Olympic Dataset Analysis

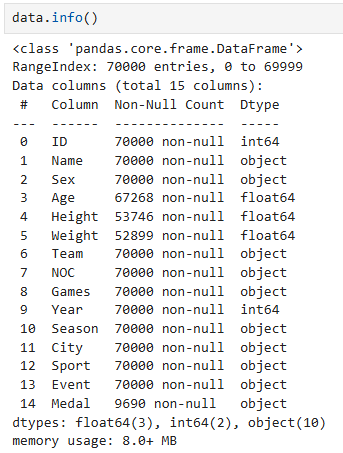
Content of this project

Necessary Libraries for the Project:

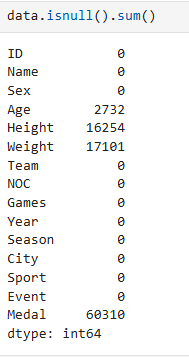
For this project the necessary libraries are Pandas, Numpy, Matplotlib and Seaborn.

Data Information:

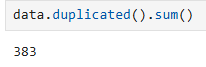
Here in this Olympic dataset there are 70,000 rows and 15 columns.



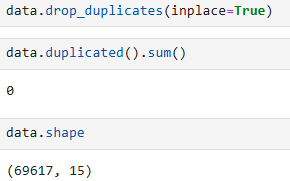
From this code we get the information of the whole dataset. Here in this dataset there are 70000 rows and 15 columns. Here in 15 columns there are 3 types of data. 3 columns contains is float64 datatype then 3 columns contains int64 datatype and 10 columns contains object datatype. After that it took 8 MB of the memory system.



From this code we get the information of the missing values containing according to the columns in the whole dataset.



This function shows how many duplicate values are present in the dataset.

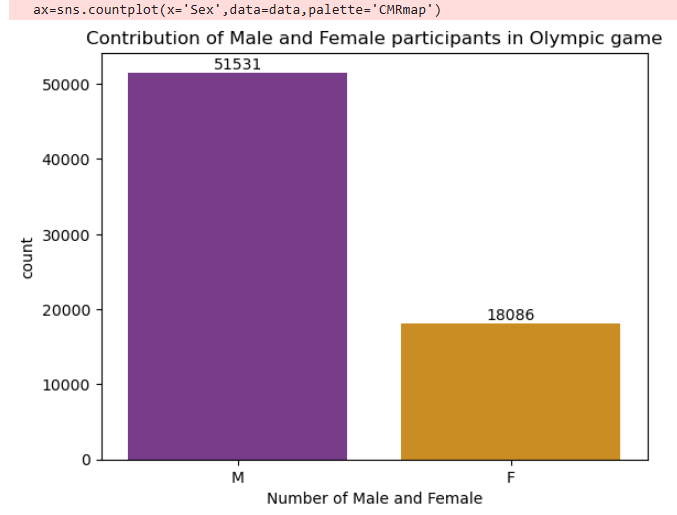


After dropping all the duplicate values the dataset contain 0 duplicated value and the number row became 69617.

Data Analysinig:

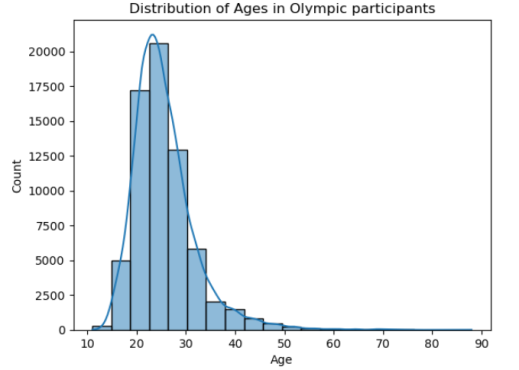
Here in the dataset Olympic game information was from 1896 to 2016.Here in the dataset there is a data of 35 years.

