Milestone 4: Presenting Your Findings

Report Sport Stats

Contents

- Review of Questions to Answer/Hypotheses/Approach;
- Discuss Technical Challenges;
- Detail: Entity Relationship Diagram (ERD);
- Initial Findings;
- Deeper Analysis;
- Hypotheses Results.

Section 1: Questions to Answer

- Q1: Is there any correlation between the performance of a country in Winter Olympics and that in Summer Olympics?
- Q2: Does country performance by year change more in Winter Olympics or Summer Olympics?
- Q3: How has the male:female ratio evolved through time?

Section 2 Initial Hypotheses

- Hypotheses 01: Yes;
- Hypotheses 02: Winter Olympics;
- Hypotheses 03: Decreased.

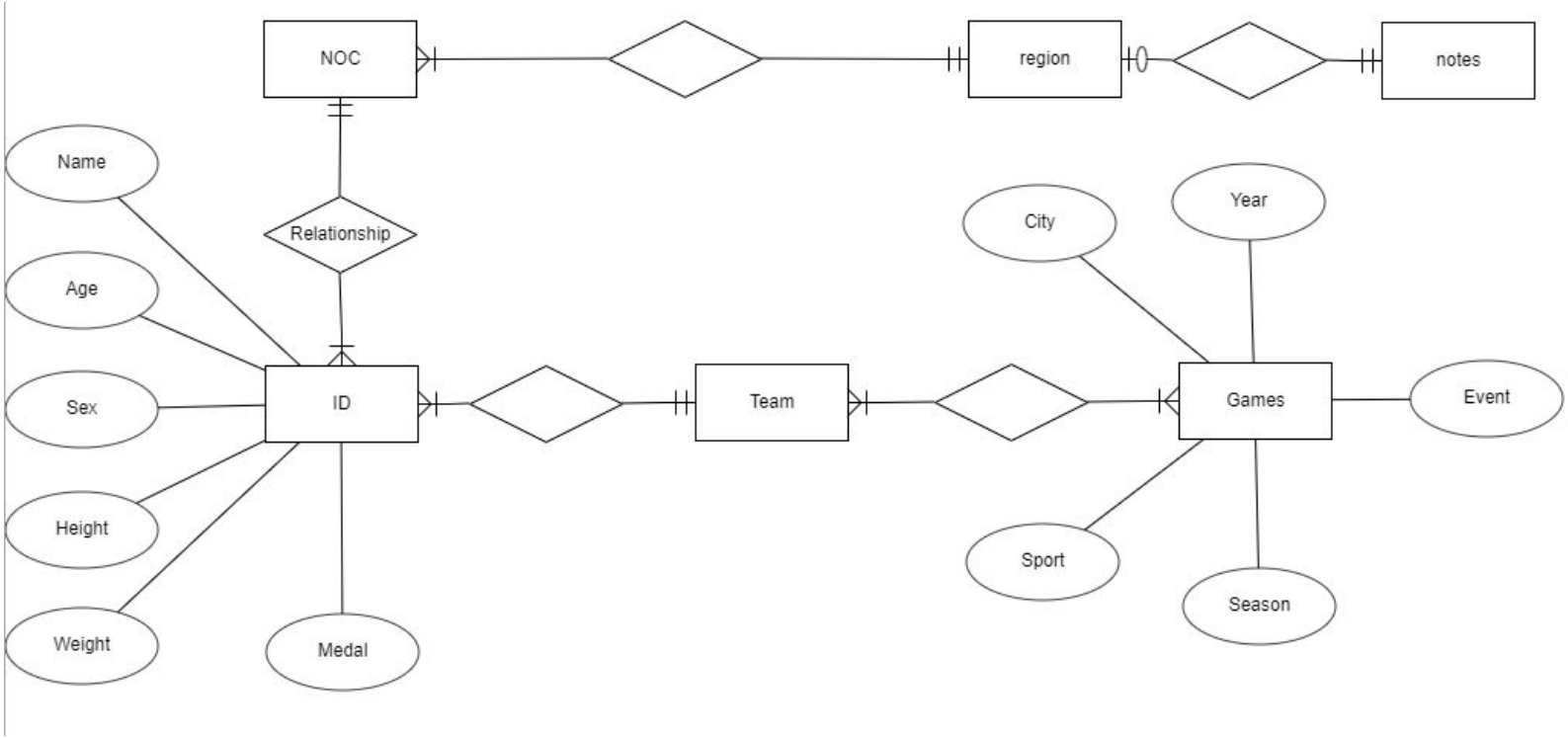
Section 3: Data Analysis Approach

- A1: to calculate the Pearson correlation coefficient.
- A2: to calculate the standard deviation in country performance through years. A Comparison between average std of Winter and that of Summer Olympics will help.
- A3: to draw a simple histogram.

