The Secret Formula

Locknar, Brandon

This project was designed to find the statistics from the 2013 Major League Baseball regular season that showed the most correlation to a team's ranking in wins. It was hypothesized that the Earned Run Average pitching statistic would show greatest correlation to wins. Data from the 2013 season was imported to Excel from a credible online statistical database. Regression, a statistical approach, was used to determine how the thirty teams rank in individual statistics to show correlation to their rank in wins and probability of the team to win the most games depending on statistical rank. The top ten teams and their rank for each statistic were analyzed to determine correlation. The difference between the sums of their rank in wins and their rank in each statistic was compared. Almost half of the tests showed relationships to wins; however, several statistics indicated a strong correlation to wins. Because it was difficult to determine correlation from all thirty teams' data, the statistics from only the top ten teams were used. When results showed ties, data from all thirty teams were considered to eliminate ties. Results displayed the On Base Percentage statistic showed the best correlation in batting statistics, while in pitching and overall statistics the On Base Plus Slugging Percentage showed the best correlation. However, the data did not support the hypothesis.

Awards Won: