

# **National Hockey League (NHL) Shots**

Grant Allvin, Melos Bekaj, Lay Len Ching, Silas Wang  
36-315: Statistical Graphics and Visualization  
Spring 2024

# Introduction

- The NHL presents a unique opportunity for sports analytics, especially with the Stanley Cup occurring
- We were interested in analyzing a collection of shots from the 2021-22 season to find several ways to optimize team performance
- 160,573 shots analyzed with 22 data points



# Research Questions

- 1) What is the impact of shot location and shot angle on frequency of shots made, and how do these impact the likelihood of making a successful shot?
- 2) How do NHL teams differ in their ability to convert shots into goals, and how does this variation in goal conversion rates correlate with shot accuracy and effectiveness across different game periods?
- 3) What is the effect, if any, of the “home-field advantage” on scoring?

# Research Question 1

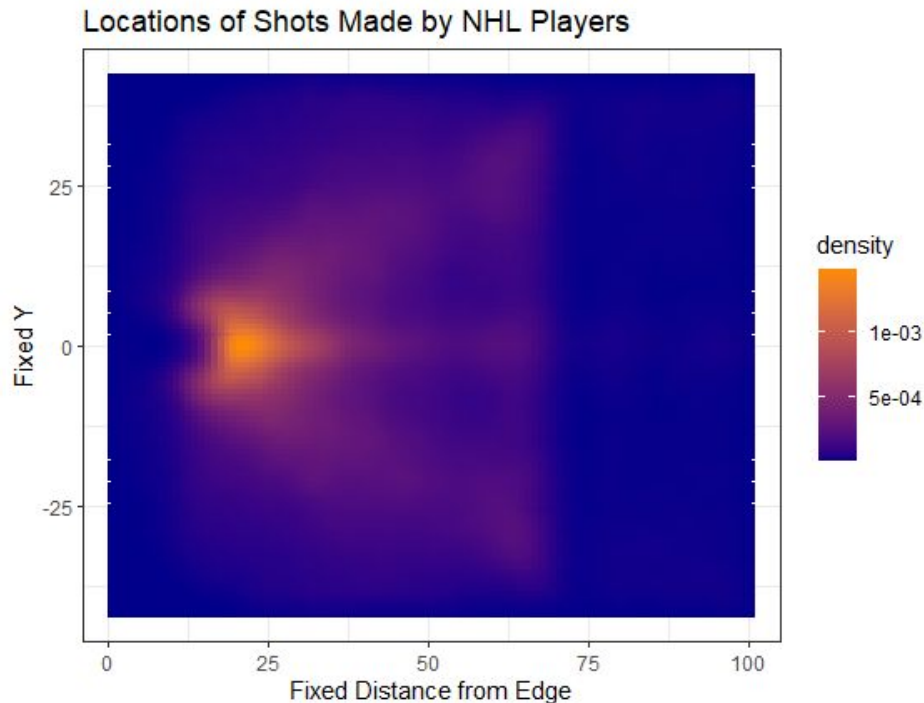
Where on the ice are goals more likely to be scored? How does the likelihood of scoring goals correlate with shot location and shot angle?

# Shot Frequency by Location

- Make new variable measuring the distance from the nearest edge to get insights for all teams in general rather than just home or away

