

A Statistical Look on a Country's Characteristics Determining Winter Olympics Success



Spyridon Mastrodimitris Gounaropoulos

Introduction

This work was created as a personal interest project caused by the domination of the gold medals from the Norwegian athletes. Norway is a small mountainous European country close to the Polar Circle, with a small population and high living standards. Somehow, against all odds, it won 37 medals. I was interested in discovering the characteristics that give a country an edge.

The first task was to create the dataset, something not so easily done. For the collection of the data, I used web scraping techniques and collected the following characteristics:

- ~ The number of gold medals
- ~ The number of silver medals
- ~ The number of bronze medals
- ~ The total number of medals
- ~ The country's average elevation
- ~ The country's yearly average temperature
- ~ The country's gross domestic product
- ~ The country's gross domestic product per capita
- ~ The country's population
- ~ The country's team size
- ~ The country's men's and women's average height
- ~ The country's men's and women's average weight
- ~ The country's men's and women's average BMI
- ~ The country's population density
- ~ The country's urbanization percentage

Contents

Introduction	1
Results of the Games.....	4
Initial approach.....	4
Environmental Factors.....	7
Temperature.....	8
Overall Population Fitness	11
Closeness to nature	13
Team effectiveness	14
Conclusions.....	16



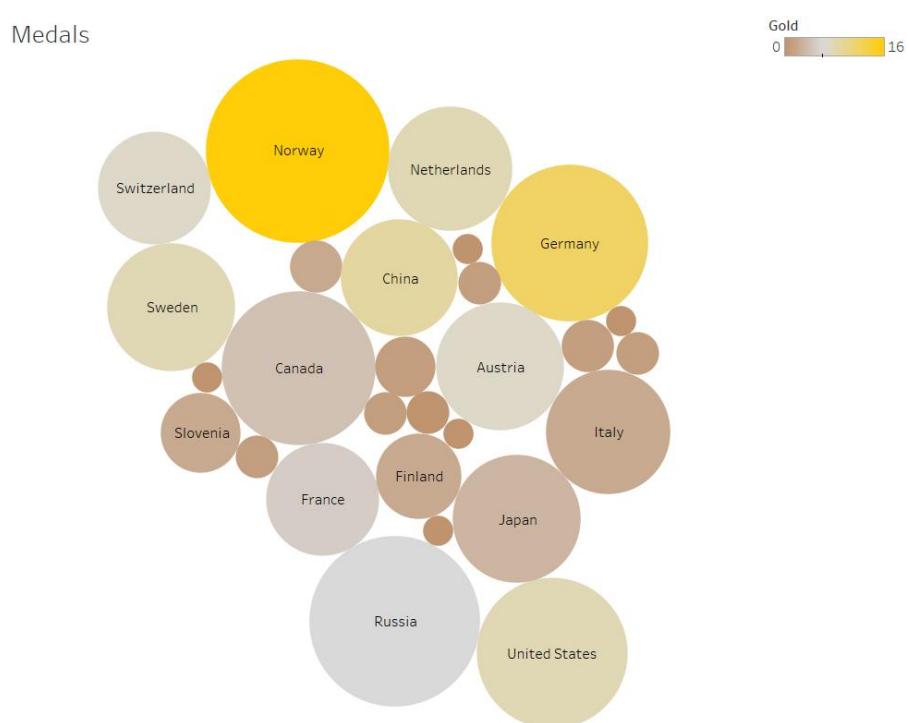
Results of the Games

For the 2022 Winter Olympic Games, Norway succeeded in being the country with the most overall medals and the country with the most gold. A trend that is not observed for other Nordic countries such as Sweden or Finland. However, Sweden managed to gather the same amount of medals as Japan while having almost three times the golds, an impressive achievement as it has one-tenth of the population. In general, small European countries are seen to punch above their weight, so to speak, with countries such as Austria, the Netherlands, and Switzerland to impress with their athlete's performance and especially for the gold to have similar success to ROC, Canada, or the USA.

My initial thought was the difference in the culture of Europe and, especially, in the countries around the Alps: Germany, Austria, Switzerland, Italy, and France. But beyond those 5, others, such as the Netherlands, Slovenia, and Hungary, also had impressive results.

