Athlete Events (Olympics) EDA Write Up

**Objective:**

Perform exploratory data analysis for the Athlete Events (Olympics) datasets and provide descriptive summaries using plots, tables, and or visualizations.

**Data Sources:**



[National Olympic Committees (NOC) - Olympic Movement (olympics.com)](https://olympics.com/ioc/national-olympic-committees)

**Data:**

ID – Unique Identifier for each athlete

Name – Name of the Athlete

Sex – Gender (M or F)

Age – Age of athlete

Height – Height of athlete (In centimeters)

Weight – Weight of athlete (In kilograms)

Team – Team athlete was on

NOC – National Olympic Committee athlete was a part of

Games – Olympic Games athlete participated in

Year – Year Olympics were held (1896 – 2016)

Season – Season of Olympics (Summer or Winter)

City – Host city for Olympics

Sport – Sport athlete participated in

Event – Event the athlete competed in

Medal – Medal the athlete won (Gold, Silver, Bronze, NaN)

**Data Completeness:**

(Percent of null data)

**A screen shot of a black screen

Description automatically generated**

Additionally, no duplicate entries were found in the dataset.

**Descriptive Statistics:**

The data provided for the Olympics ranges from 1896 to 2016 with the start of the first Winter Games beginning in 1924. The Summer and Winter Olympics were held in the same year up until 1992. Following this the Summer and Winter Olympics were staggered by 2 years. There were several years where the Olympic Games were not held. This includes 1916, 1940, and 1944. Over the course of these Olympic games 135,571 athletes (101,590 Male, 33,981 Female) competed in these games across 66 different sports and 765 events. This dataset shows there have been 230 NOCs that have been recognized with the current number being 206 (Found in datasource).

As the years progressed more sports/events, athletes, and NOCs were added to the Olympics.

(Summer)

A graph of a graph showing the growth of the company's stock market

Description automatically generated with medium confidence

The United States leads the NOCs in total with a count of 5002 and accounts for 21.57% of the Gold Medals won. Michael Fred Phelps II is currently the all time leader in medals with 28.

A graph of blue bars

Description automatically generated

A colorful circle with numbers and lines

Description automatically generated