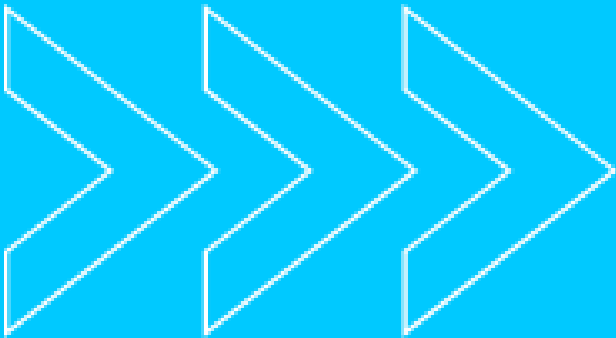Tyon Davis, Victoria Ramirez-Garcia, Cameron Davis,
Kairee Gay, Kiana Forbes

COLLEGE FOOTBALL PREDICTIONS

INTRODUCTION

The motivation behind this project stems from the need for a more reliable and data driven approach to predicting the future performance of college football players. Current methods, relying on TV sports analysts and individual assessments from sports websites like ESPN, present limitations in terms of subjectivity, emotional biases, and a focus on popular players and teams. By leveraging advanced data analytics, we aim to develop a program that provides accurate predictions, considering various performance metrics, player-specific characteristics, and game-specific factors. This initiative seeks to overcome the drawbacks of current practices and enhance the accuracy and inclusivity of player performance predictions in college football.

UPDATE DATA

The provided Python program is designed for web scraping college football player statistics from ESPN's website. It employs a modular structure, utilizing various external functions and modules such as set Driver, UpdateCSV, scrape_header, week_finder, and get_url. The main function, scrape_players_and_stats, takes advantage of the BeautifulSoup library for HTML parsing and Selenium for automated browser interactions. Notable patterns include the handling of dynamic content through a while loop that clicks the "Show More" button until it's no longer available. The code demonstrates a systematic approach to collecting player names, schools, and statistics, grouping the latter by headers and saving the data to CSV files based on the URL. Additionally, the scrape_team_future_info function is designed to extract future team information, including details like away team, home team, time, TV information, and ticket details. The use of conditional statements helps determine the type of statistics being scraped based on the URL, directing the data to the appropriate CSV file. Overall, the program showcases efficient web scraping practices, organized code structure, and automation for data collection from dynamic web pages.(Scrape data into 19 csv files)

CHARTS

The webpage starts off with three interactive charts. Allowing the user to dissect college defense by category.

+ + + Additionally, the application incorporates hidden charts (Passing, Running, Sack, Score, Total, Total Histogram) that become visible upon users selecting two teams, providing detailed insights into specific facets of the game.