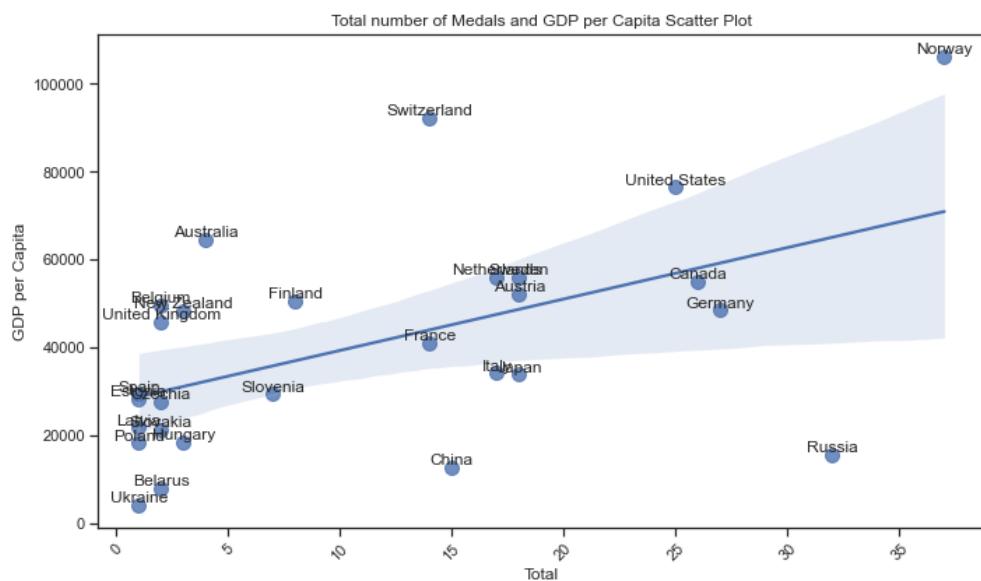
Medals Table

Country	Tot..	F	Gold	Silver	Bronze
Norway	37,00	16,00	8,00	13,00	
Russia	32,00	6,00	12,00	14,00	
Germany	27,00	12,00	10,00	5,00	
Canada	26,00	4,00	8,00	14,00	
United States	25,00	8,00	10,00	7,00	
Sweden	18,00	8,00	5,00	5,00	
Japan	18,00	3,00	6,00	9,00	
Austria	18,00	7,00	7,00	4,00	
Netherlands	17,00	8,00	5,00	4,00	
Italy	17,00	2,00	7,00	8,00	
China	15,00	9,00	4,00	2,00	
Switzerland	14,00	7,00	2,00	5,00	
France	14,00	5,00	7,00	2,00	
Finland	8,00	2,00	2,00	4,00	
Slovenia	7,00	2,00	3,00	2,00	
Australia	4,00	1,00	2,00	1,00	
New Zealand	3,00	2,00	1,00	0,00	
Hungary	3,00	1,00	0,00	2,00	
United Kingdom	2,00	1,00	1,00	0,00	
Slovakia	2,00	1,00	0,00	1,00	
Czechia	2,00	1,00	0,00	1,00	
Belgium	2,00	1,00	0,00	1,00	
Belarus	2,00	0,00	2,00	0,00	
Ukraine	1,00	0,00	1,00	0,00	
Spain	1,00	0,00	1,00	0,00	
Poland	1,00	0,00	0,00	1,00	
Latvia	1,00	0,00	0,00	1,00	
Estonia	1,00	0,00	0,00	1,00	

