



CAPSTONE PROJECT- SPORTS ANALYTICS

Abhinav Soran
DATA ANALYTICS

OVERVIEW

The capstone project focuses on analyzing olympic data, exploring athlete achievements, host city impacts, and sports trends, years frequency, medals achievement by regions. It utilizes sql, excel, and power bi to derive insights that shed light on the games' dynamics, enriching our understanding and guiding future enhancements.

The PROCESS

- *Data Retrieval from GitHub:*

Acquired the necessary dataset from a designated GitHub repository, containing essential information about Olympic Games. The dataset contained diverse Olympic events, countries, and their performances across different games.

- *Data Refinement and Enrichment:*

The procedures were done to ensure data quality and coherence. Furthermore, considered enhancing the dataset by introducing new problem scenarios that can amplify the depth of analysis.

- *Integration with Analytical Tools:*

Established connections between the dataset and various analytical tools. Linked the dataset with tools such as Power BI, Excel, and MySQL Workbench, to begin the process of data integration and processing.

- *Problem-Solving through Power BI:*

Utilized Power BI to address specific problem statements related to the Olympics. Utilized its robust functionalities for visualizing, exploring, and analyzing data, effectively unearthing insights and solutions.

- *Exploratory Data Analysis (EDA):*

Conducted exploratory data analysis using either Excel or SQL Workbench, based on the complexity of the analysis. Extracted meaningful patterns, correlations, and trends from the Olympic data to guide subsequent decision-making.

- *Compelling and Informative Presentation Creation:*

Developed an all-encompassing PowerPoint presentation that encapsulates the project's aims, methodologies, solutions to problem statements, and pivotal visualizations. Each problem statement is accompanied by a dedicated section containing relevant conclusions and insights.

- *Comprehensive Documentation:*

Compiled a detailed report that meticulously outlines the entire project lifecycle. This documentation encompasses facets such as data acquisition, refinement, formulation of problem statements, integration of analytical tools, Power BI solutions, insights derived from exploratory data analysis, and the visualizations featured in the PowerPoint presentation.

OBJECTIVES

Olympic Games hold immense historical and cultural significance, but their analysis can be complex due to various factors like athlete performance, host cities, and evolving sports trends. This project aims to comprehensively analyze Olympic data through SQL, Excel, and Power BI to unveil insights about athlete achievements, host city participants' frequencies, and sports trends across different editions.

Project Scope

The project involves the following key tasks:

- Thorough Olympic Data Analysis: Delve into athlete participation, medals, and event outcomes across multiple Games editions, highlighting trends and exceptional performances.
- Host City Influence: Investigated how host cities impact the Games' success and explored any patterns that emerge and participation, events, and sports frequency year-wise.
- Sports Evolution: Studied changes of popular sports over time, spotting new disciplines and their impact on medal rankings.

Project Objectives

- Comprehensive Data Analysis: Employed SQL and Excel to dissect Olympic data, revealing trends, statistical summaries, and insightful visualizations.
- Power BI Insights: Utilized Power BI's advanced capabilities to create interactive dashboards, providing deeper insights into athlete achievements, region-wise medals, and sports trends.
- Recommendations: Summarized analysis outcomes and derive meaningful conclusions to suggest potential enhancements for future Games editions.

Project Assessment

The success of the project will be assessed based on:

- Analytical Depth: The data has been explored and insights derived.
- Visual Impact: The quality of Power BI dashboards and visualizations in conveying complex information.
- Recommendation Practicality: The feasibility and relevance of suggestions provided for improving future Olympics.

This project holds significance as it unravels hidden patterns and trends within Olympic data, potentially guiding future Games' organization, athletes, govts, and countries and enhancing the worldwide appreciation for this celebrated event.

Significance

Analyzing Olympic data holds a crucial role in understanding the grandeur of this global sporting event. By exploring athlete achievements, host city impacts, and sports trends, this project is done by SQL, Excel, and Power BI. It brings forth valuable insights with broad implications. Let's delve into the significance of Olympic data analytics in simpler terms.

For athletes and aspiring Olympians, this analysis acts as a guiding light. By exploring athlete participation, medals, and event outcomes, athletes, sports trends, and regional strategies help them to gain a comprehensive understanding of past achievements and strategies. This knowledge will empower them to refine their training, set realistic goals, and align their efforts with evolving sports trends, increasing their chances of success on the grand Olympic stage and making themselves proud internationally.

Likewise, athletes and coaches can get benefit from Olympic data insights as they prepare for competitions. By studying past data, they can make smart choices about training and tactics. Coaches can find strategies that work and focus on the right areas for improvement. This approach helps athletes perform better and achieve their goals.

Sports organizers can also use Olympic data analysis to improve how events are managed. By looking at how host cities made Games successful and peaceful and the results and observing sports trends, organizers can plan better events, facilities, and schedules. This knowledge helps create better experiences for viewers and ensures the Olympics run smoothly.

Media and broadcasters and social platforms gain from Olympic data insights, too. They can use stats and history to tell compelling stories about athletes' journeys, unexpected wins, and rivalries, making the Games even more exciting for everyone. Sports enthusiasts gain a deeper appreciation for the Games through these insights. Analyzing sports evolution over time, athletes' evolution, and changing dynamics. This knowledge fosters a richer understanding of sports history and encourages engagement in diverse athletic pursuits.

Researchers and brands, investors, and governments can use Olympic data to learn about societal impacts and trends. By studying how different factors affect Games outcomes, they can make decisions that benefit athletes, events, and sports.

In conclusion, Olympic data analytics offers a panoramic view of the Games' significance to athletes, sports organizations, fans, and society at large. It enriches the understanding of sports dynamics, empowers athletes and coaches, optimizes event management, enhances media narratives, attracts brands, and guides the evolution of the Games. Through insightful data analysis using SQL, Excel, and Power BI, collectively contribute to shaping the future of the Olympics, fostering global unity and athletic excellence.

Data Dictionary

This data dictionary provides a concise overview of the tables and their respective fields in the Olympic Games database, facilitating understanding and reference for data analysis and exploration.

Table: Games (LOOKUP_TABLE)

Id: Unique identifier for each Olympic Games edition.

Year: Year and season of the Olympic

Season: Type of Olympics (Summer/Winter).

Table: City (Lookup_Table)

Id: Unique identifier for host cities.

City_name: Name of the host cities

Table: Games_City (Fact_table)

Game_id: Foreign key linking to the Games table.

City_id= Foreign key linking to the city table.

Table: NOC_Region (Lookup_Table)

Id= unique id for each participated regions

NOC: Unique code for each National Olympic Committee.

Region_name: Region associated with the NOC and Olympic participation.

Table: Person (Lookup_table)

Id: Unique identifier for individuals athletes in the Olympics.

Full_name: Name of the all participants.

Sex: Gender of all participants

Weights: Wights of all participating candidates.

Heights: Heights of all participating candidates.

Table: Person_Region (Fact_table)

Person_id: Foreign key linking to the Person table.

Region_id: Foreign key Region associated with the NOC_region table.

Table: Sport (LOOKup_Table)

Id: Unique identifier for each sports.

sport_name: Name of the sport.

Table: Medal (LOOKUP_Table)

Id: Unique identifier for each medal.

Medal_name: Type of medal (gold, silver, bronze).

Table: Competitor_Event (fact_Table)

event_id: Foreign key linking to the Event table.

Competitor_id = Foreign key linking to the games_competitor table.

Medal_id= Foreign key linking to the Medal table.

Table: Event (LOOKUP_table)

Id: Unique identifier for each event.

event_name: Name of the event.

schedule: Schedule details of the event.

sport_id: Foreign key linking to the Sport table.

Table: Games_Competitor (lookup_Table)

Id: Unique identifier of each competitor.

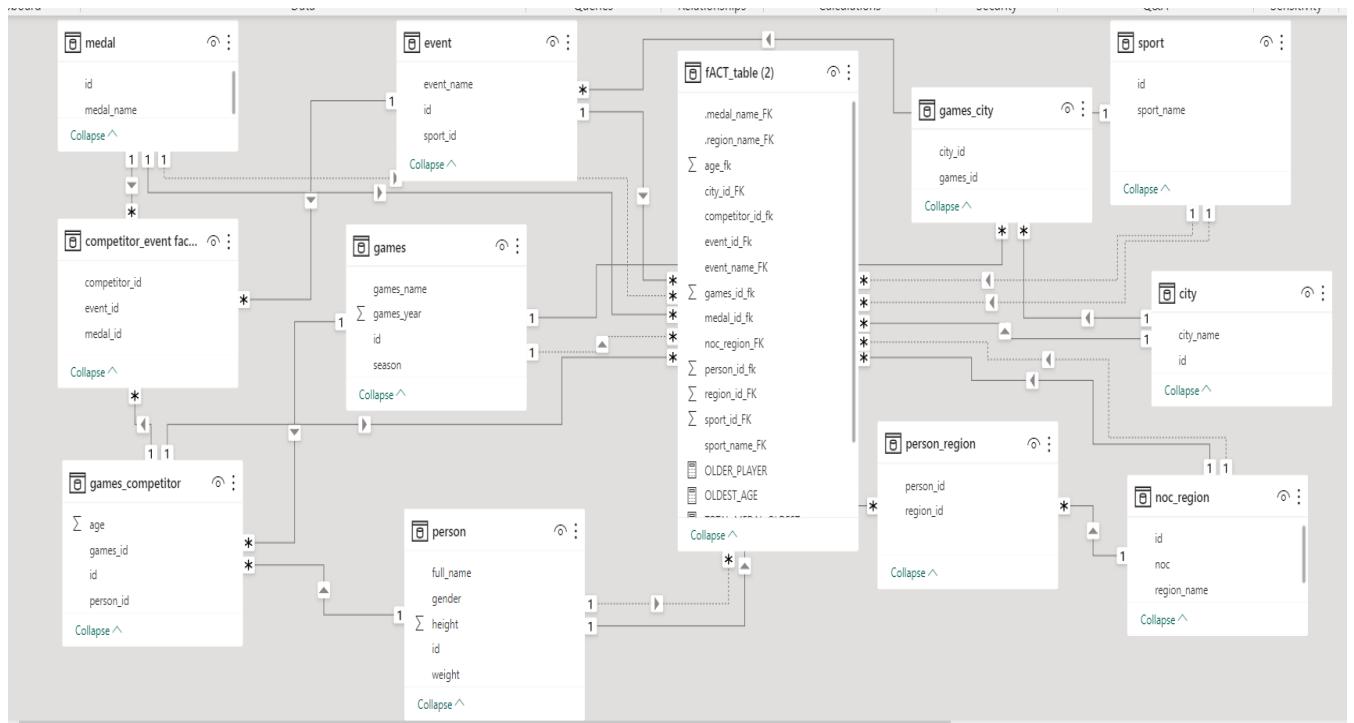
Age: The age of each participant.

Person_id: Foreign key linking to the Person table.

Games_id: Foreign key linking to the games table.

