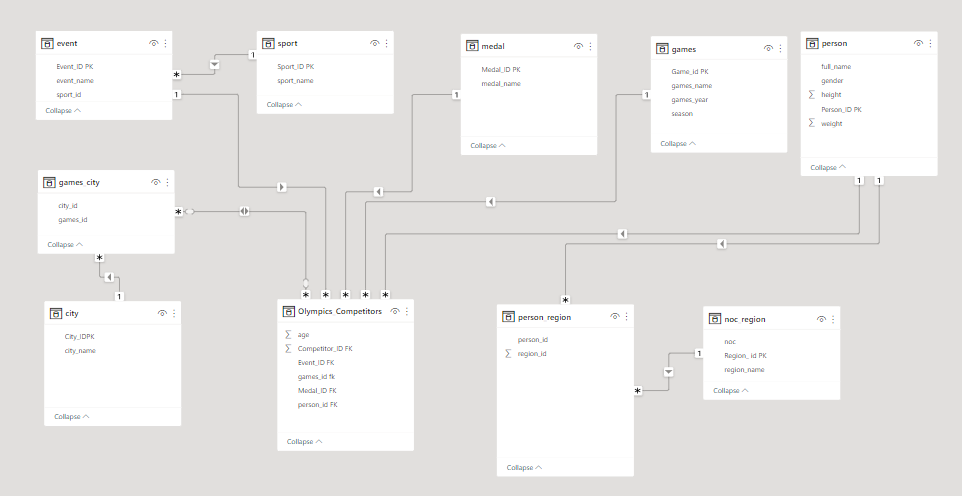
**Table: NOC Region**

* Region ID: Unique identifier for each Region
* NOC: NOC code for each Region
* Region Name: Name for Region

**Table: Olympics Competitor**

* Competitor ID: Unique identifier for each Competitor
* Event ID: Foreign Key referencing the Event ID field from the Event Table
* Game ID: Foreign Key referencing the Game ID field from the Games Table
* Medal ID: Foreign Key referencing the Medal ID field from the Medal Table
* Person ID: Foreign Key referencing the Person ID field from the Person Table

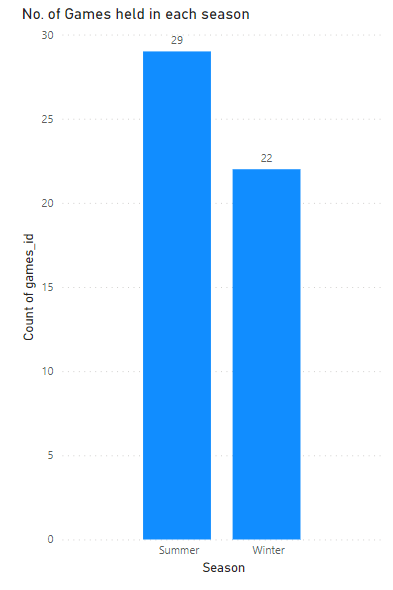
The data dictionary provides a comprehensive overview of the tables and their respective fields in the dataset. It outlines the relationships between the tables, allowing for a better understanding of the data structure and facilitating the design and implementation of the Power BI Dashboard.

**ER Diagram: **

Power BI Problem Statement:

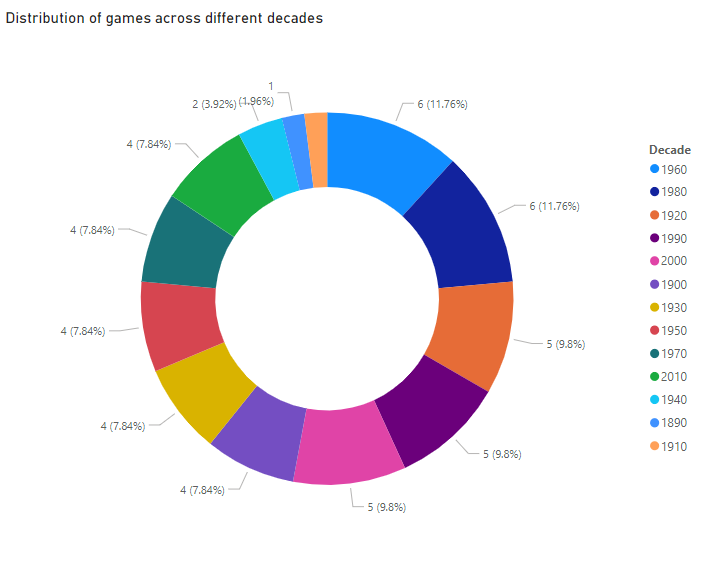
1.How many Olympic Games have been held in each season (Summer vs. Winter)?

Analysis of number of Olympic games held in each season offers the improvement of Participation in both seasons, brings equal interest on each season, Chances to spread the game over the both cold and normal regions, we can maintain more attention on Olympic since it happening in two spans as summer and winter in single year. From this visualization we can understand that both summer and winter season are conducted regularly from when the winter Olympic were introduced. Since it introduced in 1924 from 1896 to 1920 no winter Olympic takes place, So Summer Olympic has held 7 times more than the Winter Olympics.



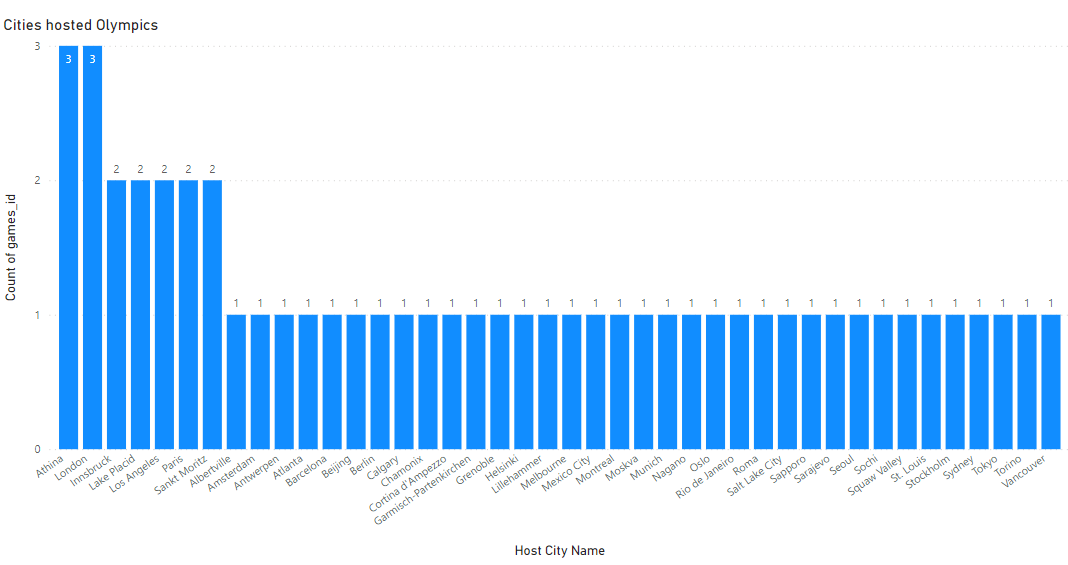
2. What is the distribution of games across different decades?

