

# A Statistical Analysis of UConn Field Hockey Penalty Corners: Evaluating Key Performance Factors

University of Connecticut – Sports Statistical Learning Internship

Julia Mazzola

2024-12-06

# 1. About Penalty Corners

A penalty corner is a strategic play that occurs when the defending team commits a foul within the attacking circle. The play consists of four defending players and the goalie. The attacking team has one inserter and around seven to eight attacking players on the top of the circle waiting to receive the ball. There are many offensive strategies that the attacking team can do making the play a very exciting and important one in the sport.

Penalty corners can often determine the outcome of matches as they offer a direct route to goal, so when a game is close being able to score on these opportunities is crucial. Moreover, much creativity can be used to design the tactics and strategies for executing the play. In high-stakes games, being able to perform and successfully score during this play can separate the good from the great teams. Understanding each aspect of the play and how a team performs can be a useful tool in evaluating team efficiency in utilizing these critical opportunities.

# 2. Experience

Throughout my time this semester, I was able to work closely with the University of Connecticut's Field Hockey Team. As one of the most successful sports programs at UConn, it was truly an amazing experience to contribute to their ongoing success. Some key responsibilities included assembling camera equipment for games to facilitate in-game APC review and coding UConn games, which aided in the collection and organization of key events throughout the game used in further video analysis.

During the 2024 season, the team achieved an impressive 17-4 overall record, secured the title of Big East Regular Season Champions, and went on to become the Big East Tournament Champions. Being part of this program allowed me to witness firsthand the dedication that drives their success, and it was an incredible opportunity to contribute to such a driven and accomplished group.

# 3. Penalty Corner Data Overview

This report focuses solely on the University of Connecticut Field Hockey's attacking penalty corner (APC) plays. It does not look at any defensive or scout data from the team. The data was taken from games in the Fall 2024 Season, dated from August 30th to October 18th. Further, the data consisted of 87 penalty corner plays in which 19 resulted in a goal. This averages to be about a 22% scoring rate for the team during these plays.

## Data Overall

Data Overall	
Total Penalty Corners:	87
Total Goals:	19
Scoring Rate:	21.84%

The data consists of multiple variables to assess UConn's performance on their attacking penalty corners. Our observations collected were start time, end time, duration, game status (who was winning/losing/drawing), quarter, score, trap, castle location, call, shot location, and pc outcome. For further analysis other variables were engineered to create a better understanding of the data. These consisted of variables such as opponent, over 15 seconds, is in top 25, one point difference, and Connecticut status. This report will analyze the different aspects of the play to assess UConn's performance throughout their 2024 Season.

## Variables Created for Analysis

Created Variable	Definition
opponent	The defending team UConn is trying to score against

Created Variable	Definition
over_15_sec	Whether the duration of the play lasted over 15 seconds
is_top_25	If the defending team is ranked in the top 25
one_point_diff	If there is a one point difference in score at the time of the play
connecticut_status	If UConn is winning, losing, drawing at the time of the play

Table 3.1: Created Variables and Definitions

## 4. Breakdown by Quarter

To look further into the timing of the APC's, the data was broken down by quarter. This breakdown provides insights into which quarters saw more attacking penalty corners and goals, offering a clearer picture of the team's performance at different stages of the match.

### Breakdown of APC Scoring Rate by Quarter

Quarter	Total PC	Goals	Scoring Rate by Quarter (%)	Quarter PC Share (%)
Q1	17	5	29.4	19.5
Q2	23	5	21.7	26.4
Q3	23	5	21.7	26.4
Q4	24	4	16.7	27.5

