EXPLORATORÝ DATA ANALÝSIS ON OLÝMPICS HISTORÝ [1896 – 2016]



INTRODUCTION

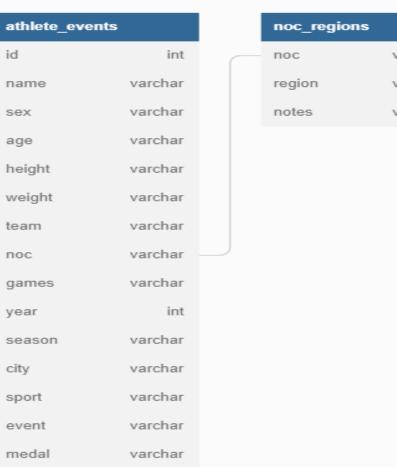
The modern Olympic games is a grand international sporting events featuring winter and summer sports competitions in which thousands of athletes and countries participate in variety of sports events. The event is held ever four years period and this exploratory analysis using SQL (structured Query Language), is to derive information about the Olympics games over the years.

The dataset contains information on the games from 1896 – 2016 and is comprised of two files: (a)athlete_events (b)noc_regions.

Athlete_events being the major file used for the analysis is made up of 271,116 rows and 15 columns while the Noc_regions contains three columns of data.

Tool Used for the Analysis: SQLite Studio

TABLES



varchar varchar varchar



EXPLORING THE DATASET

To analyze and derive information from the dataset, a few questions (source: techtfq.com & some from my own curiosity) will be asked to help know what queries to write. The questions and queries include the following:

1.How many Olympics games have been held between the given date



2.List down all Olympics games held so far



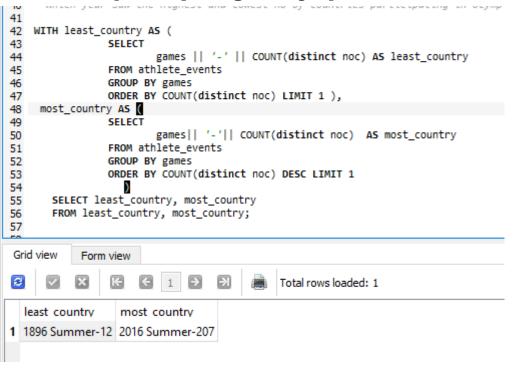
3. How many teams participated in the Olympics

```
25
26 SELECT COUNT( DISTINCT team) AS Team
27 FROM athlete_events;
28
29
30

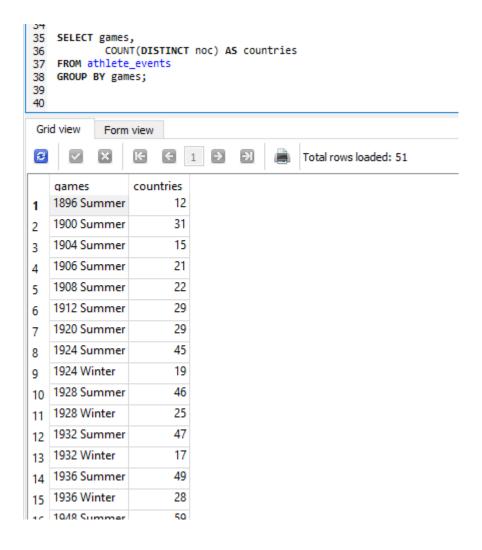
Grid view Form view

Team
1 1184
```

5. Which year saw the highest & lowest no of countries participating in Olympics



4.The total no of nations who participated in each Olympics games



6. Which Nation has participated in all the Olympics

```
WITH total_participants AS (
64
                SELECT region,
65
                        COUNT(DISTINCT games) AS total games,
66
                        DENSE_RANK() OVER(ORDER BY COUNT(distinct games)DESC) AS ranking
67
                FROM athlete_events ae
                LEFT JOIN noc regions nr
68
69
                 USING(noc)
70
                GROUP BY region
71
72 SELECT region, total games
73 FROM total_participants
74 WHERE ranking = 1;
75
Grid view
            Form view
                                            Total rows loaded: 4
             total games
  region
1 UK
                        51
                        51
2 Switzerland
                        51
3 Italy
                       51
4 France
```

