Microsoft | Power BI

Capstone Project – Author Analysis

Publications

Author Analysis

Publishers Analysis

Title Analysis

Sales Analysis

Stores Analysis

Valuable Insights: The Power BI dashboard using the Sample Publication Database provides valuable insights into the publishing company's performance and operations. It presents data-driven analytics on book sales, popular genres, and author contributions. Stakeholders can gain a deeper understanding of customer preferences, bestselling titles, and effective marketing strategies.

Improvement Focus: The dashboard helps the publishing company focus on areas that require improvement. By analyzing sales trends, store distribution, and discount effectiveness, the dashboard identifies opportunities to optimize sales and increase revenue. It enables the company to refine its marketing and distribution strategies to better cater to customer demands.

Evaluation of Effectiveness: The dashboard facilitates the evaluation of various aspects of the publishing business. It assesses the effectiveness of discounts on sales, the impact of different bookstores on revenue, and the success of marketing campaigns. With this evaluation, the company can make data-driven decisions to enhance its overall effectiveness and profitability.

Trend Identification: Using historical sales data and author royalties, the dashboard identifies trends in book sales over time. It helps identify seasonal patterns, bestselling genres in specific periods, and the success of certain authors or book titles. This trend identification empowers the company to plan for future publications and marketing efforts more effectively.

Comprehensive Understanding: The Power BI dashboard offers a comprehensive understanding of the publishing company's operations. By presenting data on book sales, author royalties, and store distribution in one centralized platform, stakeholders can have a holistic view of the business. This comprehensive understanding enables better decision-making and a more cohesive strategy across all aspects of the publishing process.

Objective: The objective of this project is to create a comprehensive Power BI dashboard utilizing the Sample Publication Database. The dashboard aims to provide valuable insights into the publishing company's book sales performance, author royalties, and store distribution, enabling data-driven decision-making and strategic planning.

Analysis Scope: The analysis will focus on various aspects of the publication process, including book sales, author contributions, store performance, and the impact of discounts. It will encompass historical sales data, author royalties based on royalty schedules, and distribution data from multiple bookstores.

Goal: The primary goal of this Power BI dashboard is to offer a holistic view of the publishing company's operations. It will provide actionable insights to optimize book sales, enhance author collaboration, improve store distribution strategies, and identify opportunities for growth and efficiency.

Insights & Recommendations: The Power BI dashboard will generate valuable insights into the top-selling book titles, bestselling genres, and sales trends over time. It will analyze the performance of different bookstores, identify bestselling authors, and calculate author royalties based on royalty schedules. Additionally, it will recommend effective discount strategies to boost book sales and customer engagement.

Report & Presentation: The final deliverable will consist of a detailed report describing the data sources, data modeling methodologies, and data cleansing processes used in creating the Power BI dashboard. The report will also include a step-by-step guide on how to interpret the insights and use the dashboard for decision-making. The presentation will showcase the key findings, visualizations, and actionable recommendations derived from the dashboard's analysis.
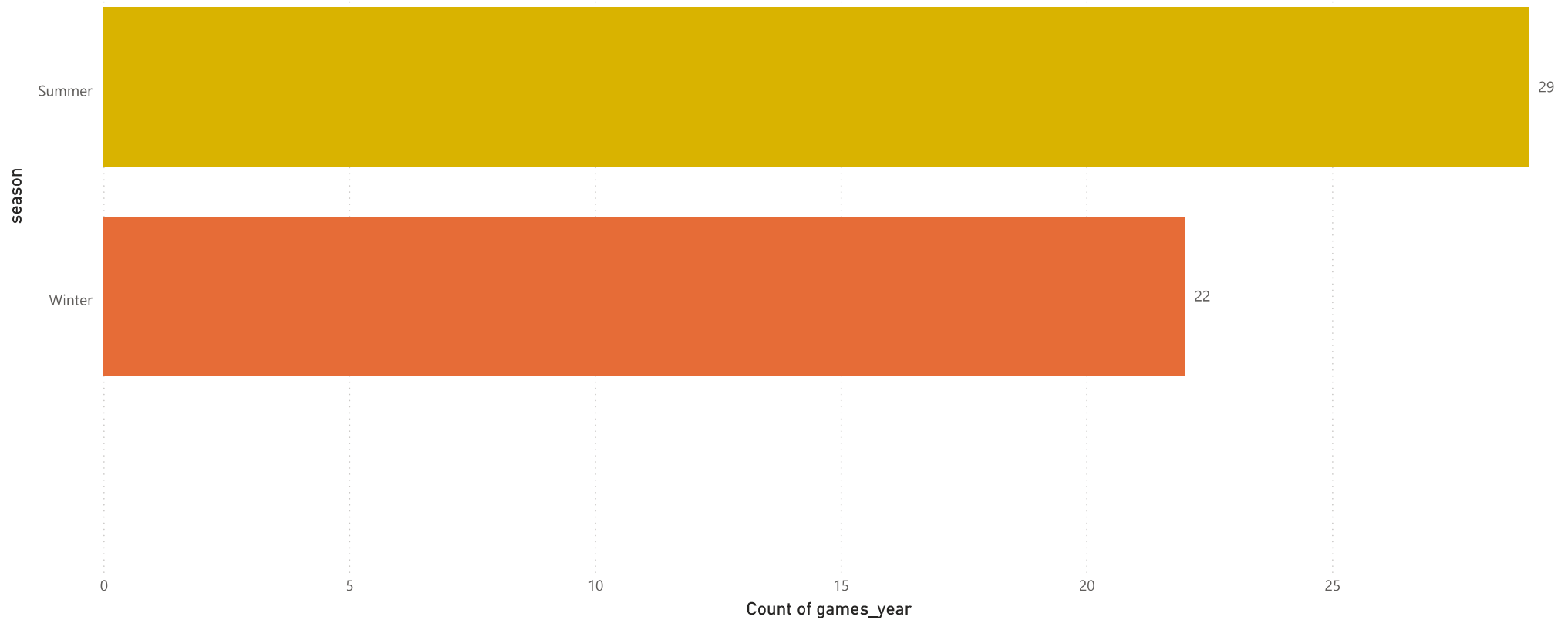
The Power BI dashboard, along with the report and presentation, will serve as a powerful tool for the publishing company's stakeholders. It will empower them to make informed decisions, optimize book sales strategies, foster author collaborations, and enhance the overall business performance in the competitive publishing landscape.
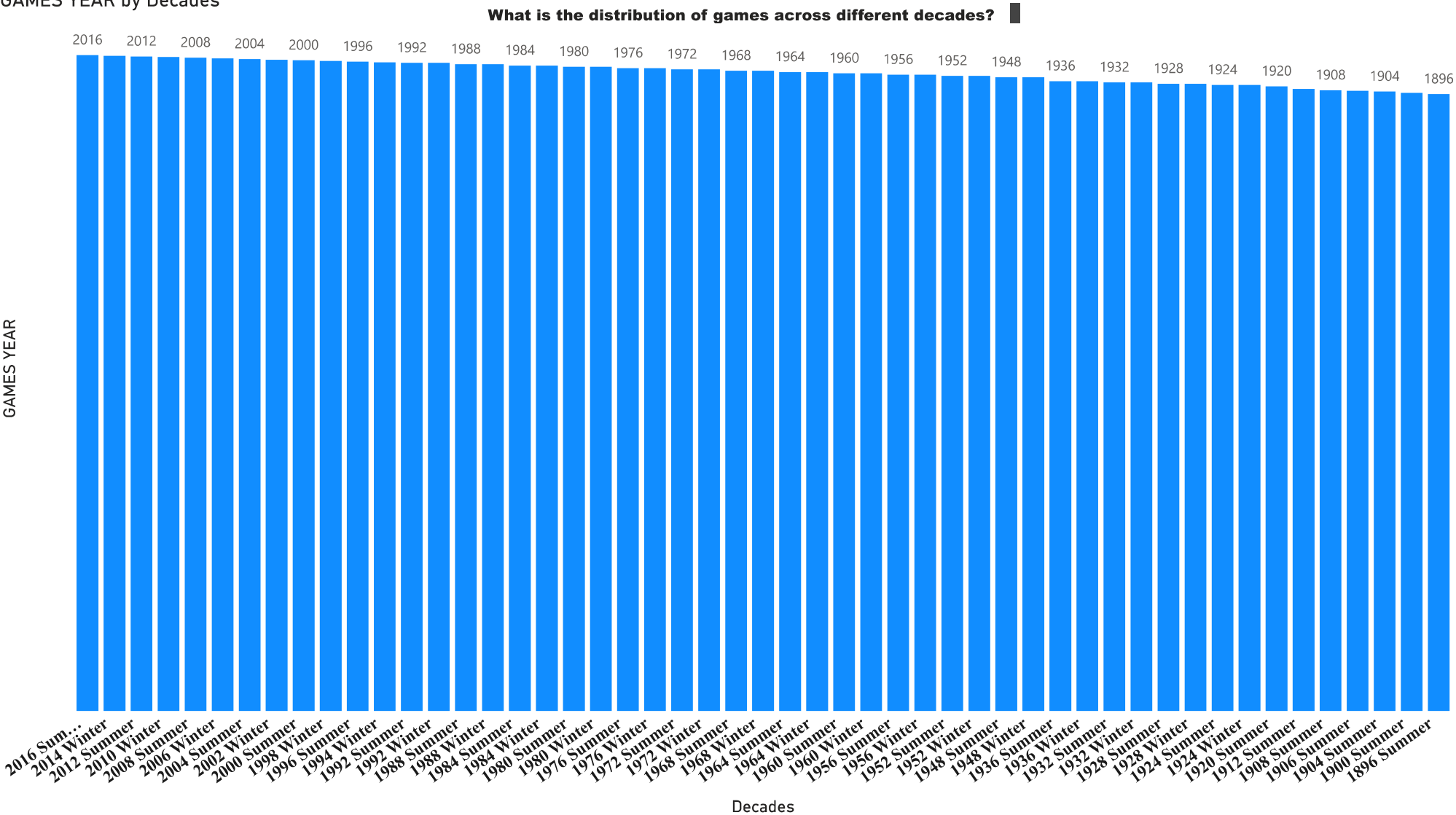
Count of games_year by season

