**SPORT ANALYSIS**

**INDEX**

|  |  |  |
| --- | --- | --- |
| **S.NO** | **LESSON** | **PAGES** |
| 1 | objective | 3 |
| 2 | MECE | 4-5 |
| 3 | EXCEL | 6-22 |
| 4 | POWER BI | 22-52 |
| 5 | summary | 52-53 |

**Objective:**

The primary goal of this report is to explore and analyse historical Olympic Games data to extract valuable insights around participation, sports, regional representation, and medal distribution. Using Power BI,Excel this dashboard brings together data from multiple dimensions to help identify trends, patterns, and across different editions of the Olympics.

Key Metrics:

* Total Olympic Games held
* Total athletes participated
* Number of countries/regions
* Total medals awarded

Visuals:

* KPI Cards: Total Games, Athletes, Countries, Medals
* Timeline Chart: Olympics over decades
* Slicer: Season (Summer/Winter)
* Narrative Box: Project overview, purpose, and data source

**1. MECE**

**Step1: Problem Statement MECE Breakdown:**

A. Historical Context & Games Overview

* Timeline of Olympic editions (Summer/Winter)
* Frequency and distribution of Games by decade/year
* Host cities and countries over time
* Joint-hosted Games (e.g., 1956 Melbourne & Stockholm)

B. Sports & Events

* Total number of sports across Summer and Winter Olympics
* Breakdown of events by sport and gender
* Trends in number of events per edition
* Introduction and discontinuation of sports/events over time

C. Athletes & Participation

* Athlete demographics (age, gender, height, weight)
* Participation trends (number of athletes per year/region)
* Representation of athletes in different sports
* Athletes competing for multiple countries

D. Medals & Performance

* Medal distribution by country and sport
* Gold, Silver, Bronze trends over time
* Most awarded sports and athletes
* Medal efficiency (medals per athlete or per event)

E. Regional Analysis

* Performance comparison by NOC/region
* Medal tally by continent or country groupings
* Participation vs. medal ratios per country
* Growth of participation from underrepresented regions

F. Organizational Insights

* Evolution in the structure of the Games
* Impact of global events (wars, pandemics) on Games continuity
* Recommendations for inclusive representation and sport diversification
* Optimal host city patterns and learnings from past editions

**Step2: Data set analysis:**

| **Category** | **Primary Tables Used** |
| --- | --- |
| Games Timeline & Hosting | Games, Games City, City |
| Sports & Events | Sport, Event |
| Athletes & Demographics | Person, Games Competitor, Person Region |
| Medal Distribution | Competitor Event, Medal |
| Regional Representation | NOC Region, Person Region, Medal |
| Participation & Performance | Games Competitor, Competitor Event, Person |
|  | |

**Step3: End Goals (Structured)**

* Descriptive Insight: What has happened over the Olympic timeline?
* Diagnostic Analysis: Why certain sports/countries excel or underperform?
* Predictive Opportunities: Which regions/sports could grow in future editions?

**2.EDA (EXCEL)**

**1)Are there any trends or patterns in the frequency of hosting Olympic?**

Sol:

SELECT g.games\_year,

COUNT(DISTINCT gc.city\_id) AS number\_of\_host\_cities

FROM games g

JOIN games\_city gc ON g.id = gc.games\_id

GROUP BY g.games\_year

ORDER BY g.games\_year;

|  |  |
| --- | --- |
| games year | number\_of\_host\_cities |
| 1896 | 1 |
| 1900 | 1 |
| 1904 | 1 |
| 1906 | 1 |
| 1908 | 1 |
| 1912 | 1 |
| 1920 | 1 |
| 1924 | 2 |
| 1928 | 2 |
| 1932 | 2 |
| 1936 | 2 |
| 1948 | 2 |
| 1952 | 2 |

**2)How has the duration of Olympic Games changed over time?**

**sol:**

SELECT games\_year, COUNT(\*) AS games\_held

FROM games

GROUP BY games\_year

ORDER BY games\_year;

Output:

|  |  |
| --- | --- |
| games\_year | games\_held |
| 1896 | 1 |
| 1900 | 1 |
| 1904 | 1 |
| 1906 | 1 |
| 1908 | 1 |
| 1912 | 1 |
| 1920 | 1 |
| 1924 | 2 |
| 1928 | 2 |
| 1932 | 2 |
| 1936 | 2 |
| 1948 | 2 |
| 1952 | 2 |
| 1956 | 2 |

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

**sol:**

WITH Medalists AS ( SELECT DISTINCT gc.games\_id, gc.person\_id

FROM medal m

JOIN competitor\_event ce ON m.id = ce.medal\_id

JOIN games\_competitor gc ON ce.competitor\_id = gc.id

WHERE m.id IS NOT NULL),

MedalistsPerGame AS

( SELECT g.games\_name, c.city\_name, g.games\_year,

COUNT(DISTINCT m.person\_id) AS unique\_medalists FROM Medalists m

JOIN games g ON m.games\_id = g.id

JOIN games\_city gct ON g.id = gct.games\_id

JOIN city c ON gct.city\_id = c.id

GROUP BY g.games\_name, c.city\_name, g.games\_year)

SELECT \*FROM MedalistsPerGame

ORDER BY unique\_medalists DESC;

|  |  |  |
| --- | --- | --- |
| **city\_name** | **games\_year** | **unique\_medalist** |
| Rio de Janeiro | 2016 | 11176 |
| Beijing | 2008 | 10895 |
| Sydney | 2000 | 10641 |
| Athina | 2004 | 10554 |
| London | 2012 | 10512 |
| Atlanta | 1996 | 10319 |
| Barcelona | 1992 | 9327 |
| Seoul | 1988 | 8341 |

