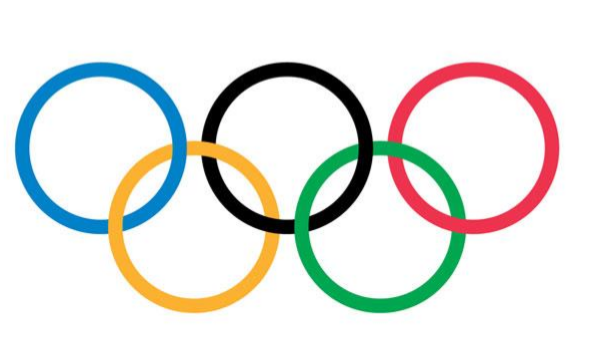
**Olympics Data Analysis Using Python**

The Olympics are leading international sports events featuring summer and winter sports competitions in which thousands of athletes from around the world participate in a variety of competitions. The Olympic Games are considered the world’s foremost sports competition with more than 200 nations participating. By using Python we will analyze this data.



**About Dataset:**

The following dataset is about the Olympics.

This is a historical dataset on the modern Olympic Games, including all the Games from Athens 1896 to Rio 2016.Here 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:

1. **ID** - Unique number for each athlete
2. **Name** - Athlete's name
3. **Sex** - M or F
4. **Age** - Integer
5. **Height** - In cm
6. **Weight** - In kgs
7. **Team** - Team name
8. **NOC** - National Olympic Committee (3-letter code)
9. **Games** - Year and season
10. **Year** - Integer
11. **Season** - Summer or Winter
12. **City** - Host city
13. **Sport** - Sport
14. **Event** - Event
15. **Medal** - Gold, Silver, Bronze, or N

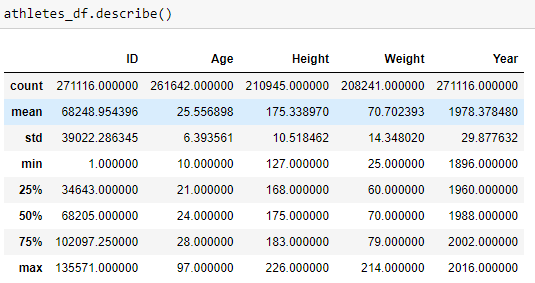
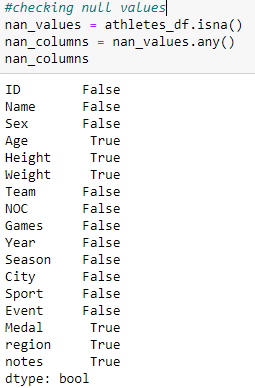
## Libraries Used:

1. **NumPy**: NumPy is for array-processing package.
2. **Pandas**: It is used for analyzing the data,
3. **Matplotlib**: It is a numerical mathematics extension NumPy
4. **Seaborn**: It is used for visualization statistical graphics plotting in Python

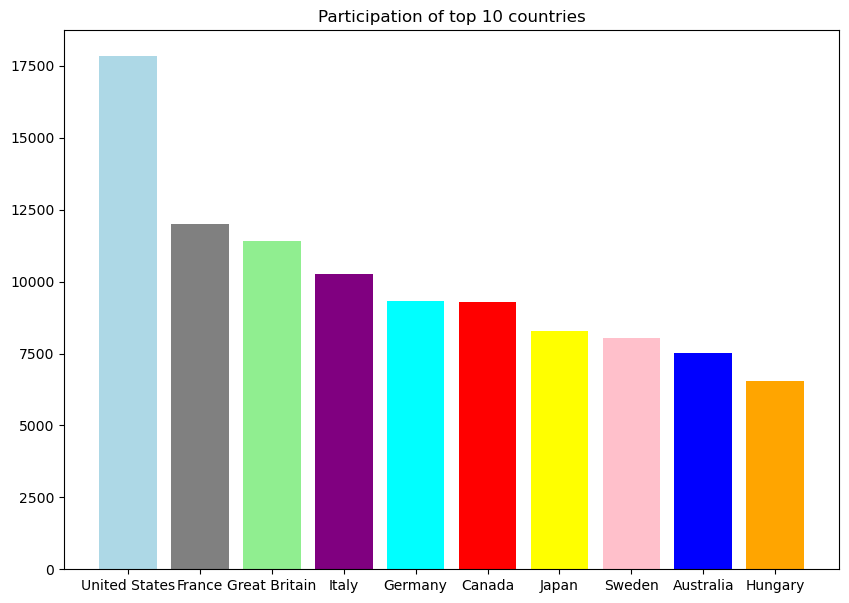
**Analysis and Visualization:**

* **Importing libraries and extract zipped files**
* **Convert dataset to csv**
* **Joining of two dataframes**
* **Age, Height, Weight , Medal, Region have the missing values**

