

# Medals by Numbers: Uncovering the Secrets of Olympic Performance



**Group 5: Thaissa Champagne, Thomas Gartley, Adam Loux, and Noah Stevens**

In our project, we aseeked to find interesting trends, unexpected outliers and unique data stories for countries using historical data containing 126 Years of Olympic competition.

## Table of Contents:

- [Data Cleanup](#)
- [Research Questions](#)
- [Question 1: Host Nation Medal Performance](#)
- [Question 2: Medal Count to Population Performance](#)
- [Question 3: Men vs. Women In Events](#)
- [Regression Plot](#)
- [Bias and Limitations](#)
- [Conclusion](#)

## Data Cleanup

### Condense Historical Countries:

- Bohemia and Czechia
- Rhodesia and Zimbabwe
- Soviet Union and Russia

### Olympic and Global Politics:

- Russian Olympic Committee (ROC)
- Korea Town
- East and West Germany

### Split Categories For Gender



- Extract Gender From Event Details

... 5 hidden cells

### Team Medals

- Condensing team medals so they are only counted once

... 4 hidden cells

## Research Questions

### Host Nation Medal Performance:

- Which nations benefited the most via medal count from hosting the Olympics?
- Are there countries who didn't have advantage as compared to their average performance?

### Medal Count to Population Performance:

- How does the medal count per capita compare to overall medal count?
- Are there any large outliers for medal count when considering a per capita count?

### Men vs. Women

- How does the ratio of male and female events change

overtime?

- Are there any countries or events that are outliers with female athlete participation?

## Question 1: Host Nation Medal Performance

---

... 4 hidden cells

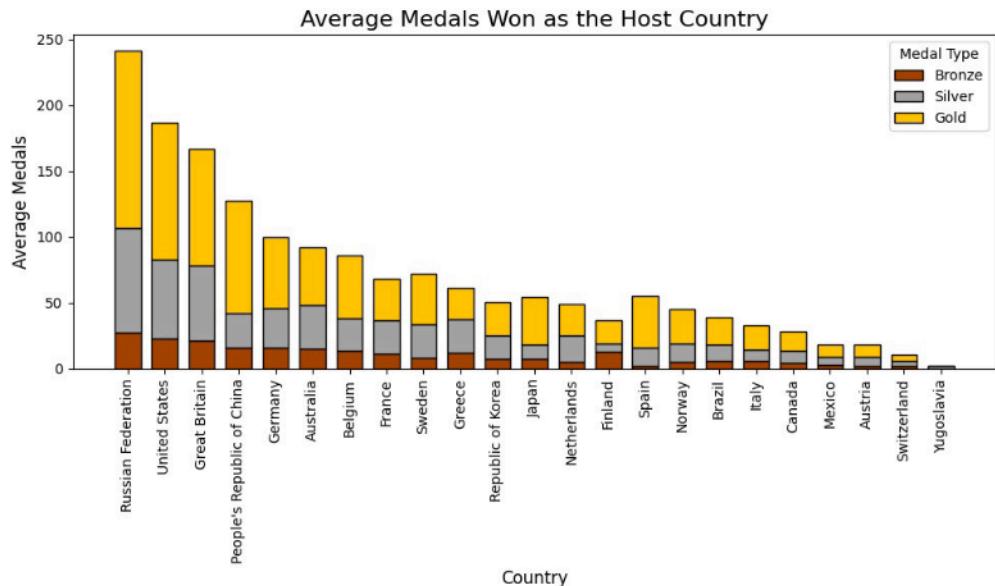
---

**The average medals won as the host country can possibly showcase outlier performances a country may have as the host country. The bar graph is sorted in descending order of the average total medal count. This will display shifts in ranking when we take a look at the average medal performance graph.**

