Statistics Group 1 Project Proposal

Ice Hockey

Bevan Stanely

Nikhil Verma

Shantanu Sengupta

Abhishek Kumar

Gubbala Hari Priya

Lakshay Sharma

Sunku Vydehi

2023-09-30

Table of Contents

[1. Executive Summary 2](#_Toc1)

[2. Problem Statement 3](#_Toc2)

[2.1 What are the critical factors predicting a player’s performance? 3](#_Toc3)

[3. Proposed Solution 4](#_Toc4)

[3.1 Variable Significance 4](#_Toc5)

[3.1.1 Individual Characteristics 4](#_Toc6)

[3.1.2 Performance Characteristics 5](#_Toc7)

[3.2 Limitations 5](#_Toc8)

[4. Timeline 6](#_Toc9)

[5. References 7](#_Toc10)

…

# 1. Executive Summary

Ice hockey is a fast-paced team sport renowned for its combination of speed, physicality, and skill, predominantly played in cold-weather nations like Canada, the USA, Russia, Sweden, and Finland. Women’s ice hockey, a variant of the sport played exclusively by female athletes, shares similarities with men’s hockey but boasts its distinct history. Since 1990, the International Ice Hockey Federation (IIHF) has organized the IIHF World Women’s Championship, showcasing the growth of women’s ice hockey.

The USA took the championship in the current year, with Canada coming in second. Since the inception of the competition, they have been regularly among the winners. The first eight championships were won by Canada, whereas the United States dominated in the more recent editions, taking home ten out of the last 14 championships([“Women’s World Championship Statistics” n.d.](#ref-WomenWorldChampionship)).

The main objective in ice hockey is to score goals by shooting a puck into the opposing team’s net. A standard game comprises three periods, each lasting 20 minutes, with intermissions for player rest and ice maintenance. Victory goes to the team with the most goals after these three periods.

In a typical lineup, each team fields six players simultaneously on the ice, featuring a goaltender (goalie) and five skaters. These skaters are typically divided into three forwards (left wing, center, and right wing) and two defensemen.

To maintain fairness and uphold the competitive spirit of the game, ice hockey employs an extensive set of rules and penalties. Players who commit infractions can be temporarily sidelined in the penalty box, leading to power plays (when one team has more players on the ice) or penalty kills (when a team is shorthanded). These regulations play a crucial role in preserving the integrity of the sport.

Our analysis involves the single-year data of the 2023 IIHF World Women’s Championship. We will address the individual player performance with linear regression and identify the key attributes affecting the performance.

# 2. **Problem Statement**

## 2.1 **What are the critical factors predicting a player’s performance?**

The challenge is to systematically identify and analyze the key factors that reliably predict a player’s performance. This includes analyzing the physical, skill-based, psychological, and strategic elements ([Nigg et al. 2020](#X9d19b3215b5ca1f18315c8df88a1bc83304c4e4)) that contribute to a player’s effectiveness in the game. There is a need for a framework that allows coaches, teams, and players to make data-driven decisions and improve the overall competitive edge in the sport.

For this study, we only looked at the official data on the physical and game statistics of 203 forwards and defenders. The rationale behind the decision to exclude goalies from our dataset stems from the disparity in their performance metrics with forwards and defenders.

In women’s ice hockey, the publicly available data is limited in terms of how important the attribute is for the players’ performance. Furthermore, there is currently no tidy dataset available. So, we are collecting and putting together information from official documents and extending it. We also want to look at the player statistics of the participants in the 2023 Women’s Ice Hockey World Championship and identify the key performance indicators for high-performing players.

# 3. **Proposed Solution**

We will compile a tidy dataset using publicly available data from the 2023 Women’s Ice Hockey World Championship. Furthermore, we will perform regression analysis to identify the dependence of the individual variables on the player’s performance. We expect to derive critical insights that could have implications for team strategizing. The insights will also help ice hockey fans understand their favorite players better.

The attributes of the data collected are as follows

|  |  |  |
| --- | --- | --- |
| No. | Attribute | Description |
|  | **Dependent variable** |  |
| 1 | Points/Game | Average of points scored across all the games played. |
|  | **Independent variables** |  |
| 2 | Position | Player-position: Forwards or Defensemen (Goalies are ignored since their performance metrics are different) |
| 3 | Goals |  |
| 4 | Assist | Goals that were assisted by focal player |
| 5 | Country | Team played under |
| 6 | Games played | Total games played by the player in the current season. |
| 7 | Height | In cms |
| 8 | Weight | In Kg |
| 9 | Penalties in minutes | Penalty duration over the championship |
| 10 | +/- Plus/minus | Player’s goal differential while they are on the ice |
| 11 | Shots on goal(SOG) | The number of attempts made by a player to score a goal that directly challenge the opposing team’s goaltender |
| 12 | Age |  |
| 13 | Shoots (left or right) | Player handedness |
| 14 | Total minutes/s played (TM) | Total player time on ice |

Our choice of points for the dependent variable covers a player’s ability to score and the team spirit assisting their teammates in scoring. As a game, hockey uses points for ranking purposes, reinforcing our choice. The remaining independent variables were picked from the player parameters available.