Table 4.1: APC Performance by Quarter

### APC by Quarter

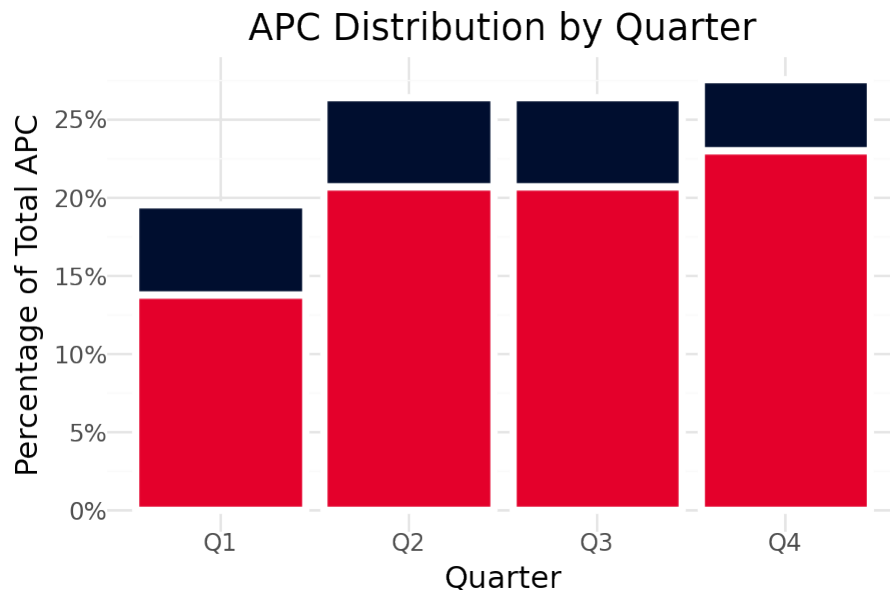


Table 4.2: APC Performance by Outcome and Quarter

Each quarter had around the same distribution of corners with slightly less occurring in Quarter 1. The distribution of goals for each quarter was around 5% of total corners which stays consistent with the 22% scoring percentage found earlier. The best performing quarter is Quarter 1 with a scoring rate of 29%, while accounting for only 19.5% of total

penalty corners. The worst performing quarter was Quarter 4 with a scoring rate of 16.7%, while accounting for the largest share of plays in the dataset of 27.5%.

## 5. APC Duration

To explore whether the duration of APCs impacts scoring rates, it is useful to analyze the average duration of APCs across different quarters. This analysis provides insight into how the timing of these plays may influence their success and overall effectiveness in the game.

Average duration of all penalty corners: **13.55** seconds

### Average Duration by Quarter

Quarter	Avg Duration (sec)	Quarter Scoring Rate (%)
Quarter 1	12.47	29.4
Quarter 2	14.73	21.7
Quarter 3	13.07	21.7
Quarter 4	13.64	16.7

*Table 5.1: Average Duration by Quarter*

The duration of UConn's APCs remained relatively consistent across the quarters, with some variability. On average, the duration of these plays was approximately 13.5 seconds. Quarter 1 saw the fastest APCs and the highest scoring rate, while the subsequent quarters (2-4) showed no clear correlation between duration and scoring rate.

## 6. Understanding APC Behavior When it is a Close Game

The intensity of game can change significantly depending on the score and the pressure of another team. Exploring whether the pressure of a close game impacts the scoring rate can be interesting to understanding how teams adjust their strategies under different game conditions. By analyzing the proportion of goals scored in these situations, we aim to answer whether the scoring behavior differs in close games compared to those with a more substantial lead. A z-test was conducted to determine if there is any statistical significance in the difference in goal proportions between these two scenarios.

Proportion of goals with one-point difference: **0.14**

Proportion of goals without one-point difference: **0.25**

Out of all the penalty corners that occurred when there was a one-point difference, 14% of them resulted in a goal and when there was no one-point difference, 25% of them resulted in a goal.

To see if there is actual statistical significance between if a close game impact scoring a z-test was conducted.

### Hypothesis:

Null Hypothesis: There is no difference in goal likelihood when it is a close game.

Alternative Hypothesis: There is a difference in goal likelihood when it is a close game.

Based on the z-statistic test conducted, with a p-value = 0.2401, we fail to reject the null hypothesis at the 5% significance level. This indicates there is insufficient evidence to suggest a significant difference in the likelihood of scoring a goal when the score is close (one-point difference) compared to when it is not close (greater than one-point difference).

For further review of these instances that resulted in "no goal" a list was constructed:

### List of PC's Where There was No Goal and a Close Game

Start Time	Opponent	Connecticut Status	Start Time	Opponent	Connecticut Status
03:27.4	New Hampshire	Winning	16:45.2	Yale	Winning
03:36.2	New Hampshire	Winning	19:23.8	Georgetown	Winning
05:41.4	Cornell	Winning	19:43.9	Georgetown	Winning
06:00.4	Cornell	Winning	25:47.4	Villanova	Winning
09:53.9	Brown	Winning	27:12.5	UAlbany	Winning
12:37.9	Harvard	Losing	28:15.0	UAlbany	Winning
13:55.0	Harvard	Losing	28:44.0	UAlbany	Winning
15:11.6	Temple	Winning	29:11.2	UAlbany	Winning
15:26.5	Temple	Winning	31:55.5	Providence	Winning
15:53.6	Temple	Winning	32:12.0	Providence	Winning
16:29.5	Yale	Winning	32:23.0	Providence	Winning

Table 6.1: APC Outcomes Resulting in 'No Goal' and Close Game

## 7. Hit Type and Success

To identify the most effective hit types, scoring rates were calculated for the top five shot types. The scoring rate for each type was determined by dividing the number of goals scored using that type by the total number of attempts of the same type. Based off of the top five most common hit types the three with the highest scoring rates were hit with 29%, flick with 30%, and pass/slip right with 20%. The least effective hit type was sweep with only 8% being successful.

### Scoring Rate by Hit Type

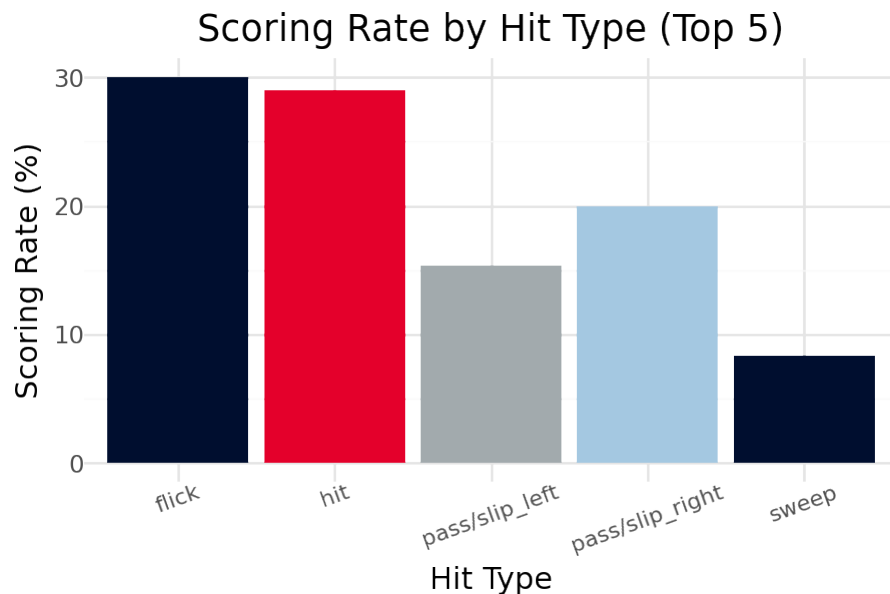


Table 7.1: Scoring Rate by Hit Type (Top 5)

### All Hit Type and Scoring Rate

Hit Type	Total APCs	Total Goals	Scoring Rate (%)
Back to Inserter	3	2	66.67
Flick	10	3	30.00
Hit	31	9	29.03
Pass/Slip Right	5	1	20.00
Pass/Slip Left	13	2	15.38
Error	7	1	14.29
Sweep	12	1	8.33

Table 7.2: All Hit Types and Score Rate

## 8. Stopper Location

During the play when the inserter passes the ball to the stopper, several zones, referred to as “castles,” were recorded in the data. These zones include Short (0.5), Castle 1 (1.0), a Middle Castle (1.5), and Castle 2 (2.0). The positioning of the stopper plays a critical role in the timing and execution of the play. Analyzing the most effective locations for the stopper can provide valuable insights into optimizing the penalty corner strategy.

