Analysis of Contributing Factors of Top Performance Countries in Olympic Games

Team Olympic Titan

Objectives:

Our mission is to figure out which country has the greatest performance over past decades and what specific factors contributed to their success, such as their height, weights, or ages in relations to their metal-worthy performance.

We have calculated the ratio of the total number of medals over the number of events they have attended each game as the performance of each country. Noting that one athlete attended in two events and won no medals is counted as no medals and two events.
Ratio = total number of medals / the number of events in each Olympic game

In order to use this information to highlights the significant insights and make recommendations for future events, the first task would be finding out the best performing countries over the years.

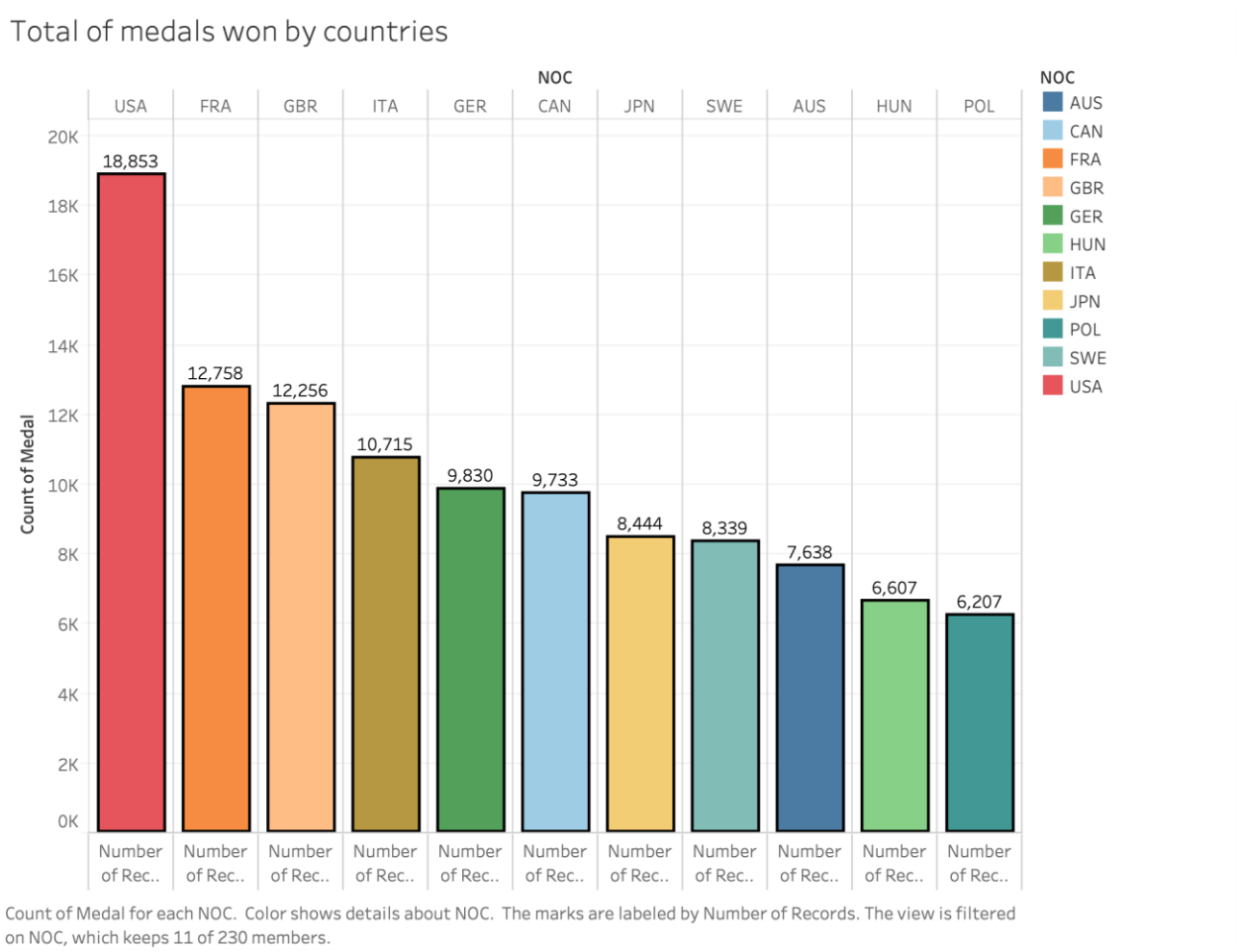
Dataset:

athlete_events.csv, noc_regions.csv

Methodology

- Total number of medals won by each country over the past Olympic Games.

