Track Results

View collegiate track meet results and create hypothetical meets

Adamson Bryant, Stephen Garrett, Twazanga Mugala, Nick Solfanelli, Keyston Fuller

Problem Description and Motivation + Related Work

The goal of this project is to create a database of collegiate track and field athletes' and teams' statistics. It offers the user the ability to query results of various events from different meets. This data will be used to compose hypothetical events; it will be used to project winning odds for teams or athletes at a given meet or in a given event. An example of this is how a true freshman on a team impacts the amount of points he or she will receive or how that athlete compares to others competing in the same event. This example serves as our motivation for our project as well. Compiling mock results for athletes and teams is something that has happened before, but it has not happened this in-depth. There are databases that are filled with information for teams but it simply does that, compiles and holds the data. Our project is a database of athletes and teams and their results from events and meets. Handling the data as such allows for comparison and this is what makes our project useful and unique. This tool could be useful for companies like ESPN and other networks that broadcast track and field events because they could use these projections to see who is most likely to win races and events, as mentioned earlier. It is also useful for fans who are interested in seeing who has the best chances of winning. They simply go to our website, type in what events, which athlete, or team they want, along with other characteristics of a meet and they have that at their disposal.

Our Solution

Our solution to creating a stylistic track database with race prediction algorithms is multifaceted, involving integrated front and back end design in order to present our data with a user-friendly experience. Our web app opens to the homepage, featuring a navbar with selectable options to accesses lists of Power 5 conference teams, and athletes. We pulled the raw data from a track database, and created tables to organize the data. From there, our data data is organized by

individual teams and athlete tables, teams displaying a school name and a list of teammates for that team accompanied with their personal best for each of their events. We chose to have one table, AthleteCompeteIn, to have the full results for everything and then build other tables off of that instead of separating the info and joining when necessary. We found the number of joins if we had made smaller tables made the database run too slowly, and adding extra records to the one table was the best solution. Our “Hypothetical Meet Creator” is a concept that pits 5 teams against each other across a specific event group in order to determine which would place first in a race. This is achieved by calling each runner for each team’s personal best time for a particular event and placing them accordingly within the race along with an accompanied score. Once scores are tallied, the places for each team in the hypothetical race as well as individual scores are calculated and the user is given their result.

New Approaches

A feature that wasn’t innovative but wasn’t necessary that we added was having every record displayed on the web page as a link. For example, when you view the results of a search and it returns a list of athletes, each athlete in the list can be clicked on as a link to their personal page displaying their best marks for each event. This also holds for teams as if they show up in the results you can click on them to view their roster of athletes that can again be clicked and links back to their page. The main new approach we took was developing an algorithm for creating the hypothetical meets. This is the main innovative feature of our project that isn’t available on other platforms for viewing similar data. The idea of the hypothetical meets is to let the user select any 5 teams for comparison, and the application will create a meet in which it is these 5 teams competing with each other and every person runs equal to their best mark. The individual results for the meet are shown, and the team scores are listed at the top. This algorithm was done in a few subqueries.

The first was to go to our Athlete table of the athletes with their best marks, and limit it down to the 5 teams for competition, the right gender race, and the event the hypothetical was supposed to take place in. The next one, the complex one, grouped the athletes by team, and in each team ranked the athletes by their marks in the event for the hypothetical. This was very difficult to figure out how to do, but necessary because scoring of these meets requires only the best 7 athletes from each team, so we had to limit the group sizes. After doing this ranking, we could take only the athletes with ranks less than 8, so the top 7 from each team. This result was then ordered again by time in general to get individual results, and to give points to each person (every place is a point, so first place gets 1 point, 15th place 15 pts, etc) we used the row_number() function. These results were displayed as the individual results, but the team results needed one additional step. For the team results, we grouped again by team, and took the sum of points from each team to get the team score and returned a sorted list of the teams in order of score (lowest score wins). This feature was very new for a database of this kind, some other ones had “performance lists” which showed individual results but didn’t create the teams of 7 and give team scores. The biggest limitation of this feature was that it required exactly 5 teams to be compared. Ideally, we would have made this dynamic to allow as many teams as the user wanted to be compared, but weren’t able to figure that out besides the proposal of hard-coding in the option for each different number.

Evaluation of Our Solution

Our solution even though it wasn’t executed to our wishes still surpasses most athletic databases in terms of innovation. The question that will always live alongside of sports is “what if”. What if Michael Jordan played Lebron James? What if Zlatan in his prime played Messi? Or Pele? What if Usain Bolt faced Maurice Greene, both in their prime? Our Hypothetical Feature is the beginning steps to this But overall our database compares well in terms of structure to other sports

databases for example ESPN and TFFRS. Our database, however, does lack in size compared to those, we were only able to scrape information about the most recent athletes, so in theory we could build a database from 2018 on but would never catch up to the likes of TFFRS.

A database we did discover that had a similar idea to our Hypothetical Feature was NFLComparisons.Com, which allows you to choose two NFL players from any era and compare their stats over the seasons you select. NFLComparisons does have some strengths we lack, in that their output page compares all stats side by side accompanied by a visual graph that allows for quick comparisons. The database is also larger which allows for a variety of comparison possibilities, but for each comparison the user in the end has to scroll through each stat and make a decision for themself, there is no ranking provided like ours. Along with that, NFLComparisons only allows head to head matchups of two where we can do five. This is significant due to the fact that comparisons only strengthen with more variables to compare. A third, fourth or fifth provides an average baseline and puts the relation in a tighter perspective. Also more variables allows the search to be completed more quickly, instead of selecting a new player two or three times we are able to complete it in one step.

In summary our solution was a step to solving the overall problem, scraping data, then creating and populating a database in itself was no easy task not to mention building a user friendly frontend. Having a larger set of data could have painted our project in a more quality light but with the data we had from 65 teams it wasn’t a detriment to our work. All in all, we are proud of our work and to have an element that well known athletic database websites lack. There is a niche in this field that has yet to be filled and if we were to collaborate with the likes of NFLComparisons and TFFRS it could be.

Possible Next Steps

In future work, drawing comparisons to similar models we can find a direction for us to follow. We took the idea of storing historical data, but currently our database only houses the current rosters of the 2018 Cross Country teams. For future progress we can look to refine data on non-cross/distance oriented athletes (in the field events as well) and look to gather this information on a more historical basis than the oldest person currently on the team. Yearly data for the database (housing multiple team rosters and years of info can provide more specific data and can help in hypothetical meets with more past results as a guide.) Since track meets are scored differently an approach to not only scoring individual events and cross country meets, but a hypothetical scoring system for say the ACC/Regional/NCAA championship in track as another point where people would be interesting in knowing how they stack up against others as they go for the win in the meet. A few other ideas that would go well with the hypothetical tests is the ability to see what would happen if a student transfers school and an option to limit the test to underclassmen and include the incoming freshman of year.