Joseph Wiens
jwiens@uncc.edu

Project Deliverable I | Use Case, Data Model, and Projections

Section I: Business Context

Recruiting talent is an important part of any organization, but proves to be immensely consequential in the world of college football.

For larger college football programs in the SEC, Big Ten, and Big Twelve, recruiting success accounts for 63% to 80% of a team's on-field outcomes (Caro, 2012). Although large football operations possess enormous budgets to scout, engage, and select potential athletes, smaller programs must be more selective in employing resources to find and bring in the right talent. For example, the University of Georgia's 2018 Recruiting budget was $4,346,403, a figure nearly eight times that of University of North Carolina at Charlotte's 2018 Recruiting budget of $548,707 (McLaughlin, 2021; US Department, 2021).

Additionally, player recruit ratings, as assigned by recruiting services, massively influences on-field results; explaining up to 36% of the variability in final Sagarin team ratings, a sufficient proxy of team success (Mankin et al. 2019). As recruit ratings routinely dictates team success, modest football programs walk a tightrope; attempting to assemble the most talented roster possible while running a lean operation, expending fractions of the resources dispensed by college football juggernauts (US Department, 2021).

Section II: Business Use Case

As a smaller football program with a restricted budget, UNC Charlotte must align their recruiting scope to fit their allocated budget. Utilizing methods to target talent hotbeds by region, position, and rating, the UNC Charlotte football program can more efficiently and effectively administer resources to these areas. Analyzing annual and multi-year trends will help the program develop a more agile, purposeful, and prosperous recruiting strategy, which translates to on-field achievement.

Another vital consideration is the stature of UNC Charlotte's football team and its relative position in the college football ecosystem, which often determines the quality of prospect a program has the capacity to attract to the university. UNC Charlotte must pursue recruits appropriately and correctly attune expectations that correspond to the caliber of player that could realistically commit to play football. Likewise, a coherent and holistic recruiting process should emphasize different facets and advantages of the football program and university most likely to entice prospects according to their respective ratings (Mirabile and Witte, 2017).

Geographic data like city, longitude, latitude, and FIPS codes are instrumental tools to prepare an efficient recruiting strategy, aiding to identify clusters of players as potential recruits. Beyond fixating on local talent, travel expenses can be reduced through concentrating on clusters of players in an area simultaneously. Cultivating connections with local coaches and maintaining a presence in specific talent rich communities are beneficial activities that can pay dividends in later recruiting cycles.

Section III: Data Description

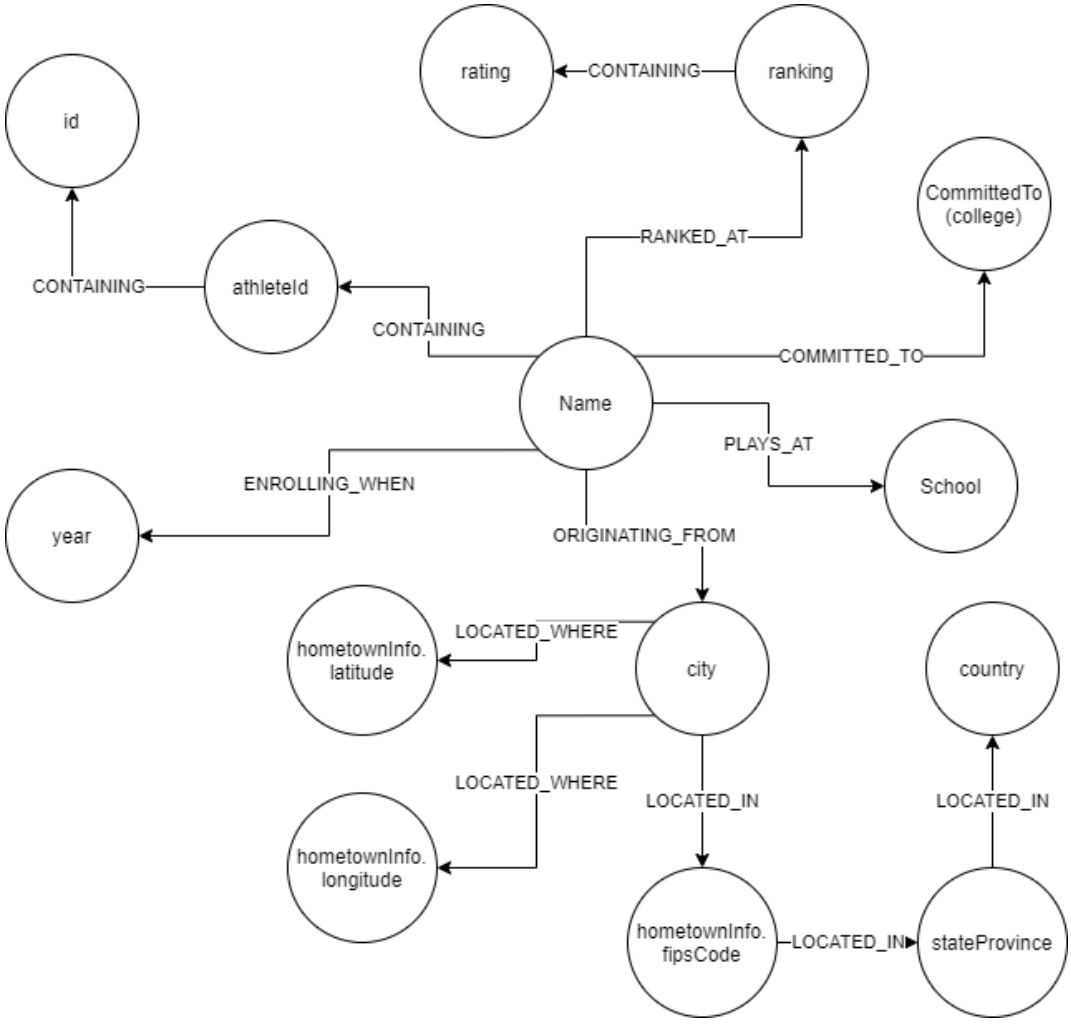
A yearly player recruit database hosted by Bill Radjewski (College Football Data website) compiles players' names, school, hometown, college commitment (if any), position, rank (by class), rating, height, weight, and more. This dataset provides player evaluation data through class rankings and a composite rating derived from various recruiting services that assess a player's skill and ability. A crucial aspect in shaping the recruiting strategy, the composite ratings component will establish a range of player quality that UNC Charlotte can feasibly draw to the football program.