### Stopper Location and Scoring Rate (%)

Stopper Zone	Total Shots	Goals	Scoring Rate (%)
0.5	6	1	16.67
1.0	21	5	23.81
1.5	12	1	8.33
2.0	48	12	25.00

Table 8.1: Stopper Location and Score Rate (%)

### Scoring Rate by Stopper Zone

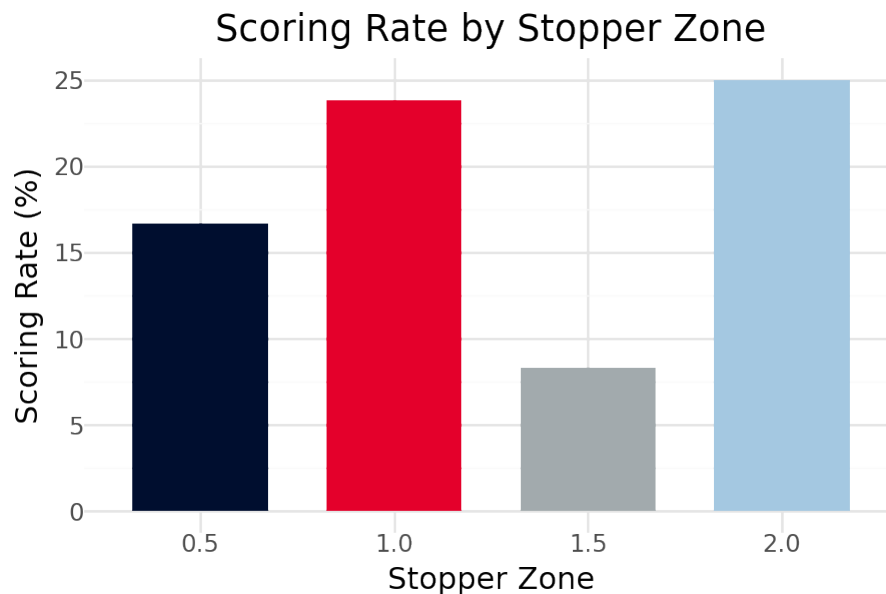


Table 8.2: Scoring Rate by Stopper Zone

The zones with the most success are Castle 1 and Castle 2 as they have scoring rates of 24% and 25%. Zone 1.5 was the least successful as only 8% of PC's inserted to that stopper location resulted in a goal.

## 9. Play Combination Analysis

Finding the most successful combinations for an APC can help distinguish which strategies are most effective. Based on the frequency of occurrences, these are the top-performing and least-effective play combinations. However, due to the limited size of the dataset, these results may not serve as definitive predictors for play strategy. Nevertheless, they offer valuable insights into the factors that may influence the outcome of a penalty corner, providing a starting point for further analysis and refinement in play planning.

### Most Successful Combinations:

Count	Quarter	Trap	Zone	Status	Hit Type
3	Q1	T2	2.0	drawing	hit
2	Q2	T2	2.0	winning	hit
2	Q4	T2	2.0	winning	flick

Table 9.1: Top 3 Most Successful Play Combinations by Goal Count

The most successful combinations occurred two to three times each in the dataset. The largest differences between the successful plays were the Quarter and Hit type used.

### Least Successful Combinations:

Count	Quarter	Trap	Zone	Status	Hit Type
5	Q2	T2	2.0	winning	hit
3	Q1	T1	1.0	drawing	pass/slip left
3	Q1	T2	2.0	drawing	hit

Table 9.2: Top 3 Least Successful Play Combinations by Count

The least successful combinations occurred three to five times in the dataset.