7.Identify the sports which was played in all summer Olympics

```
WITH all_summer_game AS (
82
                SELECT sport,
83
                         (SELECT COUNT(distinct games)
84
                            FROM athlete_events WHERE season = 'Summer') AS summer_games,
85
               COUNT(distinct games) AS total games
                FROM athlete events
86
87
                 GROUP BY sport
88
     SELECT sport,
90
            summer games,
91
            total games
     FROM all summer game
      WHERE summer games = total games ;
94
 Grid view
            Form view
8
                                             Total rows loaded: 5
   sport
              summer games | total games
1 Athletics
                            29
                                         29
2 Cycling
                            29
                                         29
3 Fencing
                            29
                                         29
4 Gymnastics
                            29
                                         29
                           29
5 Swimming
                                         29
```

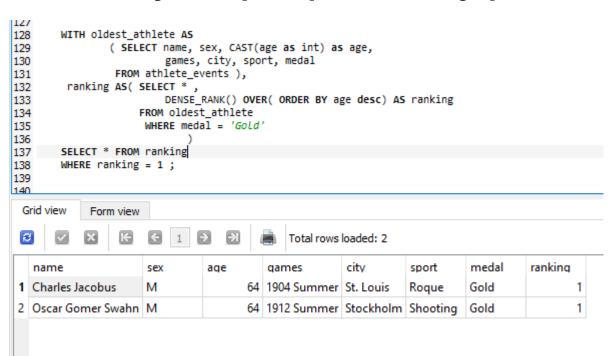
8. Which sports were played just once in the Olympics

```
WITH one_olympic_game AS (
100
101
          SELECT sport,
              COUNT(DISTINCT games) AS total_games,
102
103
              DENSE_RANK() OVER(ORDER BY count(distinct games) ) AS ranking
104
          FROM athlete_events
105
         GROUP BY sport
106
107
     SELECT sport, total_games, games
      FROM one_olympic_game
      WHERE ranking = 1
110
111
     ORDER BY sport;
112
113
114
             Form view
                                             Total rows loaded: 10
                     total games
    sport
                                  games
   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
10 Rugby Sevens
                                1 2016 Summer
```

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

```
117
118
    SELECT games,
119
             COUNT(DISTINCT sport) AS no_of_sports
    FROM athlete events
    GROUP BY games ;
122
123
 Grid view
            Form view
                 ĸ
                     € 1 →
                                            Total rows loaded: 51
                 no of sports
    games
   1896 Summer
    1900 Summer
                           20
    1904 Summer
                            18
    1906 Summer
                           13
    1908 Summer
                           24
                           17
    1912 Summer
    1920 Summer
                           25
    1924 Summer
                           20
                            10
    1924 Winter
10 1928 Summer
                           17
                            8
11 1928 Winter
 12 1932 Summer
                            18
 . - 1022 Window
```

10. How many teams participated in the Olympics



Oldest athlete to participate in the Olympics

```
142
    SELECT name, age, sport, games, medal
144 FROM athlete_events
     WHERE age <> 'NA'
146
     ORDER BY age DESC LIMIT 1;
147
148
149
 Grid view
            Form view
                                              Total rows loaded: 1
   name
                                                      games
                                                                    medal
                           age
                                     sport
                                     Art Competitions 1928 Summer NA
John Quincy Adams Ward 97
```

11.The total no of nations who participated in each Olympics games

```
153
        WITH Females AS (
154
            SELECT COUNT(sex) AS Females FROM athlete_events
155
                        WHERE Sex = (F'),
156
         Males AS (SELECT COUNT(Sex) AS Males FROM athlete events
157
                               WHERE sex = 'M'
158
159
                    --concatination of values
160
        SELECT 1|| ': ' || ROUND(CAST(Males as Float)/CAST(Females as Float),2) AS gender_ratio
161
        FROM Females, Males;
162
163
             Form view
                                              Total rows loaded: 1
   gender ratio
 1 1:2.64
```

