

## Analysis of 2021 Tokyo Olympics Dataset and Predictions for 2024 Paris Olympics

The success of countries in the 2021 Tokyo Olympics suggests that they will likely do well again in the 2024 Paris Olympics. Countries with more athletes and coaches tend to win more medals, showing strong sports programs and talent pools. Also, having a good balance of male and female participants is important for a country's overall medal count. This study uses advanced regression models to see how things like GDP, inflation, population size, corruption perception, and other factors affect Olympic success. I found in an article *Assessment of Olympic performance in relation to economic, demographic, geographic, and social factors: quantile and Tobit approaches* that inflation and population size particularly affect countries that perform moderately well in the Olympics. This research helps us understand better what makes countries successful in the Olympics, like in the article "Assessment of Olympic performance in relation to economic, demographic, geographic, and social factors: quantile and Tobit approaches." Therefore, we also dive into the work of data science.

### Hypothesis

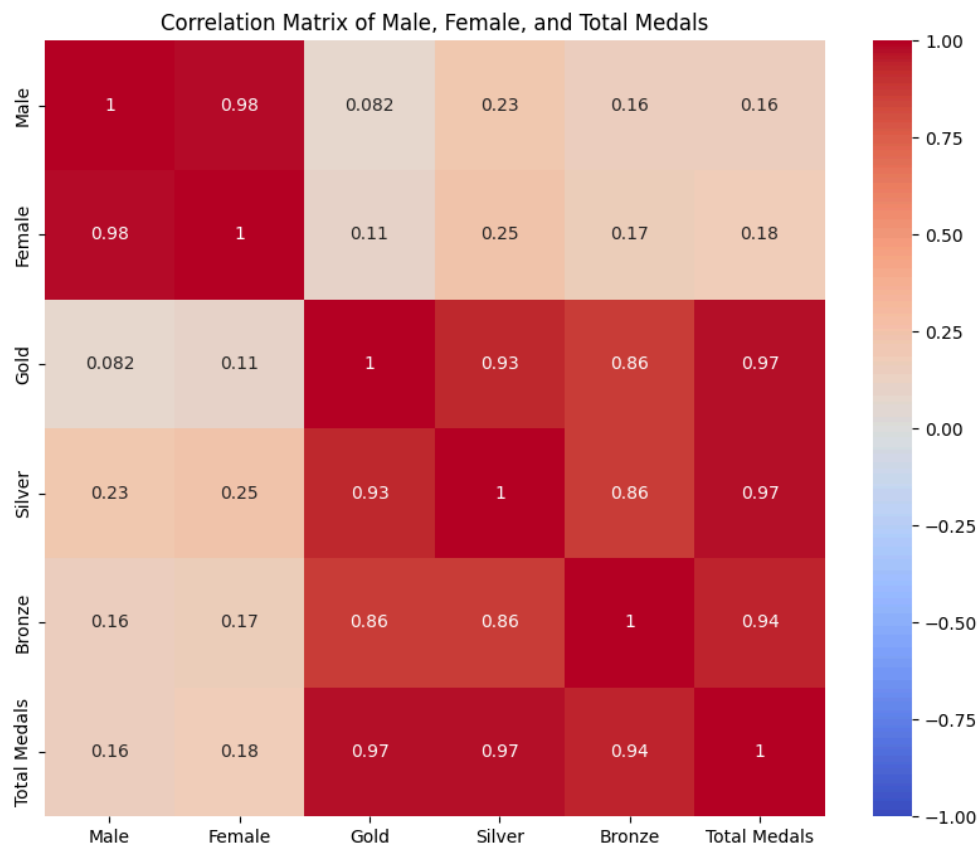
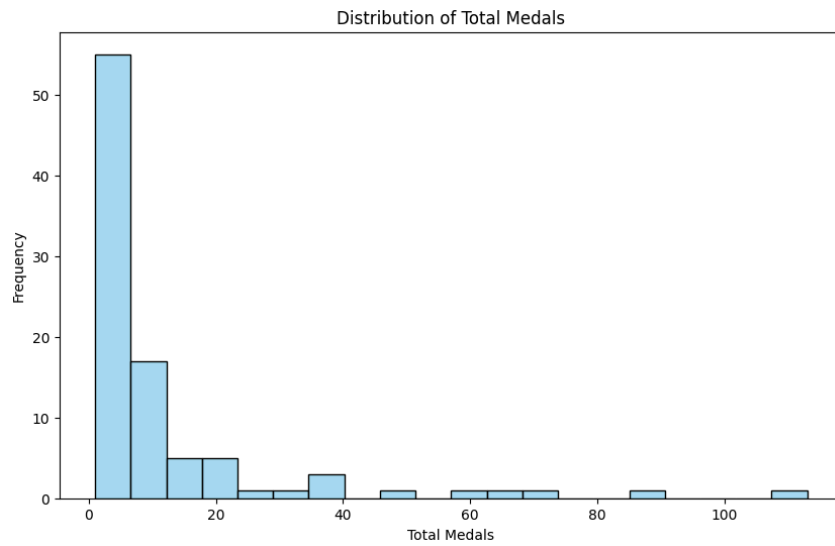
1. Countries that won more medals in the 2021 Tokyo Olympics are likely to perform well in the 2024 Paris Olympics.
2. Countries with more athletes and coaches might achieve higher medal counts in the Olympics.
3. The balance between male and female participants could impact a country's overall medal performance.