## 10. Performance Against NCAA Top 25 Ranked Teams

Understanding team performance against top-ranked opponents is essential for evaluating a team's ability to compete at the highest level. While team rank alone does not fully capture a team's skill or overall ability, it can offer valuable insight into how UConn performs under more competitive conditions. By comparing UConn's performance in penalty corners against NCAA Top 25 ranked teams to games against non-ranked teams, we can assess how the team's strategy and execution evolve when facing some of the nation's best competitors.

### Performance by NCAA Top 25 Status:

Top 25 Status	Total APC	Total Goals	Scoring Rate (%)
No	60	15	25.00
Yes	27	4	14.81

Table 10.1: APC performance against Top 25 Ranked Opponents:

### Breakdown by Opponent:

Opponent	Top 25 Status	Total APCs	Total Goals	Scoring Rate (%)
Brown	No	4	1	25.00
Cornell	No	10	1	10.00
<b>Delaware</b>	Yes	11	0	0.00
Georgetown	No	10	4	40.00
<b>Harvard</b>	Yes	3	0	0.00
New Hampshire	No	6	2	33.33
Providence	No	13	2	15.38
<b>Rutgers</b>	Yes	2	2	100.00
<b>Saint Joseph's</b>	Yes	2	1	50.00
Temple	No	5	0	0.00
<b>UAlbany</b>	Yes	9	1	11.11
Villanova	No	6	3	50.00
Yale	No	6	2	33.33

Note: **Bold teams** indicate a 2024 NCAA Top 25 Ranked Team

Table 10.2: APC performance based on Opponent:

On average the team performed better against teams not in the Top 25. This could potentially be due to teams not in the Top 25 committing more fouls against Connecticut resulting in more APC opportunities for Connecticut to score. Further, only 38% of games played were against Top 25 Teams. To understand this further, we will conduct a chi-squared test to determine if there is any statistical evidence to suggest a significant difference in the APC outcome based on whether the opponent is a Top 25 team or not.

### Hypothesis:

Null Hypothesis: There is no difference the “goal” outcomes between opponents who are in the Top 25 and those not in the Top 25.

Alternative Hypothesis: There is a difference the “goal” outcomes between opponents who are in the Top 25 and those not in the Top 25.

Chi-Square Test Output

Chi-square statistic: 0.61

P-value: 0.4334

Because the p-value is large at 0.6432, we fail to reject the null hypothesis at a 95% significance level and conclude that there is no significant statistical evidence that playing a Top 25 ranked opponent changes the outcome of the APC.

## 11. Scoring based on Connecticut Status

Further, other different circumstances that UConn experiences is the score of the game. Breaking down the status into winning, losing, and drawing we can understand how the team performs for each circumstance.

### Scoring Rate by Connecticut Status

Status	Scoring Rate (%)	Count
Drawing	20.59	34
Losing	0.00	2
Winning	23.53	51



*Table 11.1: Connecticut Status and Score Rate*

This table does not provide the most comprehensive insight, as UConn’s performance when drawing and winning is relatively similar, with a slightly higher scoring rate when winning (23.53% vs. 20.59%). The losing status has a very limited data sample, with only 2 observations, making it difficult to draw any meaningful conclusions from this status.

## 12. APC Outcome Prediction Model

To predict the outcome of an attacking penalty corner (‘goal’ or ‘no goal’), two models were constructed: Random Forest and Extreme Gradient Boosting (XGBoost). These models were chosen over regression-based approaches due to the categorical nature of the data. The final model selected was XGBoost due to its more accurate performance compared to the Random Forest Model.