Output:

**4)How has the popularity of certain sports changed over the years?**

**sol:**

WITH SportParticipation AS ( SELECT s.sport\_name,

g.games\_year,

FLOOR(g.games\_year / 10) \* 10 AS decade,

gc.person\_id FROM games\_competitor gc

JOIN competitor\_event ce ON gc.id = ce.competitor\_id

JOIN event e ON ce.event\_id = e.id

JOIN sport s ON e.sport\_id = s.id

JOIN games g ON gc.games\_id = g.id),

SportPopularity AS ( SELECT sport\_name,

decade,

COUNT(DISTINCT person\_id) AS participants

FROM SportParticipation

GROUP BY sport\_name, decade)

SELECT \*

FROM SportPopularity

ORDER BY sport\_name, decade;

|  |  |  |
| --- | --- | --- |
| **sport\_name** | **decade** | **participants** |
| Aeronautics | 1930 | 1 |
| Alpine Skiing | 1930 | 102 |
| Alpine Skiing | 1940 | 164 |
| Alpine Skiing | 1950 | 295 |
| Alpine Skiing | 1960 | 381 |
| Alpine Skiing | 1970 | 287 |
| Alpine Skiing | 1980 | 528 |

**5)How has the popularity of certain sports changed over the years?**

**sol:**

WITH SportParticipation AS ( SELECT s.sport\_name,

g.games\_year, FLOOR(g.games\_year / 10) \* 10 AS decade,

gc.person\_id FROM games\_competitor gc

JOIN competitor\_event ce ON gc.id = ce.competitor\_id

JOIN event e ON ce.event\_id = e.id

JOIN sport s ON e.sport\_id = s.id

JOIN games g ON gc.games\_id = g.id),

SportPopularity AS ( SELECT sport\_name, decade,

COUNT(DISTINCT person\_id) AS participants

FROM SportParticipation

GROUP BY sport\_name, decade)

SELECT \*FROM SportPopularity

ORDER BY sport\_name, decade;

Output:

|  |  |  |
| --- | --- | --- |
| **sport\_name** | **decade** | **participant** |
| Aeronautics | 1930 | 1 |
| Alpine Skiing | 1930 | 102 |
| Alpine Skiing | 1940 | 164 |
| Alpine Skiing | 1950 | 295 |
| Alpine Skiing | 1960 | 381 |
| Alpine Skiing | 1970 | 287 |
| Alpine Skiing | 1980 | 528 |
| Alpine Skiing | 1990 | 558 |
| Alpine Skiing | 2000 | 452 |

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

**sol:**

WITH SportRegion AS ( SELECT DISTINCT s.sport\_name,

nr.region\_name FROM competitor\_event ce

JOIN event e ON ce.event\_id = e.id

JOIN sport s ON e.sport\_id = s.id

JOIN games\_competitor gc ON ce.competitor\_id = gc.id

JOIN person p ON gc.person\_id = p.id

JOIN person\_region pr ON p.id = pr.person\_id

JOIN noc\_region nr ON pr.region\_id = nr.id),RegionCounts AS ( SELECT sport\_name, COUNT(DISTINCT region\_name) AS region\_count

FROM SportRegion GROUP BY sport\_name)

SELECT \*FROM RegionCounts

WHERE region\_count < 5

ORDER BY region\_count ASC;

|  |  |
| --- | --- |
| **sport\_name** | **region\_count** |
| Aeronautics | 1 |
| Basque Pelota | 1 |
| Croquet | 1 |
| Racquets | 1 |
| Roque | 1 |
| Cricket | 2 |
| Jeu De Paume | 2 |
| Motorboating | 2 |
| Lacrosse | 3 |
| Alpinism | 4 |

Output:

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

**sol:**

withevent\_gender\_counts AS ( SELECT gender,

COUNT(DISTINCT id) AS num\_events

FROM person GROUP BY gender)

SELECT \*

FROM event\_gender\_counts

ORDER BY gender, num\_events DESC**;**

**Output:**

|  |  |
| --- | --- |
| **gender** | **event\_count** |
| F | 33629 |
| M | 95225 |

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

**sol:**

SELECT e.event\_name, s.sport\_name, MIN(g.games\_year) AS first\_appearance

FROM event e

JOIN sport s ON e.sport\_id = s.id

JOIN competitor\_event ce ON e.id = ce.event\_id

JOIN games\_competitor gc ON ce.competitor\_id = gc.id

JOIN games g ON gc.games\_id = g.id

GROUP BY e.event\_name, s.sport\_name

ORDER BY first\_appearance DESC;

|  |  |  |
| --- | --- | --- |
| **event name** | **sport name** | **first appearance** |
| Rugby Sevens Women's Rugby Sevens | Rugby Sevens | 2016 |
| Rugby Sevens Men's Rugby Sevens | Rugby Sevens | 2016 |
| Sailing Women's Skiff | Sailing | 2016 |
| Wrestling Women's Light-Heavyweight, Freestyle | Wrestling | 2016 |
| Wrestling Women's Featherweight, Freestyle | Wrestling | 2016 |
| Biathlon Mixed 2 x 6 kilometres and 2 x 7.5 kilometres Relay | Biathlon | 2014 |
| Figure Skating Mixed Team | Figure Skating | 2014 |

