

Web-Scraping and Data Visualization of NIL Student Athlete

BY: NARAYANI BAKHATI

INTRODUCTION



Student-athlete NIL deals found on nilcollegeathletes.com extract approximately 8391 athlete's schools, sports, sponsorships, and social media accounts



Objective: To scrape information about student-athlete NIL deals from the websites and organize it in a structured manner for further analysis.

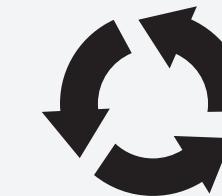


BIRD EYE VIEW ALGORITHM

Find total number of pages from the .nav tag

Construct n number of paginated urls appending page number to base url.

Iterate over each url and execute a scraping function for each.

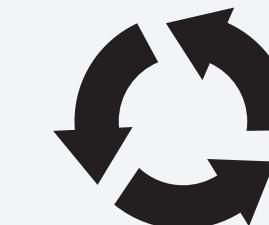


The Scraping function will go after the table data for that current page

Clean, group and visualize the data.

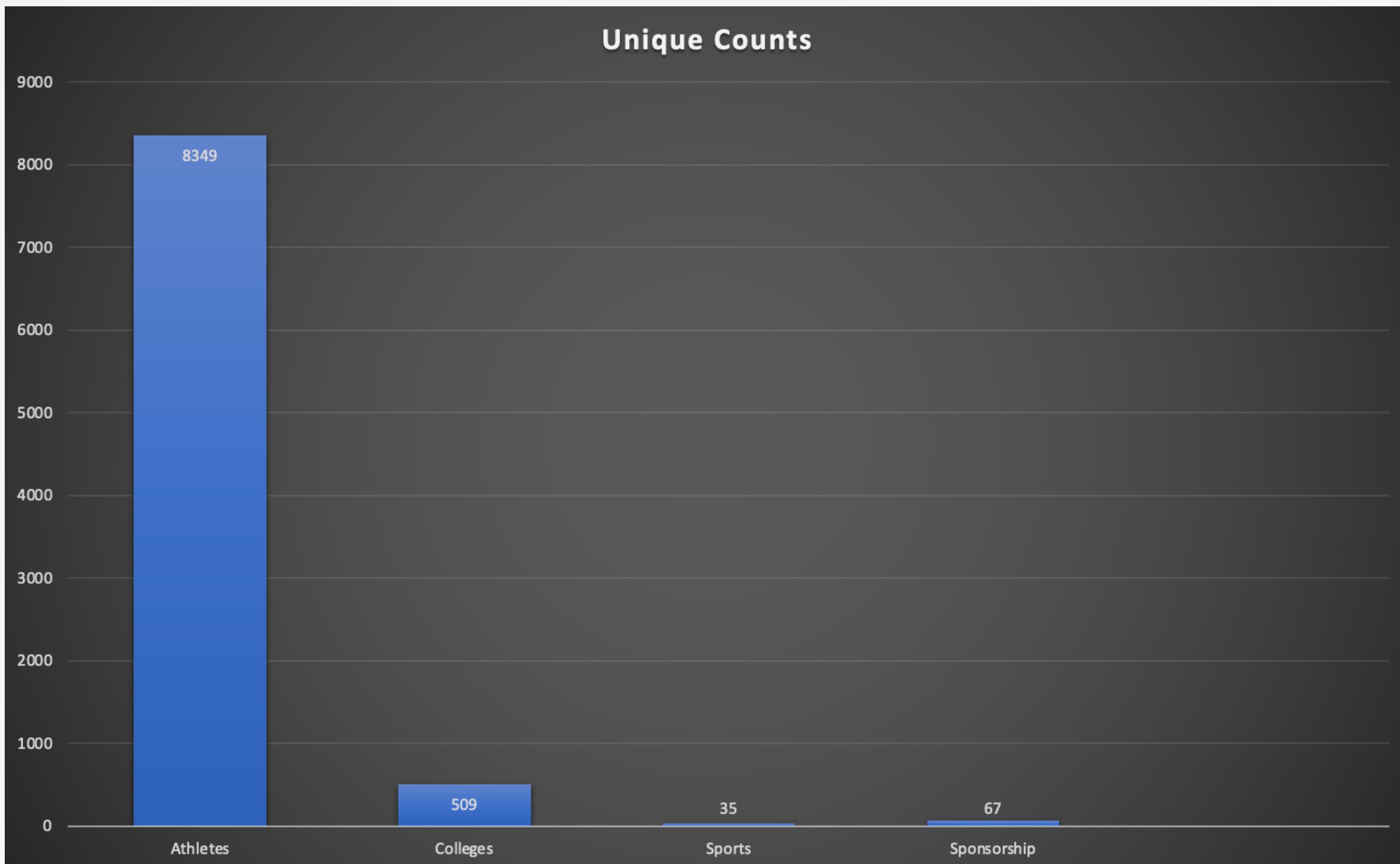
Collect the overall data in a list.