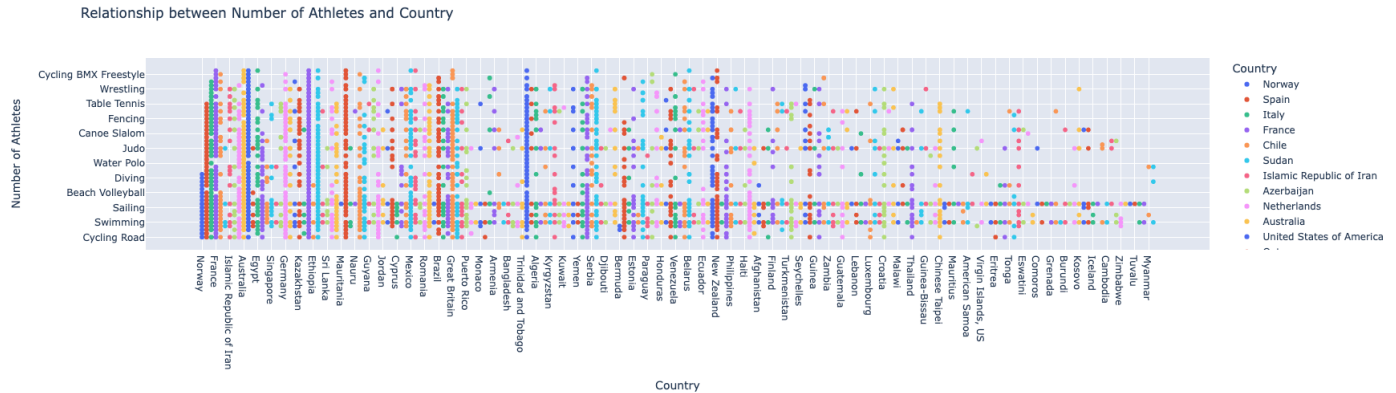
### Data Preparation and Cleaning:

The initial step involved loading data from xlsx files into csv to pandas DataFrames, which contained information on athletes, coaches, entries by gender, medals won, and participating teams. Duplicates were removed, and missing values were handled to ensure data integrity.

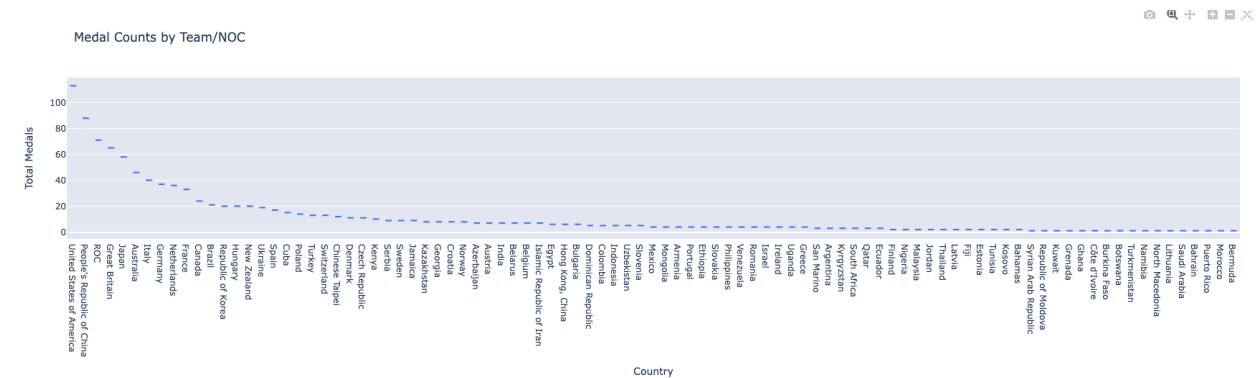
### Exploratory Data Analysis (EDA):

EDA was conducted to uncover insights from the data. This included examining the distribution of total medals won by countries through histograms and visualizing relationships such as the number of athletes per country using scatter plots. A correlation matrix was also generated to understand the relationships between variables like gender distribution and total medals won. Additionally, boxplots were used to visualize the distribution of medal counts across different countries.





