



36104- Data Visualization and Narratives

# VISUAL ANALYSIS OF OLYMPICS DATA

Abhisek Gautam

13679042



## Introduction

Olympic Games is a well-known sporting platform which is recognized all over the world, has been distinguished from the late 19<sup>th</sup> century. Its origin however can be traced back to the Greek empire, at around 3,000 years ago, which consisted only of sprint race, and was held in Greece's city Olympia only accessible to freeborn Greek people (Young & Abraham, 2020). It has grown since then and now has become a hub for all the athletes worldwide to demonstrate their abilities in more than 28 individual sporting contests. Currently, it is held every 2 years in different countries, with names Summer Olympics and Winter Olympics, both having their own set of games (Young & Abraham, 2020). It has become a place which reflects the power of the individual contestants and has become a source of pride for the countries they represent.

## Motivation and Research Questions

Olympics has been there during different revolutions consisting of industries, science, wars and the current technological era. Keeping these things in mind, this report attempts to build around the following questions:

1. What effect does the host country have in Olympics?
2. Is the performance of countries in Olympic games affected by the economic factors of the country?
3. Is the age of winning Olympics changing?

## Literature Review

The advantage of host country in any sporting activity is well known, as the participants will have familiarity of the field, and also there is a great support from the home crowd. Host countries are expected to win 3 times the medals that they were winning while playing as away (Clarke, 2000). Being a host nation and also having a communist background is also going to have a positive effect in the number of medals won (Bian, 2005).

It is found that a country's socio-economic variables, such as GDP affects the country's performance in the Olympics by a great factor. Country's population and its GDP is seen to have a correlation with the number of medals won in Olympics (Bian, 2005).

Age factor is also one of the important ones when it comes to sports and even among the athletes of the same age, relative age effect (RAE) comes into factor which determines who triumphs (Fletcher & Sarkar, 2012). RAE argues that an athlete can have more advantage as compared to another who is younger by almost a year with respect to maturity, experience and early specialization (Neill, Cotton, Cuadros & Connor, 2016).

## Approach

To answer the questions, two datasets will be used namely, "120 years of Olympic history: athletes and results<sup>1</sup>", and "Gapminder GDP per capita, constant PPP dollars- v25<sup>2</sup>". The Olympics dataset has the names of participants, their demographics, which sport they participated in and on which Olympic games. Gapminder's GDP per capita dataset consists

---

<sup>1</sup> <https://www.kaggle.com/heesoo37/120-years-of-olympic-history-athletes-and-results>

<sup>2</sup> <https://www.gapminder.org/data/documentation/gd001/>

of GDP of countries from 1960 forecasted till 2040. A custom dataset was also created that maps the cities mentioned in the Olympics dataset to country names.

Visualizations are created in Tableau and Python. Initial data cleaning is done on Excel, and visualization specific data manipulation will be carried out as needed on Tableau and Python.

## Findings

Figure 1 shows the number of times a country has hosted Olympics. The USA has hosted it the most, i.e. 8 times. After that, it is France.