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

season

Summer

29

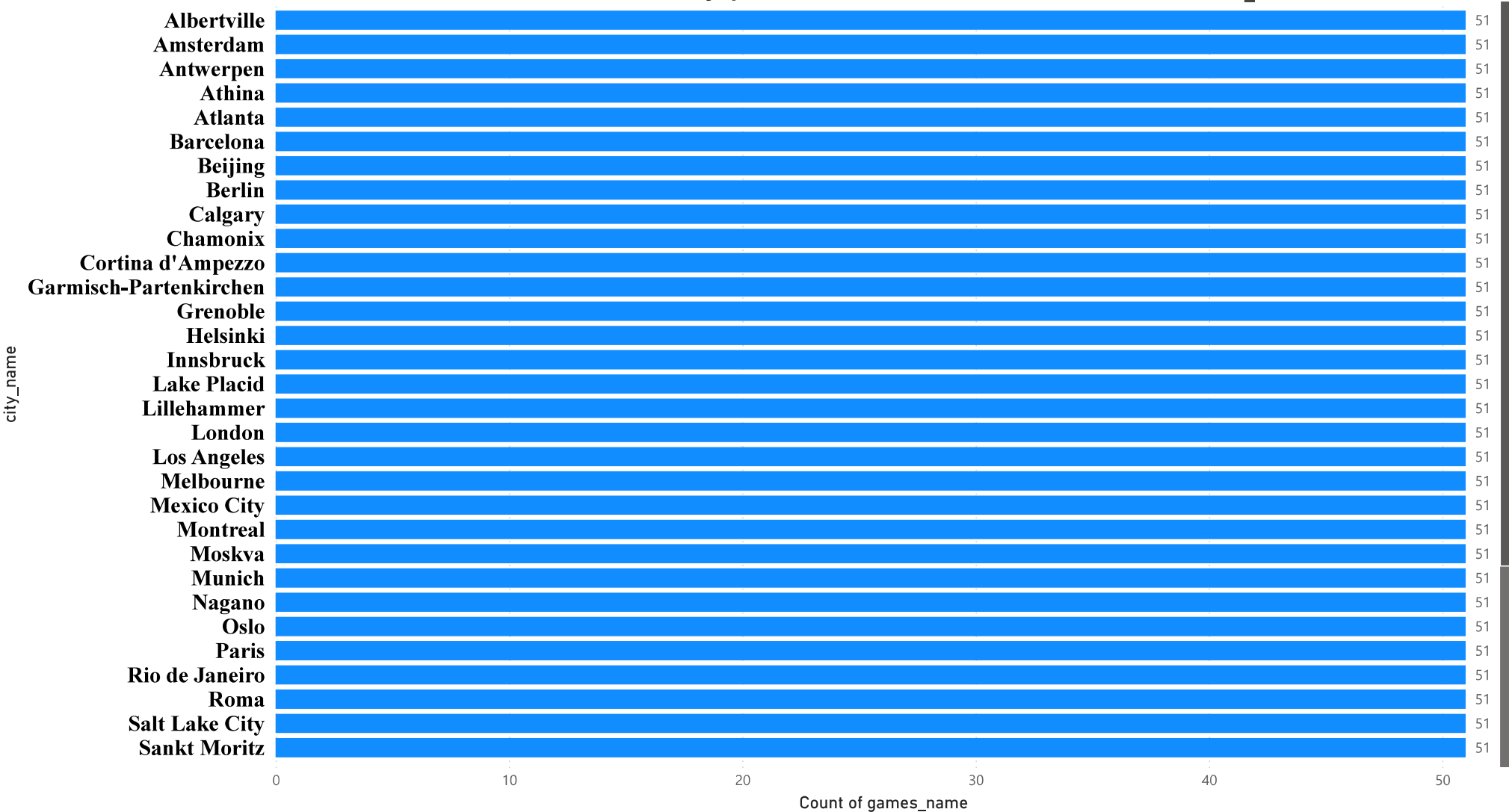Winter

22

0   5   10   15   20   25

Count of games_year

GAMES YEAR by Decades

What is the distribution of games across different decades? ▮

GAMES YEAR

2016   2012   2008   2004   2000   1996   1992   1988   1984   1980   1976   1972   1968   1964   1960   1956   1952   1948   1936   1932   1928   1924   1920   1908   1904   1896

Decades

2016 Sum… 2014 Winter 2012 Summer 2010 Winter 2008 Summer 2006 Winter 2004 Summer 2002 Winter 2000 Summer 1998 Winter 1996 Summer 1994 Winter 1992 Summer 1992 Winter 1988 Summer 1988 Winter 1984 Summer 1984 Winter 1980 Summer 1980 Winter 1976 Summer 1976 Winter 1972 Summer 1972 Winter 1968 Summer 1968 Winter 1964 Summer 1964 Winter 1960 Summer 1960 Winter 1956 Summer 1956 Winter 1952 Summer 1952 Winter 1948 Summer 1948 Winter 1936 Summer 1936 Winter 1932 Summer 1932 Winter 1928 Summer 1928 Winter 1924 Summer 1924 Winter 1920 Summer 1912 Summer 1908 Summer 1906 Summer 1904 Summer 1900 Summer 1896 Summer

Count of games_name by city_name

Which cities have hosted the most Olympic Games?

city_name	Count of games_name
Albertville	51
Amsterdam	51
Antwerpen	51
Athina	51
Atlanta	51
Barcelona	51
Beijing	51
Berlin	51
Calgary	51
Chamonix	51
Cortina d'Ampezzo	51
Garmisch-Partenkirchen	51
Grenoble	51
Helsinki	51
Innsbruck	51
Lake Placid	51
Lillehammer	51
London	51
Los Angeles	51
Melbourne	51
Mexico City	51
Montreal	51
Moskva	51
Munich	51
Nagano	51
Oslo	51
Paris	51
Rio de Janeiro	51
Roma	51
Salt Lake City	51
Sankt Moritz	51