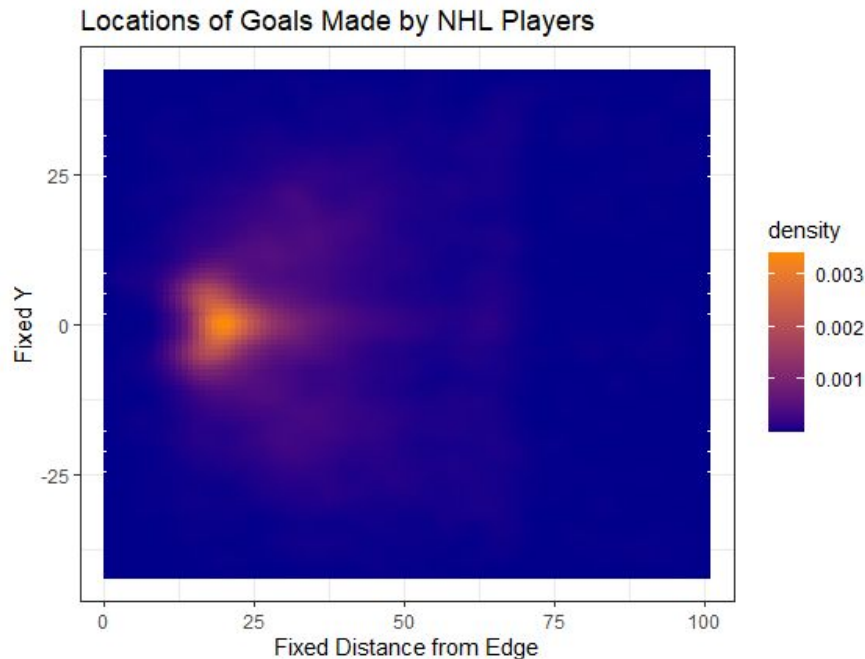
$$x\_from\_edge = 100 - |x\_fixed|$$

- Main insights:
  - Shots made most often close to the net
  - Noticeable frequency directly in front of the goal or at a 30-degree angle
  - Shots seem to cut off at ~70 units away from the edge
- Gives us a good sense of typical player behavior, may be result of training with best practices
- Preliminary data exploration before we approach our questions



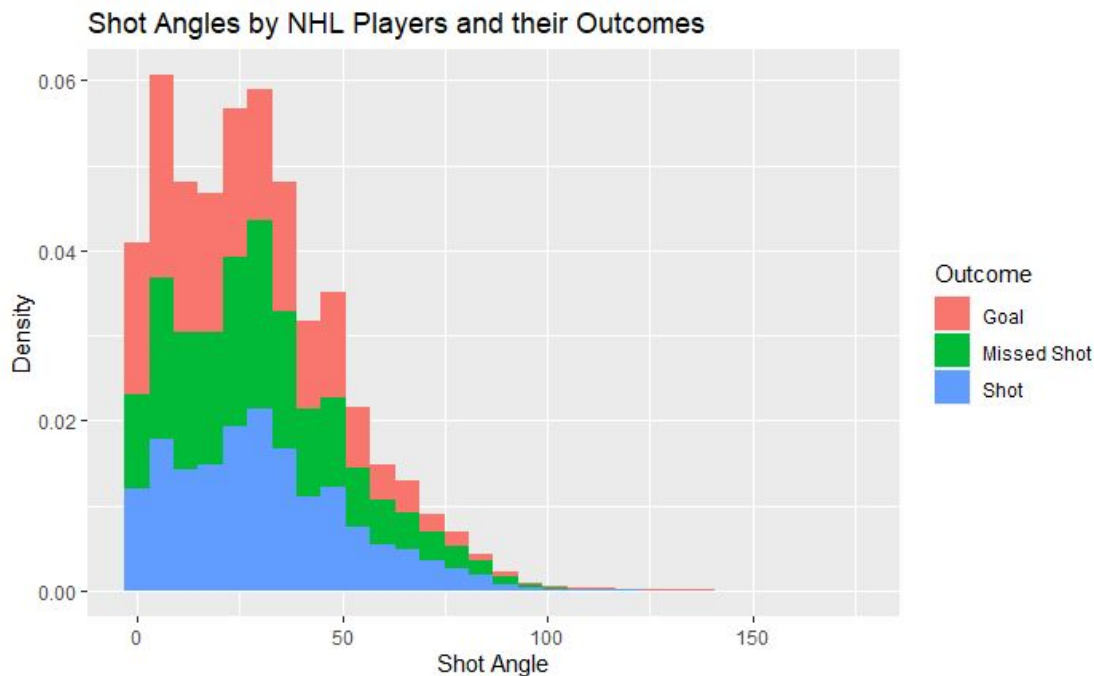
# Goal Frequency by Location

- Same premise as before, but only for goals
- Main insights:
  - Exhibits similar patterns to shot frequency
    - Most goals made near the net
    - Some goals made in front or 30 degrees away from the net
  - The cutoff happens at around 40 units
- Also gives a sense of player behavior and insights into best practices



# Shot Angle Frequency and Outcomes

- Gives a good sense of the marginal distribution of shot angles and outcomes, as well as conditional probability of outcome given shot angle
- Main insights:
  - Most shots tend to be “straight shots” (<50 degrees)
  - Conditional distribution of outcome given angle about constant
  - More goals made with straighter shots