From the details page, scrape the data and create a dictionary.
{'dt':'dd'}



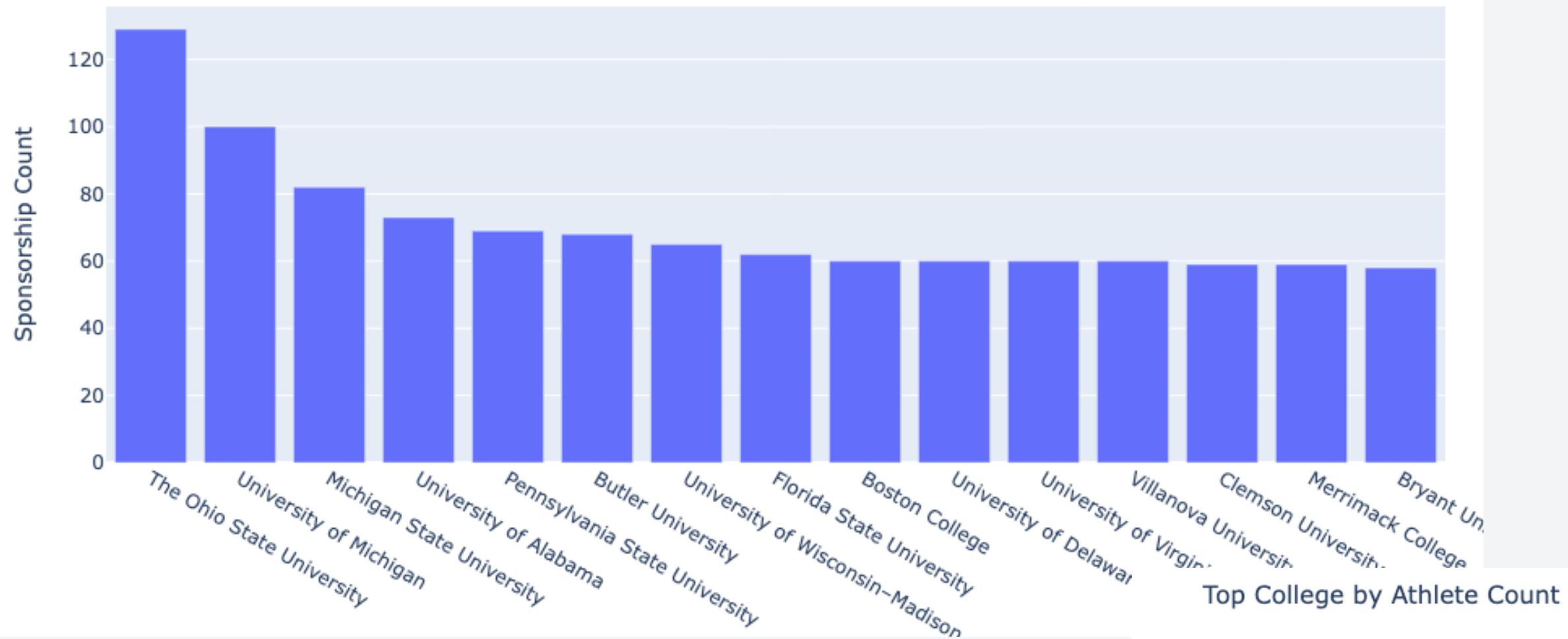
Iterate over each table row and retrieve the href for athlete details and scrape from that link.