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

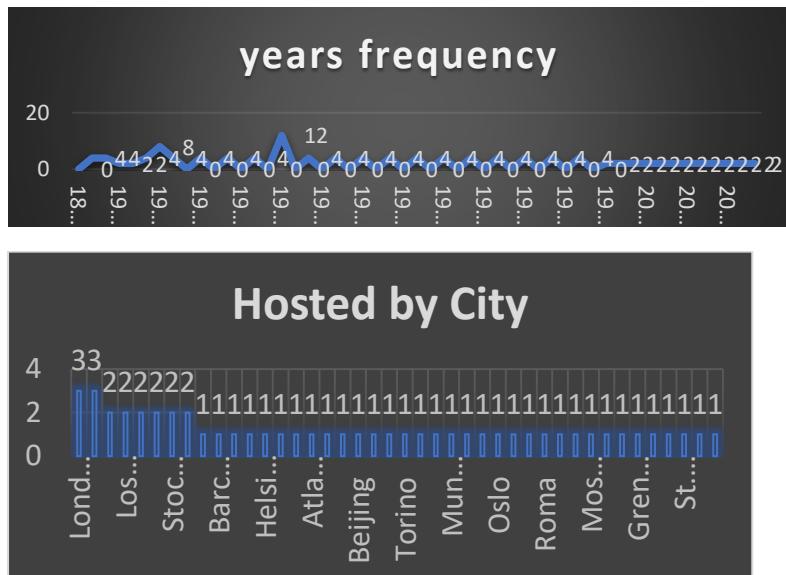


EDA PROBLEM STATEMENT

Are there any trends or patterns in the frequency of hosting the Olympic Games?

If we see the graph hosted by city we can see that the most hosted city are London and Athina.

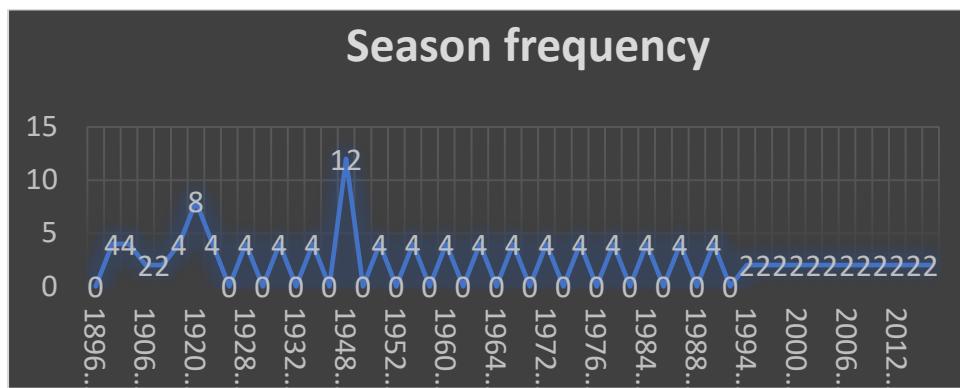
And if we see the year frequency chart we can see the year gap. In the table, there is a gap of 2 years between the 1904 to 1906 summer and the 1906 to 1908 Summer Olympics, and also 1912 summer to 1920 Summer Olympics have 8 years gap. And from the data, I found out that in 1924 Winter Olympics started, and after the 1936 summer it occurred again in the 1948 winter, with 12 years gap. Even also some winter and summer Olympics happened in the same year.



season	games_name	games_year	previous_year	duration_years
Summer	1896 Summer	1896	NULL	NULL
Summer	1900 Summer	1900	1896	4
Summer	1904 Summer	1904	1900	4
Summer	1906 Summer	1906	1904	2
Summer	1908 Summer	1908	1906	2
Summer	1912 Summer	1912	1908	4
Summer	1920 Summer	1920	1912	8
Winter	1924 Winter	1924	1920	4
Summer	1924 Summer	1924	1924	0
Winter	1928 Winter	1928	1924	4
Summer	1928 Summer	1928	1928	0
Winter	1932 Winter	1932	1928	4
Summer	1932 Summer	1932	1932	0
Winter	1936 Winter	1936	1932	4
Summer	1936 Summer	1936	1936	0
Winter	1948 Winter	1948	1936	12
Summer	1948 Summer	1948	1948	0
Summer	1952 Summer	1952	1948	4

How has the duration of Olympic Games changed over time?

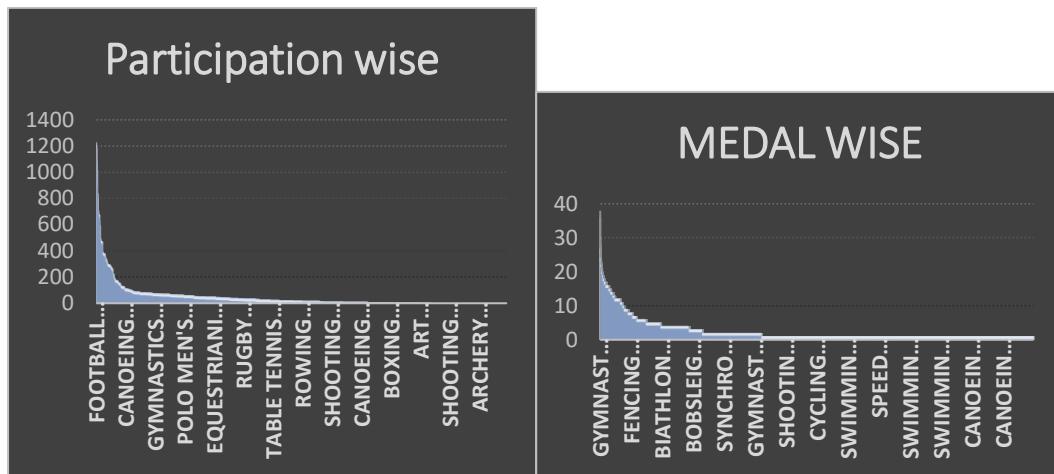
In the beginning, the duration gap between the winter olympics and summer Olympics was more significant. The winter olympics commenced in 1924, after a 24-year gap following the summer olympics. Additionally, there was another gap of 12 years between the summer Olympics and the winter Olympics.



Summer Year	Winter Year	Duration Between Summer Winter
1896	1924	28
1900	1924	24
1904	1924	20
1906	1924	18
1908	1924	16
1912	1924	12
1920	1924	4
1924	1928	4
1928	1932	4
1932	1936	4
1936	1948	12
1948	1952	4
1952	1956	4
1956	1960	4
1960	1964	4
1964	1968	4
1968	1972	4
1972	1976	4
1976	1980	4
1980	1984	4
1984	1988	4
1988	1992	4
1992	1994	2
1996	1998	2
2000	2002	2
2004	2006	2
2008	2010	2
2012	2014	2
2016	NULL	NULL

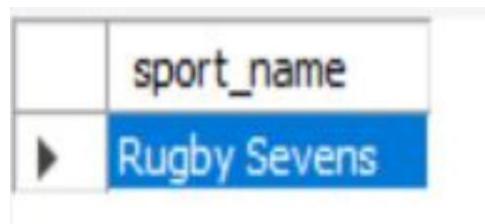
Are there any notable events or occurrences associated with specific Olympic Games?

In the summer Olympics, men's football attracted the most participants among all events, while during the winter Olympics, men's ice hockey garnered the highest number of participants. Men's Gymnastics secured the most medals during the 1896 summer games, while the winter alpinism mixed event achieved the highest medal count in the 1924 Olympics. However, when considering the entire winter Olympics, men's hockey emerged as the event with the highest medal count.



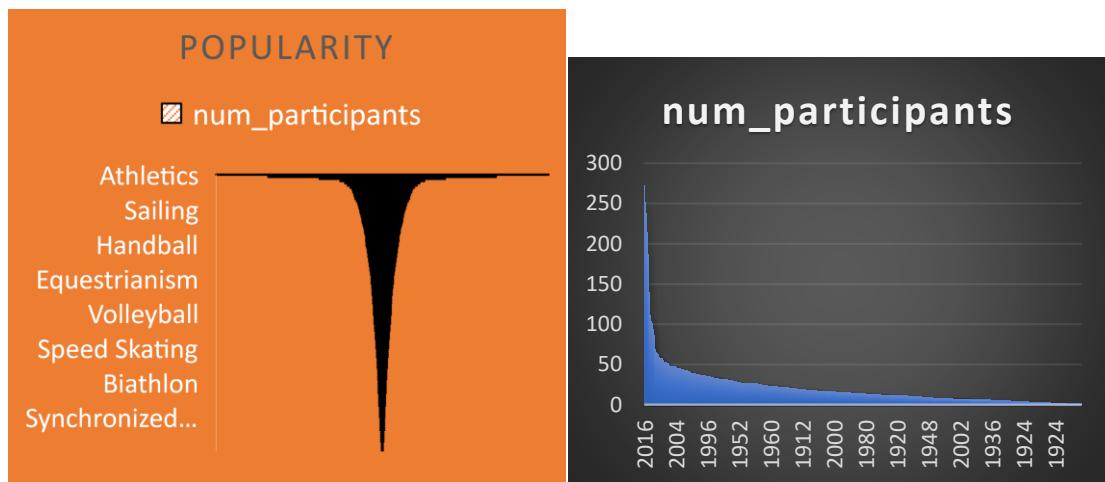
Are there any emerging sports that have been recently added to the Olympics?

From the data, have found that the new sports that have been added recently in 2016 is **Rugby Sevens**



How has the popularity of certain sports changed over the years?

The data table I have provided indicates a consistent upward trend in the number of participants for each sport over the years. Notably, the sport with the highest increase in participants is Athletics. Upon analyzing the data, it is evident that participation in Athletics has progressively risen over time. In the initial year, only one participant took part in Athletics, whereas by 2016, the number had surged to 273. This observation highlights a gradual and steady growth in participant numbers over the years.



Are there any sports that are specific to a particular region or culture?

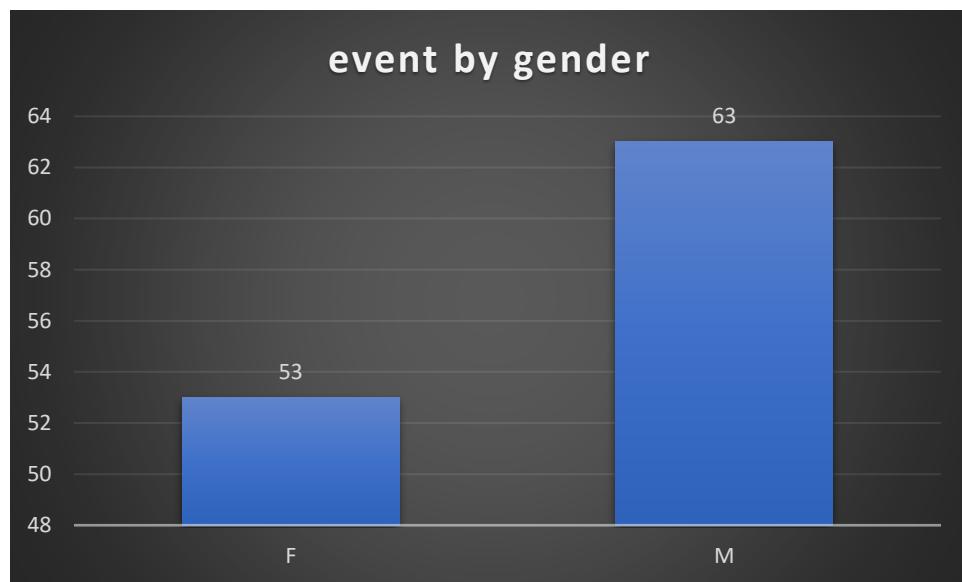
I found out that after the analysis there are 14 sports, these are particular to some particular regions.

- **Aeronautics in Switzerland,**
- **Basque Pelota in Spain,**
- **Cricket in France and UK**
- **Croquet in France**
- **Jeu De Paume in UK and USA**
- **Roque in the USA**
- **Motorboating in France and UK**
- **Racquets in the UK**

SPORTS	REGION	TOTAL SPORT ID
Aeronautics	Switzerland	1
Basque Pelota	Spain	20
Cricket	France	16
	UK	272
Croquet	France	255
Jeu De Paume	UK	288
	USA	64
Roque	USA	129
Motorboating	France	38
	UK	532
Racquets	UK	492

Are there any sports that have a higher number of events for one gender compared to others?

