PS6

christopher.p.danko-1

February 2020

1 Cleaning Data

To clean and transform the data, I dropped all values from the six columns I wanted to use in my analysis.

The transforms I made were differences between the team and the opposing team's season offensive rating, defensive rating, and pace rating. Since I got the dataset from someone who had used it to create a model already, the data was fairly clean.

2 Graphics

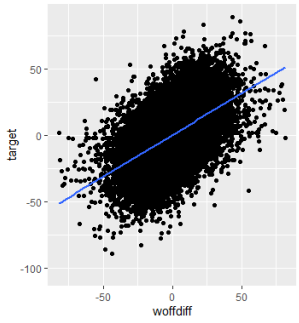

Figure 1: A simple scatterplot with a line of best fit for target score on offensive rating

This figure shows that there is a positive relationship between the difference in scores at the end of the game and the offensive rating of a team for a season. This is useful for identifying which factors influence winning-ness. This figure is valuable because it shows the negative relationship between defensive rating

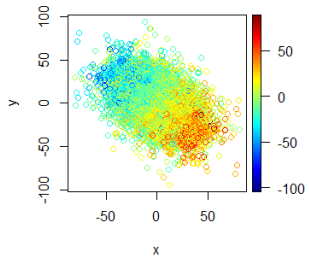

Figure 2: A heatmap scatterplot where the colors demonstrate the magnitude in difference of score, with difference in season offensive rating on the x-axis and difference in season defensive rating on the y-axis

and offensive rating, and that the trend in score difference is mostly driven by offensive rating

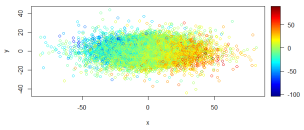

Figure 3: A heatmap scatterplot where the colors demonstrate the magnitude in difference of score, with season offensive rating difference on the x-axis and season pace difference on the y-axis

This figure is valuable because it demonstrates there is little to no correlation between pace and offensive rating. It continues to show that the difference in score is highly driven by offensive rating.