****

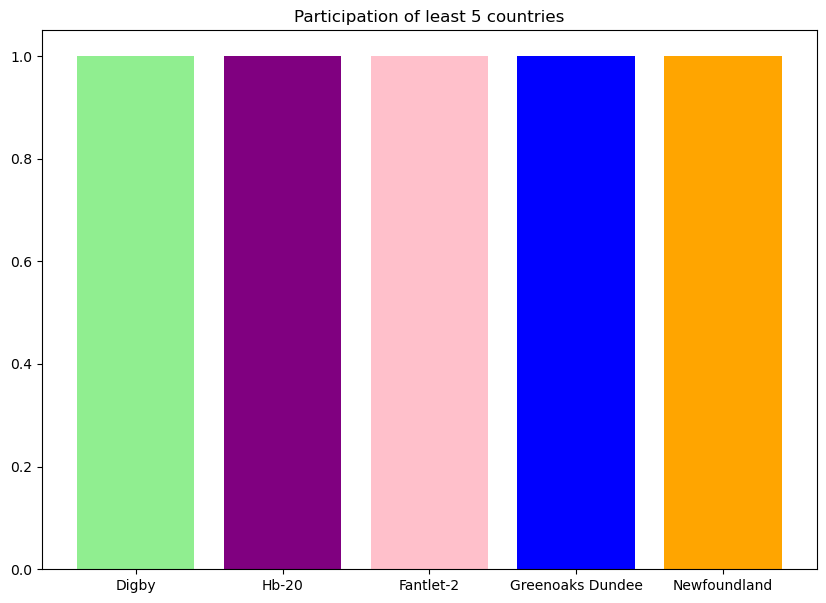
** **

**Top 10 Countries Participating in the Olympics:**

****

**The above graph shows that United States is the most participated country throughout the Olympics**

**Least 5 countries participating in the Olympics:**

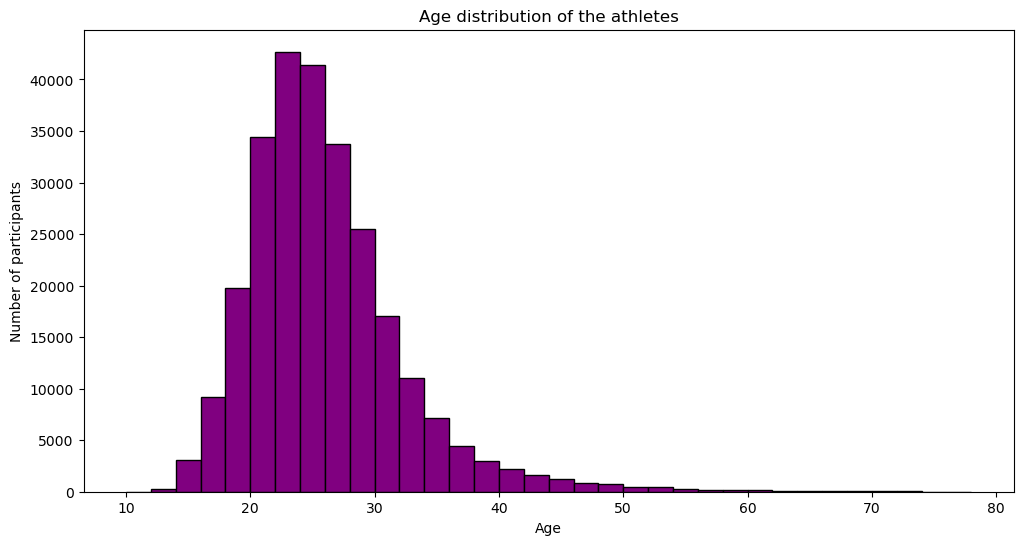
****

**The above graph shows the least 5 countries participating in the Olympics**

**Age**

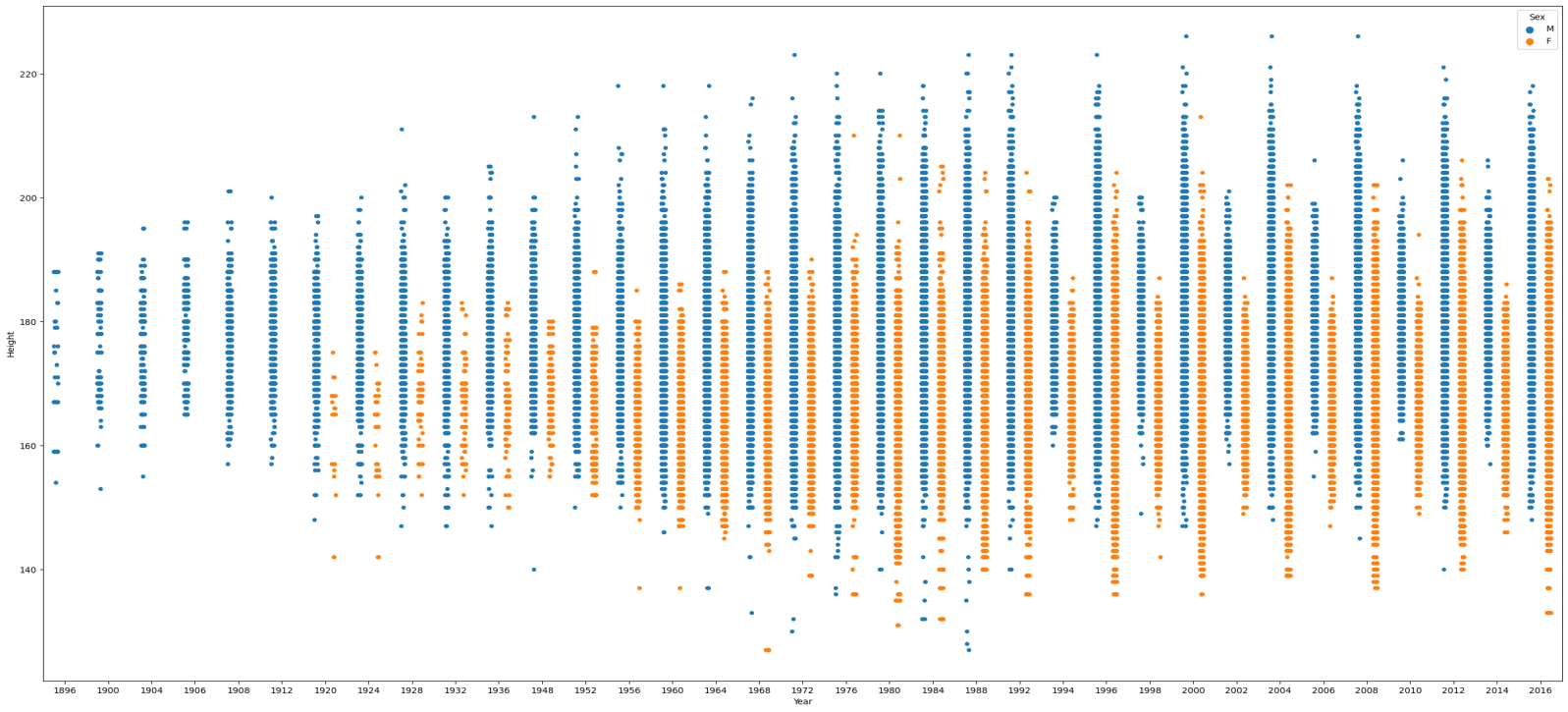
