College Football Attendance and Enrollment Demographics

Caeden Kropf

1. Introduction

Every fall, college football stadiums across the United States become some of the largest recurring gatherings in American public life. College football is consistently ranked up with the NFL as Americans' favorite sport to watch, making football one of the most powerful cultural and economic forces in higher education[1]. Yet despite its visibility, what actually drives the enormous variation in program attendance remains unclear. Some programs consistently fill 100,000-seat stadiums, while others struggle to reach 25% capacity.

This project investigates whether institutional characteristics, specifically win percentage, undergraduate enrollment, and conference affiliation, help explain differences in average home football attendance across Division I programs. Unlike fan-base narratives or historical prestige measures, these variables allow us to test whether attendance is structurally connected to the type of university a football team represents.

In this project, I plan to use 2021 average home attendance data from CFBStats[2], institutional demographics from IPEDS[3], and additional locational and branding data from CFBD[4] to evaluate whether enrollment size and gender composition contribute to differences in football crowd levels across Division I programs. By merging attendance outcomes with underlying student population characteristics, I will test whether stadium turnout reflects institutional scale and identity rather than simply on-field success or historical reputation.

2. Data

2.1 CFBStats

The dataset used in this project combines publicly available attendance figures with institutional demographic information. Average home football attendance by team was obtained from CFBStats, which provides detailed game-level box score statistics for all FBS programs. From these records, every home game from the 2024 season was extracted, and an average home attendance value was calculated for each institution. This required identifying home game designations, removing neutral-site contests, and standardizing

---

[1] https://footballfoundation.org/news/2012/3/8/_51405.aspx?utm
[2] https://cfbstats.com/
[3] https://educationdata.urban.org/documentation/colleges.html
[4] https://collegefootballdata.com/exporter/teams/fbs#google_vignette

reported attendance values. The resulting attendance variable reflects the typical draw that each football program generated over the course of the season.

2.2 IPEDS

Institutional demographic information was sourced from the Integrated Postsecondary Education Data System (IPEDS) using the Urban Institute Educational Data Portal, which reports annual enrollment profiles for all accredited U.S. postsecondary institutions. Specifically, total undergraduate enrollment, as well as male and female enrollment counts, were used to construct two variables of interest: total enrollment and the male-to-female enrollment ratio. In addition to enrollment data, IPEDS provided institution identifiers, official naming conventions, and location information (city and state). These identifiers were used to align institutions across both datasets.

2.3 Branding and Locational Dataset

In addition to demographic and attendance inputs, spatial and branding data were incorporated to support visual analysis. Stadium location coordinates and official team identifiers were obtained through the CFBD open-source college football resource (https://github.com/CFBD/cfb-web), used this repository of the CFBD), which maintains updated mappings between institution, team abbreviation, stadium name, and location. This same repository also included the official logo file set, which enabled the creation of visualization outputs in which university logos substitute for traditional scatterplot points. Incorporating these identifiers ensured that both attendance and enrollment variables could be represented clearly, while also preserving visual alignment between team brand identity and plotted value.

2.4 Cleaning and Merging

Data cleaning steps were necessary because attendance records and institutional data are not directly keyed to one another. First, city names were standardized to remove punctuation and formatting inconsistencies. Second, select institutions required naming alignment, particularly service academies and formal university naming variants (e.g., "United States Military Academy" to "Army"). Since some cities contain more than one higher education institution (e.g., Los Angeles, Miami, Houston), attendance values were matched to institutions using both standardized name mapping and athletic classification filters to ensure that FBS programs were isolated from neighboring non-FBS institutions.

Following these cleaning and merging procedures, the finalized dataset included average home attendance, total undergraduate enrollment, gender composition, athletic classification, and institutional identifiers for each program. This integrated dataset allows for descriptive assessment of attendance patterns as well as statistical and visual

comparison against institutional characteristics. By joining sports attendance metrics to demographic and structural campus attributes, this dataset makes it possible to evaluate whether the size and composition of a university meaningfully relate to its football audience levels.

Table 1 Data Dictionary

Field	Type	Source	Description
team_name	Text	CFBStats	Official name of the football program as listed by CFBStats or normalized during merge
conference	Text	CFBStats	NCAA conference affiliation of the program
avg_home_attendance	Numeric	CFBStats	Average home game attendance computed from 2021 game-level data
num_home_games	Numeric	CFBStats	Number of home games counted in the attendance average
win_pct	Numeric	CFBStats	Overall winning percentage computed from "All Games" record (wins ÷ total games)
enrollment_total	Numeric	IPEDS	Total undergraduate enrollment at the institution
enrollment_men	Numeric	IPEDS	Total enrolled undergraduate men
enrollment_women	Numeric	IPEDS	Total enrolled undergraduate women
men_women_ratio	Numeric	Computed	Ratio of male to female undergraduate enrollment (men ÷ women)

state_abbr	Text	IPEDS	Two-letter state code of the institution
city	Text	IPEDS / CFBStats	Campus city normalized for merges; used to join attendance and enrollment data
color	Text	CFBD GitHub Repository	Primary school color used for visualization
alternate_color	Text	CFBD GitHub Repository	Second color used for visualization where applicable
logo_id	Numeric	CFBD GitHub Repository	Identifier linking institutions to CFBD logo database to use for visualizations
logo	Image File (.png)	CFBD GitHub Repository	Official team logo used for scatterplots and visual outputs (found in logo file)
attendance_enrollment_ratio	Numeric	Computed	Average home attendance divided by total enrollment

3. Analysis

3.1 Win Percentage and Home Attendance

My hypothesis for the question of whether team success drives higher attendance was that teams with higher winning percentages have significantly higher average home attendance.