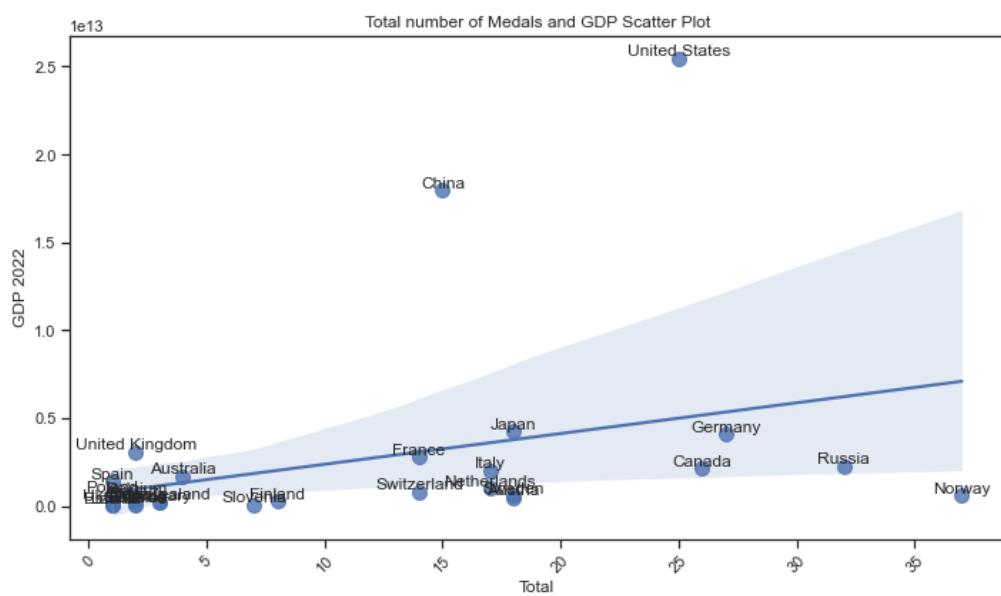


Initial approach

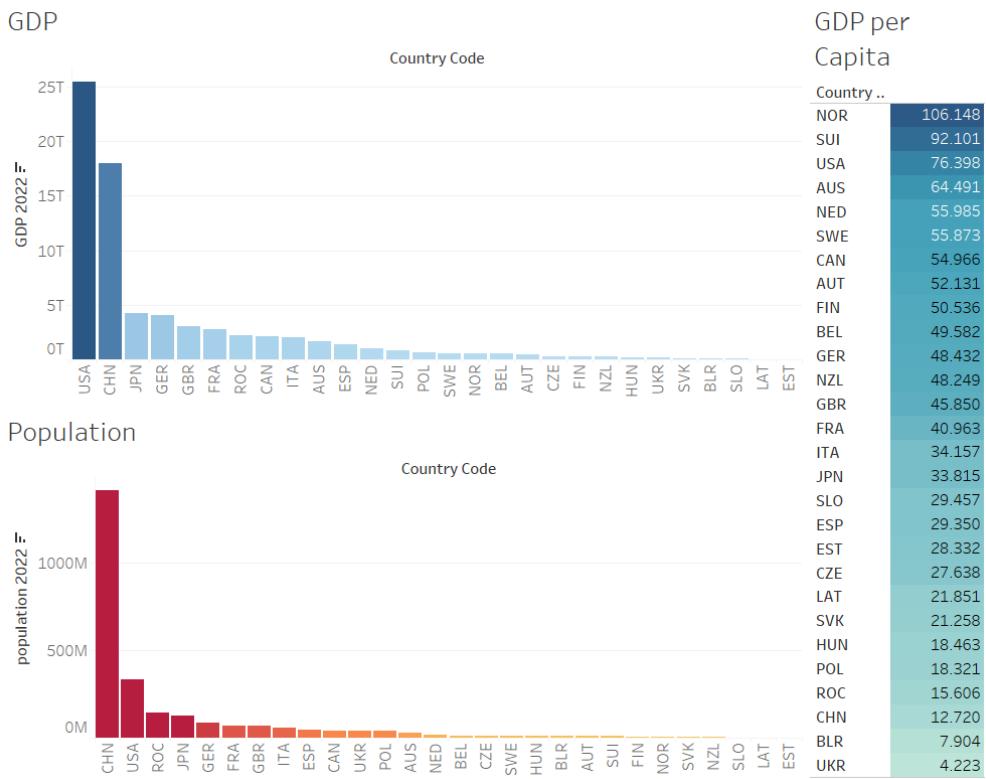
My second train of thought revolved around wealth and environmental factors of the countries of origin. Nordic and Western European countries tend to have a higher GDP per capita than the rest of the world. I hypothesized that, in theory, better life conditions and the social and economic nets protecting their citizenry allow an excellent opportunity for sports involvement from an early age. As a result, more people with the right aptitude and ability will participate and succeed in local and global competitions.



From the diagram above, we observe a possible correlation, but it is clear that not all depend on the specific factor. I wanted to see if the flat GDP had any significant impact.

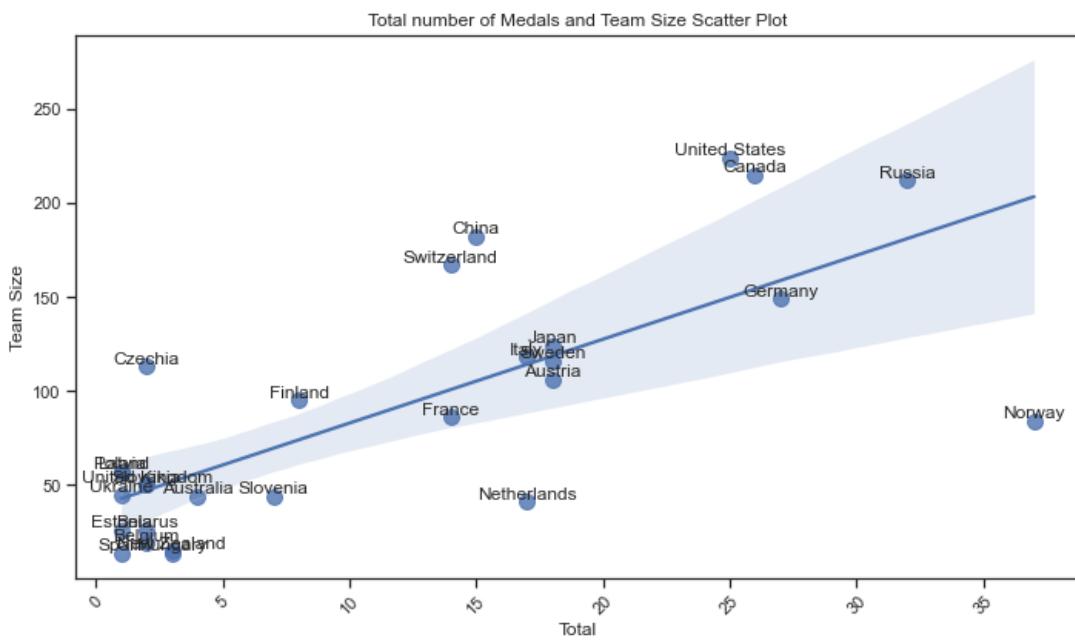


From this diagram, it is hard to make comparisons, as the two main outliers skew the plot upwards. However, if the flat GDP were a significant factor, the USA and China would have much more success.