**Height Weight Analysis**

**Age distribution of the athletes:**

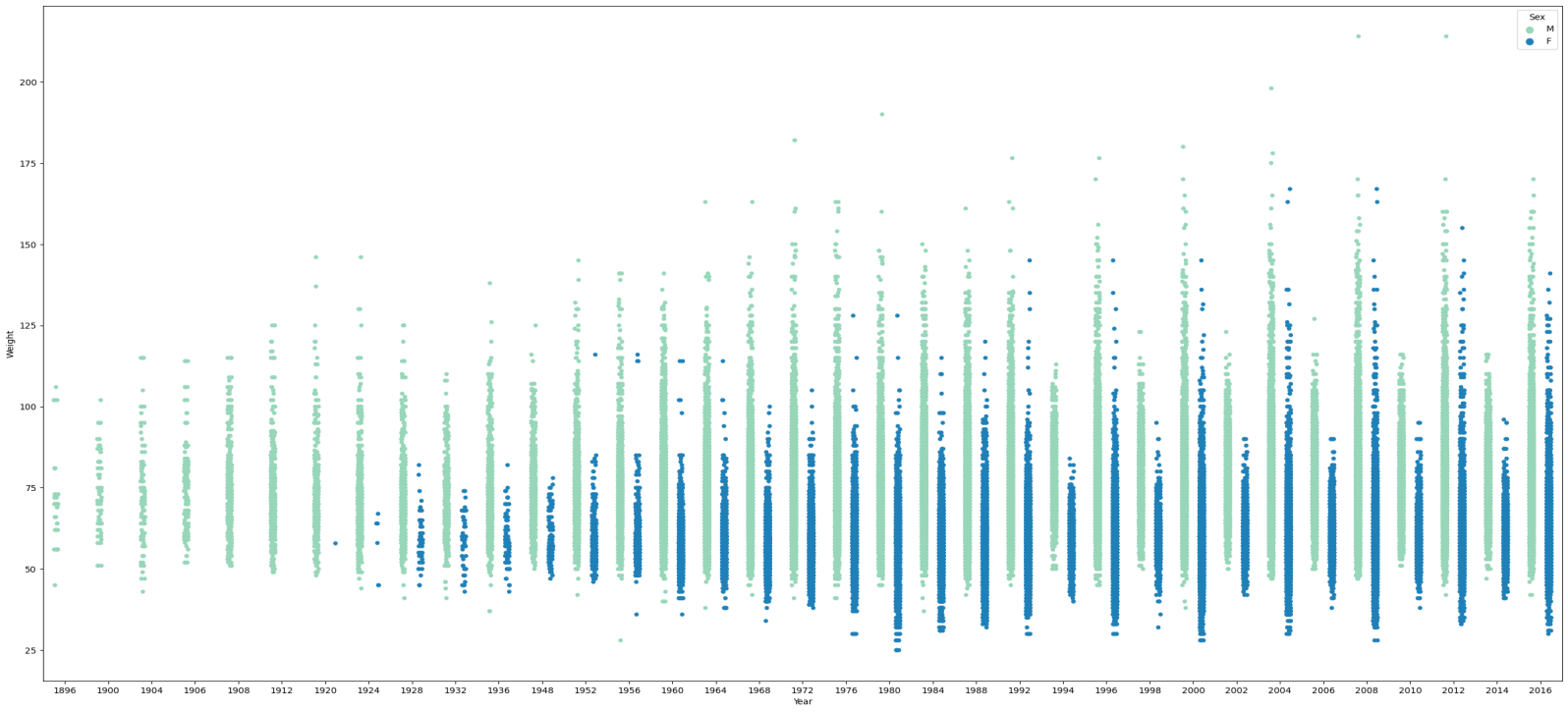
****

**Highest number of participants belongs from the age ranging from 20 -30 years**

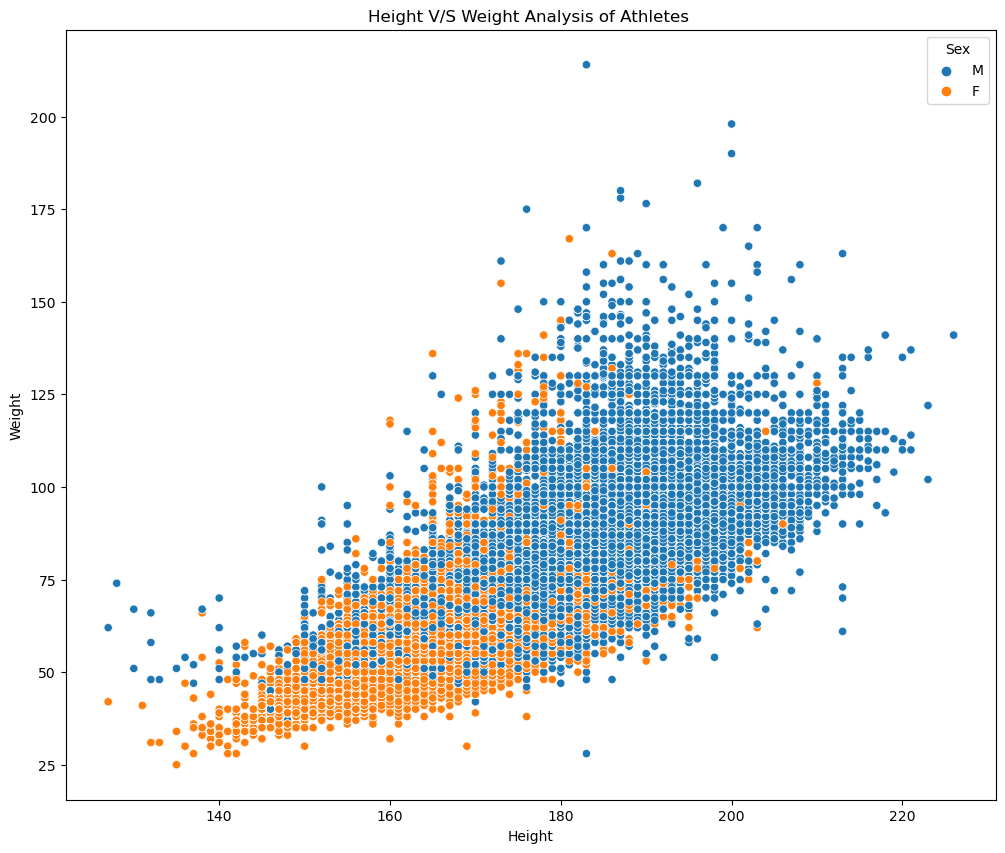
**Height Distribution of Athletes:**

****

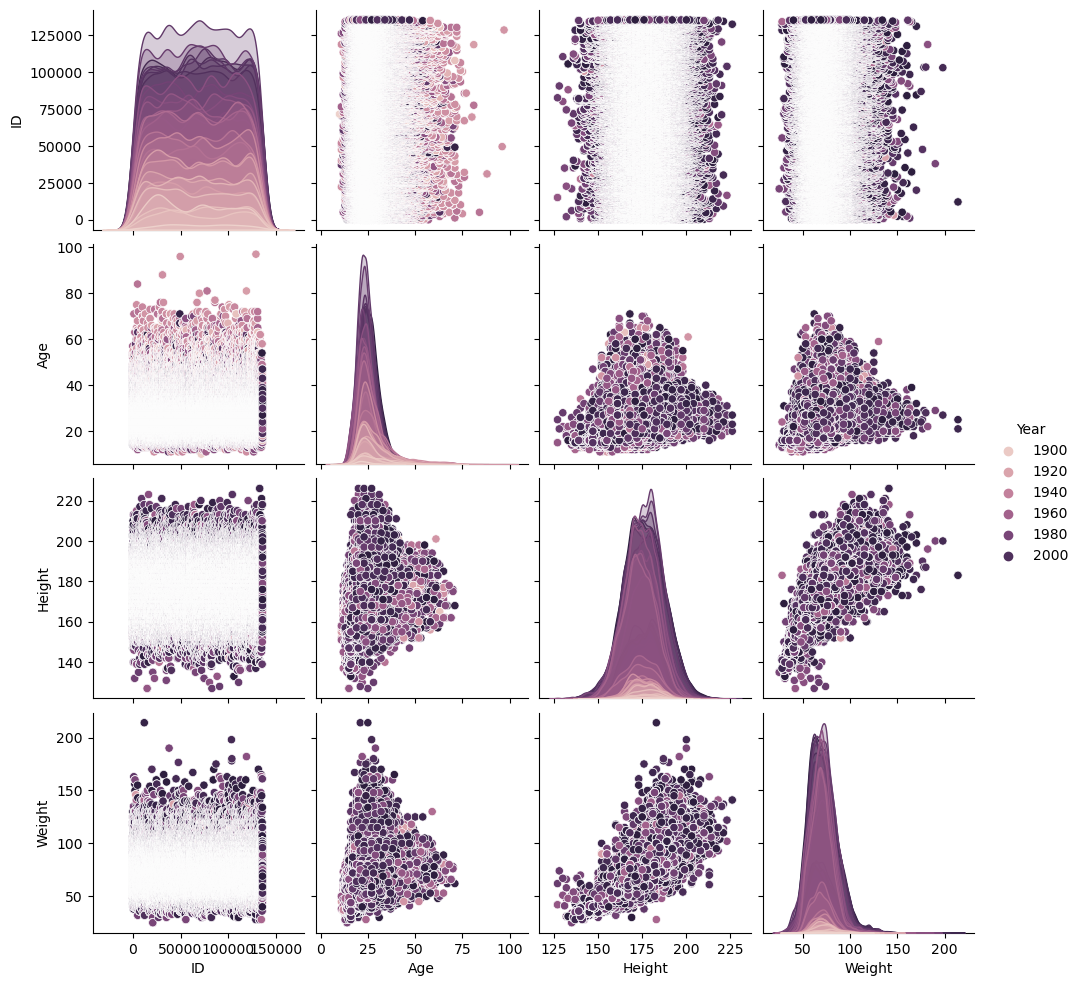