# Regression Analysis: Goal Likelihood vs. Shot Location and Angle

- Logistic regression model with the following formula:

*goal ~ x\_from\_edge + abs(y\_fixed) + shot\_angle*

- Assume “x\_from\_edge” is analogous to horizontal distance from the net (Very low frequency of shots made behind it)
- Main insights:
  - Statistically significant increases in goal likelihood with:
    - Shots made horizontally and vertically closer to the net (edge)
    - Straighter shots

**Overall:** Goals are made more often from shots made closer to the net and at straighter angles

Call:

```
glm(formula = goal ~ x_from_edge + abs(y_fixed) + shot_angle,  
     family = "binomial", data = data)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-0.719628	0.040596	-17.73	<2e-16 ***
x_from_edge	-0.033677	0.001092	-30.83	<2e-16 ***
abs(y_fixed)	-0.030293	0.001976	-15.33	<2e-16 ***
shot_angle	-0.010243	0.000822	-12.46	<2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 63970 on 122348 degrees of freedom  
Residual deviance: 59827 on 122345 degrees of freedom  
(38211 observations deleted due to missingness)  
AIC: 59835

Number of Fisher Scoring iterations: 6



# Research Question 2

How do NHL teams differ in their ability to convert shots into goals, and how does this variation in goal conversion rates correlate with shot accuracy and effectiveness across different game periods?

# Data Wrangling

## Goal Conversion Rate

```
team_conversion <- hockey |>
  group_by(event_team) |>
  summarise(goal_conversion = sum(shot_outcome == "GOAL") / n())
```

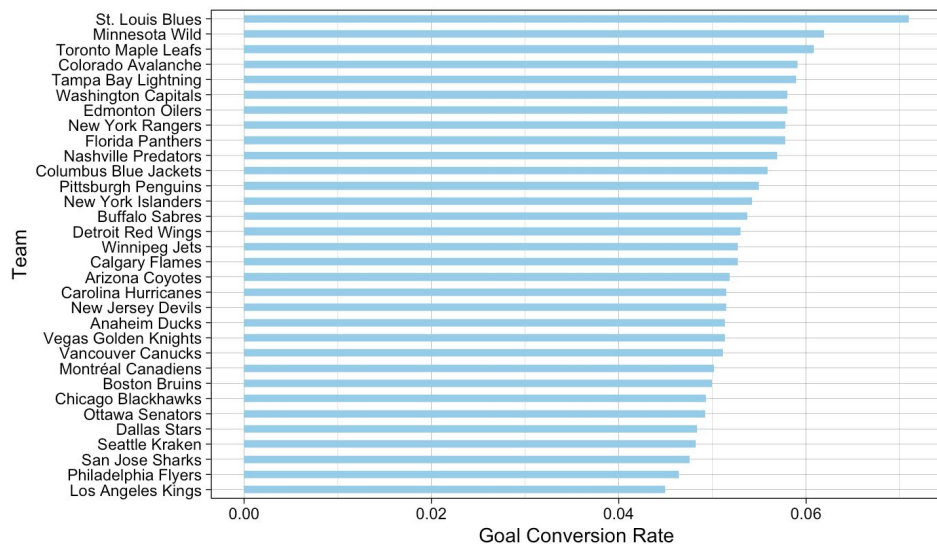
## Shot Accuracy Rate

```
hockey_data <- hockey |>
  mutate(shot_accuracy = ifelse(shot_outcome == "GOAL", 1, 0)) |>
  filter(period %in% c(1, 2, 3))