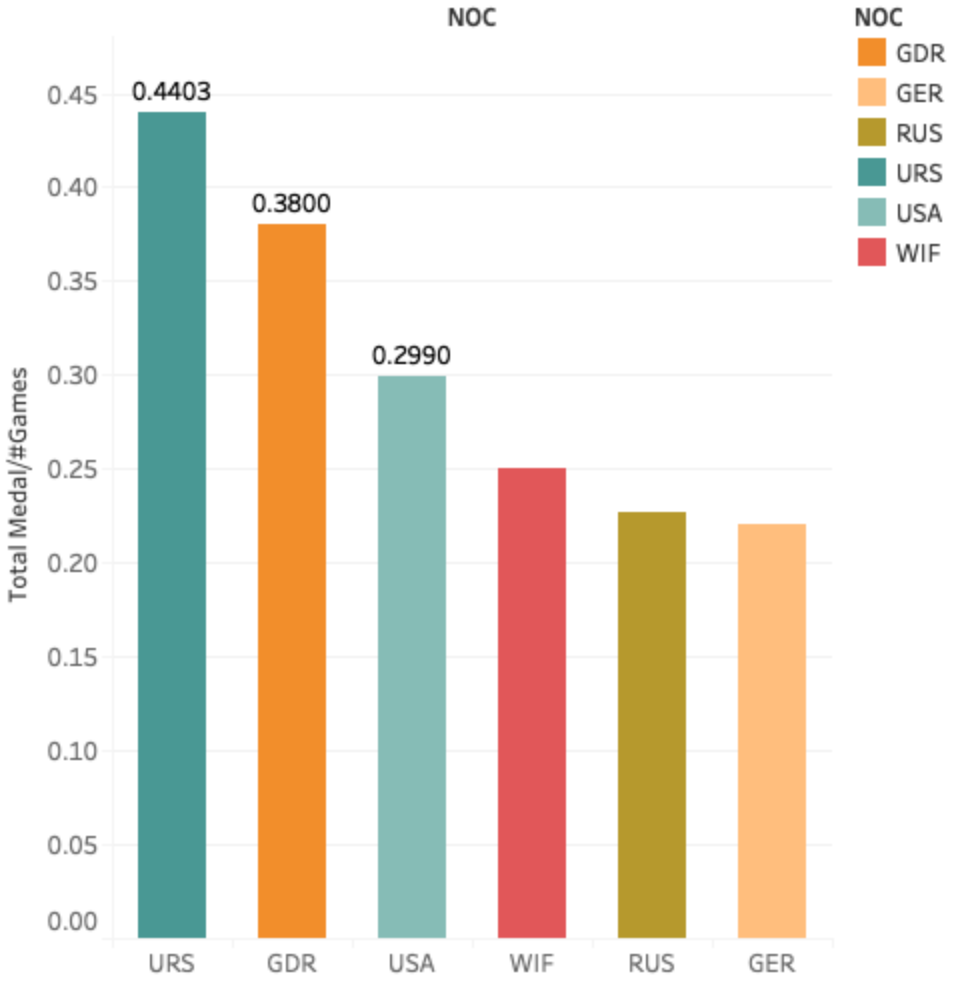
Based on our analysis, the country that has won most medals is USA, which is 18,853 in total. Following is French (FRA) with 12,758, and German (GBR) with 12,256. Since there are some country, for example, USA has attended the Olympic Games since 1896 while other countries may not attend the Games until late 20th century. Hence this data is not reliable to count which country has the best performance among all the games. Also, different country has vastly different population. USA usually send more athletes than most European countries. So only evaluate the number of each country is not very accurate.