+ + + Included data : Player performance, Tactical Adjustments, Opponent Analysis, Game Predictions, Fan Engagement, Overall Team Performance and Scouting Opponents.
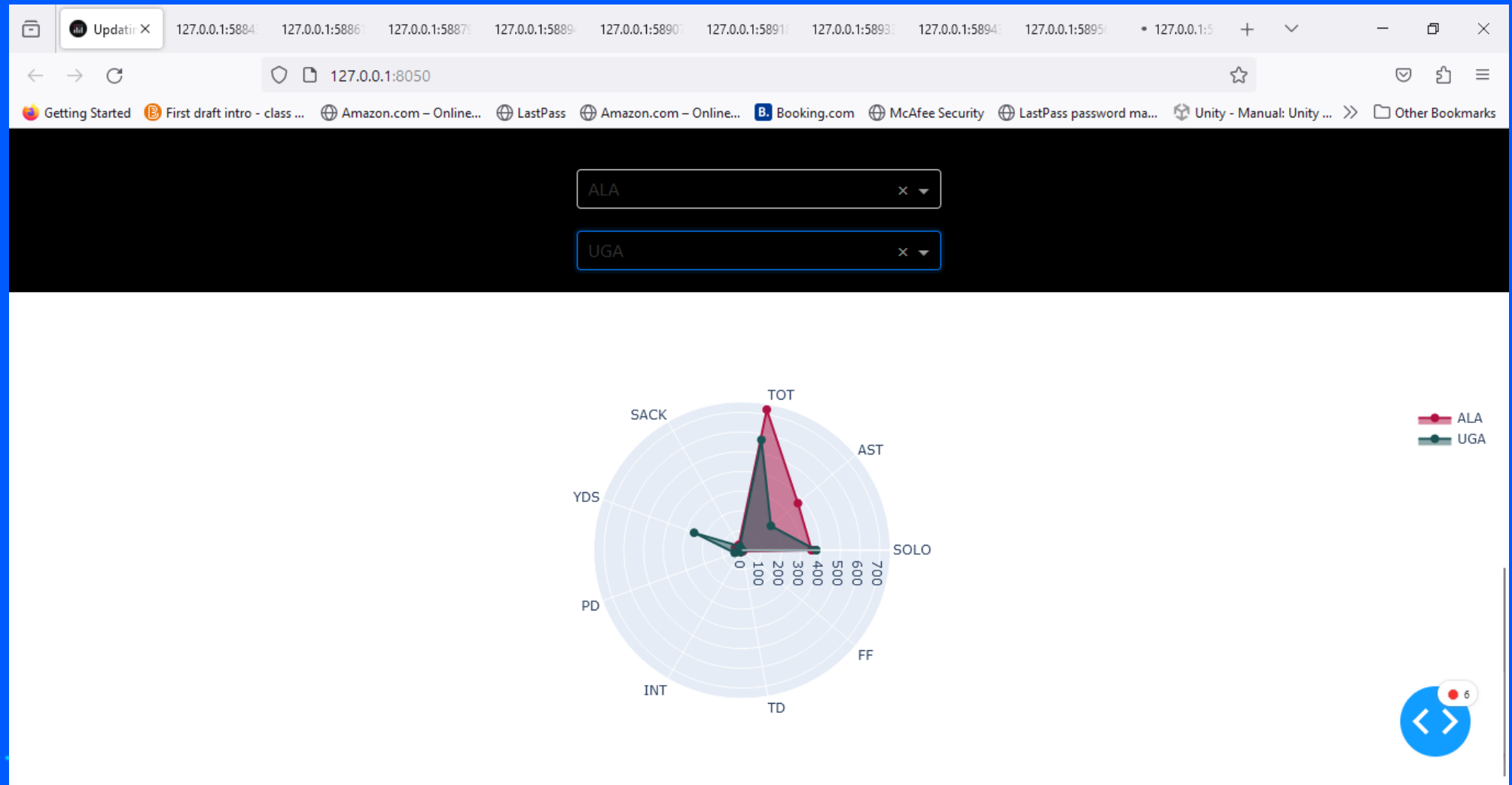
U.C.C. FOR DEFENSIVE PERFORMANCE
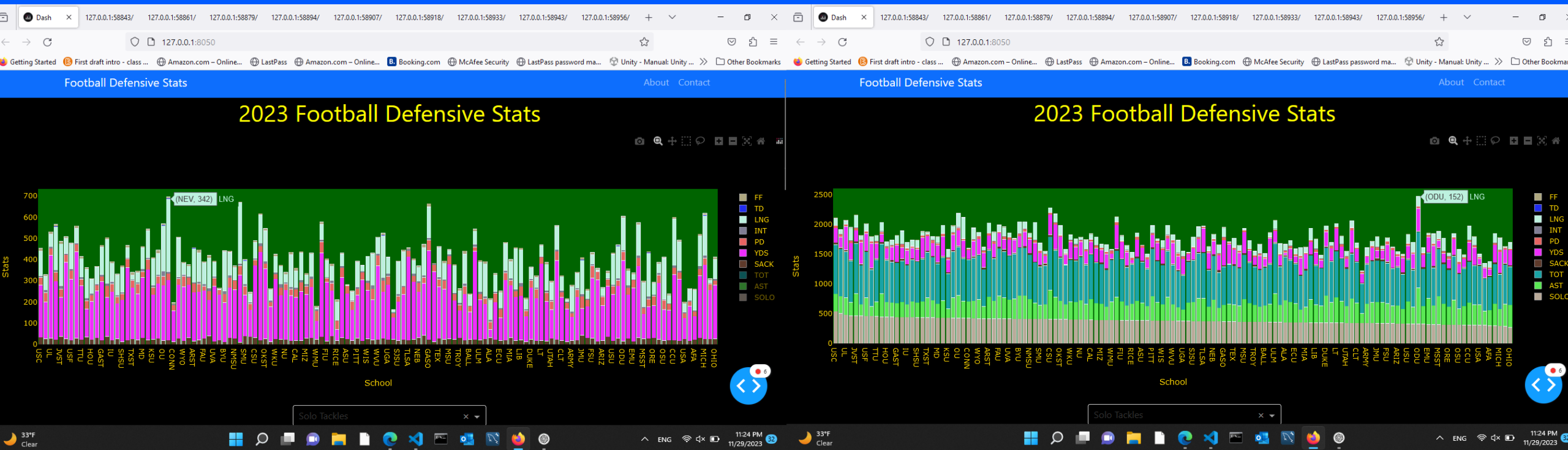
CATEGORY-SPECIFIC INSIGHTS