Analysis of the distribution of games across different decades shows the improvement of Games over the time. It helps to see the evolution of game improvement in the decades. From the below visualization we can find the valuable distribution of games in the decade.



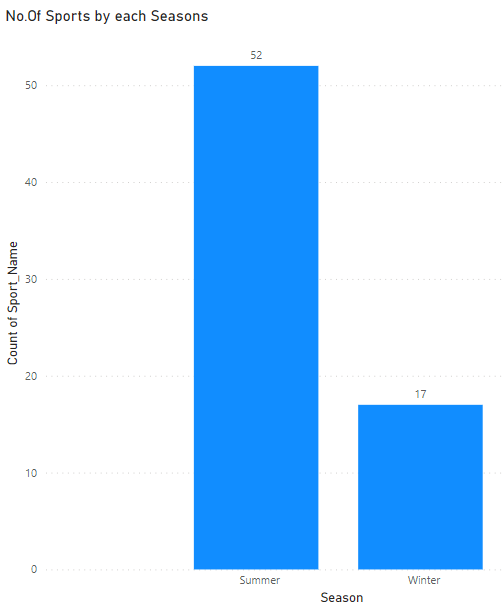
3. Which cities have hosted the most Olympic Games?

This analysis aids in understanding the evolution of the Games and the global recognition of certain cities. It provides insights into the infrastructural development and logistical expertise required for hosting such major events, highlighting successful models for future host cities. From the below visualization we can identify the City Athina, London are the most hosted city of Olympic until now and also we can see that some more country are also showing interest in hosting the Olympics



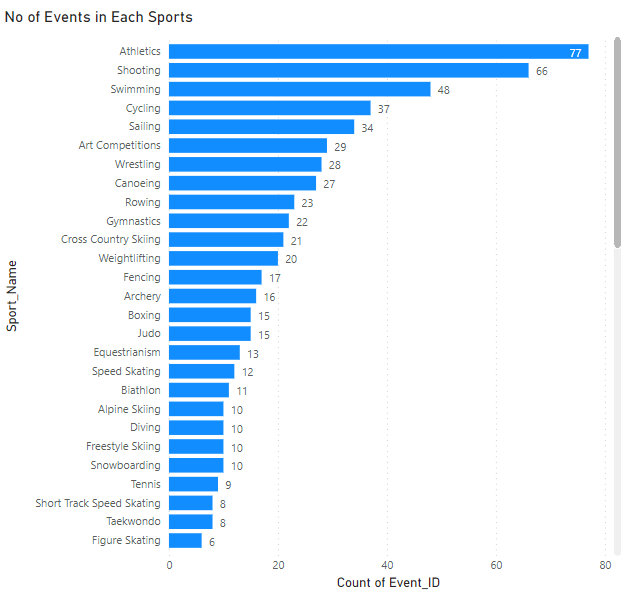
4. What is the distribution of sports between Summer and Winter Olympics?

The analysis of sports distribution between Summer and Winter Olympics reveals a distinct division based on climate and tradition. Summer Olympics encompass a wide range of sports, including athletics, gymnastics, and swimming, appealing to a broader global audience. Winter Olympics predominantly feature cold-weather sports like skiing, snowboarding, and ice hockey, catering to regions with colder climates and specific expertise. This analysis underscores the strategic allocation of sports to maximize viewership and participant diversity. It emphasizes the role of tradition and geographical considerations in shaping the sports lineup for each edition of the Games, reflecting a balance between cultural heritage and sporting innovation.



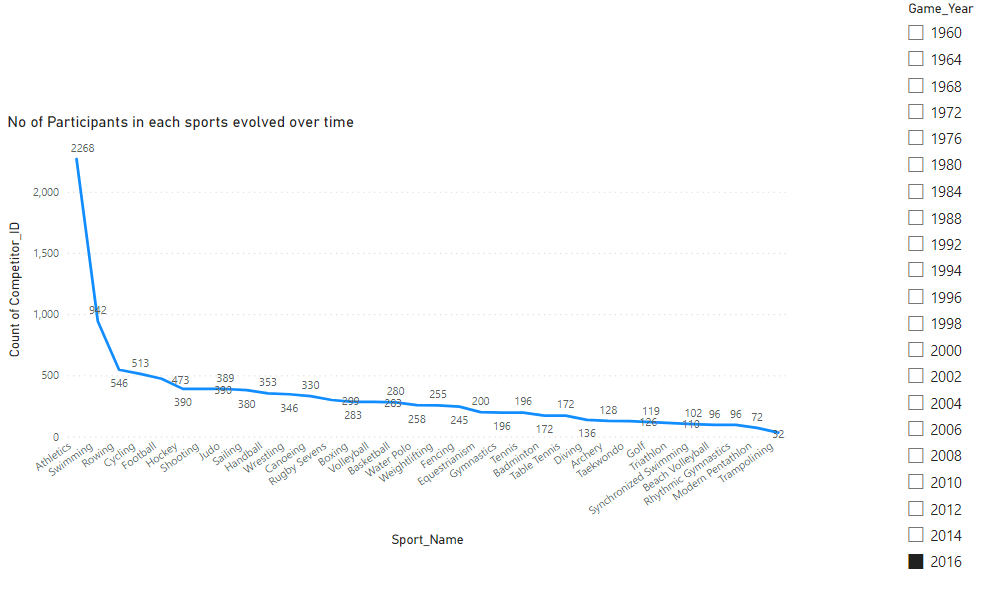
5. Which sports have the highest number of events in the Olympics?