Count of games_name

Count of sport_name by games_name and season

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

season ● Summer ● Winter
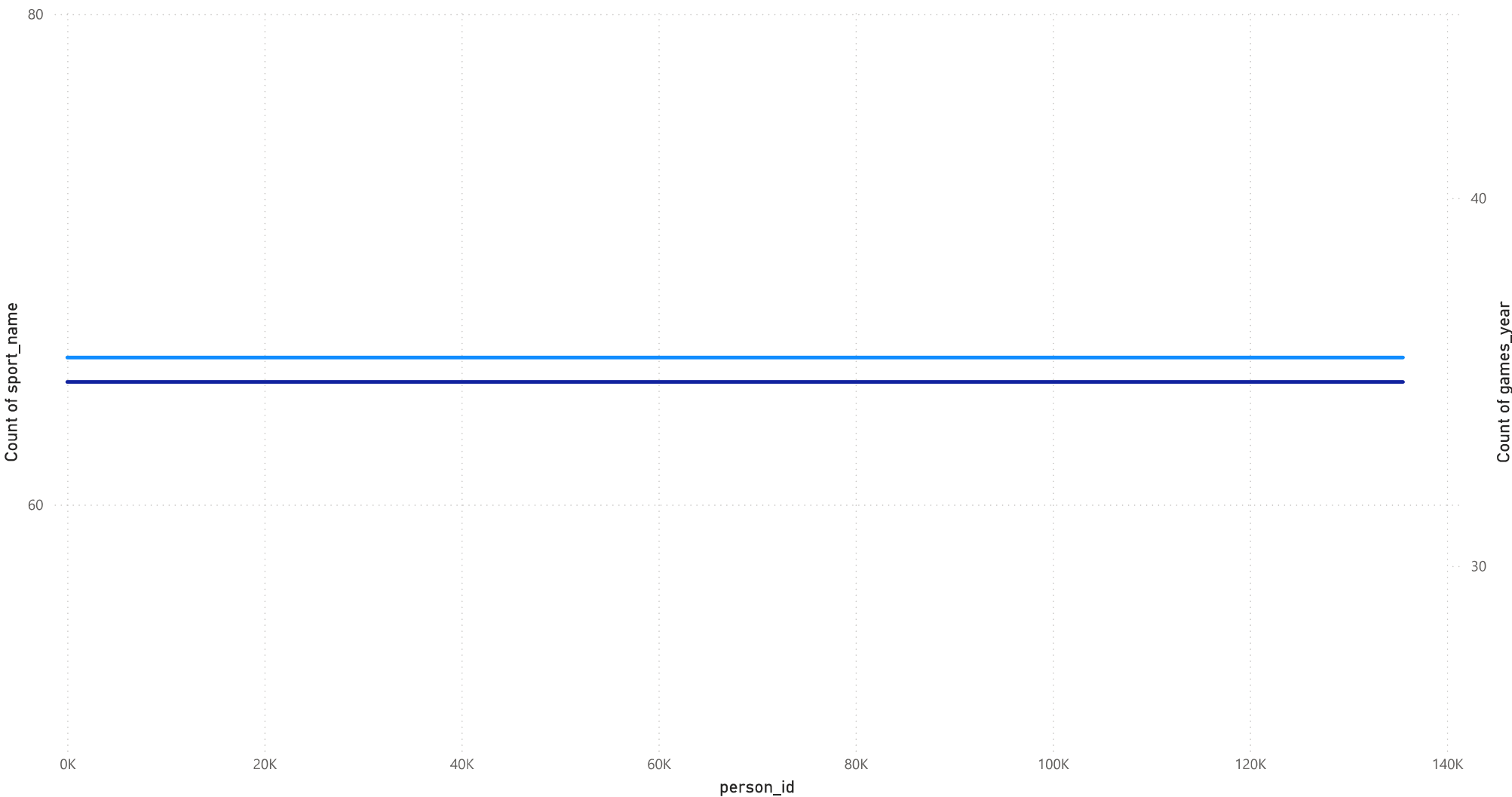

games_name (y-axis, top to bottom):
1896 Summer
1900 Summer
1904 Summer
1906 Summer
1908 Summer
1912 Summer
1920 Summer
1924 Summer
1924 Winter
1928 Summer
1928 Winter
1932 Summer
1932 Winter
1936 Summer
1936 Winter
1948 Summer
1948 Winter
1952 Summer
1952 Winter
1956 Summer
1956 Winter
1960 Summer
1960 Winter
1964 Summer
1964 Winter
1968 Summer
1968 Winter
1972 Summer
1972 Winter

Count of sport_name (x-axis): 0, 10, 20, 30, 40, 50, 60, 70

Count of id by sport_name

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

Count of sport_name and Count of games_year by person_id

● Count of sport_name ● Count of games_year   **How has the participation in each sport evolved over time?**

Sum of competitor_id by sport_name

How many events are there in each sport

sport_name	
Athletics	
Gymnastics	
Swimming	
Shooting	
Rowing	
Cycling	
Fencing	
Cross Country Skiing	
Alpine Skiing	
Wrestling	
Canoeing	
Football	
Equestrianism	
Sailing	
Speed Skating	
Ice Hockey	
Boxing	
Hockey	
Biathlon	
Basketball	
Weightlifting	
Judo	
Water Polo	
Handball	
Volleyball	
Bobsleigh	
Art Competitions	
Tennis	
Diving	
Ski Jumping	
Archery	

0%   20%   40%   60%   80%   100%

Sum of competitor_id

Count of event_name by gender

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

757 (50%) ——————

—————— 757 (50%)
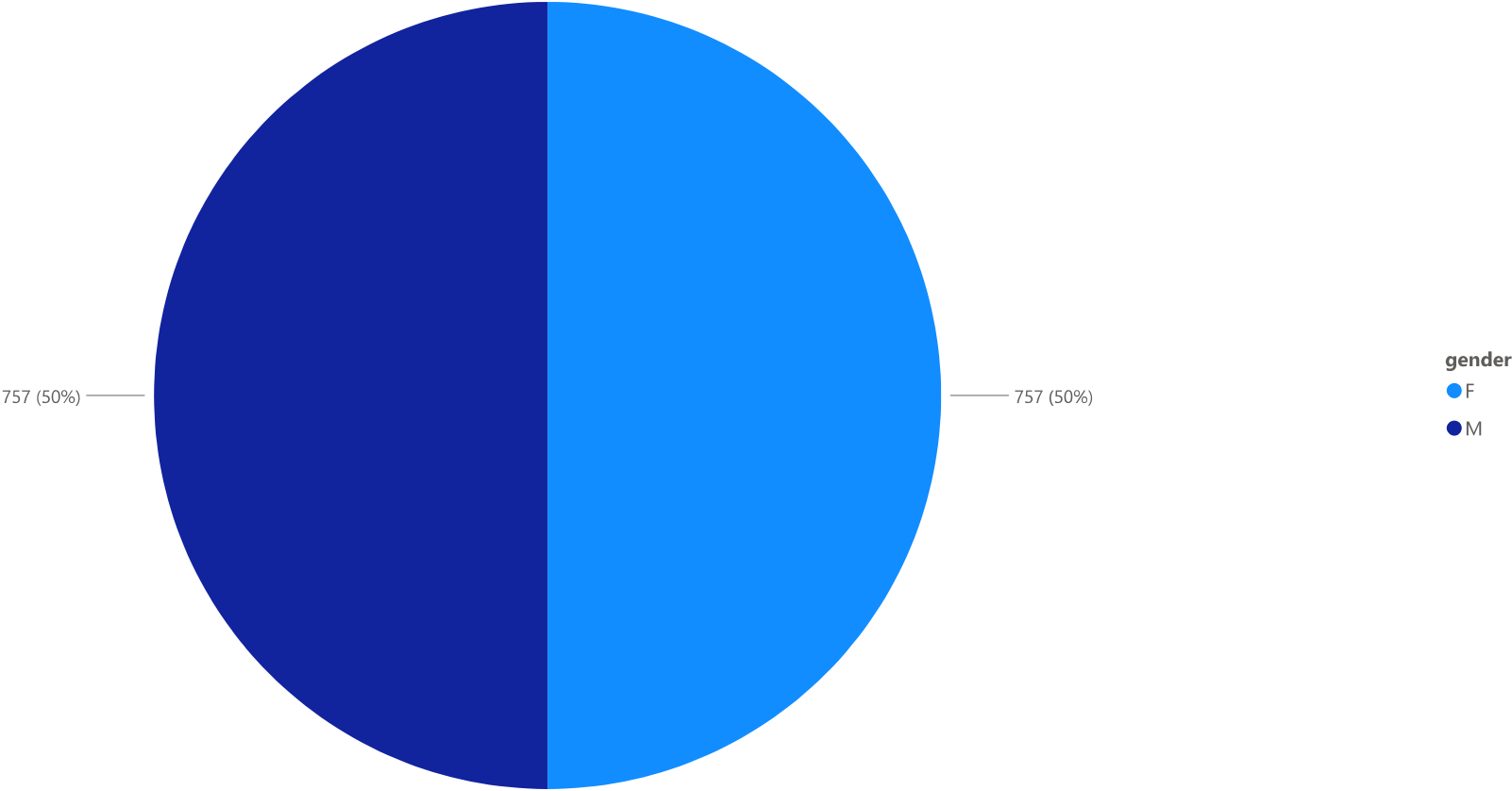
gender
- F
- M

Count of event_id by games_year   **How has the number of events changed over time?**

