

## **Using Multiple and stepwise regression to analyze the NBA data from 2022 to 2023.**

The study proposes to use 2022-2023 NBA data from Spotrac.com to assess the relationships among NBA characteristic variables such as salary caps, field goals, two-point and three-point shots, and win rates. The initial investigation will involve a correlation analysis to identify the strength and direction of the associations, where a correlation matrix and heatmap will be generated. Regression modeling will be the main analysis method. The target response variable will be win rates, and all other variables in the data set will be treated as explanatory variables. The regression analysis will be performed in two phases: multiple regression analysis with the entire data and stepwise regression. The multiple regression analysis will identify the impact of all factors in the entire data set. However, since some predictors may be highly correlated, including them in the model may lead to inaccurate predictions that may overestimate or underestimate the effect of some factors. Although the highly correlated predictors can be identified through the correlation analysis, choosing specific factors to omit in the final model may be difficult. As such, stepwise regression will be performed to identify the best model for predicting win rates in the NBA.

Ray Sun

Whapham, T.J.

Huangkun Chen