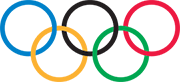
# **OLYMPICS 1986 – 2016**



**DATA SET:**

Our sample population consists of 271,116 records containing data relating to the Olympics from the years 1896 to 2016. The dataset includes Athlete Name, Sex, Age, Height, Weight, Team, NOC (National Olympic Committee designated 3-digit country code), Games, Year, Season (Winter, Summer), City (host city), Sport, Event, and Medal won.

**DATA EVALUTION:**

After performing a count of the data, we identified that out of the 271,116 events, 39,783 medals were awarded. We decided to use medals awarded as the basis for developing the plots included.

**DATA CLEAN UP:**

* In evaluating the number of medals awards, we sliced the data set to include only those records where a medal was won.
* We changed the NOC (National Olympic Committee) code name to “Country”.
* We dropped any NaN that were included in dataset.

**QUESTIONS THAT WERE ANSWERED:**

1. What was the number of male and female participants by year during the *Summer Olympics*from 1896 - 2016 and why did they differ?

Historically there have been more male athletes than female, however by 2016 the gap has narrowed substantially. In analyzing male participation, the following Olympic games have notable decreases, 1932, 1956 and 1980. The 1932 games were affected by the Great Depression. 1956 had boycotts because the Soviet Union invaded Budapest and protests by Egypt, Lebanon and Iraq due of the Israeli invasion of the Sinai Peninsula. 1980 had the US boycott of the games being held in Russia, 65 other countries also boycotted these games, which created the largest boycott in Olympic history.

1. What was the total number of male and female participants during the *Winter Olympics* across all Olympic games and how did they differ?

Historically, there have been more male athletes during the Winter Olympics as well, however the disparity between male and female participants is less in the winter due to the volume of Winter Events compared to Summer Events. In addition, world events affecting the Olympics were more prevalent closer to the Summer Games.

1. How many medals were distributed during each Olympic game? How did the splitting of winter games into their own years affect these results? Did World events have an impact?

Our data shows games were affected by world events, specifically, both World Wars, which cancelled the 1918, 1940 and 1944 games.

We also observed the Winter Games were split from taking place in the same year as the Summer games starting in 1994. Since then, they have been staggered every 4 years.

1. Which country has the most medal count in Gold, Silver and Bronze?

We wanted to observe which country had collected the most medals in Olympic history and the data shows USA, USA, USA!

1. What is the average age of all athletes?

We wanted to see the average age of athletes including oldest and youngest and how they differed for each game. We noticed that the average age seems to have hovered around 24 for the past 15 years.

Oldest - John Quincy Ward...oldest Olympic participant (97) Amsterdam Summer 1928 – Art Competitions

Youngest - Ethnikos Syllogos (10) – Athens Summer 1896 – Men’s Parallel Bars

1. Which Country has the most Gold Medals in an individual *Summer Sport*?

We were curious to see which country had won the most gold medals in a random summer event. We chose Gymnastics and found that Russia has historically dominated the gold medal count in gymnastics. After the breakup of the USSR, the medals won by other countries within the empire were inherited by Russia.

1. Which Country has the most Gold Medals in an individual *Winter Sport*?

We were curious to see which country had won the most gold medals in a random winter event. We chose Alpine Skiing and found that Austria has the most gold medals. Geographically, Austria is located in the heart of the Alps, so everyone skis!!

1. Which sport has given out the most gold medals over time?

Track and Field has more events and can offer more gold medals. Swimming also has many events and therefore more gold medals. However, we found obscure events like Roque, Aeronautics, Basque Pelota.

Personal Reflection of the project

It took us forever to settle on the right dataset. Finally, I stumbled across 120 years of Olympic Sports and it ended up being the perfect dataset to do this project on! My own personal love for sports really pushed me to dive and see what plots I could draw up. We took the whole first project day to clean up the data and planned to meet up again the next day on Skype to progress more. Once we finished that I started looking/studying more about how to use Seaborn, so we could plot our graphs. We spent more time during our project days/Skype to discuss what types of graphs and what story we wanted to tell with our project. On the Saturday before the project was due we spent two more hours finalizing everything on the project. Tuesday was just us discussing in what order did we want to speak during the presentation. I loved working with Irene and Jacob on this assignment. They always communicated with me whenever I had questions or needed to discuss something. Always making time to get on skype to work more on the project as well. From what I hear when I talk to the other classmates I think I got very lucky to be partnered up with them. Hopefully I can be partnered up with them again for project 2!