### Which countries have hosted Olympics how many times?

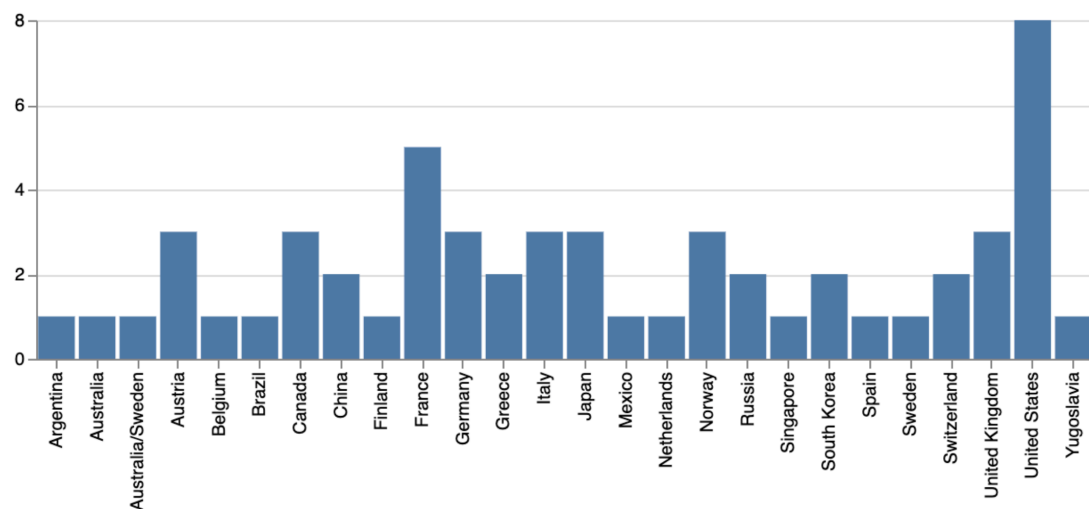


Figure 1- Countries that have hosted Olympics

When we look at the number of medals by host countries, we see that USA has hosted it the most, followed by Germany (Figure 2). So, this somewhat hints the question that home teams might have some advantage over other nations.

### Medals by Host Country

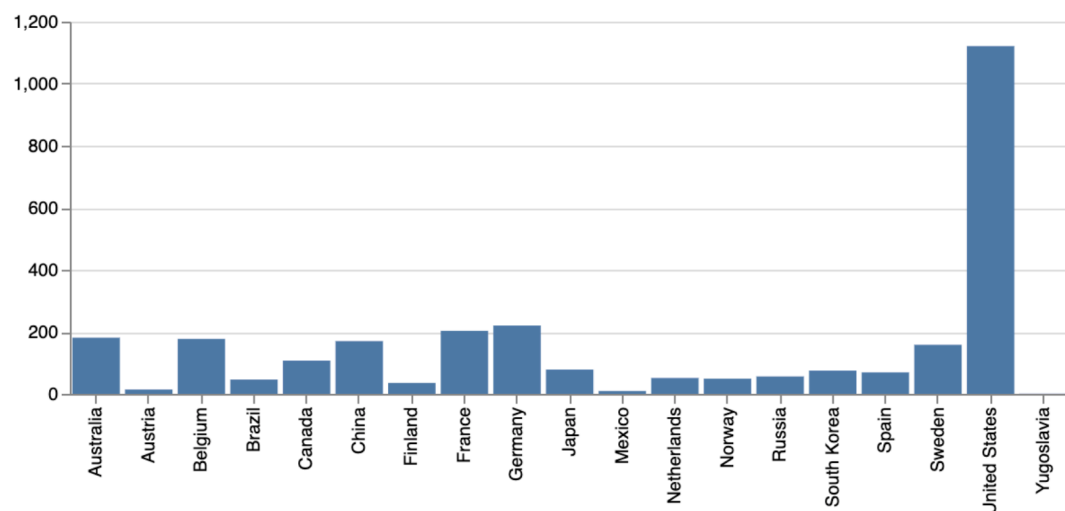


Figure 2- Medals won by Host Countries

Figure 3 shows the mean age of the medal winners from the start of Olympic Games. The graph shows that the mean age was high during the early 1900s, which was around 27-28, had some fluctuations in the middle era (which could have been due to World Wars). It decreases to 24-25 during 1960s till the 90s, from where it goes up to 26-27 in the early 21<sup>st</sup> century. The decrease in the average age during 60s till 90s may have been the aftereffect of world wars, where countries started developing their infrastructure after the destruction of the wars.

