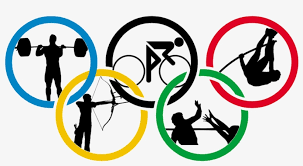
Capstone projects- Sports

analytic



Submitted by

Sanjay Kumar

Overview

The comprehensive dataset on the Olympic Games serves as a valuable resource for a thorough analysis of the event's historical evolution. By examining multiple dimensions such as sports, events, participants, medals, and regional representation, this analysis aims to uncover trends and patterns across various editions. The dataset allows us to explore the diverse range of sports and events that have been part of both Summer and Winter Olympics, highlighting the global reach and inclusivity of the Games. It also provides insights into the host cities and countries, showcasing the international collaboration behind the Olympics. By examining participants' data and their representation across different editions, the analysis sheds light on the evolving nature of sports participation. Furthermore, the dataset allows for a comprehensive understanding of medal distribution, emphasizing excellence and achievement. The insights derived from this analysis will not only enrich the understanding of the Olympic Games' history but also offer recommendations to enhance future editions by optimizing sports selection, regional diversity, and overall organization. The final report and presentation will capture these findings, providing a compelling overview of the Olympics' significance as a global sporting spectacle.

The Process

**1. Data Acquisition from GitHub:**

Obtain the requisite dataset from a designated GitHub repository, containing essential information on Olympics sports,

**2. Data Transformation and Enhancement**:

If necessary, execute data transformation procedures to ensure data quality and consistency. Additionally, consider augmenting the dataset with new problem statements to enrich the analysis potential.

**3. Connecting with Tools:**

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

**4. 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.

**5. 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.

**6. Creation of Visual and Insightful PowerPoint**:

Develop a comprehensive PowerPoint presentation that encapsulates the project's objectives, methodologies, problem statement solutions, and key visualizations. Each problem statement should be accompanied by a dedicated section with pertinent conclusions and insights.

**7. Detailed Documentation:**

Compile a detailed report that meticulously documents the entire project lifecycle. Include sections on data collection, transformation, problem statement formulation, tools integration, Power BI solutions, EDA insights, and PowerPoint visualizations.

Objective

The project aims to conduct a comprehensive analysis of historical Olympic Games data, spanning diverse dimensions such as games, sports, events, participants, medals, and global representation. By exploring this extensive dataset, we seek to uncover trends, patterns, and significant insights that have shaped the evolution of the Olympics over time. Through meticulous examination, we intend to provide a detailed overview of the Games' history, identifying key trends and milestones. This analysis will not only serve to highlight the achievements of athletes and nations but also to offer recommendations for refining future Olympics, including optimal sports selection and enhanced representation of regions. The project's culmination will be a comprehensive report and presentation, showcasing our findings and illuminating the historical and cultural significance of the Olympic Games.

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 sport

**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: Games Competitor**

• Competitor ID: Unique identifier for each Competitor

• Game ID: Foreign Key referencing the Game ID field from the Games Table

• Person ID: Foreign Key referencing the Person ID field from the Person table

• Age: Unique identifier for each Competitor

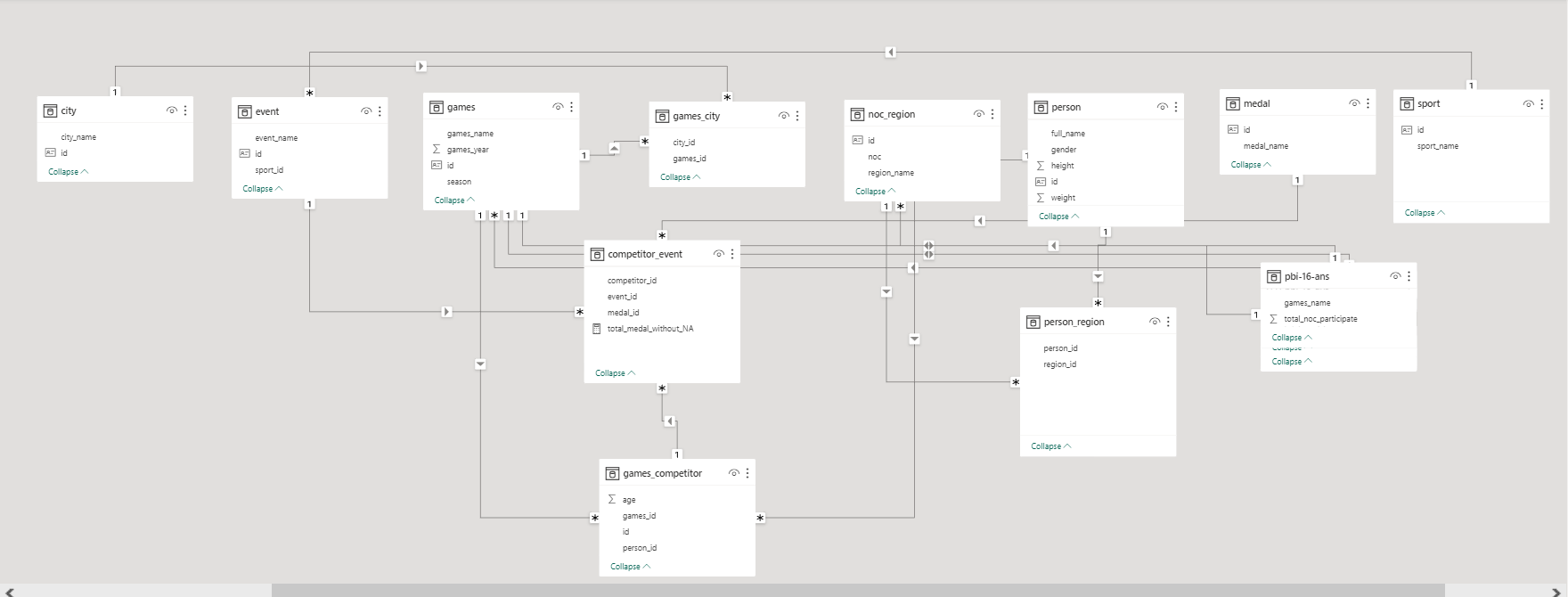
**Table: Games Competitor**

• Competitor ID: foreign key referencing the person id from the person table

• Event id: Foreign key referencing the event id from the event table

• Medal id: Foreign key referencing the id from the medal table

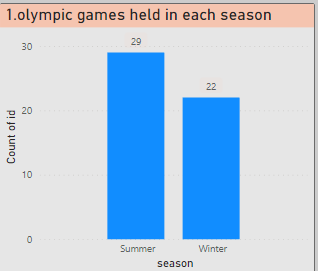
Er Diagram



Power bi problem statements

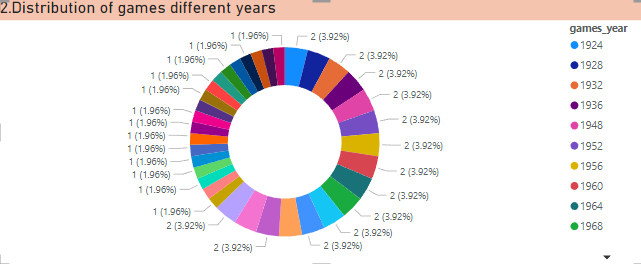
Q.1 How many Olympic Games have been held in each season (Summer vs. Winter)?

Analyzing the frequency of Olympic Games held in each season offers the opportunity to enhance participation across both types of Games, fostering balanced interest in both summer and winter editions. This approach enables a strategic distribution of sporting events across diverse climatic regions, promoting equal attention and engagement throughout the year. By adopting this approach, the Olympic Games can maintain a continuous and heightened level of global attention, as they unfold in both summer and winter seasons within a single calendar year. This visualization underscores the consistent organization of both summer and winter editions since the inception of the latter. Notably, the introduction of the Winter Olympics in 1924 transformed the landscape, filling the gap that existed between 1896 and 1920. As a result, while Summer Olympics have been conducted significantly more times, the implementation of winter editions has led to a more well-rounded and year-round Olympic experience.



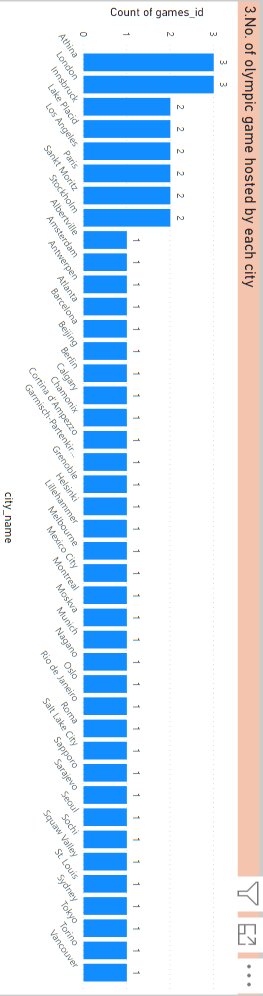
**Q.2 What is the distribution of games across different decades?**

Examining the distribution of Olympic Games across different decades provides valuable insights into the historical evolution of the event. This analysis facilitates the identification of trends, growth patterns, and shifts in participation over time. By categorizing Games into decades, we can discern periods of rapid expansion, consolidation, and innovation, helping us understand how the Olympics have adapted to changing global dynamics. This knowledge aids in evaluating the impact of historical events on the Games, such as world wars or geopolitical shifts, and highlights the overall growth in terms of participant numbers, sports diversity, and global representation. This approach contributes to a comprehensive understanding of the Olympics' journey, aiding in strategic planning for future editions and ensuring the event remains relevant and engaging on a global scale.



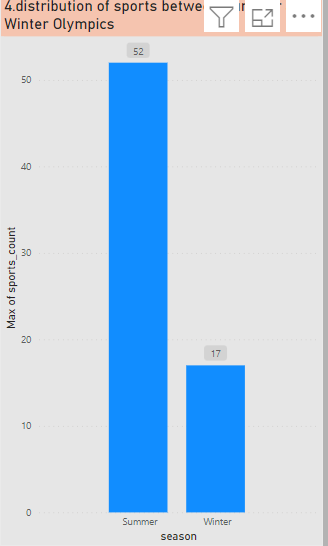
Q.3 Which cities have hosted the most Olympic Games?

Analyzing these recurring host cities offers insights into their capability to organize large-scale events and their significance in the Olympic movement. It helps in identifying patterns of successful infrastructure management, legacy creation, and public engagement. By studying these cities' repeated hosting, we can glean best practices, streamline organizational processes, and enhance event sustainability. This knowledge aids in making informed decisions when selecting future host cities, fostering long-term partnerships, and improving overall event planning and execution. Furthermore, it underscores the importance of balancing tradition with innovation to create memorable and impactful Olympic Games experiences.



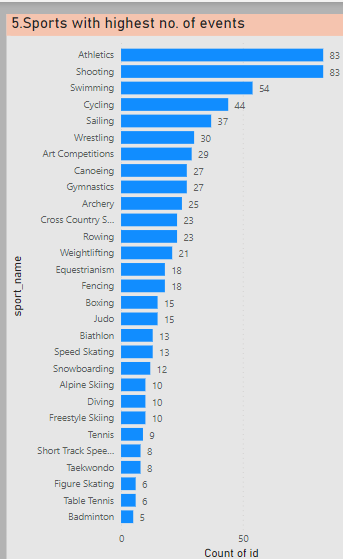
Q.4 What is the distribution of sports between Summer and Winter Olympics?

The distribution of sports between Summer and Winter Olympics reveals a contrasting array of athletic disciplines. Analyzing this distribution informs future Games' planning by showcasing popular and diverse sports. It aids in optimizing the sports portfolio, ensuring a balanced representation of both traditional and contemporary events. Learning from historical distributions helps identify gaps and potential areas for expansion, enhancing inclusivity and global appeal. By understanding preferences and trends, future Olympics can offer a well-rounded experience for participants and spectators, while also accommodating regional interests. This analysis assists in making informed decisions about sport selection, fostering innovation, and maintaining the Games' relevance to evolving athletic interests.