To investigate whether team success in a season influences home football attendance, I analyzed the relationship between winning percentage and average home attendance across FBS programs. I began by calculating a Pearson correlation coefficient using win percentage as the measure of team performance and average home attendance as the proxy for fan turnout. The resulting coefficient was r = 0.346 with a p-value < 0.001, indicating a statistically significant but moderately weak positive relationship between the two variables. While winning teams do tend to draw more fans, win percentage alone does not account for the large variance in attendance figures.

To visualize this relationship, I created a scatterplot with team logos displayed at their corresponding coordinates and added a simple linear regression trend line (Figure 1). Although the trend line slopes upward, programs with exceptional attendance, such as Nebraska, Alabama, Texas, Penn State, Michigan, and Ohio State, remain clustered well above the line regardless of fluctuations in win percentage. Conversely, teams like Houston, Coastal Carolina, and San Diego State achieved high winning percentages but did not experience the same attendance boost, suggesting that conference reputation, long-term football culture, alumni size, and regional demand likely outweigh short-term on-field success.

From the scatterplot, it becomes clear that high-performing seasons do not uniformly translate into proportionally higher attendance. Instead, attendance is strongly anchored by the brand equity and historical draw of specific programs rather than year-to-year performance. For example, Nebraska continues to pack its stadium despite a series of average records in recent seasons, while rising but smaller programs have not seen attendance climb at the same rate as their win totals. These findings align with prior research showing that Power Five schools, especially those in the SEC and Big Ten, sustain demand through entrenched fan loyalty, regional culture, and longstanding media presence rather than a single season's performance alone.