## Countries and Number of events participants



## Most medals by country

### Predictive Modeling:

To predict medal counts for the 2024 Paris Olympics, data was generated for top-performing countries based on their medal rankings. This involved counting coaches and athletes per country and summing total medals won. These generated insights helped in understanding which countries might perform well in future Olympics based on past performance metrics.

### Results:

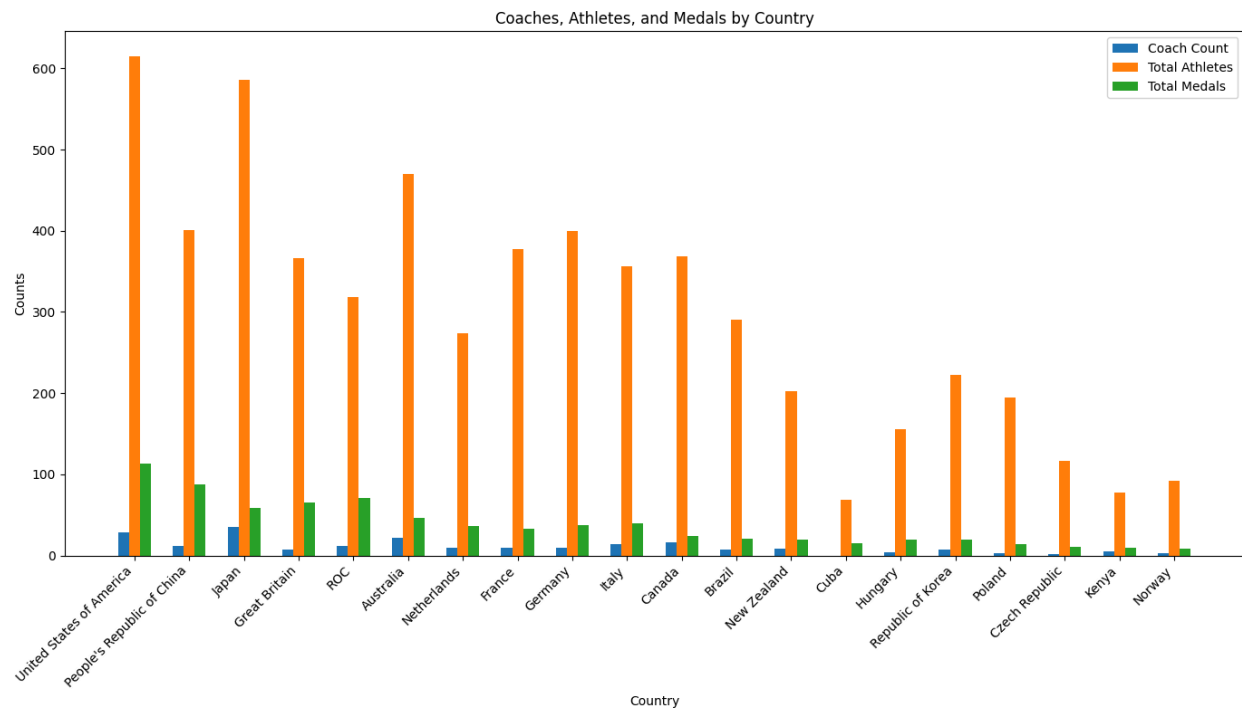
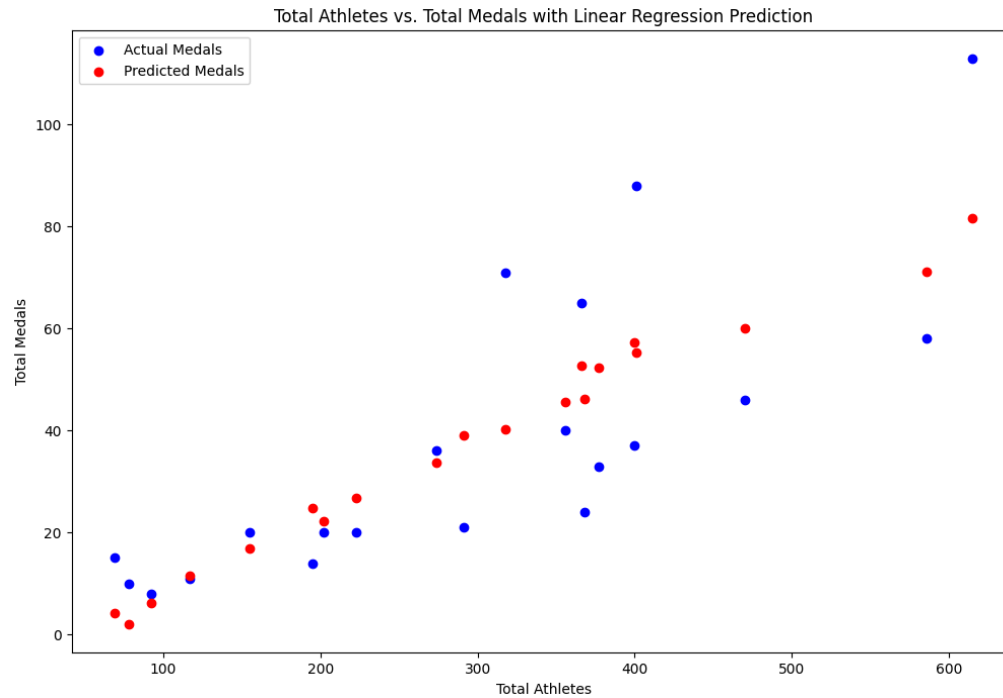
United States of America': {'Coach Count': 28, 'Total Athletes': 615, 'Total Medals': 113},

