Olympics Data Analysis

-- MySQL

```sql
-- 1. How many olympics games have been held?
Select
    Count(Distinct Games) AS Total_Olympics_Played
From Olympics_athlete_events;
```

Total_Olympics_Played
51

```sql
-- 2. List down all Olympics games held so far.
Select
    distinct Year,
    Season,
    City
From Olympics_athlete_events
Order by year;
```

Year	Season	City
1896	Summer	Athina
1900	Summer	Paris
1904	Summer	St. Louis
1906	Summer	Athina
1908	Summer	London
1912	Summer	Stockholm
1920	Summer	Antwerpen
1924	Summer	Paris
1924	Winter	Chamonix
1928	Summer	Amsterdam
1928	Winter	Sankt Moritz
1932	Summer	Los Angeles
1932	Winter	Lake Placid
1936	Summer	Berlin
1936	Winter	Garmisch-...
1948	Summer	London

```sql
-- 3. Mention the total no of nations who participated in each olympics game?
Select
    Games,
    Count(distinct Region) AS Total_Country
From olympics_athlete_events a Join olympics_noc_region n ON a.NOC = n.NOC
Group By Games
Order by Games ;
```

Games	Total_Country
1896 Summer	12
1900 Summer	31
1904 Summer	14
1906 Summer	20
1908 Summer	22
1912 Summer	29
1920 Summer	29
1924 Summer	45
1924 Winter	19
1928 Summer	46
1928 Winter	25
1932 Summer	44
1932 Winter	17
1936 Summer	49
1936 Winter	28
1948 Summer	58
1948 Winter	28
1952 Summer	66

```sql
-- 4. Which year saw the highest and lowest no of countries participating in olympics.
(Select
    Concat("Highest -",Games) AS Games,
    Count(distinct Region) AS Total_Participants_Countries
From olympics_athlete_events o Join olympics_noc_region n ON o.NOC = n.NOC
Group by Games
Order by Total_Participants_Countries DESC
LIMIT 1)
Union all
(Select
    Concat("Lowest -",Games) AS Games,
    Count(distinct Region) AS Total_Participants_Countries
From olympics_athlete_events o Join olympics_noc_region n ON o.NOC = n.NOC
Group by Games
Order by Total_Participants_Countries ASC
LIMIT 1);
```

Games	Total_Participants_Countries
Highest -2016 Summer	204
Lowest -1896 Summer	12

```sql
-- 5. Which nation has participated in all of the olympic games.
With CTE1 AS
(Select
     Region as Country,
     Count(distinct Games) AS Total_Participants,
     rank() Over (Order by Count(distinct Games) DESC) AS RN
From olympics_athlete_events o Join olympics_noc_region n ON o.NOC = n.NOC
Group By Region)
Select
     Country,
     Total_Participants
From CTE1
Where RN = 1;
```

Country	Total_Participants
France	51
Italy	51
Switzerland	51
UK	51

```sql
-- 6. Identify the sport which was played in all summer olympics.
Select
        Sport,
        Count(distinct Games) AS Played
From olympics_athlete_events
Where Season = "Summer"
Group By Sport
Order By Played DESC;
```

Sport	Played
Athletics	29
Swimming	29
Cycling	29
Fencing	29
Gymnastics	29
Wrestling	28
Rowing	28
Shooting	27
Diving	27
Football	27
Water Polo	27
Sailing	26

```sql
-- 7. Which Sports were just played only in one olympics season.
With CTE1 AS
(Select
      Sport,
      Count(distinct Games) AS No_of_season_Played
From olympics_athlete_events
Group By Sport)
Select
      Sport
from CTE1
Where No_of_season_Played = 1;
```

Sport
Aeronautics
Basque Pelota
Cricket
Croquet
Jeu De Paume
Military Ski Patrol
Motorboating
Racquets
Roque
Rugby Sevens

```sql
-- 8. Fetch the total no of sports played in each olympic games.
Select
    distinct Games,
    count(distinct Sport) AS Total_Sports
From olympics_athlete_events
Group By Games
Order by Games ;
```

Games	Total_Sports
1896 Summer	9
1900 Summer	20
1904 Summer	18
1906 Summer	13
1908 Summer	24
1912 Summer	17
1920 Summer	25
1924 Summer	20
1924 Winter	10
1928 Summer	17
1928 Winter	8
1932 Summer	18
1932 Winter	7

```sql
-- 9. Fetch oldest athletes to win a gold medal.
With CTE1 AS
(Select *,
    Case
    when age = 'NA' Then null
    Else age
    End AS New_Age
From olympics_athlete_events
Where Medal = "Gold"),
CTE2 AS
(Select *,
    dense_rank() Over (Order by New_age DESC) AS RN
From CTE1)
Select Name,
    Age,
    Games,
    Medal
 FROM CTE2
Where RN = 1;
```