In short, although winning percentage does have a measurable and statistically significant effect on average home attendance, the relationship is not strong enough to conclude that on-field performance is the primary driver. The visual and statistical results both suggest that institutional football heritage and conference prestige serve as stronger determinants of turnout than seasonal win totals.
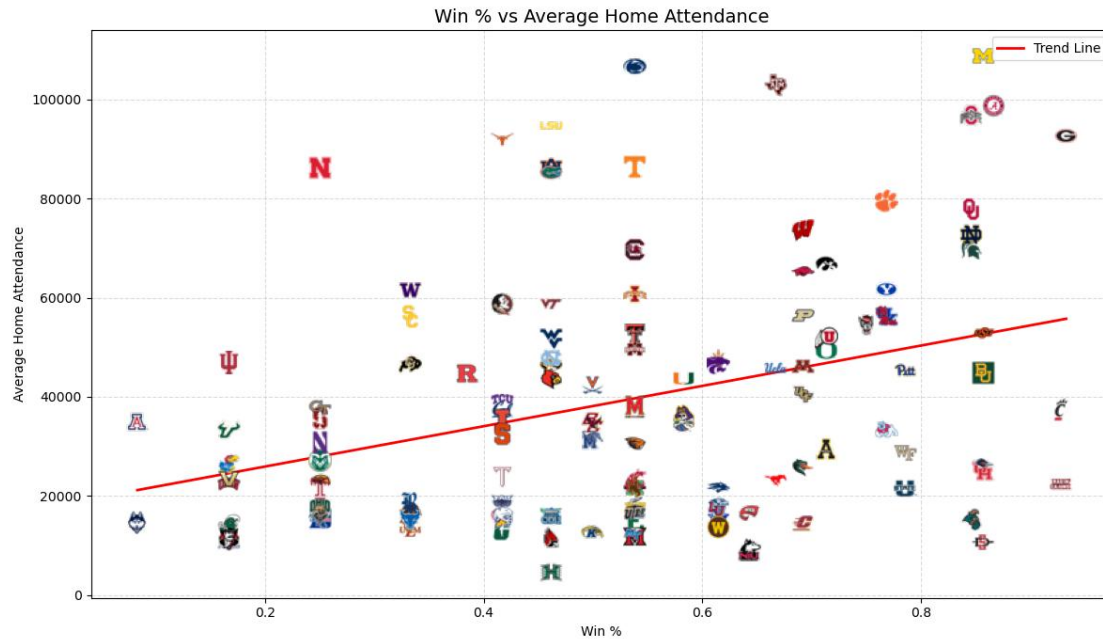
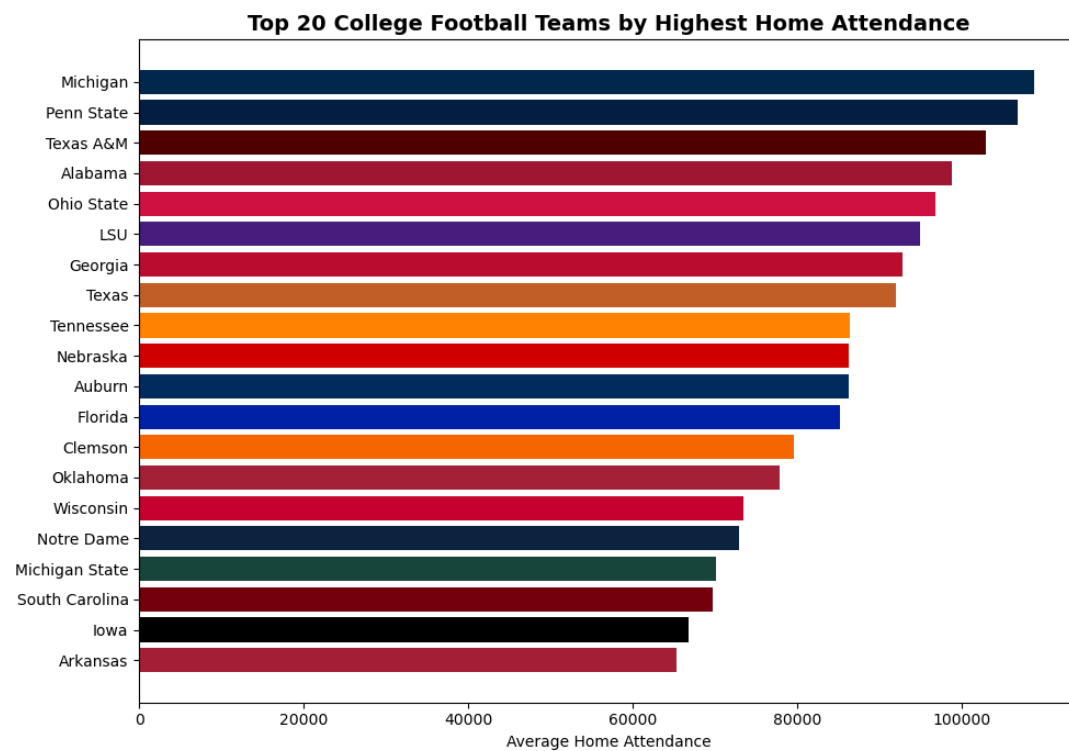Figure 1 Scatterplot of Win Percentage and Average Home Attendance with Trend Line
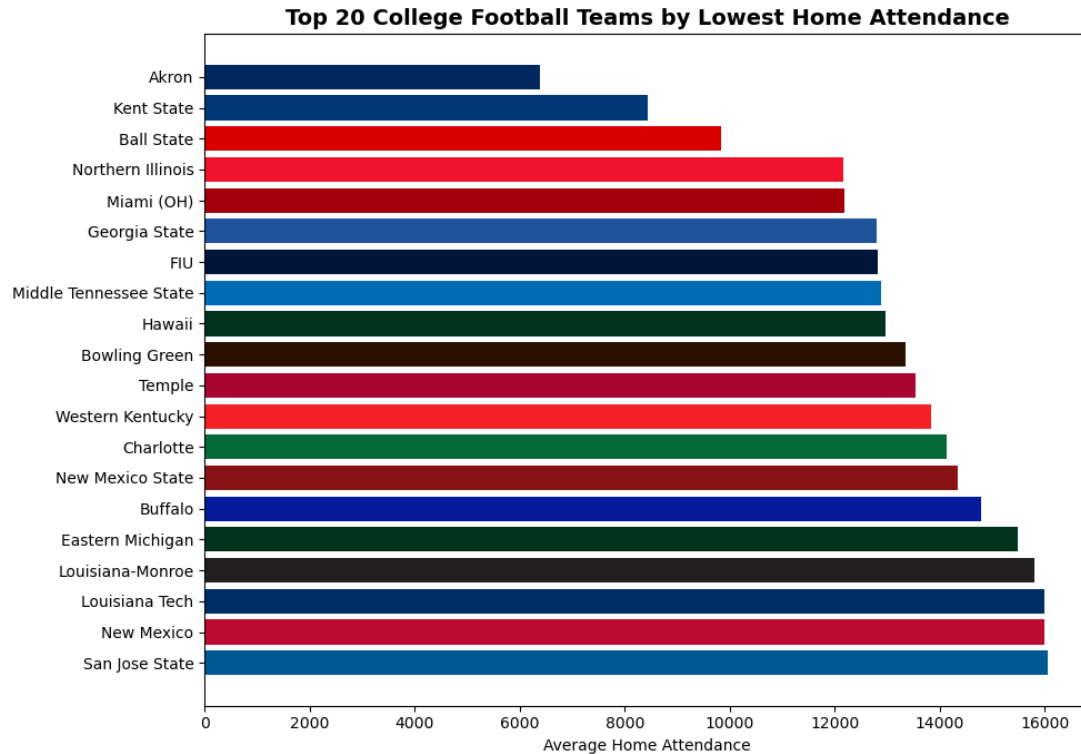

Figure 2 Bar Chart of Teams with Highest Home Attendance

Figure 3 Bar Chart of Teams with Lowest Home Attendance

3.2 Enrollment and Home Attendance

My second research question examines whether the size of a university's student body influences the level of fan turnout at home football games. Based on the expectations described in my hypothesis that larger institutions attract higher average home attendance, driven by broader alumni networks, larger student populations, and more visible athletic brands.