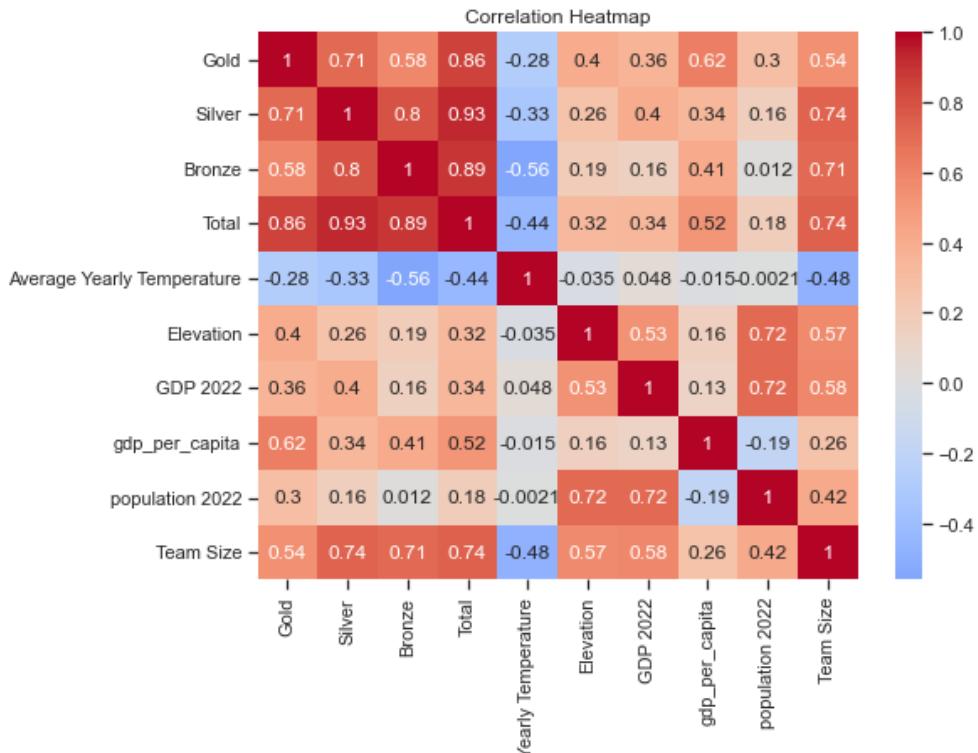
From the data, I observed that GDP impacted the team size, and it is logical that the larger a national team is, the more medals should bring home.

Indeed, this trend is evident overall with some outliers such as Switzerland, China, Czechia, the Netherlands, and again Norway.



Environmental Factors

A simple scatterplot can only reveal so much, so I created a correlation heatmap to identify possible correlations for the factors I chose to inspect.



This diagram shows multiple low and significant correlations between factors and medals. The most influential factor is the Team size, having a strong correlation for the number of medals (0.74), silvers (0.74), and bronzes (0.71) but weaker for golds (0.54), although still significant. Combined with the correlation between population and golds (0.3), it creates the image that while it is easier to win lower medals with a larger team for golds, it helps to pull from a large population to fill the team positions. Another critical factor is the GDP per capita, which also correlated with medals (0.54) and, more importantly, with golds (0.62) rather than Silvers (0.34) or Bronzes (0.41). This might reveal a trend where those athletes are able to access a better system of training, equipment, and support to gain a slight margin of advantage compared to others. Lastly, for this diagram, altitude for golds (0.4) and GDP for silvers (0.4) show a weak correlation.

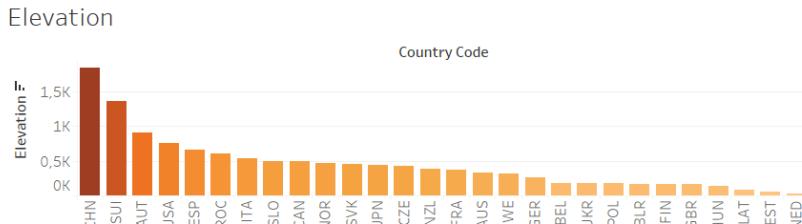
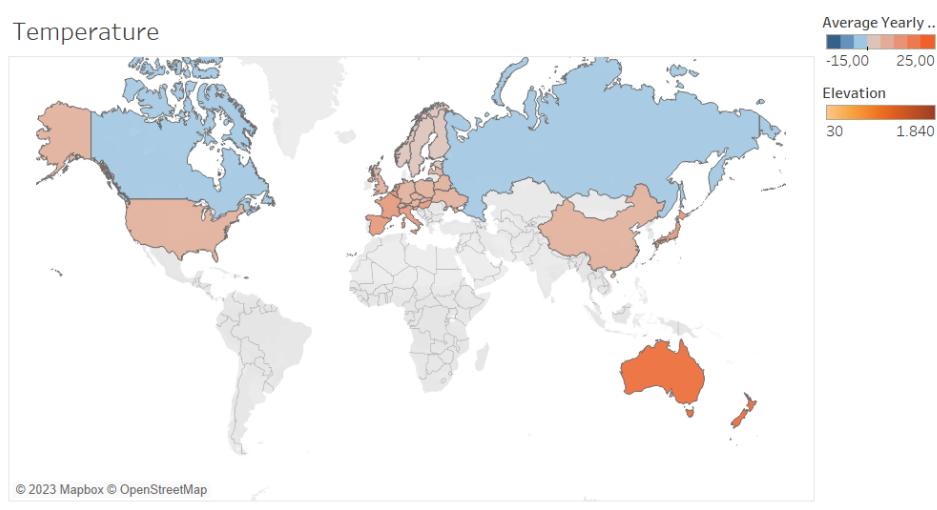
Temperature

We could not talk about the Winter Olympics without taking into account the temperature. As seen in the heatmap, a simple correlation could not provide logical results because Russia and Canada have a yearly average temperature of -5 degrees Celsius. So, I used Pearson correlations.

Average Yearly Temperature	
Gold	-0.276041
Silver	-0.325938
Bronze	-0.561312
Total	-0.439224
Average Yearly Temperature	1.000000

We see a negative correlation between high temperatures and medals, with the most significant on bronzes at (0.56) by no means insignificant. While all medals are affected to some degree, I use the word affected because there is a clear causation here, the greater the medal, the lesser the effect. Spearman's correlation showed a lesser effect and a different trend between golds and silvers.