There is a notable disparity in the number of male participants compared to female participants in sports. This trend is consistent across various sports, with a majority of them featuring a higher number of events for males than for females. This discrepancy is widespread and not tied to any specific sport; rather, it is a prevalent pattern observed across many different sports.



Are there any new events that have been introduced in recent editions of the Olympics?

From the data I fetched five events that were introduced in the Olympics in the year 2016. Prior to that, these events were not part of the Olympic program. Through careful analysis of the data, it is evident that these five events represent recent additions to the Olympic Games.

	event_name	games_year
▶	Rugby Sevens Men's Rugby Sevens	2016
	Rugby Sevens Men's Rugby Sevens	2016
	Wrestling Women's Light-Heavyweight, Freestyle	2016
	Rugby Sevens Men's Rugby Sevens	2016
	Rugby Sevens Men's Rugby Sevens	2016
	Sailing Women's Skiff	2016
	Rugby Sevens Women's Rugby Sevens	2016
	Sailing Women's Skiff	2016

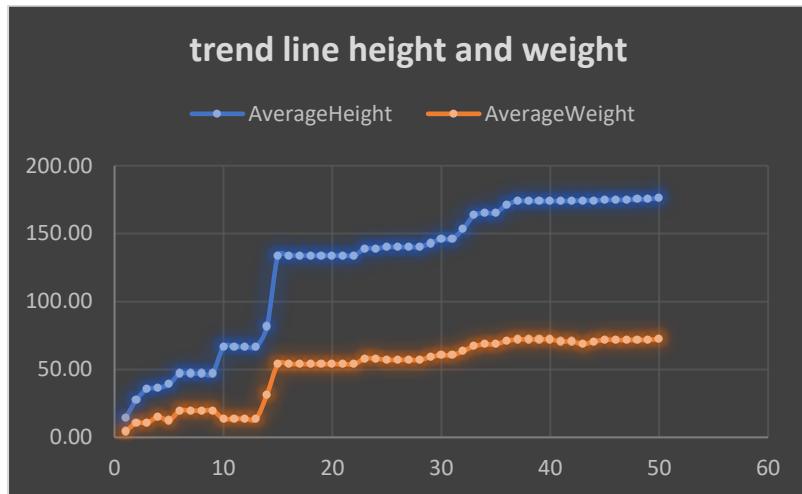
Are there any events that have been discontinued or removed from the Olympics?

After a thorough review of event names and their respective years, it's clear that many events were halted at a certain point. Interestingly, some events began and ended within the same year. Based on this detailed analysis, it can be inferred that these specific events have been discontinued and removed from the Olympics.

event_name	g_First	g_Last
Tug-Of-War Men's Tug-Of-War	1900	1920
Speed Skating Women's 500 metres	1960	2014
Speed Skating Women's 1000 metres	1960	2014
Cross Country Skiing Men's 10 kilometres	1992	1998
Cross Country Skiing Men's 50 kilometres	1924	2014
Cross Country Skiing Men's 10/15 kilometres Pursuit	1992	1998
Cross Country Skiing Men's 4 x 10 kilometres Relay	1936	2014
Cross Country Skiing Men's 30 kilometres	1956	2002
Ice Hockey Men's Ice Hockey	1920	2014
Biathlon Women's 7.5 kilometres Sprint	1992	2014
Swimming Men's 400 metres Breaststroke	1912	1920
Alpine Skiing Men's Super G	1988	2014
Alpine Skiing Men's Downhill	1948	2014
Bob-sleigh Men's Four	1932	2014
Boxing Men's Featherweight	1904	2008
Speed Skating Men's 1,500 metres	1924	2014
Cycling Women's 500 metres Time Trial	2000	2004
Alpine Skiing Men's Slalom	1948	2014
Luge Women's Singles	1964	2014
Speed Skating Men's 500 metres	1924	2014
Alpine Skiing Men's Giant Slalom	1952	2014
Canoeing Men's Canadian Doubles, 500 metres	1976	2008
Speed Skating Men's 10,000 metres	1924	2014
Sailing Mixed Two Person Heavyweight Dinghy	1956	1992
Fencing Men's Sabre, Team	1906	2012
Alpine Skiing Men's Combined	1936	2014
Art Competitions Mixed Sculpturing, Unknown Event	1932	1948
Sailing Mixed Three Person Keelboat	1948	2000
Cross Country Skiing Men's 15 kilometres	1956	2014
Shooting Mixed Trap	1968	1992
Cycling Men's Road Race, Team	1912	1956
Speed Skating Men's 5,000 metres	1924	2014
Sailing Mixed 8 metres	1908	1936
Shooting Mixed Skeet	1968	1992
Gymnastics Men's Team All-Around, Free System	1912	1920
Rowing Men's Coxed Fours	1900	1992
Wrestling Men's Light-Flyweight, Freestyle	1904	1996
Boxing Men's Light-Middleweight	1952	2000
Weightlifting Men's Flyweight	1972	1996
Judo Men's Open Class	1964	1984
Fencing Women's Foil, Team	1960	2012
Cycling Men's 1,000 metres Time Trial	1928	2004
Table Tennis Men's Doubles	1988	2004
Rowing Women's Coxed Fours	1976	1988
Sailing Mixed Windsurfer	1988	1988
Sailing Mixed One Person Dinghy	1924	2008
Alpine Skiing Women's Giant Slalom	1952	2014
Snowboarding Women's Boardercross	2006	2014
Biathlon Men's 10 kilometres Sprint	1980	2014
Figure Skating Mixed Pairs	1908	2014
Ski Jumping Men's Normal Hill, Individual	1924	2014

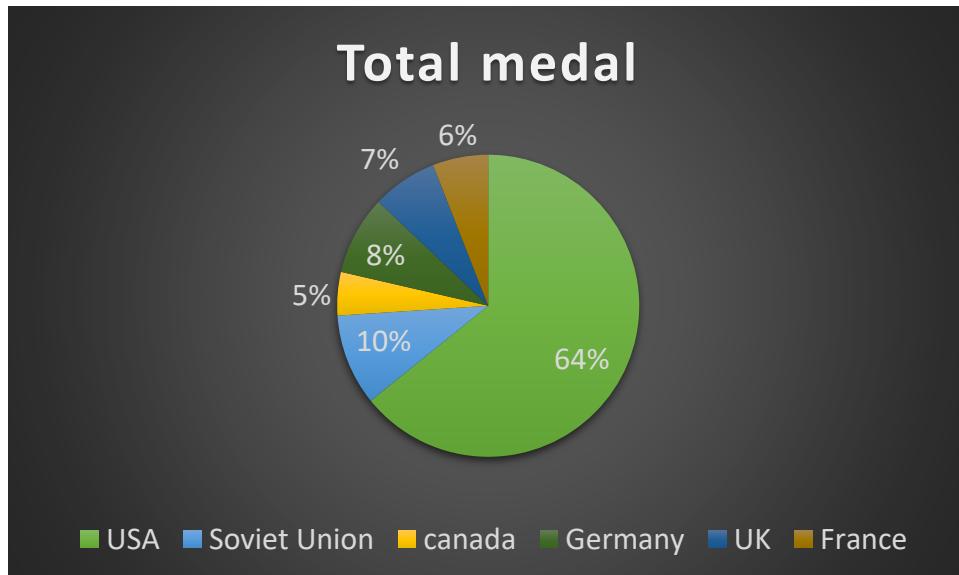
Are there any notable trends in the height and weight of participants over time?

Participants' weights and heights were initially lower but have since experienced an increase over time. This upward trend is evident in the data analysis, reflecting a progression towards higher weight and height measurements.



Are there any dominant countries or regions in specific sports or events?

From the gathered data and subsequent analysis, it becomes clear that the USA and the Soviet Union clinched Olympic victories twice each. Notably, the USA emerged victorious a remarkable 18times, underscoring its dominant performance over other regions. The provided table highlights how each region excelled in specific sports by earning the highest medal count, further reinforcing their supremacy. Noteworthy is the USA's impressive array of medals across diverse sports and mainly in athletics sports. Additionally, the data unveils various other nations, as outlined in the table, that have showcased their dominance by clinching medals in their respective sports over different regions.

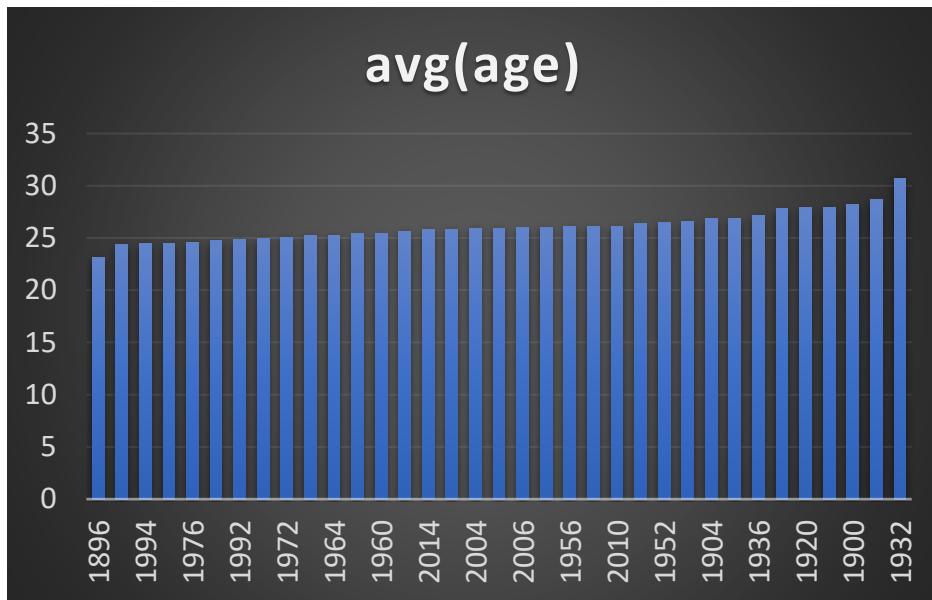


Rank	games	region_name	Bronze	Gold	Silver	Grand Total
1	1904	USA	114	115	119	348
2	1920	USA	37	106	45	188
3	1924	USA	49	93	45	187
4	1928	USA	18	51	28	97
5	1932	USA	60	84	64	208
6	1948	USA	36	86	33	155
7	1952	USA	27	80	44	151
8	1984	USA	51	181	117	349
9	1996	USA	49	153	48	250
10	2000	USA	52	132	60	244
11	2002	USA	15	10	57	82
12	2004	USA	71	116	77	264
13	2008	USA	80	128	110	318
14	2010	USA	22	12	59	93
15	2012	USA	46	145	56	247
16	2016	USA	71	138	54	263

These are the regions that have received the highest number of medals in each sport, and dominate others.						
region_name	sport_name	Bronze	Gold	Silver	Tota Med	
USA	Athletics	222	535	315	1072	
USA	Swimming	170	625	263	1058	
Italy	Fencing	74	164	137	375	
USA	Rowing	81	182	101	364	
Canada	Ice Hockey	34	220	93	347	
Canada	Ice Hockey	24	220	93	347	
USA	Basketball	33	286	24	343	
Soviet Union	Gymnastics	59	164	105	328	
Netherlands	Hockey	88	80	87	255	
USA	Shooting	40	116	36	192	
Germany	Equestrianism	45	100	46	191	
Germany	Canoeing	41	88	61	190	
Hungary	Water Polo	36	118	29	183	
UK	Cycling	27	58	66	181	
Norway	Cross Country Skiing	37	54	73	164	
UK	Sailing	32	67	49	148	
USA	Diving	47	50	50	147	
Soviet Union	Volleyball	12	81	53	146	
Soviet Union	Wrestling	30	73	35	138	
Soviet Union	Wrestling	30	73	35	138	
Brazil	Football	34	17	85	136	
Netherlands	Speed Skating	42	40	36	118	
Austria	Alpine Skiing	41	34	39	114	
Cuba	Baseball	1	64	48	112	
USA	Boxing	38	48	22	108	
Soviet Union	Handball	23	60	14	97	
Switzerland	Bobsleigh	32	28	32	92	
Germany	Biathlon	13	36	36	85	

How has the average age of Olympic athletes changed over the years? Is there any noticeable trend?