This analysis will examine player recruits across six years of recruiting classes, from 2015 to 2020, comprising of 24,442 observations. Several records have missing data, most notably the "committedTo" column, denoting the recruit's college commitment, totaling 5,293 blank fields. A few hundred recruits hailing from foreign countries, primarily Australia, lack records in geographic variable columns, a required element for this exercise. After removing 5,800 rows of missing data, the clean dataset equals 18,642 observations, averaging 3,107 recruits per year.

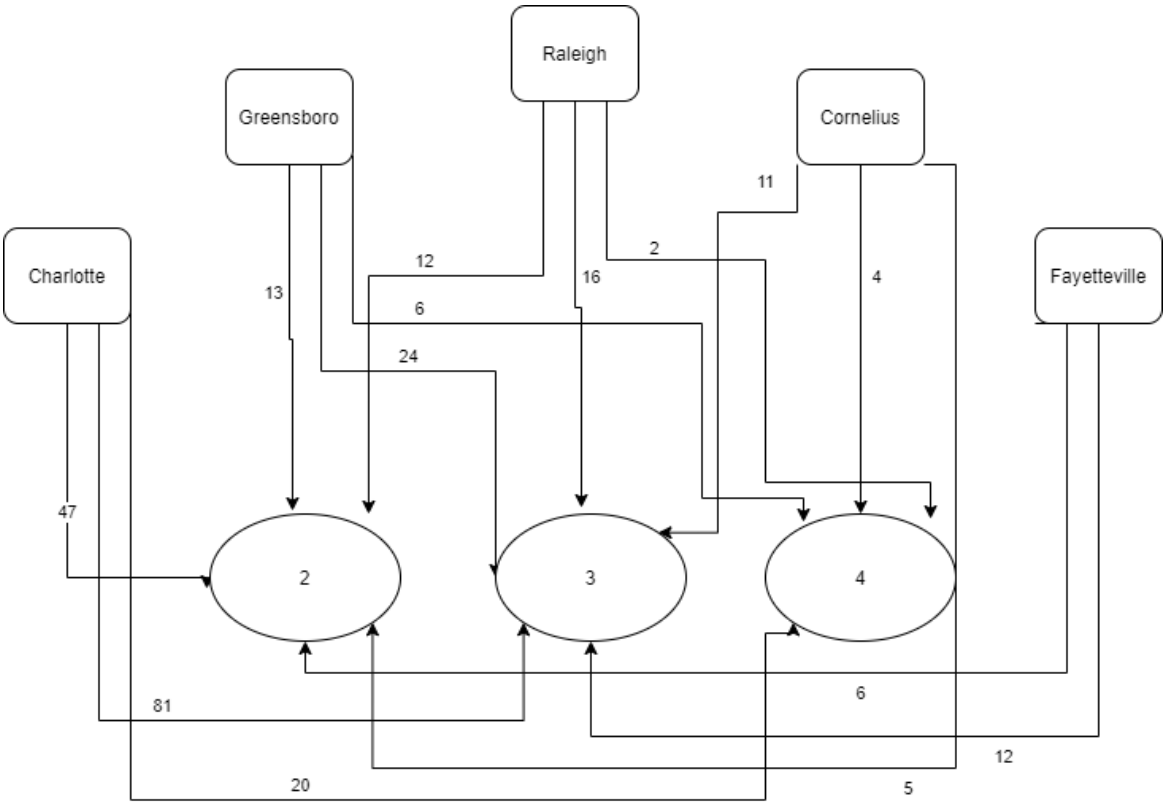
In the data's nineteen total variables, five are continuous variables including height, weight, rating, hometown longitude, and hometown latitude. While the "star" variable values range from one to five, it represents five categories of player classes corresponding to different ratings threshold. However, the "star" class is unevenly distributed: 87% of players falling into the two and three star categories, 11% designated four star status, and slightly more than one percent of all players are labeled as five star recruits.

