A visual analysis of GDP, Population and Gender in relation with performance at the Olympic games

Marco Ferraro, Stefano Foti

November 11, 2021

Abstract

How many medals should a country win? Why do some countries win more than others? Among other things, we analyze the relationship between GDP, Population, Gender and Time with the amount of medals won at the Olympic games using our data-viz web application dedicated to the Olympics.

1 Introduction

This paper has been produced as a reference for our Visual Analytics Project. In the following, we will discuss the evolution of the relationship between medals won and GDP, Population and gender. Many researches related to the field point out how the GDP is linked to each country medal amount. Moreover, men and women participation has not been equal for years, and changed a lot during the recent past. As an example, until 1968 Olympics, women could only take part to only 16 Olympic sports, and women participation was closer to 12% in the same year year. Today the participation is over 48% and the amount of sports is three times as much[4]. Our dashboard tries to show not only the well known patterns about the relationship between GDP/population/gender and the amount of medals, but offers also some features that let the user investigating in depth about those patterns in order to discover some peculiar insights.

2 The importance of GDP

We can find in the literature many studies linking medals count with GDP and population. They have been mostly used as a reference to predict country performances, understand medals distribution and so on. The papers below show that the simplest hypothesis telling medals should be proportional to population is not as reliable as it may be expected to be. Some most populated countries do not own a proportional amount of medals; it is more interesting to see the relationship between GDP and medals amount. A common result is that economic resources are essential for producing Olympic athletes and having resources to invest in human ability is important in producing success[1].

In Sydney 2000, GDP and, to a lesser degree, population has been found to be the main factor in determining how many medals each country should win at the Olympics, in fact they have been used to predict the amount of medals that countries would have won[1]. This paper was used as a basis for other studies going forward.

Further analysis of this relation was carried out by Rathke and Woitek in 2007 who used a larger dataset and some additional data to conclude that population was only a relevant factor for rich countries, which explained some of the unusual outcome from the original study while confirming the correlation with GDP [3].

In 2012, Eva and Michael Leads put an emphasis on the gender difference between athletes in their research and only considering gold medals, showing the difference that GDP had on the performance on the two genders. While confirming the correlation, their conclusion surprisingly showed that suffrage age, participation in labor and fertility were factors not only in women performance but also in men's, showing a correlation between national women's conditions and overall performance. [2].

3 Our Research

Our dashboard allows us to visually analyze the relationship between medals won and GDP/Population and how this relation has changed over the years, furthermore it allows us to show on the fly the differences of this relation for each gender, for each event (Summer/Winter), for each type of medal and for any subset of sports or countries. Although we don't augment our data with women's conditions data, we are still able to show the differences between the genders for different sports and at different points in time.

3.1 Data

Olympics datasets in machine-readable formats are not as easily available as you may think. Olympic results are published in in-depth PDF books that require an extraordinary amount of work to transcribe. The largest publicly available machine-readable dataset provides only basic data on athletes and which medal they won, making it hard to conduct research on performance growth. Given these limitations, this dataset still provides 120 years of results that can be used to show multiple insights. To complement the data on Olympics results we also used two other datasets, one for GDP by country for each year and Population by country for each year.