**Output:**

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

**sol:**

WITH event\_last\_seen AS ( SELECT

e.event\_name, s.sport\_name,

MAX(g.games\_year) AS last\_year

FROM event e

JOIN sport s ON e.sport\_id = s.id

JOIN competitor\_event ce ON e.id = ce.event\_id

JOIN games\_competitor gc ON ce.competitor\_id = gc.id

JOIN games g ON gc.games\_id = g.id

GROUP BY e.event\_name, s.sport\_name)

SELECT \* FROM event\_last\_seen

WHERE last\_year < 2020

ORDER BY last\_year DESC;

|  |  |  |
| --- | --- | --- |
| **event\_name** | **sport\_name** | **last\_year** |
| Archery Women's Individual | Archery | 2016 |
| Archery Men's Individual | Archery | 2016 |
| Archery Men's Team | Archery | 2016 |
| Archery Women's Team | Archery | 2016 |
| Athletics Women's 100 metres | Athletics | 2016 |
| Athletics Women's 4 x 100 metres Relay | Athletics | 2016 |

**Output:**

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

**sol:**

WITH height\_weight\_trend AS ( SELECT g.games\_year,

AVG(p.height) AS avg\_height,

AVG(p.weight) AS avg\_weight

FROM person p

JOIN games\_competitor gc ON p.id = gc.person\_id

JOIN games g ON gc.games\_id = g.id

GROUP BY g.games\_year)

SELECT

\* FROM height\_weight\_trend

ORDER BY games\_year;

|  |  |  |
| --- | --- | --- |
| **games\_year** | **avg\_height** | **avg\_weights** |
| 1896 | 29.0556 | 11.5417 |
| 1900 | 14.0409 | 4.3947 |
| 1904 | 29.0585 | 9.3609 |
| 1906 | 39.8875 | 14.7366 |
| 1908 | 26.4526 | 12.8149 |
| 1912 | 27.4944 | 10.5416 |
| 1920 | 35.4965 | 10.7258 |
| 1924 | 38.2198 | 11.9502 |
| 1928 | 38.851 | 12.4053 |

Output:

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

**sol:**

WITH SportMedals AS ( SELECT s.sport\_name,

nr.region\_name, COUNT(\*) AS medal\_count

FROM medal m

JOIN competitor\_event ce ON m.id = ce.medal\_id

JOIN event e ON ce.event\_id = e.id

JOIN sport s ON e.sport\_id = s.id

JOIN games\_competitor gc ON ce.competitor\_id = gc.id

JOIN person\_region pr ON gc.person\_id = pr.person\_id

JOIN noc\_region nr ON pr.region\_id = nr.id

WHERE m.id IS NOT NULL

GROUP BY s.sport\_name, nr.region\_name),

Ranked AS ( SELECT \*, RANK() OVER (PARTITION BY sport\_name ORDER BY medal\_count DESC) AS rnk FROM SportMedals)

SELECT \*FROM Ranked

|  |  |  |  |
| --- | --- | --- | --- |
| **sport\_name** | **region\_name** | **medal\_count** | **rnk** |
| Aeronautics | Switzerland | 1 | 1 |
| Alpine Skiing | Austria | 561 | 1 |
| Alpine Skiing | Switzerland | 521 | 2 |
| Alpine Skiing | USA | 518 | 3 |
| Alpinism | UK | 11 | 1 |
| Alpinism | Switzerland | 2 | 2 |
| Alpinism | Germany | 2 | 2 |
| Archery | USA | 167 | 1 |

**Output:**

**12)What factors contribute to the success or performance of participants from different countries?**

**sol:** WITH country\_stats AS ( SELECT pr.region\_id,nc.region\_name,

AVG(p.height) AS avg\_height,

AVG(p.weight) AS avg\_weight,

COUNT(DISTINCT p.id) AS total\_participants

FROM person p

JOIN person\_region pr

ON p.id = pr.person\_id

join noc\_region nc on pr.region\_id=nc.id

GROUP BY pr.region\_id)

SELECT region\_name,total\_participant

FROM country\_stats

ORDER BY total\_participants DESC;

|  |  |
| --- | --- |
| **Region\_name** | **total\_participants** |
| USA | 9225 |
| UK | 5783 |
| France | 5198 |
| Germany | 4763 |
| Canada | 4657 |
| Italy | 4654 |
| Japan | 3965 |
| Sweden | 3781 |
| Australia | 3747 |

**Output:**

**13)Are there any countries that consistently perform well in multiple Olympic editions?**

**sol:**

WITH medals\_by\_country\_year AS ( SELECT pr.region\_id,nc.region\_name,

g.games\_year, COUNT(m.id) AS total\_medals

FROM medal m

JOIN competitor\_event ce ON m.id = ce.medal\_id

JOIN games\_competitor gc ON ce.competitor\_id = gc.id

JOIN games g ON gc.games\_id = g.id

JOIN person\_region pr ON gc.person\_id = pr.person\_id

join noc\_region nc on pr.region\_id=nc.id

GROUP BY pr.region\_id, g.games\_year)

SELECT region\_name, COUNT(games\_year) AS editions\_with\_medals,

SUM(total\_medals) AS total\_medals

FROM medals\_by\_country\_year

GROUP BY region\_id

ORDER BY editions\_with\_medals DESC, total\_medals DESC;

Output:

|  |  |  |
| --- | --- | --- |
| **region\_name** | **edition\_with\_medal** | **total\_medal** |
| USA | 35 | 18127 |
| UK | 35 | 11536 |
| Australia | 35 | 7584 |
| Switzerland | 35 | 5832 |
| Greece | 35 | 2222 |
| France | 34 | 11475 |