PROJECT FINDINGS

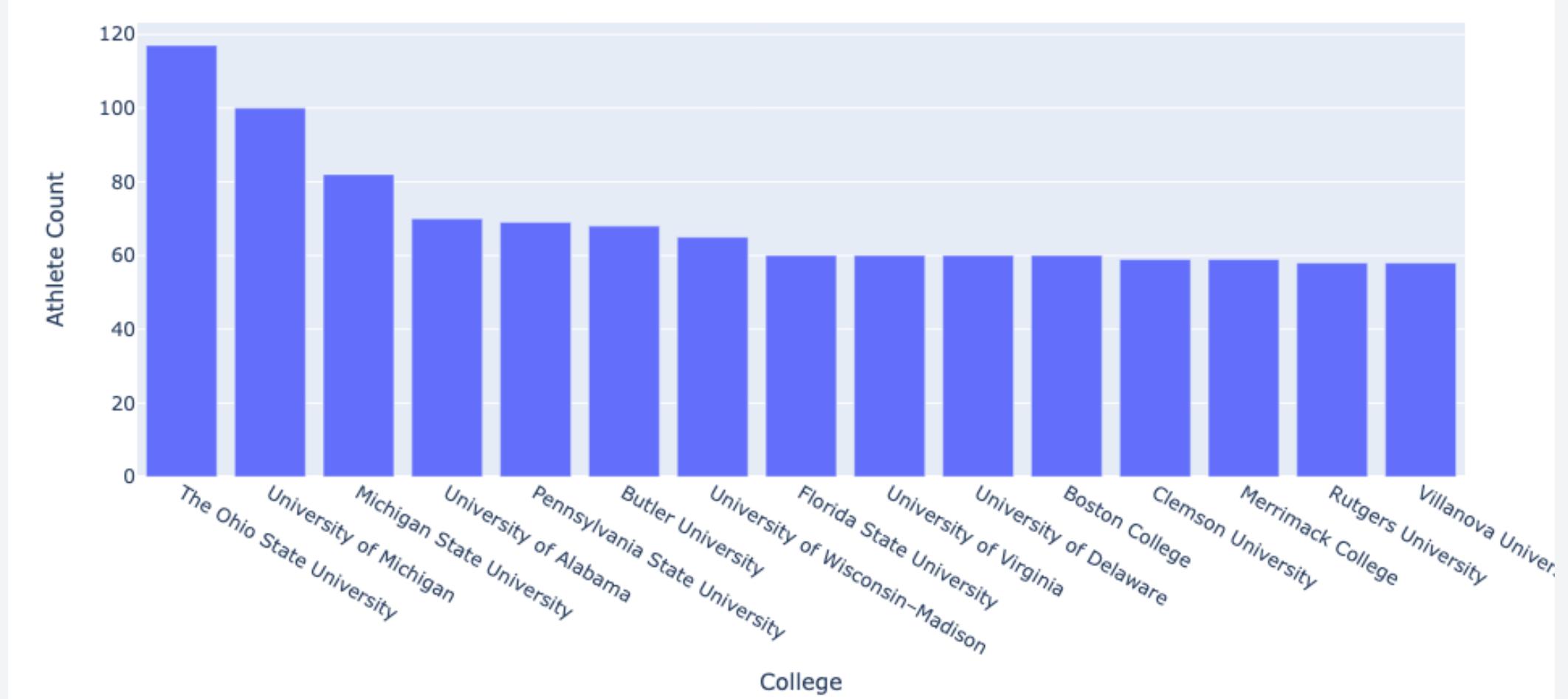


- Data for a total of 8349 NIL athletes was scrapped from the website.