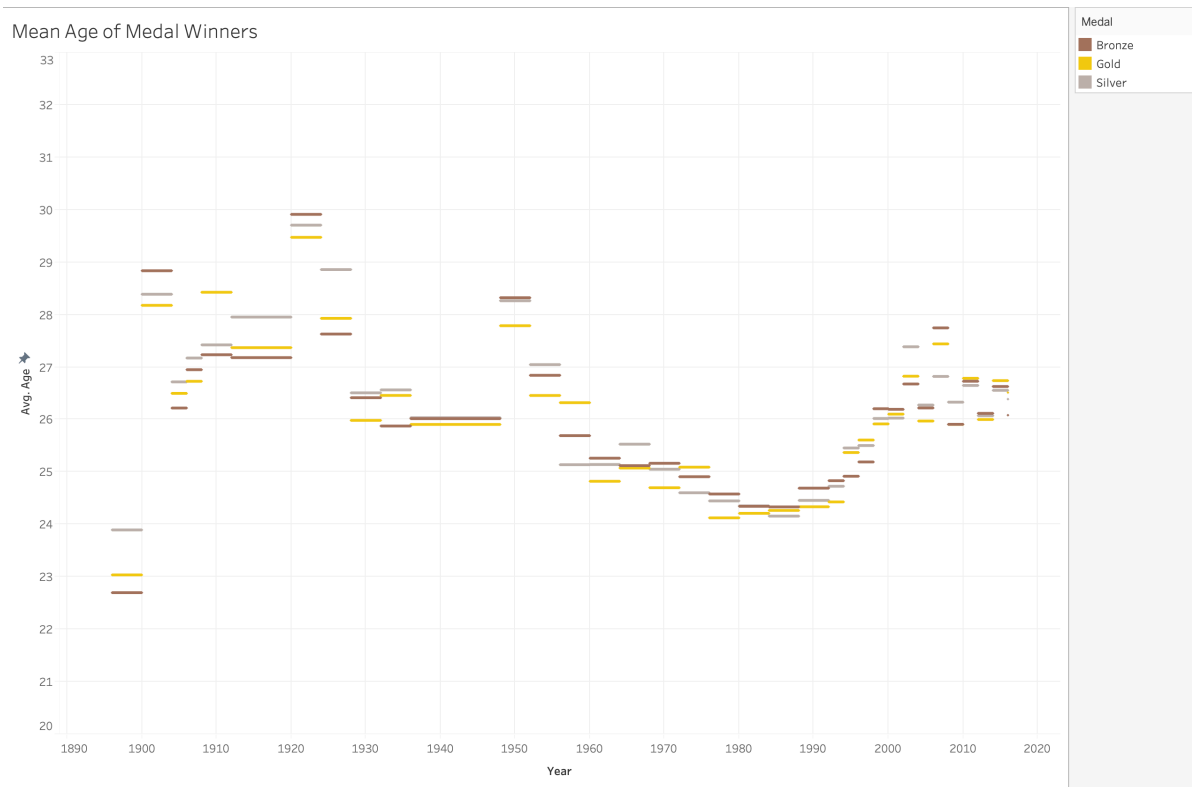


Figure 3- Mean Age of Medal Winners in Olympics

## References

Mahtani, K.R., Protheroe, J., Slight, S.P., Demarzo, M.M.P., Blakeman, T., Barton, C. A., Brijnath, B., Roberts, N. (2012). *Can the London 2012 Olympics 'inspire a generation' to do more physical or sporting activities? An overview of systematic reviews*. Retrieved from <https://bmjopen.bmj.com/content/3/1/e002058.full>

Clarke, S. R. (2000, June). Home advantage in the Olympic Games. *In Proceedings of the 5th Australian Conference on Mathematics and Computers in Sport. University of Technology Sydney, Australia* (pp. 76-85). Retrieved from <https://researchbank.swinburne.edu.au/file/3bbcf005-ec1e-4def-9602-f7dd18d0a711/1/PDF%20%28Published%20version%29.pdf>

Fletcher, D., & Sarkar, M. (2012). A grounded theory of psychological resilience in Olympic champions. *Psychology of sport and exercise*, 13(5), 669-678. Retrieved from <https://www.sciencedirect.com/science/article/abs/pii/S1469029212000544>

Bian, X. (2005). Predicting Olympic medal counts: The effects of economic development on Olympic performance. *The park place economist*, 13(1), 37-44. Retrieved from [https://www.researchgate.net/profile/Xun\\_Bian/publication/28328103\\_Predicting\\_Olympic\\_Medal\\_Counts\\_the\\_Effects\\_of\\_Economic\\_Development\\_on\\_Olympic\\_Performance/links/5525b9850cf24b822b4058f9/Predicting-Olympic-Medal-Counts-the-Effects-of-Economic-Development-on-Olympic-Performance.pdf](https://www.researchgate.net/profile/Xun_Bian/publication/28328103_Predicting_Olympic_Medal_Counts_the_Effects_of_Economic_Development_on_Olympic_Performance/links/5525b9850cf24b822b4058f9/Predicting-Olympic-Medal-Counts-the-Effects-of-Economic-Development-on-Olympic-Performance.pdf)

Bernard, A. B., & Busse, M. R. (2000). *Who wins the Olympic Games: Economic development and medal totals* (No. w7998). National Bureau of Economic Research. Retrieved from <https://www.nber.org/papers/w7998.pdf>