3.2 The Dashboard

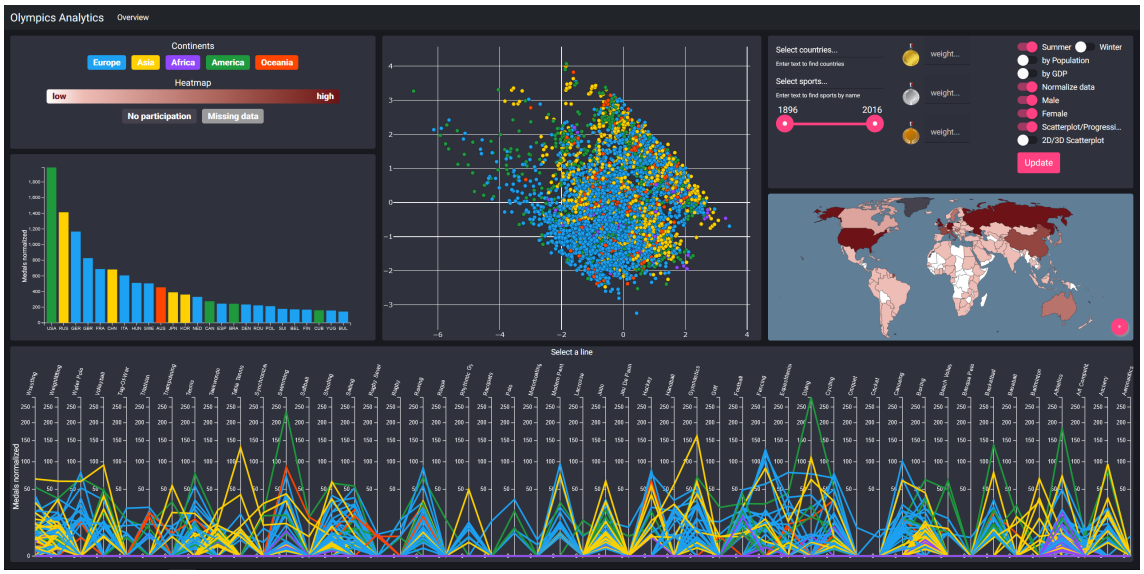

Figure 1: Our Olympics web-app

In this picture we show the dashboard, consisting of 5 different graphs (either PCA Scatterplot or Medal Progression can be shown at once) and a control panel.

- The Bar chart shows the top 25 countries by medal-count
- The Parallel-Coordinate chart shows the medal-count for each sport, with each line representing a country
- The Map colored by medal-count represents the geographical distribution of medals
- The PCA scatter-plot (2D or 3D) shows the entirety of the data projecting it along the principal axis
- The Medal progression chart shows the evolution of the medal-count over the years

Every chart is affected by the settings in the control panel, which offer multiple options:

- Country, Sport and Year range selectors, to only show the desired data
- Medal selector, to include only the selected medals or give them weights

- Summer/Winter and Male/Female filters
- GDP/Population comparison, divides medal-count by Population or GDP to show the correlation between the two
- Normalization, which computes a different value for each medal based on how many events are available for every sport
- Tradition. By clicking on the map one can enable a tradition computation that shows the best sports for that country.

Of course any combination of these options can be used at once.

3.3 GDP/Population analysis

In the following pictures we show the relation between medal-count and GDP over the years for Summer and Winter games, it is important to note the scale of the graph, which is exponential 2. These graphs show how much GDP has come to influence the medal-count more and more over the years, bundling countries closer and closer together. In particular we can see the winter games being affected by GDP the most compared to summer games. It's also apparent the tradition of winter sports for European countries colored in blue, which are performing better overall.

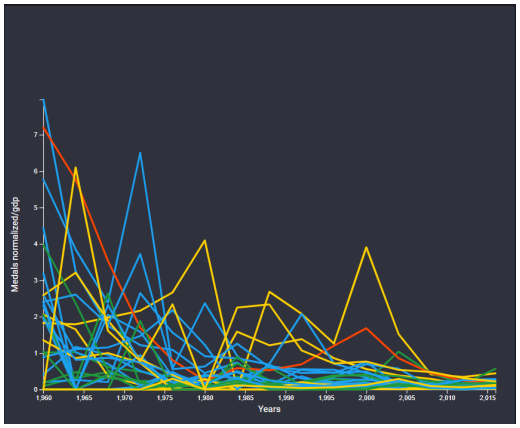

Figure 2: Medals/GDP relation for summer Olympics

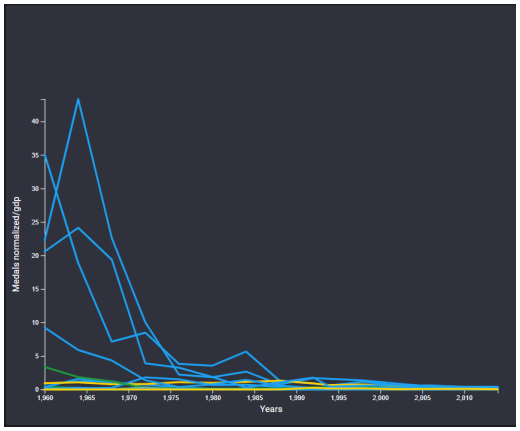

Figure 3: Medals/GDP relation for winter Olympics

The comparison with the Population data 22 shows just how much GDP is more of a determining factor when predicting medal-amounts with medal-count to population ratio varying quite a bit for different countries.