**Weight Distribution of Athletes:**

****

**Height V/S Weight Analysis of Athletes:**

****

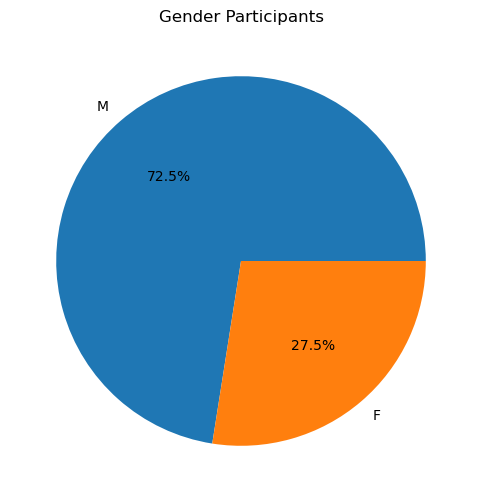
**Above graphs shows the height and weight distribution of the athletes with respect to their gender**

****

# Gender

# Analysis

**Gender Participants:**

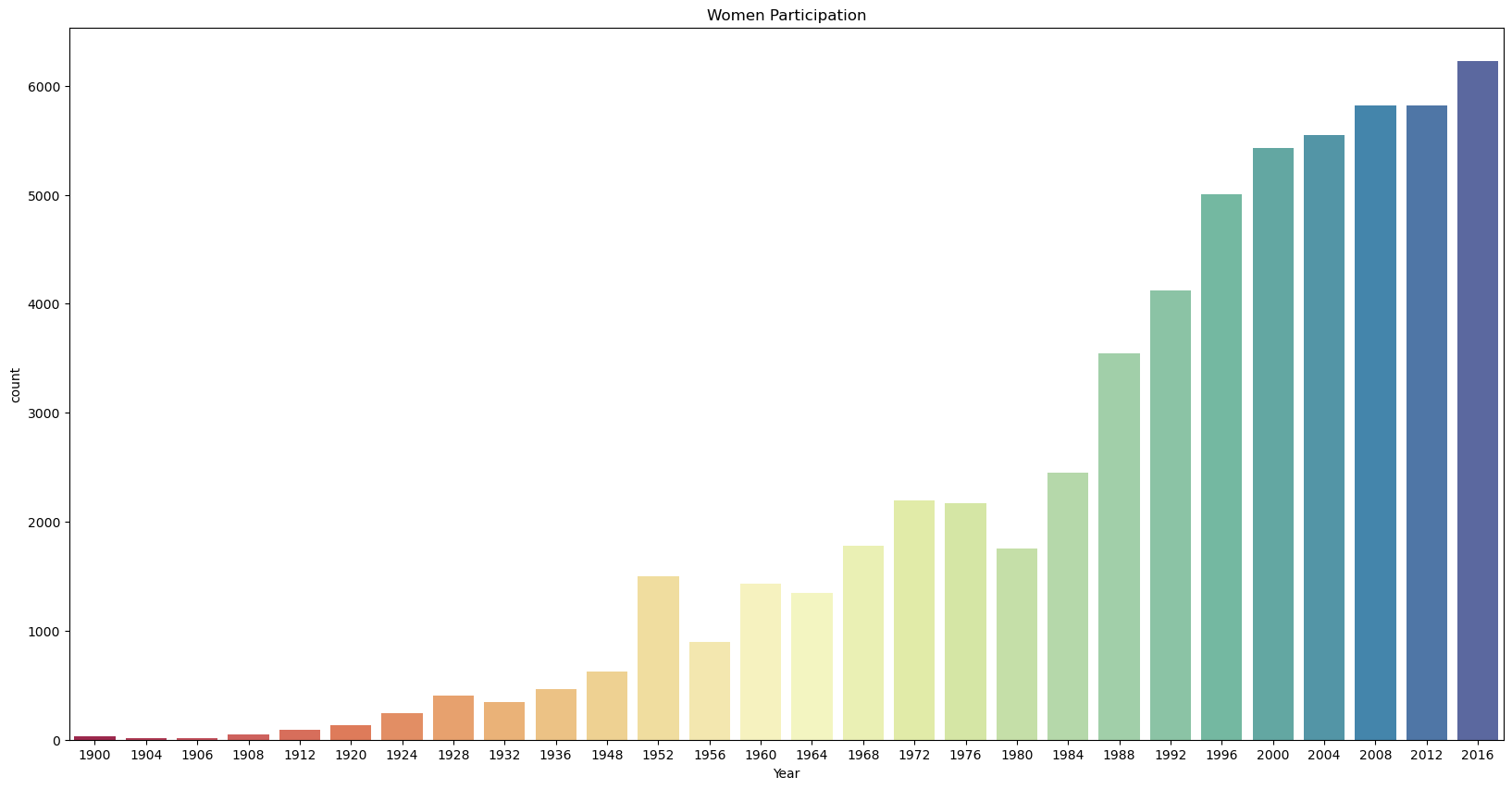
****

**The ratio of male participants is more than females. Since the beginning of the modern Olympic Games, only 27.5% of the total participation has been by females. Here is the count of gender participants:**

**Male -> 196594**

**Female -> 74522**

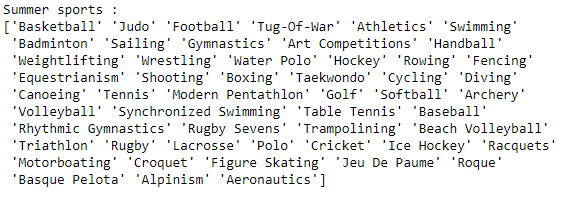
**Women Participation in Olympics Over Past years:**

****

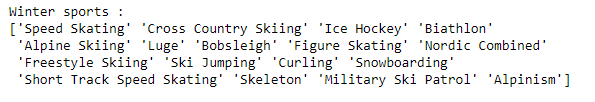
From the above graph it can be seen that over the past years the female participation in the Olympics has been increased.

