The Descriptive Statistics I looked at with the reason:

1. Which top 10 countries have won the highest number of medals in Olympics’ history?
2. What is the trend of year vs. medals won by top 5 countries found from point 1? Was there a slight change over the years? Was there a sudden rise and fall over the years? If yes, then what might be the reason?
3. The majority of countries’ medals are concentrated in a certain number of sports.
4. Which Age group the highest number of participants come from? This could answer the questions like: Was it due to awareness to this age group about the sport/Olympics or was it about physical capability needed for the particular sport?
5. Was there any sudden or high hike seen in the number of participants took part in Olympics? The number of participants in last 20 years have been increased more than ever in Olympics’ history. What might be the reason?
6. Is there any trend associated with the season and medal won by the country?

2-3 Key points about the data: New Relationships: Additional Ideas for other things to review:

1. There is a relationship between medals won by particular country and sports. It is common for all countries that majority of their medals are concentrated in a certain number of sports.
2. There is a proportional relationship between medals won by countries vs. total number of players participated in Olympics.
3. There could be a relationship between a player and medals won by him/her in different events of the same sport.
4. Additional Idea: One can check the relationship between medal-winning top 4 countries and season (winter/summer). Is there any relationship between medal-winning and season?

Prove/Disprove Hypotheses:

1. First hypothesis is True: I looked into the list of top 5 sports that have the most medals. I computed the percentage of these sports in relation to the total medals won by particular country. Based on the results, I came to the conclusion that all countries’ majority of medals are concentrated in a certain number of sports.
2. Second hypothesis is True: I studied the total number of participants of specific country in the Olympics. I looked into the percentage of the participants who won medals. Based on the results, I came to the conclusion that countries won more medals because they had higher number of participants than other countries.
3. Third hypothesis is False: I studied the participants who won the medals in different events. I grouped it according to sports. I have found different participants won in different events. This is because the requirement to win the sport is different for different events in the same sport.

Additional Questions:

1. Was there a sudden hike seen for any particular group of countries during the specific year? What was the reason behind it?
2. Which sport can be won by 35+ age group? Is there any specific group of sports/events which was won by the upper age group?

**Dive Deeper: Correlation/s I have discovered and What I learnt from those correlations.**

1. During initial data exploration, I thought to look into the relationship of top 10 countries in each Olympics event and then plot a graph of them to number of participants from those countries. But, obviously, top 10 countries are found to be different in each Olympics event. So, I decided to go with top 5 countries which have won the highest number of medals till now in Olympics history. I found ‘Soviet Union’ was also one of the countries out of top 5 countries. As it got resolved in later years, ‘Soviet Union’ was eliminated from analysis.
   1. I found the strong correlation between no. of medals won by the country and number of participants from that country.
   2. The Pearson correlation coefficients between no. of medals won by countries and number of participants from those countries are as following:

|  |  |
| --- | --- |
| **Country** | **Pearson Correlation Coefficient** |
| United States | 0.89791106 |
| Germany | 0.93401656 |
| Great Britain | 0.71517374 |
| France | 0.70679510 |

* 1. I learned that countries have won a greater number of medals because more participants came from those countries.

1. While doing gender-wise analysis, I found the correlation between no. of medals won by female sex and by male sex of particular country in top 10 sports. The list found was almost contained similar sports.
   1. I learned from this analysis that in their respective countries, the attention is being given equally to both sexes’ federation in terms of funding, training programs etc.

**Go Broader: Summary of other things I am considering in my analysis.**

1. Season-wise analysis: What impact does the season have on medals won by country?
   1. I found, in search for the answer to this question, that as there are a smaller number of sports played in Olympics in winter season, the medal-winning count was also found in a less number.
   2. Top 5 countries who won the highest number of medals in winter season are: Canada (575), United States (533), Norway (443), Sweden (428), Finland (426) and in summer season are: United States (4686), Soviet Union (2061), Germany (1687), Great Britain (1598), France (1408). The list of winter is totally differed from that of summer because winter sports might be more famous in cold countries compared to other countries.
2. Sports-wise analysis: Most of the medals are concentrated in specific sports for particular country. That is because basketball is more famous among people of United States, and cycling or fencing is more famous among people of France.
3. Age-wise analysis: No. of medals won in top 10 sports by any country was achieved by the age group of 20-27 years.

**New Metric: New metric/s to track the relationship of data I have discovered.**

1. A metric to track relationship between no. of medals won by country vs. no. of participants participated from that country is Percentage Evaluation ((No. of medal/No. of participants) \* 100). The reason I am using percentage evaluation is because I want to know that countries are winning a greater no. of medals just because they are sending a greater number of participants to Olympics.
2. To track whether a country is gender-neutral in sports area, a medal-winning count by both genders is the necessary metric.