12.Top 5 athlete who have won the most gold medal

```
WITH top_5_gold_winners AS (
169
                 SELECT name ,
170
                          COUNT(medal) AS total medal,
171
                          DENSE RANK() OVER(ORDER BY COUNT(medal) DESC) AS ranking
                 FROM athlete_events
172
173
                 WHERE medal = 'Gold'
174
                 GROUP BY name
175
        SELECT name, total medal
177
        FROM top_5_gold_winners
        WHERE ranking between 1 and 5;
179
 Grid view
             Form view
                                               Total rows loaded: 18
                                                   total medal
    name

    Michael Fred Phelps, II

                                                             23
    Raymond Clarence "Ray" Ewry
                                                             10
    Paavo Johannes Nurmi
                                                              9
   Mark Andrew Spitz
```

13.Top 5 athletes who have won the most medals [gold/silver/bronze]

Grid view

Form view

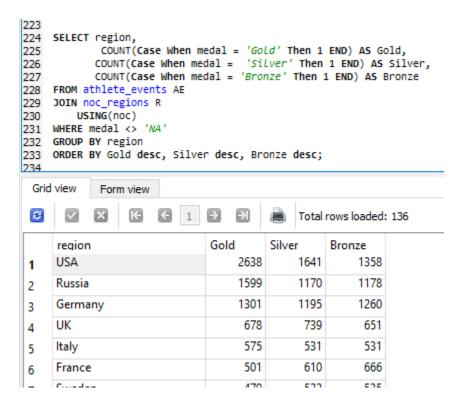
```
WITH Top 5 athletes AS (
                 SELECT name, region, team, sport,
187
                         COUNT(medal) AS total_medal,
188
                         DENSE_RANK() OVER(ORDER BY count(medal) DESC) AS ranking
                 FROM athlete events AE
190
                 JOIN noc_regions R
191
                     USING(noc)
192
                  WHERE medal <> 'NA'
193
                 GROUP BY name, region
194
195
        SELECT name, region, sport, total_medal
        FROM Top_5_athletes WHERE ranking <= 5;
197
100
```

E		otal rows load	ded: 15	
	name	region	sport	total medal
1	Michael Fred Phelps, II	USA	Swimming	28
2	Larysa Semenivna Latynina (Diriy-)	Russia	Gymnastics	18
3	Nikolay Yefimovich Andrianov	Russia	Gymnastics	15
4	Borys Anfiyanovych Shakhlin	Russia	Gymnastics	13
5	Edoardo Mangiarotti	Italy	Fencing	13
6	Ole Einar Bjrndalen	Norway	Biathlon	13
7	Takashi Ono	Japan	Gymnastics	13
8	Aleksey Yuryevich Nemov	Russia	Gymnastics	12
	DO DECEMBER 1			40

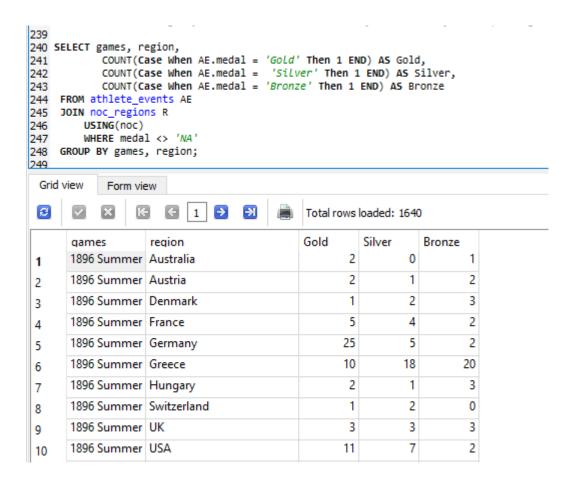
14.Fetch the top 5 most successful country in Olympics. Success is defined by the no of medals won in total.

```
WITH most successful AS (
204
                 SELECT R.region AS country,
205
                         COUNT(medal) AS no_of_medals,
                        DENSE_RANK() OVER(ORDER BY count(medal) DESC) AS ranking
206
207
                 FROM athlete_events E
208
                 JOIN noc_regions R
209
                     USING(noc)
210
                 WHERE medal <> 'NA'
211
                 GROUP BY R.region
212
213 SELECT country, no_of_medals
214 FROM most_successful
215 WHERE ranking <= 5;</pre>
216
 Grid view
            Form view
 3
                                        Total rows loaded: 5
            no of medals
   country
1 USA
                      5637
2 Russia
                      3947
3 Germany
                      3756
 4 UK
                      2068
5 France
                      1777
```

15.List the Total gold, silver & bronze won by each country.