1. Find out the Male and Female participation in the Olympic dataset?



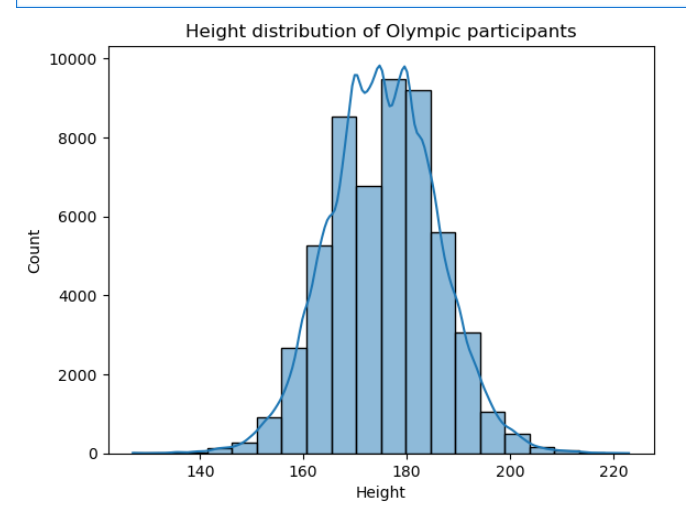
From the bar chart it is seen that from 1896-2016, in 35 years Male participation in Olympic is almost 3 times than the female participation.

1. Which age group was the highest in the Olympic game?



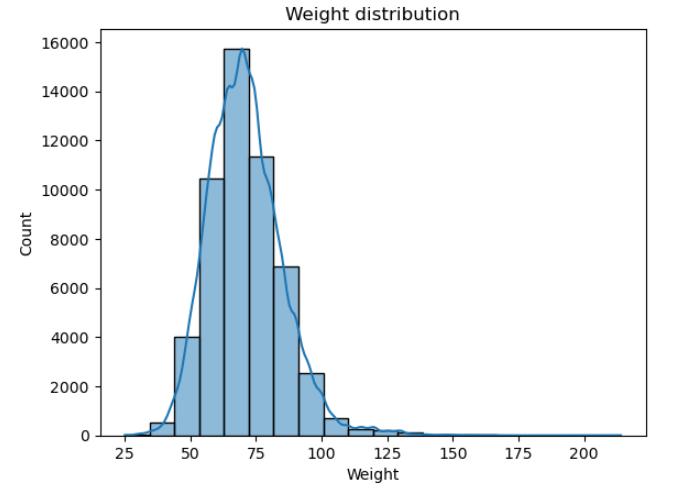
From this Distribution plot, it is seen that most of the participants was from 20-30. Years old age group.

1. Show the Height distribution in the Olympic game



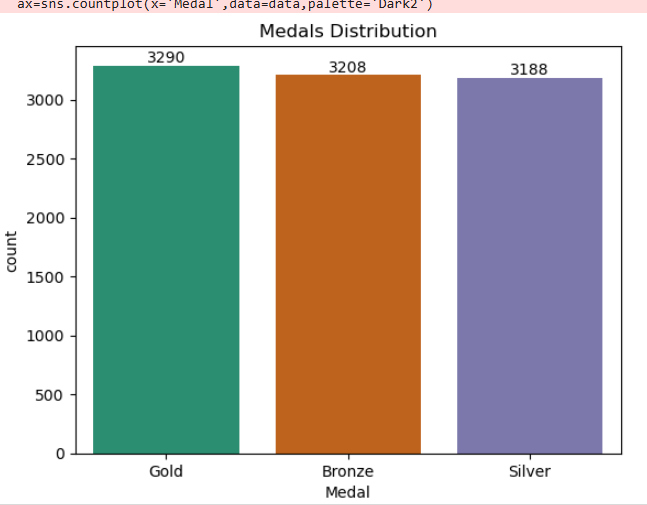
From this distribution plot, we can see that from 53746 participants most of the participant’s height was from 165-185. Here the mean height was 175.50. Here the minimum height was 127 and the maximum height was 223.