To evaluate this relationship, I calculated the Pearson correlation coefficient between total undergraduate enrollment and average home attendance across all FBS programs. The resulting value was r = 0.5853, with a p-value of 7.57×10^{-13}, indicating a moderately strong and statistically significant positive relationship between enrollment and attendance. In practical terms, this suggests that bigger schools do tend to draw larger crowds, but enrollment alone does not fully explain the substantial variation across programs.

Figure 4 illustrates this pattern. While the regression line slopes upward, indicating a general positive association, several programs sit far above or below the line. Large state flagships such as Michigan, LSU, Ohio State, Texas A&M, and Alabama dramatically outperform the attendance predicted by their enrollment numbers. These schools demonstrate that institutional size amplifies, but does not create, the historically entrenched fan cultures that define Power Five football.

Conversely, some universities with very high enrollments, such as UCF, Arizona State, and Houston, fall below the trend line. Despite having large student bodies, they do not achieve proportionally high attendance, likely due to weaker football traditions, local competition from professional sports, or more commuter-oriented student populations. This reinforces that enrollment is a structural factor, not a cultural one; football fandom depends as much on history, regional identity, and conference prestige as it does on institutional scale.

To better understand the underlying structure of enrollment across the sport, I produced an enrollment distribution boxplot grouped by conference (Figure 5). This visualization reveals that the Big Ten, SEC, and Pac-12 possess the largest median enrollments, while conferences such as the MAC, Sun Belt, and C-USA consist primarily of smaller institutions. The presence of extremely large outlier universities, such as UCF, Penn State, and Texas A&M, creates pronounced right-skew in several Power Five conferences. These structural enrollment differences help explain parts of the attendance landscape: conferences with larger schools inherently have a broader fan and alumni base to draw from.

The violin plot in Figure 6 provides a clearer view of how enrollment is distributed across conferences and highlights structural differences that help explain variation in attendance. Smaller conferences such as the Sun Belt, MAC, and C-USA show tight, low-enrollment distributions, indicating that most of their member institutions operate at a much smaller scale. In contrast, Power Five leagues, especially the SEC, Pac-12, and Big Ten, display wide, upward-skewed distributions with dense clusters of mid-sized schools and distinct peaks at the very high end, reflecting the presence of massive flagship universities like Ohio State and Michigan. These patterns reinforce that conferences differ substantially in their institutional scale, and that programs in larger-enrollment conferences have inherently broader potential fan bases. While enrollment alone does not determine attendance, the violin plot shows that the structural size of conferences provides an important backdrop for understanding why some leagues consistently draw larger crowds than others.