Top Colleges by Sponsorship



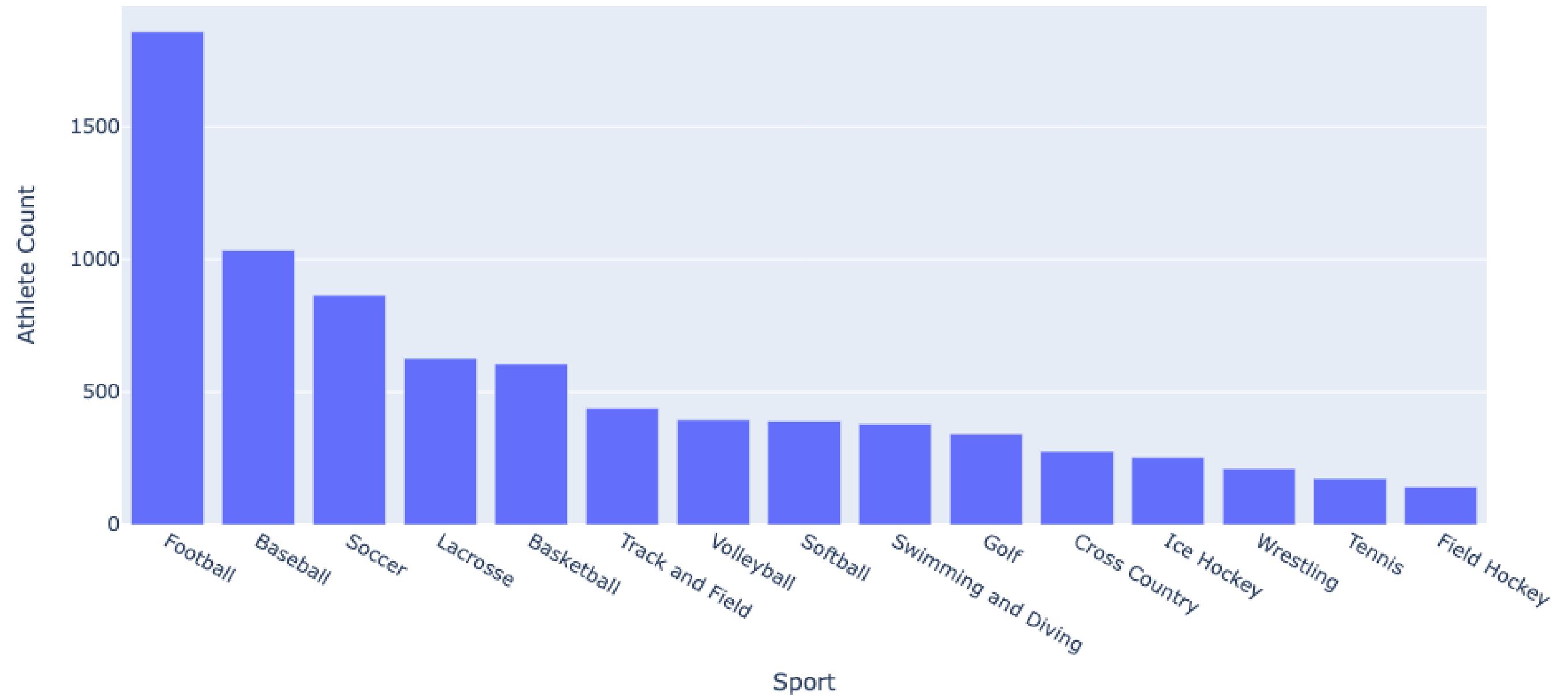
RESULTS



The Ohio State University:
129 Sponsorship counts, 117
athletes

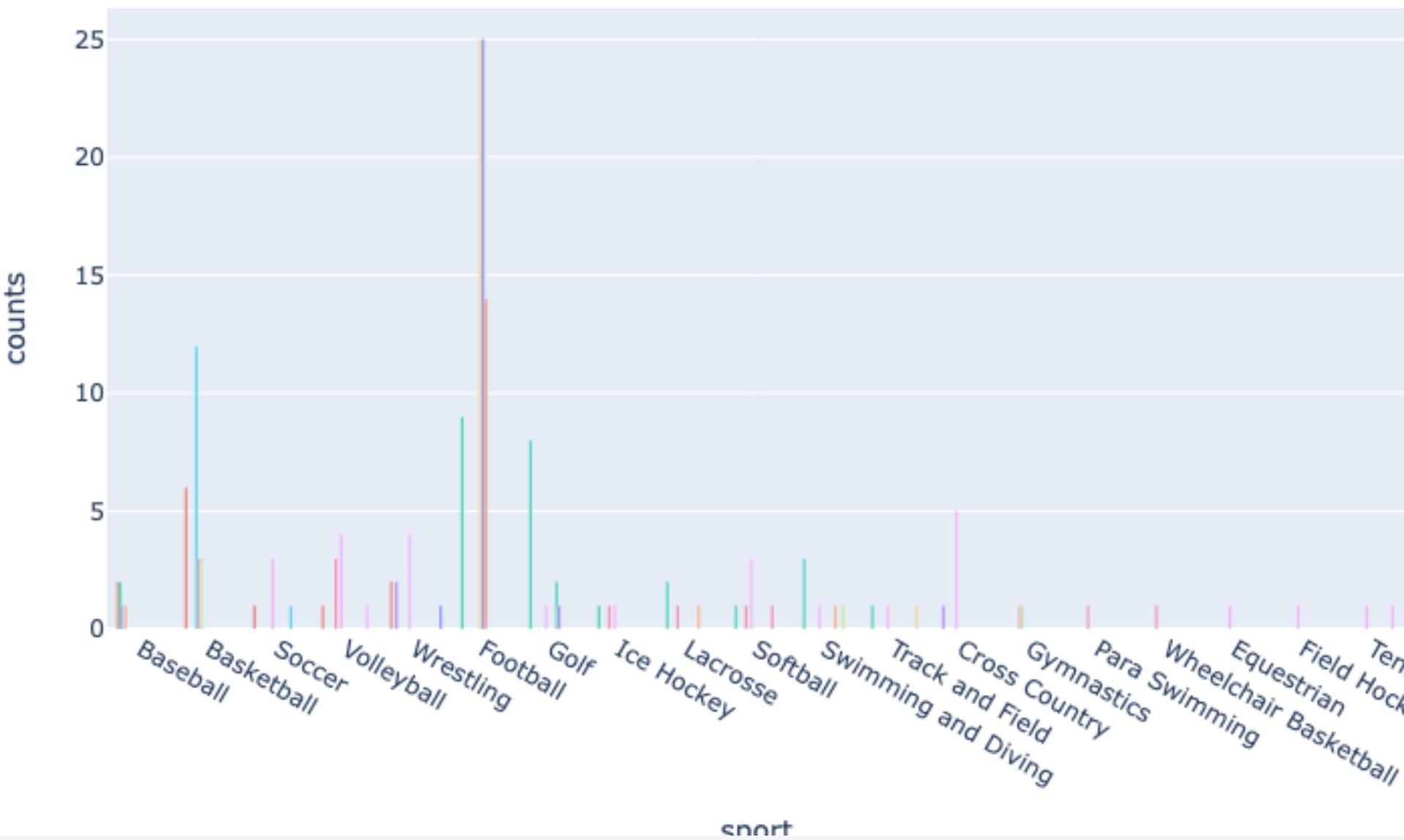
RESULTS

Top Sports by Athlete Count

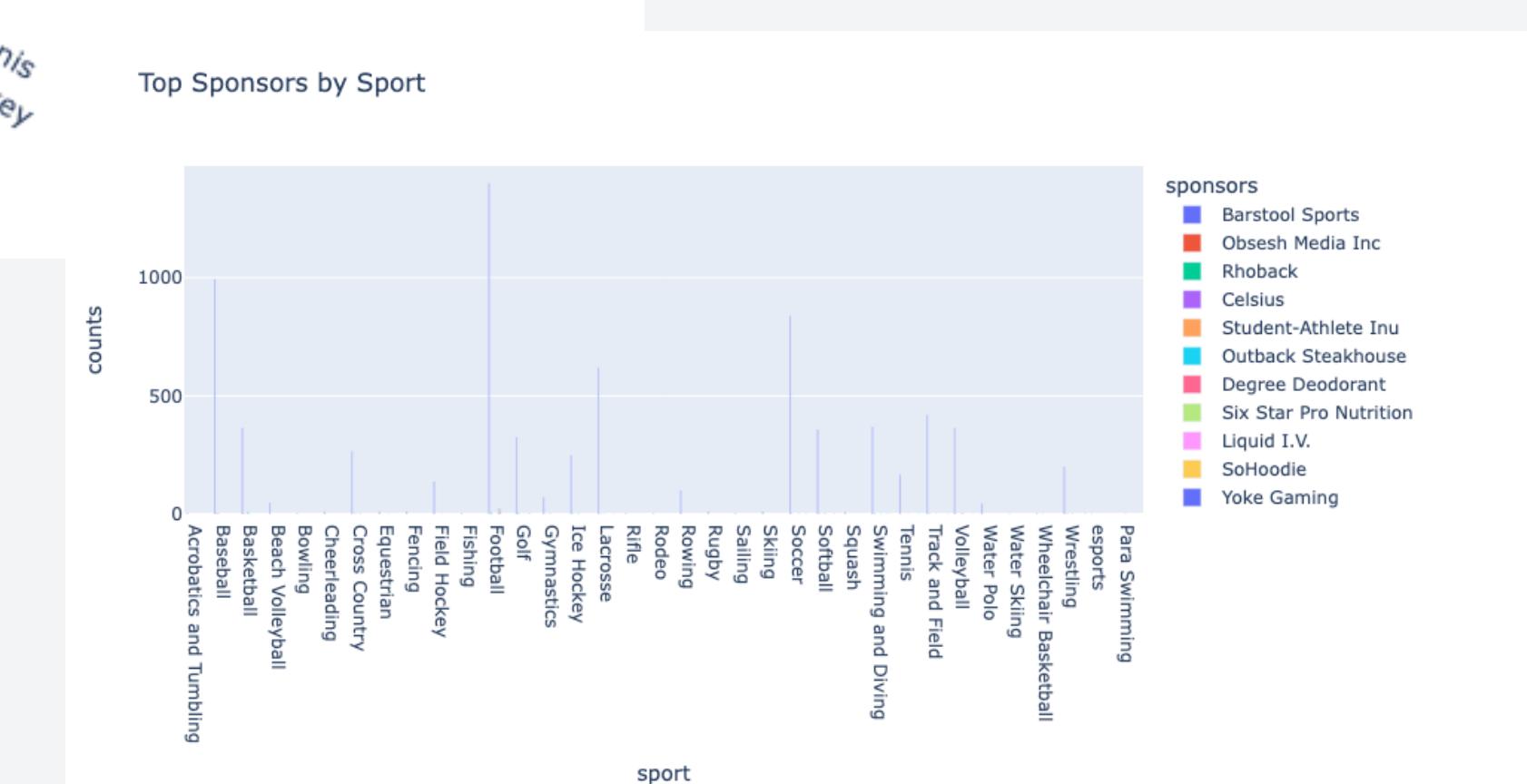


Football: Total 1860 Athletes

Top Sponsors by Sport



Football has the majority of sponsorship.
1398 football athletes receive Sponsorships from
Barstool Sports



ANALYSIS OF RESULTS



Our winners under different categories:

SPONSORSHIP

Athletes: Jason Hubbard
(12)

Sports: Football (1512)

College: Ohio State
University (98)

Sponsorship: Barstool
Sports (7438)

SPORTS

College: Ohio State
University (19 sports count)

ATHLETE

Sports: Football (1860)
College: Ohio State
University (117 athletes
count)

CONCLUSION

Challenges

- Data inconsistency
- Time and space complexity



Takeaways

- Data Visualization
- Optimize the code for efficiency

Learning Outcomes

- Got to apply course Knowledge
- Problem-solving

