Michelle Lamas Nino

SAIL Application Paper

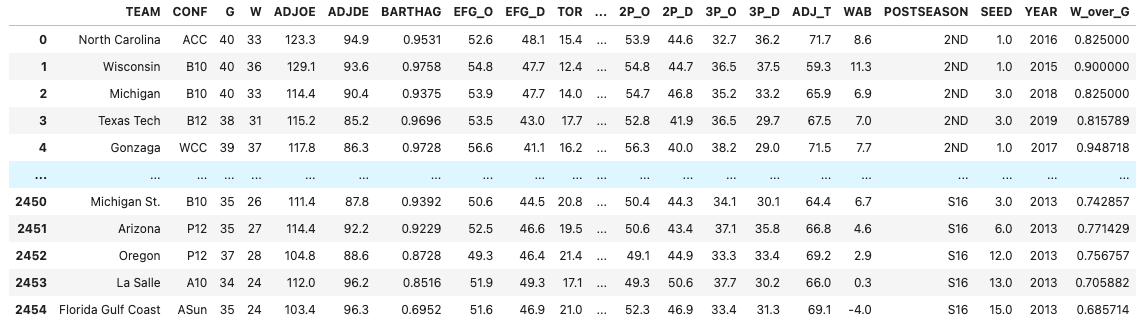
January 15th, 2023

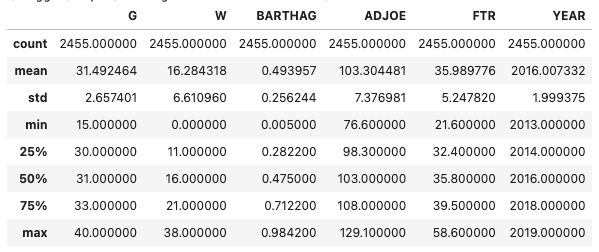
Indicators of Success:

Exploring Andrew Sundberg’s College Basketball Dataset

Andrew Sundberg’s College Basketball Dataset on Kaggle utilized data from 2013 to 2022 of Division I college basketball seasons. The dataset details 355 teams involved, the Athletic Conferences in which the schools participated in, the number of games played and won, Adjusted Offensive Efficiency, Power Rating, and more. Using this data, I will explore the relationship between multiple different indicators of success (Power Rating, Adjusted Offensive Efficiency, Free Throw Rate, and Turnover Percentage Allowed) and the number of games won in each season.

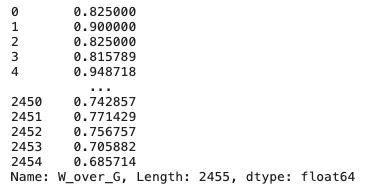
First, I will analyze the data and look for patterns. Below is the entire data set, including data on Adjusted Offensive Efficiency (ADJOE), Adjusted Defensive Efficiency (ADJDE), Power Rating (BARTHAG), Turnover Percentage Allowed (TOR), and more.

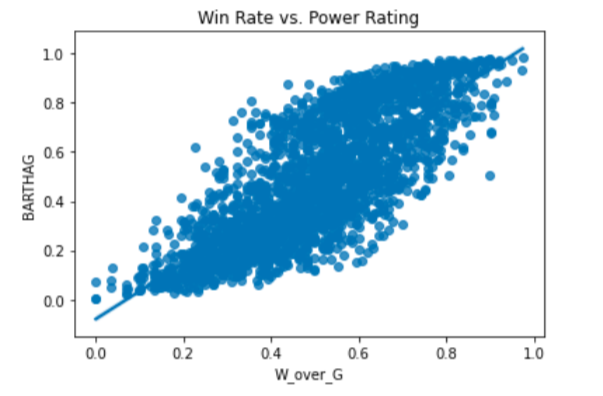
 Descriptive statistics of the data set concerning Number of Games Played (G), Number of Games Won (W), Power Rating (BARTHAG), Adjusted Offensive Efficiency (ADJOE), Free Throw Rate (FTR) are shown below:

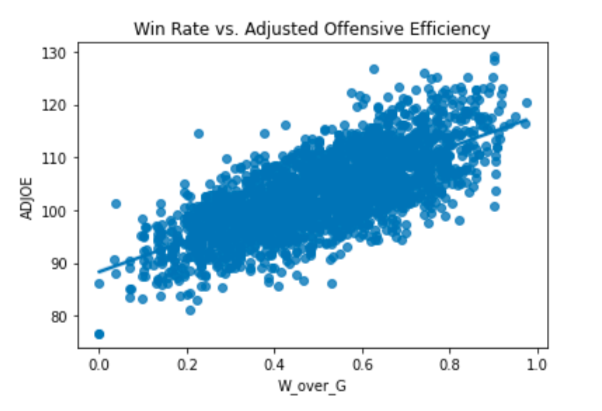


This shows that the mean number of games played is around 31 and the mean number of games won is around 16. In addition, the Power Rating, Adjusted Offensive Efficiency, and Free Throw Rate mean number is around .49, 103, and 36 respectively.

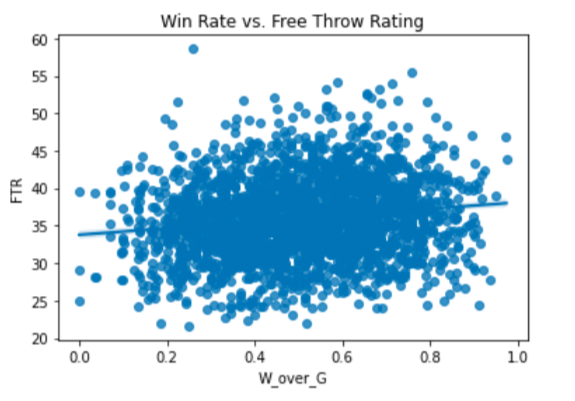
To gain a better understanding of the team’s success, success will be defined as having the highest win rate, or highest number of games won per games played. This variable is called W\_over\_G and is shown below:



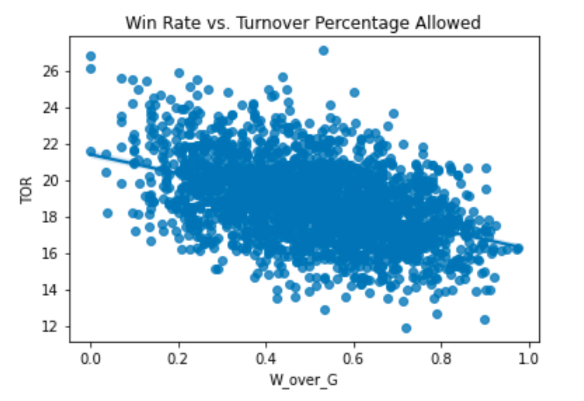
Scatterplots will be used to see the relationship between the win rate and different factors of success. First, Win Rate vs. Power Rating was plotted. This is shown below. 

Next, Win Rate vs. Adjusted Offensive Efficiency. 

Next, Win Rate vs. Free Throw Rating.



Finally, Win Rate vs. Turnover Percentage Allowed.



The plot of Win Rate vs. Power Rating shows a positive correlation between the two variables. Thus, this shows the correlation that as Power Rating increases, win rate increases. This relationship is in accordance with the definition of Power Rating, since it is a measure of a team's strength compared to other teams in the same category.

Next, the plot of Win Rate vs. Adjusted Offensive Efficiency shows a slightly lower positive correlation when compared to Win Rate vs. Power Rating. This high correlation is also understandable because it is an estimate of the offensive efficiency a team would have against the average Division I defense.

The third plot depicts a very low positive correlation between Win Rate and Free Throw Rating. In fact, the correlation is almost flat, which would indicate no correlation between win rate and how often a team shoots free throws.

Finally, the plot illustrating the relationship between Win Rate and Turnover Percentage Allowed (turnover rate) shows a negative correlation between the two variables. Thus, as win rate increases, the turnover rate decreases. Turnovers occur when a team loses possession to the opposing team before they can attempt to score, so this relationship is acceptable when compared to the definition of turnovers.

In conclusion, a team can use these results to decide which variables to focus on in their upcoming season. Since a high positive relationship is found between Adjusted Offensive Efficiency and Win Rate, it would be advisable to focus on improving a team’s offensive tactics to increase their chances of winning. However, since the frequency at which a team shoots free throws does not show any indication as to their chances of winning, a team should disregard this statistic as an indicator of their success.