



SQL for Data Science Capstone Project Olympics Dataset

Shay Hegarty

03/20/2022

OUTLINE

- Executive Summary
- Introduction
- Methodology
- Results
 - Visualization – Charts
 - Dashboard
- Discussion
 - Findings & Implications
- Conclusion
- Appendix



EXECUTIVE SUMMARY



- Review the SportsStats dataset on Olympics information for the past 120 years.
- Dive deeper into seeing how countries are consistently performing each year.
 - Compare how these countries perform from Summer games to the Winter games.
 - Are there specific countries dominating both areas.
- Look into data demographic information on the Olympians and see how they have changed over the years and more specifically how the males dominated the Olympics in the early years but its starting to level out now.

Questions

- Is there more men or women competing in the Olympics? How has this changed over time?
- Which countries are always successful?
- How has the age gap changed over the years?



Hypothesis



- Is there more men or women competing in the Olympics? How has this changed over time? I would think there has been more men every single year since the start of the Olympics.
- Which countries are always successful? I am presuming it will be the largest countries such as USA, China etc.
- How has the age gap changed over the years? I would imagine the age gap being quite significant but the overall average being in the 20-30s.

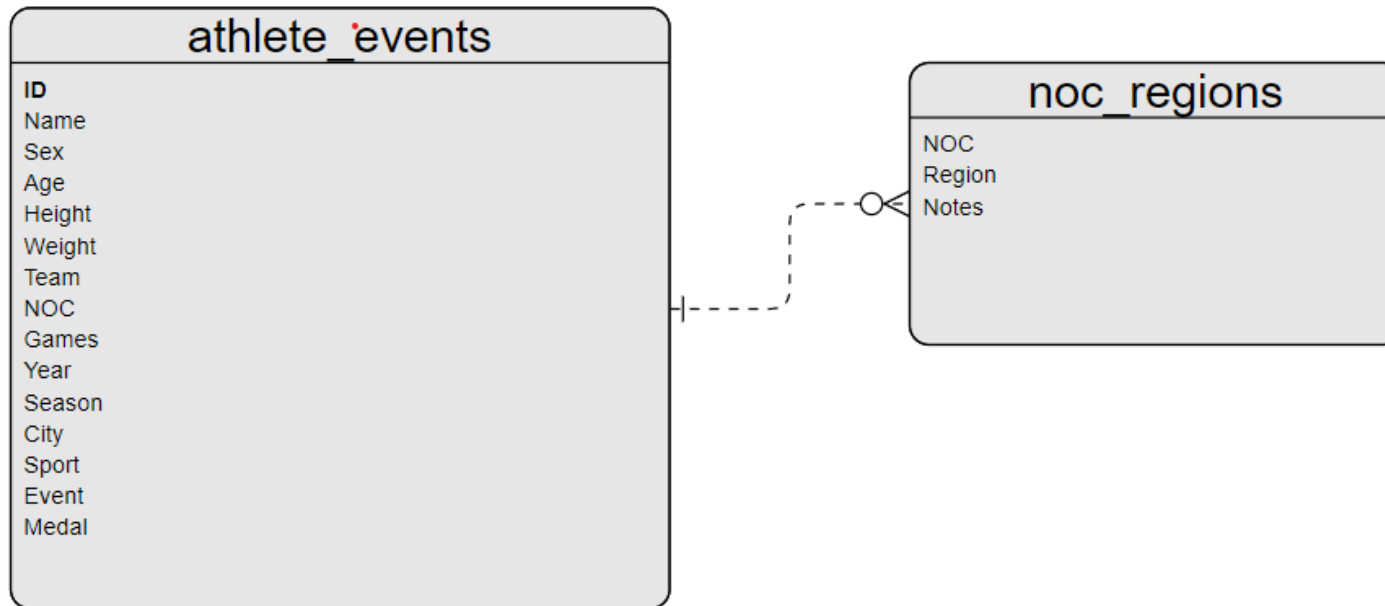
Approach/Methodology



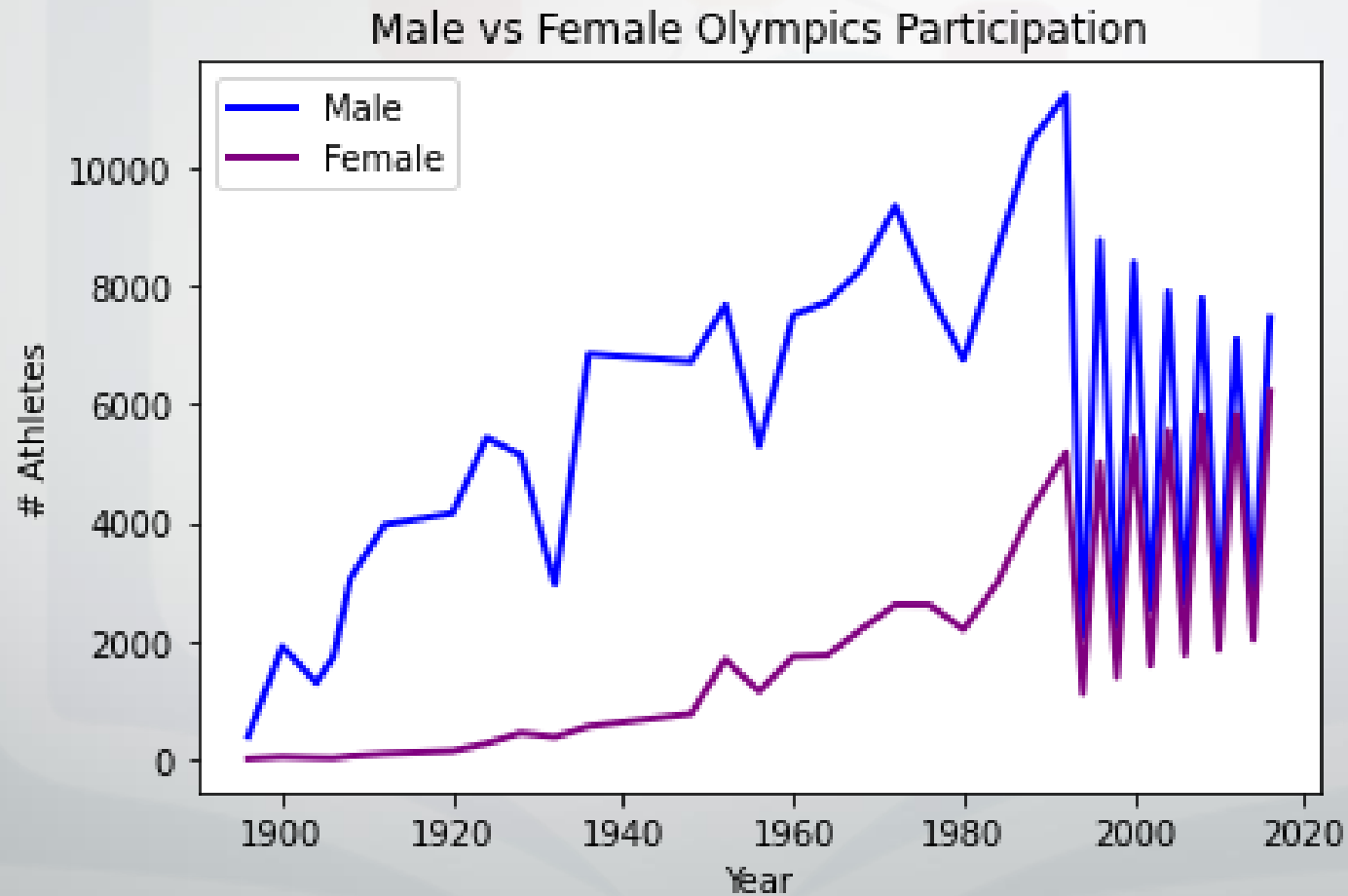
- Analysis was completed using the SportsStats 'Olympics Dataset – 120 years of data'
- Data was then wrangled/cleaned by investigating missing values, removing duplicates and normalizing the data
- Data was then explored, and any outliers were reviewed
- Visualizations were then created to determine relationships within the data
- Finally, I dove deeper into the data to find the correlation and relationships between the data