1,000

Count of event_id

800

600

1880   1900   1920   1940   1960   1980   2000   2020
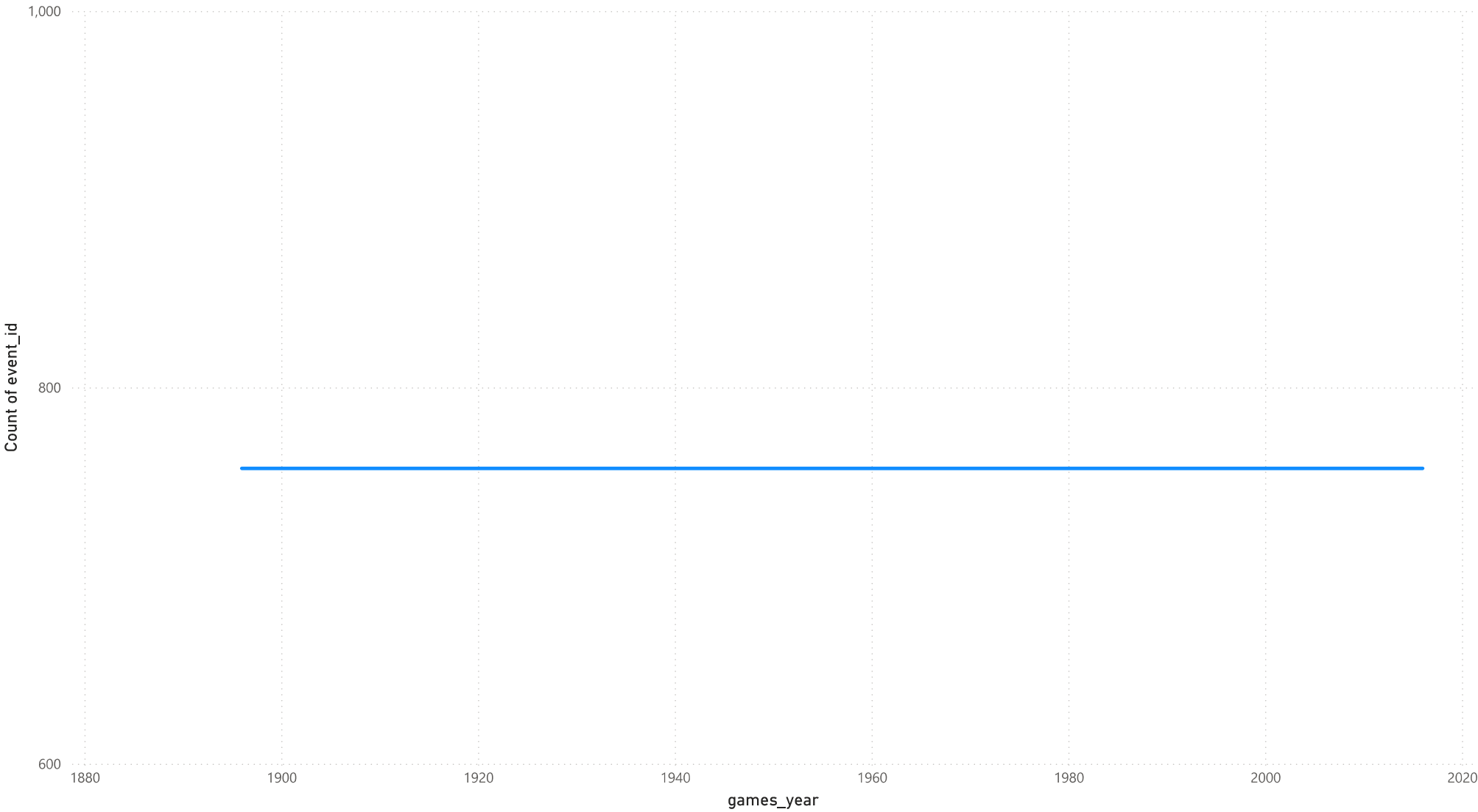
games_year

Count of gender by sport_id

What is the distribution of participants by gender?

sport_id
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
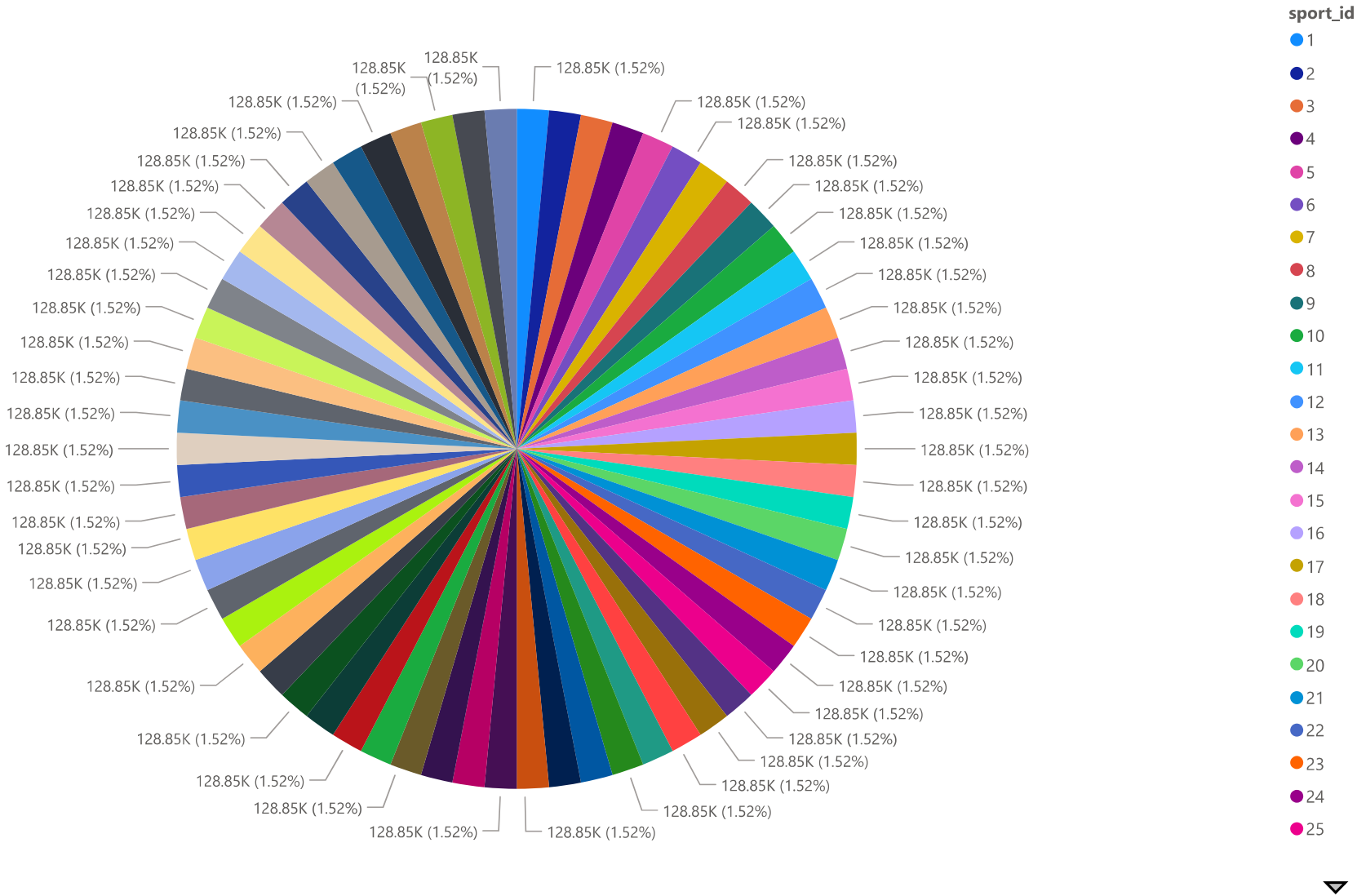
Count of sport_name and Sum of id by id