The analysis of sports with the highest number of events in the Olympics indicates a focus on disciplines with diverse skill sets and global popularity From the below visualization we can see that Athletics, with its wide range of track and field events, holds a significant position and this analysis shows the popularity of certain sports and also help the sports whit low popularity by including various events to the particular sports.

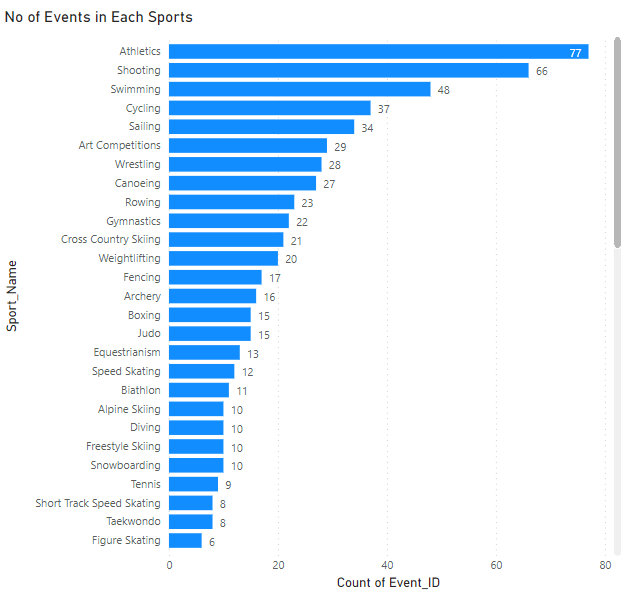


6. How has the participation in each sport evolved over time?

The analysis of the evolution of participation in each sport over time reveals shifting trends in athlete interest and global sporting preferences. Certain sports may experience fluctuations in popularity based on certain factors. From the below visualization we can track the historical data, we can identify patterns of growth or decline in athlete engagement, which can inform decisions about resource allocation, event scheduling, and sport development initiatives.

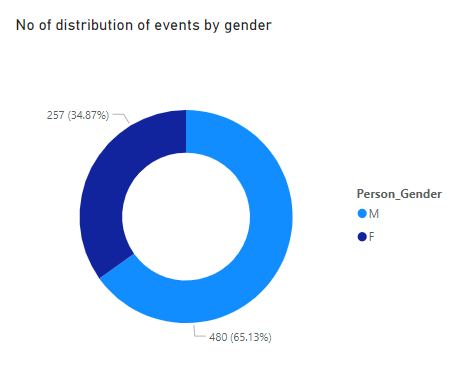
7. How many events are there in each sport?

The analysis of the number of events in each sport shows the popularity and value of the sports. From the below visualization we can examining the event distribution, organizers can ensure equitable opportunities for participation and achievement across a wide range of sports by increasing the number of events in a sports to increase the participants and to improve the opportunities and popularity of the particular sports.

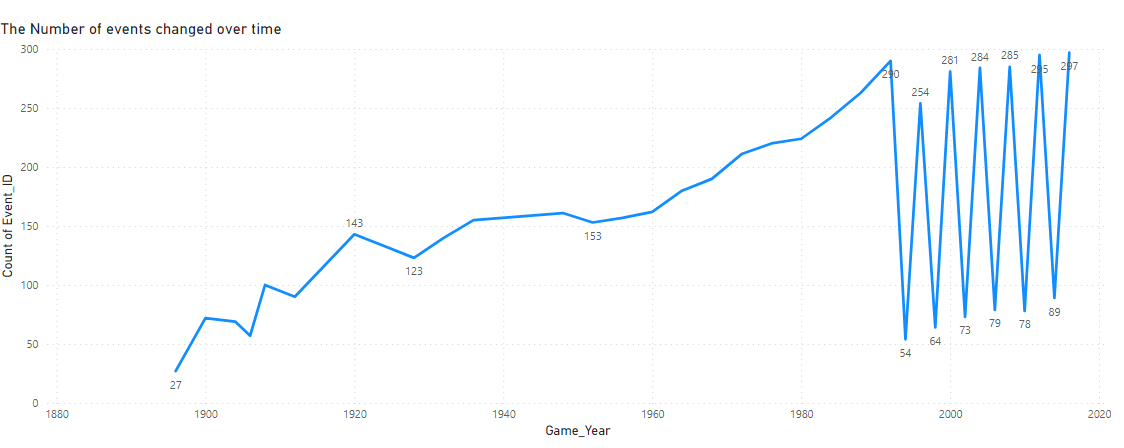


8. What is the distribution of events by gender (Men, Women, Mixed)?

The analysis of event distribution by gender (Men, Women, Mixed) underscores the progress towards gender equality and inclusivity in the Olympic Games. This analysis showcases the Games' commitment to promoting female participation and recognizing mixed-gender collaborations, fostering a more diverse and representative sporting environment. From the below visualization we can the track this distribution, organizers can ensure a fair platform for both male and female athletes, contributing to a more inclusive and empowering Olympic experience.

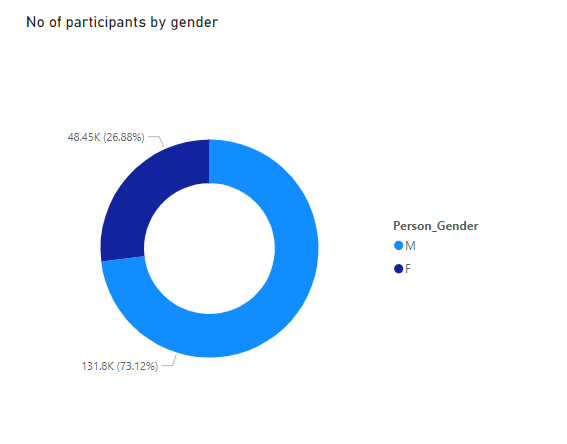
9. How has the number of events changed over time?

By this Analysis we can study the evolution, patterns of growth or contraction in specific sports by seeing thin growth rate of events. From the below visualization we can see that the number of events is in the growth direction in both summer and winter editions.