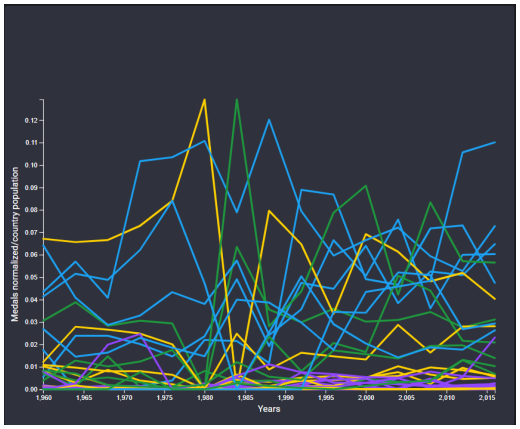

Figure 4: Medals/Population relation for summer Olympics

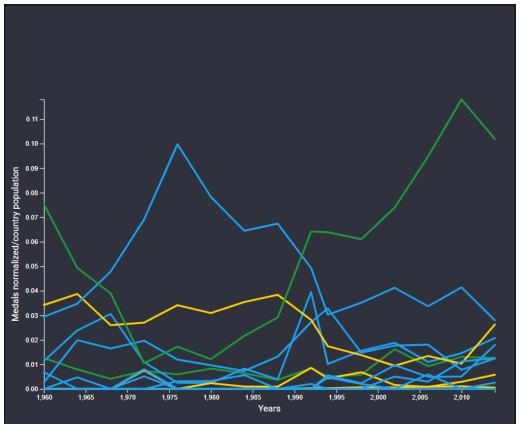

Figure 5: Medals/Population relation for winter Olympics

3.4 Gender differences

When showing the same analytics for the two genders separately, we can see how much more erratic women perform compared to their country's GDP, making it harder to predict their outcome. This acts as confirmation that women performance is also influenced by other factors as Leed's research showed.