accuracy_by_period <- hockey_data |>
  group_by(event_team, period) |>
  summarise(accuracy = mean(shot_accuracy))
```

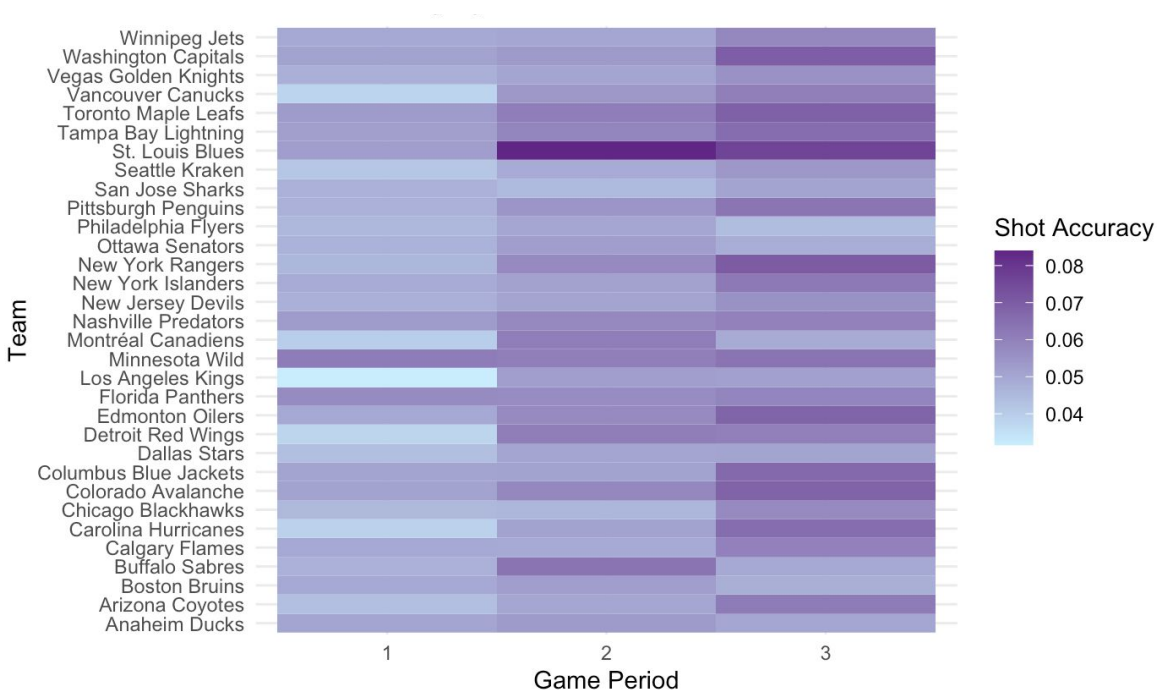
# Goal Conversion Rates by Team

- Goal conversion rates as a metric for evaluating offensive efficiency
- Insight into:
  - Shot accuracy
  - Effectiveness
  - Offensive performance
- Direct correlation between goal conversion rates and offensive performance in NHL teams.
- Teams with high conversion rates demonstrate effectiveness in turning shot opportunities into goals, suggests presence of efficient offensive strategies/skilled players.
- Lower conversion rate of teams like the Los Angeles Kings may indicate areas for improvement in their offensive gameplay.



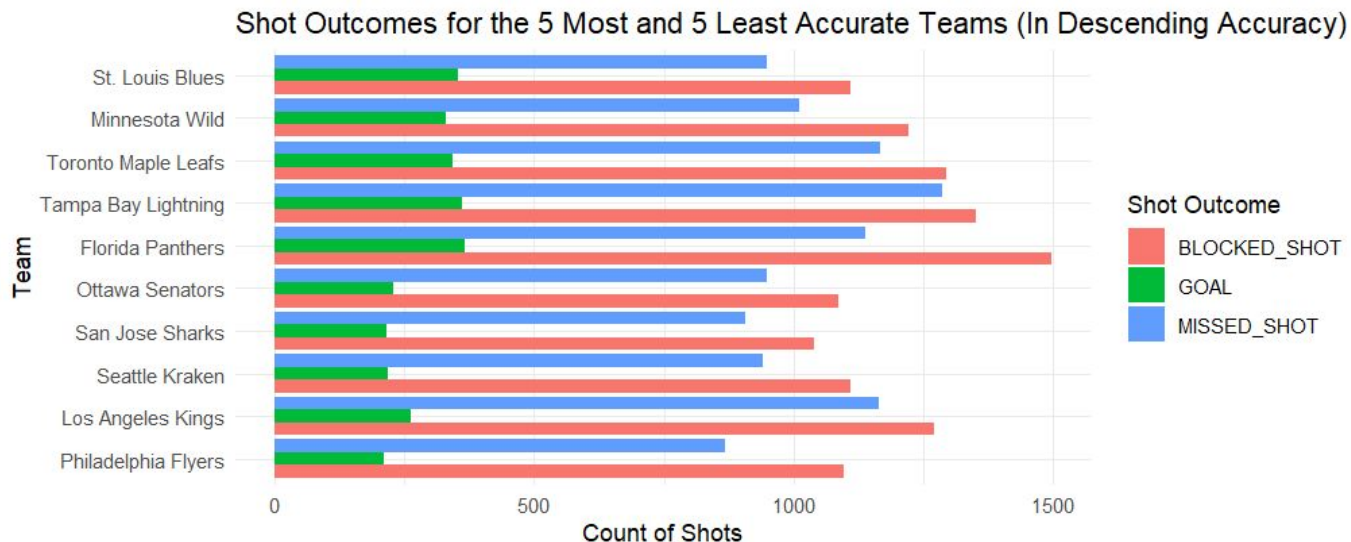
# Shot Accuracy by Team and Period

- Variability in teams' abilities to maintain consistent accuracy throughout a game
- Some insights:
  - St. Louis Blues exhibit exceptionally high shot accuracy in period 2
  - Los Angeles Kings demonstrate variability in shot accuracy throughout the game
  - Arizona Coyotes demonstrate progressive increase in shot accuracy
  - San Jose Sharks experience fluctuations, with period 2 showing lowest accuracy
  - Washington Capitals, Edmonton Oilers, Columbus Blue Jackets, and Colorado Avalanche display stronger shot accuracy in period 3
- Enables adaptation of strategies and tactics accordingly
- Opportunities for targeted interventions during periods of lower accuracy



# Shot Outcomes by Team Accuracy

- Decided to only use 10 teams in total
- Some insights:
  - Teams with higher accuracies had stronger performances in terms of proportion of goals overall, as well as fewer blocked shots and missed shots