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

sol:

WITH medals\_per\_sport\_region AS ( SELECT s.sport\_name,

pr.region\_id, nr.region\_name,

COUNT(DISTINCT ce.event\_id) AS medalists

FROM medal m

JOIN competitor\_event ce ON m.id = ce.medal\_id

JOIN event e ON ce.event\_id = e.id

JOIN sport s ON e.sport\_id = s.id

JOIN games\_competitor gc ON ce.competitor\_id = gc.id

JOIN person\_region pr ON gc.person\_id = pr.person\_id

JOIN noc\_region nr ON pr.region\_id = nr.id

GROUP BY s.sport\_name, pr.region\_id, nr.region\_name

)

SELECT \*

FROM medals\_per\_sport\_region

ORDER BY medalists DESC;

Output:

|  |  |  |  |
| --- | --- | --- | --- |
| sport\_name | region\_id | region\_name | medalists |
| Athletics | 217 | USA | 78 |
| Shooting | 70 | France | 73 |
| Athletics | 75 | UK | 72 |
| Athletics | 194 | Sweden | 71 |
| Shooting | 194 | Sweden | 69 |
| Athletics | 80 | Germany | 68 |
| Athletics | 37 | Canada | 67 |

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

sol:

WITH participation\_vs\_medals AS ( SELECT pr.region\_id,

nc.region\_name,

COUNT(DISTINCT gc.person\_id) AS participants,

COUNT(ce.medal\_id) AS medals\_won

FROM competitor\_event ce

JOIN games\_competitor gc ON ce.competitor\_id = gc.id

JOIN person\_region pr ON gc.person\_id = pr.person\_id

join noc\_region nc on nc.id=pr.region\_id

WHERE ce.medal\_id IS NOT NULL

GROUP BY pr.region\_id)

SELECT \*, ROUND(1.0 \* medals\_won / participants, 2) AS medal\_efficiency

FROM participation\_vs\_medals

WHERE participants > 0

ORDER BY medal\_efficiency DESC;

Output:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **region\_id** | **region\_name** | **participants** | **medal\_won** | **medal\_efficiency** |
| 1 | Afghanistan | 62 | 78 | 1.26 |
| 2 | Netherlands Antilles | 54 | 91 | 1.69 |
| 3 | Albania | 45 | 78 | 1.73 |
| 4 | Algeria | 368 | 555 | 1.51 |
| 5 | Andorra | 61 | 169 | 2.77 |
| 6 | Angola | 163 | 267 | 1.64 |

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

sol:

WITH participation\_by\_region\_year AS ( SELECT pr.region\_id,

g.games\_year,

COUNT(DISTINCT gc.person\_id) AS participants

FROM competitor\_event ce

JOIN games\_competitor gc ON ce.competitor\_id = gc.id

JOIN games g ON gc.games\_id = g.id

JOIN person\_region pr ON gc.person\_id = pr.person\_id

GROUP BY pr.region\_id, g.games\_year)

SELECT \*

FROM participation\_by\_region\_year

ORDER BY region\_id, games\_year;

|  |  |  |
| --- | --- | --- |
| **region\_id** | **games\_year** | **participants** |
| 1 | 1936 | 14 |
| 1 | 1960 | 12 |
| 1 | 1964 | 7 |
| 1 | 1968 | 5 |
| 1 | 1972 | 4 |
| 1 | 1980 | 11 |
| 1 | 1996 | 2 |

Output:

**17)How do cultural or geographical factors influence the performance of regions in specific sports?**

**sol:**

WITH regional\_medals AS ( SELECT pr.region\_id,

nc.region\_name, s.sport\_name,

COUNT(m.id) AS total\_medals

FROM medal m

JOIN competitor\_event ce ON m.id = ce.medal\_id

JOIN games\_competitor gc ON ce.competitor\_id = gc.id

JOIN person\_region pr ON gc.person\_id = pr.person\_id

JOIN noc\_region nc ON nc.id = pr.region\_id

JOIN event e ON ce.event\_id = e.id

JOIN sport s ON e.sport\_id = s.id

GROUP BY pr.region\_id, nc.region\_name, s.sport\_name)

SELECT region\_name,sport\_name,total\_medal

FROM regional\_medals

ORDER BY total\_medals DESC;

|  |  |  |
| --- | --- | --- |
| **region\_name** | **sport\_name** | **total\_medal** |
| USA | Athletics | 3104 |
| UK | Athletics | 2196 |
| USA | Gymnastics | 1830 |
| USA | Swimming | 1579 |
| France | Athletics | 1464 |

Output:

**18)Are there any regions that have had a notable impact on the overall medal tally?**

**sol:**

WITH region\_medal\_summary AS ( SELECT pr.region\_id,region\_name,

COUNT(m.id) AS total\_medals

FROM medal m

JOIN competitor\_event ce ON m.id = ce.medal\_id

JOIN games\_competitor gc ON ce.competitor\_id = gc.id

JOIN person\_region pr ON gc.person\_id = pr.person\_id

JOIN noc\_region nc ON nc.id = pr.region\_id

GROUP BY pr.region\_id)