ERD

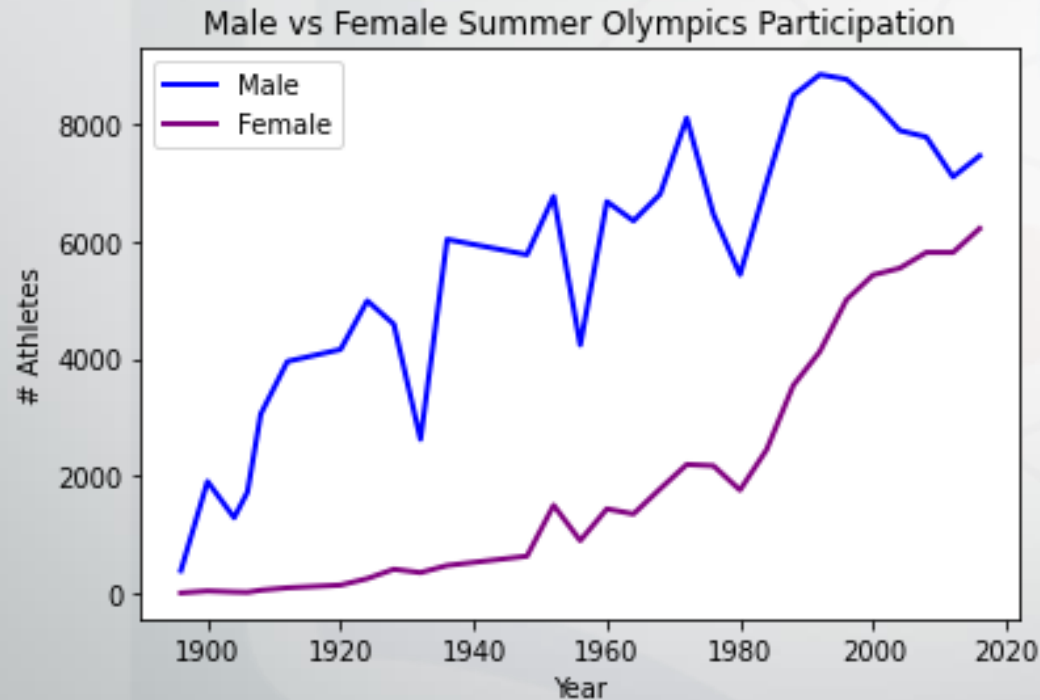
Sports Stats - Entity Relationship Diagram



