



120 YEARS OF OLYMPIC HISTORY

ATHLETES AND RESULTS



About Dataset

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.

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.

Content

- The file `athlete_events.csv` contains 271116 rows and 15 columns. Each row corresponds to an individual athlete competing in an individual Olympic event (athlete-events). The columns are:

ID - Unique number for each athlete

Name - Athlete's name

Sex - M or F

Age - Integer

Height - In centimeters

Weight - In kilograms

Team - Team name

NOC - National Olympic Committee 3-letter code

Games - Year and season

Year - Integer

Season - Summer or Winter

City - Host city

Sport - Sport

Event - Event

Medal - Gold, Silver, Bronze, or NA

List of all these 20 queries mentioned below:

- 1.How many olympics games have been held?
- 2.List down all Olympics games held so far.
- 3.Mention the total no of nations who participated in each olympics game?
- 4.Which year saw the highest and lowest no of countries participating in olympics?
- 5.Which nation has participated in all of the olympic games?
- 6.Identify the sport which was played in all summer olympics.
- 7.Which Sports were just played only once in the olympics?
- 8.Fetch the total no of sports played in each olympic games.
- 9.Fetch details of the oldest athletes to win a gold medal.
- 10.Find the Ratio of male and female athletes participated in all olympic games.
- 11.Fetch the top 5 athletes who have won the most gold medals.
- 12.Fetch the top 5 athletes who have won the most medals (gold/silver/bronze).
- 13.Fetch the top 5 most successful countries in olympics. Success is defined by no of medals won.
- 14.List down total gold, silver and broze medals won by each country.
- 15.List down total gold, silver and broze medals won by each country corresponding to each olympic games.
- 16.Identify which country won the most gold, most silver and most bronze medals in each olympic games.
- 17.Identify which country won the most gold, most silver, most bronze medals and the most medals in each olympic games.
- 18.Which countries have never won gold medal but have won silver/bronze medals?
- 19.In which Sport/event, India has won highest medals.
- 20.Break down all olympic games where india won medal for Hockey and how many medals in each olympic games.

Problem Statement: Write a SQL query to find the total no of Olympic Games held as per the dataset

Expected Output:

	total_olympic_games
▶	51

Expected Output: Query 1

Problem Statement: Write a SQL query to list down all the Olympic Games held so far.

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
1948	Winter	Sankt Moritz
1952	Summer	Helsinki

Expected Output: Query 2 (Showing few records only)

Problem Statement: SQL query to fetch total no of countries participated in each olympic games.

games	total_countries
1896 Summer	1
1900 Summer	4
1904 Summer	1
1906 Summer	5
1908 Summer	8
1912 Summer	15
1920 Summer	11
1924 Summer	19
1924 Winter	5
1928 Summer	23
1928 Winter	2
1932 Summer	12
1932 Winter	1
1936 Summer	28
1936 Winter	3
1948 Summer	31
1948 Winter	4
1952 Summer	33

Expected Output: Query 3
(Showing few records only)

Problem Statement: Write a SQL query to return the Olympic Games which had the highest participating countries and the lowest participating countries.

	Lowest_Countries	Highest_Countries
▶	1896 Summer - 1	2016 Summer - 98

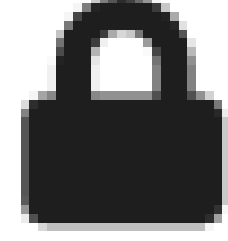
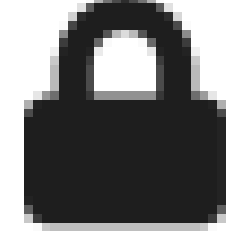
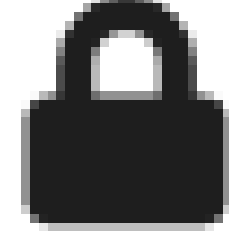
Expected Output: Query 4

Problem Statement: SQL query to return the list of countries who have been part of every Olympics games.

country character varying	total_participated_games bigint
France	51
Italy	51
Switzerland	51
UK	51

Expected Output: Query 5

Problem Statement: SQL query to fetch the list of all sports which have been part of every olympics

sport character varying 	no_of_games bigint 	total_games bigint 
Swimming	29	29
Cycling	29	29
Fencing	29	29
Gymnastics	29	29
Athletics	29	29

