

CSC343 Term Project - Discussion

Soccer has grown to be a world-wide entertainment sport, so did the FIFA games which have a real-world database on players from different leagues and clubs. As soccer leagues' policies become open world-wide, players from different countries or continents are playing together. The interaction and dynamics between them make the sport and the game more attractive. Through building the database of players and languages, we can not only help FIFA gamers and real-world clubs find their desired players, but also explore demographic distribution and identify some potential problems in the soccer realm.

First, we can do some basic queries such as finding a lineup/team of best players that speak the same language. For example, in Q1, we find 11 top players in a certain formation but all speak English (Q1 table "lineup_for_433"). They can be formed in the game or in real-world (potentially a national team). Then we investigated the number of players and average ratings in each language group or cultural background, and discovered that Spanish and Portugal language communities have the highest average rating across players (Q1 table "comparision_table"). This aligns with the real-world scenarios where there are a lot of talented players from Iberian Peninsula and South America. These two language groups not only have the most players in the world, but highest overall skills in soccer.

Moreover, we can build a squad in the database by finding players of best price-performance ratio given a certain formation. In Q2, we defined price-performance to be player's value over player's overall rating, and by finding lowest price-performance players on each position in the formation 4-3-3, with minimum rating to be 80, we successfully built a squad of best price-performance-ratio (Q2 table "lowest_price_performance_ratio_433"). This type of query is useful when a gamer or coach wants to recruit the best players with a limited budget. It is to our surprise that all chosen players are rated around 80, which suggests that lower rating players tend to have better price-performance ratio. In addition, we compared the average price-performance ratio for different positions in soccer (Q2 table "price_performance_ratio_table"), attackers have much higher price-performance ratio than defenders, which is also expected in the real-world.

Lastly, we can query the database to find out phenomena in players' communication within some clubs. Since players of different language backgrounds are playing together, their communication may not be as effective if their language is not shared, especially in team sports like soccer. Q3 is a sample query to find clubs that have common language spoken between the best goalkeeper and top rated defenders in the club, as the communication in defensive lines is usually one of the most important aspects in professional real-world soccer matches. The result shows that only 129 out of 891 clubs in the database have defensive lines speaking the same native language (Q3 table "club_communication_comparison"), and the clubs of good communication are mostly small clubs that only have local players (Q3 table "good_communication_clubs"). It is reasonable to infer that the ability to communicate in a different language has become an important factor to consider when clubs look for new players. Apparently, most players need to learn at least a new language after they start a career in a foreign country with different languages, which is a prominent issue in the soccer world.

To conclude, the FIFA database we constructed can be used to filter and select players or squads that a FIFA gamer or real-world club manager requires, and can also be queried to find out some facts in the soccer world. Topics like soccer player demographic distribution, price-performance, as well as communication issues can be thoroughly investigated in the database. Ultimately, more data and tables can be extended to the current database, to give us better insights how real-world soccer is like.