16.List the total gold, silver & bronze medals won by each country corresponding to each Olympic game.



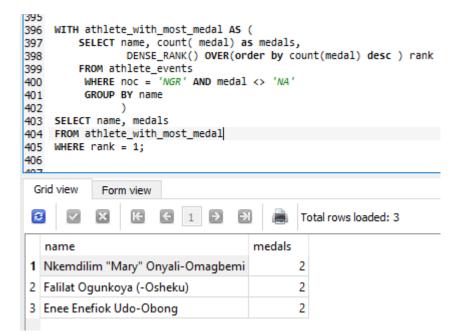
17.Identify which country won the most gold, most silver, & most bronze medals in each Olympic games.

```
254
255
        WITH Temp as (
256
                    SELECT games, region,
257
                            COUNT(Case When medal = 'Gold' Then 1 END) AS Gold,
258
                            SUM(Case When medal = 'Silver' Then 1 else 0 END) AS Silver,
259
                            SUM(Case When medal = 'Bronze' Then 1 else 0 END) AS Bronze
260
                    FROM athlete events AE
261
                    JOIN noc_regions R
262
263
                         USING(noc)
                    WHERE medal <> 'NA'
264
265
266
                    GROUP BY games, region
        SELECT distinct games,
267
                (FIRST_VALUE(region) over(partition by games order by Gold desc) | ' - '| FIRST_VALUE(Gold) over(partition by games order by Gold Desc)) AS Max_Gold,
268
                (FIRST_VALUE(region) over(partition by games order by Silver desc) | | ' - '| FIRST_VALUE(Silver) over(partition by games order by Silver Desc)) AS Max_Silver,
269
                (FIRST VALUE(region) over(partition by games order by Bronze desc) | ' - '| FIRST VALUE(Bronze) over(partition by games order by Bronze Desc)) AS Max Bronze
270
        FROM Temp
271
        ORDER BY games;
272
 Grid view
             Form view
                                              Total rows loaded: 51
                  Max Gold
                                Max Silver
                                                   Max Bronze
    games
    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 IJSΔ - 97
                                France - 51
                                                  11SA - 49
```

18.Oldest and youngest Nigerian athletes.



21. Which Nigerian athlete has won the most medal



20. Which sport/event has Nigeria won the most medals.

```
321 WITH Nigeria_most_sport_win AS (
                 SELECT sport,
323
                          count(medal) AS total_medal,
                          DENSE_RANK() OVER(ORDER BY count(medal) DESC) AS ranking
325
                  FROM athlete events AE
326
                  JOIN noc regions R
327
                      USING(noc)
328
                  WHERE region = 'Nigeria' AND medal <> 'NA'
                 GROUP BY sport
330
331 SELECT sport, total medal
332 FROM Nigeria most sport win
333 WHERE ranking = 1;
334
  Grid view
             Form view
                                               Total rows loaded: 1
   sport
              total medal
 1 Football
                       50
```

22. How many medals has Nigeria won

```
411 SELECT count(medal) AS Total_medal
412 FROM athlete_events
413 WHERE noc = 'NGR' AND medal <> 'NA';
414
415

Grid view Form view

Total medal
1 99
```

SOME INSIGHTS FROM THE ANALYSING OF THE DATA INCLUDE;

- A total of 51 games was played between 1896 2016 comprising of 1184 different teams.
- 1896 summer had the least participants of 12 countries while 2016 summer had the most of 207 countries.
- UK, Switzerland, Italy, and France took part in all 51 Olympic games.
- The oldest athlete to take part in the Olympics is John Quincy Adams Ward, 97 years old from Amsterdam.
- Michael Fred Phelps of USA have won the most medal of 28 in total.
- USA is the most successful country, 5637 medals won i.e defining success by the total medal won.
- Nigeria participated in 16 of the Olympic games and won the most medal in the football sport (50 medals). In total they have won 99 medals between 1896 2016.
- Out of the 565 Nigerian athletes, only 96 won medals.

By doing this project, I was able to properly put into use lots of SQL syntax and functions.

I'm officially a big fan of CTEs and GROUP BY

THANK YOU FOR READING