The 1932 Olympics witnessed a significantly higher average age among participants compared to other years, while the 1896 Games stood out with a notably lower average age. In contrast, most other years saw an average age close to 26.



Are there any countries that consistently perform well in multiple Olympic editions

Based on the analysis conducted, it is evident that the USA holds the highest medal count and has emerged victorious in the Olympics a remarkable 16 times. While the medal count of the Soviet Union is comparatively lower, the country's 8 Olympic victories surpass many other nations. This leads to the conclusion that both the USA and the Soviet Union have maintained a consistent and impressive performance record in the Olympics, showcasing their enduring excellence over time and also Germany, and France also constantly doing well

Rank	games	region_name	Bronze	Gold	Silver	Grand Total
1	1904	USA	114	115	119	348
2	1920	USA	37	105	45	188
3	1924	USA	49	93	45	187
4	1928	USA	18	51	28	97
5	1932	USA	60	84	64	208
6	1948	USA	36	86	33	155
7	1952	USA	27	80	44	151
8	1956	USA	51	81	117	349
9	1960	USA	49	153	48	250
10	2000	USA	52	132	60	244
11	2002	USA	15	10	57	82
12	2004	USA	71	116	77	264
13	2008	USA	80	128	110	318
14	2010	USA	22	12	59	93
15	2012	USA	46	145	56	247
16	2016	USA	71	138	54	263

region_name	games	GoldMedal	SilverMedal	BronzeMedal	TotalMedals
USA	1896	104	157	157	1626
USA	1898	158	133	157	468
USA	1904	204	210	125	541
USA	1896	277	95	98	512
USA	1894	339	205	94	1499
Germany	1932	172	136	119	478
Unified Team	1992	213	121	142	1459
Soviet Union	1980	286	245	147	1447
USA	1932	155	121	105	421
Germany	1936	153	126	50	405
USA	2000	242	106	95	404
Soviet Union	1988	312	134	179	1394
Australia	2000	108	111	97	1393
USA	2008	223	197	142	1359
USA	2016	264	94	136	1328
France	1992	27	23	84	1309
USA	2004	205	148	132	1306
Canada	1988	6	12	33	120
UK	1908	25	101	101	120
USA	1976	131	109	61	1254
Soviet Union	1976	271	180	172	1244
USA	1972	127	140	79	1236
West German	1988	64	82	85	1232
USA	2012	255	108	83	1229
West German	1972	81	38	77	1214
UK	2012	91	56	83	1209
West German	1960	53	136	96	1154
United States	1984	44	69	59	1152
Canada	1976	2	14	27	1162
South Korea	1988	48	69	24	1137
Canada	1992	57	59	29	1127
USA	1968	160	62	60	1126
Germany	1996	67	57	78	1114
Soviet Union	1972	244	93	111	1108
Spain	1992	94	34	6	1107
Brazil	2016	60	13	10	1105
Russia	2008	79	61	101	1104
Germany	1936	168	131	107	1100
UK	1992	14	5	72	1099
Russia	2000	113	118	99	1099
China	2008	113	79	80	1097
Italy	1992	43	45	39	1087
Russia	2004	88	91	168	1082

Are there any sports or events that have a higher number of medalists from a specific region?

With a remarkable tally of 18 Olympic triumphs, the USA stands out as a frontrunner. The accompanying tables spotlight the years and sports, Athletics included, in which the USA has emerged victorious. This comprehensive analysis underscores the USA's enduring dominance in Athletics, along with its prowess across a range of other sports. As a result, it's clear that a significant number of medalists hail from the USA.

Rank	games	region_name	Bronze	Gold	Silver	Grand To
1	1904	USA	114	115	119	348
2	1920	USA	37	106	45	188
3	1924	USA	49	93	45	187
4	1928	USA	18	51	28	97
5	1932	USA	60	84	64	208
6	1948	USA	36	86	33	155
7	1952	USA	27	80	44	151
8	1984	USA	51	181	117	349
9	1996	USA	49	153	48	250
10	2000	USA	52	132	60	244
11	2002	USA	15	10	57	82
12	2004	USA	71	116	77	264
13	2008	USA	80	128	110	318
14	2010	USA	22	12	59	93
15	2012	USA	46	145	56	247
16	2016	USA	71	138	54	263

region_name	games	GoldMedal	SilverMedal	BronzeMedal	TotalMedals
USA	1992	171	104	157	1731
USA	1988	168	133	97	1628
USA	1904	204	210	195	1514
USA	1996	277	95	98	1512
USA	1984	339	205	94	1499
Germany	1992	172	136	119	1478
Unified Team	1992	213	121	142	1459
Soviet Union	1980	367	286	245	1447
USA	1932	155	121	105	1421
Sweden	1912	198	126	50	1405
USA	2000	242	109	85	1405
Soviet Union	1988	312	134	179	1394
Australia	2000	108	111	97	1393
USA	2008	223	197	142	1359
USA	2016	264	94	136	1328
France	1992	27	23	84	1303
USA	2004	205	148	132	1306
Canada	1988	6	12	33	1260
UK	1908	251	181	131	1257
USA	1976	131	109	61	1254
Soviet Union	1976	271	180	172	1244
USA	1972	127	140	79	1236
West German	1988	64	82	85	1232
USA	2012	255	108	83	1229
West German	1972	81	38	77	1214
UK	2012	91	56	83	1209
West German	1984	59	136	96	1194
Canada	1976	44	65	53	1193
Canada	1976	2	14	27	1162
South Korea	1988	48	69	24	1137
Canada	1992	57	59	29	1127
USA	1968	160	62	60	1126
Germany	1996	67	57	78	1114
Soviet Union	1972	244	93	111	1108
Spain	1992	94	34	6	1107
Brazil	2016	60	13	10	1106
Russia	2008	78	81	101	1104
Germany	1936	168	131	107	1100
UK	1992	14	5	72	1100
Russia	2000	113	118	99	1099
China	2008	113	79	80	1097
Italy	1992	43	45	39	1087

What are some notable instances of unexpected or surprising medal wins?

Analyzing the data reveals notable Olympic trends. Some regions, like North Burnero and Newfoundland, had early successes but limited subsequent participation. The surprise 1964 medal for Zimbabwe (Rhodesia) and Azerbaijan's spike in 2016 stand out. East and West Germany's post-2012 decline and Bohemia's 151 medals after 1924 add further intrigue to the Olympic narrative. Since 1836, the US, UK, and Switzerland have maintained consistent participation and achieved noteworthy success. Slovakia's 11 appearances resulted in a modest 11 medals, while Romania's 1936 debut yielded an impressive 52 medals. Similarly, Belarus secured 34 medals despite its shorter Olympic history.

only region who joined in 2016 and able to get 12 medals.		
region	first_year	medal
Refugee Olympic	2016	12

region	count(ce.medal_id)	first_year	lastparticipation	year
USA	1593	1896	201	
France	1287	1900	201	
Italy	1155	1904	201	
Sweden	1106	1900	201	
UK	1068	1896	201	
Germany	961	1896	201	
Canada	816	1904	201	
Hungary	692	1900	201	
Spain	666	1920	201	
Australia	665	1906	201	
Soviet Union	665	1952	200	
Norway	609	1900	201	
Switzerland	559	1900	201	
Russia	553	1912	201	
Romania	432	1924	201	
Brazil	428	1920	201	
Netherlands	399	1900	201	
Argentina	388	1924	201	
Poland	372	1924	201	
West Germany	353	1964	201	
Denmark	342	1900	201	
Bulgaria	325	1924	201	
Egypt	323	1920	201	
Czechoslovakia	316	1912	200	
Japan	316	1928	201	
Mexico	310	1924	201	
Belgium	285	1908	201	
Cuba	270	1928	201	
Austria	253	1900	201	
Turkey	251	1924	201	
Finland	245	1900	201	
East Germany	214	1964	199	
Yugoslavia	209	1924	200	
Nigeria	205	1952	201	
Ukraine	201	1988	201	
Venezuela	197	1977	201	

This regions has lowest medal . These regions started to play after 1956 summer,it has shown in the table with their participation.	
Guyana	14
Indonesia	42
Iraq	43
Kenya	84
Lebanon	40
Morocco	144
Netherlands Antilles	5
Sudan	17
Tunisia	140
Uganda	45
West Indies Federation	5
Zimbabwe (Rhodesia)	3

This 2 region got lowest trophy,i analyzed that they not played anymore after this years.					
region	nam	1904	Su	1906	Su
Newfoundland		1		1	
North Borneo		4	1956	Su	

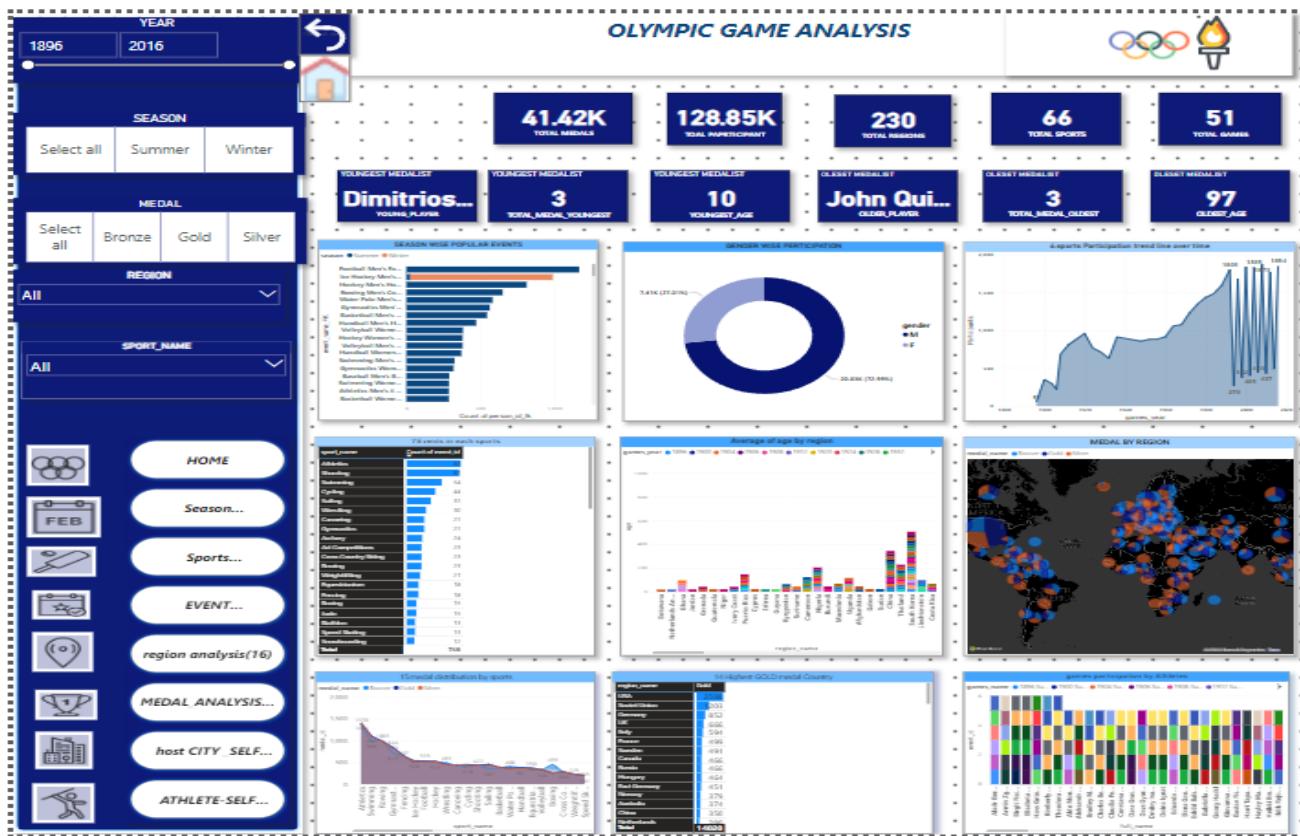
Are there any regions that have experienced significant growth or decline in Olympic participation?