### Extreme Gradient Boosting (XGBoost)

XGBoost Evaluation:

Test Accuracy: 0.8333

Precision: 0.2500

Recall: 1.0000

F1-Score: 0.4000

Mean CV accuracy: 0.7862

#### Model Evaluation:

After running the Extreme Gradient Boosting (XGBoost) model, several metrics were calculated to assess its performance. The test accuracy, which measures how often the model correctly predicts an outcome, was 83%. Additionally, the cross-validation accuracy, which evaluates the model’s reliability on different subsets of the data, was 79%. These results suggest that the model performs reasonably well on new, unseen data. However, since the test accuracy is slightly higher than the cross-validation accuracy, there is a possibility that the model may be overfitting to the test data.

While accuracy gives a general sense of performance, it doesn’t reveal how well the model handles specific outcomes—in this case, whether a penalty corner results in a “goal” or “no goal.” For instance, the model’s precision for predicting “goal” was 25%, meaning it correctly identified only 25% of actual goals. On the other hand, its recall was 100%, indicating that the model successfully captured all true goals. However, this perfect recall might signal overfitting, where the model memorizes patterns in the training data but struggles with generalization. Factors such as a small sample size could contribute to this behavior.

Finally, the F1-Score, which balances precision and recall, was 0.4. This score highlights that while the model excels at identifying all actual goals (recall), its low precision significantly impacts overall performance.

## Confusion Matrix

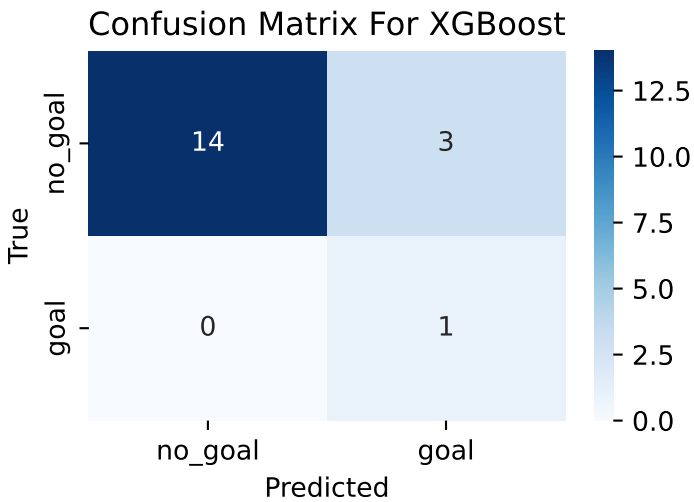


Table 12.1: XGBoost Confusion Table

The confusion matrix is a visual tool to understand how the model is performing predicting each class (“goal” or “no goal”). Here the model correctly predicted “no goal” fourteen times and “goal” one time. However, the model had three false positives and zero false negatives.