**Sport Season Analysis**

**Summer Olympics Games:**

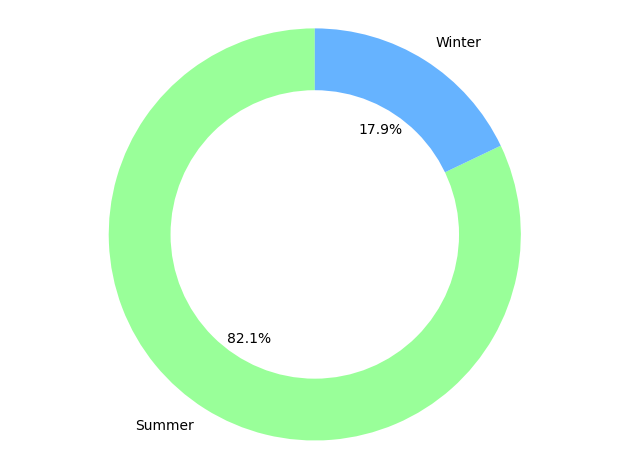
****

**Winter Olympics Games:**

****

**Number of Summer games is more compared to Winter games**

**Athletes Participation across season:**

****

**It can be seen that the number participants in the summer Olympics sports is more compared to the winter Olympics sports as the no. of games is more in Summer compared to Winter**

**Medal Analysis**

**Total no. of medals won by the athletes:**

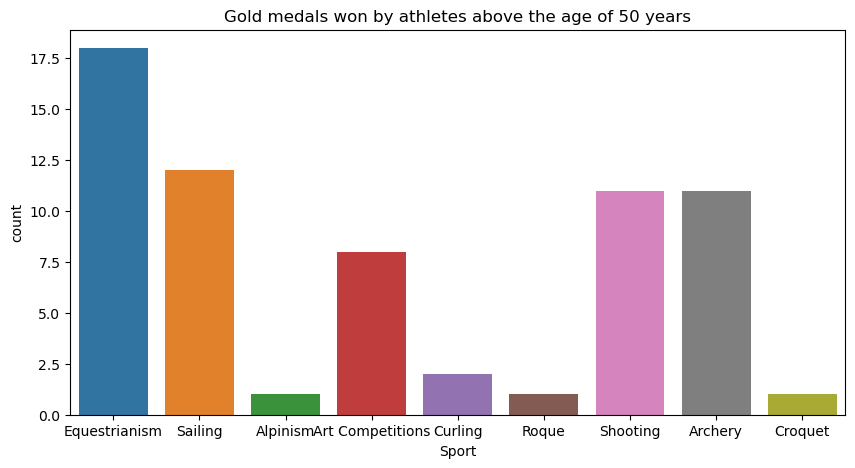
Medal Counts :