- Moving forward, to avoid the factors of different attendance among different countries. We counted the ratio of the number of medals one country gains versus the number of games one attended over the past years.

The results is obvious that (Soviet Union)URS ranked top 1 among over participants with almost 44% of all the games they have attended they win medals. Following is German (GDR), with around 38% of winning, Australia and New Zealand (ANZ) with 33%, Unified team (EUN) with around 32%.
Noting that we have exclude ANZ & EUN, because ANZ only attended 2 games and they won several medals. In our case, they can counted as outliers since the pattern we need to find should be the trend represented by all past years' events. Unified Team (EUN) is even older than the former Soviet Union (URS), it does not make any sense to include a country that is not exist anymore for our analyze hence we exclude the outdated data.

Winning medals per game of each country



Sum of Total Medal/#Games for each NOC. Color shows details about NOC.
The view is filtered on NOC, which keeps 6 of 230 members.

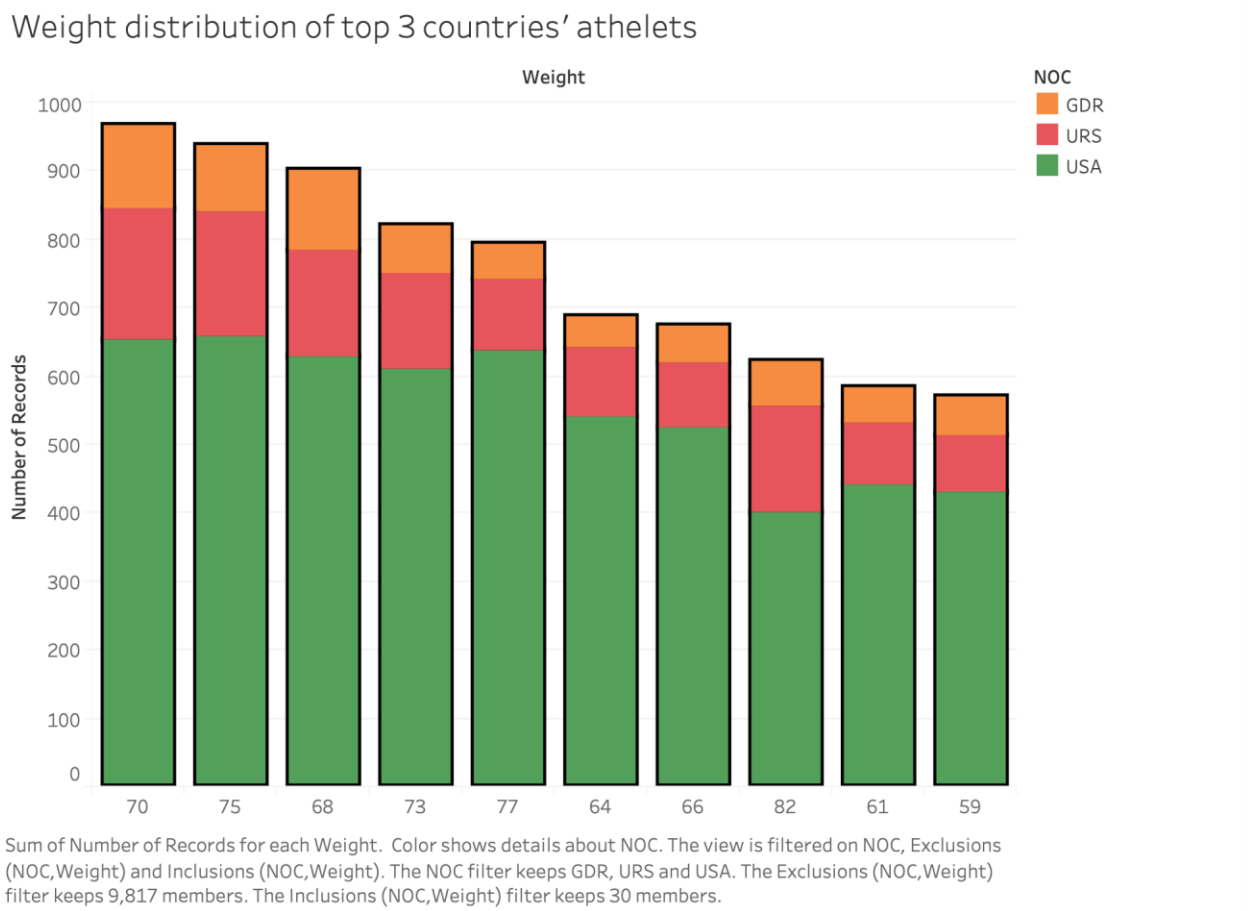
Note:
The most large number of height of the players in top three country was 180cm. But 178cm was only a few number.