● Count of sport_name ● Sum of id

1. **Which countries have the highest number of participants in the Olympics?**
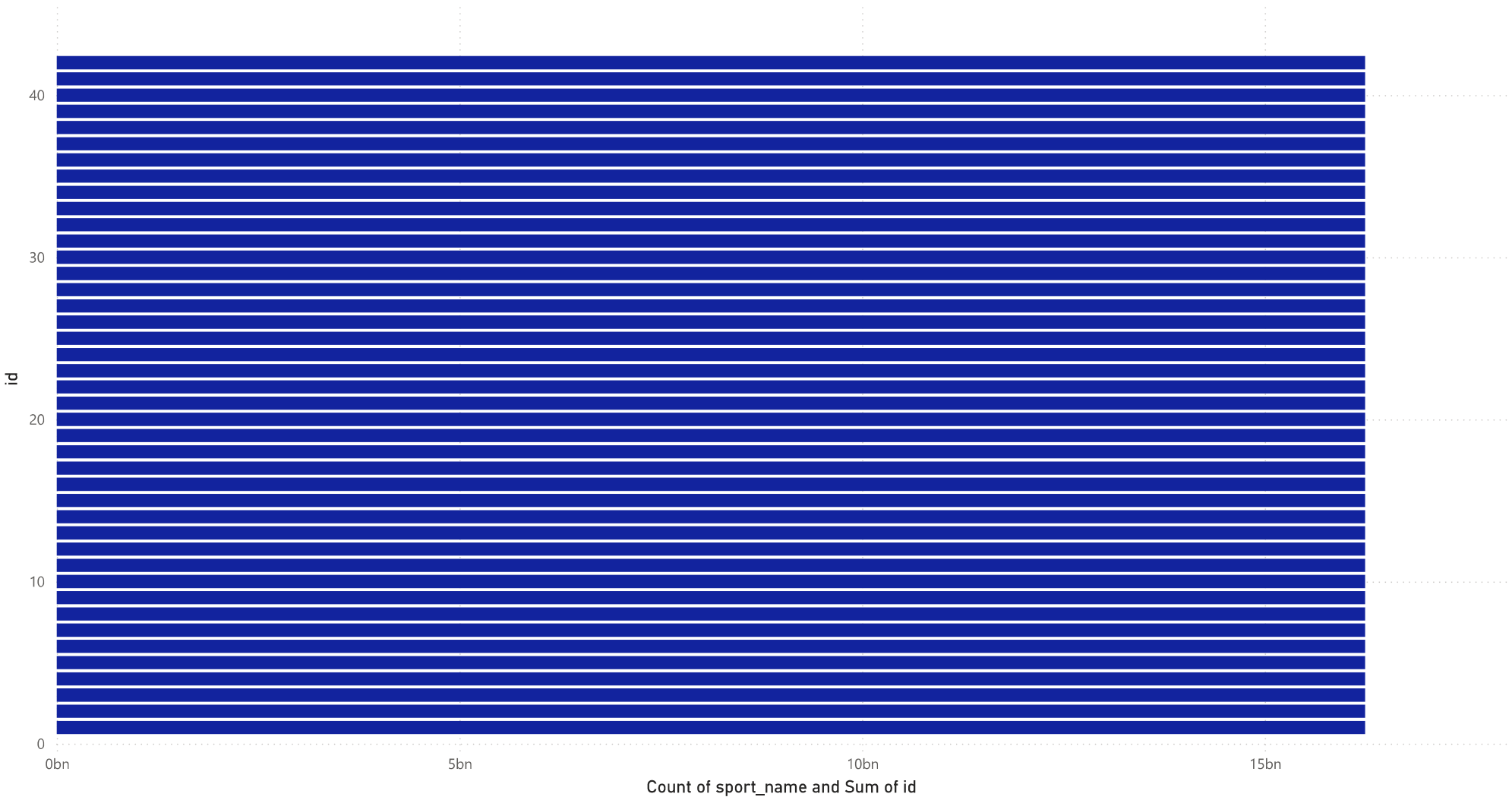
games_year, age

- [ ] 1896
- [ ] 1900
- [ ] 1904
- [ ] 1906
- [ ] 1908
- [ ] 1912
- [ ] 1920
- [ ] 1924
- [ ] 1928
- [ ] 1932
- [ ] 1936
- [ ] 1948
- [ ] 1952
- [ ] 1956
- [ ] 1960
- [ ] 1964
- [ ] 1968
- [ ] 1972
- [ ] 1976
- [ ] 1980
- [ ] 1984
- [ ] 1988
- [ ] 1992
- [ ] 1994

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

Sum of games_year by medal_name

How many medals have been awarded in each Olympics?

Sum of games_year

140K

120K

100K

80K

Bronze   Gold   NA   Silver
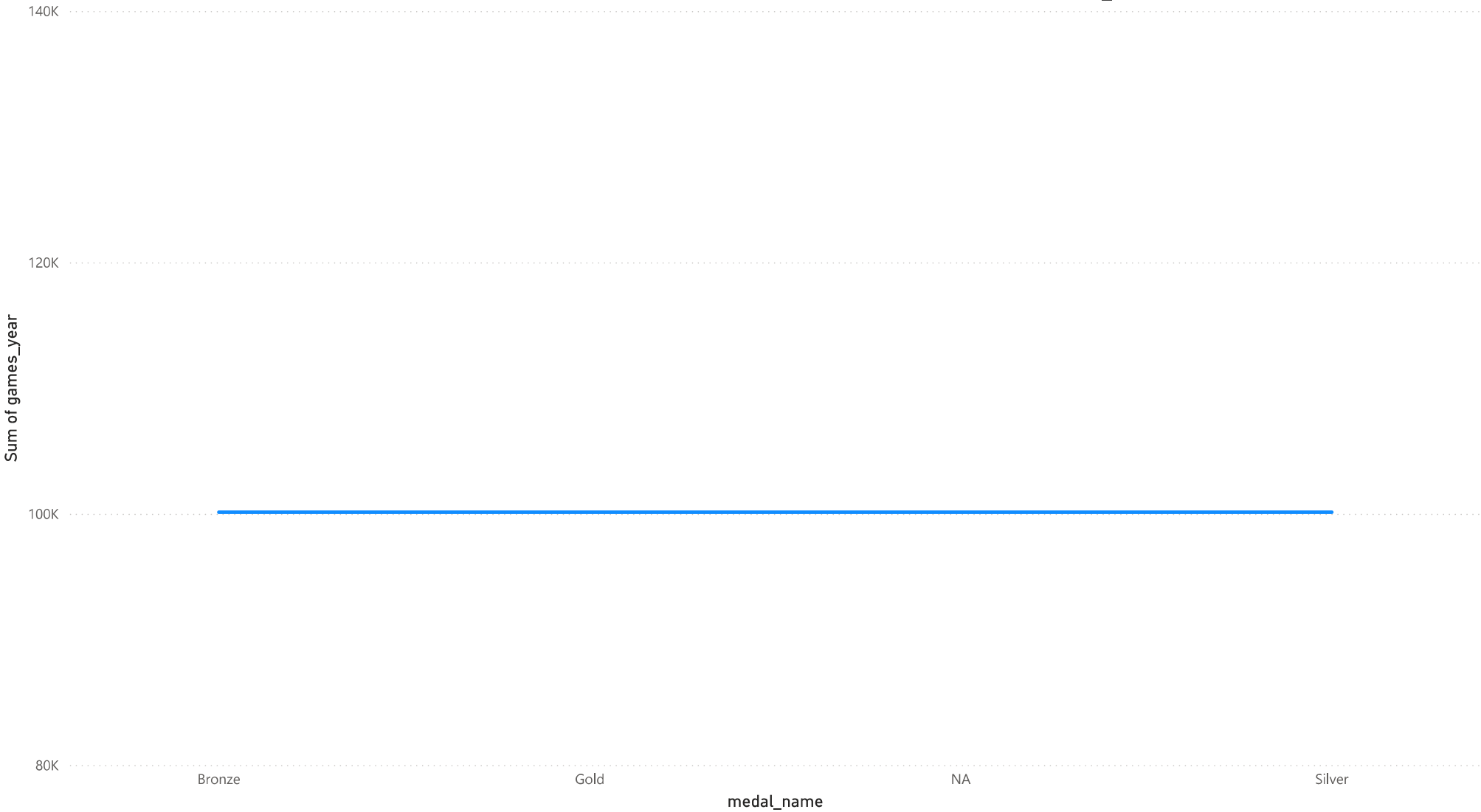
medal_name

Sum of region_id by medal_name (groups)

Which countries have the highest number of gold medals?

medal_name (groups) ● Bronze ● Gold ● NA ● Silver

Count of sport_name by medal_name (groups) 2

How does the medal distribution vary across different sports?