SELECT region\_name,total\_medals

FROM region\_medal\_summary

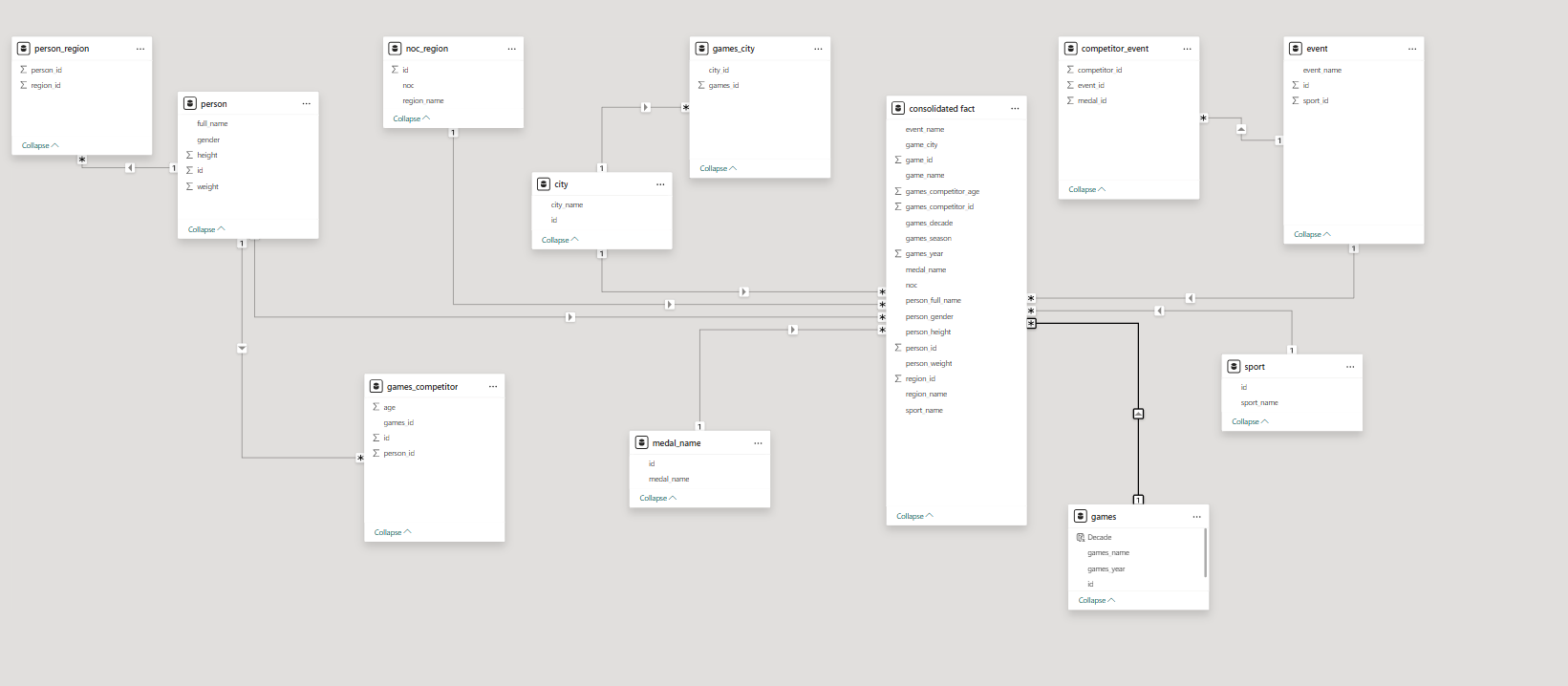
ORDER BY total\_medals DESC;

Output:

|  |  |
| --- | --- |
| **region\_name** | **total\_medal** |
| USA | 18127 |
| UK | 11536 |
| France | 11475 |
| Germany | 10496 |
| Italy | 10355 |

**2.Power bi**

**ER DIAGRAM**

****

Created a relationship to every tables one to many connections in power bi.

**Author analysis:**

In the Sports Analysis, the tables used include Sport, Event, City, Games, Games City, NOC Region, Person, Person Region, Games Competitor, Medal, and Competitor Event.

**Publication:**

The visuals provide insights into how many participants represented each country and city, how participation fluctuated year by year, and how some athletes competed for different countries. They also show the distribution of medals—gold, silver, bronze—and highlight that some participants did not win any medals.

**Games Analysis:**

explores the historical timeline of the Olympic Games, distinguishing between Summer and Winter editions since 1896. It highlights trends in the frequency of events, locations of host cities, and notable changes in hosting patterns, such as multi-city editions. This analysis helps uncover the global expansion and organizational evolution of the Olympics over time.

**Sports and Events :**

examines the evolution of sports and events in the Olympics, including additions, removals, and shifts in popularity. It highlights gender-based classifications—Men's, Women's, and Mixed events—and tracks participation trends across different disciplines. This analysis reveals how the Olympic program has adapted to global sporting interests over time.

**Objective:** To examine the variety and evolution of sports featured in the Olympic Games by identifying trends in sport inclusion across editions, participation rates, and gender distribution. The goal is to uncover shifts in global interest and guide recommendations for future sport selection. In POWER BI overview should given

**Analysis scope:** This analysis explores all sports featured in the Olympic Games across different editions. It focuses on the number of sports, their seasonal classification (Summer/Winter), gender-based event availability, and trends in sport additions or removals over time.

**Goal:** The goal is to identify key trends in the inclusion, popularity, and gender balance of Olympic sports over time. This insight will help support strategic decisions on sport selection, event design, and equitable representation in future Olympic Games.

A group of pictures of graphs

AI-generated content may be incorrect.