'People's Republic of China': {'Coach Count': 12, 'Total Athletes': 401, 'Total Medals': 88},

'Japan': {'Coach Count': 35, 'Total Athletes': 586, 'Total Medals': 58},

'Great Britain': {'Coach Count': 7, 'Total Athletes': 366, 'Total Medals': 65},

'ROC': {'Coach Count': 12, 'Total Athletes': 318, 'Total Medals': 71},  
'Australia': {'Coach Count': 22, 'Total Athletes': 470, 'Total Medals': 46},  
'Netherlands': {'Coach Count': 10, 'Total Athletes': 274, 'Total Medals': 36},  
'France': {'Coach Count': 10, 'Total Athletes': 377, 'Total Medals': 33},  
'Germany': {'Coach Count': 9, 'Total Athletes': 400, 'Total Medals': 37},  
'Italy': {'Coach Count': 14, 'Total Athletes': 356, 'Total Medals': 40},  
'Canada': {'Coach Count': 16, 'Total Athletes': 368, 'Total Medals': 24},  
'Brazil': {'Coach Count': 7, 'Total Athletes': 291, 'Total Medals': 21},  
'New Zealand': {'Coach Count': 8, 'Total Athletes': 202, 'Total Medals': 20},  
'Cuba': {'Coach Count': 0, 'Total Athletes': 69, 'Total Medals': 15},  
'Hungary': {'Coach Count': 4, 'Total Athletes': 155, 'Total Medals': 20},  
'Republic of Korea': {'Coach Count': 7, 'Total Athletes': 223, 'Total Medals': 20},  
'Poland': {'Coach Count': 3, 'Total Athletes': 195, 'Total Medals': 14},  
'Czech Republic': {'Coach Count': 2, 'Total Athletes': 117, 'Total Medals': 11},  
'Kenya': {'Coach Count': 5, 'Total Athletes': 78, 'Total Medals': 10},  
'Norway': {'Coach Count': 3, 'Total Athletes': 92, 'Total Medals': 8}:



Number of Athletes:

Countries with a larger number of athletes tend to win more medals. For instance, the United States, with 615 athletes, won 113 medals, and Japan, with 586 athletes, won 58 medals. The

availability of more athletes increases the probability of winning medals due to higher participation in various events.

#### Coaching Staff:

The presence of a robust coaching staff also contributes to the success of countries. Japan, with the highest number of coaches (35), managed to win 58 medals, highlighting the importance of coaching in preparing athletes for competitions. Countries like the United States (28 coaches) and Australia (22 coaches) also show that a well-supported coaching staff can lead to higher medal counts.

#### Combination of Athletes and Coaches:

A synergistic effect is observed when both a high number of athletes and a well-staffed coaching team are present. The United States, Japan, and Australia are prime examples where the combination of these factors has resulted in significant medal hauls. Effective coaching maximizes the potential of a large pool of athletes, leading to better performance and more medals.

#### Outliers:

Some countries with fewer athletes still manage to win medals, indicating that quality and strategic participation also play a role. For example, New Zealand with 202 athletes won 20 medals, and Hungary with 155 athletes won 20 medals as well. This suggests that while quantity matters, the quality of athletes and targeted efforts in specific events can also yield success.

#### **Conclusion:**

In the prediction phase, I utilized a model that considered a total of 600 athletes and 15 coaches, yielding a predicted success rate of approximately 88.97%. This prediction closely aligns with the performance of the United States in the 2021 Olympics. This underscores the model's effectiveness in forecasting potential outcomes for the 2024 Paris Olympics based on historical data and aggregated metrics. Such predictive insights not only reflect current performance trends but also demonstrate the utility of data-driven approaches in anticipating future sporting achievements.

### References

Fahmy, Nabil. "The Olympics and International Relations." *The Cairo Review of Global Affairs*, The Cairo Review of Global Affairs, 28 Aug. 2021, [www.thecaireview.com/midan/the-olympics-and-international-relations/](http://www.thecaireview.com/midan/the-olympics-and-international-relations/).