After applying graph analytics techniques and algorithms to this dataset, new relationships can be illustrated between player ability, position, region, time, and other attributes, providing valuable insights to maximize recruiting results.

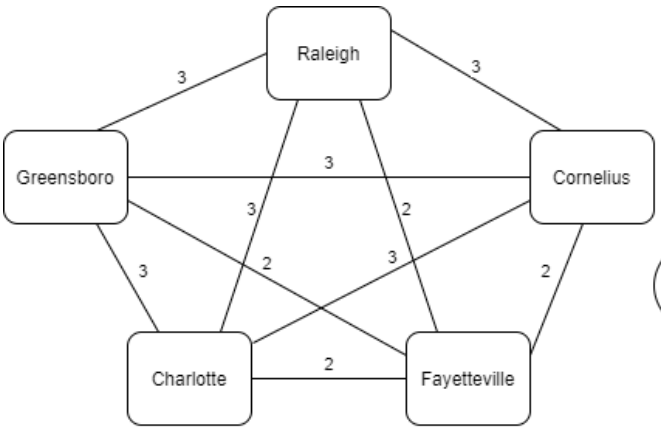
Section IV: Graph Data Model



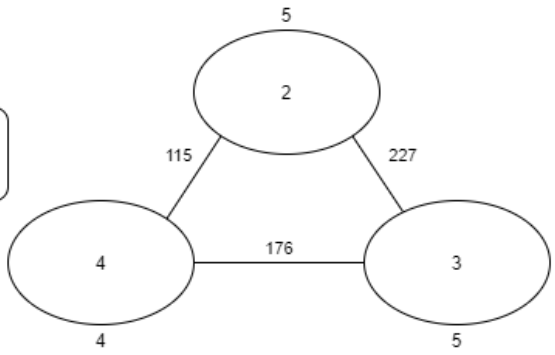
Section V: Graph Projections



Cities and Stars
Bipartite Graph
Relationship Weights = Number of recruits



Projection of Cities
Monopartite Graph
Relationship Weights = Number stars in common



Projection of Stars
Monopartite Graph
Node Weights = Number of cities with star value
Relationship Weights = Combined number stars by cities in common

References

Caro, C.A. (2012). *College Football Success: The Relationship between Recruiting and Winning* https://www.semanticscholar.org/paper/College-Football-Success%3A-The-Relationship-between-Caro/a715dec1d9fa4efcd2e6559f528bb4b222f803f8

Mankin, J., Rivas, J.A., & Jewell, J. (2019). The Effectiveness of College Football Recruiting Ratings in Predicting Team Success: A Longitudinal Study. *Behavioral & Experimental Finance (Editor's Choice) eJournal*.

McDonald Paul Mirabile & Mark David Witte, 2017. "A Discrete-Choice Model of a College Football Recruit's Program Selection Decision," *Journal of Sports Economics, vol. 18*(3), pages 211-238, April.

McLaughlin (2021). *College Football Recruiting and the Correlation to Success* https://digitalcommons.coastal.edu/cgi/viewcontent.cgi?article=1408&context=honors-theses

Radjewski, Bill. "College Football Data." *CollegeFootballData.com*, https://www.collegefootballdata.com/

US Department of Education. (2021). *Equity in Athletics Data Analysis: Fiscal Year 2018-2019.* https://ope.ed.gov/athletics/#/