Milestone 4: Presenting Your Findings

Shaula Marquez
September 21,2024

Contents

- Review of Questions to Answer Hypotheses/Approach
- Discuss Technical Challenges
- Detail: Entity Relationship Diagram (ERD)
- Initial Findings
- Deeper Analysis
- Hypotheses Results

Questions to Answer

1. Who are the top performing countries over the years?
2. Which country won the most medals?
3. Is there a correlation between the total participants and the number of medals won by the top performer in the Olympics?

Initial Hypotheses

1. The top countries probably who are rich countries
2. United States of America
3. The more athletes the more chances they could get medals

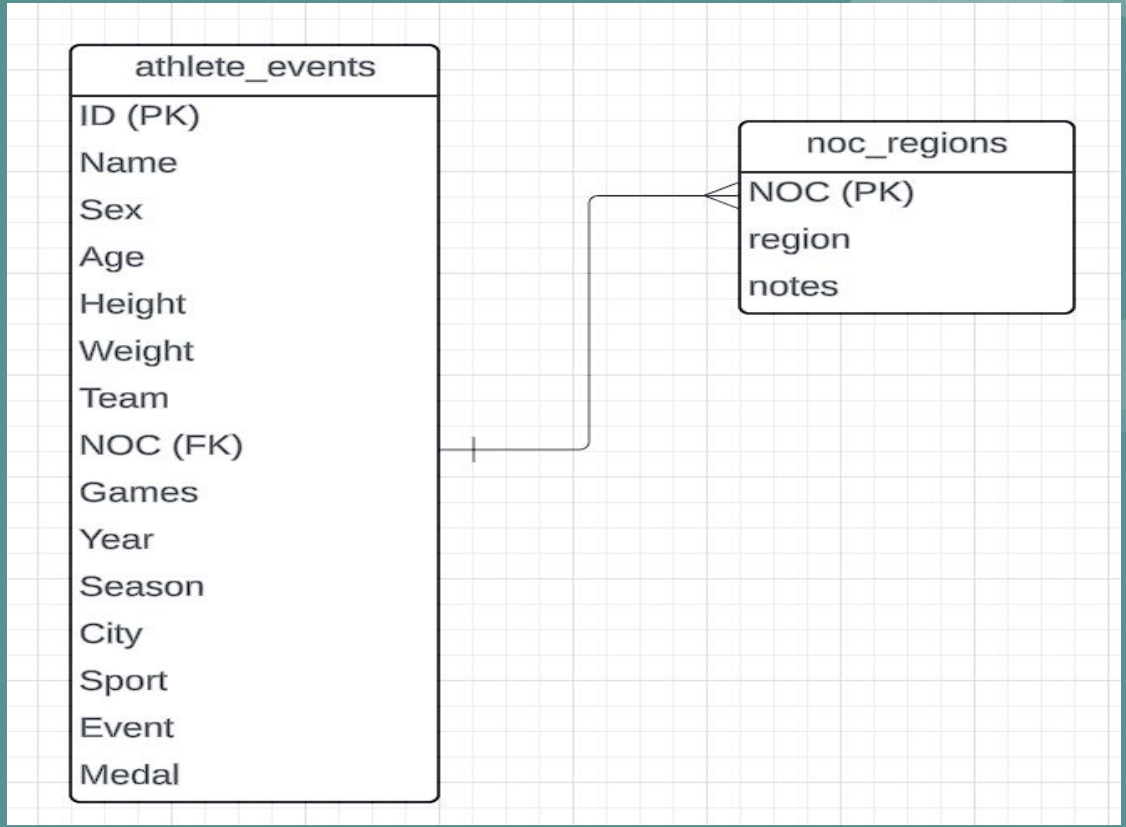
Data Analysis Approach

1. To calculate the total medals of each countries
2. Calculate the correlation between the number of athletes and medal won by the top performer of the Olympics

Technical Challenges

- I’m still learning python that’s why I’m not familiar with the commands. But later on I get used to it.
- Pandasql manage to perform what I wanted on my dataset.

Entity Relationship Diagram



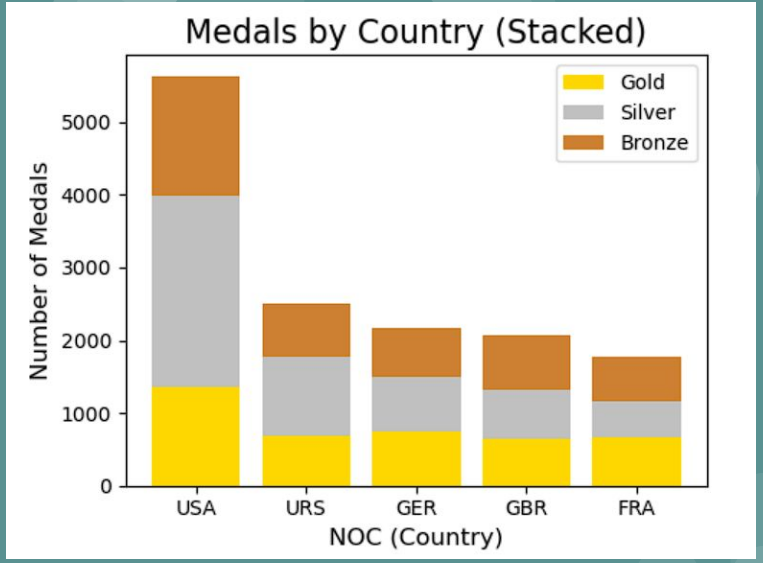
Initial Findings

The initial hypotheses were evident In the data. The top countries who top the Olympics are the rich countries.

	NOC	Gold	Silver	Bronze	Total
0	USA	1358	2638	1641	5637
1	URS	689	1082	732	2503
2	GER	746	745	674	2165
3	GBR	651	677	739	2067
4	FRA	666	499	602	1767
5	ITA	531	575	531	1637
6	SWE	535	479	522	1536
7	CAN	451	463	438	1352
8	AUS	517	348	455	1320
9	RUS	408	390	367	1165

Initial Findings

The top 5 countries are United States of America (5,637 medals), Russia (2,503 medals), Germany (2,165 medals), United Kingdom (2,067 medals) and France (1,767 medals).



Deeper Analysis

The correlation between the total number of medals in the Olympics of United States from 1896 to 2016, is 0.91, which a strong positive correlation. This suggests, that when more athletes participate, the number of medals won tends to increase as well.

	total_count	medal_count	gold_count	silver_count	bronze_count
total_count	1.000000	0.914259	0.816552	0.800579	0.876949
medal_count	0.914259	1.000000	0.920362	0.888292	0.893183
gold_count	0.816552	0.920362	1.000000	0.676242	0.708602
silver_count	0.800579	0.888292	0.676242	1.000000	0.800822
bronze_count	0.876949	0.893183	0.708602	0.800822	1.000000

Deeper Analysis

- Total Count and Medal Count (0.91): This strong positive correlation indicates that as the number of athletes representing the USA increases, the number of medals won also rises significantly. This could be due to a greater pool of opportunities across different sports, increasing the likelihood of success.

Deeper Analysis

- Total Count and Gold (0.82), Silver (0.80), and Bronze (0.88) Medals: Each medal type shows a strong positive correlation with the total athlete count. Notably, the correlation with bronze medals is the highest, suggesting that an increase in participation may have a particularly strong association with winning bronze medals, though all medal types benefit from more athlete representation.

Deeper Analysis

- Medal Count and Gold (0.92), Silver (0.89), and Bronze (0.89) Medals: The correlation between total medals and individual medal types shows that the overall medal count depends heavily on each medal type, with gold having the most significant impact. This makes sense as gold medals are often more celebrated and can serve as an indicator of overall dominance in sports.

Deeper Analysis

- Gold and Silver (0.68), Gold and Bronze (0.71), Silver and Bronze (0.80): The moderate correlations between gold and the other medal types suggest that success in winning gold medals does not strongly predict winning silver or bronze medals. However, the relationship between silver and bronze is stronger, indicating that athletes who perform well but fall short of gold often win multiple lower-tier medals.

Final Findings

- Overall, there is a strong relationship between the total number of athletes and the number of medals won, indicating that increasing athlete participation tends to increase medal counts.
- All medal types (gold, silver, and bronze) contribute significantly to the total medal count, though gold has the strongest influence.
- The moderate correlations between the different medal types suggest that while winning one type of medal increases the likelihood of winning others, it’s not a perfect predictor.
- This matrix reveals that higher athlete participation tends to positively impact the USA’s success in winning medals across all types.

Recommendations

- A multi-pronged approach that increases athlete participation, focuses on underperforming sports, and optimizes both top-tier and broader medal success could enhance the USA’s overall Olympic performance. Data-driven strategies for selection, resource allocation, and preparation will maximize medal potential.