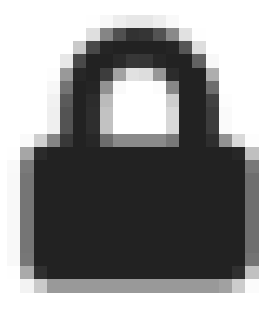
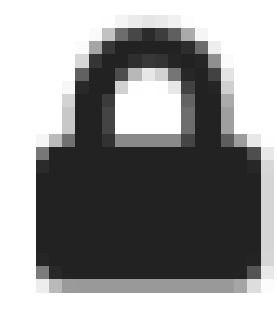
Expected Output: Query 6

Problem Statement: Using SQL query, Identify the sport which were just played once in all of olympics.

sport character varying 	no_of_games bigint 	games character varying 
Aeronautics	1	1936 Summer
Basque Pelota	1	1900 Summer
Cricket	1	1900 Summer
Croquet	1	1900 Summer
Jeu De Paume	1	1908 Summer
Military Ski Patrol	1	1924 Winter
Motorboating	1	1908 Summer
Racquets	1	1908 Summer
Roque	1	1904 Summer
Rugby Sevens	1	2016 Summer

Expected Output: Query 7

Problem Statement: Write SQL query to fetch the total no of sports played in each olympics

games character varying 	no_of_sports bigint 
2000 Summer	34
2004 Summer	34
2016 Summer	34
2008 Summer	34
2012 Summer	32
1996 Summer	31
1992 Summer	29
1988 Summer	27
1920 Summer	25
1984 Summer	25

Expected Output: Query 8 (Showing few records only)

Problem Statement: SQL Query to fetch the details of the oldest athletes to win a gold medal at the olympics.

name	age	team
character varying	integer	character varying
Oscar Gomer Swahn	64	Sweden
Charles Jacobus	64	United States

games	sport	medal
character varying	character varying	character varying
1912 Summer	Shooting	Gold
1904 Summer	Roque	Gold

Expected Output: Query 9

Problem Statement: Write a SQL query to get the ratio of male and female participants

ratio

text



1 : 2.64

Expected Output: Query 10

Problem Statement: SQL query to fetch the top 5 athletes who have won the most gold medals

name	total_gold_medals
character varying	bigint
Michael Fred Phelps, II	23
Raymond Clarence "Ray" Ewry	10
Mark Andrew Spitz	9
Larysa Semenivna Latynina (Diriy-)	9
Frederick Carlton "Carl" Lewis	9
Paavo Johannes Nurmi	9
Ole Einar Bjrndalen	8
Matthew Nicholas "Matt" Biondi	8
Sawao Kato	8

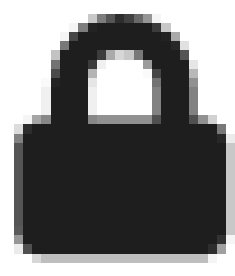
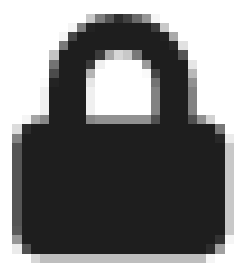
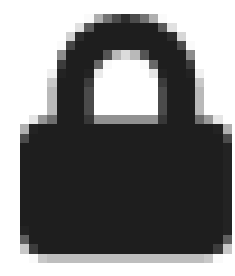
Expected Output: Query 11
(Showing few records only)

Problem Statement: SQL Query to fetch the top 5 athletes who have won the most medals (Medals include gold, silver and bronze).

name character varying	team character varying	total_medals bigint
Michael Fred Phelps, II	United States	28
Larysa Semenivna Latynina (Diriy-)	Soviet Union	18
Nikolay Yefimovich Andrianov	Soviet Union	15
Edoardo Mangiarotti	Italy	13
Takashi Ono	Japan	13
Borys Anfiyanovych Shakhlin	Soviet Union	13
Ole Einar Bjrndalen	Norway	13
Dara Grace Torres (-Hoffman, -Minas)	United States	12
Paavo Johannes Nurmi	Finland	12
Natalie Anne Coughlin (-Hall)	United States	12