## 3.1 Variable Significance

We can group the things that affect how a player plays a game into different categories.

### 3.1.1 Individual Characteristics

The individual player possesses distinct characteristics such as height, weight, age, position, and player handedness (shooting). Since they could affect individual performance, they would be useful for our model.

### 3.1.2 Performance Characteristics

Resulting from the players’ participation in the championship are variables pertaining to the outcome of their participation. Goals, assists, games played, penalties in minutes, +/- differential, shots on goals, and total minutes played are performance variables. The average of total goals scored with assists is our dependent variable. The remaining variables indicate how the players’ time on the ice was spent.

Finally, the variable country groups the players with their teammates. Ice hockey being a team sport, the player’s performance is linked to their teammates. Therefore, we could use it to assess the player performance distribution within the team.

We will use the following data sources for the data.

| Sr. no. | Data |
| --- | --- |
| 1 | Quant hockey Website ([n.d.](#ref-IIHFWorldChampionship)) |
| 2 | IIHF Official Website ([n.d.](#ref-WomenWorldChampionship)) |
| 3 | Team USA ([2023](#ref-USAPlayerStatistics2023)) |
| 4 | Team Japan ([2023](#ref-JapanPlayerStatistics2023)) |
| 5 | Team Canada ([2023](#ref-CanadaPlayerStatistics2023)) |
| 6 | Team Switzerland ([2023](#ref-SwitzerlandPlayerStatistics2023)) |
| 7 | Team Czech Republic ([2023](#ref-CzechiaPlayerStatistics2023)) |
| 8 | Team France ([2023](#ref-FrancePlayerStatistics2023)) |
| 9 | Team Finland ([2023](#ref-FinlandPlayerStatistics2023)) |
| 10 | Team Germany ([2023](#ref-GermanyPlayerStatistics2023)) |
| 11 | Team Sweden ([2023](#ref-SwedenPlayerStatistics2023)) |
| 12 | Team Hungary ([2023](#ref-HungaryPlayerStatistics2023)) |

## 3.2 Limitations

We foresee two limitations of the dataset, which should be accounted for in the analysis.

1. Small sample size compared to other tournaments
2. The tournament had two parts. 1st part was a group stage and 2nd a knockout competition, giving more match time and the opportunity to score more points for qualified teams.

# 4. Timeline

The anticipated time frame for the project tasks is as follows.

|  |  |
| --- | --- |
| Task | Due Date |
| Finalize Data Collection | 5th September 2023 |
| Mid Term Review | 3rd October 2023 |
| Final Submission | 29th November 2023 |

# 5. References

“Canada Player Statistics.” 2023. April 16, 2023. <https://www.iihf.com/pdf/498/ihw4980can_83_8_0>.

“Czechia Player Statistics.” 2023. April 16, 2023. <https://www.iihf.com/pdf/498/ihw4980cze_83_8_0>.

“Finland Player Statistics.” 2023. April 16, 2023. <https://www.iihf.com/pdf/498/ihw4980fin_83_7_0>.

“France Player Statistics.” 2023. April 16, 2023. <https://www.iihf.com/pdf/498/ihw4980fra_83_4_0>.

“Germany Player Statistics.” 2023. April 16, 2023. <https://www.iihf.com/pdf/498/ihw4980ger_83_6_0>.

“Hungary Player Statistics.” 2023. April 16, 2023. <https://www.iihf.com/pdf/498/ihw4980hun_83_6_0>.

“IIHF World Championship (Women) 2023 - Player Stats.” n.d. QuantHockey. Accessed August 30, 2023. <https://www.quanthockey.com/whc-women/en/seasons/whc-women-players-stats.html>.

“Japan Player Statistics.” 2023. April 16, 2023. <https://www.iihf.com/pdf/498/ihw4980jpn_83_7_0>.

Nigg, Claudio R., Andreas Gessner, Carina Nigg, Marco Giurgiu, and Rainer Neumann. 2020. “Demographic, Physiological, Psychological, and on-Ice Performance Indicators Predict Plus/Minus Status of Recreational Ice Hockey Players Across a Season.” *German Journal of Exercise and Sport Research* 50 (4): 463–69. <https://doi.org/10.1007/s12662-020-00679-2>.

“Sweden Player Statistics.” 2023. April 16, 2023. <https://www.iihf.com/pdf/498/ihw4980swe_83_9_0>.

“Switzerland Player Statistics.” 2023. April 16, 2023. <https://www.iihf.com/pdf/498/ihw4980sui_83_9_0>.

“USA Player Statistics.” 2023. IIHF International Ice Hockey Federation. April 16, 2023. <https://www.iihf.com/pdf/498/ihw4980usa_83_8_0>.

“Women’s World Championship Statistics.” n.d. IIHF International Ice Hockey Federation. Accessed August 30, 2023. <https://www.iihf.com/en/events/2023/ww/teamstats/scoringefficiency>.