

Final Project Presentation

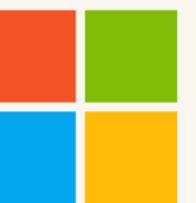
Team: Nexus Nucleus

SCTP - Junior Data Engineer 2024

9 October 2024



Generation
SINGAPORE



ABOUT US



ASMR Research

ASMR Research formed by a group of data enthusiasts specializing in Data Engineering and Analytics.

- **Asher**
- **Serchen**
- **Min**
- **Richard**

Previous portfolio includes :

Analysis on Singapore Bicycle Parking Locations





Interim Project

Analysis on Singapore's Bicycle Parking Spaces

► Objective

From the 3 estates (Jurong, Woodland, Tampines)...

1

Most HDB bicycle parking lots

Answer : Tampines

2

Most number of sheltered bicycle lots

Answer: Jurong

3

Most bus-stops with bicycle parking facilities

Answer : Tampines

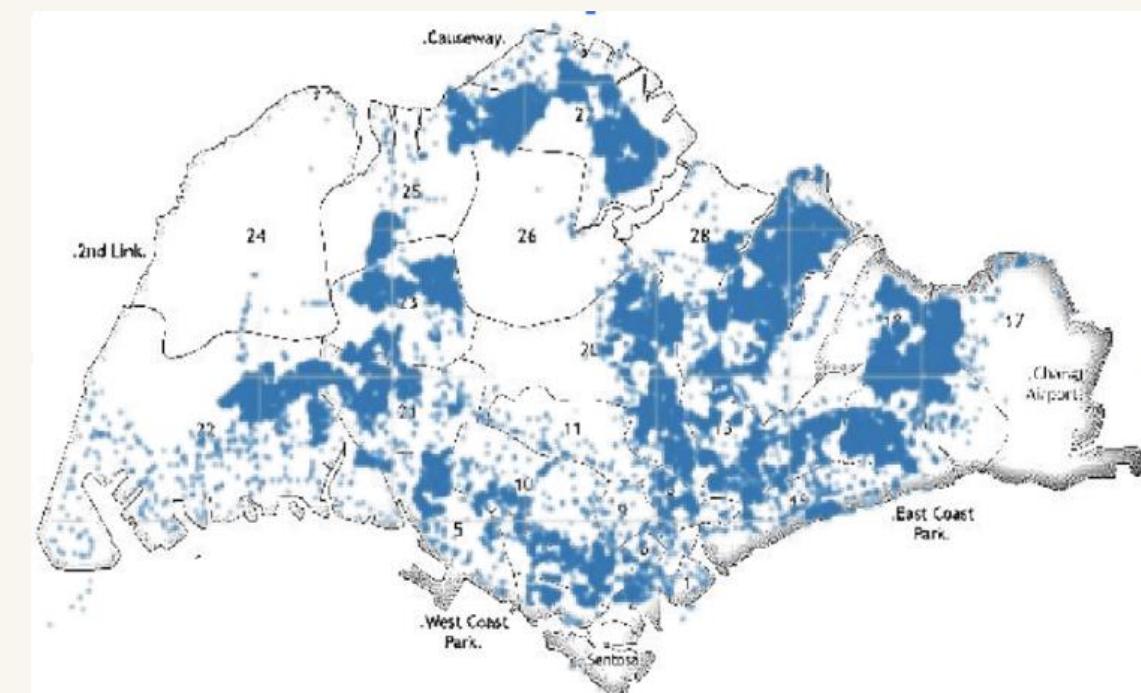
► Method/Approach



► Tools & Libraries

Jupyter Notebook
PostgreSQL
Matplotlib.pyplot
SQLAlchemy
Requests
Pandas
NumPy

► Visualization



NHL Game Data Analysis



TABLE OF CONTENTS

01

INTRODUCTION

Background/
Problem Statement

02

OBJECTIVE

Project Objectives

03

METHODOLOGY

Architecture Overview
Approach

04

ANALYSIS

NHL 101
X-Y Coordinates
NPC Dream Team

05

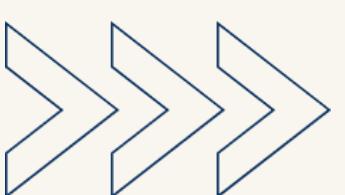
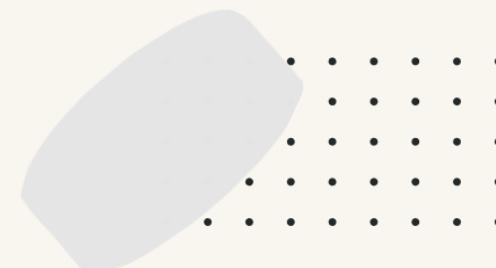
CHALLENGES

Challenges and
Workarounds

06

CONCLUSION

Learnings



INTRODUCTION

Background

In late 2020, an undisclosed game studio started developing their next NHL computer game.