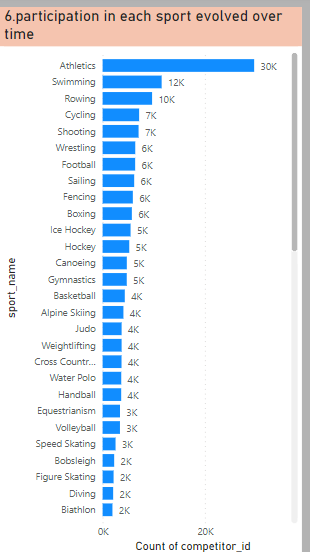
Q.5 Which sports have the highest number of events in the Olympics?

Sports like Athletics, Swimming, and shooting typically have the highest number of events in the Olympics. This insight holds significance for future Games by highlighting sports with broad appeal and opportunities for diverse athlete participation. Understanding the popularity of these sports can guide decisions on resource allocation, scheduling, and venue planning. By emphasizing events with high participation and viewership, future Olympics can ensure sustained interest and engagement. Moreover, this knowledge assists in strategic marketing and sponsorships, enhancing revenue generation. Additionally, recognizing which sports drive fan enthusiasm helps in creating an engaging atmosphere and promoting global sportsmanship. While maintaining traditional events, it's also essential to consider emerging sports to cater to evolving interests and demographics, fostering a dynamic and inclusive Olympic experience.



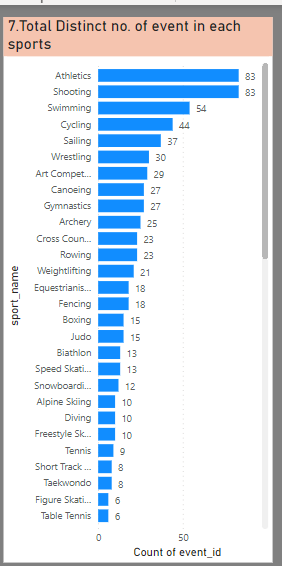
Q.6 How has the participation in each sport evolved over time?

Studying the evolution of participation in each sport over time provides invaluable insights for organizing future Olympic Games. This analysis reveals shifting trends in athlete interest and global representation, guiding decisions on sports selection and event scheduling. By understanding historical patterns, organizers can anticipate changes in popularity and adapt the sports lineup accordingly to maintain viewer engagement. This knowledge aids in allocating resources effectively, optimizing venues, and enhancing the overall spectator experience. Furthermore, observing the growth or decline of certain sports can inform investment decisions, ensuring that facilities and infrastructure align with athletes' and audiences' preferences. Overall, this analysis facilitates strategic planning, helping organizers create a dynamic and relevant Olympic program that resonates with both athletes and a diverse global audience.



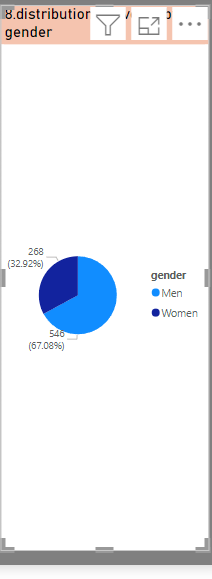
Q.7 **How many events are there in each sport?**

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.



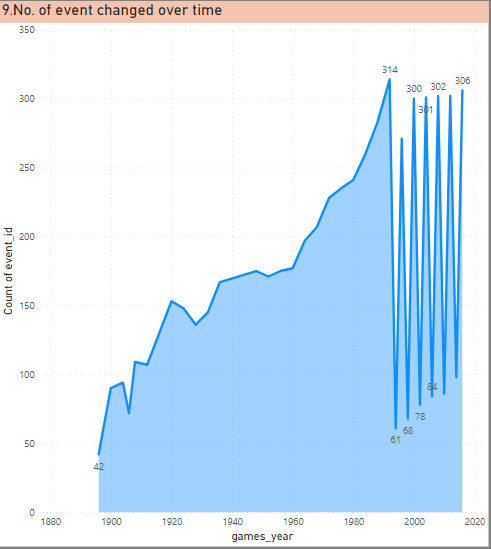
Q.8 What is the distribution of events by gender (Men, Women, Mixed)