What teams excel in specific defensive categories, such as Solo Tackles, Sacks, Interceptions, or Forced Fumbles? - Comparative Analysis: (Dynamic)
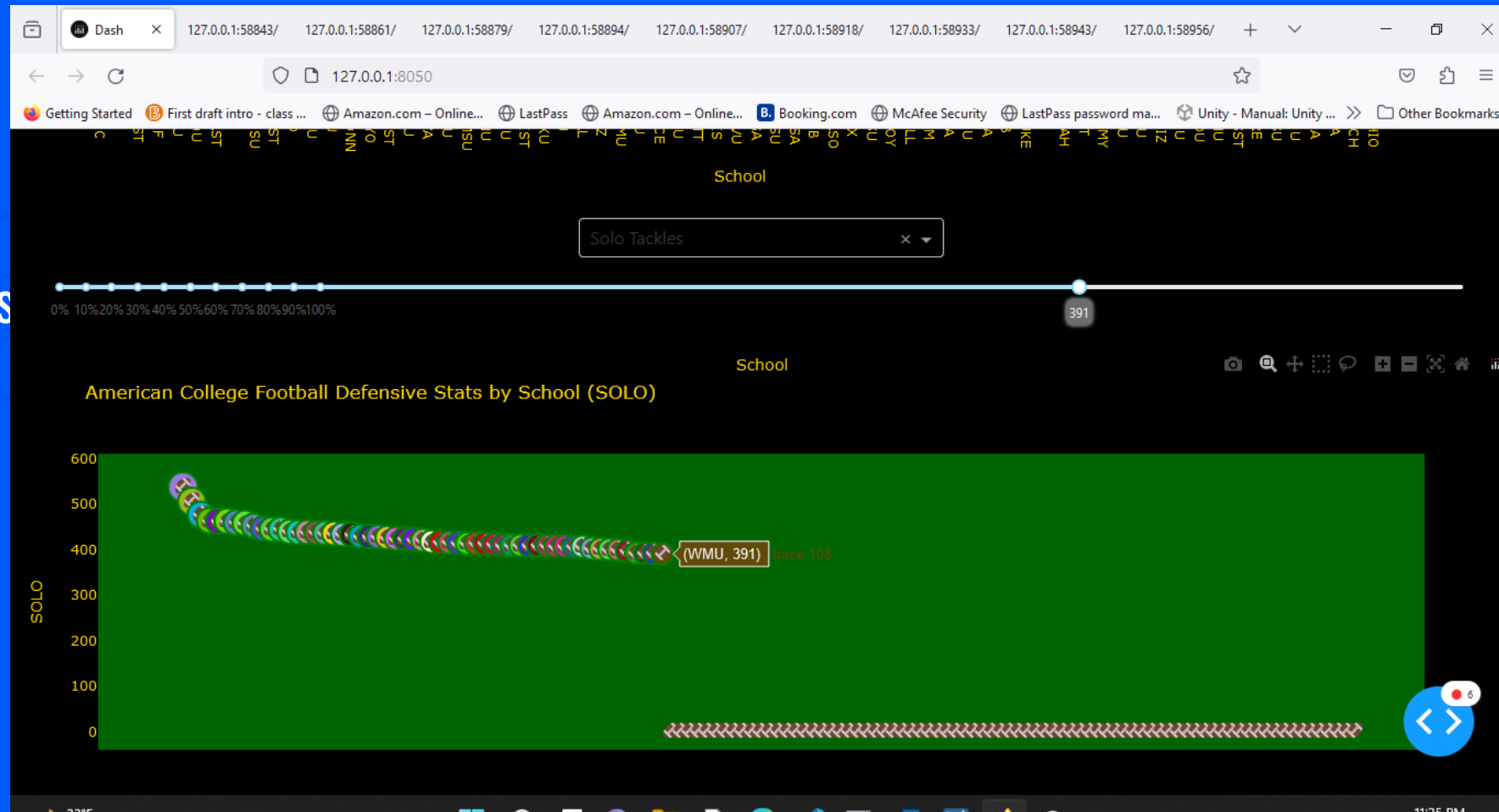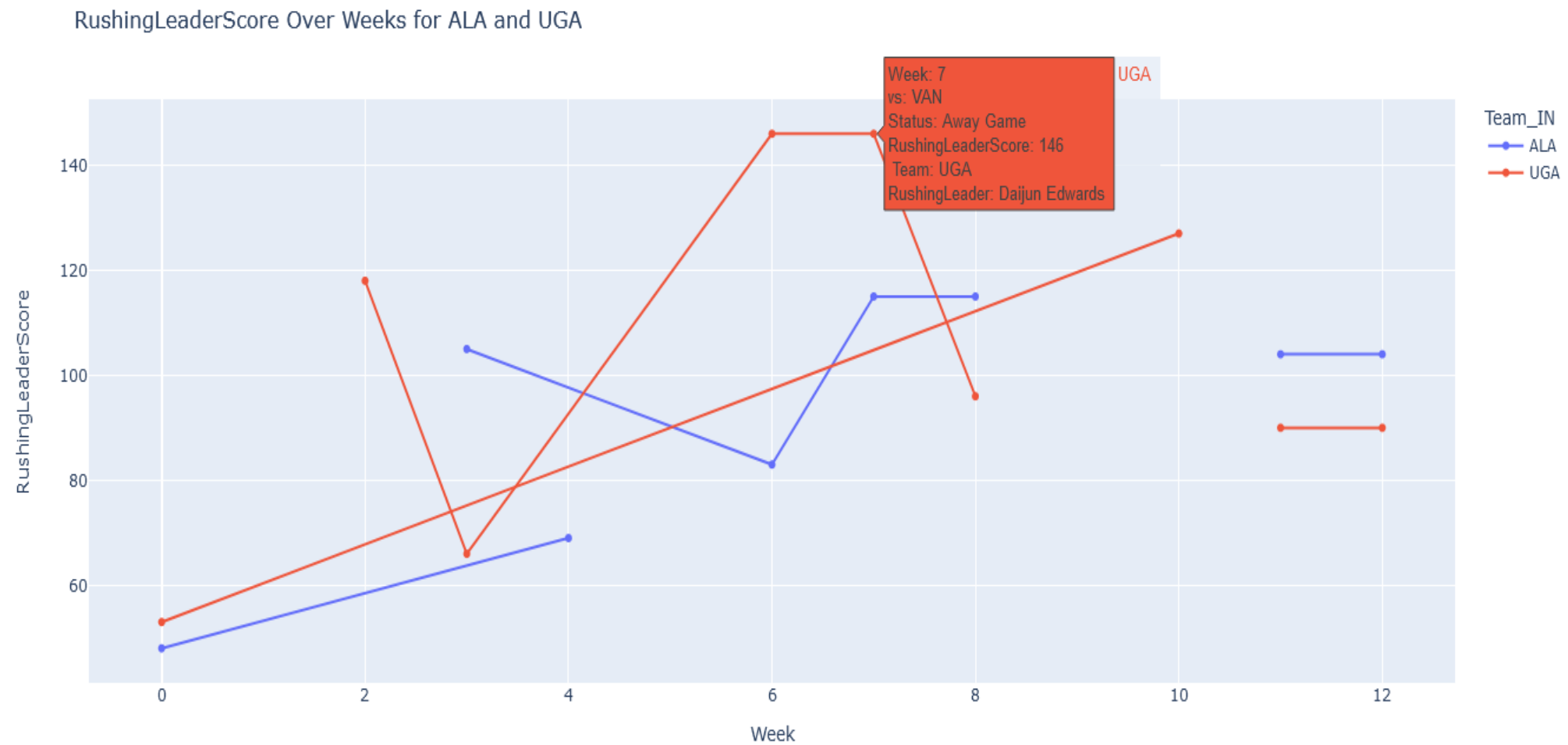
How do two selected teams compare in terms of overall defensive performance and specific defensive categories

