**DATA ANALYSIS ON OLYMPICS DATA: Report**

**Inspiration:**

I was reading an article(<https://nuunlife.com/blogs/news/a-timeline-of-women-in-the-olympics>): Women in Olympics and there were so many fun facts that the author shared. I was very much intrigued by the study and wanted to do analysis on the Olympics data myself. Also, I have always been interested in sports, thus, it was so much fun to treat the data and find patterns and anomalies.



**INTERESTED FINDINGS AND VISUALISATIONS**:

1. **SOURCE OF DATA:** It is a dataset on the modern Olympic Games, including all the Games from Athens 1896 to Rio 2016.The data have been scraped from www.sports-reference.com in May 2018.
2. **The dataset has 271116 rows and 15 columns.**
3. **The youngest athlete to have participated in Olympics was at the age of 10 and oldest was at the age of 97.**
4. **The minimum height of the athlete was 127 cm and maximum height was 226 cm**
5. **The minimum weight of the athlete was 25 kg and Maximum weight of the athlete was 214kg.**
6. **The name of an athlete that appears the most is Robert Tait McKenzie, about 58 times. Verified this information on Wikipedia as McKenzie participated in every Olympic Summer Games from 1896 in Athens to the 1932 L.A. Olympics**
7. **Out of the two genders, male has dominated the Olympics**
8. **The athletes from United States have participated the most in the Olympics till now.**
9. **Out of 39783 medals awarded, 13372 were gold medals.**
10. **Men's Football as the game and London as city appeared the most in the data.**
11. **Most Gold Medals**

**Graphical user interface, text, application

Description automatically generated**

**Chart, bar chart

Description automatically generated**

1. **A picture containing chart

   Description automatically generated**
2. **Country that won maximum Gold Medals**

Graphical user interface, text, application, Word

Description automatically generated

**Chart, bar chart

Description automatically generated**

1. **Which country has the biggest difference between their summer gold medal counts and winter gold medal counts relative to their total gold medal count?**

Graphical user interface, application

Description automatically generated

**Chart, histogram

Description automatically generated**

1. **Who was the youngest and oldest player to won gold in Olympics?**

Graphical user interface, application

Description automatically generated

**The oldest player to win Olympics gold medal is William Herbert "Bill" Northam in the event "Sailing Mixed 5.5 meters" at the age of 59 and the youngest player to win gold is Fu Mingxia in the event of "Diving Women's Platform".**

### How many athletes have won the gold medals above the age of 50 and below the age of 15? In which sports?[¶](http://localhost:8888/notebooks/Documents/Python%20files/olympics_data_analysis.ipynb#Question-6:-How-many-athletes-have-won-the-gold-medals-above-the-age-of-50-and-below-the-age-of-15?-In-which-sports?)

**Chart, bar chart

Description automatically generated**

**From the above visualisations, we can clearly see that people above age 45 scored gold in sports like sailing, shooting and even horse riding, whereas for athletes below 15, swimming and gymnastics are very popular. The sports are very contrasting in nature as swimming is about dynamics and gymnastics is all about flexibility.**

1. **Which city hosted the Olympics the most?**

**Chart, bar chart

Description automatically generated**

**London is clearly a winner in terms of how many times Olympics were hosted with Athena on second, and, on the other hand, Sydney, Atlanta, and Rio de Janeiro were very close.**

1. **Women Participation in Olympics**

**A group of women in a race

Description automatically generated with low confidence**

**Chart, histogram

Description automatically generated**

1. **Comparison of Women Participation in Summer and Winter Olympics**

**Chart, bar chart

Description automatically generated**

1. **Comparison of Men and Women Participation**

**Chart, bar chart, histogram

Description automatically generated**

### Distribution of Gold Medals among the Age of Athletes

### Chart, histogram Description automatically generatedFrom the plot we can see that most of the athletes that won the gold medals have the age window of 20-30 and 24 is the peak age at which most of the athletes have won the gold medals.

### Distribution of Gold Medals among the Height of Athletes

### Chart, histogram Description automatically generated

### By looking at the distribution plot, There seems to be a pattern of a wave, looks like optimal height to win a gold medal is 170 to 190 cm, however there are few spikes at certain heights.

### Distribution of Gold Medals with the Age, Weight and Height of athletes in 100m men

### Chart, bar chart Description automatically generated

### Gold medal count among countries in 100m

### Chart Description automatically generated

### There is so much more analytics and visualisation techniques that can be used to find more interesting facts and patterns. TO BE CONTINUED……….