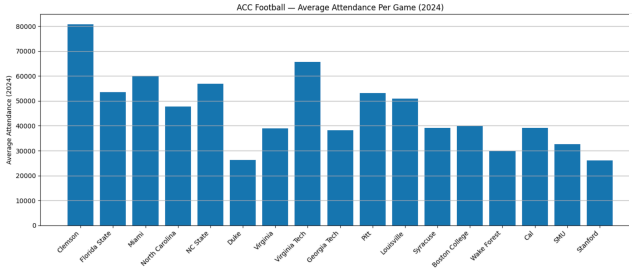
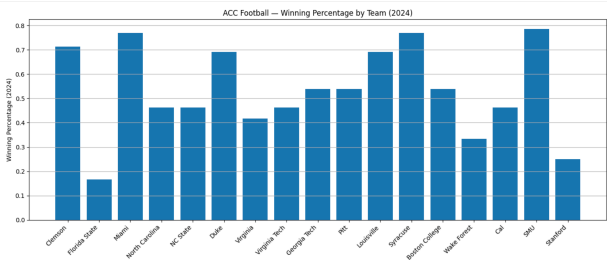
Introduction

College football attendance is routinely looked at as a way to gauge fan interest in a team. Coaches often want the stadium to be full of their fans as they see it as a way that provides a competitive advantage to them. We would likely expect there to be a positive relation between the average attendance for a team and their winning percentage because of this. I am focusing specifically on the ACC in this example because it is one of the power four conferences in college football. Do college football teams in the ACC who have a higher average attendance per game also have a higher winning percentage in 2024?

Data methods

I first wanted to know what my null hypothesis and alternative hypothesis would be. My null hypothesis does not believe that there is a relationship between average a attendance per game for an ACC team in 2024 and winning percentage. My alternative hypothesis believes the opposite. It does think there is a relationship between winning percentage and average attendance per game for an ACC team in 2024. I decided to first start to look at each team and their winning percentage before I ran a regression model. I wanted to see the variation between winning percentages and average attendance for the seventeen teams. What I found showed variability in both categories and specifically, two outliers in the winning percentage model with those being Florida State and Stanford.





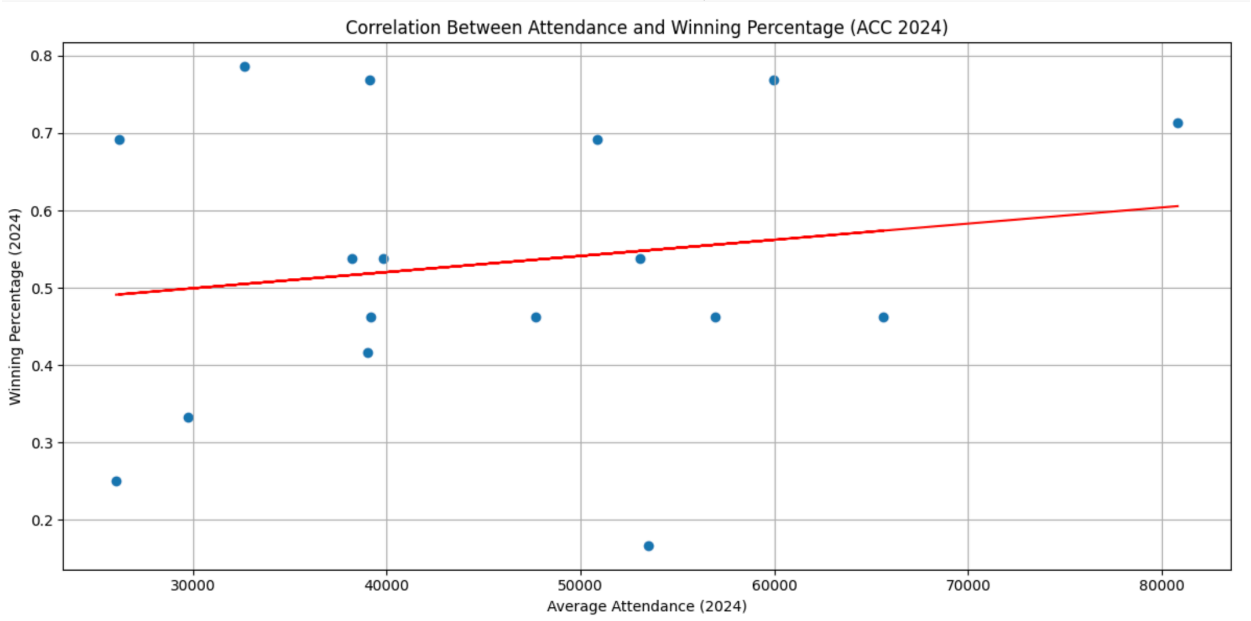
Without running a linear regression model, I can see some parts where there is a correlation. This can be specifically found in the first point, Clemson. Their attendance spikes as the highest and their winning percentage is among the best in the conference.

Statistical methods

For this research question, I decided that running a linear regression model would be best. This would also give me a line of best fit to see if there is correlation, if any, to my research question. This would also let me see visually how strong the p-value was if it was strong at all. The linear regression equation would be:

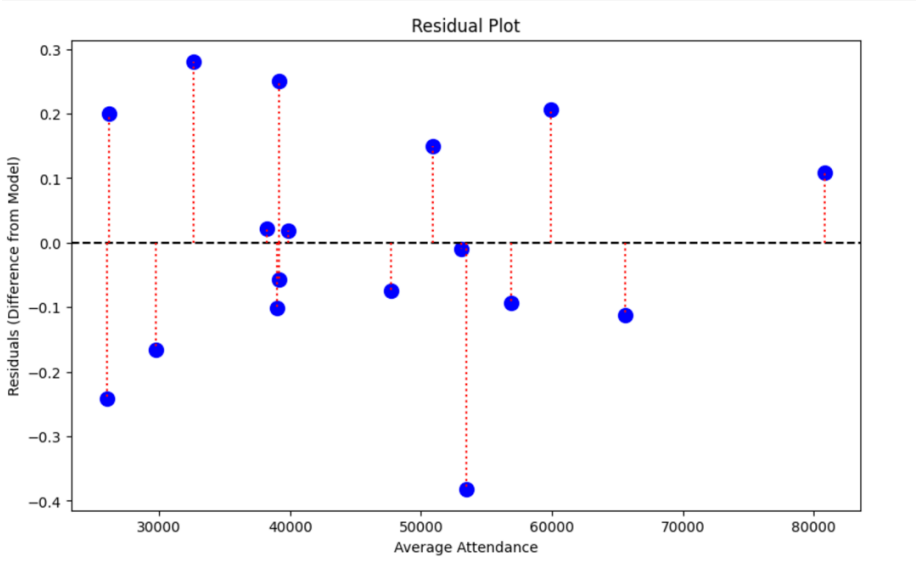
$$\text{WinningPercentage} = B_0 + B_1\ \text{AverageAtendance} + e$$

My next figure helps show the relationship between the two variables with the red line showing the regression. I decided that this would be best because it would show me that if there was a relationship, my data points would likely be scattered around the red line.



This allowed me to see if my model or research question was supported and if my alternative hypothesis believing that there would be a relationship between high average attendance and high winning percentage.

This next figure is a residual plot showing each team's relationship to the dotted line. It represents how overestimated or underestimated my model was. As you can see, there were a few points that very much disagreed with the relationship that y research question was attempting to prove.



Results

From looking at the data and my p-value being relatively high at 0.520, there is no definitive relationship between average attendance and winning percentage for a football team in the ACC. From looking at the linear regression model, there are only a few teams that may have some correlation between average attendance and winning percentage. When looking at my r-squared statistic, it provides a very low value of 0.0028. This means that on average, only about 3% of a team's average attendance can be attributed to their success or failure. My f-statistic is also low at 0.443 which shows that there is little to no variability in my data. The residual plot provided many differences in the dataset as well. From looking at my linear regression model in addition to my p-value, it provided me with a clear conclusion.

Conclusion

After running my data attempting to reject the null hypothesis that there is no correlation between a team's success and their average attendance in the ACC, I came to a conclusion that I could not reject the null hypothesis. Instead, I had to reject my alternative hypothesis. There is no clear relationship between fan turnout and success of a team. This can put many rumors to rest that having more fans at your games can significantly impact the results on the field.

Potential Shortcomings and Limitations

I would be remised if I did not discuss several shortcomings that I noticed along the way. For one, there are a few teams that logically cannot have a high number at fans per game, a team like SMU has a much smaller stadium than one like Clemson. Additionally, my data included both home and away attendance for each team, meaning that if a team played at a school with a smaller stadium (like SMU) that naturally, even if it was full, it could have an impact on their average attendance per game in 2024. Lastly, this only focused on one season. The data pool was relatively small. It would be a good idea to, if I ran this research question again, to include more than one season, and even branch out to including all power four conferences instead of only the ACC.

Sources

- https://www.d1ticker.com/2024-fbs-attendance-trends/
- https://theacc.com/standings.aspx?standings=114