They had access to detailed NHL data, including play-by-play information with x, y coordinates which they planned to use in their game development

However....



INTRODUCTION



Problem Statement

The game studio lacked the expertise to analyze the NHL dataset.

And thus, they approached **ASMR Research** to collaborate with them on data analysis and leverage data-driven insights to enhance gameplay mechanics and player dynamics



OBJECTIVES

ETL Pipeline

Effective, scalable ETL pipeline
for NHL data

- Ingest
- Transform
- Load

X-Y Coordinates

To propose the best attack
angles, distances, and shot types
with the highest probability of
scoring goals in the rink

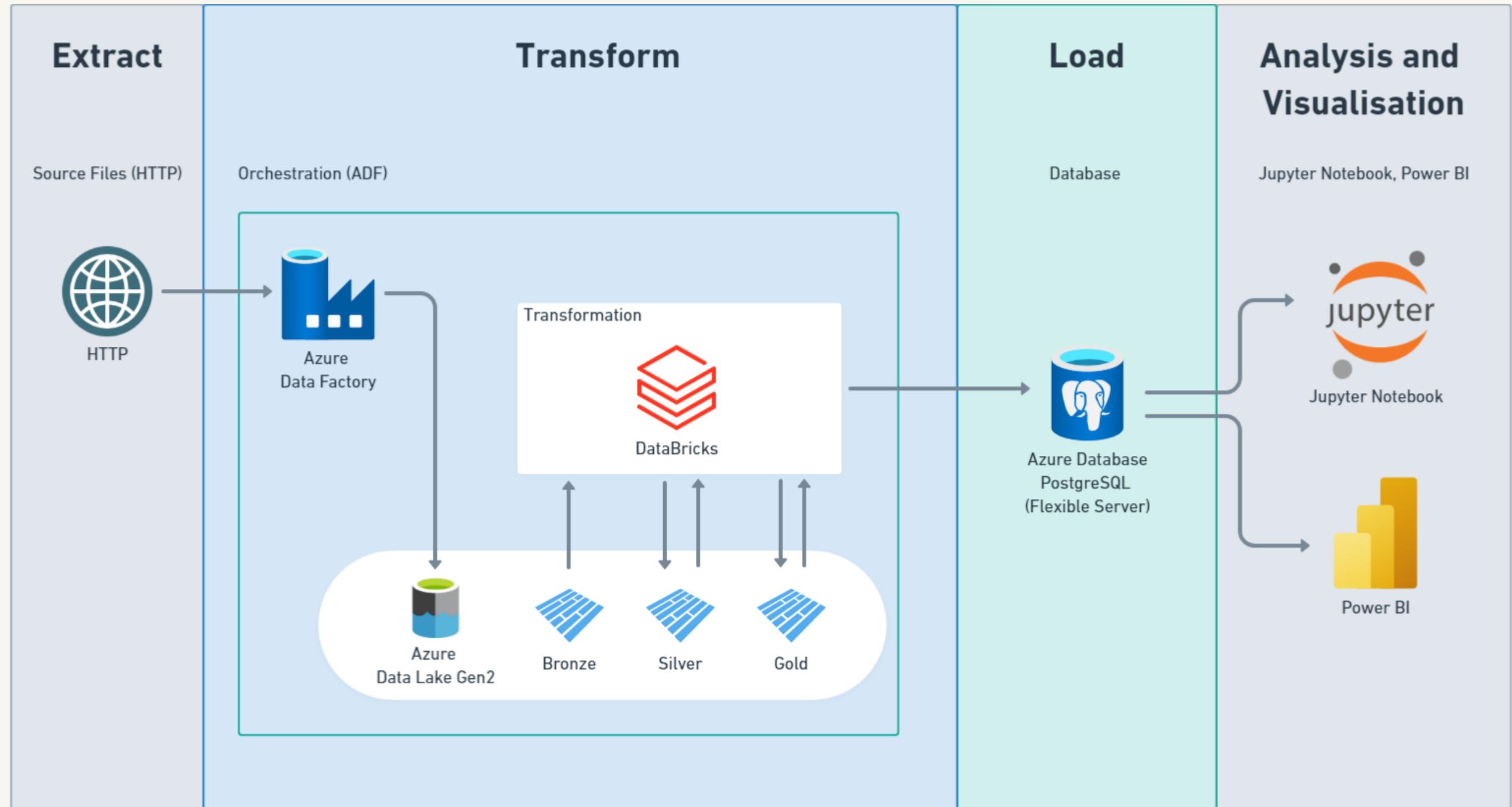
NPC Dream Team

To identify and propose a dream
team of 23 non-playable
characters (NPCs), comprising 12
forwards, 8 defensemen, and 3
goalies.