**1)How many Olympic games have been held in each Season (summer vs winter ?**

**A blue circle with a number of numbers

AI-generated content may be incorrect.**

An evaluation of the distribution of Olympic Games by season, based on the count of unique games grouped by games. Season, reveals a notable imbalance in seasonal occurrence. Specifically, 56.86% of the events were held during the summer season, while 43.14% took place in the winter. This indicates a historical skew toward the Summer Olympics, which have been hosted more frequently over time. The difference may be attributed to the earlier inception of the Summer Games in 1896, compared to the Winter Games, which began in 1924. Additionally, broader global participation, favourable weather, and logistical considerations have likely contributed to the dominance of the summer editions. Overall, this trend underscores a persistent seasonal preference in the organization of the Olympic Games.

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

**A graph with blue lines

AI-generated content may be incorrect.**

The chart displays the distribution of Olympic Games across various decades. From the 1930s to the 1990s, each decade consistently hosted two editions of the Games, indicating a stable pattern. However, starting from the 2000s, the number of Games per decade reduced to one. This change could be due to scheduling shifts, global events, or adjustments in the hosting pattern. The earlier decades also show some irregularities, likely due to world wars or organizational changes. Overall, the chart highlights both consistency and disruption in the historical timeline of the Olympic Games.

**3) Which cities have hosted the most Olympic Games?**

**A map of the world with blue circles

AI-generated content may be incorrect.**

An analysis of Olympic Games hosting frequency based on games\_id counts shows that London has hosted the most Olympic Games, followed closely by Athina (Athens) and Atlanta. These cities have established strong historical associations with the Olympics, serving as prominent hosts across different decades. London's recurring role reflects its global stature, infrastructure readiness, and consistent international appeal. Similarly, Athina holds symbolic significance as the birthplace of the ancient and modern Olympic Games, which has earned it repeated recognition. Atlanta’s prominence stems from its successful organization of the 1996 Summer Olympics and ongoing engagement in athletic development. The repeated selection of these cities underscores their enduring legacy, robust support systems, and alignment with the Olympic movement’s values. Overall, their multiple appearances highlight a blend of historical prestige and logistical capability that continues to influence host city selection.

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

A pie chart with numbers and a few words

AI-generated content may be incorrect.

The distribution of sports between the summer and winter Olympics appears to be relatively equal when comparing based on sport\_id Additionally, some sports are not assigned to any specific season, and these are represented as blank values in the dataset . Overall ,sports are nearly evenly distributed among the three categories summer, winter and blank(unassigned),indicating a balanced representation across Olympic seasons

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

**A graph of blue and white lines

AI-generated content may be incorrect.**

Athletics has the highest number of events in the Olympic Games, followed closely by Shooting. These two sports hold the top positions in terms of event count and are the most featured disciplines across Olympic editions. Following them are Swimming, Cycling, Sailing, Wrestling, and Art Competitions, which also have a significant number of events. In contrast, Tennis has comparatively fewer events, placing it among the sports with the lowest event counts.

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

**A graph on a computer screen

AI-generated content may be incorrect.**

The participation in each sport has evolved significantly over time. For instance, sports like Basketball, Speed Skating, and Athletics saw notable participation in the 1992 Olympics, with approximately 11.13k participants, and in 1988, with around 9.76k participants. Sports such as Handball and Luge also had high participation in 1976, reaching about 7.15k participants. These figures highlight how the popularity and scale of participation in different sports have varied over the years, reflecting changing trends, global interest, and inclusion of new events.Count of games year raise to fall Athletics ,shooting, rowing, football etc

**A screenshot of a graph and a map

AI-generated content may be incorrect.**

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

**A graph of a number of events

AI-generated content may be incorrect.**

Each sport in the Olympics consists of different events, and the number of events varies significantly from one sport to another. The event distribution ranges from high to low based on event \_ id counts. Sports like Athletics and Shooting have the highest number of events, each with more than 100 events, making them the most diversified disciplines in terms of competition categories. Following them are Cycling, Sailing, and Wrestling, which also have a considerable number of events. On the other end of the spectrum, sports like Tug of War have the fewest, with only 1 recorded event, indicating limited representation in the Olympic program. This variation illustrates how the structure and prominence of sports within the Olympics differ greatly.

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

**A blue pie chart with text and numbers

AI-generated content may be incorrect.**

The distribution of Olympic events by gender shows a significant difference in participation. Based on the count of event \_ id, men account for approximately 26.1%, with around 33.63k events. In contrast, women represent a much larger share, accounting for about 73.9%, with 95.23k events. This indicates a strong emphasis on women's events in the dataset, which may reflect changes in gender inclusion policies over time.

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

A graph on a white background

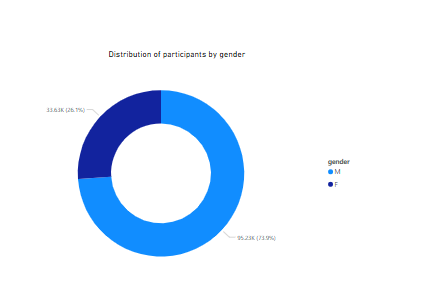
AI-generated content may be incorrect.

The number of events in the Olympic Games varies from year to year, reflecting changes in the sports program. For example, there were 314 events in 1992, while 2016 featured 306 events. These fluctuations occur as new sports and events are added or removed. The changes reflect evolving global interests, gender inclusion, and strategic decisions by the Olympic Committee. Overall, the event count highlights the dynamic nature of the Olympic Games.