medal_name (groups) 2

Bronze

Gold

NA

Silver

0   10   20   30   40   50   60   70

Count of sport_name
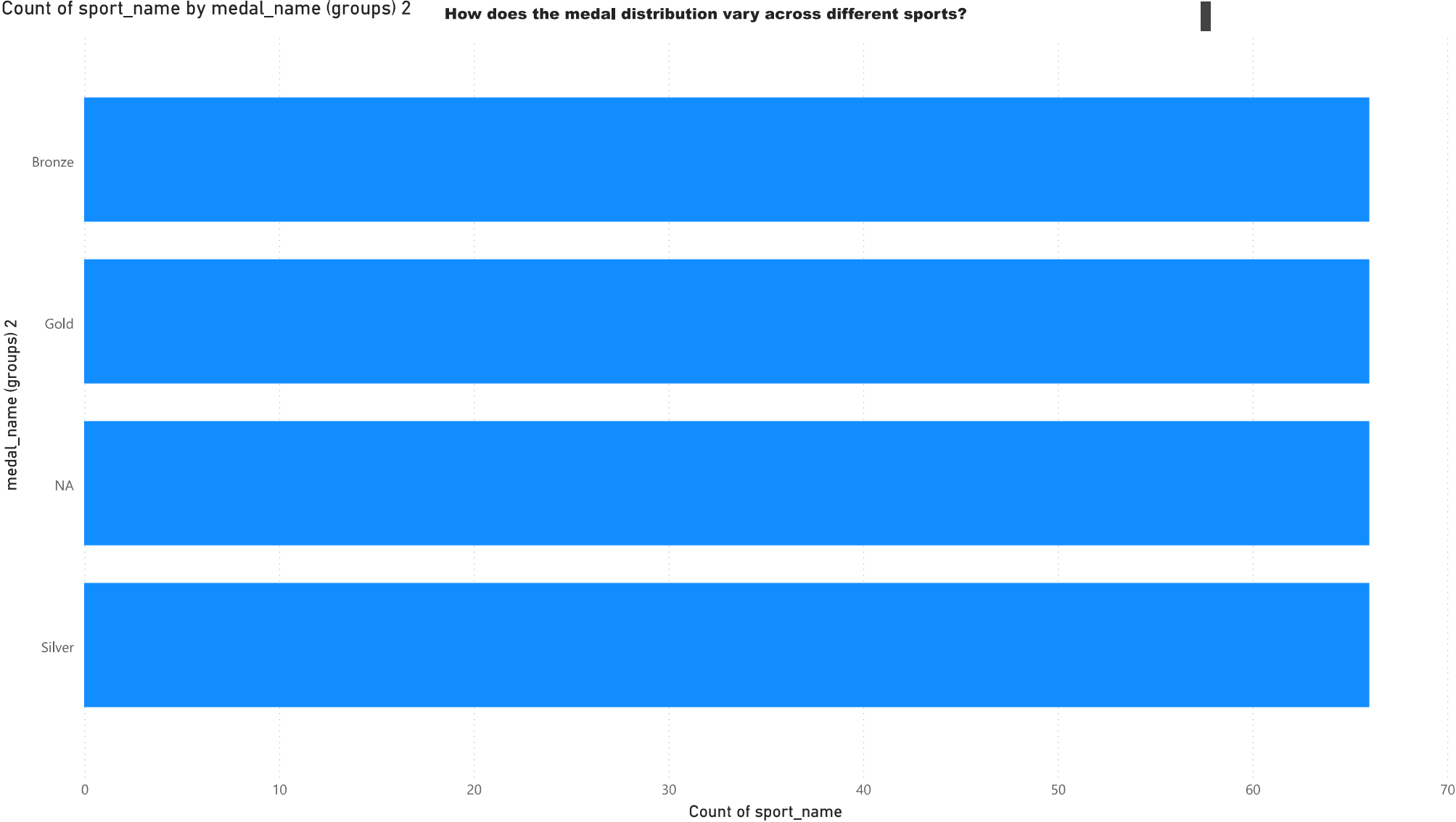
Count of games_year by medal_id

How many regions or NOCs participate in each Olympic Games?

Count of games_year

40

30

1   2   3   4

medal_id
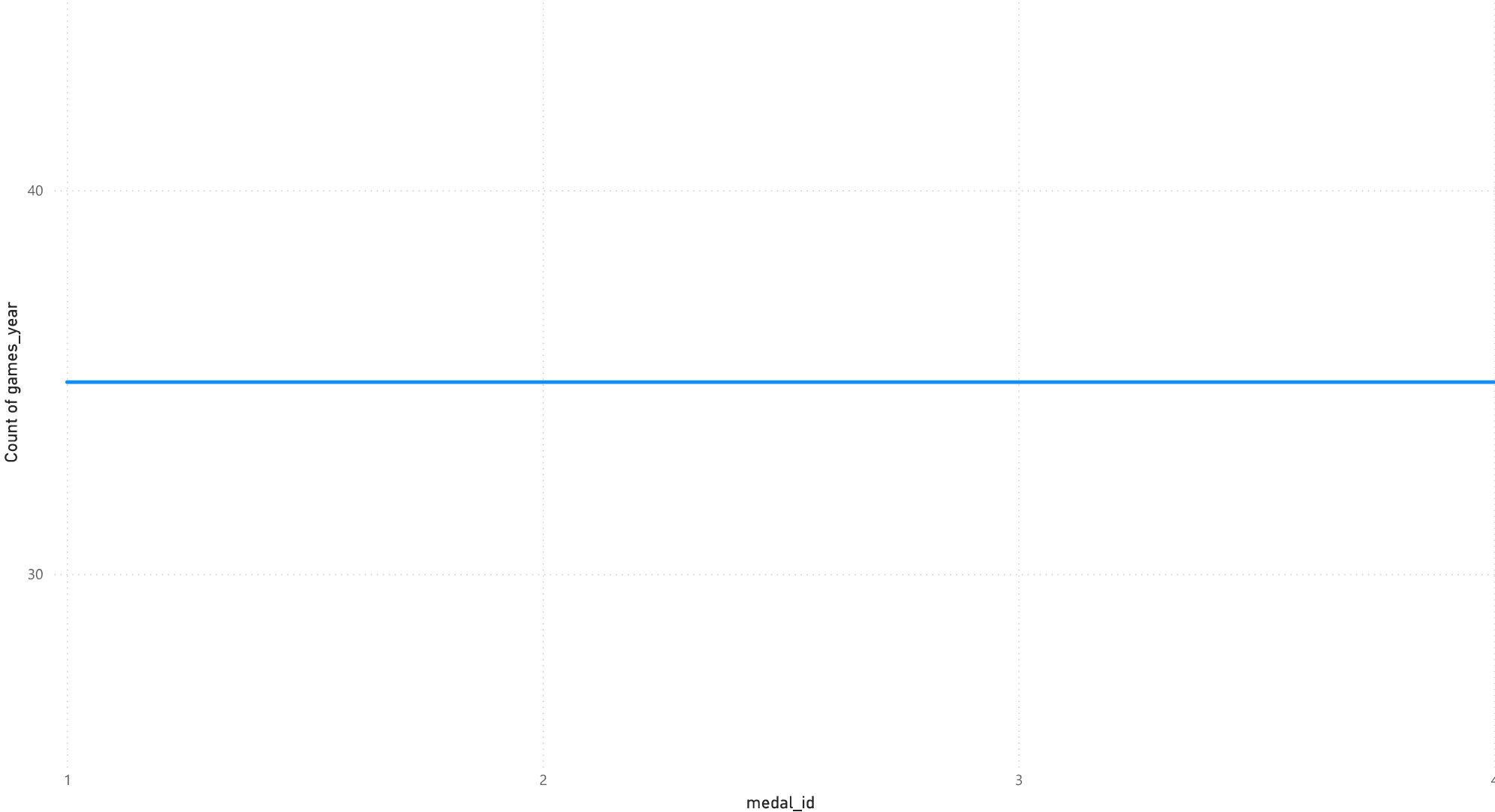
Count of region_id by event_id

Which regions have the highest number of participants in the Olympics?
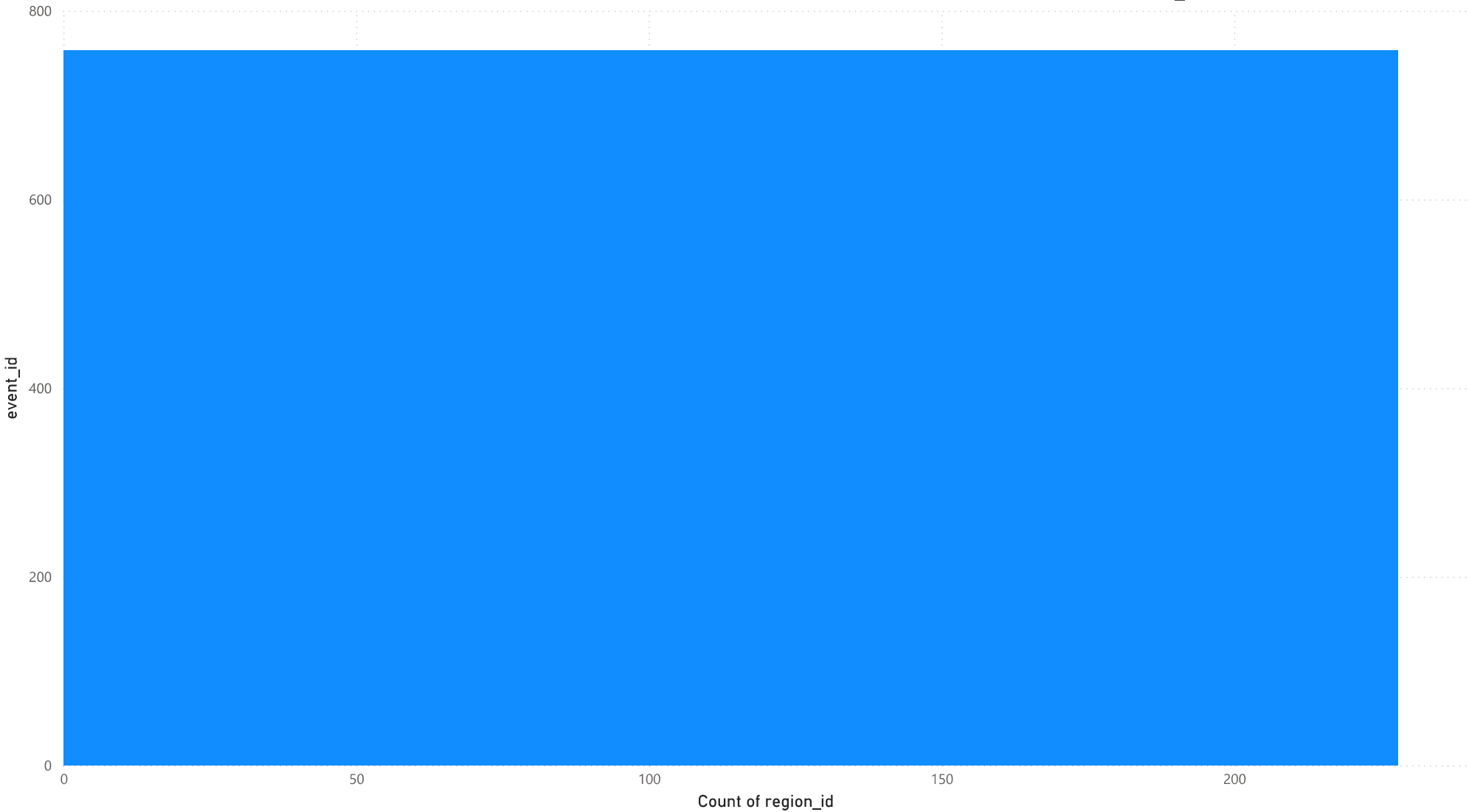
Count of region_id by medal_name