  - Blues have higher proportion of goals than Flyers, and more goals as well
  - Opportunities to analyze specific areas of improvement, whether it is whether your shots are blocked or missed
- Teams exhibit similar trends in shot conversion, shot accuracy, and proportion of each category of shots (Flyers have lowest accuracy + second-lowest GC)
- Combining this data with RQ1 may lead to targeted improvement for teams in several ways (location on ice, angle, player formation)



# Research Question 3

Do there exist any significant differences between the score distributions of home vs. away games and by what amount?

Motivation:

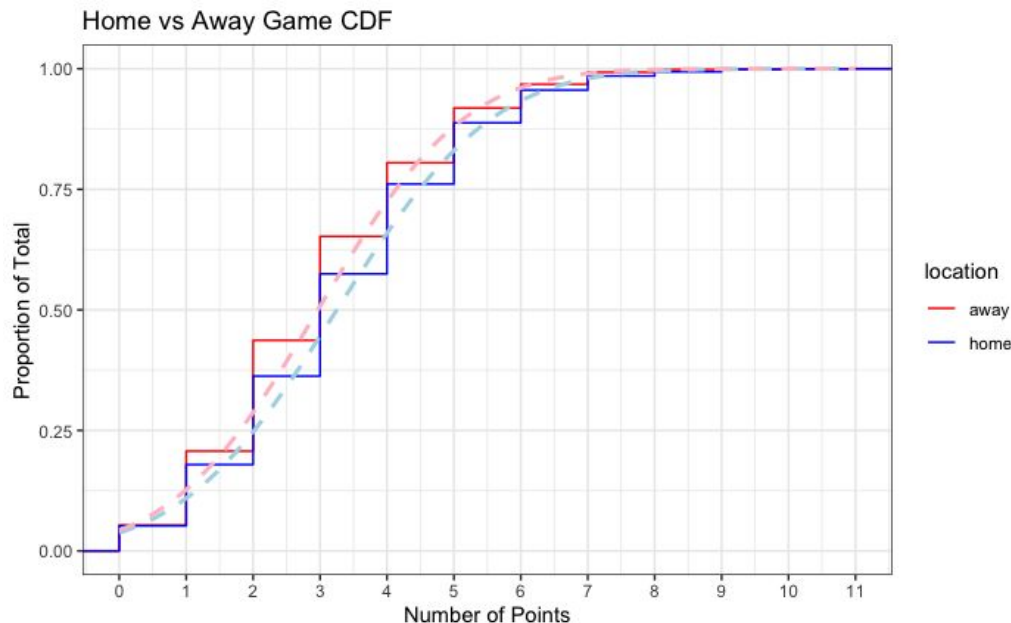
- Recently coming out of the pandemic
- Lower league wide fan attendance
- Many players temporarily prevented from playing the game

# Home versus Away Cumulative Point Densities

Main insights:

- Teams score between 0 and 3 points in most games
- Both home and away score distributions followed the normal curve
- Fitted normal distribution curve for home games (in light blue) lies to the right of away games (in pink)

Seems to indicate that home games still have a higher average score than their away counterparts



# Home vs Away Distributions

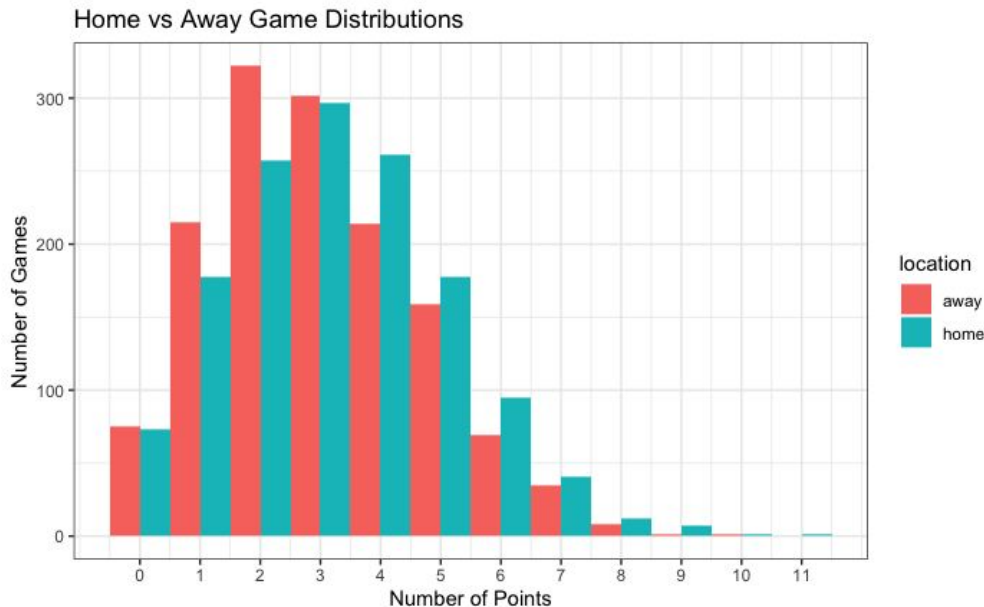
Main insights:

- Higher mode for home games
  - 2 points for away games
  - 3 points for home games
- Higher maximum for home games
- Smaller variance in away games than in home games

Using a t-test to examine differences in means:

- Differential means:
  - Home: 3.25 Away: 2.97
- 95% Confidence Interval of Difference in Means: 0.15 - 0.41
- P-value:  $2.683 \times 10^{-5}$

Allows us to compare games pre-pandemic to see if there are any differences in the magnitude of this separation in scores