---

>Show hidden code

---

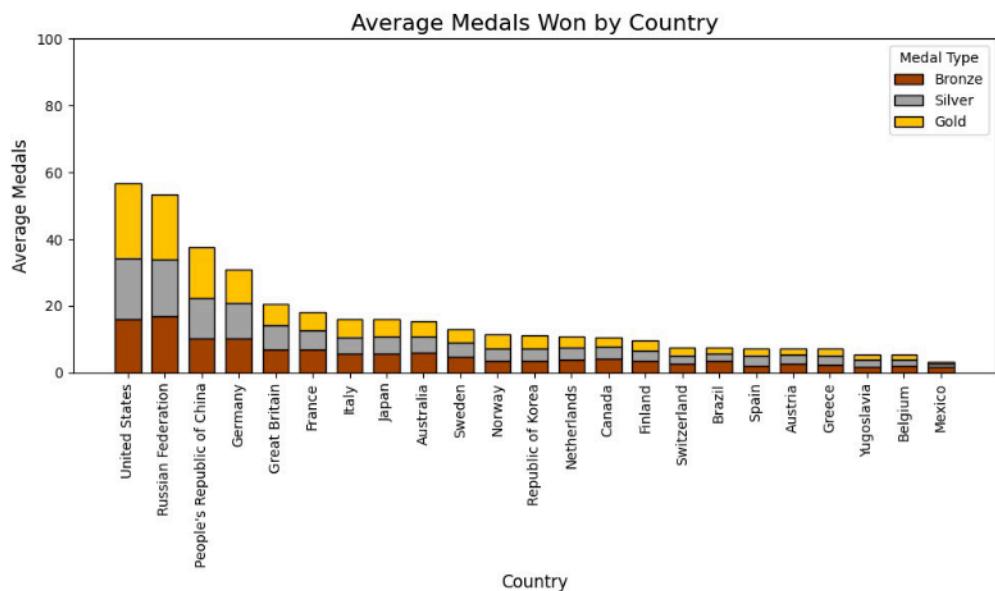


>Show hidden code

>Show hidden code

Show hidden code

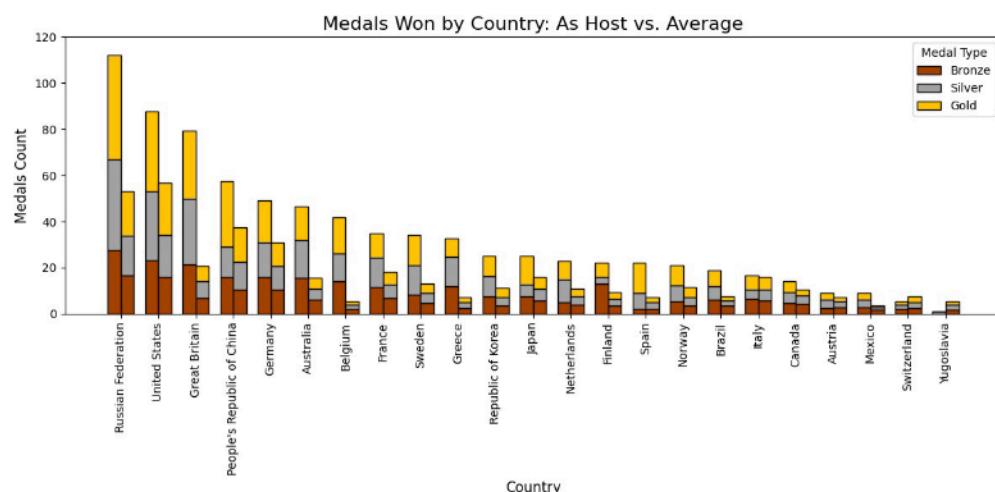
Show hidden code



Show hidden code

**Now, looking at the double bar graph below, there is a stark contrast between how almost all countries outperform their average Olympic performance when they are the host. Some countries have a more extreme difference than others.**

⌄ Show hidden code



## Question 2: Medal Count to Population Performance

**By comparing medal counts to population size, we can discover insights into which countries are outperforming and achieving extraordinary outcomes despite their size, and which larger nations might not be reaching their potential.**

---

... 10 hidden cells

---

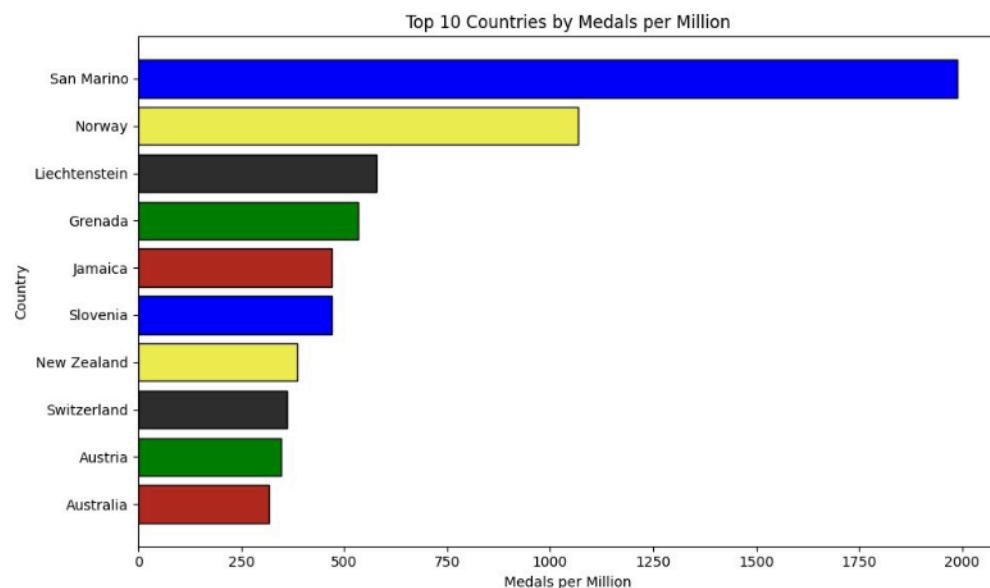
**San Marino is a large outlier for punching above their weight. The small microstate has only have won medals this year in Paris 2024.**

**Countries such as San Marino and Lichtenstein are smaller than most cities.**

---

>Show hidden code

---

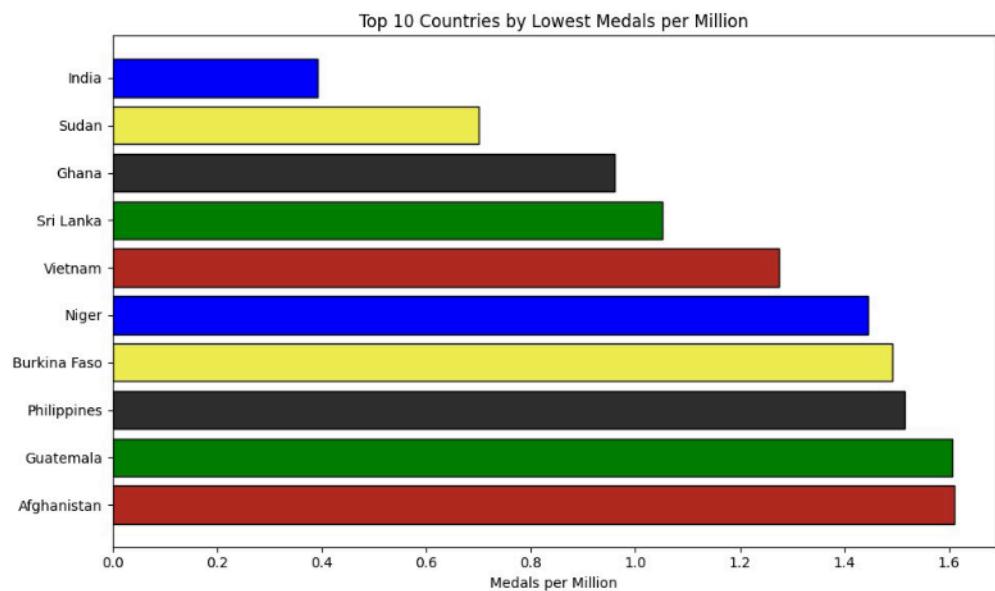


**India performs the worst on a basis of population to medals. For some of these countries, this data could direct towards other longterm extraneous factors such as possible lack of athletic investment, population interest, or even political instability.**