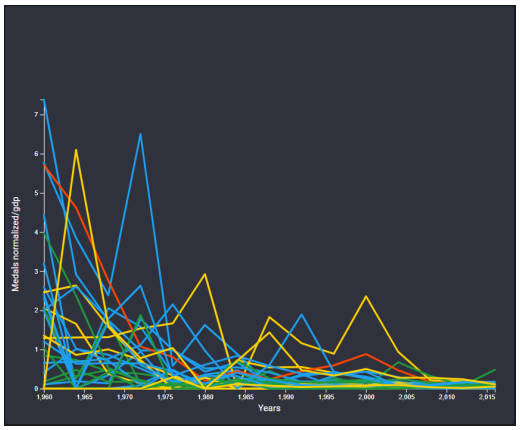

Figure 6: Medals/GDP relation for men at summer Olympics

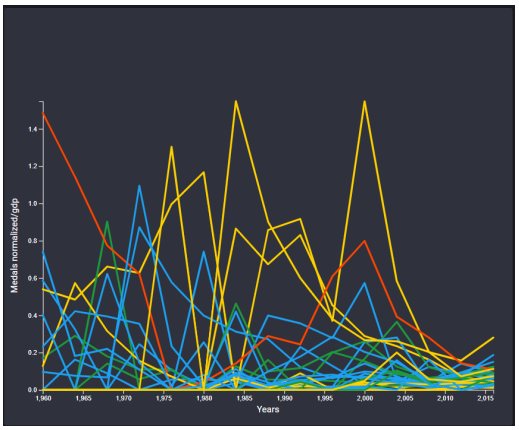

Figure 7: Medals/GDP relation for women at summer Olympics

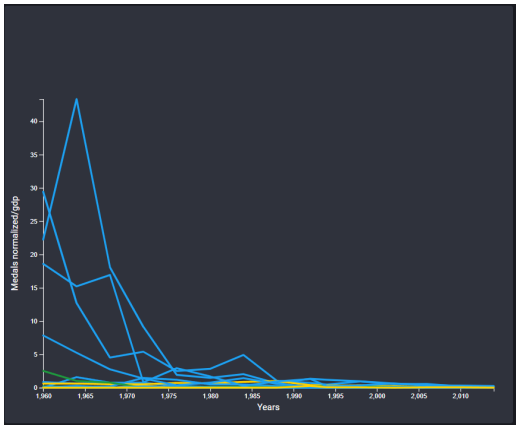

Figure 8: Medals/GDP relation for men at winter Olympics

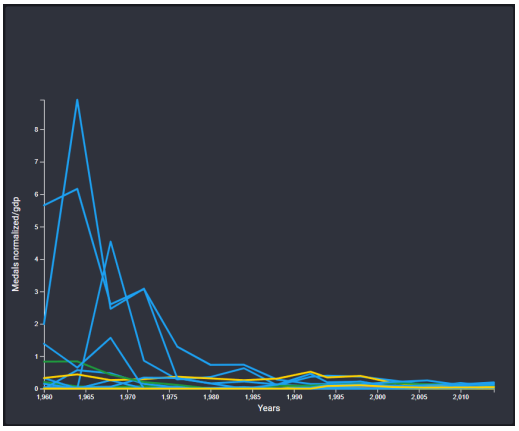

Figure 9: Medals/GDP relation for women at winter Olympics

3.5 GDP impact in different sports and the power of tradition

Having seen the striking impact of GDP on national performance at the Olympics, now we shift our focus on trying to understand which sports are effected by GDP the most. To be able to conduct such research, firstly we need to elaborate on the normalization process we introduced in the calculations. It wouldn't be fair to compare directly sports like Football and Swimming because swimming simply offers a lot more medals divided in multiple events, therefore, each medal has its value calculated as 1/number of events for that sport, year by year.

We start by looking at the most popular and iconic sport category at the Olympics, Athletics. To no surprise, we find that athletics national performance is extremely correlated to GDP with an average of 0.006 medals every 10 billion dollars of GDP in recent, but to which extent is this sports correlated compared to others?

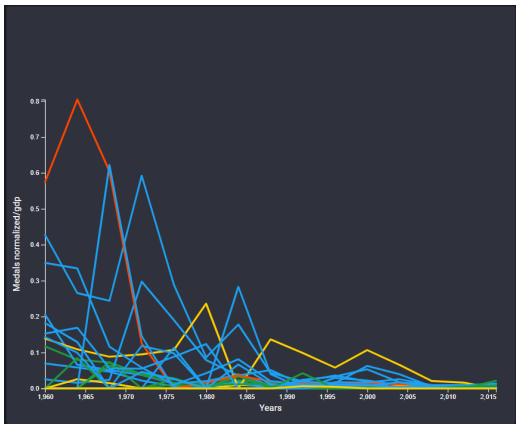

Figure 10: Medals/GDP relation in Athletics

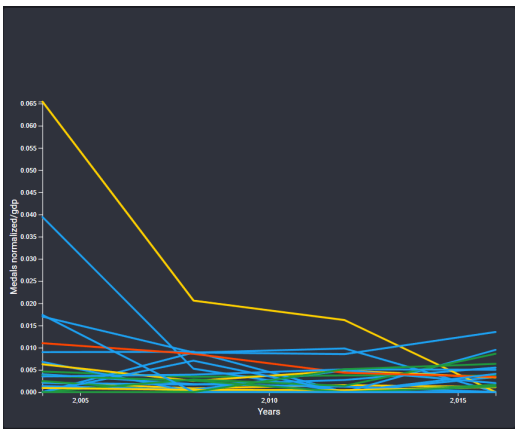

Figure 11: Medals/GDP relation in Athletics 2004+

Let's now compare it with Fencing, it's immediately obvious that medals are awarded to fewer countries who are specialized in such discipline and GDP is a looser contributor in determining success with countries to GDP ratios oscillating much more compared to Athletics.

