120 years of Olympic history: athletes and results

Reda EL JABIRY

CONTEXT

The present dataset embodies the complete history of the modern Olympic Games, encompassing all editions, from Athens 1896 to Rio 2016.

DATASET : https://www.kaggle.com/datasets/heesoo37/120-years-of-olympic-history-athletes-and-results

About Dataset

There are two tables, one of which is named "athlete". Each record within this table comprises relevant information about a particular athlete. The second table consists of the names of National Olympic Committees (NOC) and their corresponding countries.

Table 1 : athlete

1. **ID - Unique number for each athlete**
2. **Name - Athlete's name**
3. **Sex - M or F**
4. **Age - Integer**
5. **Height - In centimeters**
6. **Weight - In kilograms**
7. **Team - Team name**
8. **NOC - National Olympic Committee 3-letter code**
9. **Games - Year and season**
10. **Year - Integer**
11. **Season - Summer or Winter**
12. **City - Host city**
13. **Sport - Sport**
14. **Event - Event**
15. **Medal - Gold, Silver, Bronze, or NA**

Table 2 : country

1. **NOC - National Olympic Committee 3-letter code**
2. **country_name : name of the country**

In which year did the Olympics have the highest and lowest number of participating countries?

```sql
WITH  all_countries AS
) (SELECT Games, COUNT(DISTINCT noc) as count
  FROM olympics.athlete
  GROUP BY Games)
  SELECT DISTINCT CONCAT(first_value(Games) over (ORDER BY count),' _ ',first_value(count) over (ORDER BY count)) AS MIN
      CONCAT(first_value(Games) over (ORDER BY count DESC),' _ ',first_value(count) over (ORDER BY count DESC)) AS MAX
  FROM all_countries
```

MIN	MAX
1896 Summer _ 12	2016 Summer _ 207

Which country has participated in all of the olympic games

```sql
# 5. Which nation has participated in all of the olympic games
WITH no_olympic_games AS(
SELECT COUNT(DISTINCT Games) as count
FROM olympics.athlete)
SELECT c.country_name
FROM olympics.athlete a
JOIN olympics.country c
ON a.noc = c.noc
GROUP BY c.country_name
HAVING COUNT(DISTINCT Games) = (SELECT * FROM no_olympic_games)
```

	country_name
▶	France
	Italy
	Switzerland
	UK

Sport which was played in all summer olympics

```sql
WITH summer_olympics AS (SELECT COUNT(DISTINCT Year,Season)
FROM  olympics.athlete a
WHERE a.Season = 'Summer')
SELECT Sport,COUNT(DISTINCT Year,Season)
FROM  olympics.athlete a
WHERE a.Season = 'Summer'
GROUP BY Sport
HAVING COUNT(DISTINCT Year,Season) = (SELECT * FROM summer_olympics)
```

Sport	COUNT(DISTINCT Year,Season)
Athletics	29
Cycling	29
Fencing	29
Gymnastics	29
Swimming	29

Which sports that were played just once

```sql
WITH cte AS (
SELECT Sport,Games
FROM olympics.athlete
GROUP BY Sport,Games
),
cte1 AS (SELECT Sport,COUNT(DISTINCT Games) AS count
FROM cte
GROUP BY Sport
HAVING COUNT(DISTINCT Games) = 1)
SELECT cte.Sport,Games,count
FROM cte
JOIN cte1
ON cte.Sport = cte1.Sport
ORDER BY Sport
```

Sport	Games	count
Aeronautics	1936 Summer	1
Basque Pelota	1900 Summer	1
Cricket	1900 Summer	1
Croquet	1900 Summer	1
Jeu De Paume	1908 Summer	1
Military Ski Patrol	1924 Winter	1
Motorboating	1908 Summer	1
Racquets	1908 Summer	1
Roque	1904 Summer	1
Rugby Sevens	2016 Summer	1

Total number of sports played in each olympics

```sql
SELECT Games,COUNT(DISTINCT Sport) AS count
FROM olympics.athlete
GROUP BY Games
ORDER BY count DESC
```

Games	count
2004 Summer	34
2016 Summer	34
2008 Summer	34
2000 Summer	34
2012 Summer	32
1996 Summer	31
1992 Summer	29
1988 Summer	27
1920 Summer	25
1984 Summer	25
1908 Summer	24
1936 Summer	24
1980 Summer	23
1976 Summer	23
1972 Summer	23
1964 Summer	21
1900 Summer	20
1924 Summer	20
1968 Summer	20
1948 Summer	20

Finding the oldest athletes to win a gold medal

```sql
SELECT DISTINCT Name,Sex,Height,Weight,Team,NOC,Games,City,Sport,Event,Medal
FROM olympics.athlete
WHERE  Age = (SELECT MAX(Age) FROM olympics.athlete WHERE Age NOT IN ('NA') AND Medal = 'Gold') AND Medal = 'Gold'
```

Name	Sex	Height	Weight	Team	NOC	Games	City	Sport	Event	Medal
Charles Jacobus	M	NA	NA	United States	USA	1904 Summer	St. Louis	Roque	Roque Men's Singles	Gold
Oscar Gomer Swahn	M	NA	NA	Sweden	SWE	1912 Summer	Stockholm	Shooting	Shooting Men's Running Target, …	Gold

TOP 5 athletes who have won the most medals

```sql
SELECT Name,Team,COUNT(Medal) as count
FROM olympics.athlete
WHERE Medal NOT IN ('NA')
GROUP BY Name,Team
ORDER BY count DESC
```

	Name	Team	count
▶	Michael Fred Phelps, II	United States	28
	Larysa Semenivna Latynina (Diriy-)	Soviet Union	18
	Nikolay Yefimovich Andrianov	Soviet Union	15
	Ole Einar Bjrndalen	Norway	13
	Edoardo Mangiarotti	Italy	13

The top five countries that have earned the highest number of medals.

```sql
SELECT DISTINCT country_name,COUNT(Medal) as count
FROM olympics.athlete a
JOIN olympics.country c
ON a.NOC = c.NOC
WHERE Medal NOT IN ('NA')
GROUP BY country_name
```

	country_name	count
▶	USA	5637
	Russia	3947
	Germany	3756
	UK	2068
	France	1777