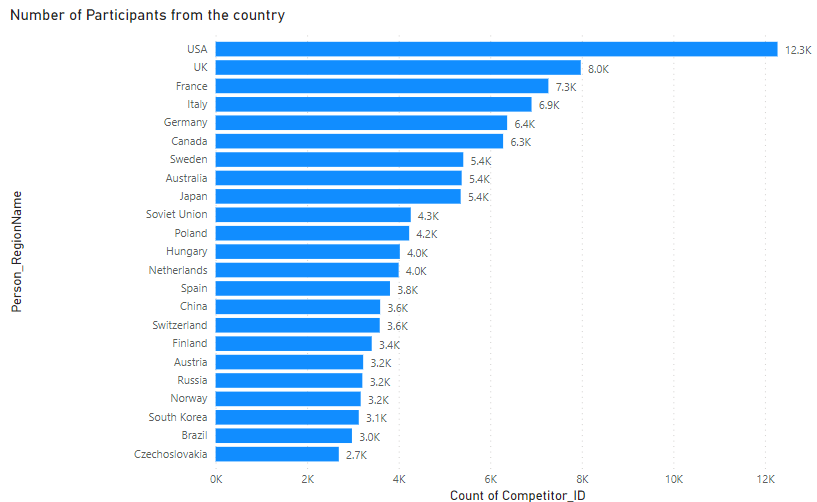
10. What is the distribution of participants by gender?

The analysis of participant distribution by gender in the Olympics shows the progress towards gender equality in sports. It showcases a historical shift from male-dominated participation to a more balanced representation of male and female athletes and This analysis reflects the increasing recognition of women's contributions to sports. From the below visualization we can track the gender distribution, So that organizers can ensure equal opportunities for both male and female athletes are available.



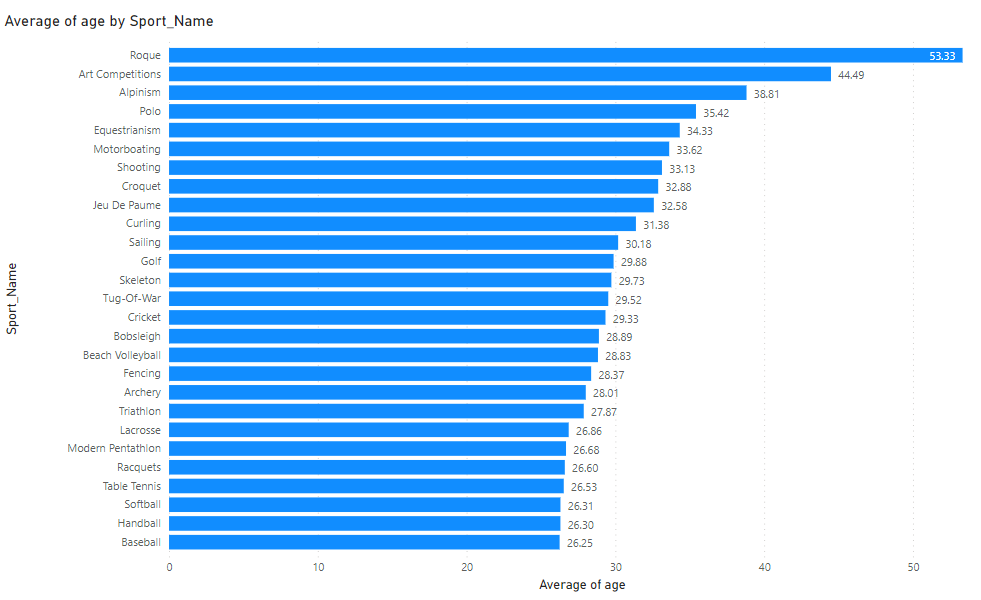
11. Which countries have the highest number of participants in the Olympics?

The Analysis of the higher no of participants from the country shows the power and facility of a individual countries to encourage the participants to participates in the Olympic. By doing this analysis we can give a awareness of the other emerging countries to overtake the dominant countries in the following years. So that the country can show their wealth, ability and their growing phase to the world. From the below visualization we can see the number of participants over the countries and we can suggest accordingly.



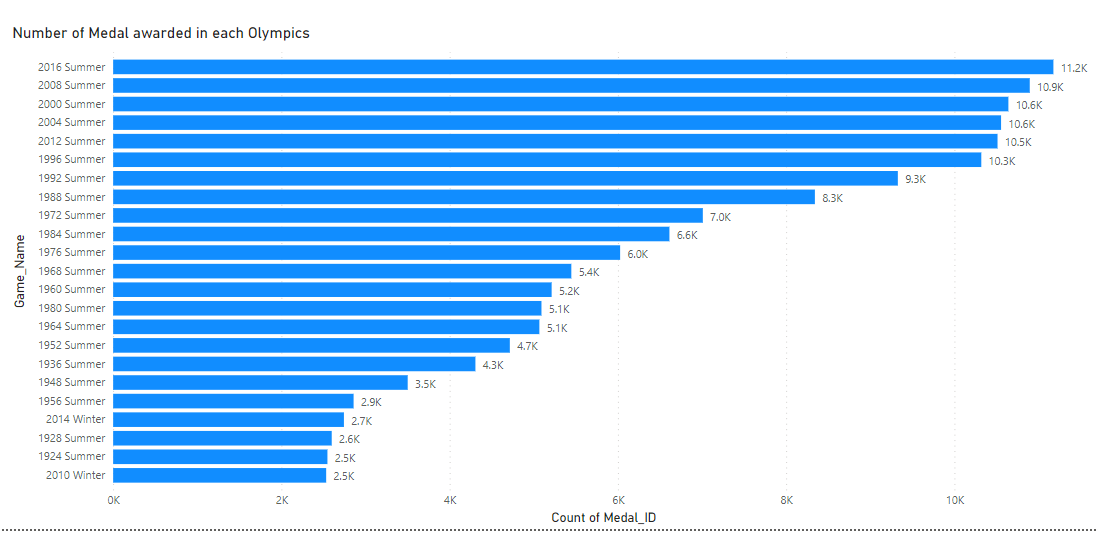
12. How does the age distribution of participants vary across different games?

The analysis of age distribution among participants across different Olympic Games reveals patterns of generational engagement for the sports and This analysis highlights the dynamic balance between experience and youthful Vigor. It aids in understanding sports' appeal to different age groups from the below visualization we can see the distribution of age and we can give a valuable insight to the balance between the all age group of sports.