Technical Challenges

- Encountered challenges with getting the starting year of the Summer Olympics different from that of the Winter Olympics;
- Limitation of Pandas for SQL (Sqlite) made some SQL difficult to execute but manageable.

Entity Relationship Diagram (ERD)

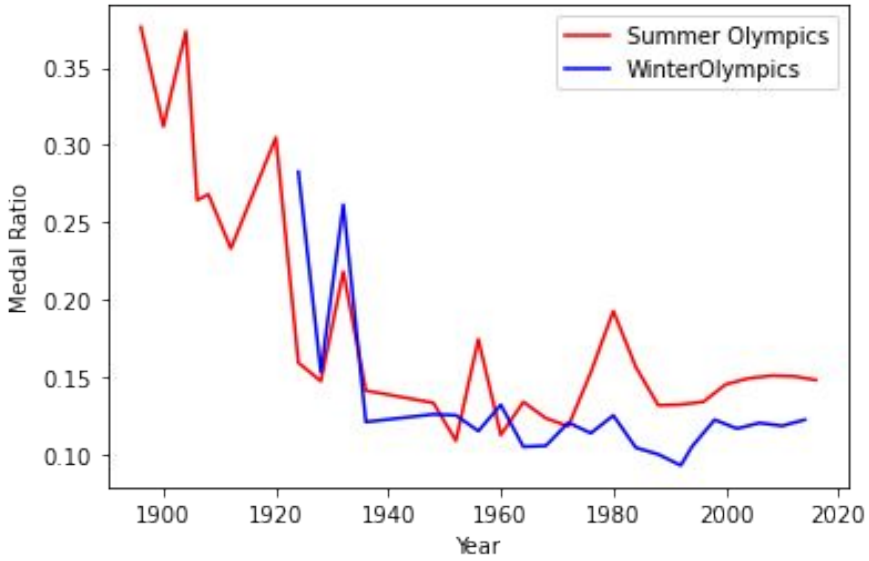


Initial Findings

- Although the ratio between the Summer Olympics and the Winter Olympics is indeed different, men happen to be dominant. My first assumption is that the ratio of women to men has increased over time. I began to dive into it.
- There are significant differences between male and female participants not only in terms of expected height and weight, but also in terms of age. The first two differences can be attributed to biology. Although the latter may require more than just: it is worth considering social factors at the same time.
- Another interesting fact is that the age gap in the Winter Olympics is much smaller (~2.8 years old and 1.5 years old)
- Another analysis of the number and ratio of medals is needed. I checked the ratio of total medal winners and the changes in the ratio of different medals:

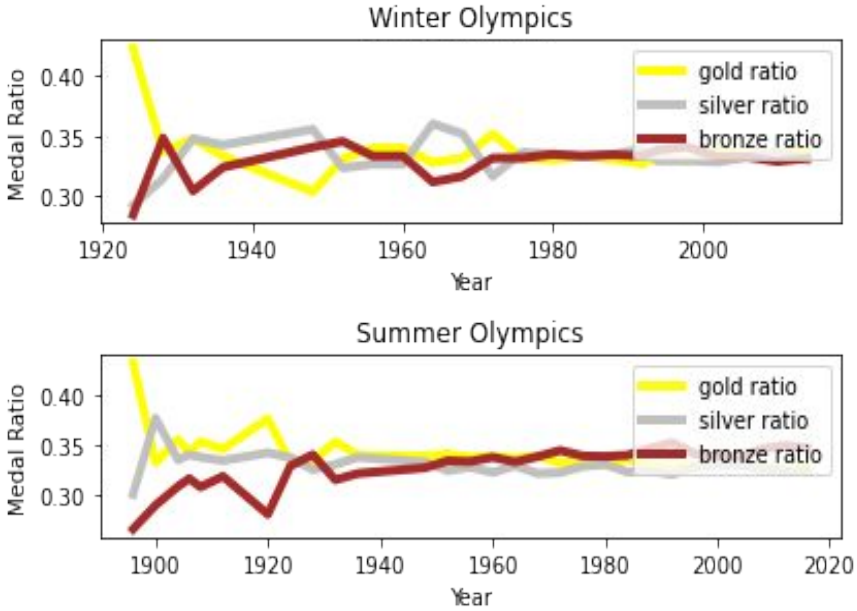
Initial Findings

In the last century, the medal ratio fluctuated greatly in the two competitions, but eventually stabilized. This can be interpreted as establishing norms on these issues.