After studying the data table, it's clear that the 16 regions mentioned have stopped participating after their last indicated years. Notably, Czechoslovakia, East Germany, and Yugoslavia achieved considerable total medals, showcasing positive progress. However, there is no information available regarding their participation beyond the specified last years. This implies that these regions either chose to stop participating or experienced a decrease in their involvement in the Olympics.

region	Medal Total	last participated year
North Yemen	7	1988
Saar	7	1958
United Arab Republic	36	1964
Czechoslovakia	316	2000
East Germany	214	1998
South Yemen	1	1988
Yugoslavia	209	2002
Bohemia	151	1936
Australasia	94	1924
Vietnam (pre)	51	1972
West Indies Federation	44	1972
Malaya	5	1960
Zimbabwe (Rhodesia)	3	1960
North Borneo	2	1956
Newfoundland	2	1906
Crete		

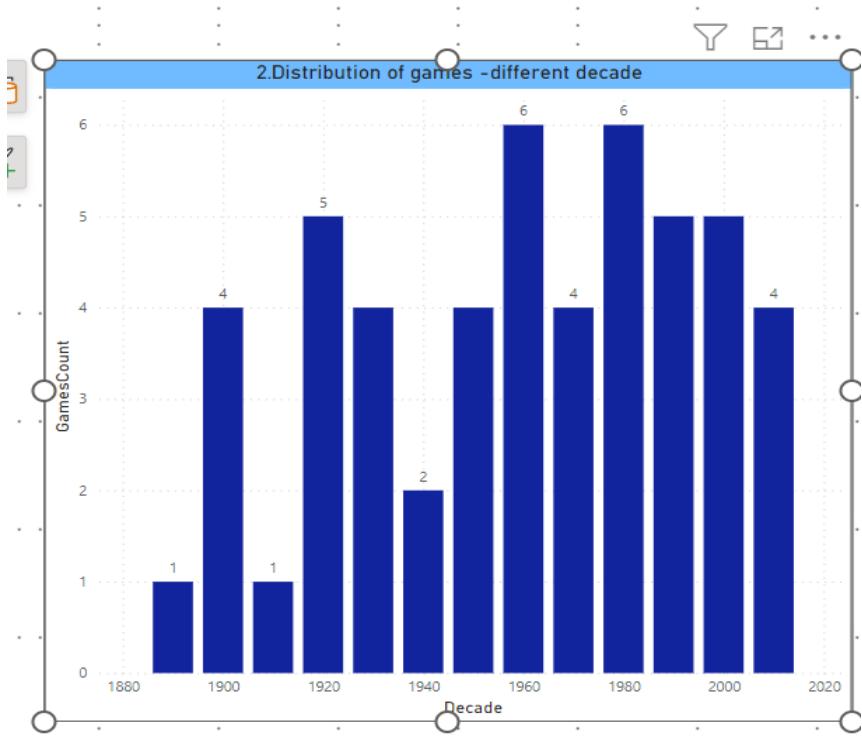
Power BI Problem Statements

DASHBOARD



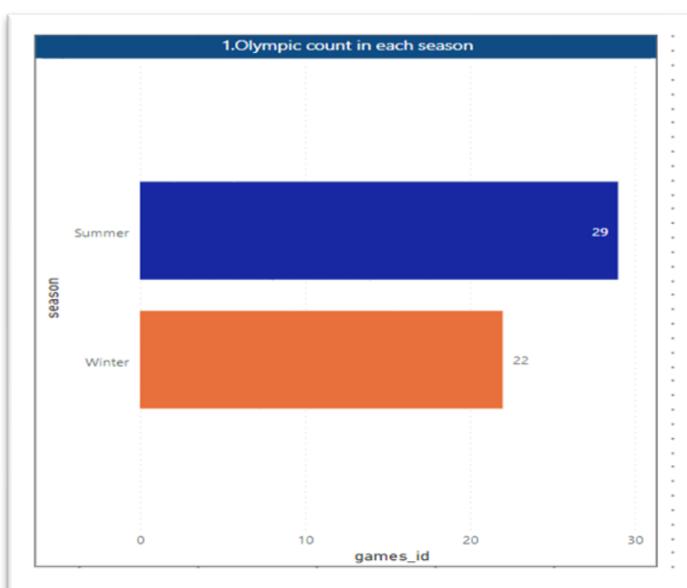
What is the distribution of games across different decades?

The data indicates that the year 1960 witnessed the most prominent occurrence of the Olympic Games, while the 2000s experienced four editions of the Games throughout that decade.



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

29 games in the summer season and 22 games in the winter season.



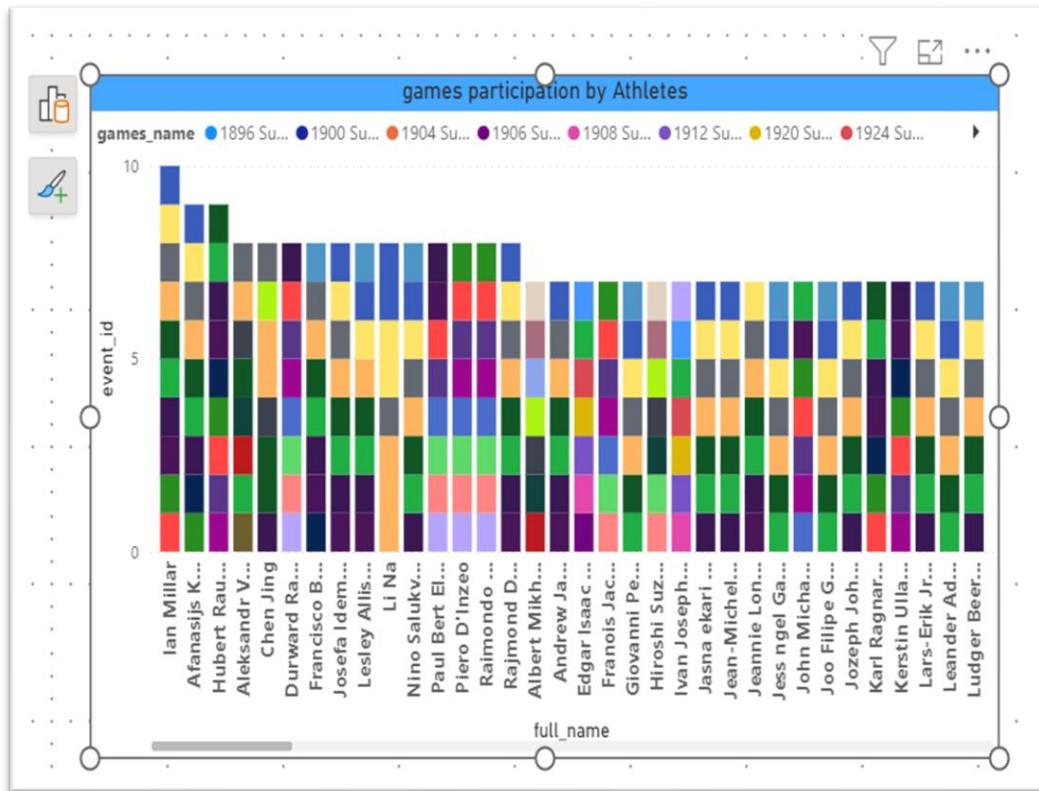
The name of the Athlete who got highest gold medal.

Michael Fred Philips II got the highest gold medal.

full_name	Gold	Total
Michael Fred Phelps, II	23	23
Birgit Fischer-Schmidt	16	16
Hans Ginter Winkler	14	14
Larysa Semenivna Latynina (Dirly-)	13	13
Lyubov Ivanovna Yegorova	12	12
Reiner Klimke	12	12
Viktor An	12	12
Vitaly Venediktovich Shcherbo	12	12
Aleksandr Vladimirovich Popov	11	11
Larisa Yevgenyevna Lazutina (Piltsyna-)	10	10
Raymond Clarence "Ray" Ewry	10	10
Viktor Ivanovych Chukarin	10	10
Aleksandr Aleksandrovich Karelin	9	9
Andrey Ivanovich Lavrov	9	9
Anfisa Anatolyevna Reztsova (Romanova-)	9	9
Borys Anfianovych Shakhlin	9	9
Frederick Carlton "Carl" Lewis	9	9
genes Keleti-Srkny (Klein)	9	9
Mark Andrew Spitz	9	9
Dominique Monami	8	8
Total	14020	14020

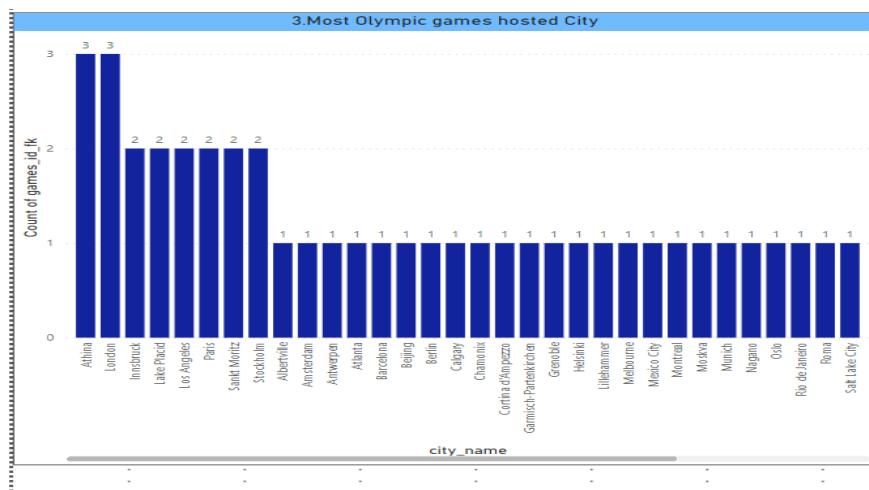
Which athletes took part in specific years?

The data illustrates the participation of athletes in various years, with Ian Millar having the highest number of participations compared to other athletes.