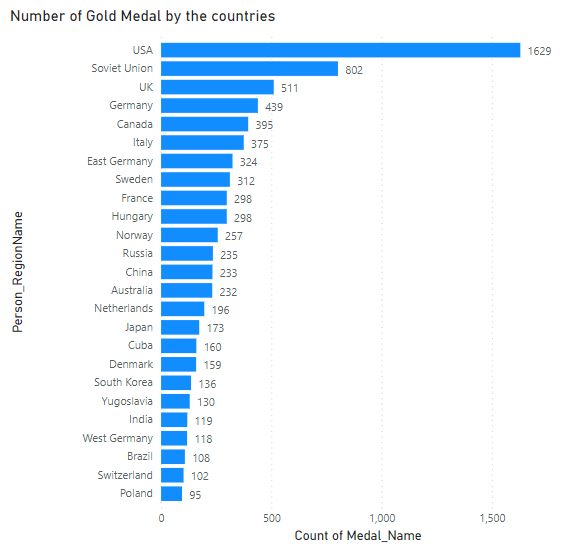
13. How many medals have been awarded in each Olympics?

The analysis of the total number of medals awarded in each Olympic Games provides a historical perspective on the growth and scale of the event. By tracking medal counts across editions, trends in increased participation and diversification of sports become evident. From the below visualization we can understanding medal distribution, organizers can gauge the popularity of sports and allocate resources accordingly.



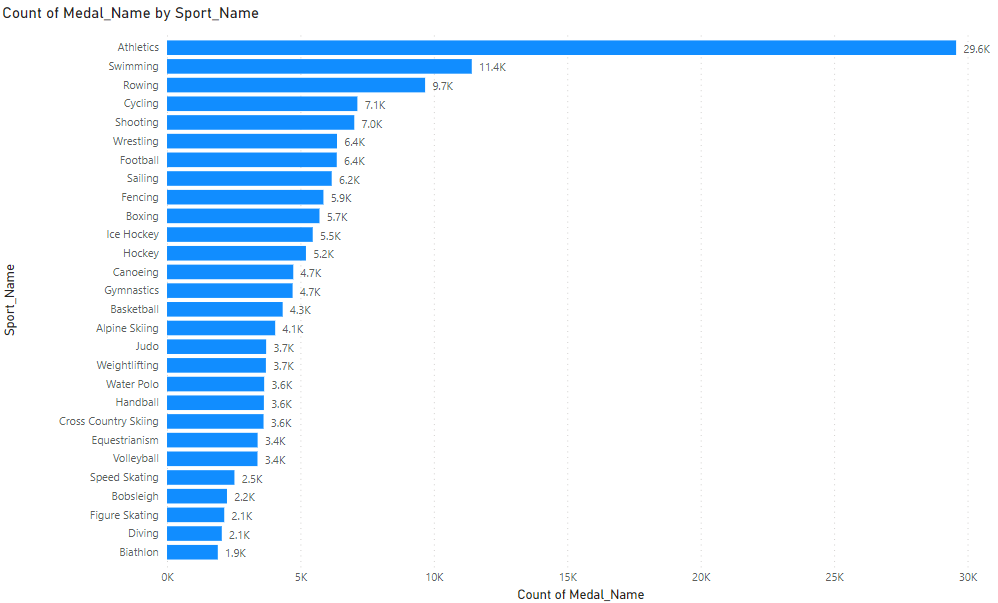
14. Which countries have the highest number of gold medals?

The analysis of countries with the highest number of gold medals reveals the powerhouse nations in the Olympic Games. Countries like the United States, China, and Russia consistently rank at the top due to their robust sports development programs and extensive athlete pools. From the below visualization we can understand the enduring dedication and talent of athletes from these countries, contributing to the prestige and global impact of the Olympic Games.



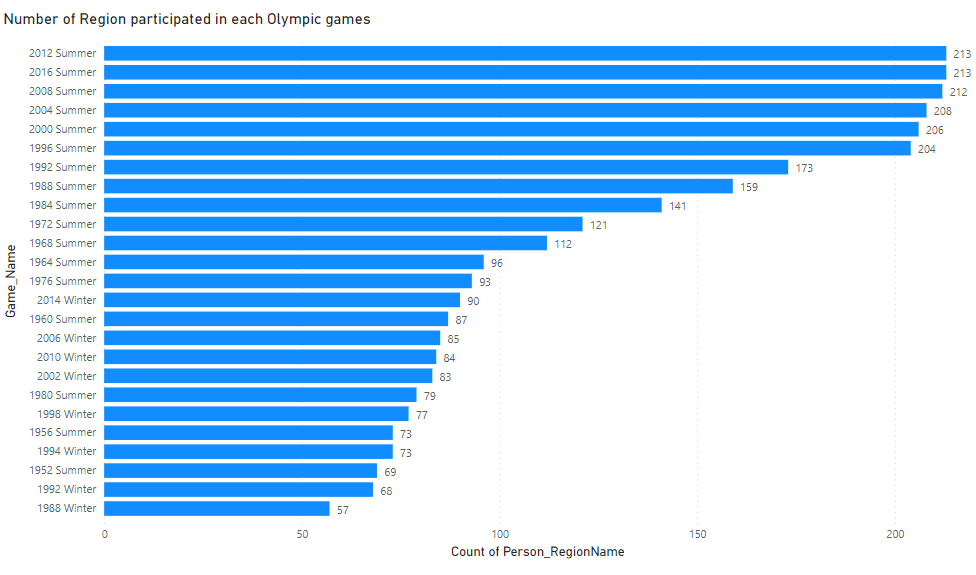
14. How does the medal distribution vary across different sports?

The analysis of medal distribution across different sports reveals varying degrees of competitive dominance and diversity. Certain sports may exhibit concentrated medal distribution among a few dominant nations, reflecting historical strengths and expertise. In contrast, other sports might demonstrate a wider distribution of medals, showcasing a more inclusive and globally competitive field. This analysis offers insights into the level of parity and accessibility within different sports, guiding discussions on sports development, resource allocation, and potential strategies for fostering greater competitive balance. Ultimately, it highlights the dynamic and evolving nature of the Olympic Games, where certain disciplines may witness shifting patterns of dominance and emerging contenders over time.