**A screenshot of a graph

AI-generated content may be incorrect.**

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

****

The distribution of Olympic events by gender shows a significant difference in participation. Based on the count of event \_ id, men account for approximately 26.1%, with around 33.63k events. In contrast, women represent a much larger share, accounting for about 73.9%, with 95.23k events. This indicates a strong emphasis on women's events in the dataset, which may reflect changes in gender inclusion policies over time

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

**A graph of numbers and names

AI-generated content may be incorrect.**

An analysis of Olympic participation by region indicates that the United States leads with the highest number of athletes. This can be attributed to the country's strong emphasis on sports and consistent investment in athletic development, resulting in nearly 9225 participants over the years. The United Kingdom ranks second with approximately 5783 . Other countries with significant participation include France, Germany, Canada, Italy, and Japan. These regions have demonstrated a consistent and long-standing commitment to the Olympic Games, reflecting their active engagement in international sports competitions.

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

**A graph of a number of age

AI-generated content may be incorrect.**

The chart displays the age distribution of Olympic participants based on the count of person IDs. Most athletes fall within the age range of 18 to 30, with a sharp peak around age 22, indicating this is the most common age for Olympic competition. The number of participants declines significantly after age 30, with very few athletes above 40. This suggests that Olympic-level performance is concentrated in younger age groups, reflecting the physical demands of elite sports.

**A screenshot of a graph

AI-generated content may be incorrect.**

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

**A graph of a number of people

AI-generated content may be incorrect.**

The chart shows the distribution of Olympic medals by sport. Gymnastics stands out with the highest number of medals awarded, far surpassing other sports. Art Competitions, Swimming, and Athletics follow but with significantly lower medal counts. Most other sports have comparatively smaller medal tallies, indicating a more limited number of events or participation. This highlights how certain sports offer more medal opportunities than others in the Olympics.

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

**A graph of blue and black lines

AI-generated content may be incorrect.**

The chart displays the distribution of Olympic gold medals by region. The USA leads significantly, with more than double the gold medals of the next closest region, the Soviet Union. Other strong performers include Germany, UK, and Italy, though their counts are considerably lower. This reflects historical dominance and consistent athletic excellence from certain countries across Olympic history.

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

**A screenshot of a computer

AI-generated content may be incorrect.**

The chart illustrates the number of Olympic medals awarded across different sports. Gymnastics dominates with the highest medal count by a large margin, reflecting its numerous events and strong global participation. Art Competitions, Swimming, and Athletics also hold notable positions but with significantly fewer medals. Sports like Football, Nordic Combined, and Table Tennis show minimal medal counts, likely due to fewer events or more recent inclusion in the Games. This visualization emphasizes the varying scale of competition across Olympic sports.

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

**A graph of different countries/regions

AI-generated content may be incorrect.**

An analysis of Olympic participation by region indicates that the United States leads with the highest number of athletes. This can be attributed to the country's strong emphasis on sports and consistent investment in athletic development, resulting in nearly 9225 participants over the years. The United Kingdom ranks second with approximately 5783 . Other countries with significant participation include France, Germany, Canada, Italy, and Japan. These regions have demonstrated a consistent and long-standing commitment to the Olympic Games, reflecting their active engagement in international sports competitions.

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

**A graph of a number of people

AI-generated content may be incorrect.**

An analysis of Olympic participation by region indicates that the max number participant United States leads with the highest number of athletes. This can be attributed to the country's strong emphasis on sports and consistent investment in athletic development, resulting in nearly 9225 participants over the years. The United Kingdom ranks second with approximately 5783 . Other countries with significant participation include France, Germany, Canada, Italy, and Japan. These regions have demonstrated a consistent and long-standing commitment to the Olympic Games, reflecting their active engagement in international sports competitions.

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

A graph with a line going up

AI-generated content may be incorrect.

The total number of medals awarded is 260,971, distributed across various countries and regions around the world. For example, Afghanistan has received 89 medals, Albania has been awarded 74, and American Samoa has earned 46 medals. Algeria has a higher count with 492 medals distributed among its athletes. Each country’s medal tally reflects its athletes’ performances over many Olympic Games, highlighting the global diversity in sports excellence. This wide distribution of medals showcases the competitive spirit and achievements of nations from all corners of the globe. Overall, the medal counts serve as a testament to the dedication and skill of athletes representing their countries on the international stage.

**Hosting Trends and Games Analysis**

**Objective:**

Understand how hosting trends have evolved, highlighting countries and cities that have frequently hosted, and assessing the geographical spread.

**Visuals:**

* Bar Chart: Count of Games by host city
* Line Chart: Games held by decade (with disruptions like WWII shown)
* Map: Geographic spread of host countries
* Table: Cities with multiple hosting instances
* Slicer: Filter by year, city, or host country

**Insights:**

* Identify which countries have hosted most often
* Joint-hosting instances and geopolitical influences

**Sports and Events Analysis**

**Objective:**

Dive into the structure of sports and events across Olympics. Analyze which sports dominate, how events are distributed by gender, and how they've evolved.

**Visuals:**

* Bar Chart: Event count per sport
* Pie Chart: Sports representation share
* Column Chart: Events by gender (Men/Women/Mixed)
* Line Chart: Total events across Olympic years
* Slicer: Filter by sport or event

**Insights:**

* Gymnastics has the highest number of medal events
* Growth in gender diversity over time
* Introduction and discontinuation of specific sports

**Athlete Demographics & Participation**

**Objective:**

Explore the distribution of athletes based on age, gender, and their participation trends.