Average Yearly Temperature	
Gold	-0.154084
Silver	-0.116158
Bronze	-0.335388
Total	-0.194013
Average Yearly Temperature	1.000000



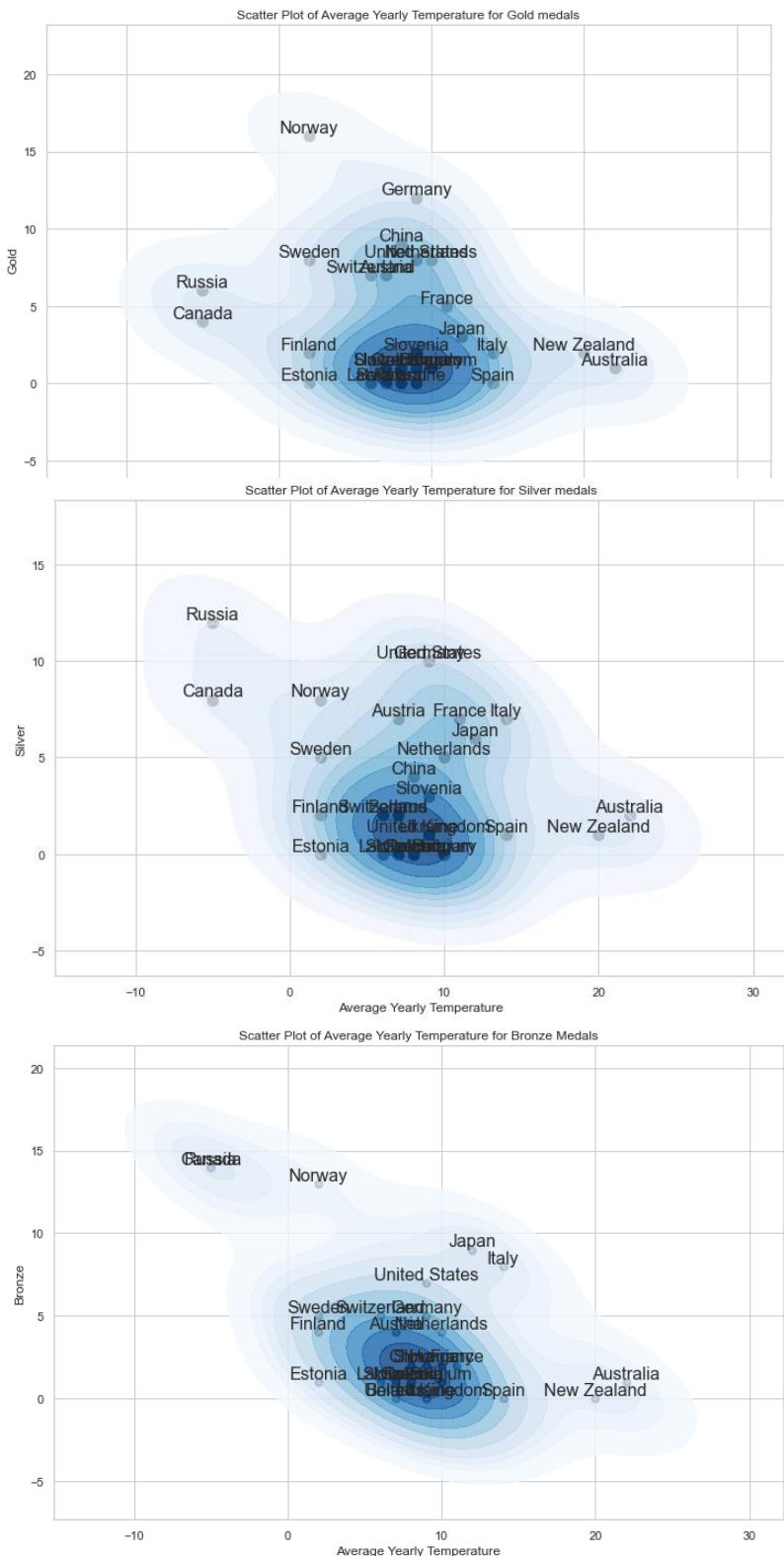
The following diagrams better visualize the connection between wins and the average yearly temperature for each medal.

The effect might have been more pronounced if Norway, a country with an average yearly temperature of 2 degrees, or Germany, with a temperature of 9, took fewer golds.

Russia and Canada did very well for the silvers, and the shape changed to an ellipsoid, revealing a negative correlation. As for the bronzes, the pattern repeats.

Countries like Estonia, Belarus and Ukraine underperform if analyzed in a bubble.

