What the Puck: Building a Successful NHL Team Using Data Science

Dustin Casey, Data Science Student, Bellevue University

Abstract

This project will explore how data science and analytics can be used to try to make an existing National Hockey League (NHL) franchise successful by optimizing their use of funds on player salaries. This will be done by exploring how a franchise can build the best roster while staying under budget and create an exciting team to watch.

Introduction

Creating a well-functioning, winning, and profitable sports franchise can be a challenge even in the best markets like Los Angeles or New York, but even more challenging in a small market or when a team does not have a large budget to work with. Optimizing and utilizing every dollar can make an average team good and a good team great. There is also a level of balance that needs to be achieved so that your team doesn't win every game and bore people.

Why is this Data Science?

- Bringing in outside data, processing it and exploring it, and then using it to make decisions for a sports organization
- Evaluating options for players and making meaningful decisions based on that analysis.

Acknowledgements

I want to thank my mom and my dad for supporting me with this project and teaching me to love the great game of hockey from a young age. Some of my favorite and most cherish childhood memories were at Minnesota Wild games and they are the reason/inspiration for this project.

Methods and Materials

- Using NHL player statistics, salary data, and scholarly sources on the subject.
- Creating a model to evaluate and rate players on their ability compared to their salary.
- Finding the players with the best value in each position.
- Using multiple machine learning algorithms to evaluate players.

Results

- Provide a full 23-man NHL roster
- Under the salary cap restrictions of the National Hockey League.
- Goal - best players who will already either be on the team, picked up via free agency, or drafted in the 2019 NHL season.
- Without being able to actually build the team and have them compete, testing the "best" roster possible will be tough and somewhat objective.
- Evaluation will be based on prior stats and college stats for players who are drafted.

Figure 1. NHL Logo

Discussion

- The idea of sabermetrics and using statistics to make decisions on who to sign comes from baseball.
- Does not translate easily to other sports that have more complicated team dynamics compared to baseball (i.e. hockey and basketball).
- Data is not always readily available for decision making for a multitude of reasons ranging from disputes between management and players to a variety of different intangible/unmeasurable statistical categories.

Basic NHL Stats Used Currently

- Corsi – Plus/minus for shot attempts instead of goals.
- Fenwick – same as Corsi, but only shots on goal and missed shots count
- (any statistic) for Percentage – (Corsi, Fenwick, Goals) as a ratio.
- (any statistic) relative to Teammates – how did a player perform in a given statistic compared to his teammates.

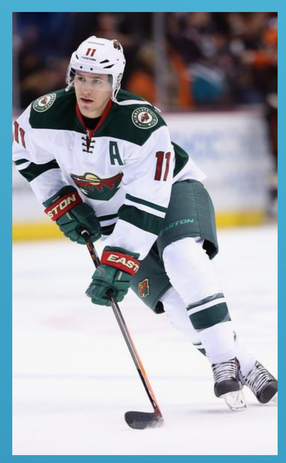

Figure 2: Zach Parise of the Minnesota Wild

Conclusions

- There is no easy answer to the questions we are asking
- Our output and analysis will be better than blindly making decisions off of what we think are "meaningful" metrics.
- We need to find areas we need to grow in that are also undervalued throughout the NHL and capitalize on this information.

Contact

Dustin Casey
Bellevue University
Email: dcasey@my365.bellevue.edu

References

1. Thomas W. Miller. 2016. Sports Analytics and Data Science: Winning the Game with Methods and Models (1st Edition). Pearson Education, Inc., Old Tappan, NJ.
2. Skinner, B., & Guy, S. J. 2015. A method for using player tracking data in basketball to learn player skills and predict team performance. PLoS One, 10(9) doi:http://dx.doi.org.ezproxy.bellevue.edu/10.1371/journal.pone.0136393
3. Brian Macdonald. 2012. An Expected Goals Model for Evaluating NHL Teams and Players. MIT Sloan Sports Analytics Conference, Boston, MA.
4. Ballard, C. (2005, October 24). Measure of success. Sports Illustrated, 103(16), pp. 78-82.Morton L. Heilig. 1962. Sensorama Simulator, U.S. Patent 3,050,870, Filed January 10, 1961, issued August 28, 1962.
5. Daniel S. Mason and William Foster. 2007. *Putting Moneyball on Ice?*, International Journal of Sport Finance (November 2007).
6. Dater, A. (2006, October 15). A new ice age: NHL plays numbers game Hockey taking broader approach in analyzing statistics to determine a player's net worth in team game. Denver Post, p. BB.01
7. Davidi, S. (2006, October 5). Baseball-style statistical analysis starts creeping into hockey world. The Canadian Press.
8. Charlie O'Connor (2017, October 9). *An advanced stat primer: Understanding basic hockey metrics.* The Athletic.