ANALYSIS OF HOCKEY'S GRITTY STATISTICS

State University of New York at Fredonia
By: Jakob Pachucinski
Mentor: Dr. Powell

Abstract

Hockey is known for its gritty nature and physicality. Players are constantly getting hit, fighting, and taking pucks to the body. However, in today's league these attributes are often overlooked in favor of shifty goals and speed. Do physical teams truly get rewarded with wins and points? Team statistics from the National Hockey League (2007 to 2021) such as hits, blocked shots, and penalty minutes were analyzed with linear regression tests, time series plots, and correlation matrices to see if there was significant relationship to gritty statistics and good teams. The data used was gathered from https://stathead.com/sport/hockey.

Introduction

The National Hockey League (NHL) has always been an exciting sport to watch for people of all ages. This is partly caused by the physical nature of the league and sport itself. In order to analyze *grittiness* of the sport, I chose three statistics that best represented this word: hits, penalty minutes, and blocked shots. For those of you that do not watch hockey often, hits are classified when a player throws his body into another player on the ice, penalty minutes happen when a rule is broken such as fighting, tripping, or hitting another player with a stick, and blocked shots occur when a player puts his body in front of the puck in order to prevent a shot from going towards the net.

Methods

- NHL team data was downloaded from https://stathead.com/sport/hockey. This data included almost every measurable hockey statistic for each team per season. Cleaning the data was necessary to combine some statistics into one data frame and to change some of the columns data types.
- A linear regression hypothesis test was done in order to see if there was a positive relationship between wins and hits, blocks, and/or penalty minutes.
- Linear regression models were created to predict the amount of wins a team gets in a season based on hits, penalty minutes, and blocked shots.
- To account for the lockout year and COVID year, a time series plot was created for the average amount of hits, blocked shots, and penalty minutes *per game*, for each season.
- Lastly, a linear regression model was created to predict the rate at which mean penalty minutes are declining in the NHL over time.

Results

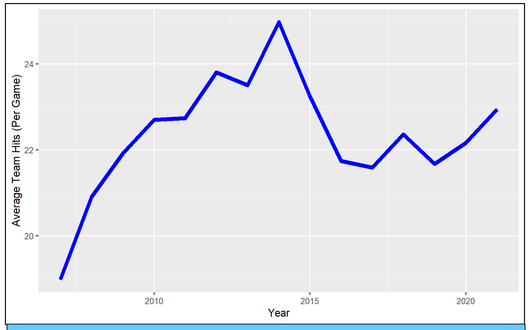

Figure 1. The average hits in the NHL per game from 2007 to 2021. As you can see there is a decline around 2015 which may be caused by the new 3 on 3 overtime which shortens overall games. However, hits have began to increase again around 2019.

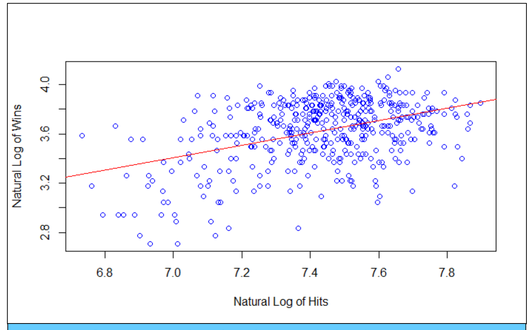

Figure 4. Scatter plot showing the correlation between the natural log of a team's hits in a season compared to the natural log of a team's wins in a season (2007 to 2021).

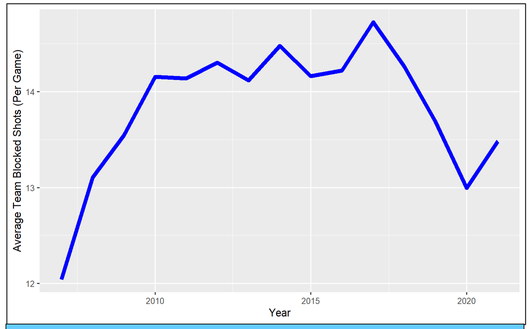

Figure 2. The average blocked shots in the NHL per game from 2007 to 2021. There appears to be a decrease around 2018. However, after COVID the amount increased slightly.