What is the distribution of medals among different regions?

228 (25%)   228 (25%)

228 (25%)   228 (25%)
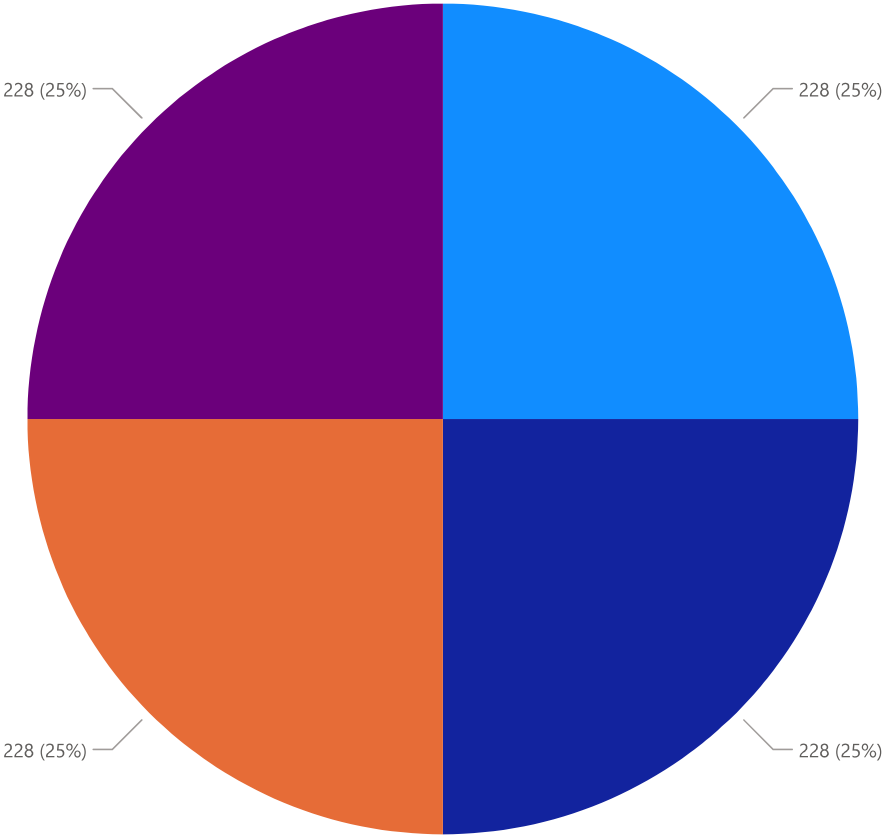
medal_name
- Bronze
- Gold
- NA
- Silver

EDA Problem Statements

Count of sport_name by full_name

full_name
- A Dijiang
- A Lamusi
- A. Aanantha S...
- A. J. Tyronne ...
- A. Joshua "Jos...
- A. W. Nancy "...
- Aa Bela Joaqu...
- Aadam Ismae...
- Aadjijatmiko F...
- Aadolf Fredrik...
- Aafke Hament
- Aage Albert L...
- Aage Avaldorf...
- Aage Berntsen
- Aage Birch
- Aage Brge Po...
- Aage Carl Chr...
- Aage Emil Brix
- Aage Emil Kir...
- Aage Ernst La...
- Aage Even He...
- Aage Heimann
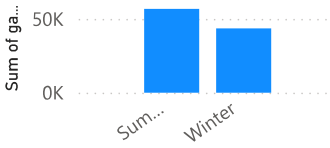- Aage Hermann
- Aage Hy Pede...
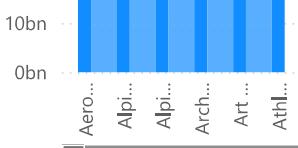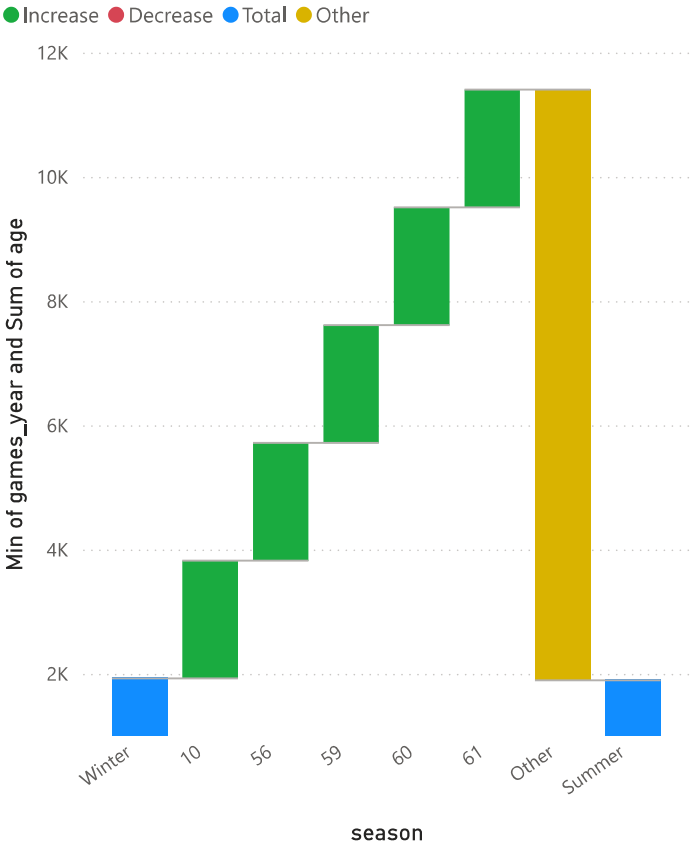- Aage Ingvar E...
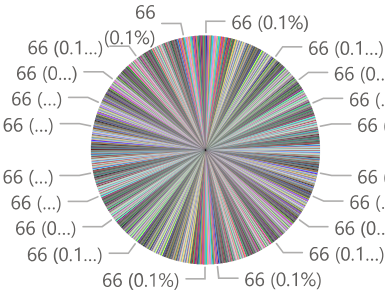
Sum of competitor_id by season

seas... ● Summer ● Winter
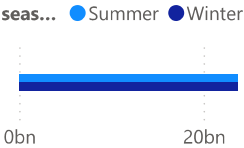
Sum of competitor_id by city_name
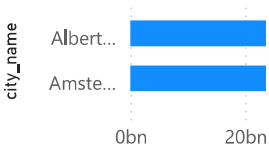
Sum of games_year by season

Sum of id by sport_name
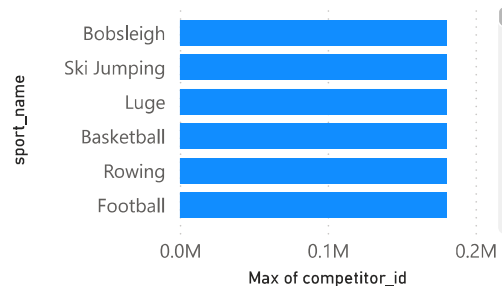
Min of games_year and Sum of age by season and age