The graph below is the frequency graph of the athletes from those 3 countries with the best performance. All the height are in centimeter. From the graph, it is obvious that the most frequent height is 183 and the total number of athletes from these countries decreases from the left column to the right ones. Therefore, it is recommended that athlete with a height of 180, 183 and 178 are more likely to succeed in the Olympic games.

Height of top 3 countries distribution

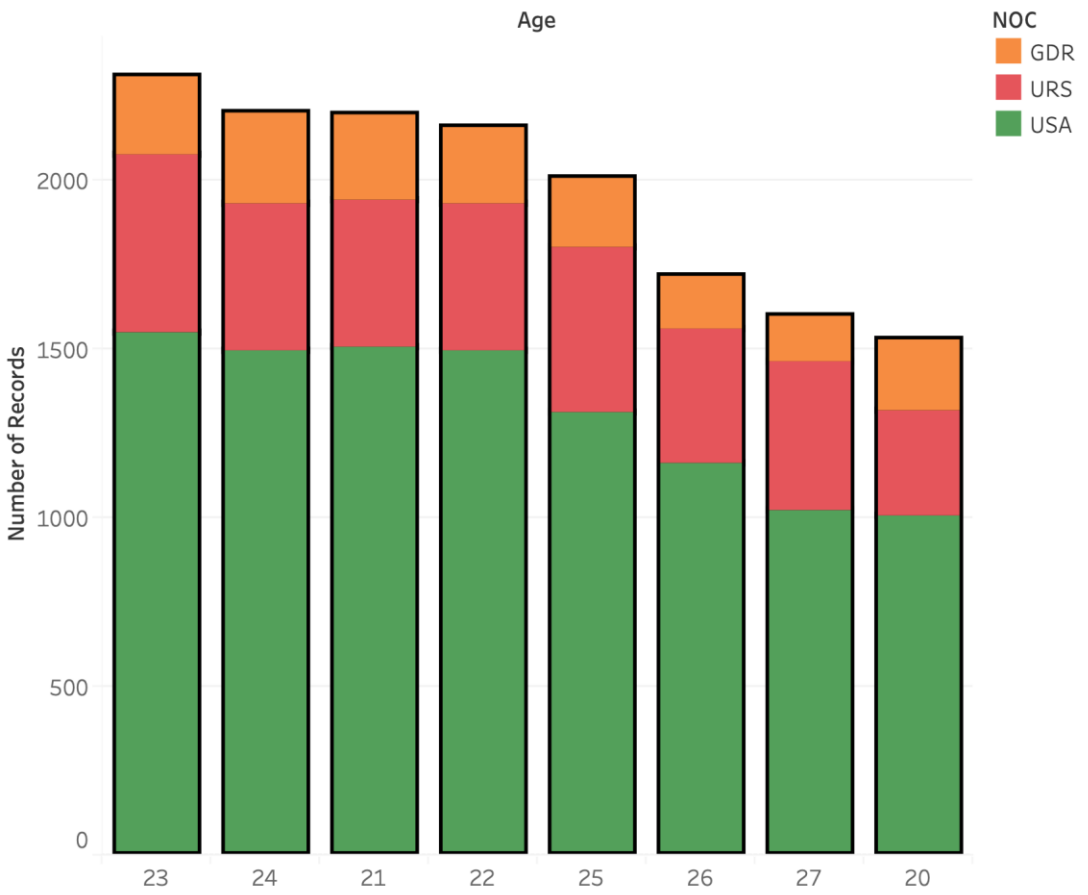
Sum of Number of Records for each Height. Color shows details about NOC. The view is filtered on NOC, Exclusions (Height,NOC) and Inclusions (Height,NOC). The NOC filter keeps GDR, URS and USA. The Exclusions (Height,NOC) filter keeps 7,683 members. The Inclusions (Height,NOC) filter keeps 24 members.