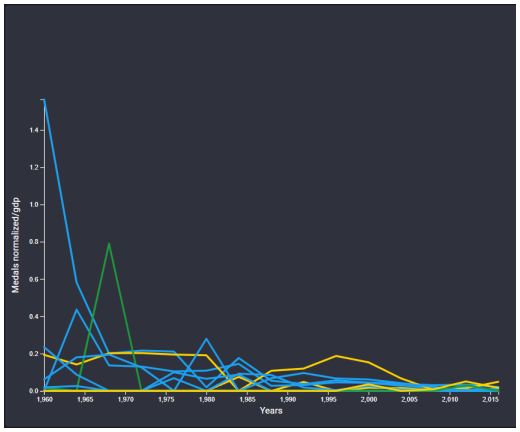

Figure 12: Medals/GDP relation in Fencing

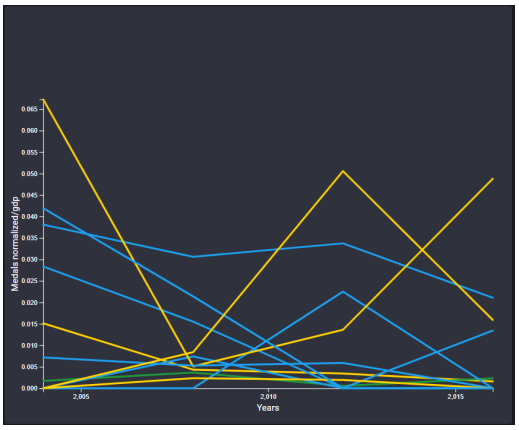

Figure 13: Medals/GDP relation in Fencing 2004+

In fact, we can show that the top countries in fencing, namely Russia, South Korea, Italy and France have fencing as one of their best sports, showing how much tradition can influence expected performance.

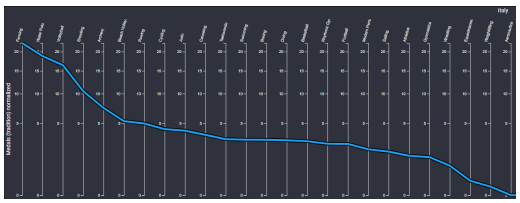

Figure 14: Italy's best sport is fencing

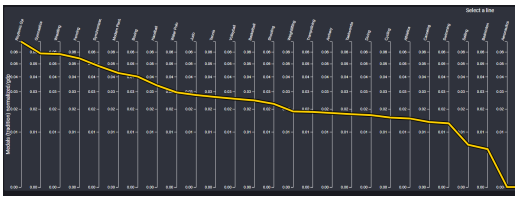

Figure 15: Russia's fourth best sport is fencing

Another striking example can be made of Sailing. In this case, medals are awarded to quite a few countries, but the relation to GDP is still looser than Athletics and all top performing countries have a long tradition of Sailing, like Greece, Australia, Sweden and Argentina.

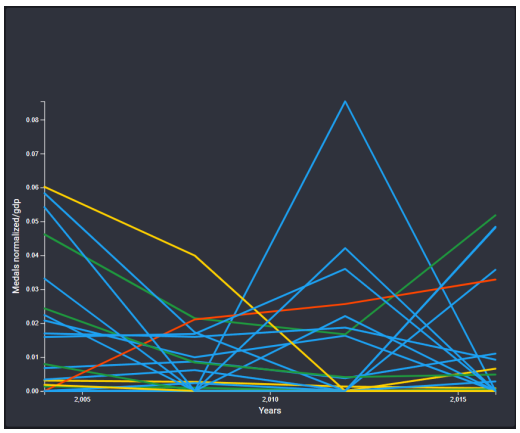

Figure 16: Medals/GDP relation in Sailing 2004+

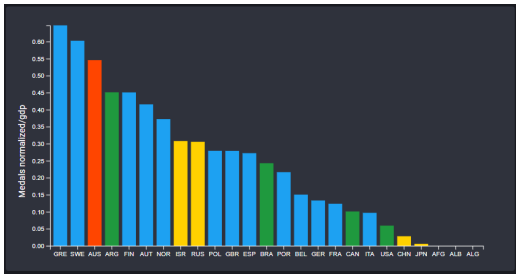

Figure 17: Best Countries for Medals/GDP ratio

4 Further insights

Having built a feature rich application, other insights have emerged that weren't related to our main research focus, here we can find some of them.

