Isabel Marshall

730610651

The Predictability of ACC Success

**Overview**

Andrew Sunberg’s “College Basketball Dataset” on Kaggle provides statistics on over 300 Division 1 men’s basketball teams from the 2013 to 2021 seasons. Given all this information, the question I posed was: Do stronger defensive or offensive factors seem to be more predictive of a team’s success in the ACC conference during the 2020 season? I chose to analyze the 2020 season, since there was no postseason due to COVID. This way, the numbers I analyzed would only be from the regular season and not skewed to that of postseason success. I also decided to use the ACC conference due to my constant favoritism for UNC basketball!

To begin with, it is important to define what “success” is considered during this season. In terms of this analysis, I chose to define success as the winning percentage of a team. All the comparisons made in the following analysis are being made between certain offensive and defensive factors compared with the winning percentage of a team throughout the season. In general, the strongest correlations between the “success” of a team were with their defensive factors. However, though there are some visual and numerical correlations evident in the data, they are weak. No complete conclusions can be made from these graphs alone, but they can provide some evidence into what is the most important for a team to work on to find success in a season.

**Data and Models**

Graph 1:

Chart, histogram

Description automatically generatedSince many of the comparison below will be used against the winning percentages, Graph 1 provides an overview of the winning percentages among the ACC teams in the year of 2020. Coming in at the top is Florida State University, winning 26 of its 31 games that season. Following them is Duke University, which won 25 of its 31 games. The shortest bars on the graph correlate to Boston College and Wake Forest, each winning 13 of their 32 and 31 games, respectively. While analyzing the following graphs, I will highlight the statistics of these four teams as examples of each of these factors.

Chart, scatter chart

Description automatically generatedGraph 2:

Sunberg provides the average number of points each team scored for and against per 100 possessions. In Graph 2, the direct comparison between the two values and the winning percentage is made. The average points scored for are provided as the green dots, while the average points scored against are in red.

Graph 3: Graph 4:

Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated

Graph 3 takes the data points for the average points scored per 100 possessions and isolates them into one graph. When put in a correlation test, the points demonstrated a correlation coefficient of 0.476. Florida State had an average of 111.2 points, Duke had 115.3, Boston College had 99.7, and Wake Forest had 108.8. Graph 4 holds the data of winning percentage vs. the average number of points scored against per 100 possessions. The correlation coefficient for the two variables in Graph 4 is -0.733. Florida State had an average of 92.5 points against, Duke had 91.9, Boston College had 99.2, and Wake Forest had 101.6.

Graph 5:

Chart, scatter chart

Description automatically generated

In Graph 5, the winning percentage is compared to the average field goal percentages scored for and against each of the ACC teams. In red is the average field goal percentage scored for, while in green is the average field goal percentage scored against.

Graph 6: Graph 7:

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

Similar to Graphs 2 and 3, Graphs 6 and 7 were created to isolate the variables presented in graph 5. For graph 6, winning percentage is compared to the average field goal rate. The correlation coefficient of these variables is 0.625. Florida State had a 51.9% field goal rate, Duke had 52.6%, Boston College had 46.5%, and Wake Forest had a 49.0%. The visual outlier that is below the slightly linear positive trend is UVA, who had a field goal rate of 46.9% while holding a winning percentage of 76.7%. For graph 7, the winning percentages were compared to the field goal rates allowed by each team. The correlation coefficient between these values was -0.779. Florida State had a field goal rate allowed of 47.0%, Duke had 45.7%, Boston College had 50.5%, and Wake Forest had 49.7%.

Graph 8: Graph 9:

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

Graphs 8 and 9 compare the variables from above, relating the points per 100 possessions and field goal rates. Graph 8 shows the correlation between the two offensive variables, demonstrating a correlation coefficient of 0.729. Graph 9 shows the correlation between two defensive variables, demonstrating a correlation coefficient of 0.900.

**Results**

I chose to compare winning percentages with points scored for and against per 100 possessions and field goal rates for and against to compare offensive and defensive statistics. In graphs 2, 3, and 4, the average number of points scored per 100 possessions for and against teams was used in comparison. The points scored for give an idea into the offensive tactics of the team. The sampled teams—Florida State, Duke, Boston College, and Wake Forest—offered supporting evidence that a more successful team is going to score more points more often. Duke and Florida State had the greatest number of points scored, while Boston College fell behind greatly. The correlation coefficient between these points is weak, being 0.476. This can explain why Wake Forest’s points scored for did not vary as greatly from the two greater teams even though its winning percentage falls lower than Boston College.

For graph 4, the four teams sampled and our correlation coefficient, -0.733, suggest that there is a stronger connection between points scored against and the success of a team. Duke and Florida State hold much lower average amounts of points scored per 100 possessions, being 91.9 and 92.5 points, than Boston College and Wake Forest, being 99.2 and 101.6 points.

Similar to the results from above, the graphs comparing winning percentages to the field goal rates for and against also provide a stronger correlation between scoring against than scored for. For example, the correlation coefficient between the field goal rate for and winning percentage is 0.625, while the field goal rate against and winning percentage coefficient is 0.779. This drastic difference is also demonstrated visually on the graph, as UVA being an outlier skews the least square regression line.

Graphs 8 and 9 demonstrate another comparison between offensive and defensive variables. The correlation between points scored against and field goal rate allowed had a high correlation coefficient of 0.900, offering evidence that these two factors work together with each other. While the offensive traits of points scored for and field goal rate offered a mediocre correlation coefficient of 0.729, it is weaker than the other.

**Conclusions**

The strongest correlations recorded in the data were between the defensive data and the success of a team. The correlation coefficient of the offensive factors did not hold strong enough value, meaning that the offensive performance in these two factors were not predictive of the success of that team. On the other hand, many of the defensive variables showed significant correlation to one another and to the success of the team that season. For this reason, I have concluded that defensive factors of a team in the ACC offer a better predictor to the success of that team.