1. Show the Weight distribution in the Olympic game



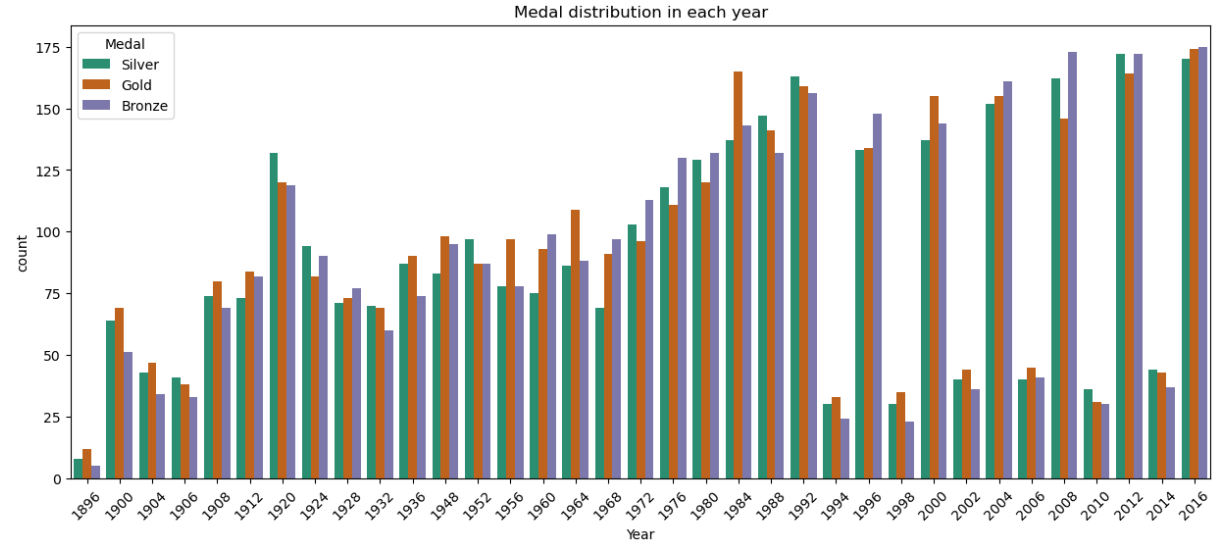
From this distribution plot, average weight of the participants was 70.90. Here the minimum weight was 25 and the maximum height was 214.