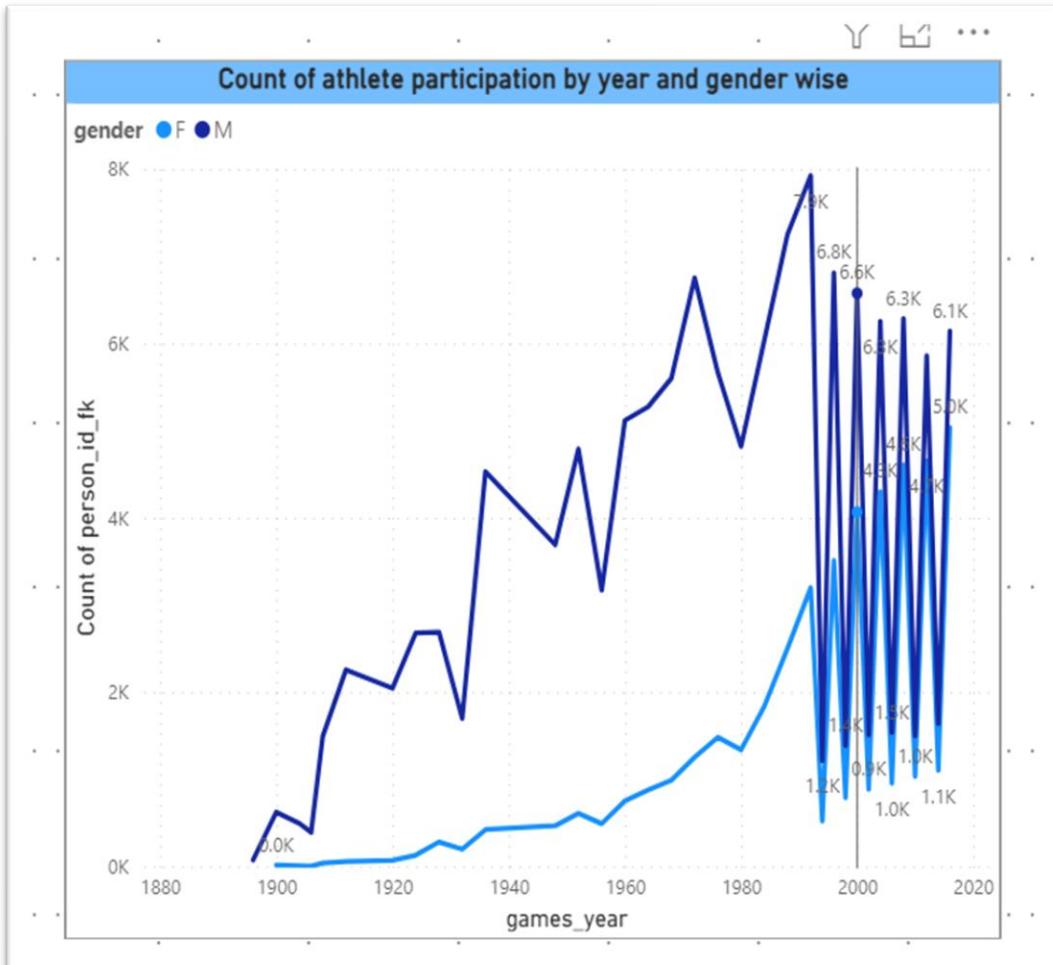
Which cities have hosted the most Olympic Games?

London and Athina hosted the most Olympic games.



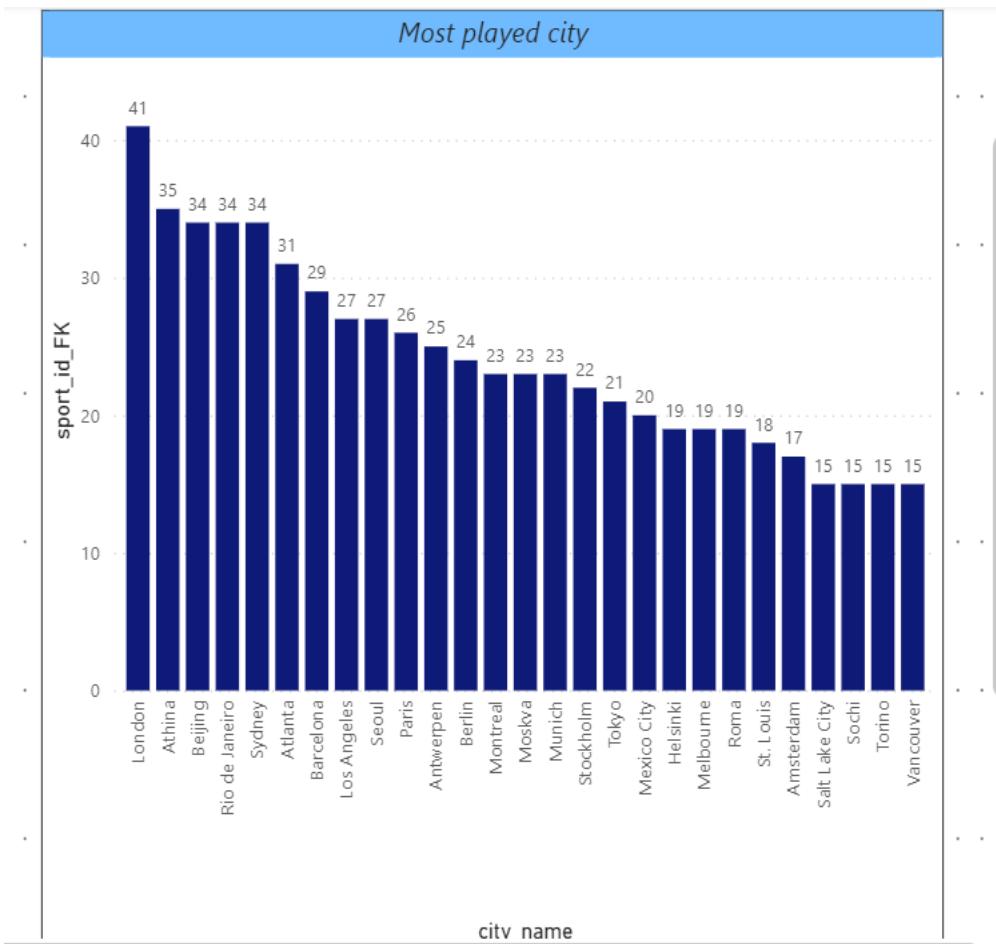
What is the distribution of athletes year-wise and gender-wise?

The data indicates a consistent disparity between the number of male and female athletes, with historically lower participation of females. However, there is a noticeable trend of increasing female participation over time.



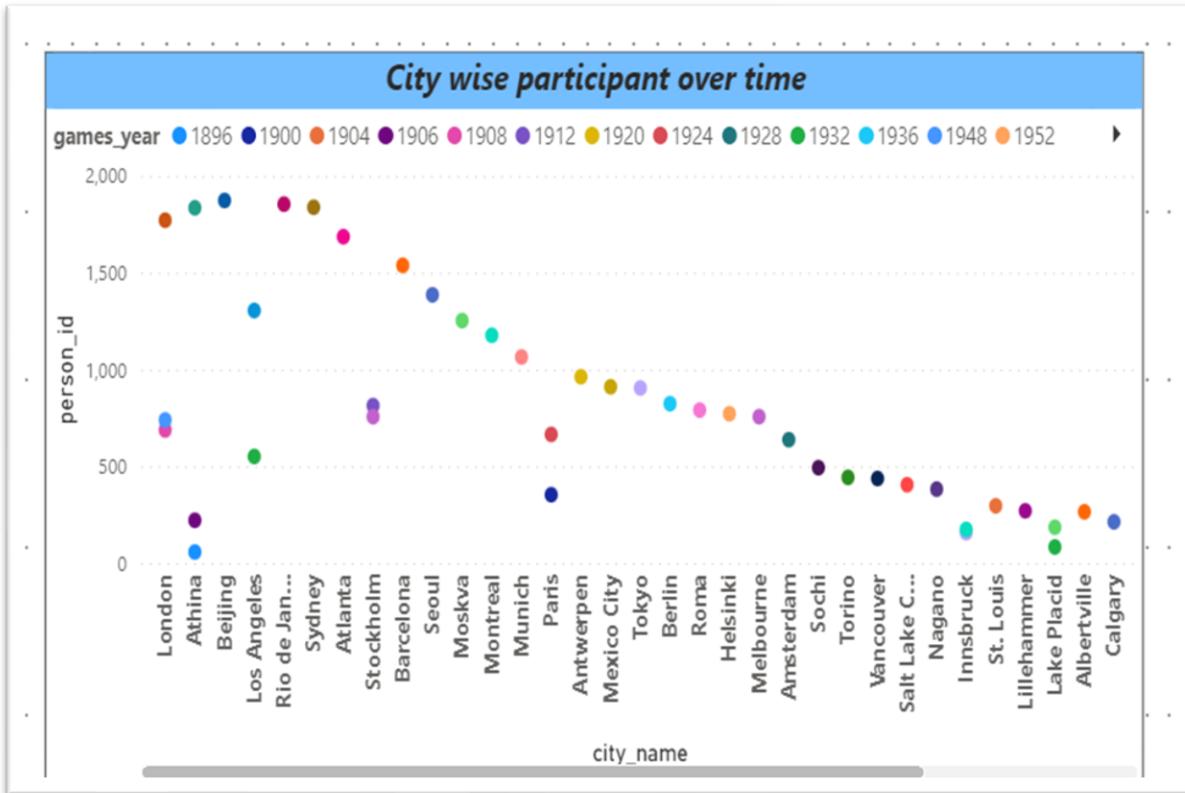
What host city saw the highest number of sports events occurring