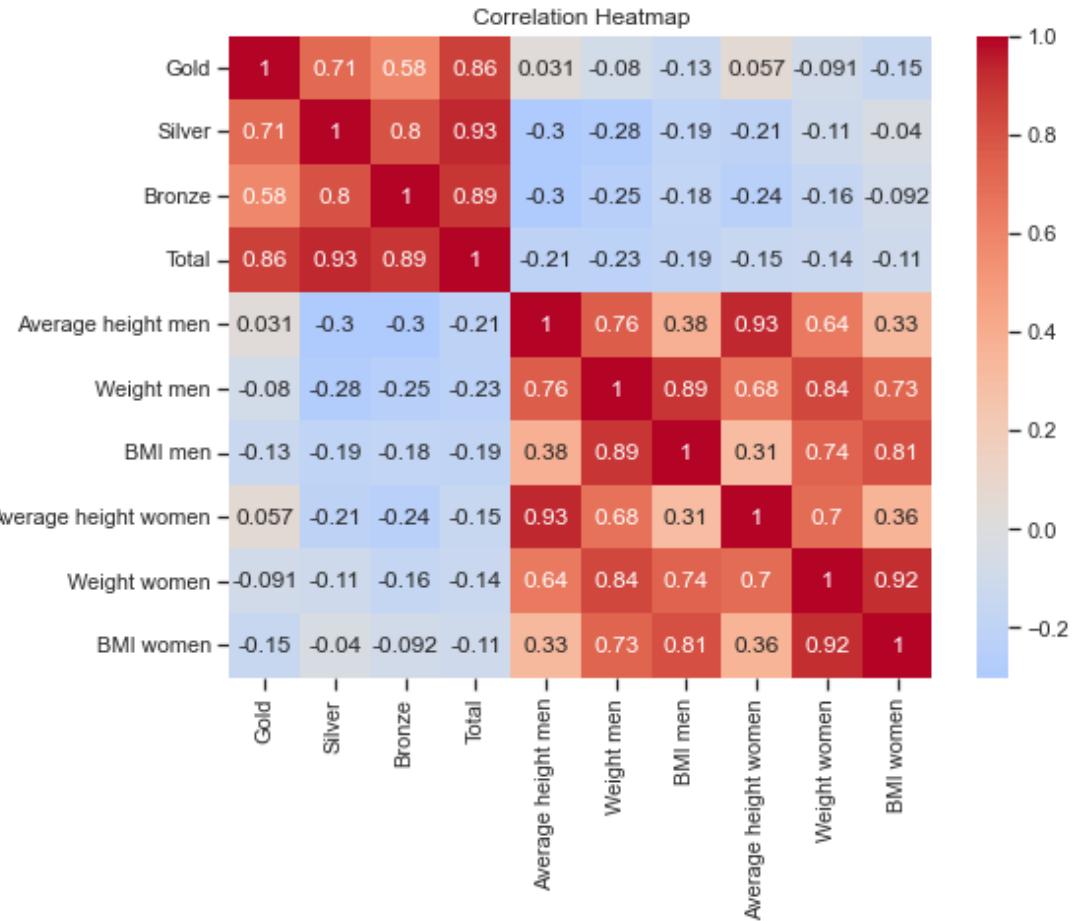
However, if considering the population, team size and gdp per capita, their performance can be rationalized.





Overall Population Fitness

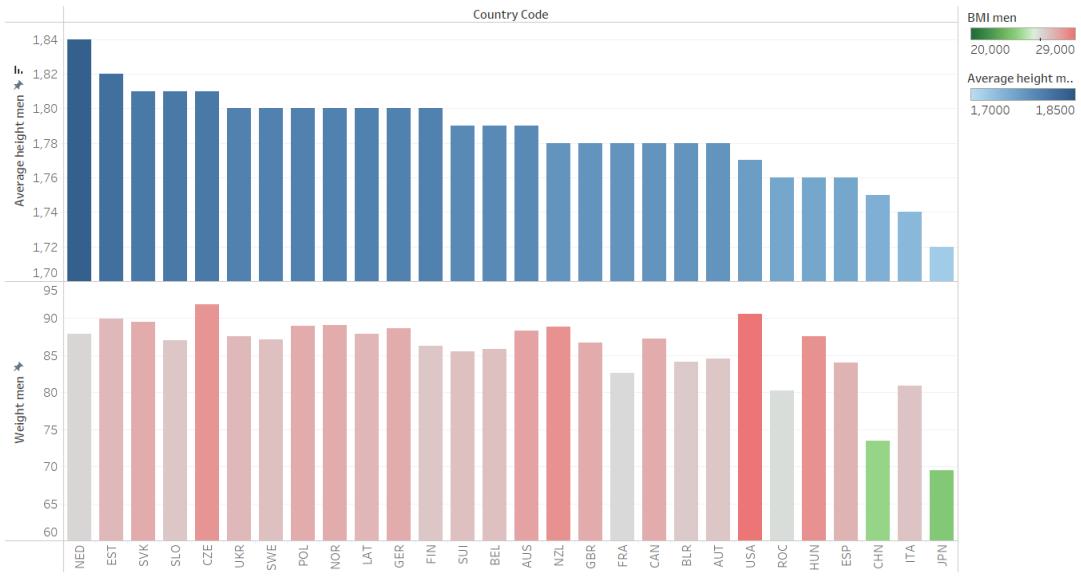
One different approach tested was the hypothesis that a healthier, athletic, and sports-involved population could be a hotspot for Olympic winners. For the hypothesis, I analyzed the correlation between height, weight, and BMI for both men and women.



The correlation results for women were not noteworthy as it mattered only between height and Bronze medals with an unconvincing (-0.24), revealing a questionable negative correlation. If the conditions were more favorable, I would like to test it again based on sport, as not all value the same characteristics equally.