---

>Show hidden code

---



## Question 3: Men vs. Women in Events

>Show hidden code

>Show hidden code

>Show hidden code

>Show hidden code

0

0

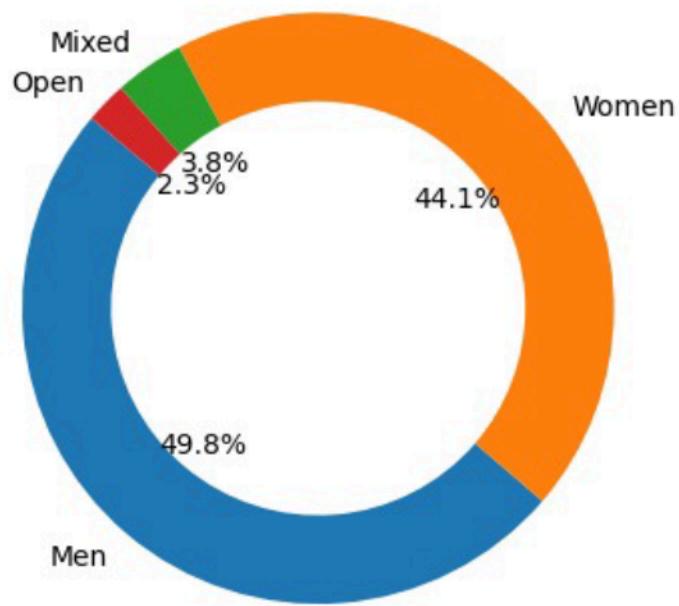
Out[36]:

	noc	country
--	-----	---------

... 4 hidden cells

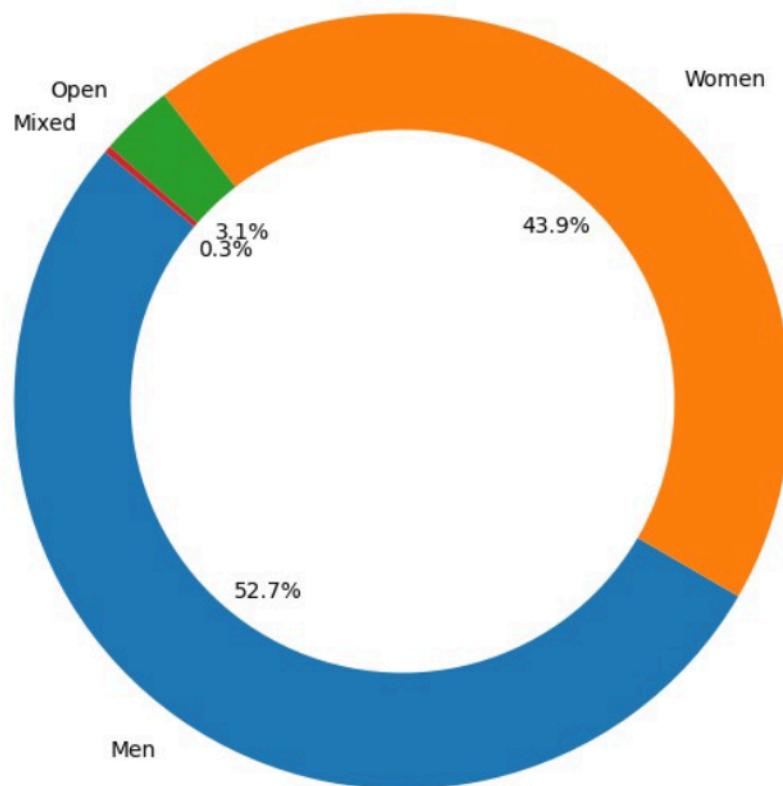
>Show hidden code

Total Medals by Gender Distribution



Show hidden code

Gender Distribution in 2004 Olympics

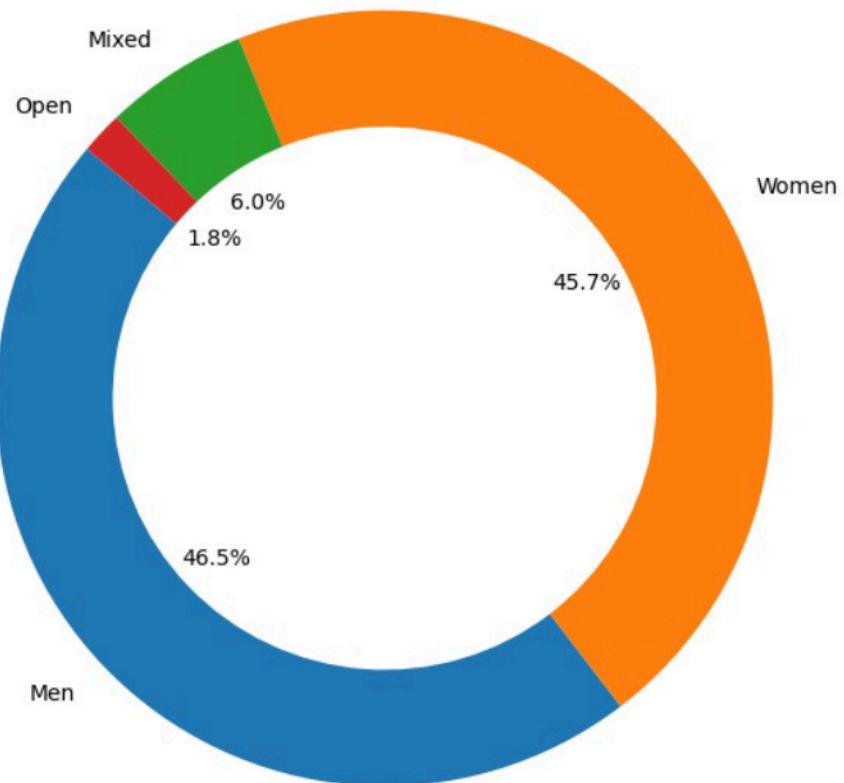


---

⌄ Show hidden code

---

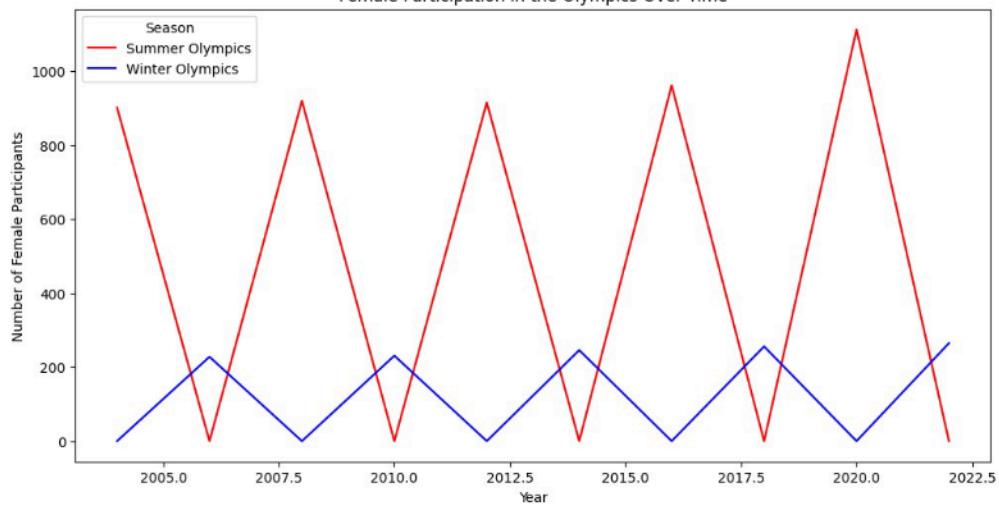
### Gender Distribution in 2020 Olympics