This is the graph for the weight of the same 3 countries. All the units are in kilograms. Like the height graph, athletes with the weight of 70, 75 and 68 kilograms have the best performance overall, so it's recommended that athletes with those weights have more potential to win a medal.



This bar graph is the age distribution of the athletes from the top 3 performance country. It is apparent that all the athletes are in their 20s, with the ages 23 on the top, 24, and 21 followed. The countries should send athletes in their 20s to attend the Olympics game, particularly 23, 24 and 21 with a better chance of winning.

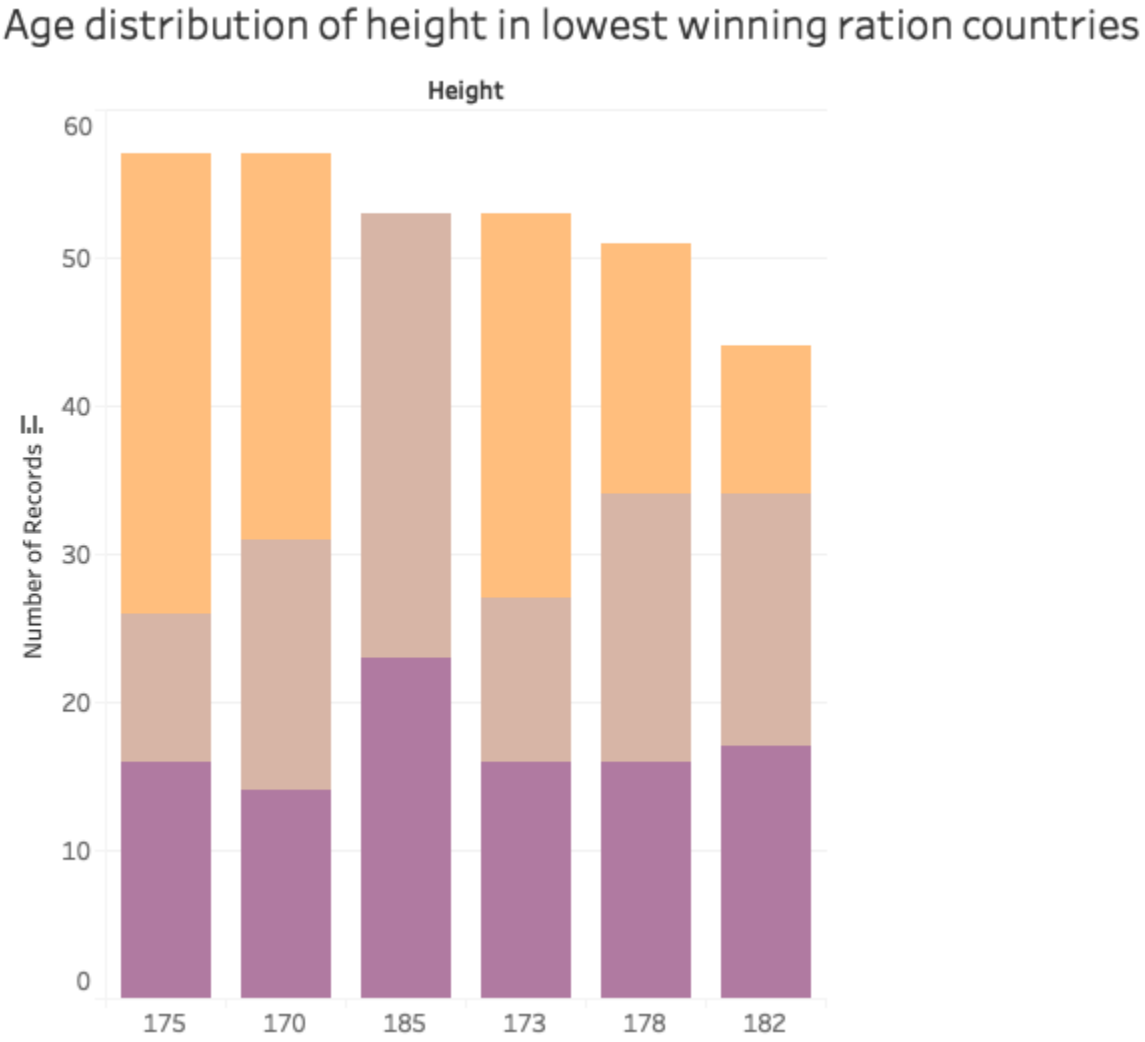
Age distribution of top 3 countries' athelets



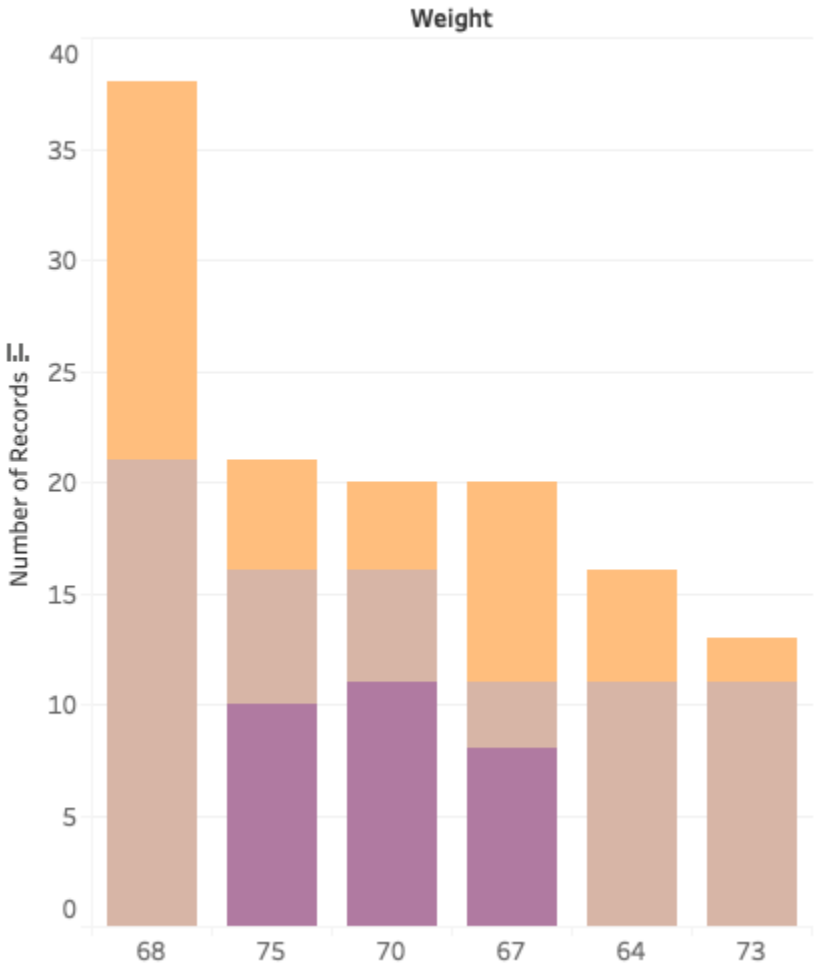
Sum of Number of Records for each Age. Color shows details about NOC. The view is filtered on NOC and Inclusions (Age,NOC). The NOC filter keeps GDR, URS and USA. The Inclusions (Age,NOC) filter keeps 24 members.