# Conclusions

- Discovered results that may be used to inform team managers, general managers, and offensive and defensive coordinators
- Targeted and localized results in terms of shot outcome by location on the ice, and a team's individual performance may be impacted by even where they play
- Realized trends in goal conversion and shot accuracy, identifying areas of overall performance that may be optimized
- Quantified existence of a "home-field advantage", suggesting teams should expand their outreach to more rinks

# Future Topics

## **Optimizing Shot Placement on the Ice**

- Building on our investigation of shot location and its impact on goal-scoring likelihood
- Develop model through exploration of specific zones or angles within rink where shots are most effective
- Consider variables such as goalie positioning and defensive strategies

## **Analyzing Player Positioning and Movement Patterns**

- Examine how player positioning and movement patterns contribute to creating goal-scoring opportunities
- Track player movements during game to understanding positioning relative to goal-scoring opportunity
- Insights into effective offensive strategies based on player positioning and movement patterns

## **Examining Goalie Performance Metrics**

- Explore goalie performance metrics in relation to shot outcome to uncover defensive strategies
- Collect data on goalie positioning, reaction times, and block percentage in different game situations
- Conduct statistical analysis to identify correlations between goalie performance and defensive success
- Understand effective goalie strategies/positioning techniques to optimize team defensive capabilities