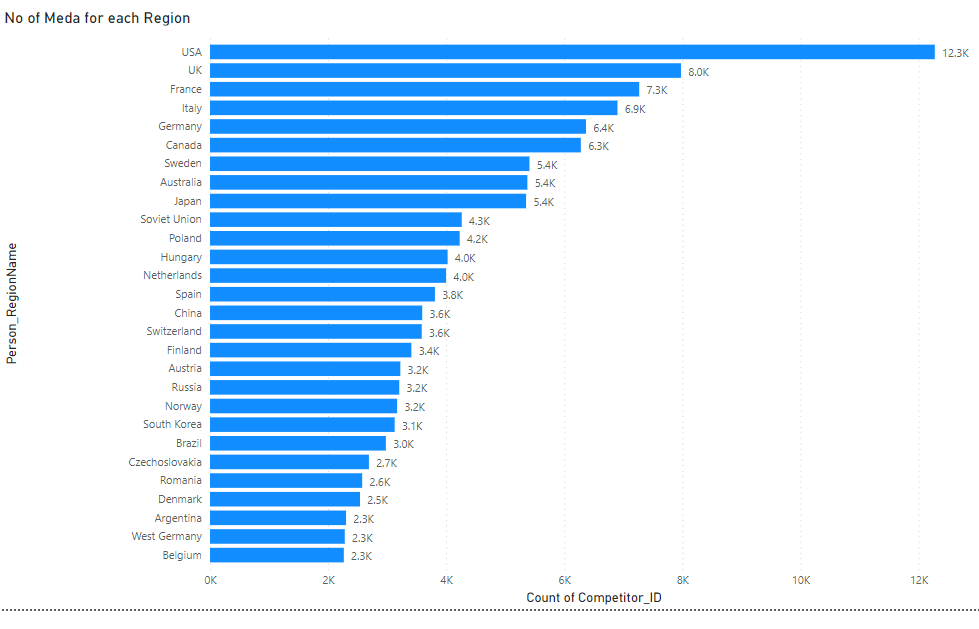
15. How does the medal distribution vary across different sports?

The analysis of medal distribution across different sports highlights the diverse landscape of athletic excellence in the Olympic Games. Some sports may witness concentrated medal accumulation by a handful of dominant nations, reflecting historical strengths and investments. Conversely, other sports might showcase a more even distribution, underscoring a global competitive platform. This analysis provides insights into the evolution of sporting prowess, guiding discussions on resource allocation, athlete development, and potential measures to foster inclusivity. Ultimately, it emphasizes the Games' role in celebrating a wide array of talents while encouraging equitable opportunities for nations of varying sporting backgrounds.



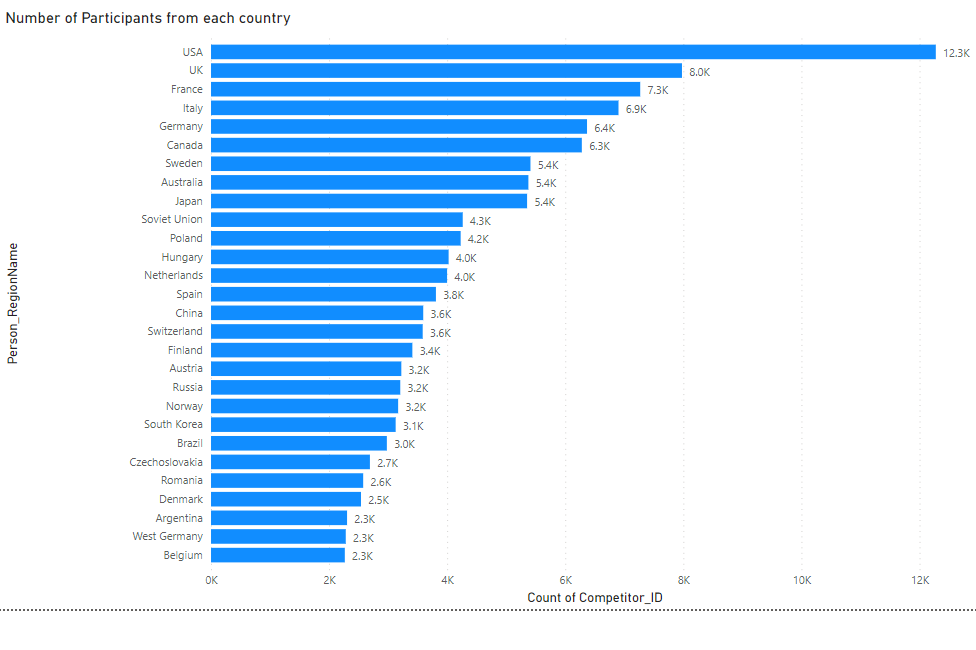
16. How many regions or NOCs participate in each Olympic Games?

The analysis of the number of regions or National Olympic Committees (NOCs) participating in each Olympic Games underscores the event's global reach and inclusivity. Tracking the participation reveals fluctuations in the diversity of represented nations over time. This analysis showcases the expanding international appeal of the Games, fostering connections between different regions and cultures. By understanding participation trends, organizers can evaluate the success of outreach efforts, promote unity, and identify potential areas for increased representation. Ultimately, it emphasizes the Olympics' role in bringing together nations from around the world to celebrate the spirit of sportsmanship and camaraderie.



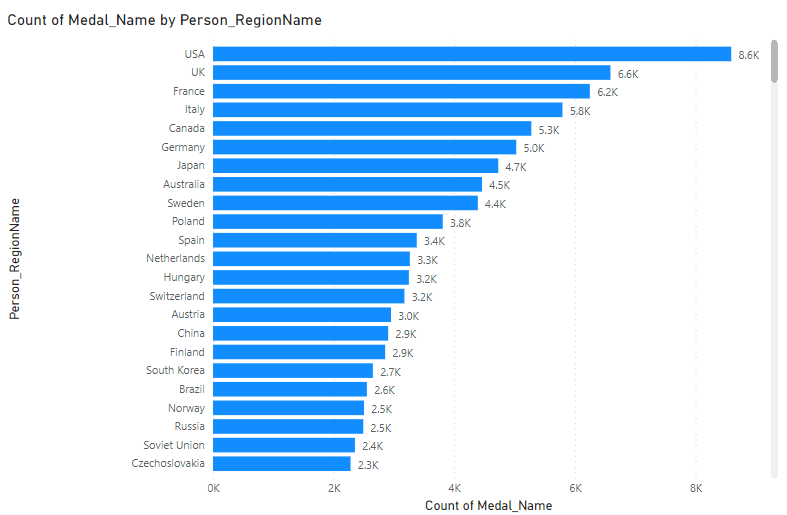
17. Which regions have the highest number of participants in the Olympics?