Gold 13372

Bronze 13295

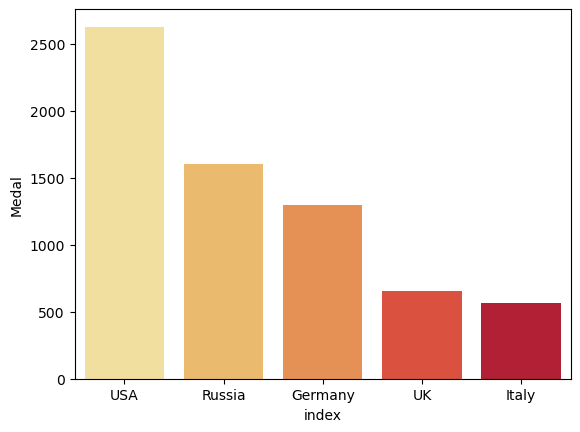
Silver 13116

**Gold medals of athletes above age 50:**

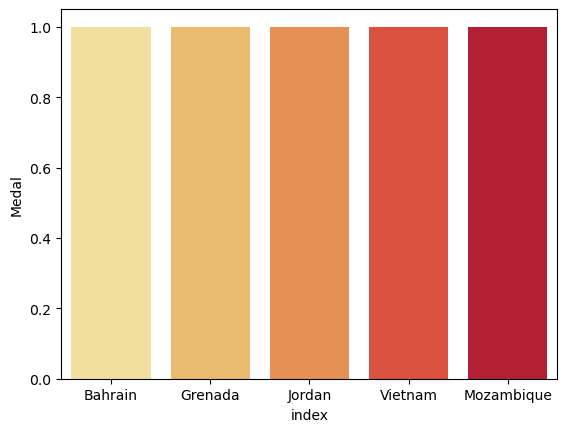
****

**Above graph shows the gold medal won by the athletes age above 50 in the following above games**

**Top 5 countries winning the most number of gold medals**

****

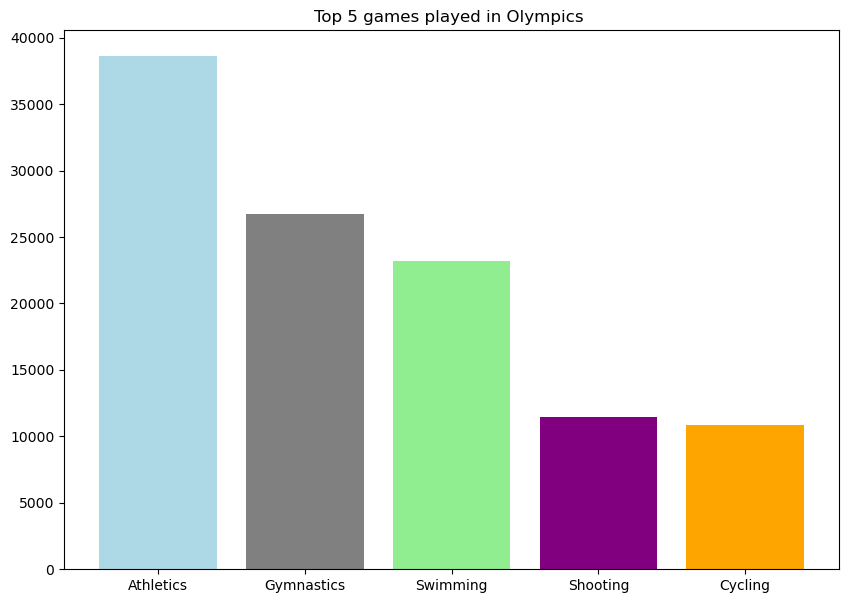
**Least 5 countries winning the most number of gold medals**

****

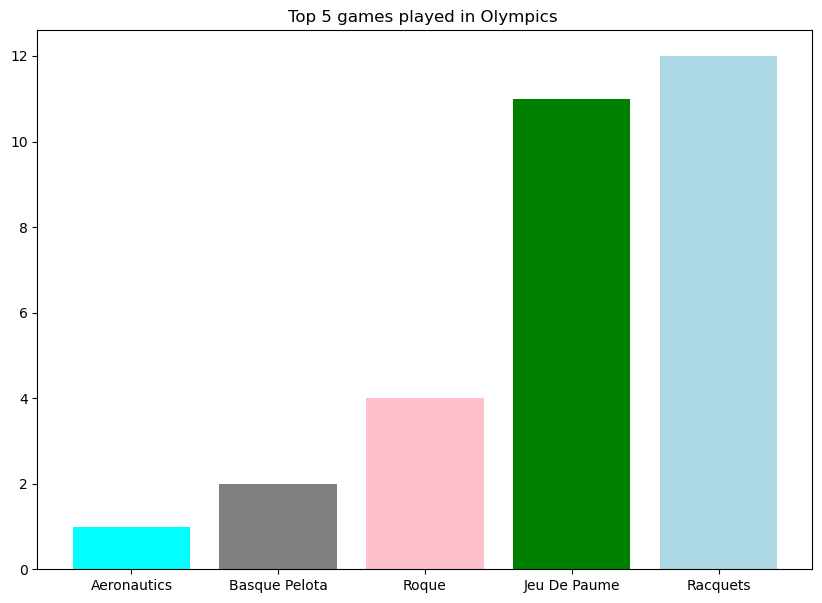
**USA is the most winning country with the gold medals and Bahrain, Greneda, Vietnam with least gold medals in olympics**

