

120 years of Olympic history: Athletes and results

Introduction

This dataset provides an opportunity to ask questions about how the Olympics have evolved over time, including questions about the participation and performance of women, different nations, and different sports and events.

Context

This is a historical dataset on the modern Olympic Games, including all the Games from Athens **1896 to Rio 2016**. I scraped this data from www.sports-reference.com in May 2018. The R code I used to scrape and wrangle the data is on GitHub. I recommend checking [my kernel](#) before starting your own analysis.

Note that the Winter and Summer Games were held in the same year up until 1992. After that, they staggered them such that Winter Games occur on a four year cycle starting with 1994, then Summer in 1996, then Winter in 1998, and so on. A common mistake people make when analyzing this data is to assume that the Summer and Winter Games have always been staggered.

1. How many Olympics games have been held?

Query		Query History
1	--//How many olympics games have been held?	
2	select count (distinct(games)) as Total_Olympic_Games	
3	from olympics_history	
Data Output		Messages Notifications
	total_olympic_games bigint	
1	51	

2. List down all Olympics games held so far.

```

1  --//List down all Olympics games held so far.?
2  select distinct year, season,city
3  from olympics_history
4  order by year

```

Data Output

Messages

Notifications

≡+

📄

▼

📋

▼

🗑

🗑

📶

⬇

📈

	year integer	season character varying	city character varying
1	1896	Summer	Athina
2	1900	Summer	Paris
3	1904	Summer	St. Louis
4	1906	Summer	Athina
5	1908	Summer	London
6	1912	Summer	Stockholm
7	1920	Summer	Antwerpen
8	1924	Summer	Paris
9	1924	Winter	Chamonix
10	1928	Summer	Amsterdam
11	1928	Winter	Sankt Moritz
12	1932	Summer	Los Angeles
13	1932	Winter	Lake Placid

Total rows: 52 of 52

Query complete 00:00:00.251

3. Mention the total no of nations who participated in each olympics game?

```

1 select games, count(distinct(noc))
2 from olympics_history
3 group by games
4 order by count(distinct(noc))

```

Data Output

Messages

Notifications

≡ +

📄 ▼

📋 ▼

🗑️

🗄️

⬇️

📈

	games character varying	count bigint
1	1896 Summer	12
2	1904 Summer	15
3	1932 Winter	17
4	1924 Winter	19
5	1906 Summer	21
6	1908 Summer	22
7	1928 Winter	25
8	1948 Winter	28
9	1936 Winter	28
10	1912 Summer	29
11	1920 Summer	29
12	1960 Winter	30
13	1952 Winter	30

Total rows: 51 of 51

Query complete 00:00:02.371

4. Which year saw the highest and lowest no. of countries participating in Olympics?

Query	Query History
1	<code>select games, count(distinct(noc))</code>
2	<code>from olympics_history</code>
3	<code>group by games</code>
4	<code>order by count(distinct(noc)) desc</code>
5	<code>limit 1</code>
Data Output	Messages Notifications
<div> <div>games</div> <div>character varying</div> </div> <div> <div>count</div> <div>bigint</div> </div>	
1	2016 Summer 207

Query	Query History
1	<code>select games, count(distinct(noc))</code>
2	<code>from olympics_history</code>
3	<code>group by games</code>
4	<code>order by count(distinct(noc))</code>
5	<code>limit 1</code>
Data Output	Messages Notifications
<div> <div>games</div> <div>character varying</div> </div> <div> <div>count</div> <div>bigint</div> </div>	
1	1896 Summer 12

5. Which nation has participated in all of the Olympic games?

Query	Query History
3	<code>join olympics_history_noc_regions H</code>
4	<code>on O.NOC=H.NOC</code>
5	<code>group by region</code>
6	<code>having(count(distinct games))=51</code>
7	
Data Output	Messages Notifications
<div> <div>region</div> <div>character varying</div> </div> <div> <div>total_games</div> <div>bigint</div> </div>	
1	France 51
2	Italy 51
3	Switzerland 51
4	UK 51

6. Identify the sport which was played in all summer olympics.

Query	Query History
1	<code>with CTE as(</code>
2	<code>select count (distinct games) as total_games,sport</code>
3	<code>from olympics_history</code>
4	<code>where season= 'Summer'</code>
5	<code>group by sport)</code>
6	<code>select sport, total_games</code>
7	<code>from CTE</code>
8	<code>where total_games=29</code>
Data Output	Messages Notifications
<div> <div>sport</div> <div>character varying</div> </div> <div> <div>total_games</div> <div>bigint</div> </div>	
1	Athletics 29
2	Cycling 29
3	Fencing 29
4	Gymnastics 29
5	Swimming 29

7. Which Sports were just played only once in the olympics.

Query		Query History	
1	with CTE as(
2	select Sport,		
3	count (distinct games) as total_games		
4	from olympics_history		
5	group by sport)		
6	select sport, total_games		
7	from CTE		

Data Output		Messages	Notifications
	sport character varying	total_games bigint	
1	Aeronautics		1
2	Basque Pelota		1
3	Cricket		1
4	Croquet		1
5	Jeu De Paume		1
6	Military Ski Patrol		1
7	Motorboating		1
8	Racquets		1
9	Roque		1
10	Rugby Sevens		1

8. Fetch the total no of sports played in each Olympic games.

1	select count (distinct Sport) as No_of_games_played,games		
2	from olympics_history		
3	group by games		
4	order by No_of_games_played desc		

Data Output		Messages	Notifications
	no_of_games_played bigint	games character varying	
1	34	2008 Summer	
2	34	2000 Summer	
3	34	2004 Summer	
4	34	2016 Summer	
5	32	2012 Summer	
6	31	1996 Summer	
7	29	1992 Summer	
8	27	1988 Summer	
9	25	1920 Summer	
10	25	1984 Summer	
11	24	1908 Summer	
12	24	1936 Summer	
13	23	1972 Summer	

Total rows: 51 of 51	Query complete 00:00:00.980
----------------------	-----------------------------