Male vs Female Participation

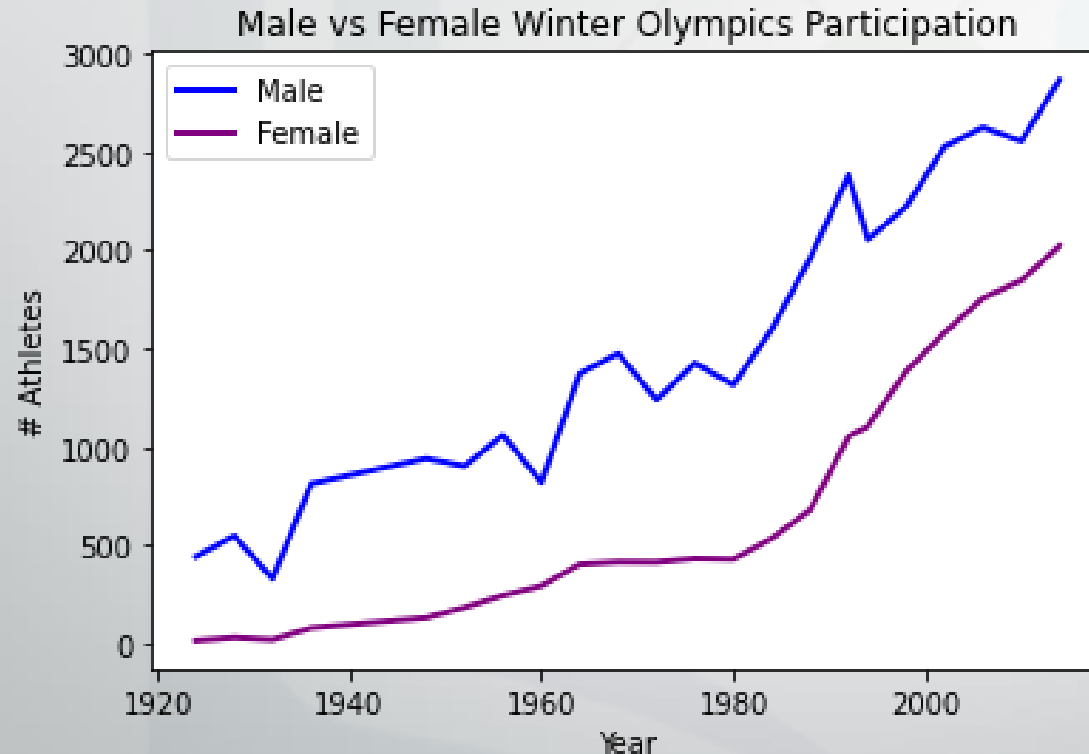


Male vs Female Participation Summer Games



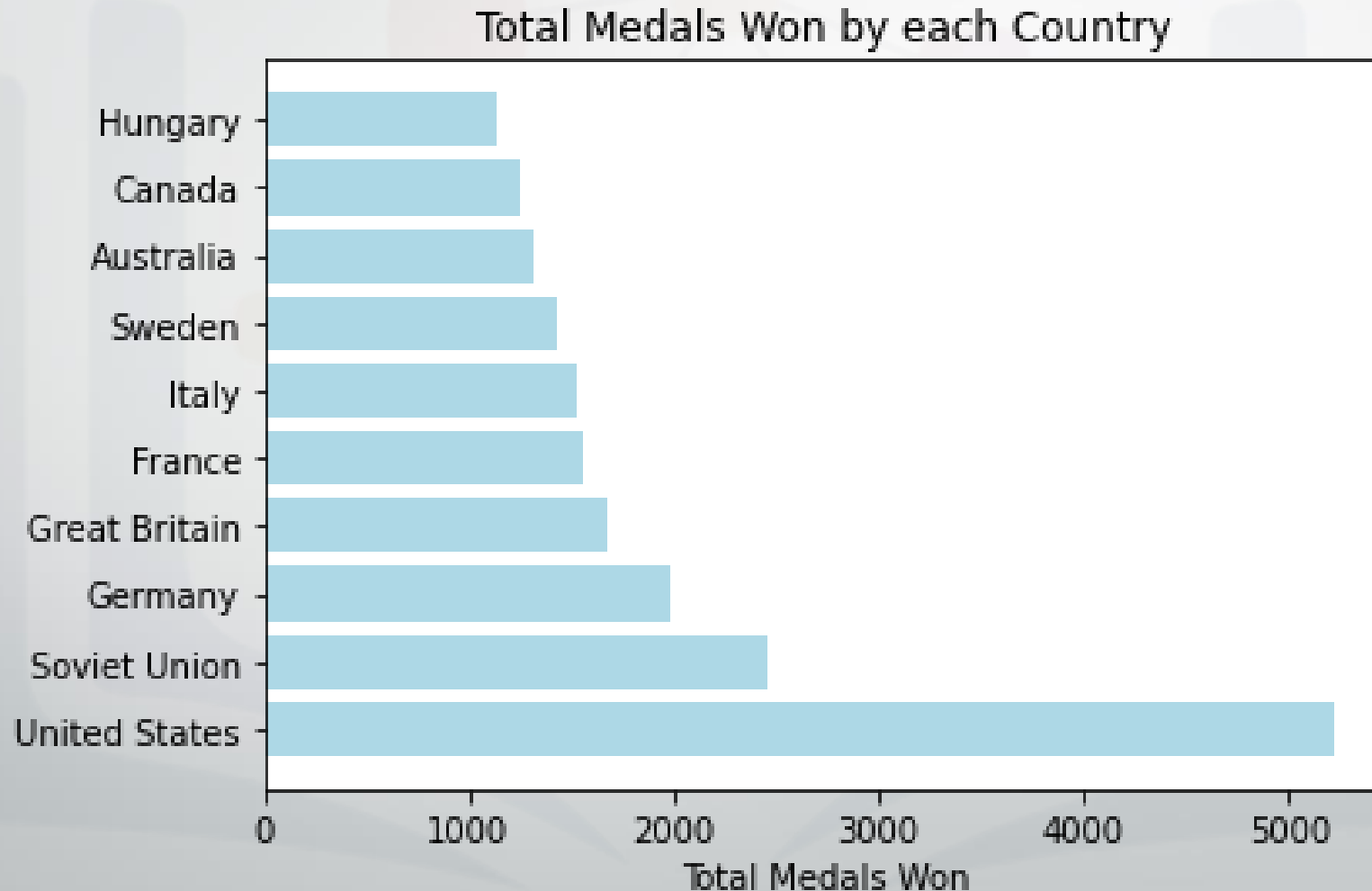
- Large discrepancies between male and female participation in the early years
- 1980s was a huge turning point
- Participation rate for females is steadily going up since the 1980s and will hopefully level out completely in the years to come.