METHODOLOGY

Pipeline Architecture



Services Used

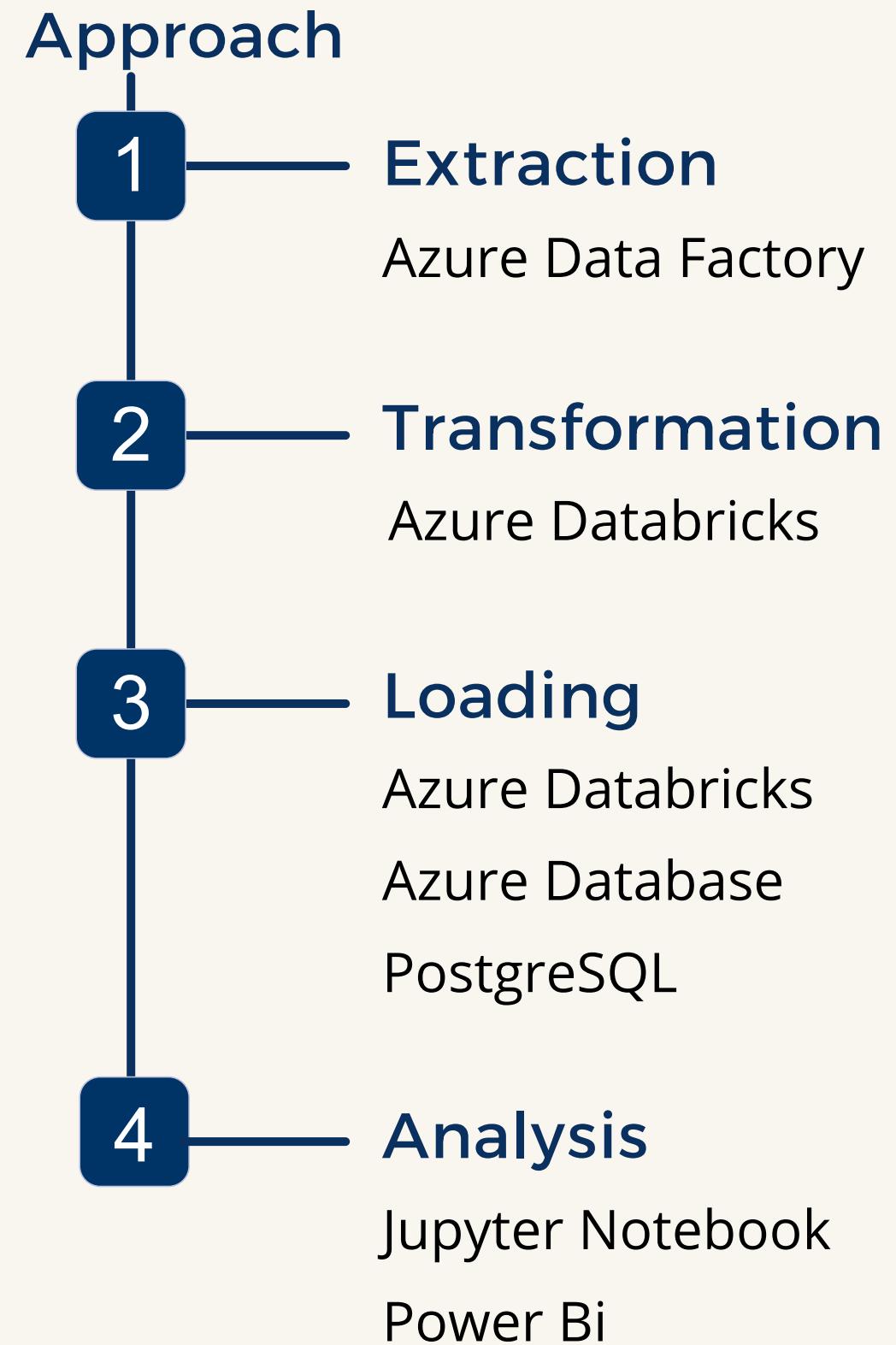
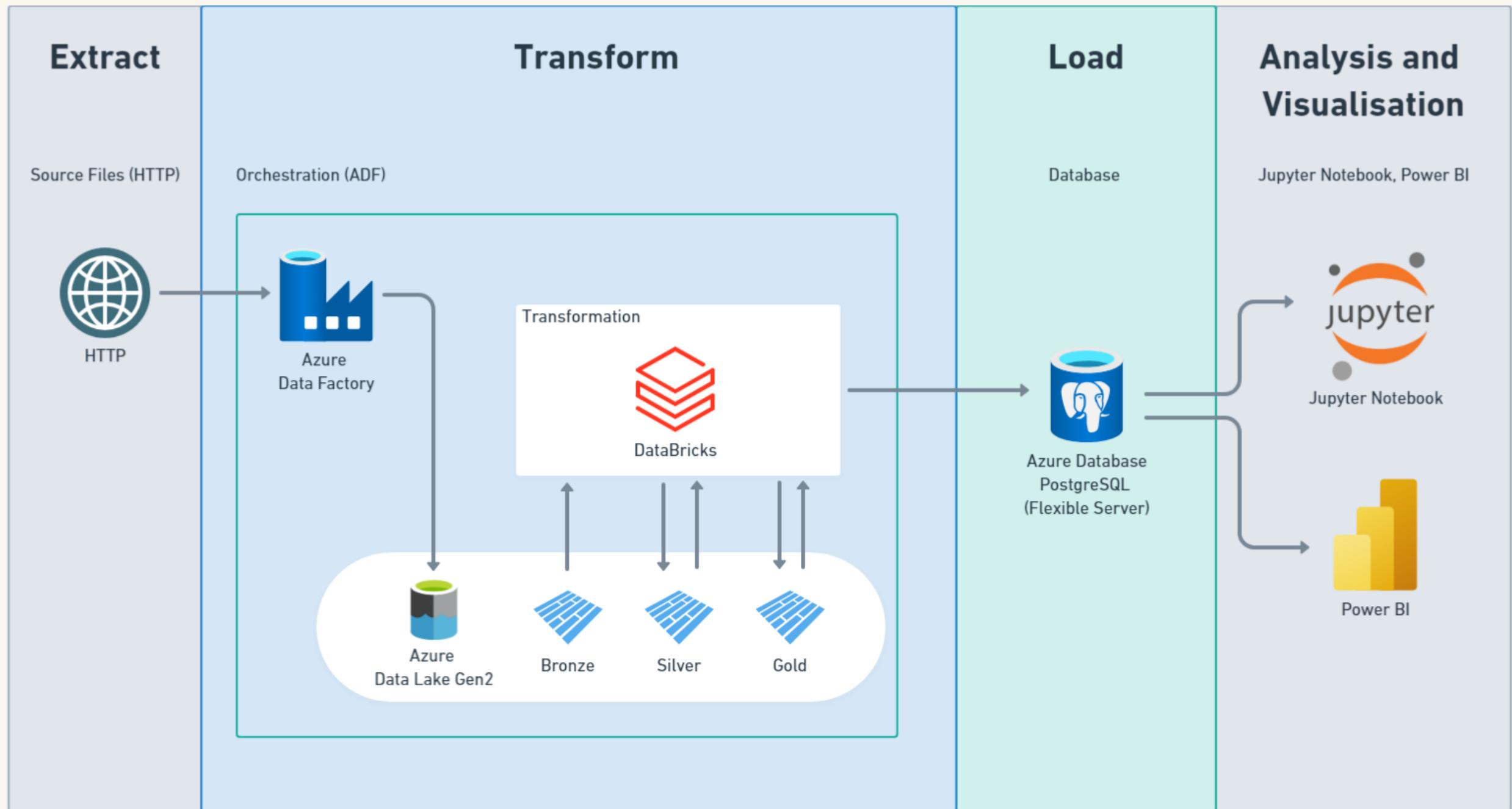
- Azure Data Factory (ADF)
- Azure Data Lake Gen2 (ADLS)
- Azure Databricks
- Azure Database PostgreSQL

