Jack R. O’Connor

09/14/2019

Project Deliverable Milestone 1

Database for Olympic Swimming Results

Project Summary:

I will be creating a database containing the Olympic championship results for the swimming events. This will include the Olympic Games and year, events, athletes, Olympic records (OR), place won, time, and country of the athlete. Users will be able to see the downward trends in the Olympic records as technology and training methods improve as well as individual athletes’ time and ranking progressions. This database will be useful to equipment manufacturers and coaches as it will allow them to see what technologies resulted in faster times and which training methods were most effective.

Stakeholders:

Equipment manufacturers (Speedo, Tyr, and Arena) will be able to compare this data to their records to see where their equipment made the most difference in times.

Coaches will be able to use the data to compare training regiments to one another based on the Olympic results. They can figure out what worked with their training and weed out what did not.

Governing bodies, such as FINA, can use this database to see the effects of new rules and standards on Olympic results.

The Athletes, themselves, will be stakeholders in this system insomuch that they will be able to benefit from knowing past results and Olympic records when setting goals for themselves.

Sports casters during current Olympics will be able to utilize this database to pull past results ahead of time to use during the casting of current races for comparison purposes.

Glossary:

**Olympic Games** is the specific swim meet in which the swimmers competed.

GamesID is the surrogate key for the Olympic Games Table

Year is the year in which the Olympic Games took place and references the Year table.

City is the location of the Olympic Games.

CountryID is the country location of the Olympic Games and the Foreign Key to reference the Country table.

**Event** is the specific race in which an athlete competed.

EventID is the surrogate key for the events table.

Distance is how far the athletes must swim to complete the race.

Stroke is the event’s stroke (Freestyle, Backstroke, Butterfly, Breaststroke, or IM)

Gender is the gender of the athletes in the event and references the Gender table.

**Games Event** is the reference table for the events in a particular Olympic Games since every Games will have the same events.

Games EventID is the surrogate key for the Games Event table.

GamesID is the first foreign key referencing the Olympic Games table.

EventID is the second foreign key referencing the events table.

**Olympic Record** is the current fastest time swum at an Olympic Games to that point in time. This will be a table referencing historical data and may not be directly related to every athlete event instance since many years may pass before the record is broken.

Olympic RecordID is the surrogate key that identifies the record.

Time is the time in which it took the swimmer to complete the race, assuming that the time was faster than the previous record.

Date set is the date on which the previous record was broken.

Athlete eventID is the foreign key that references the athlete event table if a swimmer breaks an Olympic record.

**Athlete** is the swimmer.

AthleteID is the surrogate key for the athletes.

First name is the first name of the swimmer.

Last name is the last name of the swimmer.

CountryID is a foreign key referencing the country table and identifies for which country the swimmer is competing.

Date of birth is the athlete’s date of birth and will be used to calculate the age of the swimmer at the time of a specific event.

Gender is the gender of the athlete and references the Gender table.

**Athlete event** is a bridge table between the events table and the athlete table which will identify which athlete competed in which event since swimmer may compete in multiple events over multiple Games.

Athlete eventID is the surrogate key for the athlete event table.

AthleteID is the first foreign key that references the athlete table and identifies that athlete in the athlete event table.

Games eventID is the second foreign key that references the games event tables since many athletes swim the same events at multiple Olympic Games.

Event time is the time in which it took the athlete to complete the race.

Ranking is the place in which the swimmer finished the race compared to the other athletes in the same race.

Age is the age of the swimmer at the time of the event which will be calculated off of the swimmer’s date of birth and the year of the Olympic Games.

**Country** is the list of countries that participated in the Olympic Games.

Country name is the name of the country.

**Gender** is the gender (Male of Female).

**Year** is the year in which an event occurred (either an Olympic Games, Event, or Olympic Record).

Data Questions:

1. In which year did the most Olympic Records get broken?

A screenshot of a social media post

Description automatically generated

A screenshot of a cell phone

Description automatically generated

1. Which country has won more medal, the US Men or the US Women?

A screenshot of a social media post

Description automatically generated

A screenshot of a cell phone

Description automatically generated

1. Which country has won more medals, the US or Russia?

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

1. What is the age of the oldest person to win a medal?

A screenshot of a computer

Description automatically generated

A screenshot of a cell phone

Description automatically generated

1. Which counties have fewer total medals than Michael Phelps does alone?

A screenshot of a cell phone

Description automatically generated

A picture containing screenshot

Description automatically generated

A screenshot of a cell phone

Description automatically generated

Conceptual Model:

A close up of text on a white background

Description automatically generated

Logical Model:

A picture containing screenshot

Description automatically generated

Reflection: Going into this project, I was eager to try and get the most out of it. I wanted to toe the line between do-ability and complexity. As I look back to those early days when all I knew how to do was create a table, I think it is kind of funny that I wanted to really push myself with this project. I assumed that I would be able to handle some of the aspects of the project better than I was actually able to; however, I can now code SQL statements more efficiently than I could even two weeks ago. I had to dig deeper into some of the topics we covered to understand them better so I could apply them to my project. I am thankful for this even though it has been a crazy end of term. I would have like to have spent a bit more time with either RStudio or Access and understanding how to use those tools more effectively in answering my data questions. For me, this was the most difficult part. I would have also liked to have spent a bit more time inserting more datapoints into my tables to be able to really see the answers to my data questions. All-in-all, this was an excellent project and I really did enjoy working on it. I have never done anything this cool or complex, so it gave me a great sense of accomplishment (which will either be reinforced or destroyed by the grade).

Summary: I had to change some of my tables around and restructure. It was nothing major, but it made answering my data questions much easier. Other than those few minor changes, this document is pretty much the same thing as my part 1 deliverable.