London experienced the highest number of occurring sports events



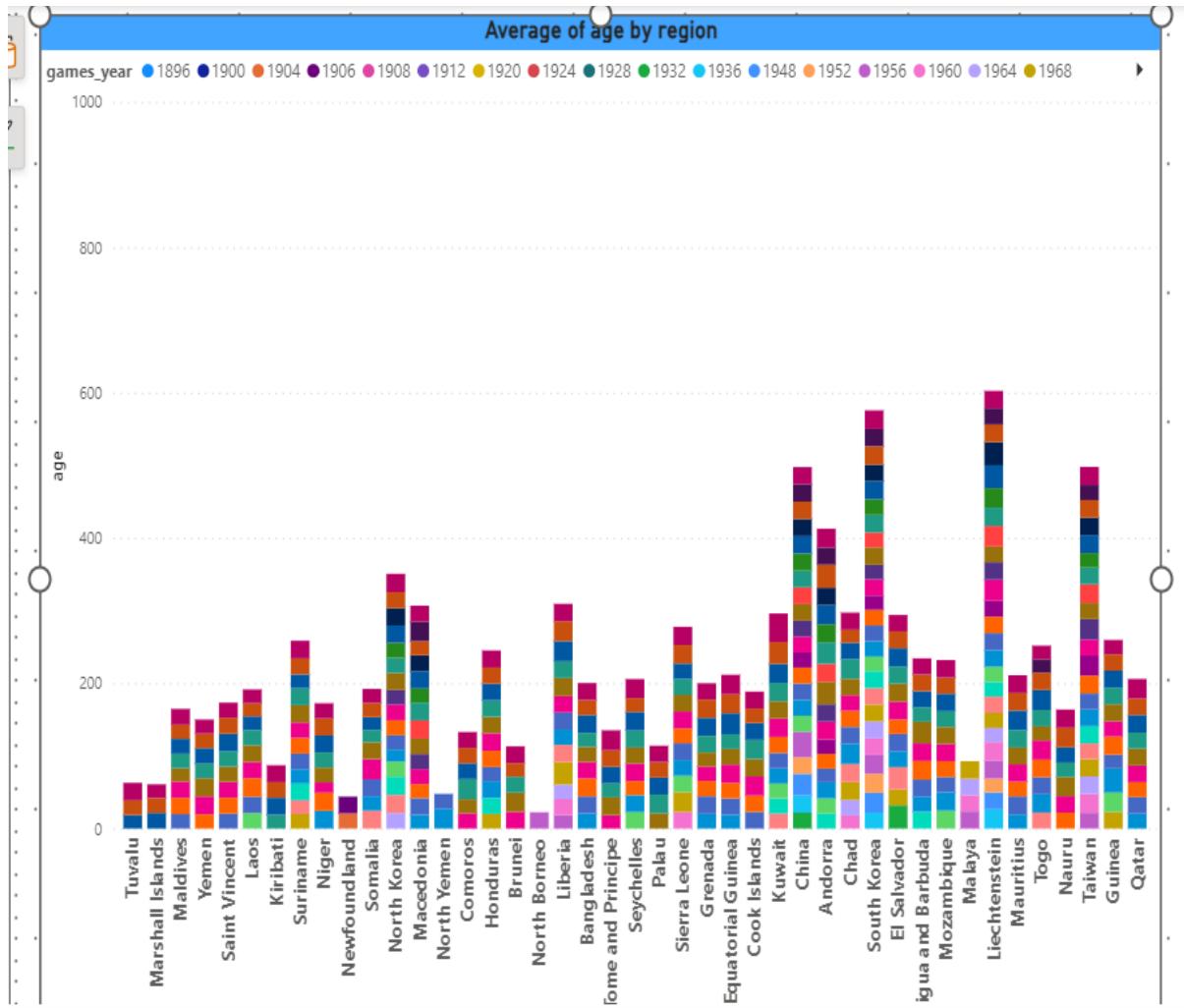
What is the distribution of participation city wise and year wise?

The graph displays the number of participants for each Host-city in every year.



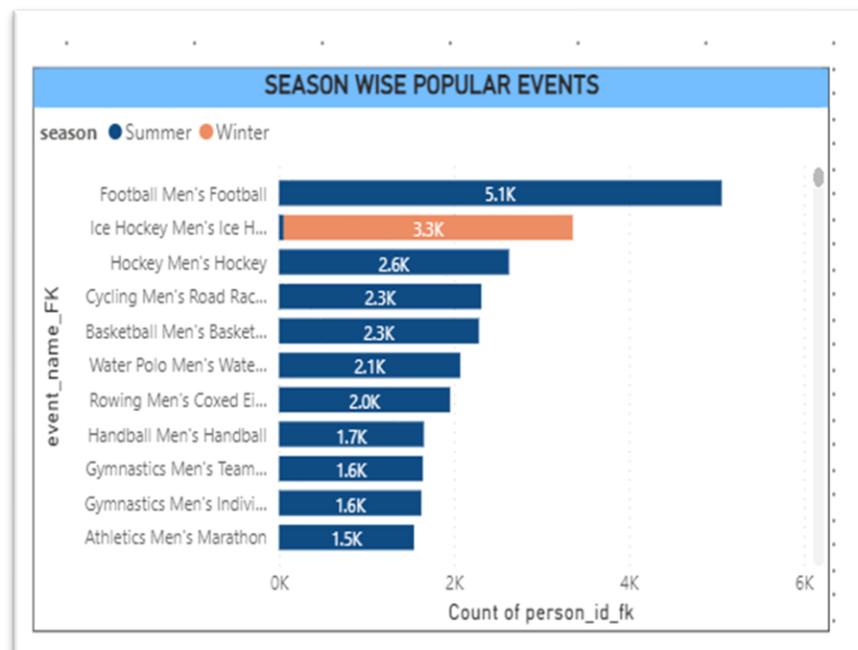
What is the distribution of age by regions year wise?

In the graph, the data portrays the average age of athletes' participation for each region over the years.



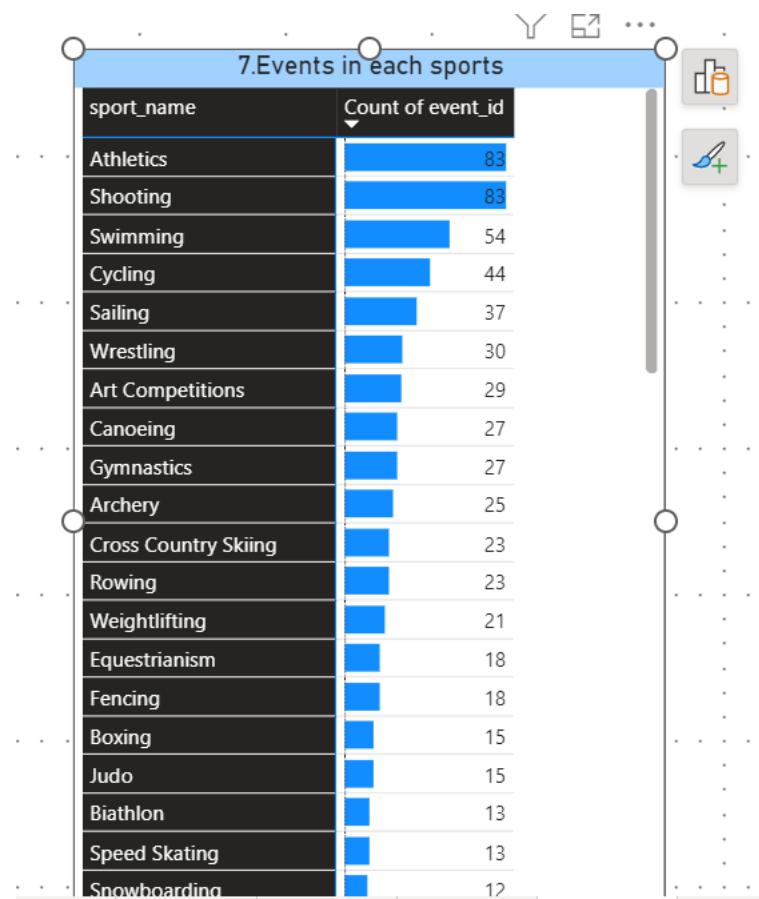
Identify which sports are more popular in Summer Olympics compared to Winter Olympics, and vice versa

According to the data depicted in the graph, Men's Football emerged as the most popular event during the Summer season, while Men's Ice Hockey held the distinction of being the most popular event in the winter.



How many events are there in each sport?

In the graph, the data is presented, athletics has the most events



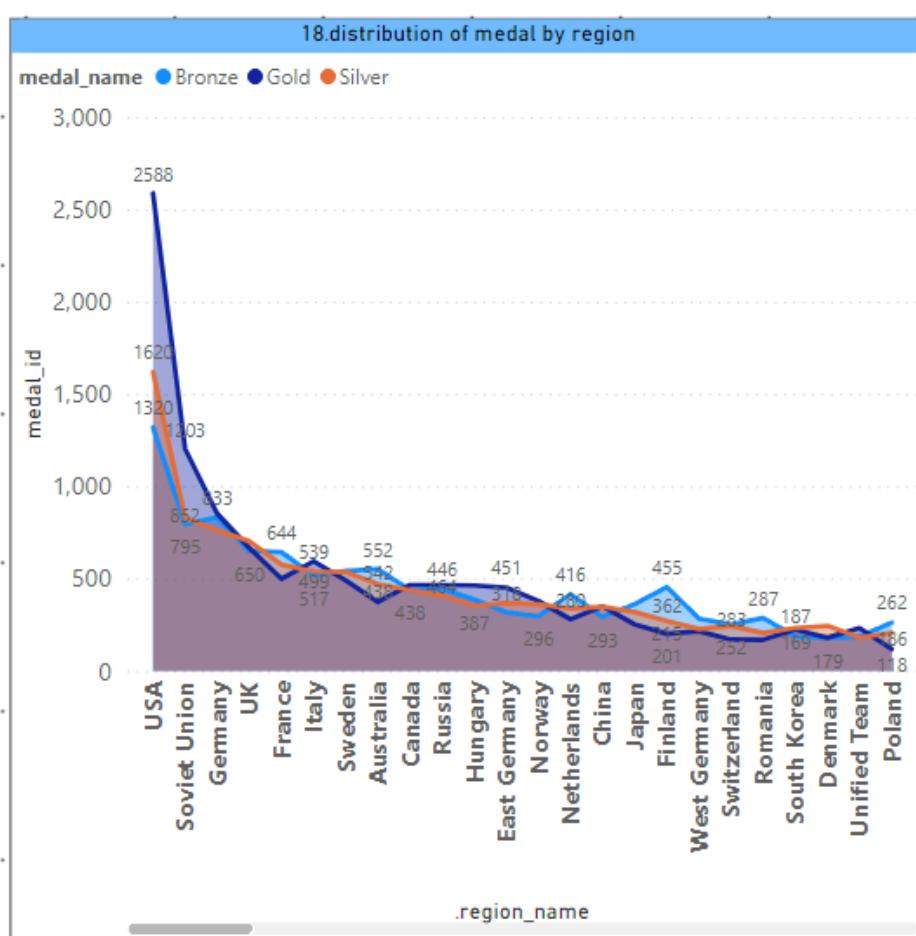
Which countries have the highest number of gold medals?

USA is the highest gold medal won country.

14.Highest GOLD medal Country	
region_name	Gold
USA	2588
Soviet Union	1203
Germany	852
UK	666
Italy	594
France	499
Sweden	491
Canada	466
Russia	466
Hungary	464
East Germany	451
Norway	379
Australia	374
China	350
Netherlands	280
Japan	252
Unified Team	233
South Korea	224
West Germany	215
Finland	201
Denmark	180
Switzerland	170
Romania	169
Cuba	164
Yugoslavia	136
Total	14020

What is the distribution of medals among different regions?

In the graph, the data is presented and the USA got the highest medals and also won 18 times.