Medallion Architecture

- **Bronze:** Raw, unprocessed
- **Silver:** Cleansed, conformed
- **Gold:** Curated, ready for business

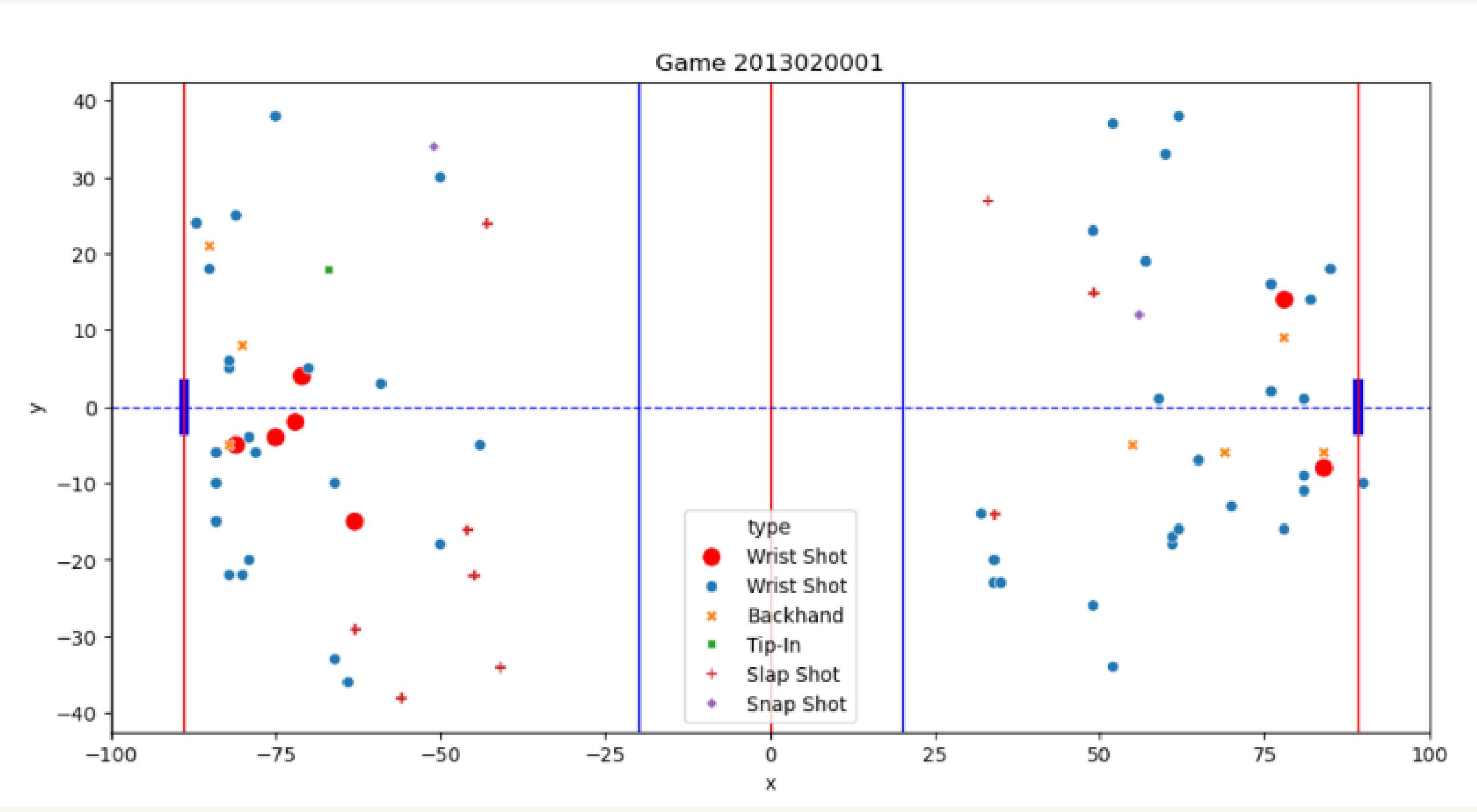
METHODOLOGY

Pipeline Architecture





NHL 101



Game

60 minutes long: 3 periods of 20 minutes each.