The distribution of events by gender (Men, Women, Mixed) illustrates the progress towards gender equity in the Olympics and informs future Games planning. Analyzing this distribution helps in ensuring a fair representation of opportunities for all genders. By identifying imbalances, organizers can strive for a more equal division of events, fostering inclusivity and reflecting changing societal norms. This insight also aids in promoting female athletes' visibility and participation, inspiring greater involvement and fostering role models. Moreover, understanding mixed-gender events' success encourages innovation, promoting collaborative competition. By maintaining a gender-diverse and balanced event portfolio, future Olympics can uphold their commitment to inclusivity, amplify social impact, and cater to a wide array of athletic talents, enriching the overall Games experience.



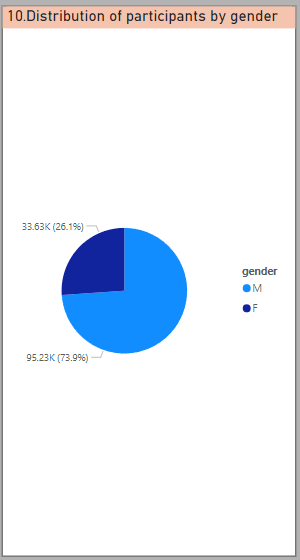
Q.9 How has the number of events changed over time?

The historical analysis of the changing number of events in the Olympics offers valuable insights for future Games' planning and execution. This examination tracks the growth and diversification of events, helping organizers anticipate trends and allocate resources effectively. By understanding the evolution of events, they can adapt the program to accommodate emerging sports and evolving athlete interests, ensuring continued global engagement. This knowledge informs venue selection, scheduling, and logistical arrangements, aligning with viewers' preferences. Additionally, recognizing the historical fluctuations in event numbers assists in creating a balanced and captivating program that appeals to diverse audiences, fostering long-term sustainability and excitement. By observing the past, future Olympics can strike a harmonious balance between tradition and innovation, providing a dynamic and relevant showcase of athletic excellence.

****

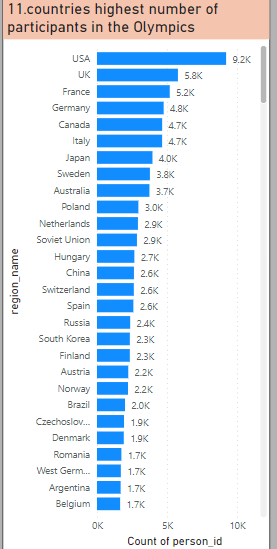
Q.10 What is the distribution of participants by gender?

Understanding the distribution of participants by gender in the Olympics is crucial for fostering gender equality and enhancing the event's overall impact. Analyzing this distribution sheds light on progress towards inclusivity and can identify disparities that need attention. By ensuring a balanced representation of male and female athletes, the Olympics can serve as a platform for promoting diversity and empowering women in sports. This insight also aids in designing equitable qualification processes, encouraging nations to prioritize female athlete development. Addressing gender imbalances fosters greater global interest and engagement, enhancing sponsorships, media coverage, and revenue generation. Ultimately, this analysis guides strategic decision-making, propelling the Olympics towards a more inclusive future, where all athletes, regardless of gender, can compete on an equal and celebrated stage**.**

****

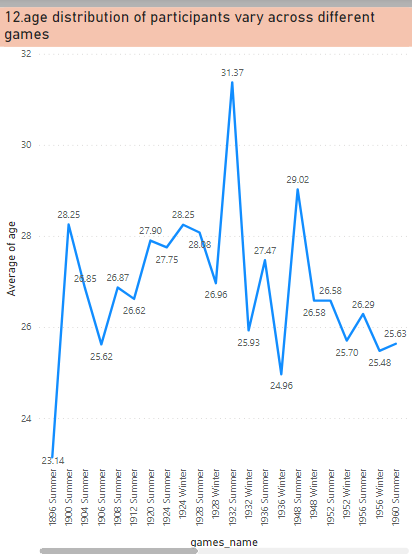
Q.11 Which countries have the highest number of participants in the Olympics?