Male vs Female Participation Winter Games

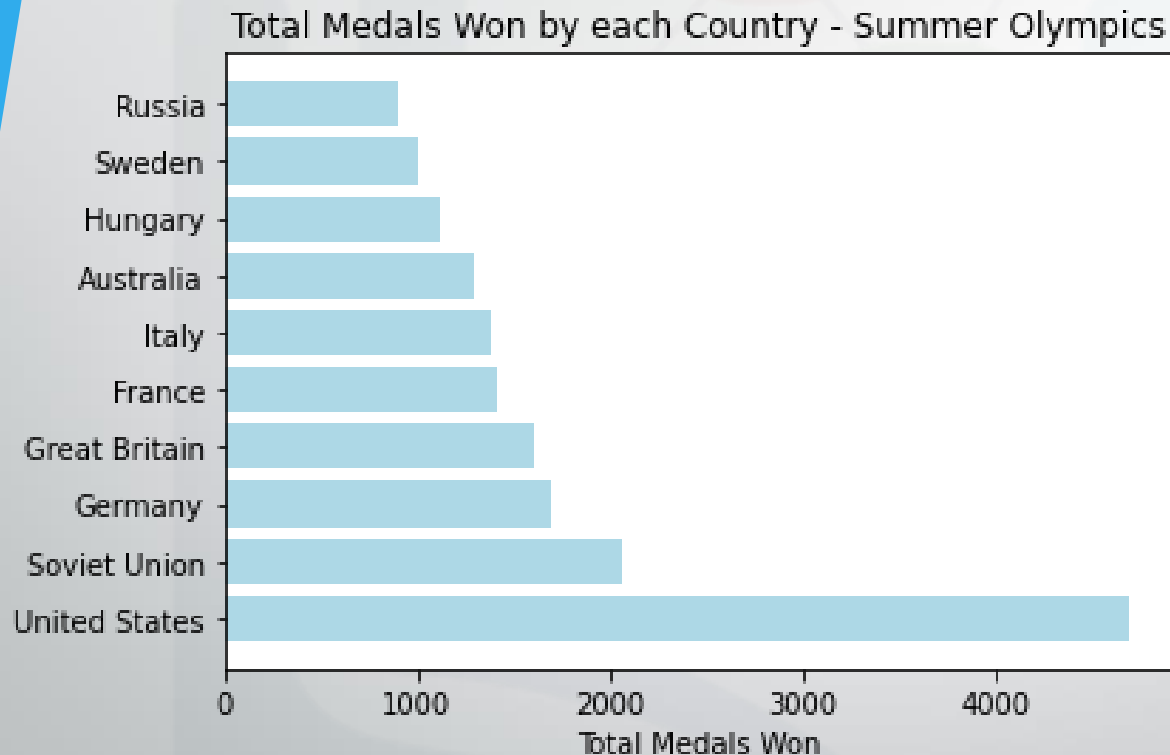


- Similar to the summer games there is a large discrepancies between male and female participation in the early years
- 1980s -1990s was a huge turning point for the winter games also.
- Participation rate for females is steadily going up since the 1980s and will hopefully level out completely in the years to come.