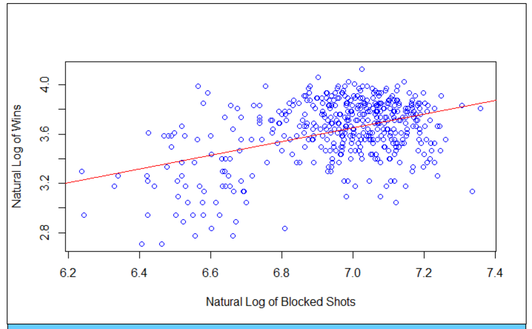

Figure 5. Scatter plot showing the correlation between the natural log of a team's blocked shots in a season compared to the natural log of a team's wins in a season (2007 to 2021).

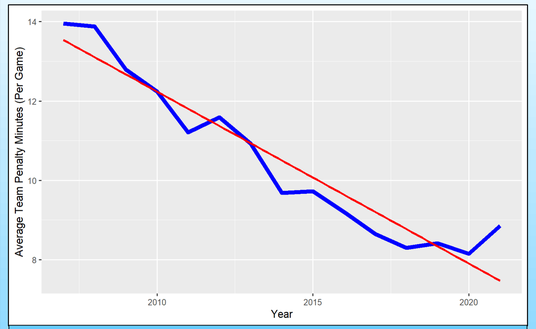

Figure 3. The average penalty minutes in the NHL per game from 2007 to 2021. This was the most consistent declining graph, showing a true negative slope. After extensive research, penalty minute declines can be attributed to the decrease of fighting in recent years. The linear regression equation shown in red is:

$$\hat{y} = 883.96203 - 0.43369x$$

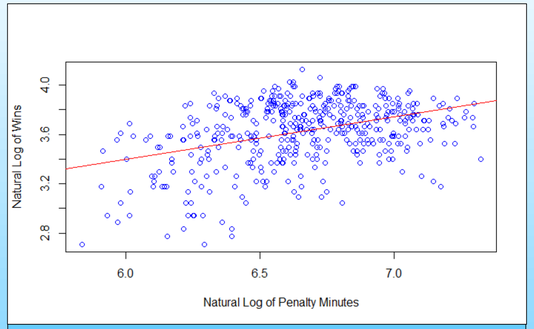

Figure 6. Scatter plot showing the correlation between the natural log of a team's penalty minutes in a season compared to the natural log of a team's wins in a season (2007 to 2021).

Conclusion

There are two main takeaways that can be discussed from the data analysis. First off, penalty minutes have substantially declined over the years from 2007. This can be attributed to the rapid decline of fights and the NHL's attempt to decrease violence and injuries throughout the league. Since fights account for 10 total minutes of penalties (5 for each team), they add 5 times more penalty minutes compared to a regular 2-minute minor penalty (tripping, delay of game, etc.). Also, the linear regression model in Figure 3 has an R^2 value of 0.9118. This shows extremely strong correlation between the year and average number of penalty minutes per game. As many hockey fans have realized, this demonstrates that the NHL is getting softer over the years. The second main result is that grittier teams appear to perform better in the regular season. Hits, blocked shots, and penalty minutes all had significant p-values when used as explanatory variables in a linear model with wins. In fact, each linear regression equation has a positive slope, reinforcing the fact that grittiness is a relevant asset in the National Hockey League and can help attribute to a team's success. Surprisingly, penalty minutes had the strongest correlation to wins out of the three statistics. This is extremely interesting because taking penalty minutes should put a team at a disadvantage. Penalties result in a team having to continue the game with one less player for at least 2 minutes. However, sometimes penalties are taken to hype up the bench, which may result with more effective play. Overall, the NHL appears to be shying away from grittiness even though these statistics can make (or break) a team's season.

Future Work and Limitations

- Originally, I wanted to analyze which players are underrated in gritty statistics. Now knowing how important hits, penalty minutes, and blocked shots are I could investigate certain players that are efficient with these three statistics. This could then help teams add cheaper players that can still improve their team.
- Getting data from more years of the NHL (not just 2007 to 2021), would give us more accurate information for the conclusion. Also, it would be interesting to see how much the league has changed throughout its existence.

Acknowledgements

This data was provided by Hockey Reference's website and their affiliate https://stathead.com/sport/hockey. I would also like to thank Dr. Amber Powell for her help throughout this project, semester, and my entire career at SUNY Fredonia.