**Sports Analysis**

**Top 5 games played in Olympics:**

****

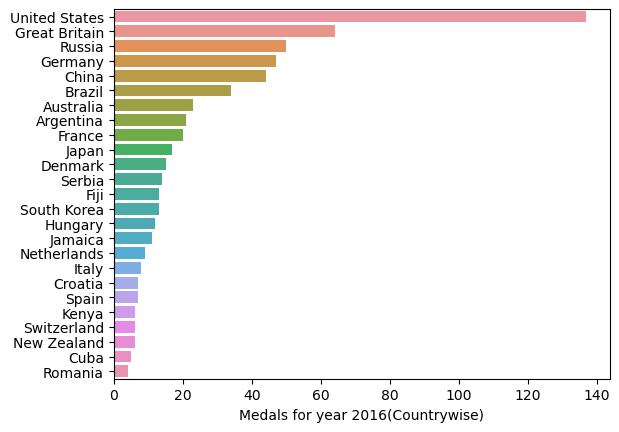
**Least 5 games played in Olympics:**

****

**RIO OLYMPICS:**

**Rio Olympics was played in 2016. The below graph shows the country wise the Medals won.**

**USA has won the most number of medals.**

****

**(2016 is the last year of the given dataset)**

**STATISTICS**

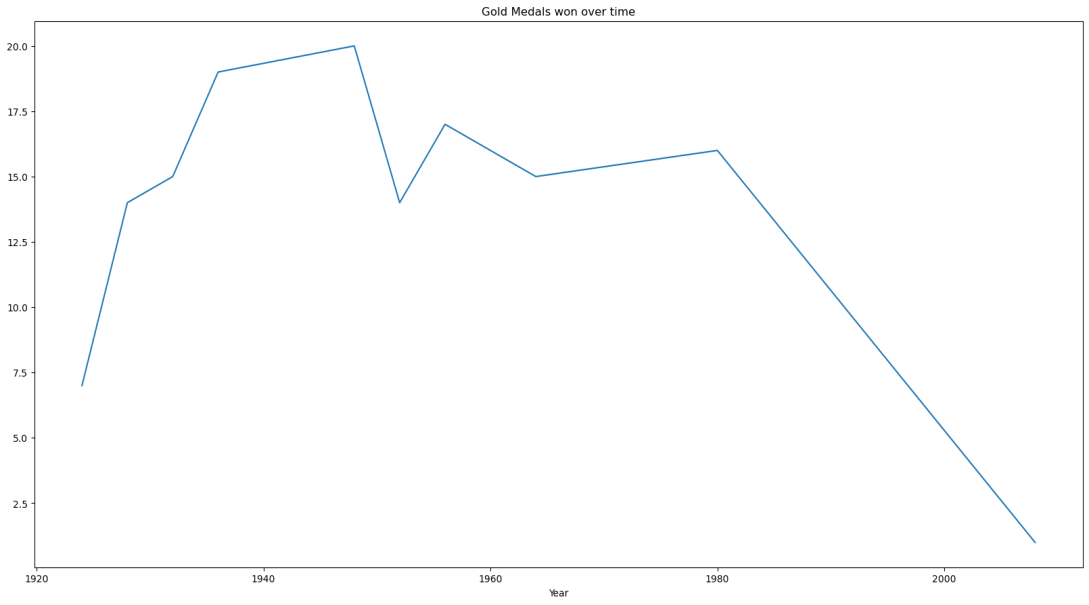
**FOR**

**INDIA**

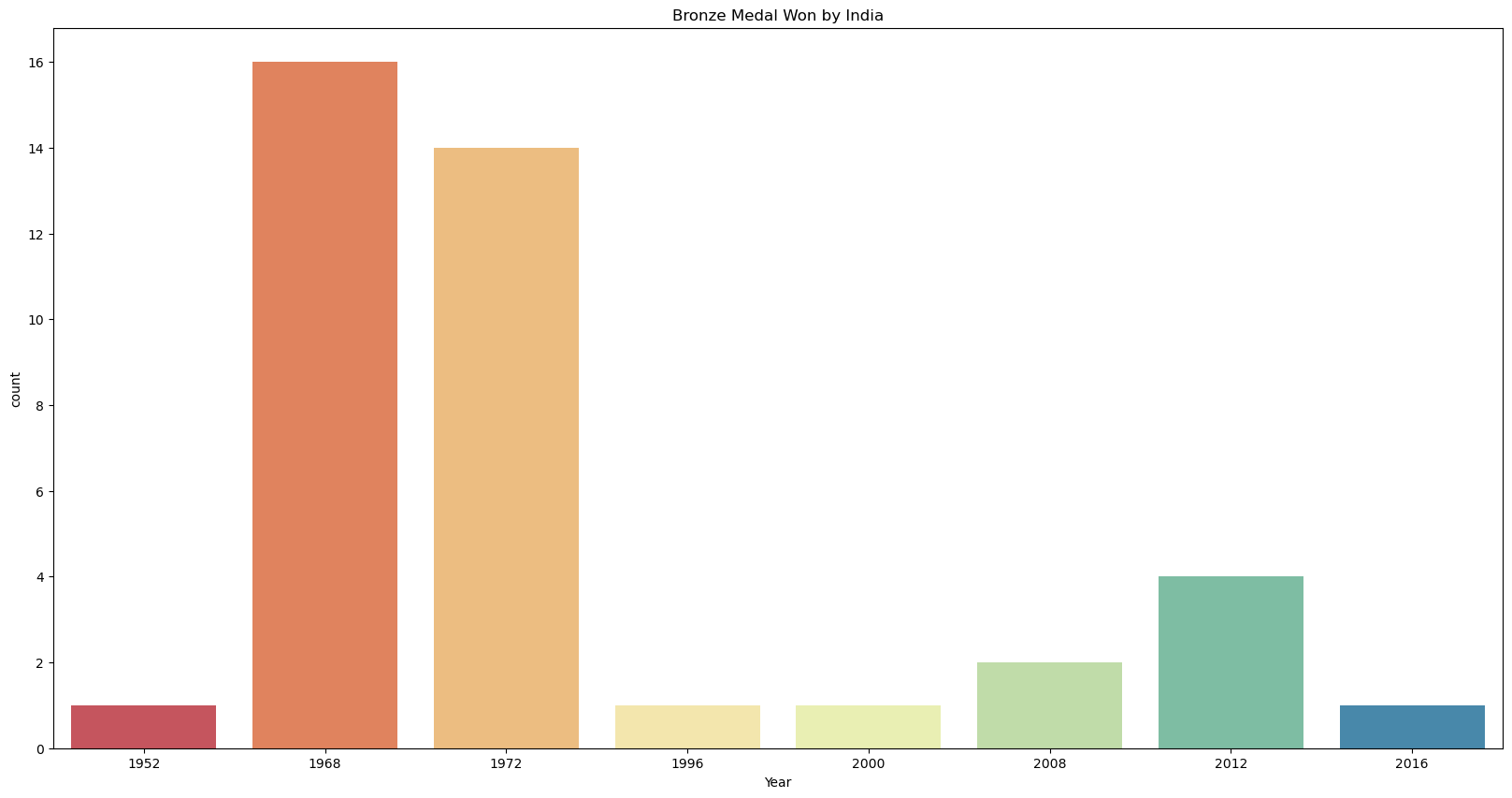
* India has participated from 1900- current
* Total number of all Medals India won 197
* Sports in which India won a Gold are Hockey, Shooting, Alpinism
* India won its first gold medal in Indian Men’s hockey team – Amsterdam 1928

* The person who won the first **individual** **Gold** for India was Abhinav Bindra in 2008 for sports Shooting

**Gold Medal won by India over the past year**



**Bronze Medal won by India over the past year**



**Conclusions:**

We’ve drawn many inferences from the survey. Here’s a summary of a few of them:

* US seems to dominants in terms of participation of maximum gold as well as overall participation in games.
* We observe athletes from the age of 12 till the age of 58 years winning medals.
* Summer Olympics have higher no. of events and sports as compared to the winter Olympics.
* In the history of 120 years of Olympics, Michael Fred Phelps, II has won maximum medals for his country i.e. 28 Medals
* We see a trend that woman participants across the years are in upward trend.