Top 10 Countries by Medals Won

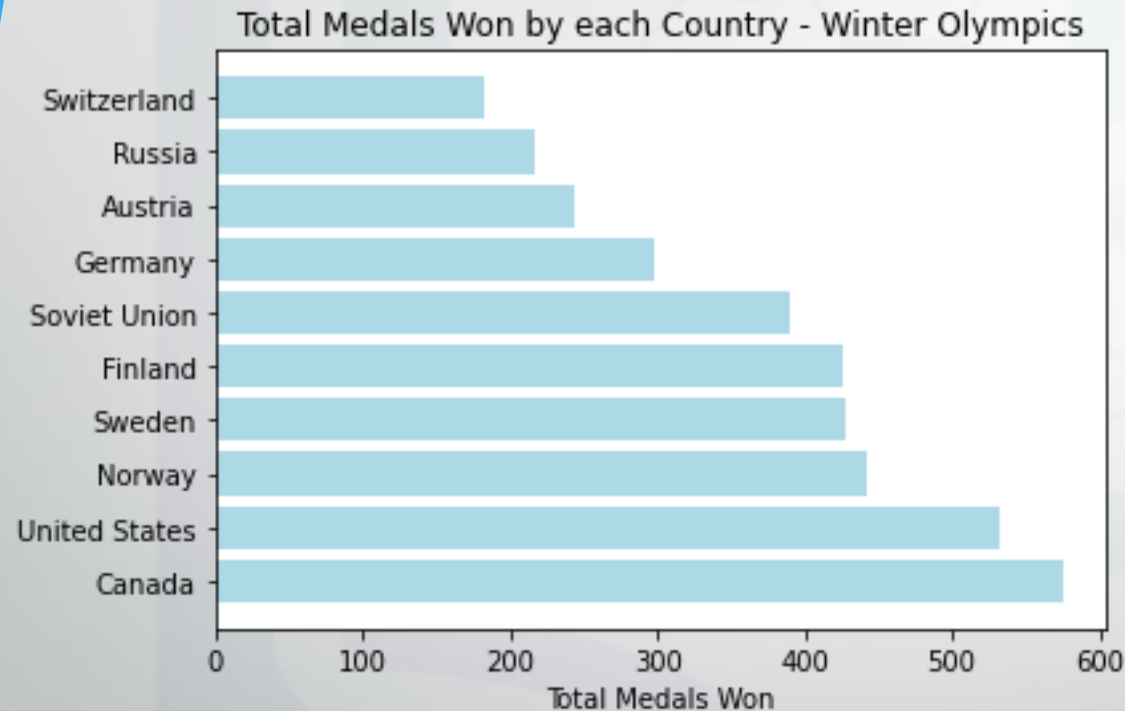


Countries by Medals Won Summer Games



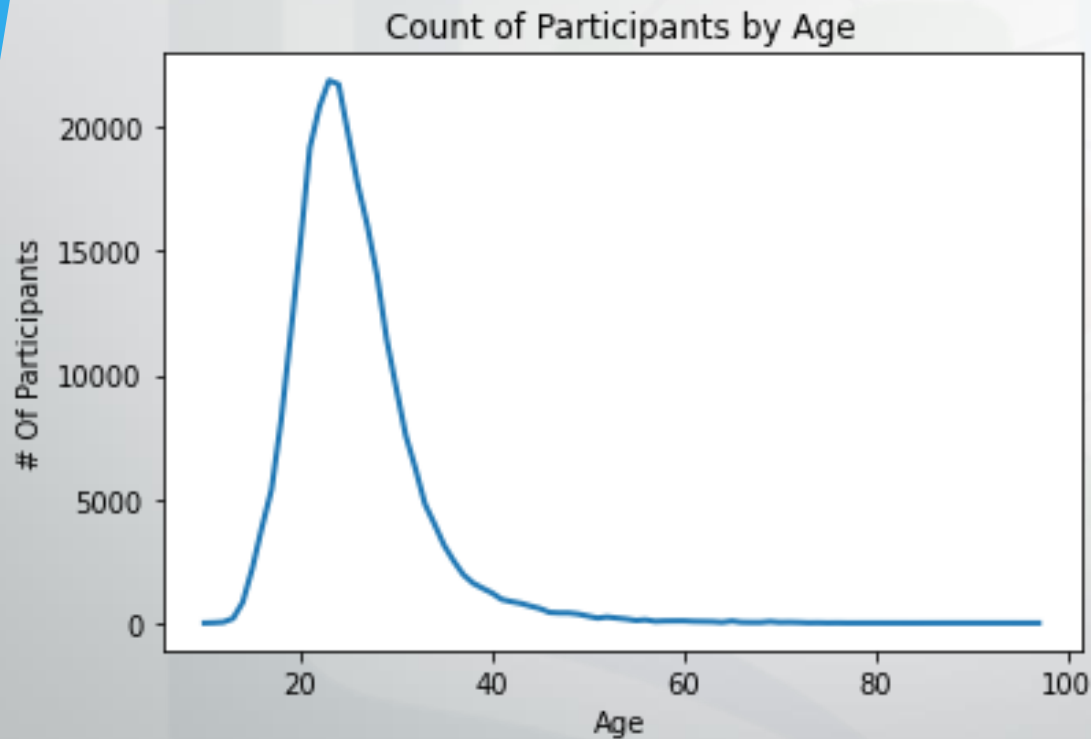
- As expected in our hypothesis USA is dominating the field when it comes to medals being awarded for the Summer Games.