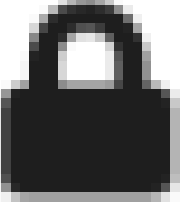
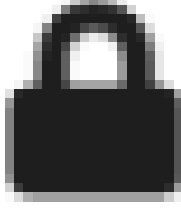
Expected Output: Query 12 (Showing few records only)

Problem Statement: Write a SQL query to fetch the top 5 most successful countries in olympics.
(Success is defined by no of medals won)

region character varying 	total_medals bigint 	rnk bigint 
USA	5637	1
Russia	3947	2
Germany	3756	3
UK	2068	4
France	1777	5

Expected Output: Query 13

Problem Statement: Write a SQL query to list down the total gold, silver and bronze medals won by each country.

country character varying 	gold integer 	silver integer 	bronze integer 
USA	2638	1641	1358
Russia	1599	1170	1178
Germany	1301	1195	1260
UK	678	739	651
Italy	575	531	531
France	501	610	666
Sweden	479	522	535
Canada	463	438	451
Hungary	432	332	371

Expected Output: Query 14 (Showing few records only)

Problem Statement: Write a SQL query to list down the total gold, silver and bronze medals won by each country corresponding to each olympic games.

games text	country text	gold integer	silver integer	bronze integer
1896 Summer	Australia	2	0	1
1896 Summer	Austria	2	1	2
1896 Summer	Denmark	1	2	3
1896 Summer	France	5	4	2
1896 Summer	Germany	25	5	2
1896 Summer	Greece	10	18	20
1896 Summer	Hungary	2	1	3
1896 Summer	Switzerland	2	0	1
1896 Summer	UK	3	3	3

Expected Output: Query 15 (Showing few records only)

Problem Statement: Write SQL query to display for each Olympic Games, which country won the highest gold, silver and bronze medals.

games text	max_gold text	max_silver text	max_bronze text
1896 Summer	Germany - 25	Greece - 18	Greece - 20
1900 Summer	UK - 59	France - 101	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

Expected Output: Query 16
(Showing few records only)

Problem Statement: Similar to the previous query, identify during each Olympic Games, which country won the highest gold, silver and bronze medals. Along with this, identify also the country with the most medals in each olympic games.

games text	max_gold text	max_silver text	max_bronze text	max_medals text
1896 Summer	Germany - 25	Greece - 18	Greece - 20	Greece - 48
1900 Summer	UK - 59	France - 101	France - 82	France - 235
1904 Summer	USA - 128	USA - 141	USA - 125	USA - 394
1906 Summer	Greece - 24	Greece - 48	Greece - 30	Greece - 102
1908 Summer	UK - 147	UK - 131	UK - 90	UK - 368
1912 Summer	Sweden - 103	UK - 64	UK - 59	Sweden - 190
1920 Summer	USA - 111	France - 71	Belgium - 66	USA - 194
1924 Summer	USA - 97	France - 51	USA - 49	USA - 182
1924 Winter	UK - 16	USA - 10	UK - 11	UK - 31

Expected Output: Query 17
(Showing few records only)

Problem Statement: Write a SQL Query to fetch details of countries which have won silver or bronze medal but never won a gold medal.

country character varying	gold integer	silver integer	bronze integer
Paraguay	0	0	17
Montenegro	0	0	14
Fiji	0	0	13
Namibia	0	0	4
Afghanistan	0	0	2
Kuwait	0	0	2
Sri Lanka	0	0	2
Tanzania	0	0	2
Curacao	0	0	1



Expected Output: Query 18 (Showing few records only)

Problem Statement: Write SQL Query to return the sport which has won India the highest no of medals.

<div>sport</div> <div>character varying</div> <div></div>	<div>total_medals</div> <div>bigint</div> <div></div>
Hockey	173

Expected Output: Query 19

Problem Statement: Write an SQL Query to fetch details of all Olympic Games where India won medal(s) in hockey.

team character varying 	sport character varying 	games character varying 	total_medals bigint 
India	Hockey	1948 Summer	20
India	Hockey	1936 Summer	19
India	Hockey	1956 Summer	17
India	Hockey	1980 Summer	16
India	Hockey	1968 Summer	16
India	Hockey	1932 Summer	15
India	Hockey	1964 Summer	15
India	Hockey	1928 Summer	14
India	Hockey	1972 Summer	14

Expected Output: Query 20