This analysis underscores the correlation between strong sporting development and higher athlete representation. By identifying regions with significant participation, Olympic organizers can tailor initiatives to further promote sports engagement, diversity, and equal opportunities, enriching the global Olympic experience. From the below Visualization we can see that USA has a highest number of Participants.



18. What is the distribution of medals among different regions?

This Analysis shows the medal distribution, Olympic stakeholders can assess the effectiveness of sports development initiatives, encourage broader participation, and work towards fostering equitable opportunities for regions worldwide. Ultimately, this analysis showcases the Games' ability to celebrate excellence while inspiring diverse regions to thrive on the Olympic stage.



19. How to enhance the Future Olympics with help of use full aspects?

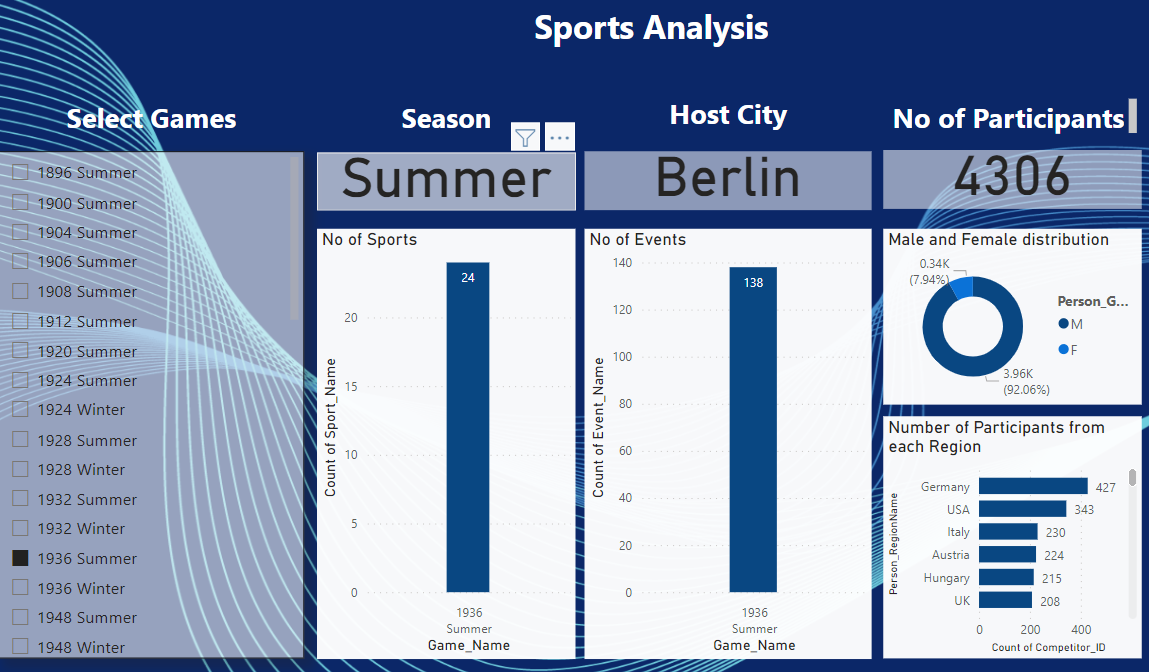
Introduce fresh, captivating sports or innovative adaptations of existing ones to deeply engage athletes and audiences, nurturing inventiveness while preserving the event's integrity. Through insightful sports analysis, evident is the commendable growth in participation rates; although, specific regions exhibit heightened interest, others lag. To address this, introducing captivating events within chosen sports, along with an expanded roster, fostering balanced competition and equitable representation across genders, becomes imperative. Utilizing this data-driven evaluation, a targeted approach can be crafted, focusing on regions witnessing both ascents and declines in Olympic participation, including prominent nations, for strategic promotion. By inspiring and empowering emerging nations, a blueprint arises to amplify their excellence, thus elevating the trajectory of future Olympics.