>Show hidden code

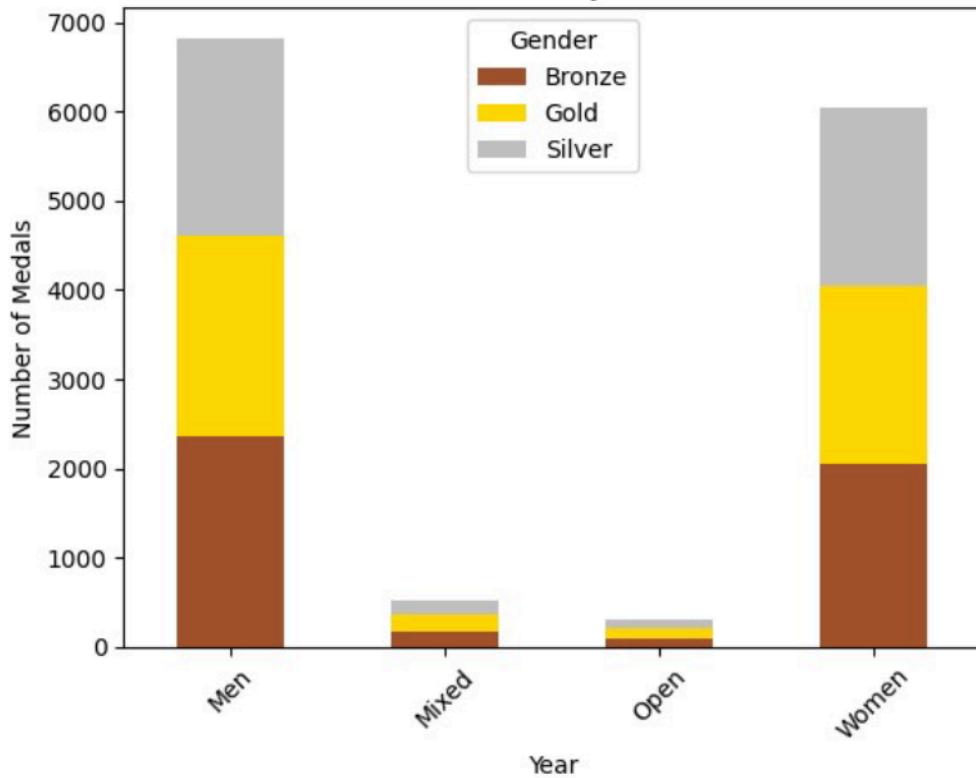
>Show hidden code

### Female Participation in the Olympics Over Time



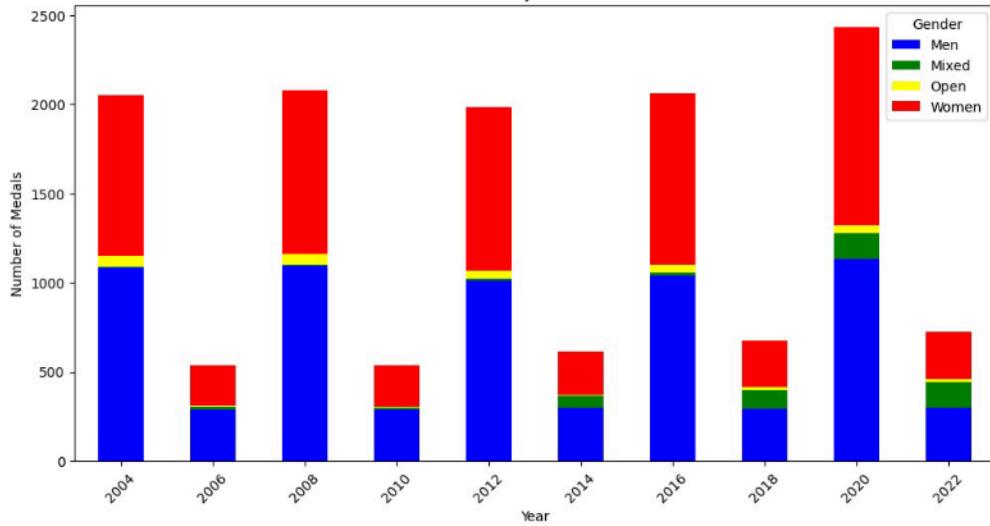
▼ Show hidden code

Total Medals by Gender

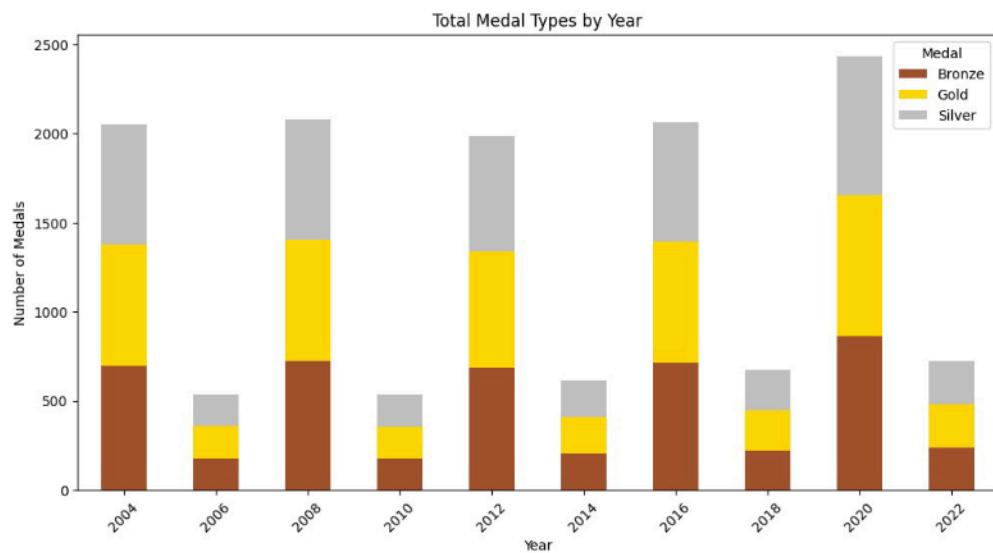


▼ Show hidden code

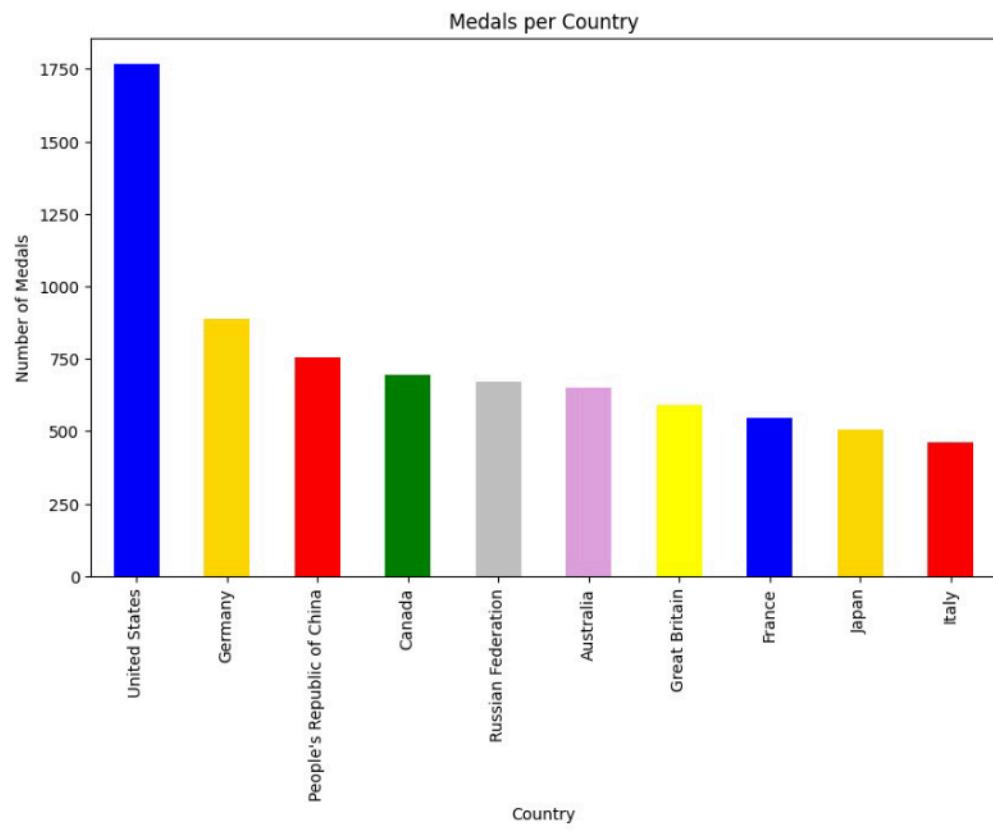
Total Medals by Gender vs Year



▼ Show hidden code

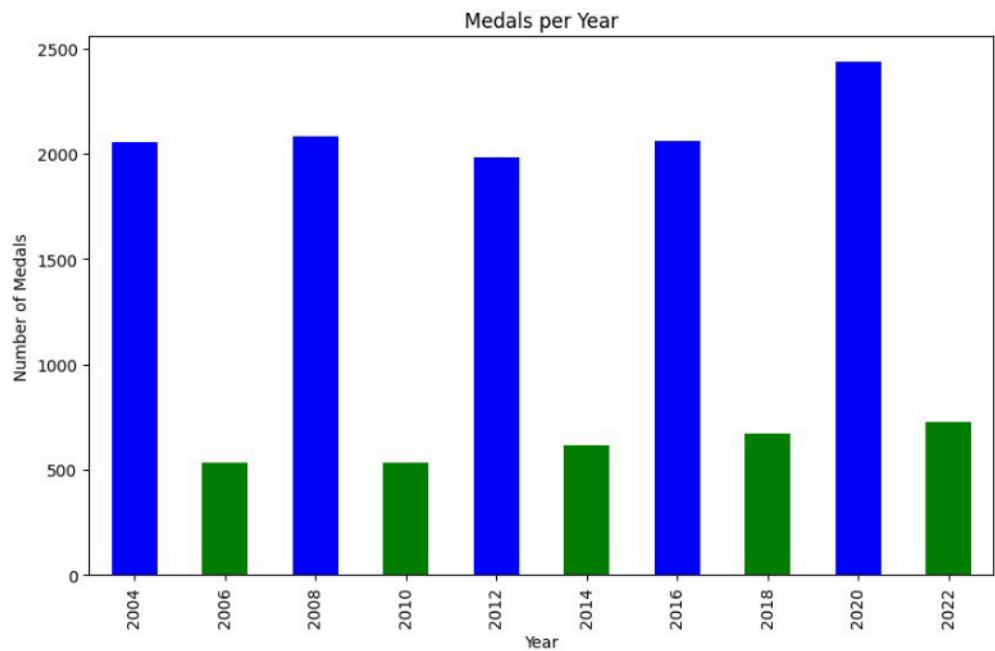


>Show hidden code

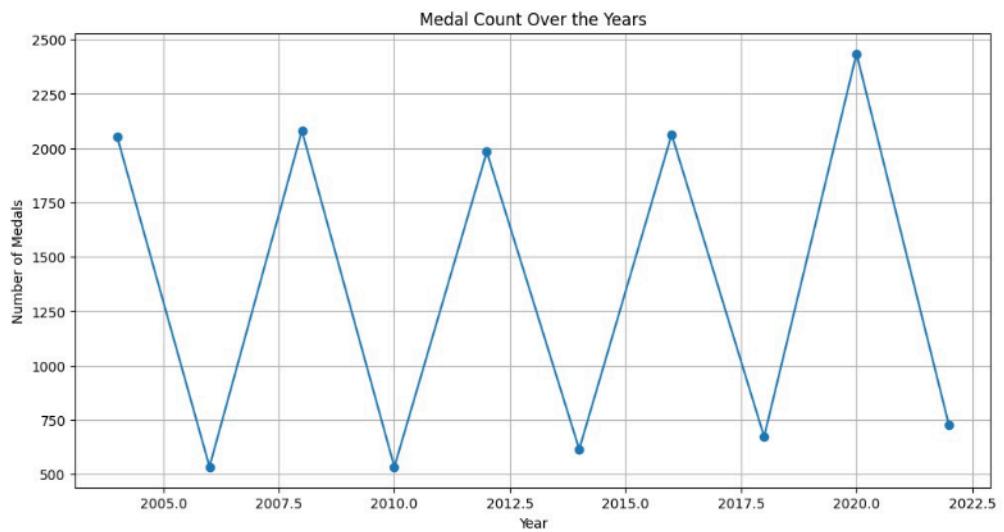


Show hidden code

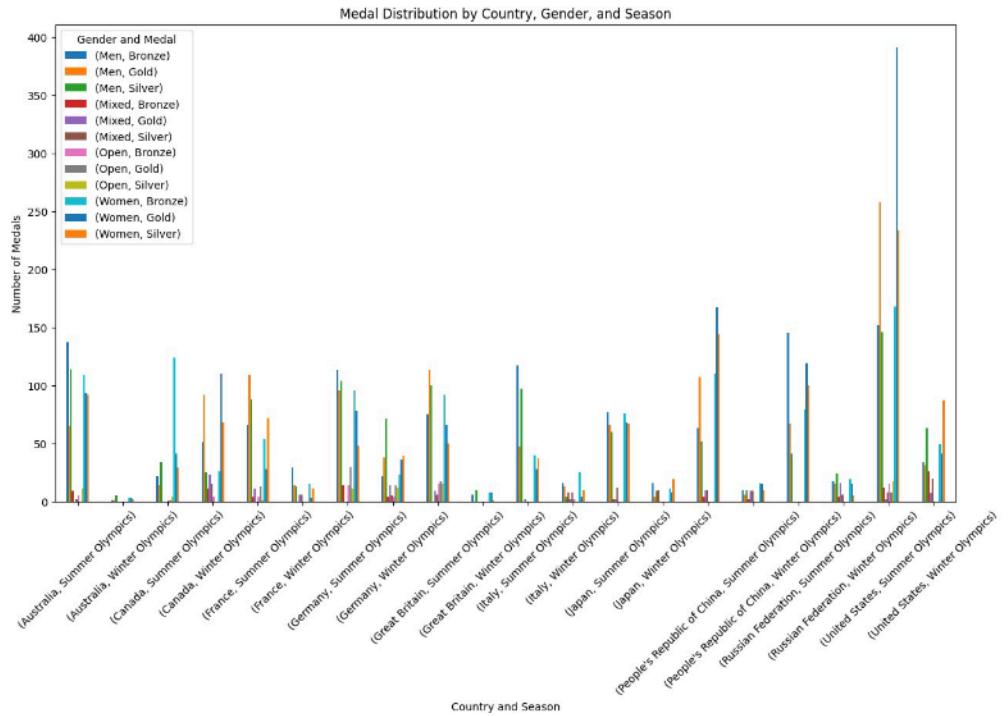
Show hidden code



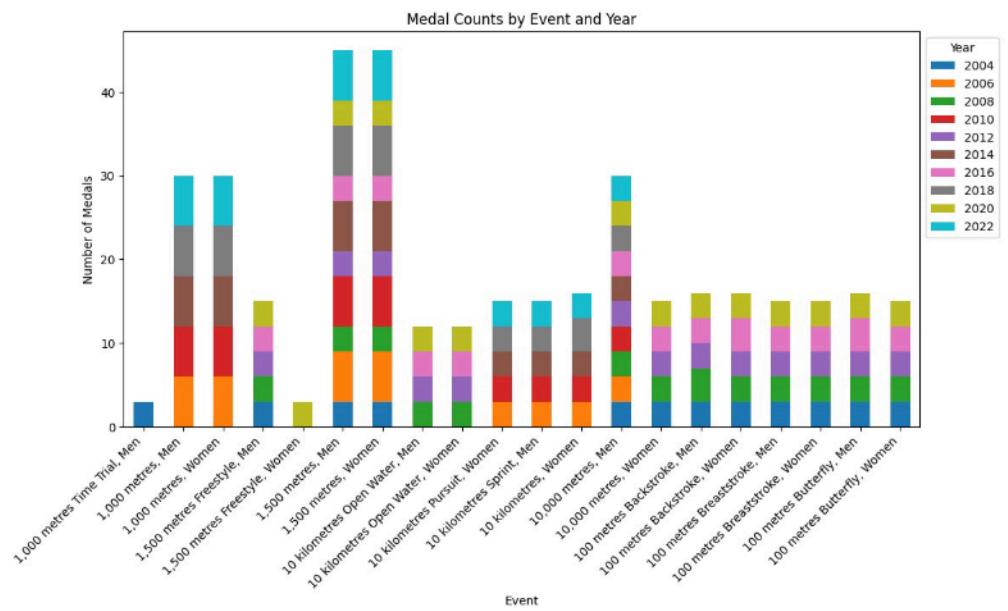
⌄ Show hidden code



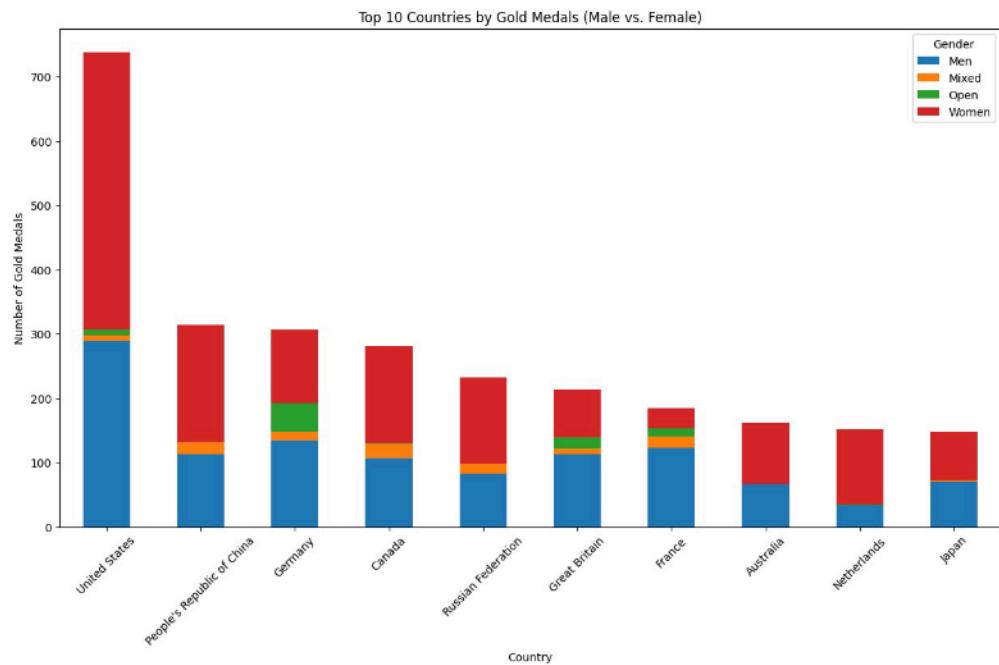
⌄ Show hidden code



