Olympic Swimming and Running Trends

Comparing time results across sports from 1912 to 2016

FURMAN COMPUTER SCIENCE

Grace Houser

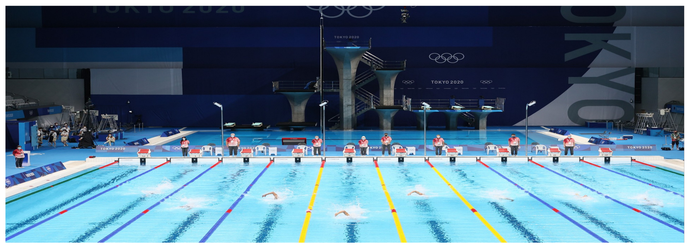
INTRODUCTION

Swimming and running have been an Olympic sport since the first modern Olympic Games of 1896. Since then, both sports have evolved over time with changes in rules, techniques, nutritional information, and overall improvement of athletes over time.

This study takes a look at the trends of swimming results over the years of the Olympic Games since 1912, which is the earliest Olympic Games that has the recorded information. Additionally, this study compares the Olympic swimming and running results over time of the same distances.

DATA

For this project, two datasets were outsourced: one for swimming and one for running.

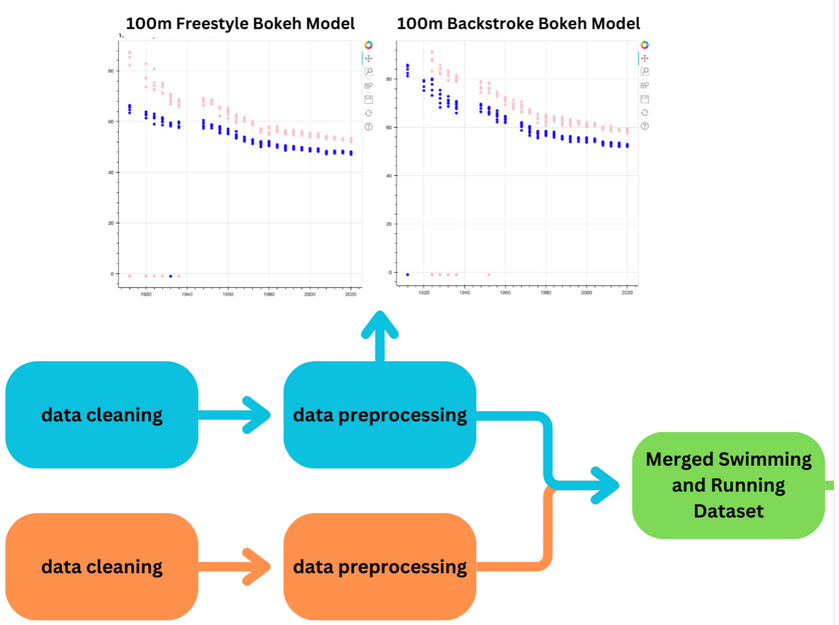


Both datasets include all known Olympic result times from every operating Olympic Games. Thus, it is important to note that because the Olympic Games were canceled due to World War I, there is no data from the year 1916. Additionally, there is no data from the years 1940 and 1944 because the Olympic Games were canceled due to World War II. However, because we have 25 years of Olympic data, these three years of no information should not hinder any trend lines that we may see in the data.

Swimming	Running
Location	Location
Year	Year
Athlete	Name
Gender	Gender
Team	Nationality
Rank	Medal
Distance	Event
Stroke	Result
Relay?	
Results	

Attributes

PIPELINE



After finding the “Olympic Swimming History 1912 to 2020” dataset, initial data cleaning needed to be completed. First, eight instances had to be dropped because the given swimmers were disqualified or had a blank time entered under the “Results” attribute. Furthermore, the “Distance (in meters)” and “Results” attributes were implemented as strings, so they were changed into numerical values.

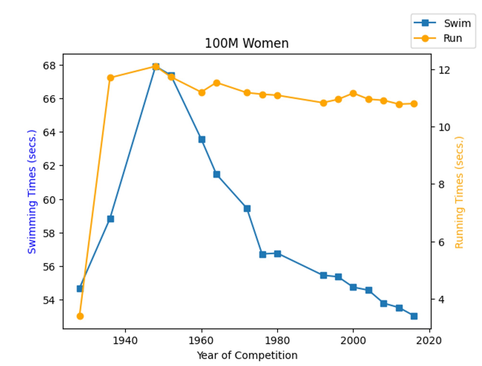
Once data cleaning was finished with the swimming dataset, preliminary findings were computed with Bokeh, which is a Python library. Preliminary findings of the swimming dataset revealed that the rate in which women improved in a swimming event over Olympic years, men improved at a similar rate. This is most noticeable and comparable between 100m events, since all stroke have a 100m distance event.

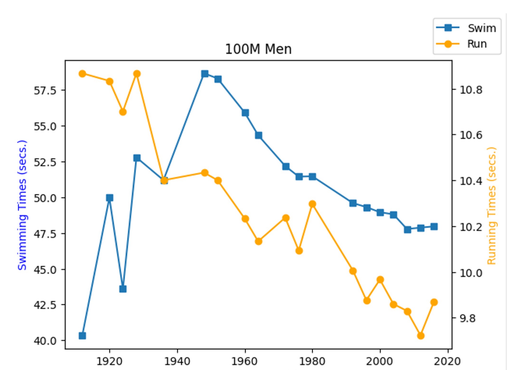
After this, the swimming and running dataset was merged into one, so that similar distance events of swimming and running could be graphed and compared.

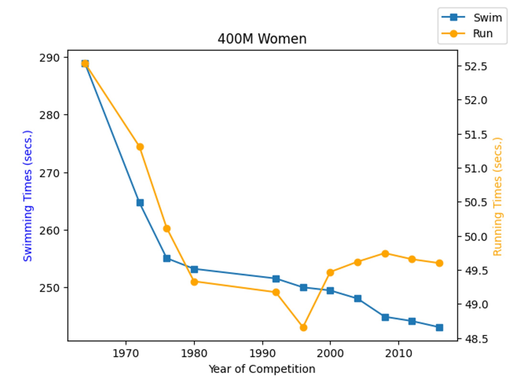
RESULTS

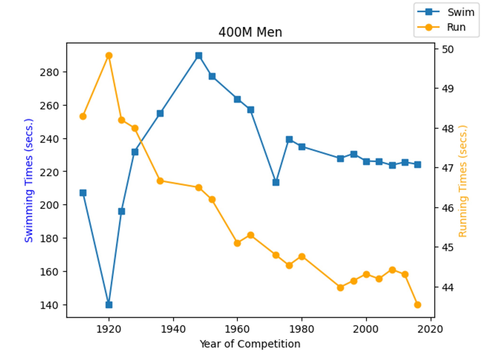
It is interesting to note that when comparing men’s running to swimming, the men’s running trend line tends to be below the respective swimming trend line. Whereas for women, their running trend lines tend to be similar to or above their respective swimming trend line. This could be because womens’ bodies are better built for swimming than men’s; women have biological advantages like lower hemoglobin levels and oxygen supply, which is why women are actually better than men at marathon swimming. Even though this is a small example of such and the distances are much lower than a marathon, this could still perhaps explain this pattern.

GRAPHS









ACKNOWLEDGMENTS

This project was able to be completed with the help of the Furman University Computer Science Department. Special thanks to Furman professor Dr. Sultan for his mentorship and guidance in providing me with the resources needed to conduct this research.