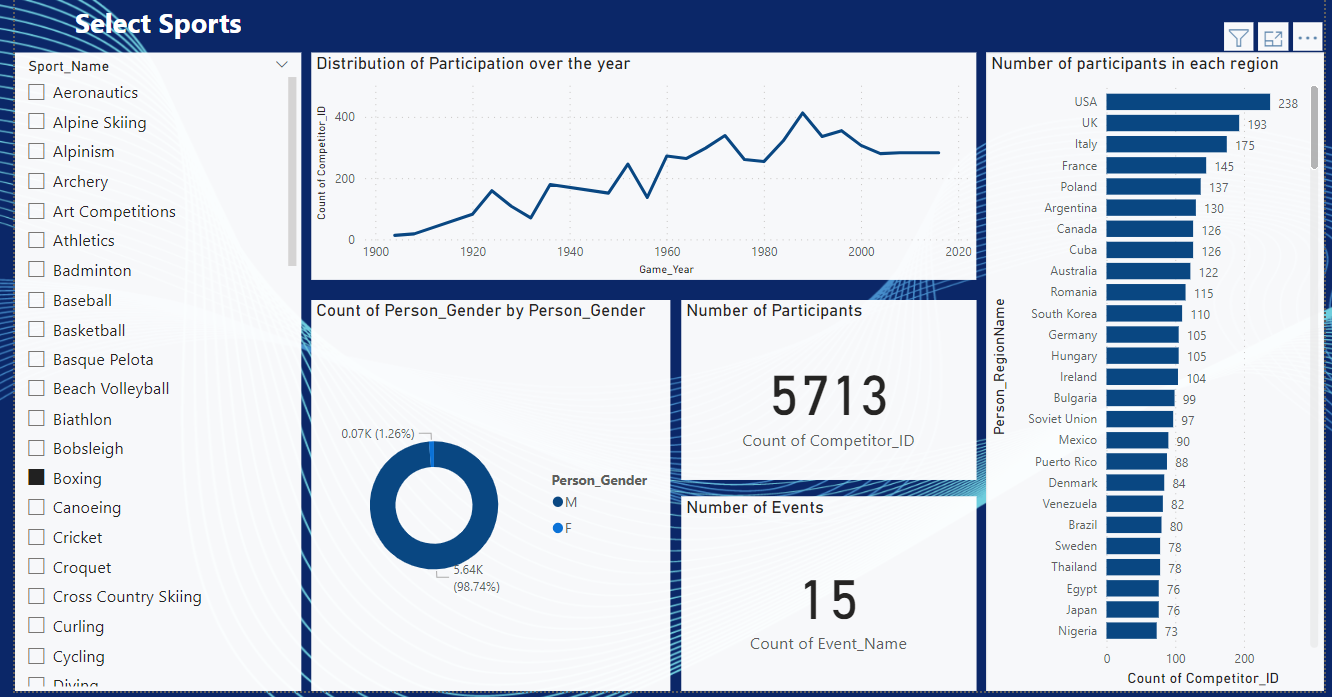
**Power BI Reports:**

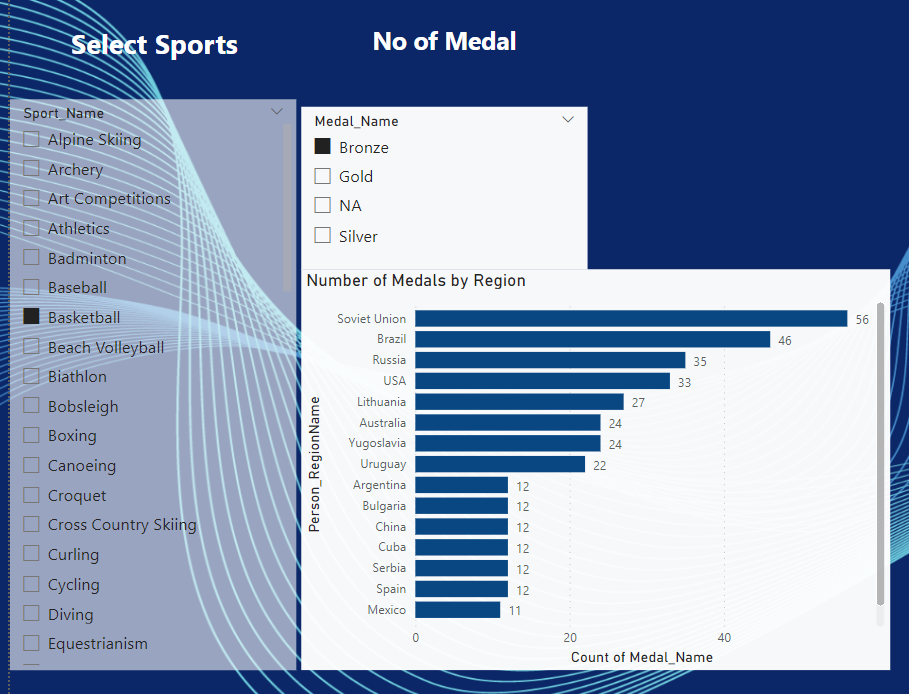
Created a Power BI Dashboard based on the four important factors such as Sports, Participants, Regions and Medals this will breakdown the complexity by dividing a big problems and Analysis to easier way. With the help of this we can find the pros and cons of each factor and we can suggest exact thing which comes under this individual factor to enhance the future Olympics and also, we can understand the current scenarios of the Olympics.

**Power BI Dashboard:**

****

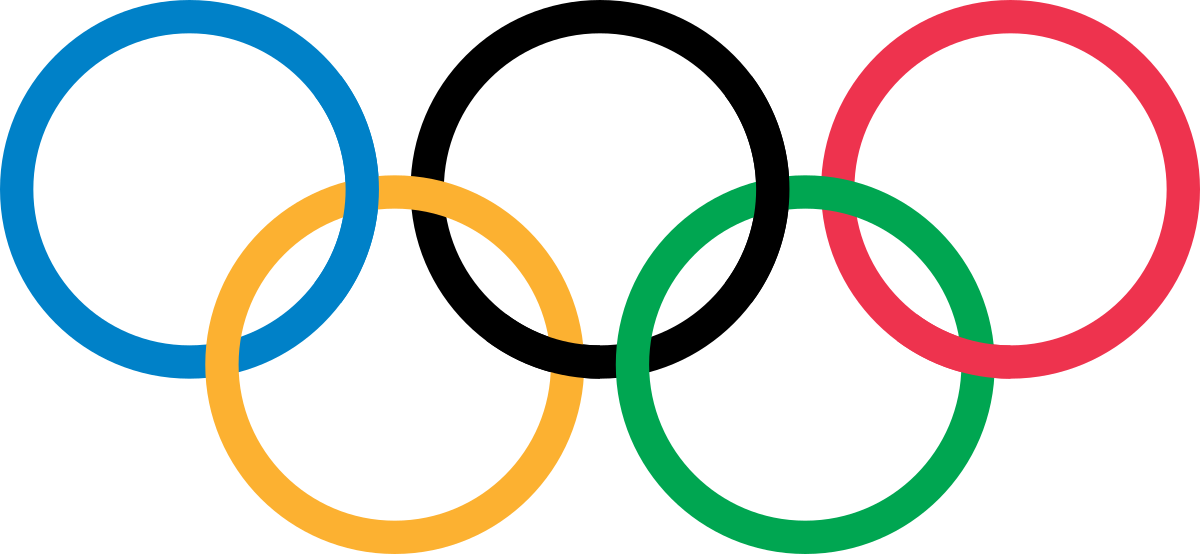
****

****

****

**Outcomes From this analysis:**

* From this Analysis I try to show the exact popularity of individual Sports and the sports which need popularity, attention among the participants and the viewers.
* Identified the distribution of the gender equality to show the ability of all gender in a particular sport. So that in future we can enhance the population of participants in all gender.
* Identified the distribution of Medals over the sports, countries to see the domination and providing a view to the other emerging countries to became them also a dominant in the Olympics in future.
* Identified the distribution of number of participants from each country over the period. So that we can see the countries which need attention to spread the sports in all countries.



Thank You