**Visuals:**

* Histogram: Age distribution of athletes (with peak around 20-25)
* KPI Cards: Average age, height, weight
* Bar Chart: Gender-wise participation across Games
* Area Chart: Participation growth over time
* Slicer: Filter by gender or year

**Insights:**

* Most athletes fall in the age range of 18-30
* Gender participation has become more balanced over time
* Consistency in height/weight for specific sports

**Medal Distribution Analysis**

**Objective:**

Evaluate how medals are distributed across sports and regions, highlighting top-performing countries and medal-rich sports.

**Visuals:**

* Bar Chart: Medal count by sport (Gymnastics leads significantly)
* Bar Chart: Gold Medals by region (USA dominates)
* Table: Top athletes with highest medal count
* Pie Chart: Share of Gold, Silver, Bronze
* Slicer: Filter by country, sport, year

**Insights:**

* USA leads in medal tally, followed by Soviet Union and Germany
* Some countries specialize in specific sports
* Gold medal counts concentrated among few regions

**Regional Representation & Recommendations**

**Objective:**

Understand geographic diversity, identify underrepresented regions, and propose future strategies.

**Visuals:**

* Map: Countries by total participation and medal count
* Bar Chart: Region-wise medal-to-participation ratio
* Table: Countries with low participation but potential
* Text Box: Recommendations for IOC and organizers

**Insights & Recommendations:**

* Increased participation from emerging regions (e.g., Africa, Southeast Asia)
* Need for better representation and investment in less-involved countries
* Encourage diversity in sports selection

**Gender Participation Analysis**

**Title**: *Olympic Participation by Gender*

**Visuals**:

* Line chart: Male vs. Female participants over the years
* Pie chart: Total share of male vs. female athletes
* Bar chart: Gender split by top 5 sports

**Key Insights**:

* Track the evolution of gender participation over time
* Identify sports with greater gender equality
* Analyse when and where female participation surged

**Summary**:  
This page highlights the growth in female participation across Olympic events, showcasing how gender balance has evolved and identifying gaps and leading sports in terms of inclusivity.

**Country-Wise Medal Efficiency**

**Title**: *Medal Efficiency by Country*

**Visuals**:

* Scatter plot: Number of athletes vs. medals won (by country)
* KPI tiles: Top 3 most efficient countries
* Heatmap: Continent-wise medal efficiency

**Key Insights**:

* Which countries win more medals per athlete sent
* Detect over- and under-performing nations
* Uncover hidden strong performers (e.g., small teams, high results)

**Summary**:  
Focuses on medal conversion efficiency rather than volume, helping reveal countries that make the most out of their athlete pool regardless of size.

**Athlete Longevity & Performance**

**Title**: *Athlete Career & Longevity Analysis*

**Visuals**:

* Histogram: Age distribution of medal winners
* Line chart: Repeat medalists across years
* Table: Athletes with participation in 3+ editions

**Key Insights**:

* Average age of medalists by sport
* Patterns of long athletic careers
* Sports with higher longevity (e.g., equestrian, shooting)

**Summary**:  
This analysis highlights how long athletes stay competitive in the Olympics and which disciplines allow longer careers, offering insight into physical and skill-based demands.

**Recommendations & Insights**

**Title**: *Strategic Recommendations*

**Visuals**:

* Bullet points or infographic-style layout
* Icon-based cards summarizing insights
* Comparative bar chart (e.g., medal gaps by gender, region, or sport)

**Recommendations Could Include**:

* Encourage participation in underrepresented sports and regions
* Promote gender balance in events and funding
* Focus development programs on sports with higher medal potential
* Use historical data for better city/region selection for hosting

**Summary**:  
Concludes the analysis with data-driven recommendations for Olympic committees, sports federations, and host cities to enhance inclusivity, performance, and global reach.

**Summary:**

A detailed exploration of historical Olympic Games data, as structured in the dataset, highlights distinct patterns in both seasonal occurrence and city-level hosting frequency. Based on the grouping and count of games by season, approximately 56.86% of the Olympic events have taken place during the summer season, while 43.14% were held during the winter season. This reveals a clear seasonal imbalance that Favors the Summer Olympics, likely influenced by a combination of factors such as climate accessibility, global athlete participation, and broader logistical feasibility. The Summer Games, first held in 1896, have had a longer timeline to accumulate editions, whereas the Winter Olympics only began in 1924, contributing to the numerical gap.

When examining host cities, London emerges as the most frequent host, having successfully organized the Olympics three times (1908, 1948, and 2012). It is followed by Athina (Athens), which holds historical prestige as the birthplace of the Olympic Games, and Atlanta, which hosted the centennial games in 1996. These cities have established a lasting legacy, not just through official hosting but also through consistent athletic participation and infrastructural readiness. Their repeated selection reflects a blend of symbolic importance, geopolitical influence, and proven capability to support the complex demands of hosting a global event.

The data suggests that certain cities and regions continue to dominate Olympic narratives due to historical connections, recurring investments in sports infrastructure, and strategic positioning within the Olympic movement. For instance, Athens' cultural significance grants it a unique status, while cities like London benefit from modern facilities, global appeal, and political stability. Meanwhile, newer hosts often emerge based on rotational fairness, continental balance, or the Olympic Committee’s intent to expand global representation.

Overall, the Olympic Games’ evolution demonstrates a strong interplay between tradition and transformation. While summer events and historically significant cities maintain dominance, there is also an ongoing effort to diversify the Olympic experience across seasons and geographies. This dynamic continues to shape the legacy and global reach of the Games in each new edition.

END