Overall, the results demonstrate that enrollment has a meaningful, statistically significant relationship with home attendance. However, the scatterplot and inter-conference comparisons show that enrollment is only one piece of the puzzle. Fan engagement in college football is strongly intertwined with tradition, conference identity, geographic culture, and program brand strength. Enrollment provides the underlying scale, but history and culture determine how fully that scale converts into filled stadiums.
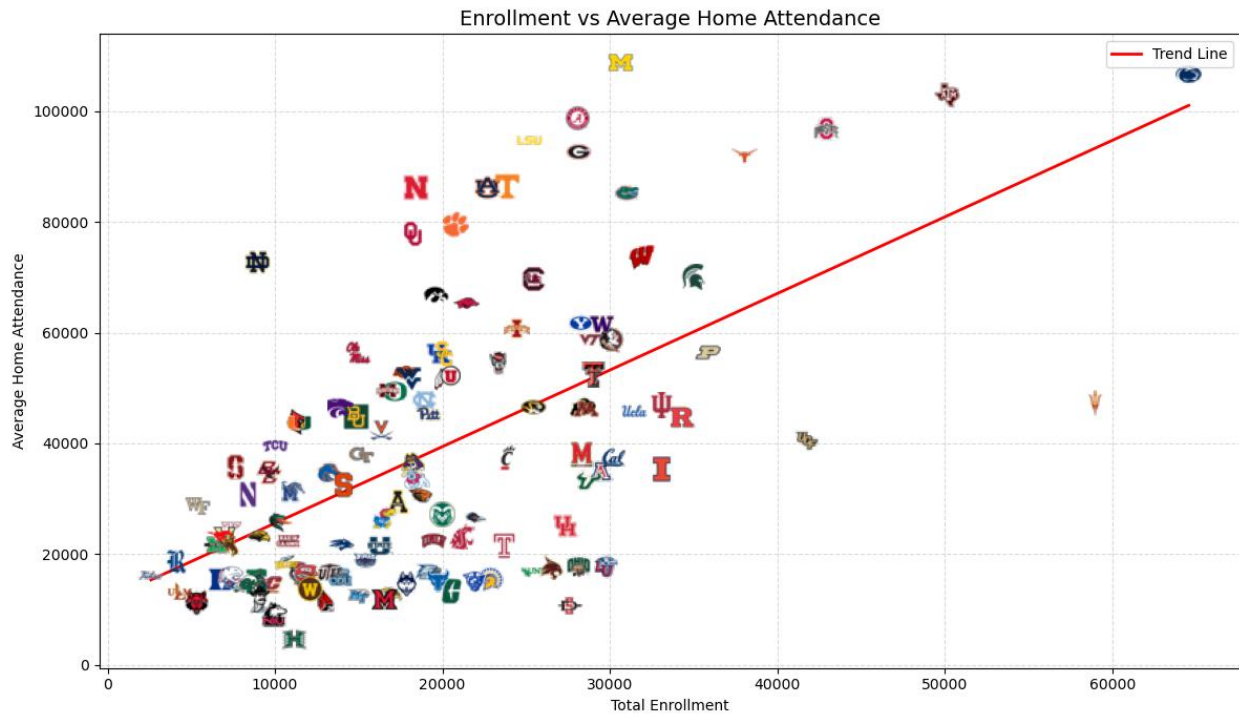
Figure 4 Scatterplot of Enrollment and Average Home Attendance with Trend Line
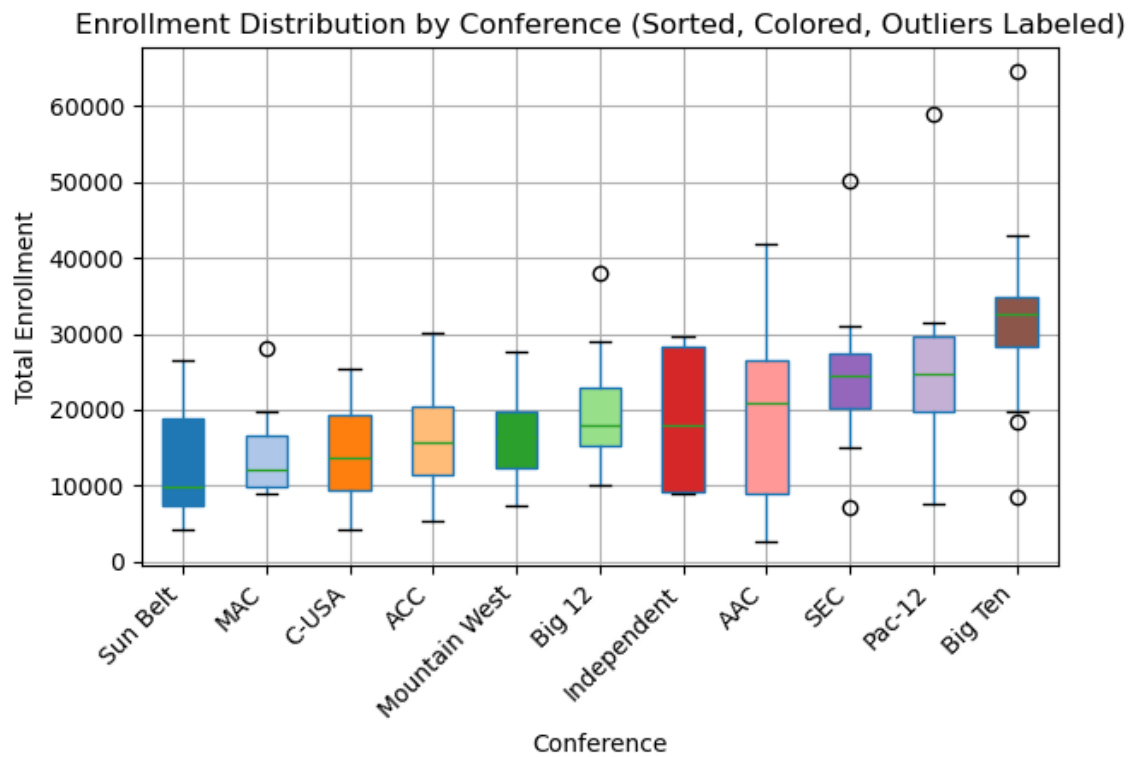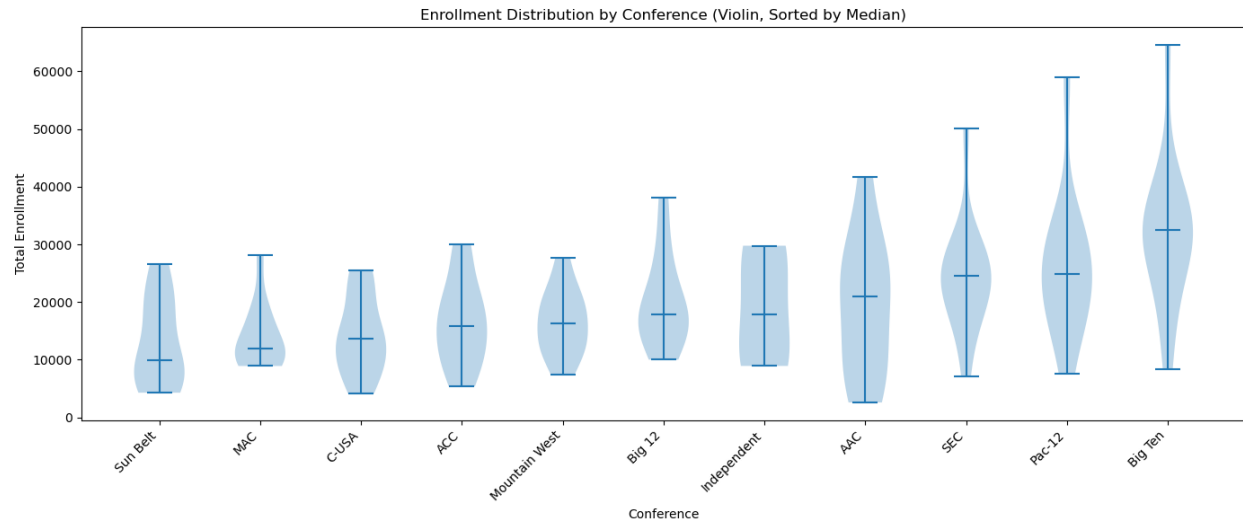

