Data Descriptions and Context

Background and Data:

The NWHL has finished their 2021 season, and are getting ready to build new rosters for 2021/2022. Your Data Science team is hired by Toronto Six to help assemble a team for the new season. You can choose any current player from any current team in the league. Some of the questions you may want to address are:

1. The coaches would like to select five top players. The players should, ideally, excel in all areas of the game. Since scoring goals is the ultimate objective, at least three should be excellent goal scorers and at least two should be excellent passers. In addition, you need at least two faceoff specialists and one takeaway specialist. It would be especially beneficial if the passers were familiar with the shooting specialists (i.e., had a good track record of completing passes to these players).
2. The coaches would like to select five power play specialists and four penalty kill specialists. Note that the main objective of the power play is to score goals, while the main objective of the penalty kill is to prevent the opponent from doing so. Thus, the key skills for a penalty killer are takeaways and precise passes. For a power play specialist the key skills are successful shots and precise passes.
3. Pick a player you feel should definitely be included on the team and explain why. Provide advice to this player on
    1. What types of shots should the player focus on from each position in the field?
    2. What types of passes should the player be attempting in the defensive and offensive zones?
    3. What are the key areas for improvement?

Please select only one question to focus on (this does not have to be any of the questions listed above – feel free to make up your own question). **While the 2021/2022 season has already taken place, the goal is not to match what happened but rather use the data to form your own best team.**

Your submission should be in the form of a presentation to any NWHL General Manager addressing one or more of these points. The presentation should be limited to 10 slides and be non-technical in nature (hockey staff are not known for their data science skills). It can be followed by an Appendix containing more technical descriptions of your work. While you can also submit your code in a separate file, the presentation should be self-contained (i.e., it should not be necessary for the reader to open your code).

The data is provided in a separate CSV file (**Rotman MMA Summer Datathon NWHL.csv (https://q.utoronto.ca/courses/320820/pages/data-excel-file)**). Definitions are provided below.

Hints / Suggestions: For further reading, check out this article from the IIHF: **Applying analytics to 2019 Women's Worlds (https://www.iihf.com/en/news/18514/applying-analytics-to-2019-women-s-worlds)** (https://www.iihf.com/en/news/18514/applying-analytics-to-2019-women-s-worlds).

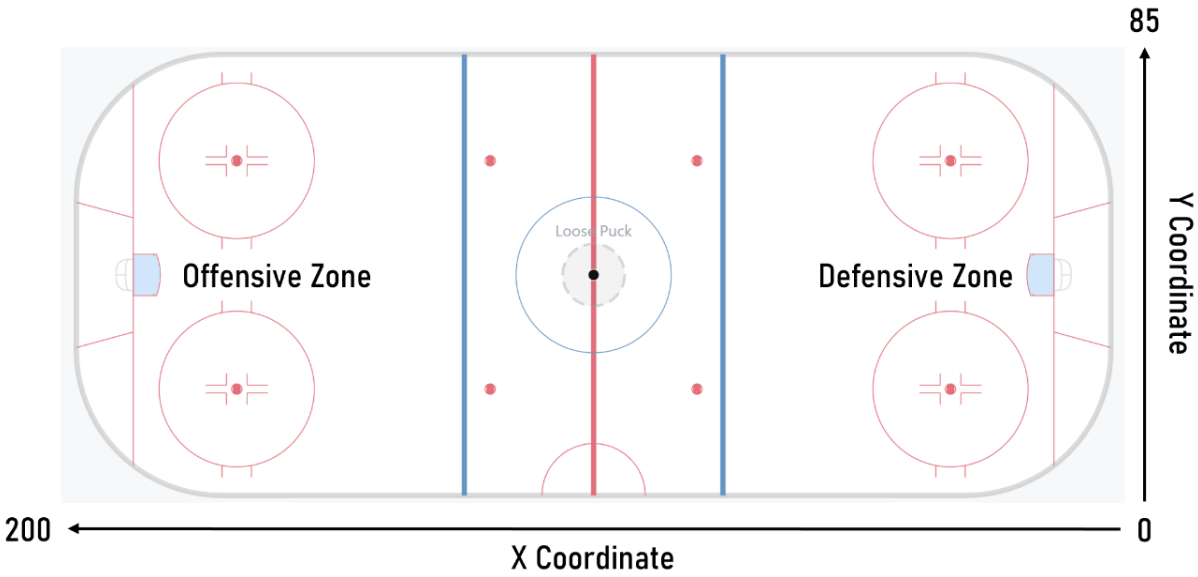
Dataset

Summary

The dataset is comprised of Stathletes-tracked women's hockey data from the NWHL. The included events have been translated from Stathletes' raw data to enhance accessibility and interpretability. The various event types include shots, plays, takeaways, puck recoveries, dump ins, dump outs, zone entries, faceoffs and penalties. Event definitions may slightly differ from other sources. For each event, expanded details are provided and the relevant skaters and teams involved are indicated when necessary.