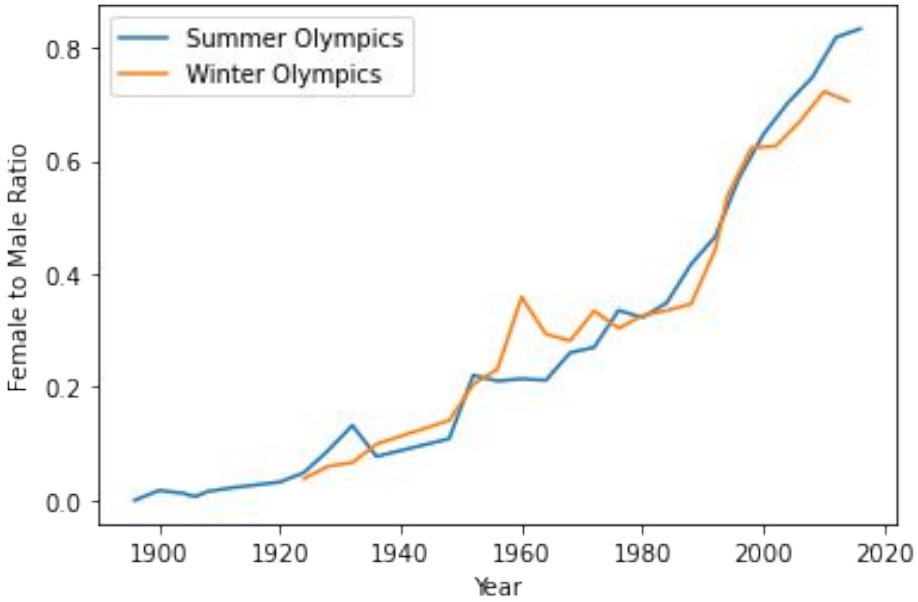
Initial Findings

The relative percentages of gold, silver and bronze medals have also stabilized, which may be due to the reasons mentioned above.



Initial Findings

This assumption seems to be correct. Over time, the ratio of women to men has indeed increased. However, there is an interesting detail: during the Second World War, the proportion of the Summer Olympics dropped sharply, but then it resumed its growth momentum. Without further analysis, I cannot explain this phenomenon.



Deeper Analysis

The length of the array of the number of medal count in the Winter Olympics and Summer Olympics are different because Winter Olympics started in 1924, but Summer Olympics started in 1896. Therefore, I have to create a new shortened table of the Summer Olympics started in 1924 to match the length of the Winter Olympics.

Deeper Analysis

- The Pearson correlation coefficient between the total number of medals in the winter and Summer Olympics, from 1924 to 2016, is 0.94, which is highly positive. Therefore, the performance of a country in Winter Olympics is highly correlated to that in Summer Olympics;
- I will then calculate the standard deviation in country performance through years. A Comparison between average std of Winter and that of Summer Olympics will help.

Deeper Analysis

- std_medal_count_summer_olympics = 475;
- std_medal_count_winter_olympics = 153;
- From 1924 to 2016, as the standard deviation in the Summer Olympics is about 3 times that in the Winter Olympics, country performance by year change more in Summer Olympics.

Final Findings (Result of Hypotheses)

- Yes, the performance of a country in Winter Olympics is highly correlated to that in Summer Olympics;
- Yes, the country performance by year change more in Winter Olympics than that in Summer Olympics;
- The male:female ratio has decreased from 1896 t o 2016.

Recommendations

The Olympiad Organizing Committee should devote more resource in the weather prediction to help organize the Olympics, as the weather affects the performance of athletes.

The Olympiad Organizing Committee should advocate the equality between male and female and keep encouraging more female to join the Olympics.