## XGBoost Model Feature Importance

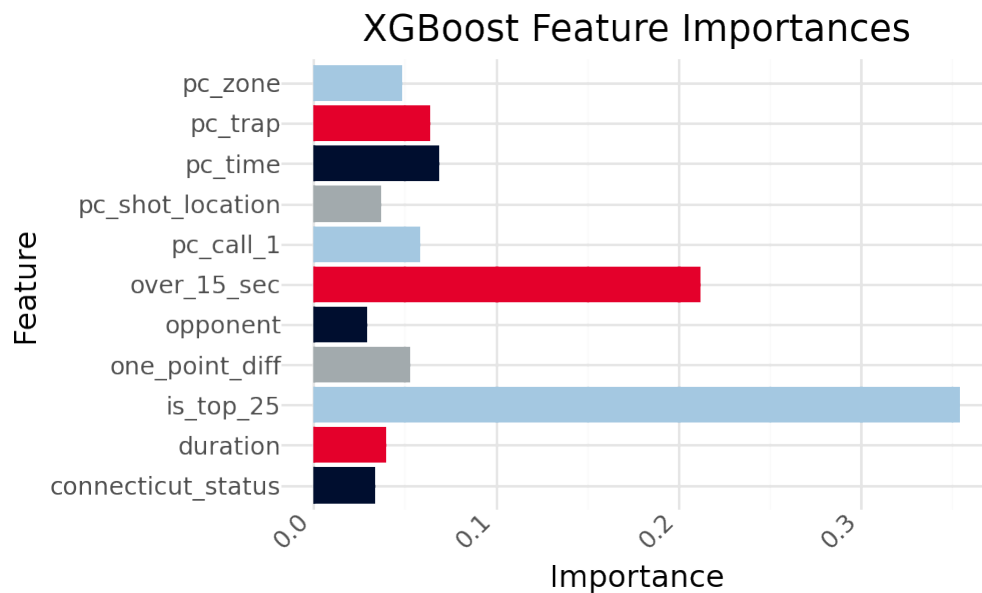


Table 12.2: XGBoost Feature Importance

The features that were most important for the XGBoost classifier were the variables `is_top_25`, `over_15_sec`, and `pc_time`. Those were the data points that carried the highest significance for the model to make its decision.

While this model was not the most effective in predicting the outcome of a penalty corner, it provided valuable insights into the factors influencing the model’s decisions. Additionally, it helped us better understand the key elements that contributed to its predictions. Potentially adding more observations or numerical variables could be influential in helping the model make more accurate and precise predictions.

### Predicted Probability of Goal by Duration

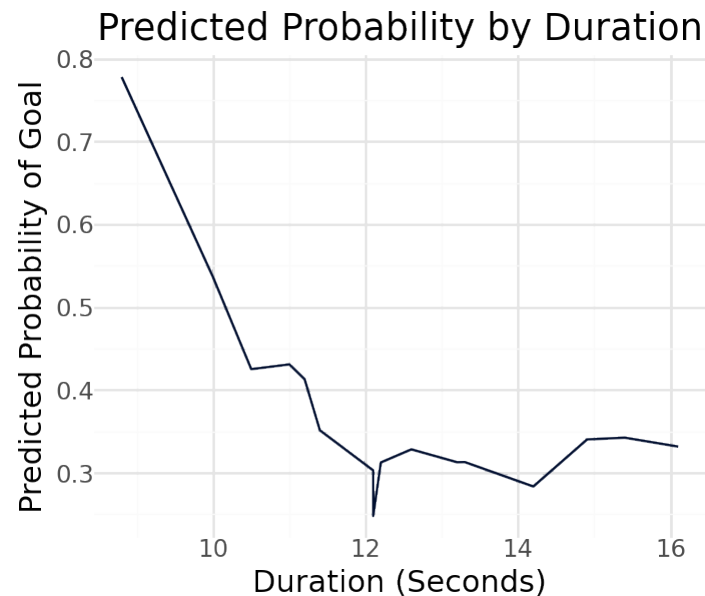


Table 12.3: Predicted Probability of Goal by Duration

Based off of the model created the predicted probability of a goal decreases as the duration of the play goes on. Then when the play hits around 12+ seconds the predicted probability levels out at around 30%.

## 13. Conclusion

This analysis of UConn Field Hockey's performance in attacking penalty corners hopes to provide useful insights into how various factors, such as game status, opponent ranking, and stopper location, impact their success. By examining these elements, we try to better understand the dynamics of penalty corners under different game circumstances. These findings contribute to a deeper understanding of the team's performance and can inform future strategies to optimize penalty corner execution, further enhancing the team's overall approach.

## 14. Acknowledgments

I am very grateful for UConn Field Hockey Head Coach Paul Caddy, and the rest of the coaching staff and team. My time with the team has truly been an incredible experience, allowing me to learn so much about the sport and grow both personally and professionally. It was an honor to be part of such a successful and driven team, and I'm so grateful for the opportunity to contribute to this amazing program.

I would also like to express my sincere thanks to Dr. Jun Yan and Nan Zhang for their guidance and support throughout this process. Their expertise was invaluable in providing me with the statistical tools necessary to complete this report, making this a truly rewarding learning experience.