Contextual Data

- Date (e.g. '2020-12-23'. Format = 'yyyy-mm-dd')
- Home Team (e.g. 'Toronto Maple Leafs')
- Away Team (e.g. 'Boston Bruins')
- Period (range from 1-3 for regulation, 4+ for overtime)
- Clock (e.g. '19:34'. Format = 'mm:ss')
- Home Team Skaters (range from 3-6 for home skaters currently on the ice)
- Away Team Skaters (range from 3-6 for away skaters currently on the ice)
- Home Team Goals (current goals scored by the home team at the time of the event)
- Away Team Goals (current goals scored by the away team at the time of the event)
- Team (name of the team responsible for the event)
- Player (name of the player responsible for the event)
- Event (type of event, e.g. 'Play', 'Shot', …)
- X Coordinate (x-coordinate of where an event occurred on the ice, between 0 and 200)
- Y Coordinate (y-coordinate of where an event occurred on the ice, between 0 and 85)
    - Coordinates are always from the perspective of the eventing team



- Detail 1-4 (up to 4 supplementary details for each event, varies by event type)
- Player 2 (name of a secondary player involved in an event, varies by event type)
- X Coordinate 2 (x-coordinate of a secondary event detail, varies by event)
- Y Coordinate 2 (y-coordinate of a secondary event detail, varies by event)

Events

Shot

Shot attempts that are unsuccessful (block, miss or save)

Players Involved

- Player: Shooter

Coordinates

- X,Y Coordinate: Release location

Event Details

- Detail 1: Shot Type (Deflection, Fan, Slapshot, Snapshot, Wrap around, Wristshot)
- Detail 2: Shot destination (on net, missed or blocked)
- Detail 3: Traffic (true or false)
- Detail 4: One timer (true or false)

Goal

Shot attempts that are successful (goal)

Players Involved

- Player: Shooter

Coordinates

- X,Y Coordinate: Release location of the puck

Event Details

- Detail 1: Shot Type (Deflection, Fan, Slapshot, Snapshot, Wrap around, Wristshot)
- Detail 2: Shot destination (on net, missed or blocked)
- Detail 3: Traffic (true or false)
- Detail 4: One timer (true or false)

Play

Pass attempts that are successful

Event Types

- Direct (e.g. a tape-to-tape pass)
- Indirect (e.g. a pass that is rimmed along the boards)

Players Involved

- Player: Passer
- Player 2: Intended pass target

Coordinates

- X,Y Coordinate: Pass release location
- X,Y Coordinate: Pass target location

Event details

Detail 1: Pass Type

- Direct (eg. a tape-to-tape pass)
- Indirect (eg. a pass that is rimmed around the boards)

Incomplete Play

Pass attempts that are unsuccessful

Event Types

- Direct (e.g. a tape-to-tape pass)
- Indirect (e.g. a pass that is rimmed along the boards)

Players Involved

- Player: Passer
- Player 2: Intended pass target

Coordinates

- X,Y Coordinate: Pass release location
- X,Y Coordinate: Pass target location

Event details

Detail 1: Pass Type

- Direct (eg. a tape-to-tape pass)
- Indirect (eg. a pass that is rimmed around the boards)

Takeaway

Steals, pass interceptions and won battles that lead to a change in possession

Players Involved

- Player: Skater credited with the takeaway

Coordinates

- X,Y Coordinate: Location where the skater gained possession when taking the puck away

Puck Recovery

Possession gains initiated by retrieving a loose puck that was created by a missed/blocked/saved shot, an advance (e.g. dump-out/dump-in), a faceoff or a broken play

Players Involved

- Player: Skater who recovered the puck

Coordinates

- X,Y Coordinate: Location where skater gained possession

Dump In/Out

Actions in which a skater intentionally concedes possession by advancing the puck up ice

Players Involved

- Player: Skater who dumped/advanced the puck

Coordinates

- X,Y Coordinate: Location where skater released the puck

Event details

- Detail 1: Possession Outcome (Retained, Lost)

Zone Entry

Attempts to move the puck into the offensive zone from the neutral zone

Players Involved

- Player: Entry skater
- Player 2: Targeted defender

Coordinates

- X,Y Coordinate: Point of release for dumps/advances, point where puck crossed the blueline for passes and carries

Event details

- Detail 1: Entry Type (Carried, Dumped, Played)

Faceoff Win

Faceoffs

Players Involved

- Player: Skater who won the draw
- Player 2: Skater who lost the draw

Coordinates

- X,Y Coordinate: Location of faceoff dot

Penalty Taken

Infractions

Players Involved

- Player: Skater who took the penalty
- Player 2: Skater who drew the penalty

Coordinates

- X,Y Coordinate: Location of infraction

Event Details

- Detail 1: Infraction Type (e.g. Slashing, Tripping, Roughing, Hooking, ...)