4.1 Women tradition

In the first editions of Olympic games women could not take part to many sports and so could not heavily contribute to build up the national medal table. During last editions of Olympic games the thing changed and the participation is today akin. As a general result, some countries got their position in the ranking mostly with the medals won with the contribute of one specific gender; in other words, if we take the charts about medals won by men and women separately, many countries do not hold the same position in the rankings. Let us consider, as an example, Italy and China during the last 30 years of Summer Olympic games (1986 - 2016). If we consider only medals won by men, Italy would be in position 6, while china in position 3; if we consider medals won by women, Italy would be in position 14, while China in position 1.

Upon further inspection, what we could gather is that Chinese women not only are the leading force of their nation's success, but they are also winning in many more sports then men, showing an unparalleled example of rich sport tradition for women. Comparing these results with Italy's traditions, the difference is very evident, women win in many less sports then men.

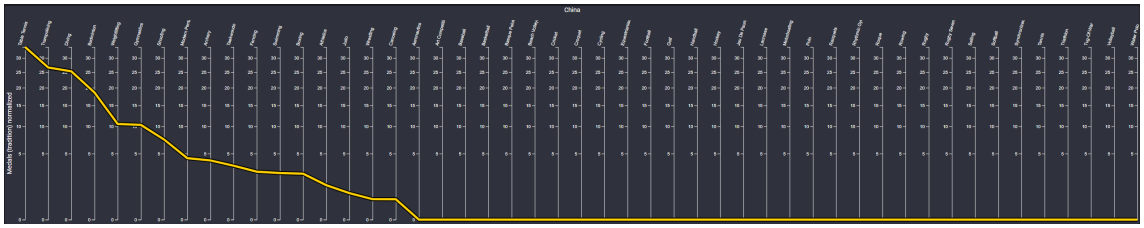

Figure 18: Best Sports for Chinese men

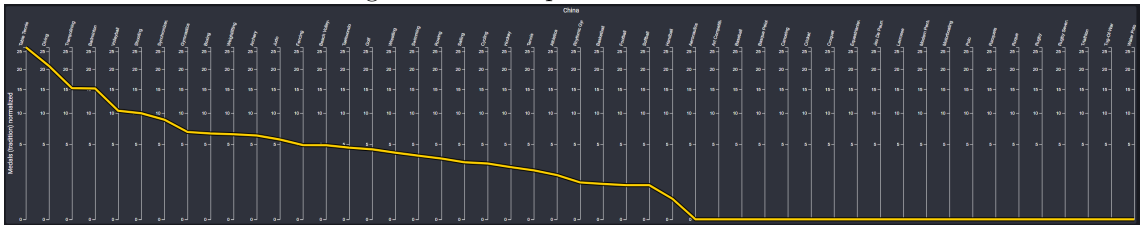

Figure 19: Best Sports for Chinese women

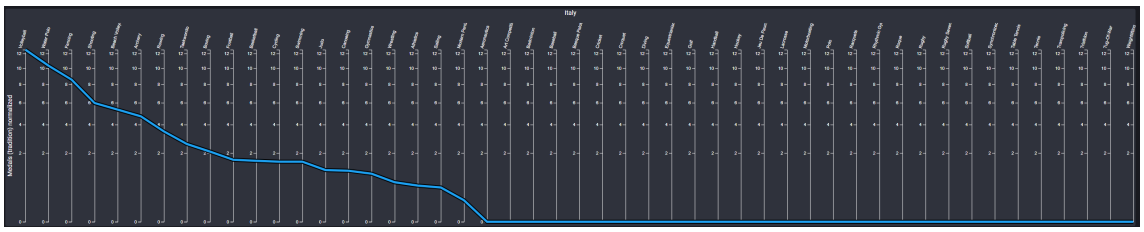

Figure 20: Best Sports for Italian men

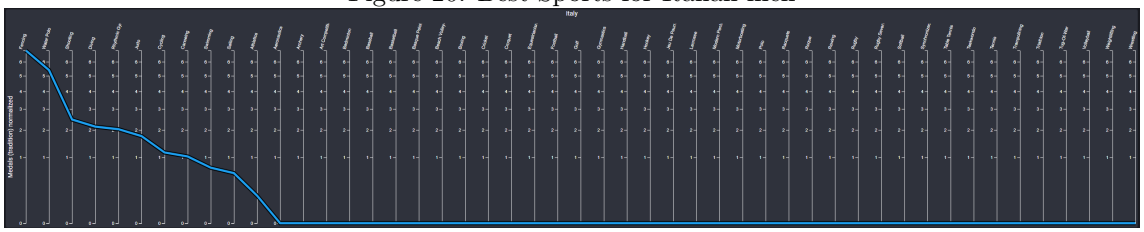

Figure 21: Best Sports for Italian women

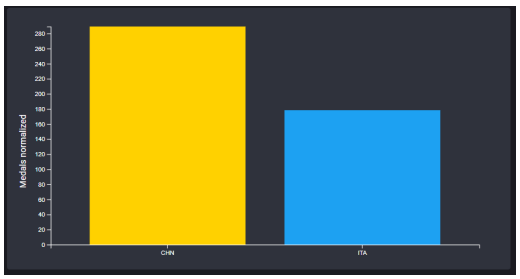

Figure 22: Men medals, China and Italy

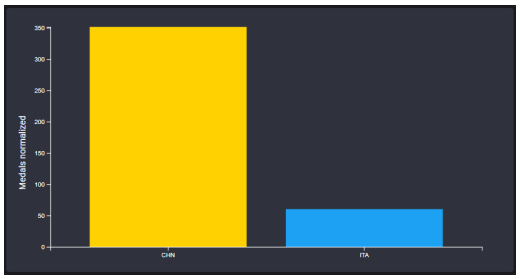

Figure 23: Women medals, China and Italy

From the barchart we can also see that Italian men win almost 3 times as many medals as women, while Chinese women win more then Chinese men.

We're not going to propose a reason for this behaviour as there could be a number of possible factors that range from socio-political to economic ones, it would be beyond our scope.