Name	Age	Games	Medal
Oscar Gomer Swahn	64	1912 Summer	Gold
Charles Jacobus	64	1904 Summer	Gold

```sql
-- 10. Find the Percent of male and female athletes participated in all olympic games.
SELECT
    CONCAT(ROUND(SUM(CASE WHEN Sex = 'M' THEN 1 END) / COUNT(*) * 100, 1), '%') AS male_Pct,
    CONCAT(ROUND(SUM(CASE WHEN Sex = 'F' THEN 1 END) / COUNT(*) * 100, 1), '%') AS female_Pct
FROM
    olympics_athlete_events;
```

male_Pct	female_Pct
72.4%	27.6%

```sql
-- 11. Fetch the top 5 athletes who have won the most gold medals.
With CTE1 AS
(Select
        Name,
        Count(Medal) AS Won,
        Dense_Rank() Over (Order by Count(Medal) DESC) AS RN
From olympics_athlete_events
Where Medal = "Gold"
Group By Name)
Select * From CTE1
Where RN <=5;
```

Name	Won	RN
Michael Fred Phelps, II	23	1
Raymond Clarence "Ray" Ewry	10	2
Larysa Semenivna Latynina (Diriy-)	9	3
Frederick Carlton "Carl" Lewis	9	3
Paavo Johannes Nurmi	9	3
Mark Andrew Spitz	9	3
Jennifer Elisabeth "Jenny" Thompson (-Cumpelik)	8	4
Matthew Nicholas "Matt" Biondi	8	4
Ole Einar Bjrndalen	8	4
Usain St. Leo Bolt	8	4
Birgit Fischer-Schmidt	8	4
Sawao Kato	8	4
Nikolay Yefimovich Andrianov	7	5
Vra slavsk (-Odloilov)	7	5
Viktor Ivanovych Chukarin	7	5
Aladr Gerevich (-Gerei)	7	5
Donald Arthur "Don" Schollander	7	5

```sql
-- 12. Fetch the top 5 athletes who have won the most medals (gold/silver/bronze).
With CTE1 AS
(Select
    Name,
    Count(Medal) AS Total_Medals,
    Dense_Rank() Over (Order by Count(Medal) DESC) AS RN
From olympics_athlete_events
Where Medal IN ("Gold","Silver","Bronze")
Group By Name)
Select
    *
From CTE1
Where RN <= 5;
```

Name	Total_Medals	RN
Michael Fred Phelps, II	28	1
Larysa Semenivna Latynina (Diriy-)	18	2
Nikolay Yefimovich Andrianov	15	3
Ole Einar Bjrndalen	13	4
Takashi Ono	13	4
Borys Anfiyanovych Shakhlin	13	4
Edoardo Mangiarotti	13	4
Sawao Kato	12	5
Ryan Steven Lochte	12	5
Aleksey Yuryevich Nemov	12	5
Paavo Johannes Nurmi	12	5
Birgit Fischer-Schmidt	12	5
Natalie Anne Coughlin (-Hall)	12	5
Jennifer Elisabeth "Jenny" Thomp...	12	5
Dara Grace Torres (-Hoffman, -Mi...	12	5

```sql
-- 13. Fetch the top 5 most successful countries in olympics. Success is defined by no. of medals won.
With CTE1 AS
(Select
        Region AS Country,
        Count(Medal) AS Won,
        Dense_Rank() Over (Order By Count(Medal) DESC) AS RN
From olympics_athlete_events o Join olympics_noc_region n ON o.NOC = n.NOC
Where Medal IN (select Medal From olympics_athlete_events Where Medal != "NA")
Group By Country)
Select * From CTE1
Where RN <=5;
```

Country	Won	RN
USA	5637	1
Russia	3947	2
Germany	3756	3
UK	2066	4
France	1757	5

```sql
-- 14. List down total gold, silver and bronze medals won by each country.
Select
    n.Region AS Country,
    Count(Case when Medal = "Gold" Then 1  End) AS Gold,
    Count(Case when Medal = "Silver" Then 1  End) AS Silver,
    Count(Case when Medal = "Bronze" Then 1  End) AS Bronze
From olympics_athlete_events o Join olympics_noc_region n ON o.NOC = n.NOC
Group By Country
Order by Gold DESC, Silver DESC, Bronze DESC;
```

Country	Gold	Silver	Bronze
USA	2638	1641	1358
Russia	1599	1170	1178
Germany	1301	1195	1260
UK	676	739	651
Italy	575	531	531
France	497	594	666
Sweden	479	522	535
Canada	463	438	451
Hungary	432	332	371
Norway	378	361	294
Australia	368	459	522
China	351	349	293
Netherlands	287	340	413
Japan	247	309	357
South Korea	221	232	185
Finland	198	270	432

```sql
/*15. List down total gold, silver and bronze medals won by each country corresponding to each
      olympic games.*/
Select
    Games,
    Region AS Country,
    Count(Case when Medal = "Gold" Then 1  End) AS Gold,
    Count(Case when Medal = "Silver" Then 1 End) AS Silver,
    Count(Case when Medal = "Bronze" Then 1 End) AS Bronze
From olympics_athlete_events o Join olympics_noc_region n ON o.NOC = n.NOC
Group By Country, Games
Order by Gold DESC, Silver DESC, Bronze DESC;
```