Countries by Medals Won Winter Games



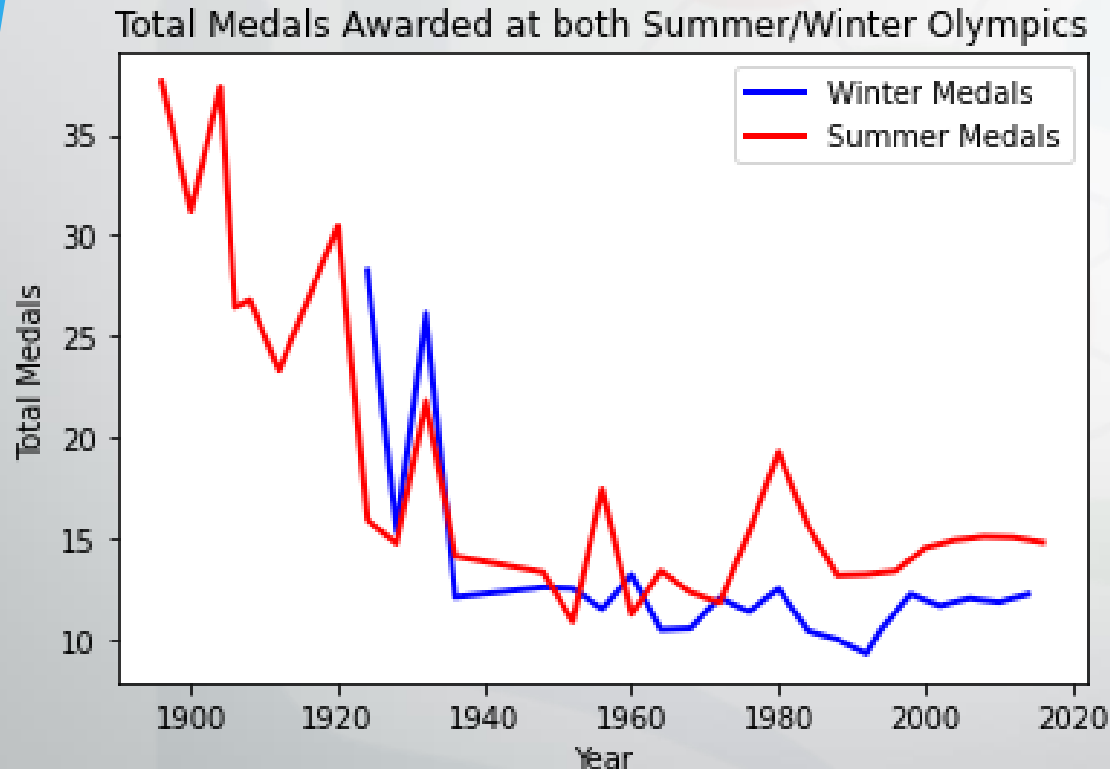
- When it comes to the Winter games USA is still a massive competitor but its not comparable to their lead in the summer Olympics.
- Seeing the Scandinavian countries such as Norway, Sweden and Finland which is expected for the Winter games.