● Increase ● Decrease ● Total ● Other

Max of region_id and Count of event_name by sport_name
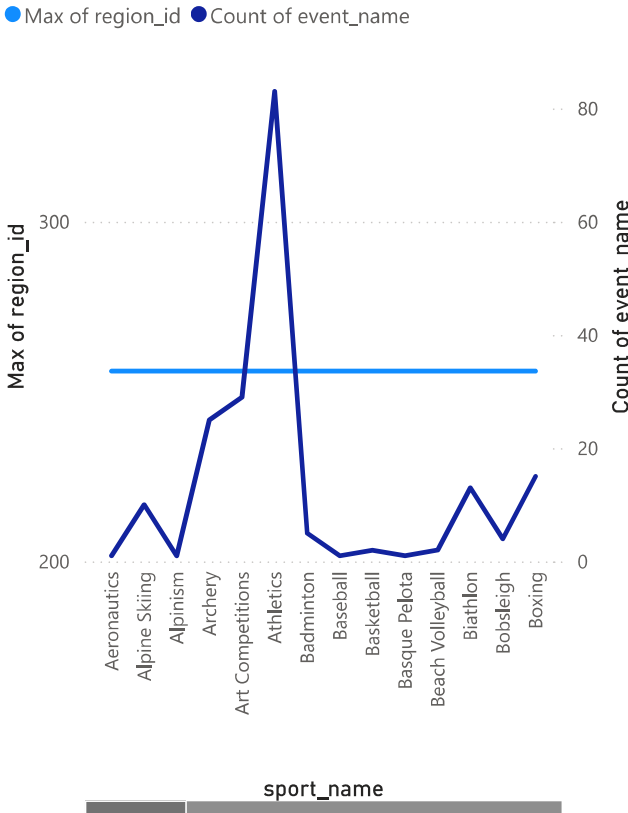
● Max of region_id ● Count of event_name

Max of competitor_id by sport_name

Count of sport_name by event_name

event_name
- Aeronauti...
- Alpine Skii...
- Alpine Skii...
- Alpine Skii...
- Alpine Skii...
- Alpine Skii...
- Alpine Skii...

66
Count of sport_name

full_name	Count of age
Ian Millar	10
Afanasijs Kuzmins	9
Hubert Raudaschl	9
Chen Jing	8
Durward Randolph Knowles	8
Francisco Boza Dibos	8
Josefa Idem-Guerrini	8
Lesley Allison Thompson-Willie	8
Li Na	8
Nino Salukvadze (-Machavariani)	8
Paul Bert Elvstrm	8
Piero D'Inzeo	8
Total	**74**

Sum of games_year by medal_name

100K (25%)
100K (25%)
100K (25%)
100K (25%)

medal_name
- Bronze
- Gold
- NA
- Silver

Sum of games_year by sport_name

Count of city_name by height and gender

gender ● F ● M

Count of games_name and Count of city_name

100%

Count of game... 51
Count of city_n... 42

82.4%
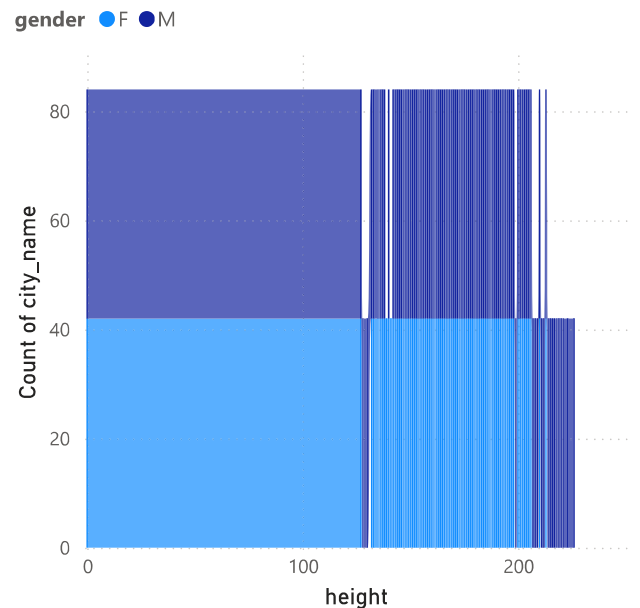
Count of season by city_name and games_name

Microsoft Bing

© 2025 TomTom, © 2025 Microsoft Corporation, © OpenStreetMap
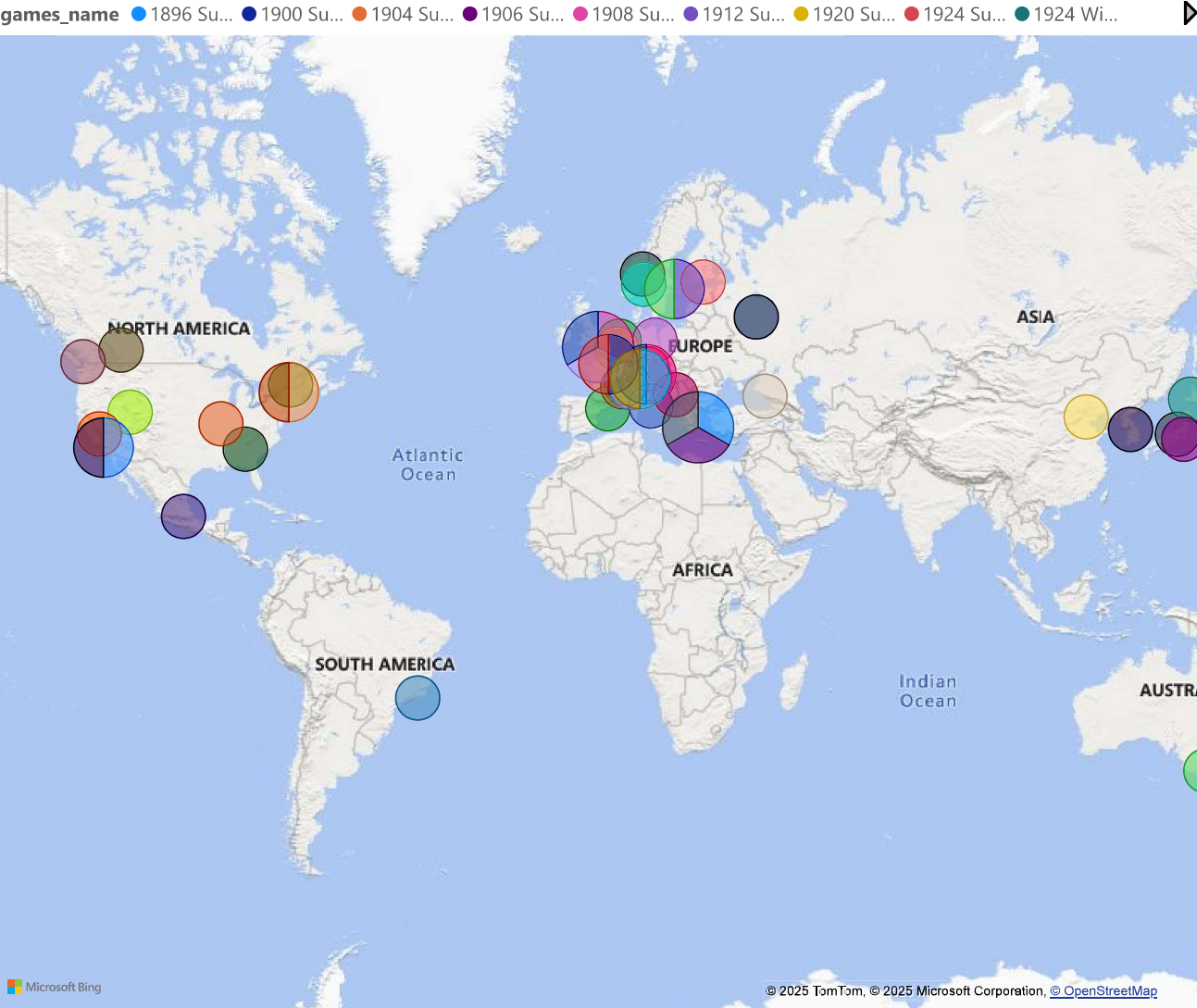
Sum of competitor_id by event_name and medal_name

medal_name ● Bronze ● Gold ● NA ● Silver

event_name (top to bottom):
Football Men'...
Ice Hockey M...
Hockey Men's ...
Water Polo M...
Basketball Me...
Cycling Men's ...
Gymnastics M...
Gymnastics M...
Rowing Men's...
Handball Men...
Gymnastics M...
Gymnastics M...
Gymnastics M...
Gymnastics M...
Athletics Men'...
Volleyball Me...
Gymnastics M...
Gymnastics M...
Athletics Men'...
Athletics Men'...
Bobsleigh Me...
Athletics Men'...
Volleyball Wo...
Gymnastics W...
Gymnastics W...
Gymnastics W...
Gymnastics W...
Alpine Skiing ...
Gymnastics W...
Fencing Men's...

0.0bn   0.5bn
Sum of competitor_id