It must also be said that these trends stand out from the average traditions of the most performing countries, in fact, for countries such as France, USA, Australia, Russia, Germany or UK, to name a few of the most performing and most diverse, all have roughly the same amount of sports in their men and women tradition even though the relation between total medals won by men and total medals won by women varies wildly from nation to nation. Also it's not the case the Italian men win in more sports than other countries but rather Italian women win in far less sports than other countries.

5 Conclusions

We have visually shown how GDP remains the best predictor of Olympic success, but we have also shown that GDP is not as good of a predictor when dealing with women and single sports, these tend to vary a lot more and sports mainly depend on country specializations, which are by far the best predictors when it comes to single sport results. We have also shown that winter Olympics are even more GDP dependent than summer Olympics and that Population is not as big of a factor as it seems.

Some unexpected results also arose from our research, such as the the peculiar lack of success in diversification for Italian women and the striking success of Chinese women instead, which is something that seems to contradict Leads' research about women success being dependant on national women's condition, in fact, the Gender Inequality Index (GII) for Italy is quite smaller than China's, and it's the same index used in Leads' research paper.

References

[1] Bernard, A. B., and Busse, M. R. Who wins the olympic games: Economic resources and medal totals. *Review of Economics and Statistics* (2000).

[2] Marikova, L. E., and Leeds, M. Gold, silver, and bronze: Determining national success in men's and women's summer olympic events. *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik) 232*, 3 (2012), 279–292.

[3] Rathke, A., and Woitek, U. Economics and olympics: An efficiency analysis. *Institute for Empirical Research in Economics* (2007).

[4] Smith, M., and Alison M. Wrynn, M. Women in the 2000, 2004 and 2008 olympic and paralympic games. *A Women's Sports Foundation Research Report* (2009).