Identifying countries with the highest number of participants in the Olympics offers valuable insights for both past and upcoming Games. This knowledge highlights nations that prioritize sports and showcases their commitment to athletic excellence. For future Games, understanding these prominent countries can inform collaboration and support strategies, fostering strong partnerships and enhancing global participation. Analyzing this data aids in resource allocation, ensuring facilities and services accommodate larger delegations. For smaller nations, observing the participation of more successful countries can provide inspiration and benchmarks for athlete development programs. Additionally, recognizing the dynamics of participation can influence marketing and broadcasting decisions, catering to the interests of a diverse global audience. Overall, this analysis contributes to the creation of an engaging and impactful Olympics, celebrating a wide array of cultures, talents, and athletic achievements.



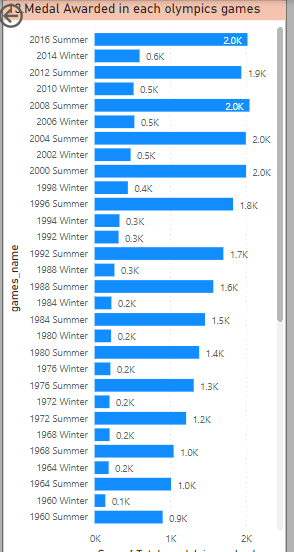
Q.12 How does the age distribution of participants vary across different games?

The age distribution of participants varies across different Olympic Games, reflecting trends in athlete development and societal factors. Analyzing this distribution provides insights for both past and upcoming Games. Studying age patterns in past Games helps in understanding shifts in athlete demographics, training methodologies, and generational interests. This knowledge aids in tailoring athlete support programs, enhancing performance optimization strategies, and ensuring fair representation for athletes of all ages. For future Games, this analysis guides resource allocation, such as athlete accommodations and medical support. It also informs event scheduling, accounting for age-related performance variations. Recognizing age trends contributes to the Games' inclusivity, as different age groups bring diverse strengths and narratives. Ultimately, this insight enriches the Olympic experience, fostering a dynamic and competitive environment that resonates with athletes and spectators alike, honouring both tradition and progress.



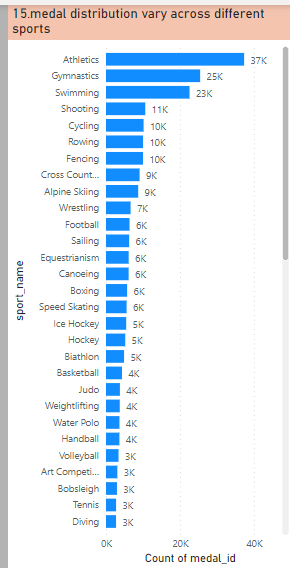
Q.13 How many medals have been awarded in each Olympics?

Understanding the number of medals awarded in each Olympics is vital for both past assessment and future planning of the Games. Analyzing this data offers insights into historical trends in medal distribution, highlighting the growth of sports and nations' competitive strengths over time. For past Olympics, this information aids in evaluating the event's scale and significance, providing context for assessing athletes' achievements. Looking ahead, this analysis guides resource allocation, including medal production, security, and ceremonies planning. By recognizing the evolution of medal counts, organizers can identify potential gaps in diversity or dominance and work to ensure equitable opportunities for all participants. This knowledge also informs marketing and promotion efforts, highlighting medal milestones and narratives that resonate with global audiences. Ultimately, understanding the medal distribution enriches the Games' legacy, fostering a balanced and compelling showcase of human excellence, unity, and sporting achievement.