Figure 5 Boxplot of Enrollment by Conference

Enrollment Distribution by Conference (Violin, Sorted by Median)

Figure 6 Violin Plot of Total Enrollment

3.3 *Conference and Home Attendance*

My final research question examines whether certain conferences consistently lead in home football attendance and whether conference membership helps explain the variation in turnout that cannot be attributed to team success or enrollment alone. My hypothesis was Southern and Midwestern conferences (SEC, Big Ten) have higher mean attendance. The first step in assessing whether certain conferences consistently lead in home attendance is to compare average turnout across the FBS. As shown in Figure 7, the differences between conferences are substantial. The SEC once again sits at the top, averaging 72,437 fans per game, followed by the Big Ten at 64,785. These two leagues form the clear upper tier of college football demand. The Big 12 (54,447) and ACC (44,572) make up the next level, while the Pac-12 (43,364) trails the rest of the Power Five despite its large markets and historically competitive programs. The Independent category averages 35,721, though this value is heavily influenced by Notre Dame. Group of Five conferences occupy the lower end of the distribution: the AAC averages 27,895, the Mountain West 20,832, C-USA 18,547, the Sun Belt 18,089, and the MAC just 13,746. These updated averages show a clear and persistent hierarchy, reinforcing that conference affiliation remains one of the strongest predictors of home attendance in college football.

The pie chart of total season attendance (Figure 8) reinforces this pattern, demonstrating that the SEC and Big Ten alone account for over 40% of all FBS home attendance, despite representing a minority of total teams. This concentration of fan turnout aligns with broader cultural and historical factors: the SEC and Big Ten house many of the largest stadiums, most tradition-rich programs, and most football-oriented fan cultures in the country. Their dominance in both average and total attendance cannot be explained solely by enrollment,

success, or geography; it reflects deep, persistent conference-level differences in football identity. These conference-level effects were also anticipated in my project proposal, which hypothesized that southern and midwestern conferences would exhibit the strongest attendance patterns.

To further understand how conference membership shapes the relationship between enrollment and attendance, I calculated the correlation between these variables separately for each conference. The results reveal striking disparities. ACC (r = 0.769), SEC (r = 0.764), Big 12 (r = 0.715), and AAC (r = 0.699) show strong positive correlations, suggesting that larger schools in these conferences reliably draw larger crowds. In contrast, the Big Ten (r = 0.496) shows a much weaker relationship, driven by the fact that several of its most dominant attendance programs, such as Nebraska, Penn State, and Michigan, dramatically outperform what their enrollment alone would predict. At the lower end, Group of Five leagues such as the MAC (r = 0.362), Sun Belt (r = 0.224), Pac-12 (r = 0.203), and C-USA (r = –0.119) exhibit little to no relationship between size and turnout. Mountain West and Independent programs likewise show near-zero correlations, indicating that enrollment plays almost no role in shaping attendance within those conferences.