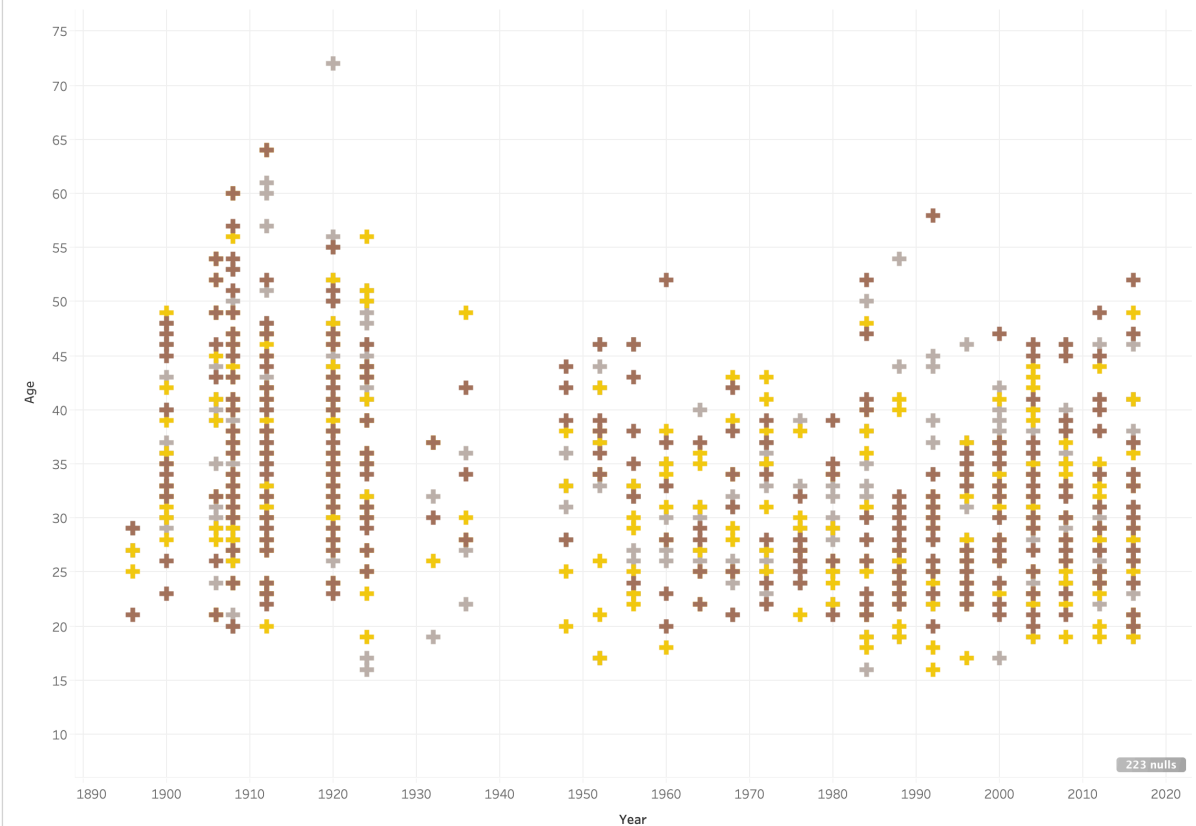
O'Neill, K. S., Cotton, W. G., Cuadros, J. P., & O'Connor, D. (2016). An investigation of the relative age effect amongst Olympic athletes. *Talent Development & Excellence*, 8(1), 27-39. Retrieved from [https://www.researchgate.net/profile/Juan\\_Cuadros2/publication/317582562\\_An\\_Investigation\\_of\\_the\\_Relative\\_Age\\_Effect\\_amongst\\_Olympic\\_Athletes/links/5cb4a254299b12097682856/An-Investigation-of-the-Relative-Age-Effect-amongst-Olympic-Athletes.pdf](https://www.researchgate.net/profile/Juan_Cuadros2/publication/317582562_An_Investigation_of_the_Relative_Age_Effect_amongst_Olympic_Athletes/links/5cb4a254299b12097682856/An-Investigation-of-the-Relative-Age-Effect-amongst-Olympic-Athletes.pdf)

Young, D. C., Abrahams, H. M. (2020). *Olympic Games*. Retrieved from <https://www.britannica.com/sports/Olympic-Games>

## Appendix I

### RAE analysis of Shooting

RAE among various sports - Shooting



Not clear, may need to craft variables in order to show the RAE