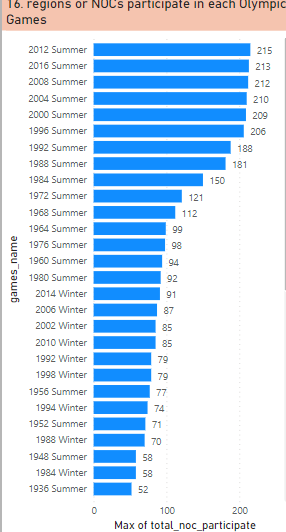
Q.15 How does the medal distribution vary across different sports?

The variation in medal distribution across different sports offers valuable insights for both past Olympic assessment and future Games planning. Analyzing this distribution provides a snapshot of each sport's competitive landscape, highlighting disciplines where certain nations excel and others are emerging. For past Games, this knowledge helps evaluate the popularity and competitiveness of sports, guiding decisions on resource allocation and promotional efforts. Looking ahead, understanding medal distribution guides the inclusion of sports that encourage diverse global participation and success, enriching the Games' narrative. It aids in strategic investment, ensuring support for sports with untapped potential and those that align with evolving audience preferences. This insight also fosters inclusive planning, promoting equity in athlete development, funding, and opportunities. By comprehending the varying medal dynamics, organizers can curate an engaging Olympic program that reflects the evolving nature of sports and celebrates a broad spectrum of talents, cultures, and national achievements.



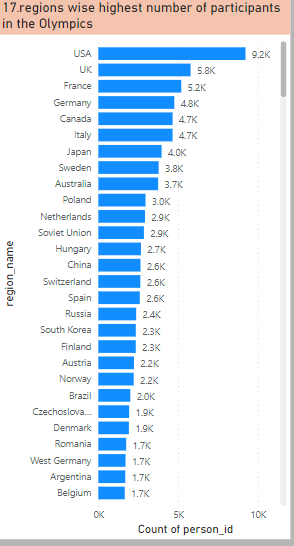
Q.16 How many regions or NOCs participate in each Olympic Games?

The number of regions or National Olympic Committees (NOCs) participating in each Olympic Games can vary from edition to edition. The participation count is influenced by factors such as geopolitical dynamics, the growth of NOCs, and regional interests. In recent Games, the number has exceeded 200 NOCs. This diversity of participation showcases the global nature of the Olympics, promoting unity and friendly competition among nations. The variation in participation underscores the evolving international landscape and the Olympic Movement's ability to foster international camaraderie through sports. Analyzing this data over different Games helps identify trends in NOC growth, regional representation, and the broader reach of the Games' ideals. It also informs future Games' planning, highlighting potential areas for expanding participation and promoting inclusivity in the Olympic experience.



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

Identifying regions with the highest number of participants in the Olympics offers essential insights for both past Games analysis and future event planning. Analyzing regional participation highlights areas of strong athletic development, promoting global diversity and competition. For past Games, this information aids in assessing regional dominance, identifying emerging talents, and acknowledging the broader impact of the Olympics on sporting cultures. Looking ahead, understanding participation trends informs strategies for cultivating athletic talent, fostering partnerships, and tailoring outreach efforts to regions with untapped potential. It also guides logistical planning, accommodating larger delegations and providing appropriate support systems. This insight contributes to fostering inclusivity by promoting engagement from regions that may have historically been underrepresented. By recognizing the regions with significant participation, organizers can curate an Olympic experience that truly represents a worldwide celebration of unity, sportsmanship, and excellence.



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

Understanding the distribution of medals among different regions is crucial for assessing past Olympic Games and shaping future events. Analyzing this distribution provides insights into the competitive strengths of various regions, highlighting dominant and emerging sporting nations. For past Games, this knowledge offers a holistic evaluation of success, helping nations recognize their achievements and areas for improvement. Looking ahead, understanding medal distribution informs resource allocation, athlete development programs, and strategic partnerships to enhance performance. It guides the promotion of underrepresented regions and fosters a balanced representation of global talent. Moreover, recognizing trends aids in fostering fair competition, encouraging equitable opportunities for athletes from all regions. By comprehending the regional medal dynamics, organizers can curate an engaging and dynamic Olympic program that resonates with a diverse global audience, celebrating excellence and unity on an international stage.