These findings are visualized in the faceted scatterplots (Figure 9), which illustrate how attendance behaves within each conference. The SEC and ACC panels display clear upward trends, reflecting strong, size-driven conference dynamics. The Big Ten, despite having some of the largest crowds in the country, shows scattered patterns, demonstrating that fan demand is rooted more in legacy and institutional football culture than in institutional size. Group of Five conferences exhibit tight clusters at low attendance levels regardless of enrollment, reinforcing that structural and cultural factors associated with conference identity exert far more influence than institutional characteristics. Overall, these visual and statistical results confirm that conference membership is one of the most powerful predictors of football attendance, shaping both the level and the structural pattern of fan engagement across the FBS.
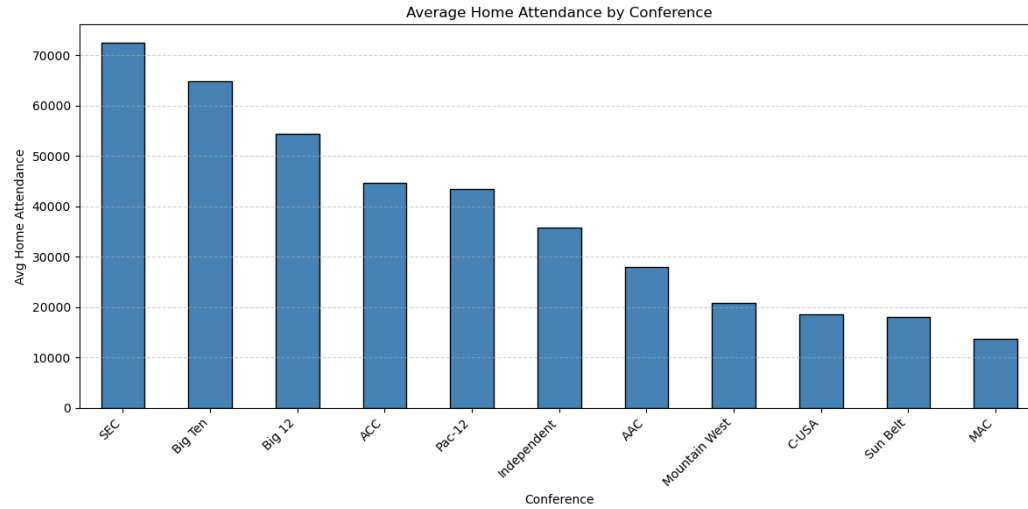
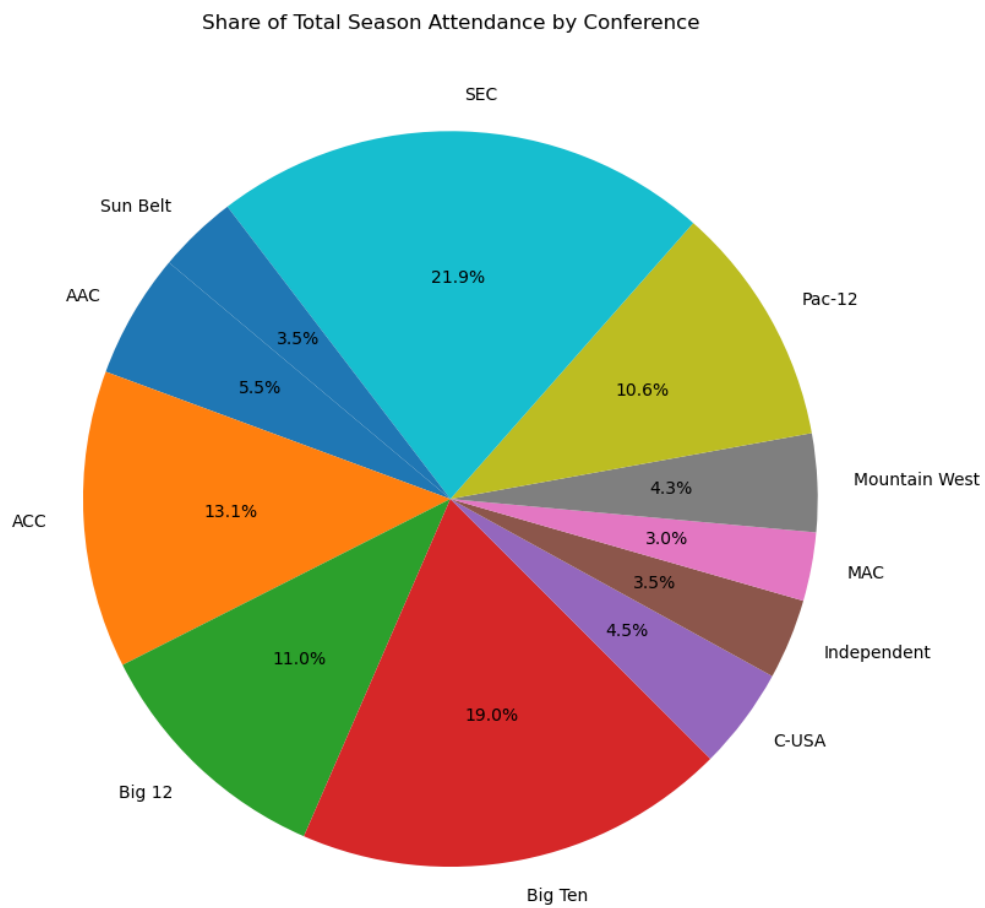*Figure 7 Bar Chart of Conference Average Home Attendance*