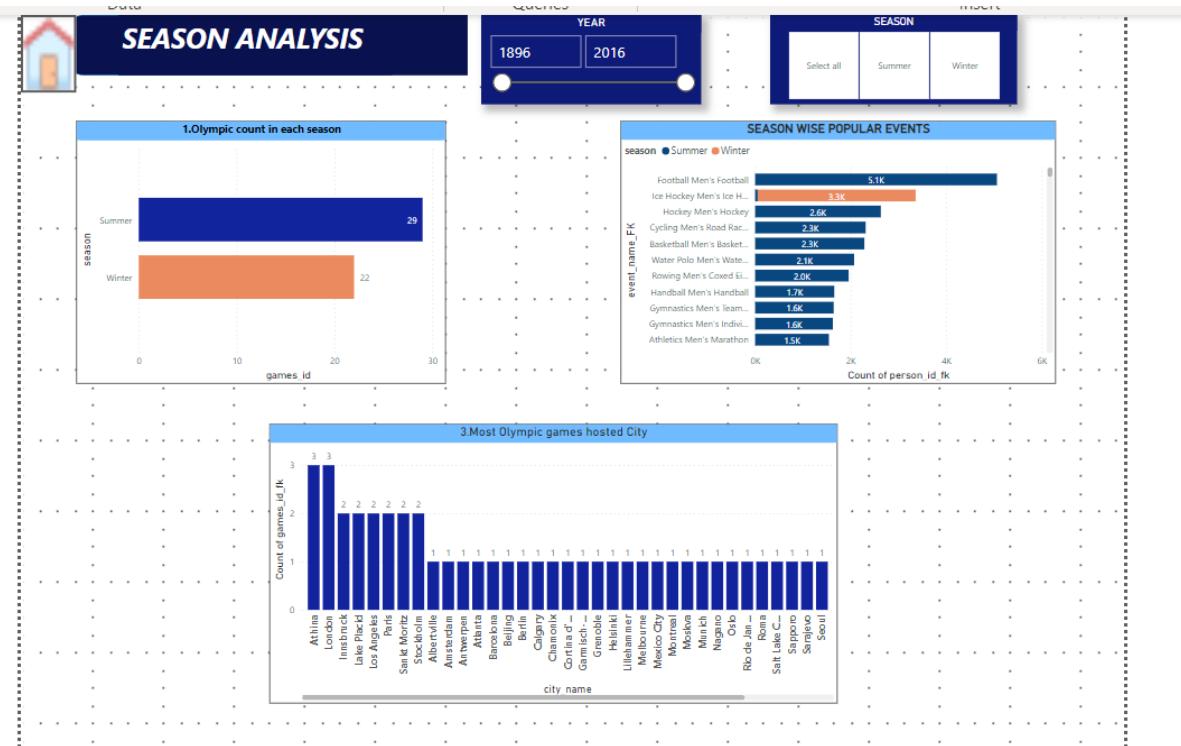


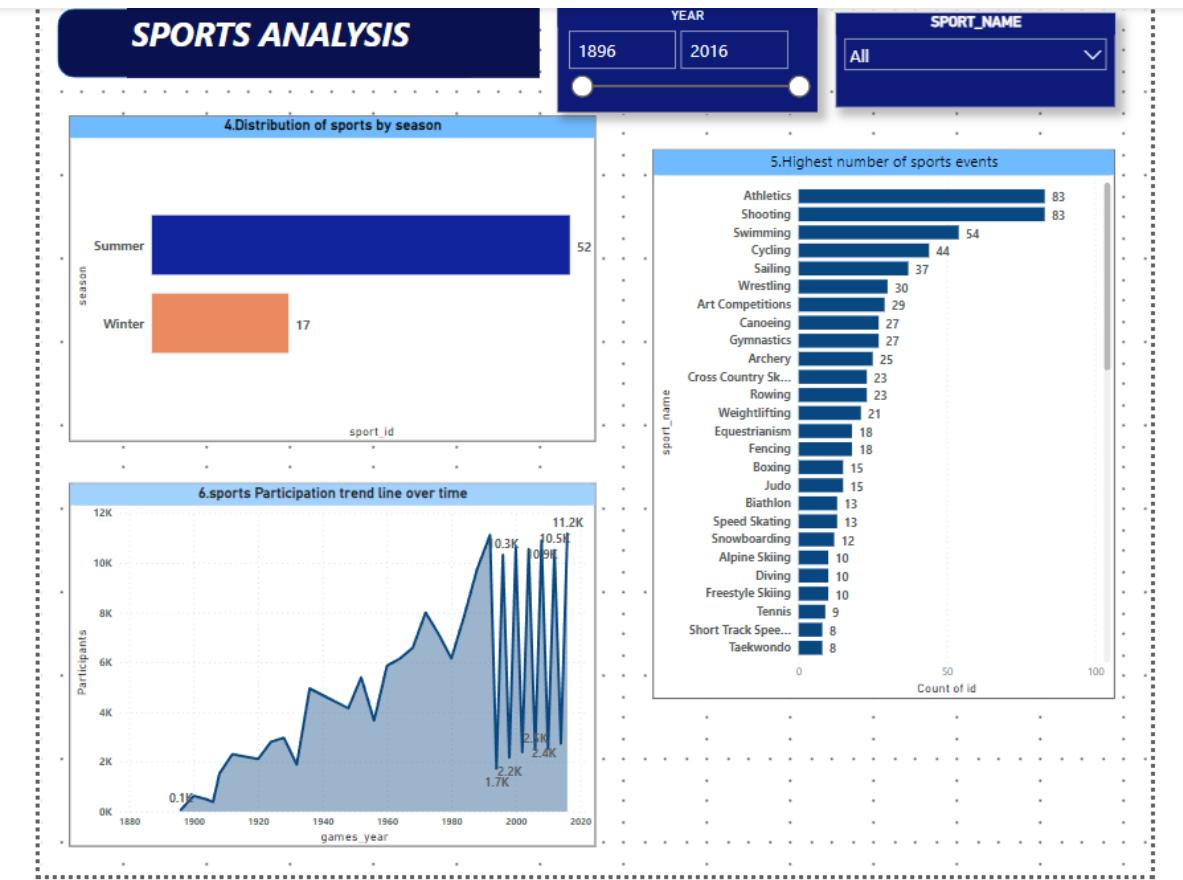
How many regions or NOCs participate in each Olympic Games?

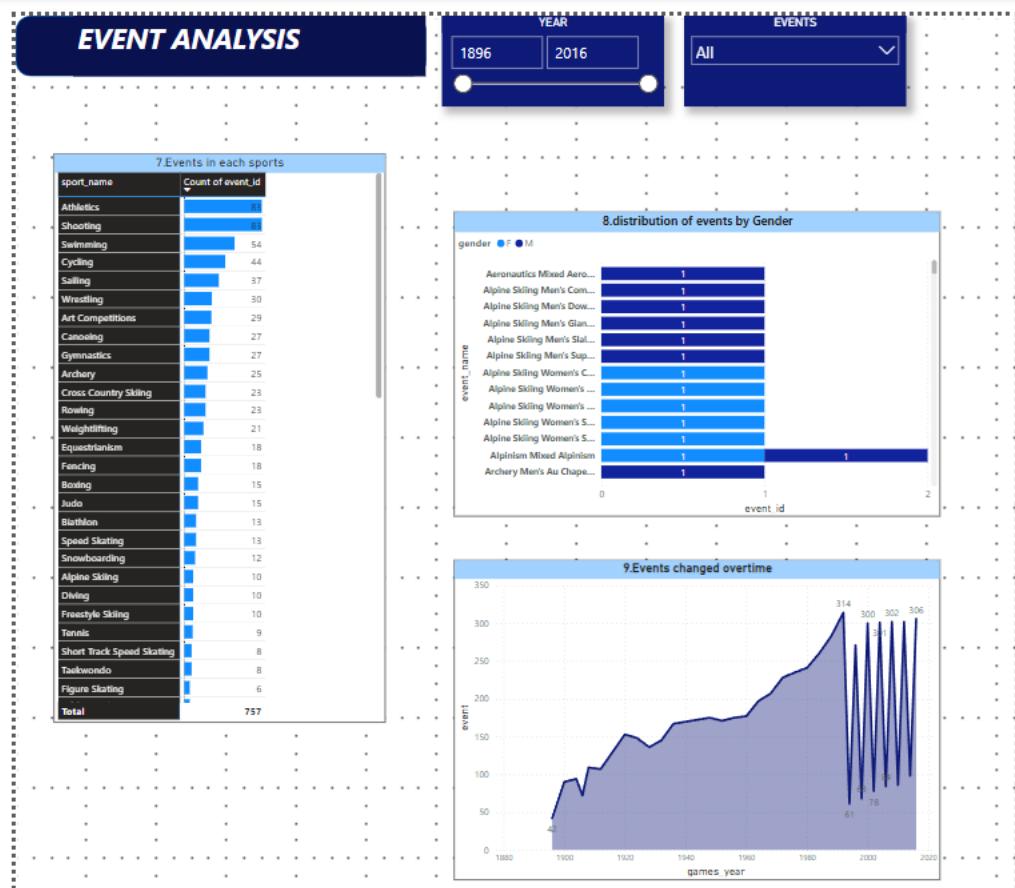
Most of the participants from different regions took part in the 2016 Summer Olympics.

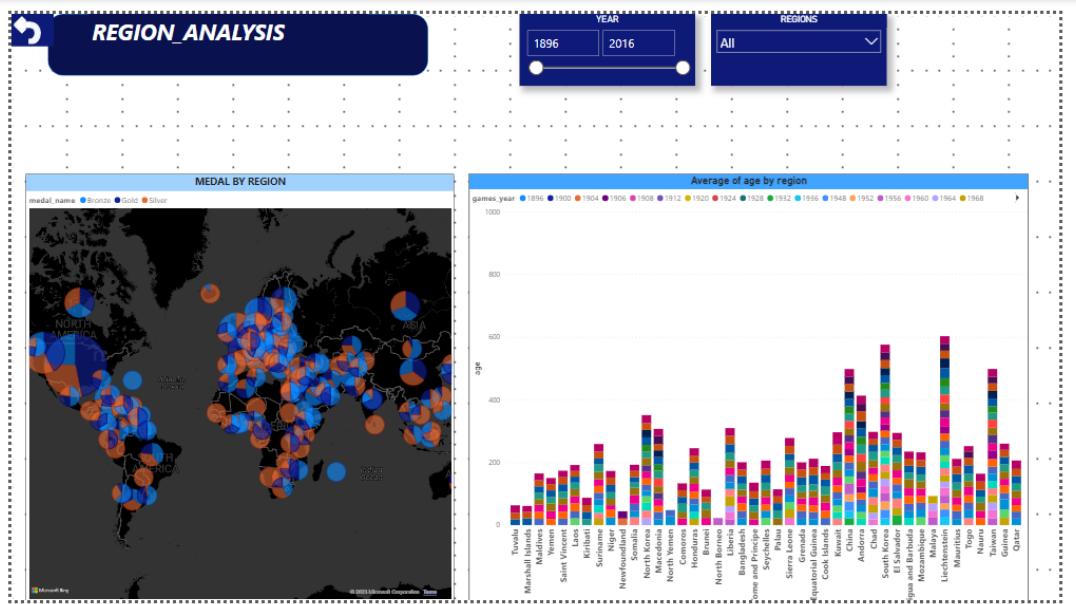
16.Total Participated Region	
games_name	Count of region_id_FK
2016 Summer	213
2014 Winter	91
2012 Summer	215
2010 Winter	85
2008 Summer	212
2006 Winter	87
2004 Summer	210
2002 Winter	85
2000 Summer	209
1998 Winter	79
1996 Summer	206
1994 Winter	74
1992 Winter	79
1992 Summer	188
1988 Winter	70
1988 Summer	181
1984 Winter	58
1984 Summer	150
1980 Winter	41
1980 Summer	92
Total	228

Power BI Report Snapshots









MEDAL ANALYSIS

YEAR
1896 2016

MEDAL

Select all	Gold	Silver
Bronze	NA	

13.Total medal in each Olympic

games_name	Count of medal_id_fk
2016 Summer	13806
2014 Winter	4936
2012 Summer	13103
2010 Winter	4440
2008 Summer	13842
2006 Winter	4450
2004 Summer	13821
2002 Winter	4237
2000 Summer	14347
1998 Winter	3819
1996 Summer	14655
1994 Winter	3528
Total	272672

14.Highest GOLD medal Country

region_name	Gold
USA	2598
Soviet Union	1203
Germany	852
UK	666
Italy	594
France	499
Sweden	491
Canada	486
Russia	486
Hungary	464
East Germany	451
Norway	379
Australia	374
China	350
Netherlands	280
Japan	252
United Team	233
South Korea	224
Total	14020

MEDAL DISTRIBUTION BY GENDER

medal_name: ● Bronze ● Gold ● Silver

gender	Bronze	Gold	Silver
M	9.8K	10.0K	9.7K
F	4.0K	4.0K	3.9K