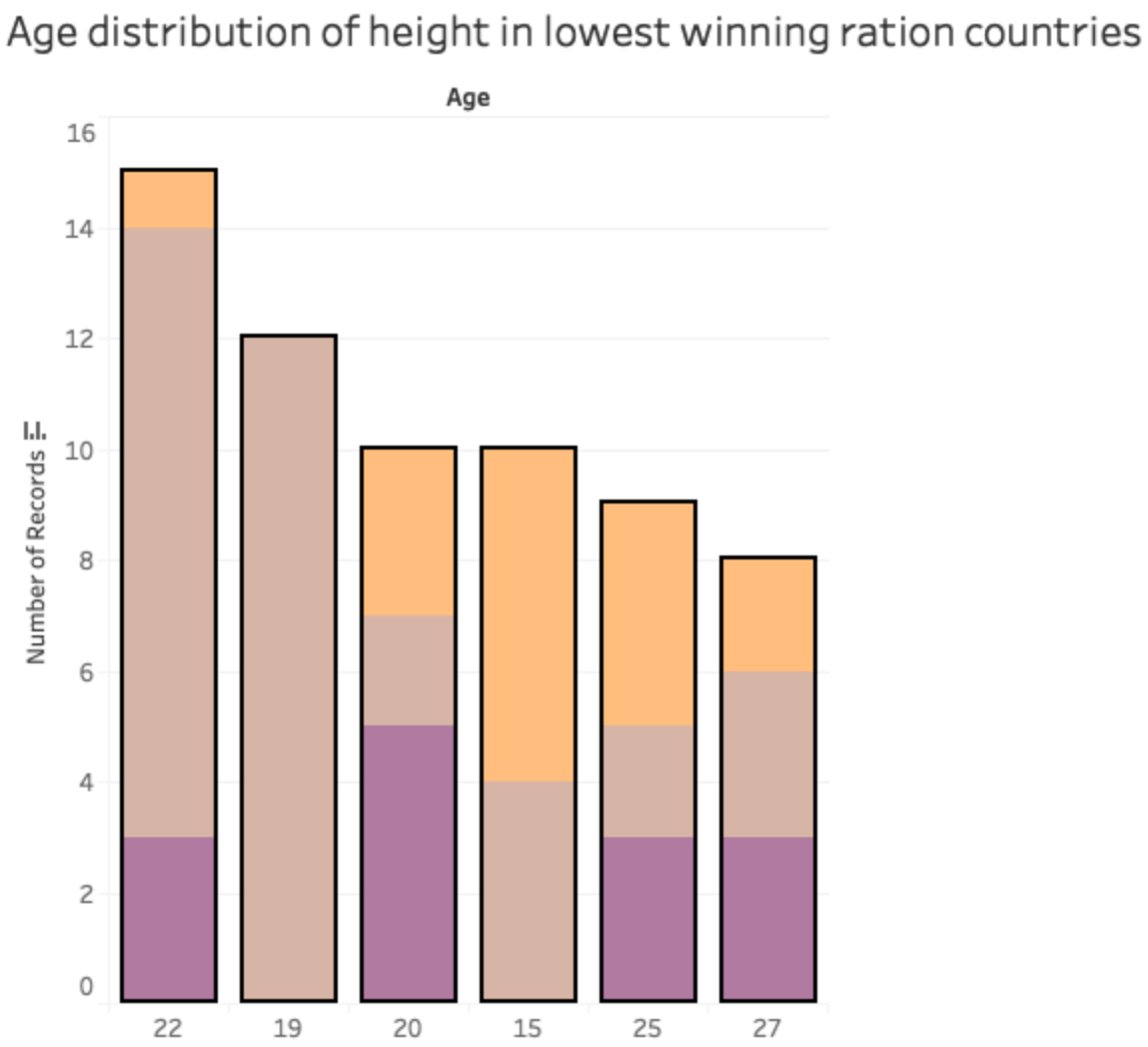
Eventually, we compare those ratio of countries in the low medal

Now let us see the lowest 3 winning ration per game countries's data:



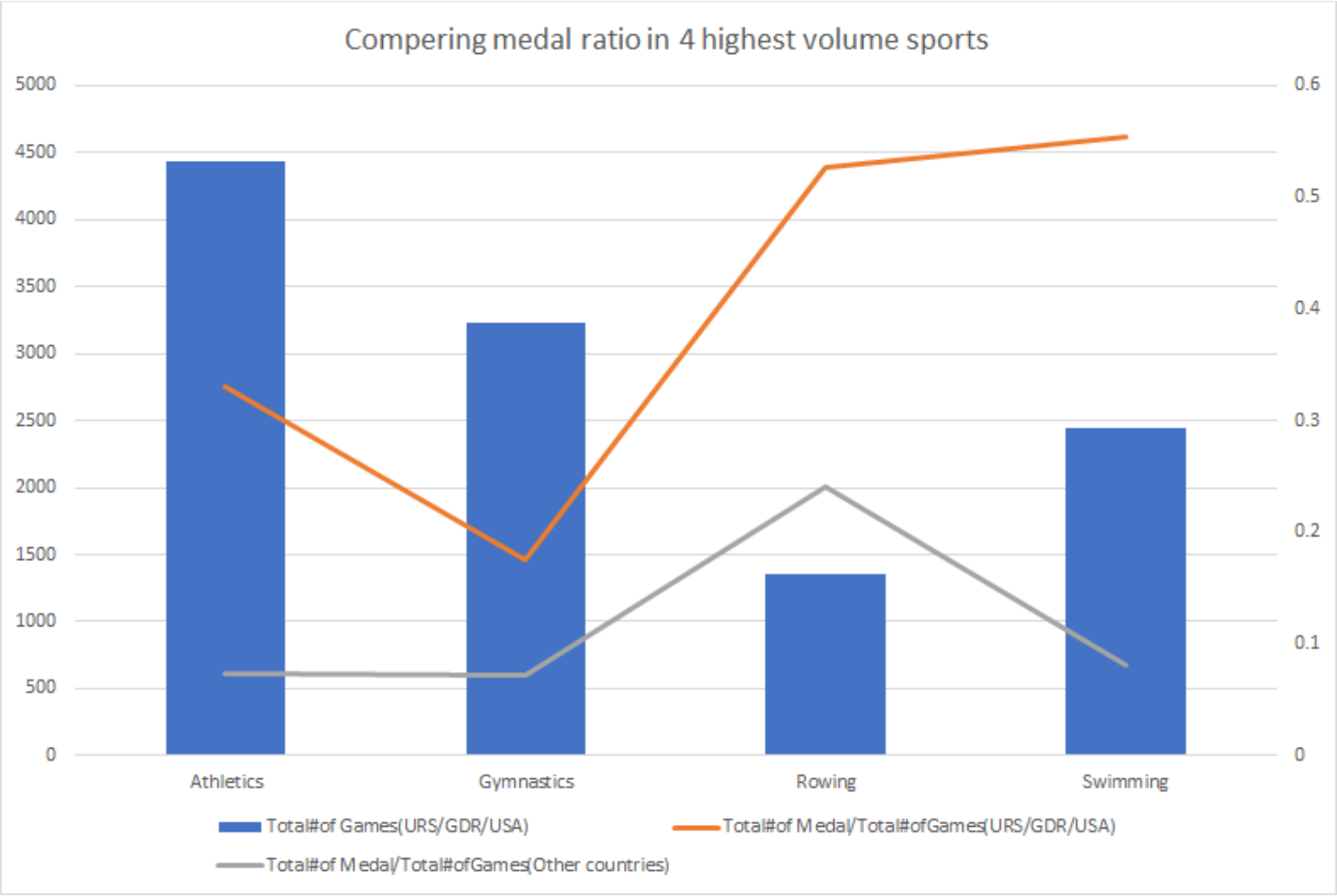
Weight distribution of height in lowest winning ration countries





Hence our question moved toward to Whether proportional sending athletes to events make the chance of winning a medal higher?

An analysis of the number of athletes sent to an event compared to the medals won. In the top three countries, an analysis of the top four highest volume sports (Athletics, Gymnastics, Rowing, and Swimming) showed a greater ratio of medals won. An increase in the volumes of athletes sent to an event is correlated to a greater ratio of medals won, in high performing nations.



Conclusion

Generally, based on the analysis on those higher performance nations, we determined that there was a difference in height and weight compared to all other nations, while age was not a significantly different factor. For the ideal winning athlete, it should be around 183 cm, age 23, weight 70 kg. An increased volume in sending athletes to sports generally result in increased ratio of medals for winning nations.