Which team is a running team or passing team? Is the team balanced with running and passing stats?

RushingLeaderScore Over Weeks for ALA and UGA

Week: 7
vs: VAN
Status: Away Game
RushingLeaderScore: 146
Team: UGA
RushingLeader: Daijun Edwards

UGA

Team_IN
— ALA
— UGA

RushingLeaderScore

Week

Week-by-Week Score Difference for Selected Teams

Distribution of Total Tackles for All Schools

DETAILS OF THE EXPERIMENTS; OBSERVATIONS

During our experimental sessions on PrizePicks, where users utilized our program in groups of two players specializing as quarterbacks (QB), running backs (RB), or wide receivers (WR), we implemented a betting strategy based on the numerical values provided by PrizePicks. Users were able to make predictions on whether the actual scores would be higher or lower than the given numbers. The graphical representation of our outcomes showed green lines for correct predictions and red lines for incorrect ones. Remarkably, out of the 16 players assessed, users employing our program achieved accurate predictions for 13 instances. This success rate suggests that our program demonstrated effectiveness in assisting users in making informed decisions. Notably, our observations indicated that PrizePicks often set scores within a range where outcomes could swing either way for certain games. This aligns with our approach, showcasing that users, with the assistance of our program, were able to capture the unpredictable nature of the scores

HOW PRIZE PICKS WORK
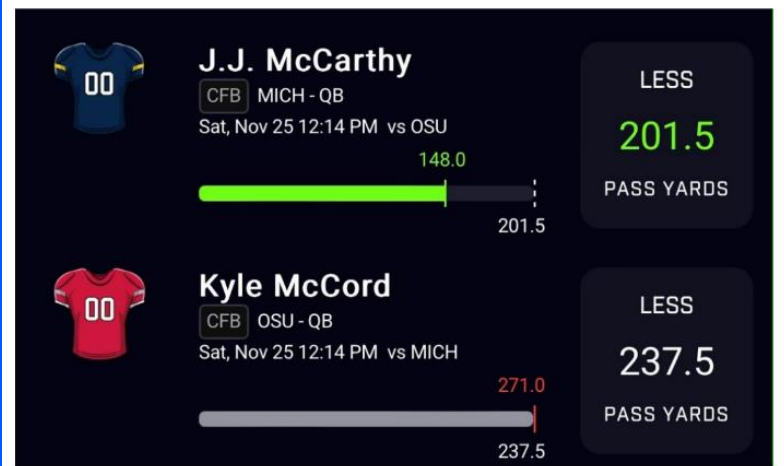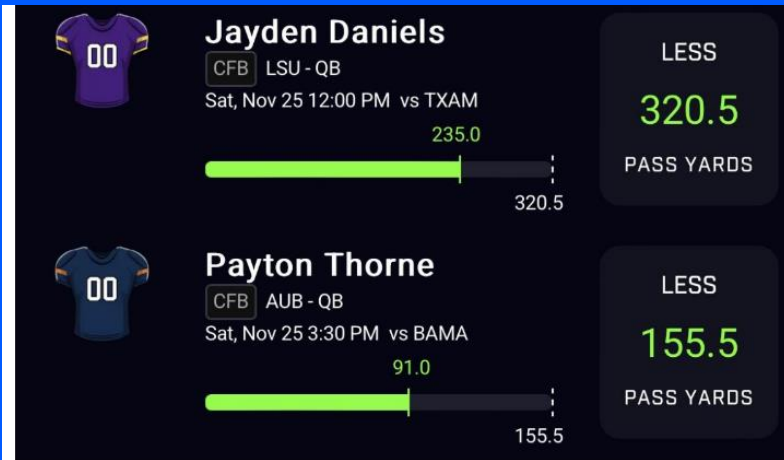
CONCLUSION

- In conclusion, for those interested in furthering this project, a logical next step would be to enhance the program's capabilities to autonomously generate predictions for users. Our experience highlighted the challenges associated with web scraping, with messy data retrieval being a common hurdle. Leveraging tools like Selenium proved invaluable for programmers in streamlining the process. Once the data was successfully scraped and stored in CSV files, the subsequent analysis became significantly more manageable.

- From the collected data, valuable insights emerged, particularly regarding the Total (TOT) chart. It was observed that the TOT chart, resembling a bell curve, provides a useful indicator for predicting potential winners. Teams positioned to the left with lower TOT values were found to have a higher likelihood of winning. This observation implies that TOT serves as a significant predictor of game outcomes.

CONCLUSION

- Overall, our findings suggest that integrating the program into the prediction process enhances accuracy for users. However, to further improve the program, programmers should consider incorporating additional variables that account for scenarios where score differentials become substantial, or key players are benched, impacting overall team performance. It's crucial to acknowledge that while predictions contribute valuable insights, the inherent unpredictability of sports, including off days for teams, implies that correctness cannot be guaranteed in every instance