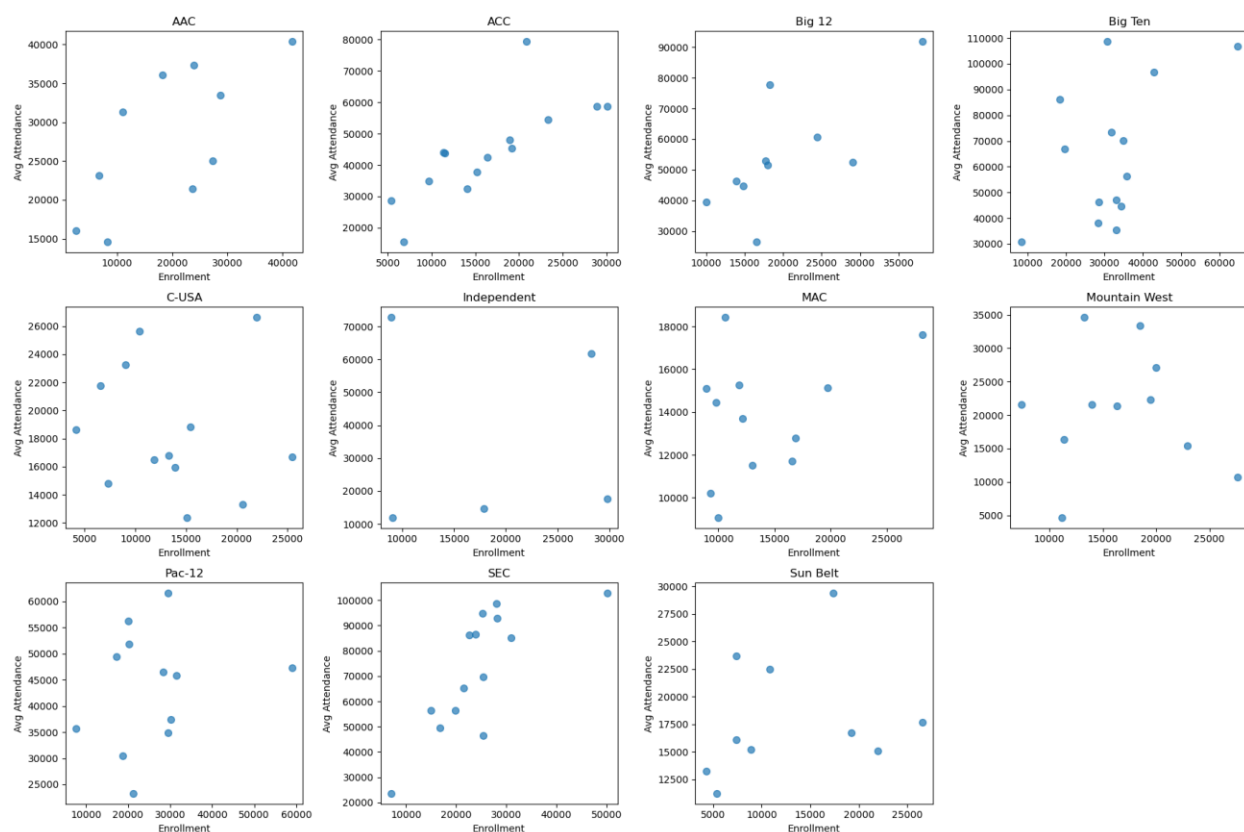*Figure 8 Pie Chart of Conference % of Total Football Attendance*

Figure 9 Scatterplots of Each Conference Enrollment to Average Attendance

4. Conclusion

In this project, I analyzed three factors that may influence average home football attendance: team success, institutional enrollment, and conference affiliation. In summary, the results of the three research questions from my proposal indicate the following:

1. Does team success drive higher attendance?
There is a statistically significant but moderately weak relationship between win percentage and home attendance (r = 0.346). Although winning does correlate with larger crowds, the effect is not strong enough to conclude that short-term performance is a primary driver of turnout. Several programs, such as Nebraska, Michigan, Ohio State, and Alabama, attract exceptionally high attendance regardless of seasonal outcomes, suggesting that long-term brand equity and football culture play a stronger role than yearly success.

2. Do larger universities draw bigger crowds?
Institutional size shows a meaningful relationship with home attendance. The correlation

between total undergraduate enrollment and average home attendance (r = 0.585, p < 0.001) indicates that larger schools generally host larger crowds. However, the relationship is far from deterministic. Power Five programs consistently outperform their enrollment-based expectations, while some large universities, such as UCF, Houston, and Arizona State, draw lower-than-expected crowds. Visualizations showed that conference context and football heritage interact with enrollment, making institutional size an important but incomplete predictor of attendance.

3. Do certain conferences consistently lead in attendance?

Conference affiliation emerged as one of the strongest predictors of home football attendance. The SEC (72,437) and Big Ten (64,785) form a clear top tier, followed by the Big 12 and ACC. Group of Five conferences average far lower attendance, with the MAC at just 13,746. The SEC and Big Ten together account for more than 40% of all FBS home attendance, demonstrating how deeply conference identity shapes fan turnout. Correlation patterns within conferences further highlight this effect: enrollment strongly predicts attendance in the ACC, SEC, Big 12, and AAC, but shows weak or negligible predictive value in the Big Ten, Pac-12, and all Group of Five conferences. These results show that structural and cultural differences across conferences, such as stadium size, football tradition, and regional fan culture, play an outsized role in determining crowd levels.

This project has several limitations, including its focus on a single season of attendance data, the exclusion of student ticketing policies and stadium capacities, and the inability to account for regional economic factors or variations in opponent quality. Future work could include incorporating multiyear attendance trends, modeling the effect of stadium capacity constraints, integrating ticket price or marketing data, or evaluating how conference realignment impacts fan engagement over time. Expanding the dataset would also make it possible to build predictive models that estimate expected attendance for programs based on institutional, geographic, and performance-related inputs.