Games	Country	Gold	Silver	Bronze
1980 Summer	Russia	187	129	126
1984 Summer	USA	186	116	50
1996 Summer	USA	159	48	52
1908 Summer	UK	147	131	90
2012 Summer	USA	145	57	46
2016 Summer	USA	139	54	71
1988 Summer	Russia	134	67	99
2000 Summer	USA	130	61	51
1904 Summer	USA	128	141	125
2008 Summer	USA	127	110	80
1976 Summer	Germany	123	76	74
2004 Summer	USA	117	75	71
1980 Summer	Germany	115	88	61
1976 Summer	Russia	114	95	77
1988 Summer	Germany	111	91	94

```sql
-- 17. Identify which country won the most gold, most silver, most bronze medals in each olympic games.
WITH cte AS (
SELECT
    games,
    region,
    medal,
    COUNT(*) AS medal_count,
    DENSE_RANK() OVER (PARTITION BY games, medal ORDER BY COUNT(*) DESC)
    AS _rank
FROM olympics_athlete_events o JOIN olympics_noc_region n ON o.NOC = n.NOC
WHERE medal IN ('Gold', 'Silver', 'Bronze')
GROUP BY games, region, medal)
SELECT
    games,
    MAX(CASE WHEN medal = 'Gold' AND _rank = 1 THEN CONCAT(region, ' ', medal_count) END) AS max_gold,
    MAX(CASE WHEN medal = 'Silver' AND _rank = 1 THEN CONCAT(region, ' ', medal_count) END) AS max_silver,
    MAX(CASE WHEN medal = 'Bronze' AND _rank = 1 THEN CONCAT(region, ' ', medal_count) END) AS max_bronze
FROM cte
GROUP BY games
ORDER BY games;
```

games	max_gold	max_silver	max_bronze
1896 Summer	Germany 25	Greece 18	Greece 20
1900 Summer	UK 57	France 85	France 82
1904 Summer	USA 128	USA 141	USA 125
1906 Summer	Greece 24	Greece 48	Greece 30
1908 Summer	UK 147	UK 131	UK 90
1912 Summer	Sweden 103	UK 64	UK 59
1920 Summer	USA 111	France 71	Belgium 66
1924 Summer	USA 97	France 51	USA 49
1924 Winter	UK 16	USA 10	UK 11
1928 Summer	USA 47	Netherlands 29	Germany 41
1928 Winter	Canada 12	Sweden 13	Switzerland...
1932 Summer	USA 81	USA 47	USA 61
1932 Winter	Canada 14	USA 21	Germany 14
1936 Summer	Germany 93	Germany 70	Germany 61
1936 Winter	UK 12	Canada 13	USA 14

```sql
-- 18. Which countries have never won gold medal but have won silver/bronze medals?
With CTE1 AS
(Select
    Region As Country,
    Count(Case when Medal = "Gold" Then 1  End) AS Gold,
    Count(Case when Medal = "Silver" Then 1 End) AS Silver,
    Count(Case when Medal = "Bronze" Then 1 End) AS Bronze
From olympics_athlete_events o Join olympics_noc_region n ON o.NOC = n.NOC
Group By Country)
Select * From CTE1
Where Gold = 0 and (Silver > 0 or Bronze > 0);
```

Country	Gold	Silver	Bronze
Afghanistan	0	0	2
Barbados	0	0	1
Bermuda	0	0	1
Botswana	0	1	0
Curacao	0	1	0
Cyprus	0	1	0
Djibouti	0	0	1
Eritrea	0	0	1
Gabon	0	1	0
Ghana	0	1	22
Guatemala	0	1	0
Guyana	0	0	1
Iceland	0	15	2

```sql
-- 19. In which Sport/event, India has won highest medals.
Select
    Region as country,
    Sport,
    Count(Medal) AS Count_Medal
From olympics_athlete_events o Join olympics_noc_region n ON o.NOC = n.NOC
Where Region = "India" and Medal != "NA"
Group By Country, Sport
Order By Count_Medal DESC
Limit 1;
```

```sql
/*20. Break down all olympic games where India won medal for Hockey and how many medals in
        each olympic games.*/
Select
      Region AS Country,
      Games,
      Sport,
      Count(medal) AS Hockey_Medal
From olympics_athlete_events o Join olympics_noc_region n ON o.NOC = n.NOC
Where Region = "India" and Sport = "Hockey" and Medal != "NA"
Group By Country, Games
Order by Hockey_Medal DESC;
```

Country	Games	Sport	Hockey_Medal
India	1948 Summer	Hockey	20
India	1936 Summer	Hockey	19
India	1956 Summer	Hockey	17
India	1980 Summer	Hockey	16
India	1968 Summer	Hockey	16
India	1964 Summer	Hockey	15
India	1932 Summer	Hockey	15
India	1928 Summer	Hockey	14
India	1972 Summer	Hockey	14
India	1952 Summer	Hockey	14
India	1960 Summer	Hockey	13

Thank you