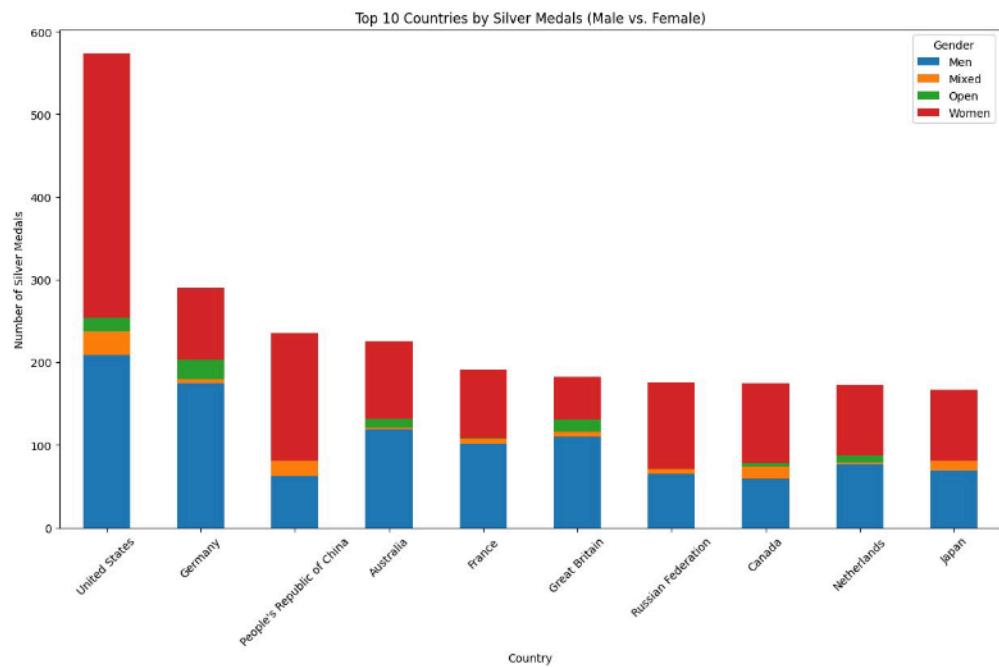
>Show hidden code



>Show hidden code



^ Show hidden code



^ Show hidden code

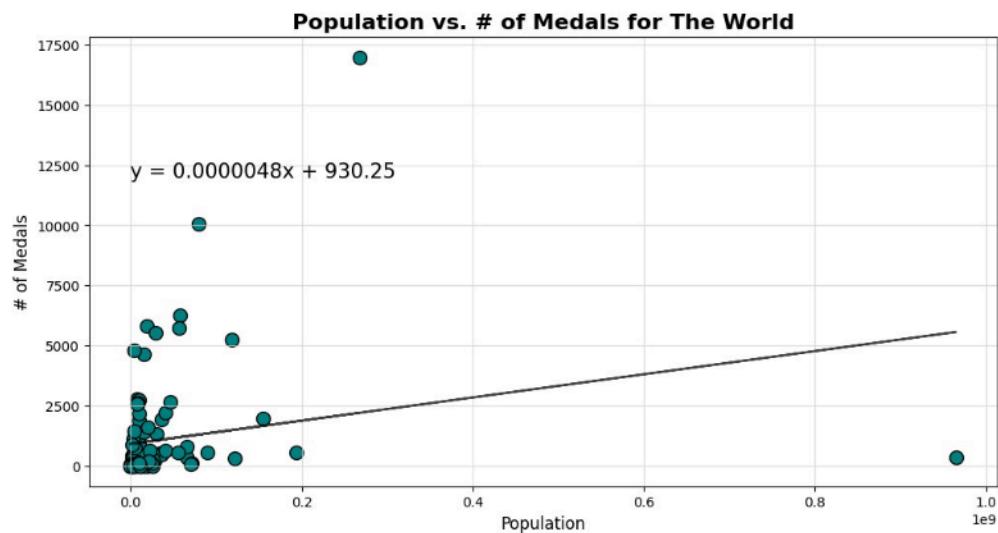


## Regression Plot

>Show hidden code

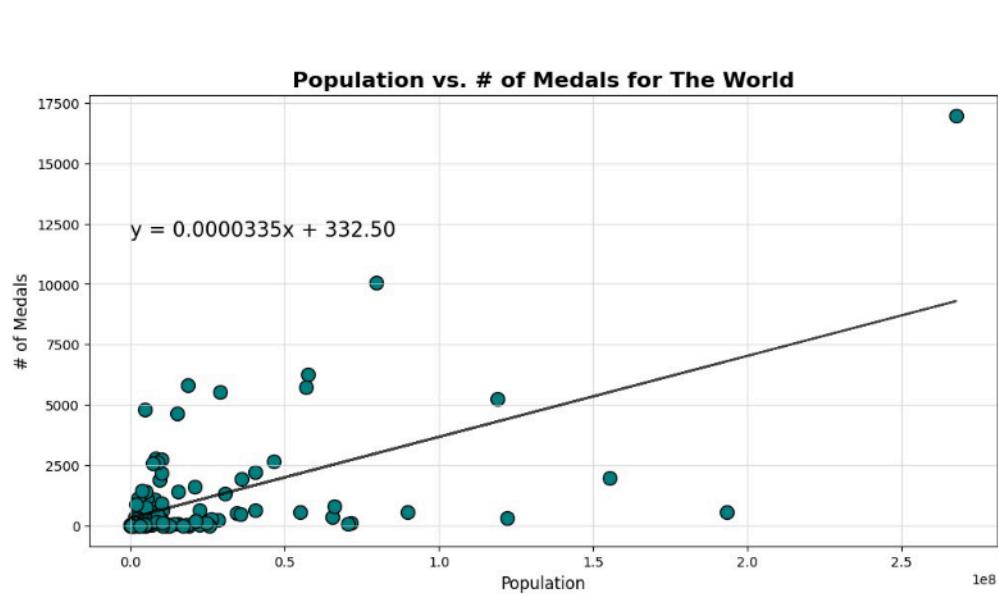
For the most part there is a correlation between population and number of medals won. India as a large outlier might have greatly impacted the slope-intercept formula and deterred positive correlation.

Show hidden code



>Show hidden code

Here is the plot without India for comparison.



>Show hidden code

## Bias and Limitations

---

**Bias:**

- Olympic data might not be an accurate depiction of world athlete distribution.
- World politics impacting participating countries.
- Judged events having country favoritism.

**Limitations:**

- Limited accessibility to regions that may harbor hidden talent.
- Methodology in Olympic Qualifications.
- Although we maintained focus on population, gender and host location, the extraneous factor of country GDP could be a large contributor for determining overall Olympic performance.

## Conclusion

**Regarding country performance in the Olympics, we covered two primary factors, population and participation as the host. Our data reflects that host countries by far outperform their own average records when they are not the host. Population also has a positive relationship with**

**medal count performance, although it is weaker, there is a positive correlation. This suggests that a complex combination of extraneous factors may be contributing to a nation's performance on the world stage like funding for sports, country GDP, and internal political climate among others.**

**The analysis of the Olympic dataset reveals a positive shift towards greater inclusivity and opportunity for female athletes. This expanding participation and recognition of women's achievements have the potential to inspire future generations and further solidify gender equality in sports.**



---

## **License**

This Notebook has been released under the [Apache 2.0 open source license](#).

## **Continue exploring**