**ECON 3990- Spring 2024-Final Report**



**Muriel Kroflin**

**Dr. Samantha Johnson**

**Abstract:**

The success of college soccer teams depends on various variables. This study investigates the influence of the variables, attendance, location as in home or away, corners, total number of yellow cards, yellow card comparison to opponent, shots, and goals. A dataset describing 95 soccer games was created using the University of Oklahoma (OU) stats website that offers game statistics from all official conference and non-conference games from 2018 to 2023. Throughout a comprehensive statistical analysis using Stata, variables that affect winning, scoring, and yellow cards were investigated. The findings reveal significant variables that influence winning such as corners, in conference play, and shots. The investigation concluded that while corners significantly impact winning, they are not a significant variable leading to goals. Moreover, the research revealed that attendance significantly affects OU having more yellow cards than the opponent. By understanding the relationship of these variables, one can develop a better understanding of college soccer, which provides valuable insights to coaches and players involved in the sport. Furthermore, the findings provide information that can be used for performance optimization and development of tactical strategies.

**Introduction:**

College soccer is a competitive landscape in which the variables that affect success are mainly unknown. Factors like player quality, player work rate, practice quality and frequency are certainly variables that play an important role. While these factors are assumed to be impactful, their significance is unknown. Empirical research uncovers concrete variables, their relationship, and their significance to success in college soccer. The significance of exploring these variables lies in the curiosity of current college soccer players. Coaches justify their tactical instruction with hypothesis like “corners will bring us success in this game, therefore we will train them a lot” or “we need to take many shots.” As college soccer players tend to question tactics wondering if that is the most efficient way to win a game, the question of whether these hypotheses are true arises. The structure of this paper aims to answer this question. More specifically the three hypotheses will be tested; 1. “Corners are a great chance to score goals” 2. “High attendance motivates players and leads to wins” 3. “Taking many shots leads to many goals.” Therefore, the introduction is followed by an outline of the methodology, describing the data set, and the statistical tools used. This is followed by a discussion of the results and a conclusion that highlights the key components of the findings and further research questions.

**Methodology**

The data source for the dataset is the OU soccer website. The stats section on the website includes easily measurable stats for both teams. The data set includes 95 official OU soccer games. Using this resource eight variables were defined. The first variable is called “win”. The data type is Boolean, indicating this is a yes or no variable. “Yes” is considered a win for OU, while “no” represents ties and losses. 37% of the games in the data set were won. The second variable “attendance” is measured in an absolute number of people in attendance. The average attendance across the data set is 711 people in attendance per game. Thirdly, home, or away was defines as a Boolean variable. Home is considered a game that was played at John Crain field in Norman. 43% of the games in the data set are home games. The total number of corners of OU per game is the fourth variable which averaged at 5 corners per game. The total number of yellow cards per game were on average 1.8 cards per game. Moreover, the variable more cards than opponents indicates whether OU received more cards than the opponent each game. Variables seven and eight are shots per game and goals per game. While shorts per game averaged at 13.6 the goal average per game is 1.4. For the analysis different regressions were run in Stata. The significance of any independent variable in reference to the dependent variable was tested in combination with other variables to test the significance in different scenarios. The data is limited to OU soccer games and therefore does not represent all college women’s soccer teams. All findings refer to the OU soccer team.

**Results:**

All results are based on the analysis of 95 OU soccer games from 2018-2023 using regressions at the 95% confidence level. As descriptive statistics the average per game of variable was recorded:

**A table with numbers and text

Description automatically generated**

Hypothesis 1, “Corners are a great chance to score goals” turned out to be false in this data set. Considering only the relationship between goals and corners, corners impact goals significant in a positive manner which is indicated by a positive coefficient of 0.129. Adding shots to the regression, the significance of corners regarding goals changes. Corners are no longer statistically significant and shots have a greater impact on goals and therefore are significant. This pattern repeats itself when adding more variables.

Considering all variables but more yellow cards than the opponent, it turns out that shots, home location, conference and wins are the only variables that have a significant impact on goals. While shots and home games positively impact goals scoring, in conference play has a negative impact on goals scoring. Corners remain insignificant in this regression, indicating that more corners do not lead to more goals.

Hypothesis 3, “Taking many shots leads to many goals” was proven to be true by this regression. Shots have a positive impact on goal scoring and are the most significant variable. While shots led to more goals, it is important to mention that on average 13.58 shots were taken per game only leading to 1.36 goals per game. These findings are represented in table 1.1 below.

A table with numbers and a list of goals

Description automatically generated

Hypothesis 2, “Playing at home is an advantage to winning games”, was rejected after running four regressions. Considering attendance and home location, playing at home is a most significant to winning. Adding conference play, home location remains significant to winning. Once game variables like shots and corners are added, home location becomes insignificant. While, home location is the most significant variable not directly connected to game play, it is insignificant when considering the actions in the game and therefore, does not impact winning. Overall, the home advantage is not significant as the variables of shots and corners have a stronger, significant impact on winning. These findings are represented in table 1.2 below.

A table with numbers and a number of numbers

Description automatically generated

Outside the scope of the hypothesis, the research revealed that attendance has a positive relationship with OU having more yellow cards than the opponent, but not with the total number of yellow cards. More specifically, this means that the higher attendance leads to OU having more yellow cards than the opponent. There are no other variables that have a significant impact on more yellow cards than the opponent.

A table with numbers and a few words

Description automatically generated with medium confidence

**Conclusion:**

This analysis highlights the importance of game analysis. With limited practice time, training should focus on variables that can be adjusted and that have an impact on the success of the team. The research suggest that corners do not lead to goals and therefore, offensive training should focus on shots which is the most significant variable for goal scoring. Furthermore, it turned out that only ever 24th shot leads to a goal, indication that there is a big development gap for accuracy that can be addressed in training. Finally, the home advantage turned out to be insignificant. Consequently, there is no need to focus on location in schedule planning and training. Moving forward, this research can be complemented by adding more in game variables to clearly define which strategies lead to success in college soccer. Examples for this are analysis substitution behavior as college soccer allows unlimited substitution. Coaches hypothesize that regular changes in the line-up keep players fresh and increase the chance of winning. Moreover, a general belief in college soccer is that heading long balls is essential to winning. There are many more coaching related hypotheses that can be tested. Further research should expand the data set using data for all NCAA D1 teams, to generate results that are applicable to the entire league.