1. Which medal got the highest in number in the 35 years of Olympic?

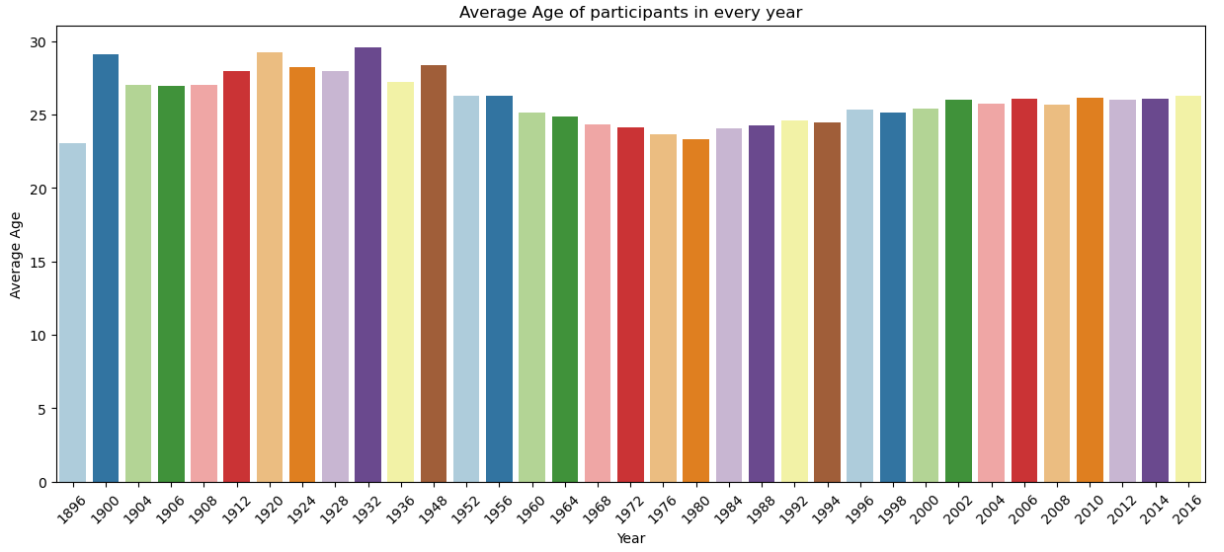


From the bar chart it is seen that, the maximum amount of medal winning was the gold medal and the number was 3290 and the lowest number of medal winning was the silver medal. At the same time 3208 number of medal was Bronze medal which was won in the 35 years Olympic game.

1. In which year medal winning was highest in Olympic game?

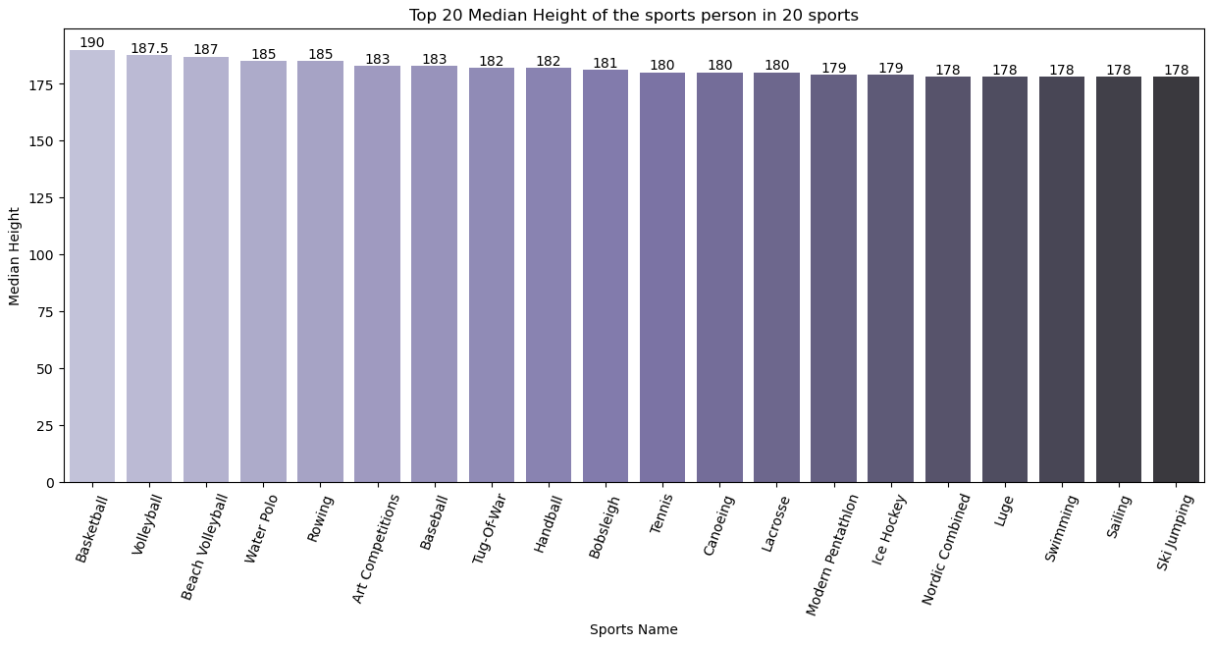
 From this bar chart it is seen that, in 2016 all types of medal (Silver, Gold and Bronze) winning was highest in number. In the beginning year of 1896 this medal winning in number was the lowest. Silver winning in number fluctuated from 1896 to 1968. After 1968 to 1992 this medal winning increased gradually and after 1992 to 2016 this number again shows a fluctuated trend. In terms of gold winning in number, also shows the same trend as silver winning. The highest number of gold won in the year of 2016. In terms of Bronze, highest amount of bronze won in the year of 1896 and the lowest amount of bronze won in the year of 2016. In the year of 2004, 2008, 2012 and 2016 bronze winning was comparatively higher than the other two types of medal.

1. What was the average age in every year in Olympic game?



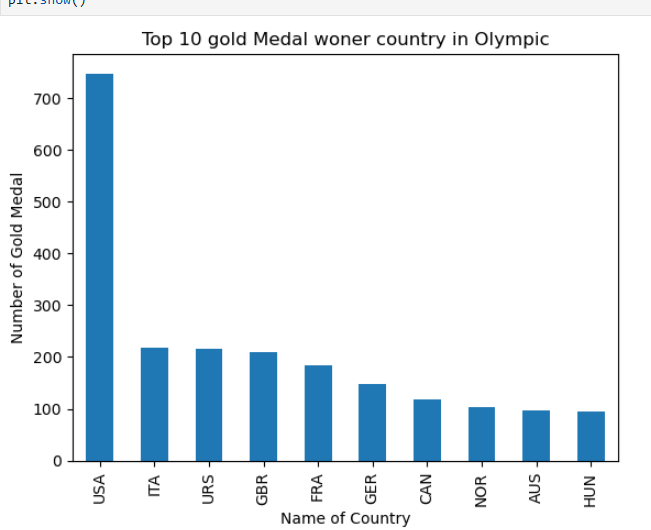
From 1896 to 2016, in 35 years Olympic game, in 1932, the Olympic participants mean age, was highest and it was 32.64 and on the contrary, participants average age was lowest in 1896 and the age 23.03. Average age of the participants was fluctuated from 1896 to 1980. After a very slow increase the mean age of the participants reached at 26.26 at the year of 2016.

1. Find out the top 20 Median Height of sports person of 20 different sports .



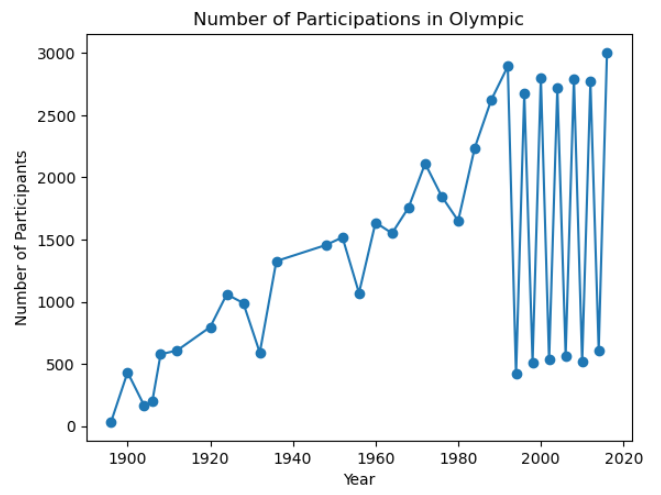
Around 65 types of different games in Olympic, the mean height of the players was 181.66. On the contrary, according to the bar chart, the tallest median height was Basketball and the median height was 190 and at the same time the lowest median height player was from ski jumping where their median height was 178. Whereas, according to the Olympic dataset, the lowest median height was from Gymnastics and their median height was 164.

1. Find out the top 10 medal winner country in Olympics.



According to the bar chart, USA got the highest number of medal in altogether of 35 Olympic games and the total number of medal they got was 747 and the lowest number of medal according to the chart was Hungary and the medal they got was 94.

10. Show the sports participation in every Olympics



According to the line chart, the highest number of sports person was participated in the Olympic in 2016 and the number of player was 3003. On the contrary, the lowest number of player was participated in 1896 and the number of player was only 34. From 1896 to 2016, the number of players fluctuated the whole 35 events.