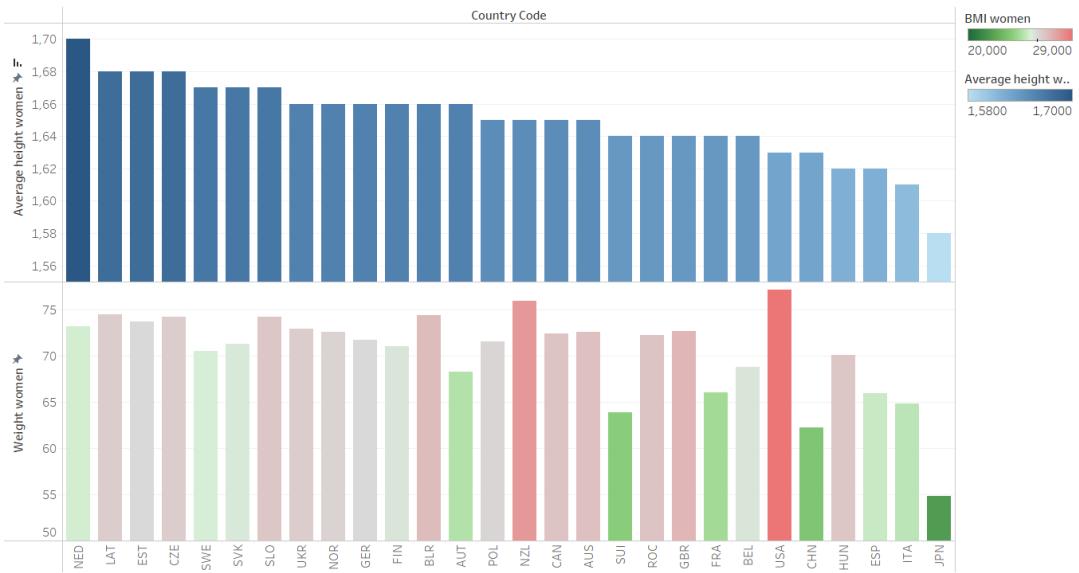
While there is a negative correlation for both height and weight for men, the correlation for BMI seems relatively weak, if not insignificant. Perhaps this is because, in contrast with women in most countries, men exceed the healthy BMI limits, enhancing the effect.

Average Height and Weight of Men Comparison



Sum of Average height men and sum of Weight men for each Country Code. For pane Sum of Average height men: Color shows sum of Average height men. For pane Sum of Weight men: Color shows sum of BMI men.

Average Height and Weight of Women Comparison

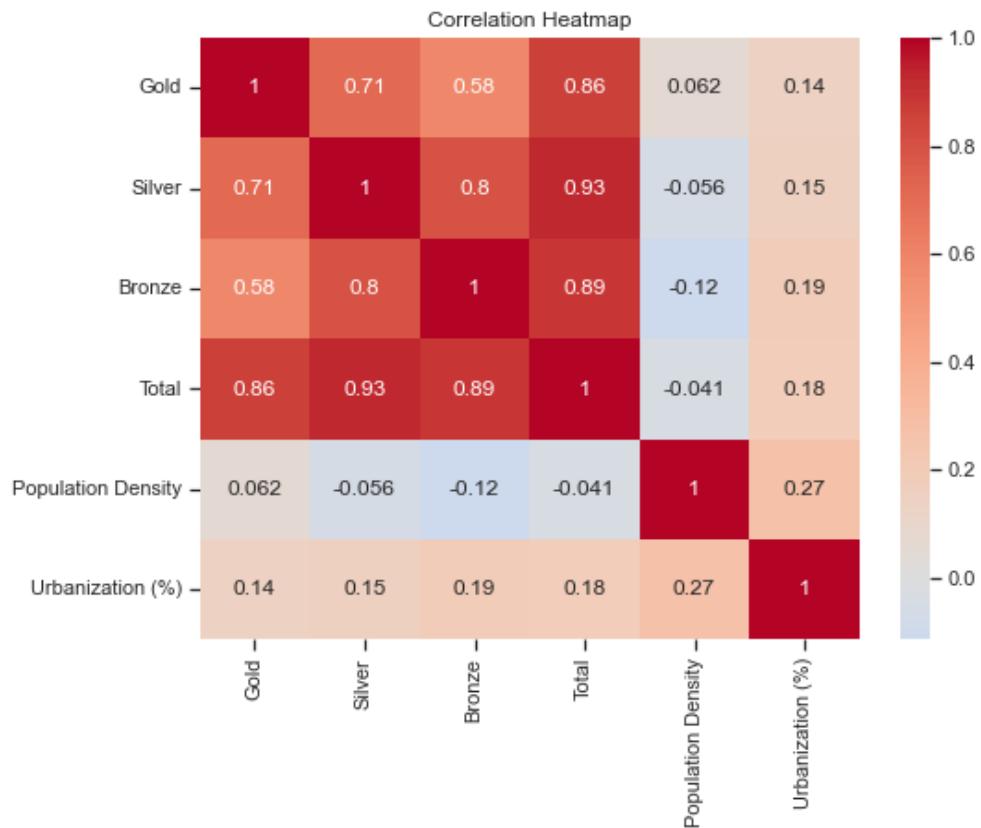


Sum of Average height women and sum of Weight women for each Country Code. For pane Sum of Average height women: Color shows sum of Average height women. For pane Sum of Weight women: Color shows sum of BMI women.

