**Project proposal – visualization of swimming ranking**

Neta Shiff

Ns1695@mynsu.nova.edu

Nova Southeastern University

CSIS 060: Data visualization

Summer 1 2024

Professor Lazlo

May 29, 2024

This project aims to analyze the performance of the top 100 swimmers in various competitive swimming events. The dataset is taken from USA swimming, which includes detailed records of swimmers’ performance, in the various yards size events. This project will investigate trends, patterns, and factors that might influence swimming performance. During the analysis and visualization the main focus will be to explore how the various aspects such as age, foreign, club team , distance and meet characteristics to derive meaningful insights and recommendations.

The dataset is taken from the USA swimming which is the national governing body for competitive swimming in the United States. This ensures that the dataset is reliable and comprehensive as the official body overseeing swimming events nationally. The data used is accurate and consistent since USA swimming provides data collection using standardized methods and criteria making it the main trusted source used by coaches, swimmers, fans and researchers.

The dataset includes extensive detailed records of swimmer performances, including wide range of performance from various years and competitions, allowing the data have a longitudinal analysis and trend identification over multiple years. The encompassing columns such as FullName, Distance, Event, SwimmerAge, EventCompetitionCategoryKey, SwimTime, SwimTimeAdj, Standard, MeetName, SwimDate, ClubName, LscCode, IsForeign, Rank, PowerPoints and Sanctioned. These records are crucial for examining trends, patterns, and factors influencing swimming performance as the power points and ranks offers a deeper nuanced view of the performance than swim times.

The prime objectives of the project are layered, the project will start and investigate the relationship between the performance and the swimmers ages including both the ranks, power points and the time. another comparison will be between the achievements of foreign and local swimmers . Secondly, there will be a visualization showing which months and years more swimmers entered into the ranking , my assumption is that on march since it is when the NCAA competition is held. Thirdly, the project will analyze how swim times vary with distance in different events, examining whether longer distances consistently result in higher swim times.

Moreover, another objective that will hopefully achieve is a comparison of the top performance across the different event such as the 50 yard compare to the mile. Additionally,, it will assess the impact of high-profile meets versus local meets on swimmer performance times and examine whether there are seasonal trends in swimmer performance. Lastly, the project aims to identify clubs that consistently produce top-ranked swimmers and analyze their performance trends over time. The insights gained from this study will help in understanding key factors influencing performance and could inform training and competition strategies for swimmers and coaches.

**References:**

Plotly. (n.d.). *Plotly.express.scatter\_geo*. plotly.express.scatter\_geo - 5.22.0 documentation. https://plotly.com/python-api-reference/generated/plotly.express.scatter\_geo

USA swimming. (n.d.). *TIMES OVERVIEW*. USA swimming. https://www.usaswimming.org/times/