Age of Participants



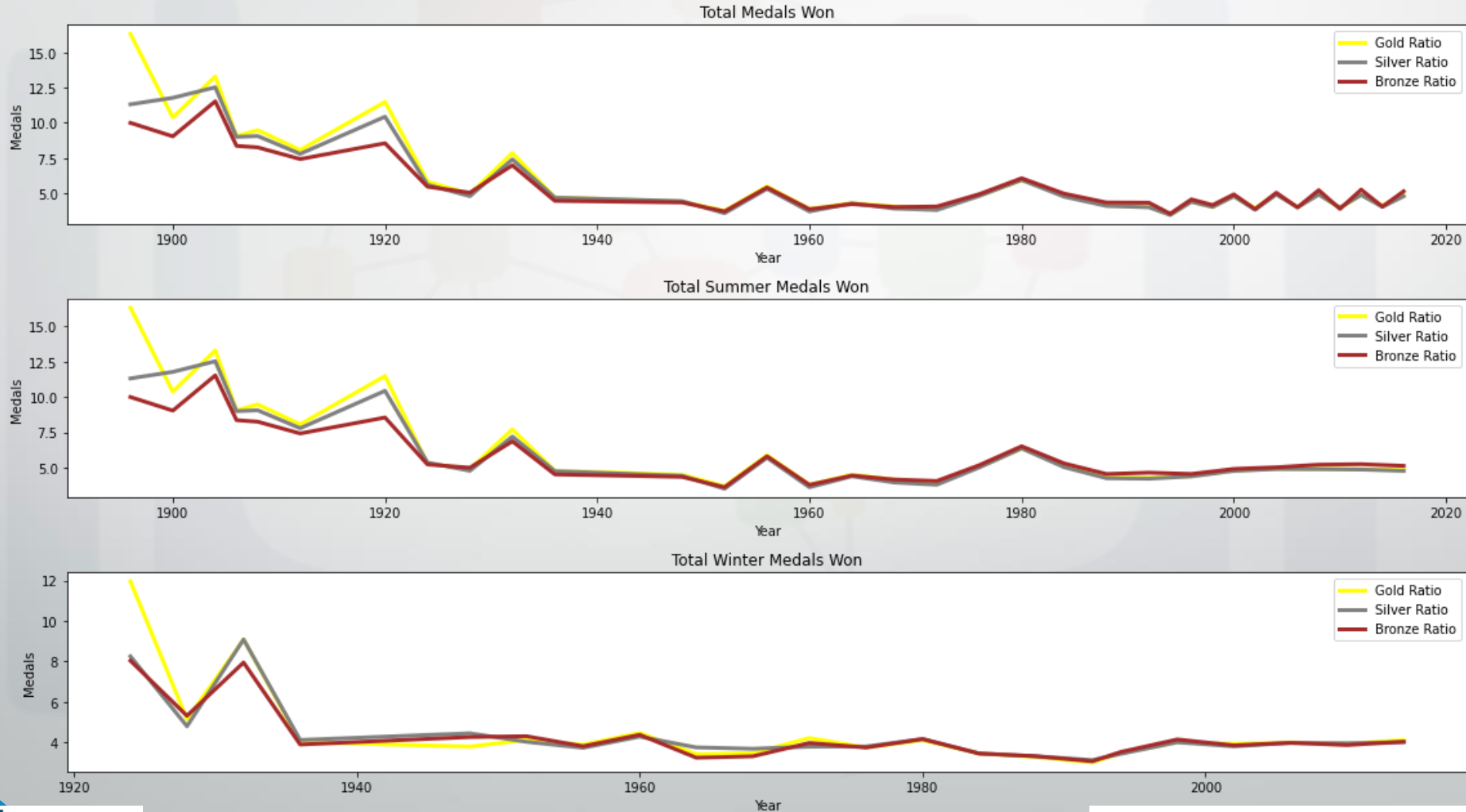
- Youngest Olympian is listed at 10years old.
- Average ages are between 20 – 30 which was predicted.

Total Medals Awarded Summer vs Winter games



- Early years there were different amounts of medals distributed for each class
- Eventually this levelled out and it looks like the same amount of medals get distributed

Total Medals Awarded Summer vs Winter games



FINDINGS

STD and CORRELATION

Standard Deviation:

Standard Deviation for Summer Olympics Medals = 475.323015441357

Standard Deviation for Winter Olympics Medals = 152.56899942903493

From 1924 to 2016, as the standard deviation in the summer Olympics is about 3 times that in the winter Olympics, country performance by year change more in Summer Olympics.

The Pearson correlation coefficient between the total number of medals in the winter and summer Olympics from 1924 to 2016, is 0.94, which is highly positive. Therefore, the performance of a country in winter Olympics is highly correlated to that in summer Olympics.

DISCUSSION

- What further insights can we determine from this dataset?
- Investigate into the spike in the 1980s for the amount of medals that were awarded.



CONCLUSION

- Is there more men or women competing in the Olympics? How has this changed over time? *I would think there has been more men every single year since the start of the Olympics.*
- Which countries are always successful? *I am presuming it will be the largest countries such as USA, China etc.*
- How has the age gap changed over the years? *I would imagine the age gap being quite significant but the overall average being in the 20-30s.*