The most probable reason for both height and weight to be negatively correlated with success is that populations with those characteristics also lack in GDP per capita, for example, Estonia, Latvia, Czechia, Slovakia, Slovenia, Ukraine, and Poland. Other countries with populations with high GDP per capita and height, for example, Norway, Germany, Netherlands, and Sweden, have excellent results, meaning that population height and weight are independent of an athlete's success.

Closeness to nature

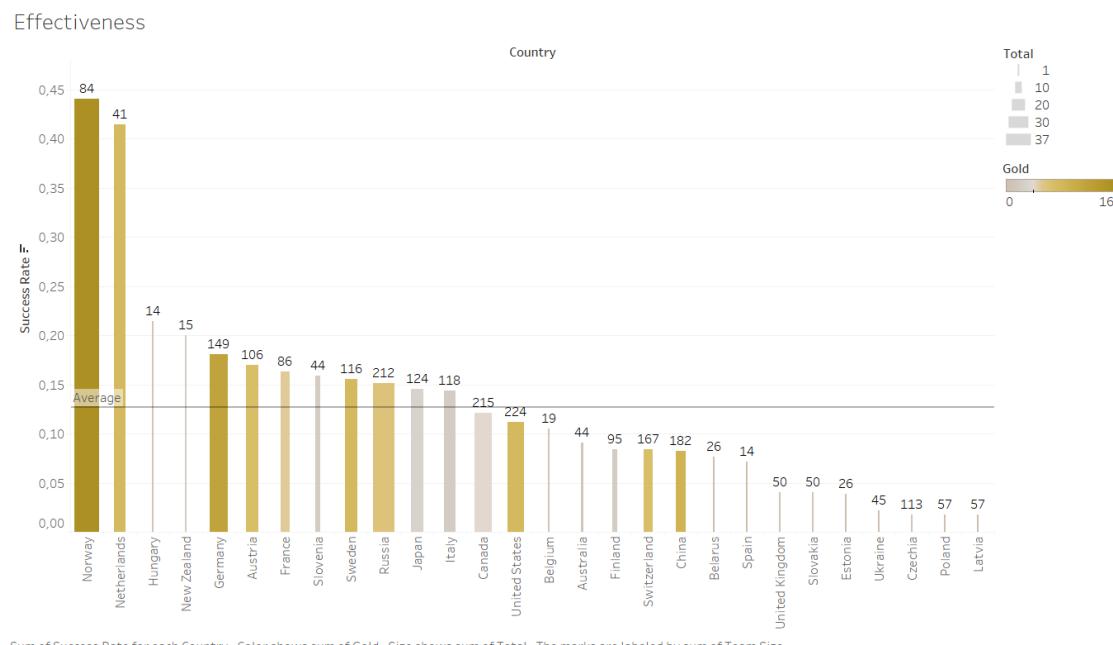
Many modern Winter Olympics sports can trace their routes to survival skills in the wild, such as the Biathlon on activities in the frozen outdoors such as curling, ice skating, and hockey. I wanted to observe any correlation left between a population with lower urbanization and, to an extent, a closer relationship with nature and a greater affinity for success in games.



My initial hypothesis was completely turned around as a more urbanized population correlates very weakly with medals. At the same time, population density seems to be completely independent. By chance, the most interesting statistic here is that the correlation between population density and urbanization is so weakly correlated.

Team effectiveness

One very utilitarian approach for sure, but it should shed a little different shade on the problem. The idea behind the "effectiveness" of the national team here is the successful chance of an athlete to get a medal if they are part of this country's team. This statistic penalizes a team if it has many athletes and increases its score if it has many medals. For example, the team from the Netherlands counted 41 athletes, winning 17 medals from those 8 gold.



Unfortunately, because many sports require an entire team, countries that compete and fail to win are penalized unfairly. If this diagram was analyzing countries sport for sport and, more specifically, for individual sports, it would be much more valuable.

Interestingly, even this less refined look shows that countries with low elevation, not low enough temperature, low GDP, and low GDP per capita are at the bottom of the list.



FIGURE SKATING
NASHVILLE

Conclusions

Norway's success is indeed an outlier from most optics. Nevertheless, even if this research cannot conclusively quantify the reasons behind Norway's success, it definitely picks up on trends and correlations for the average country.

Team size seems to be the most significant determining factor (0.74) highly influenced by GDP and population. Bigger countries with high GDPs and large populations, such as the USA, Canada, Russia, and Germany, have a strong presence, while others with similar characteristics, such as China and the United Kingdom, rank lower due to other factors such as temperature, GDP per capita or elevation.

The second most influential factor for gold medals is GDP per Capita (0.62), and countries such as Norway, Sweden, Switzerland, and the Netherlands have a solid hold on the highest podium position.

The last important factor is the temperature (0.44). It becomes apparent from examples such as Australia and New Zealand, with very high GDP per Capita relatively high GDP and average elevation, getting outperformed from Slovenia, a country significantly lower in all metrics but temperature.

Other factors, such as population characteristics and population density, are shown to be independent. At the same time, urbanization seems to have a minuscule positive correlation (0.18) with game success.

As a prediction for the following games: since temperature and altitude are somewhat fixed, countries with increased GDP per Capita, GDP, and population will show greater success. While countries with lower rates of economic development, as seen in the European economic space for 2022-2023, might see fewer medals in the following years.