Team

23 players in a team
6 players consisting
3 attackers, 2 defenders, 1
goalie

Rink

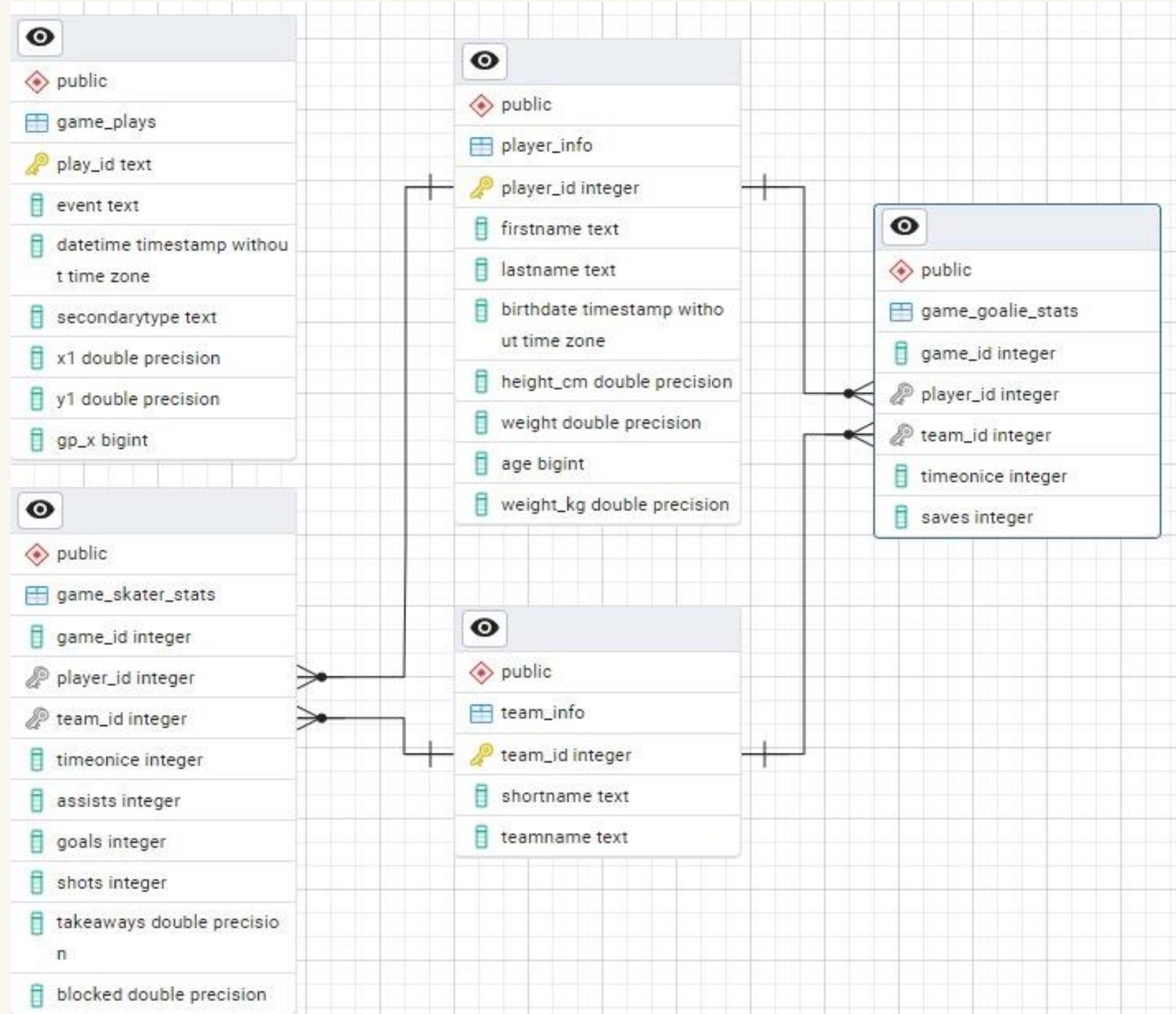
200 ft x 85 ft
Centre: 0,0
Goal 0,-89 and 0,89
Left/Right wing

Shot Types

Wrist, Slap, Snap, Tip-In,
Backhand, Wrap around,

ANALYSIS

Database ERD



(A) X-Y Analysis

(B) Dream Team Analysis

5 Tables

- **game_plays**
- **game_skater_stats**
- **game_goalie_stats**
- **player_info**
- **team_info**

Uploaded from Gold layer for analysis
in PostgreSQL

ANALYSIS: X-Y



PROBLEM STATEMENT

Which positions on the rink are most likely to result in a goal?

What types of shots are most effective based on distance and angle from the goal post?

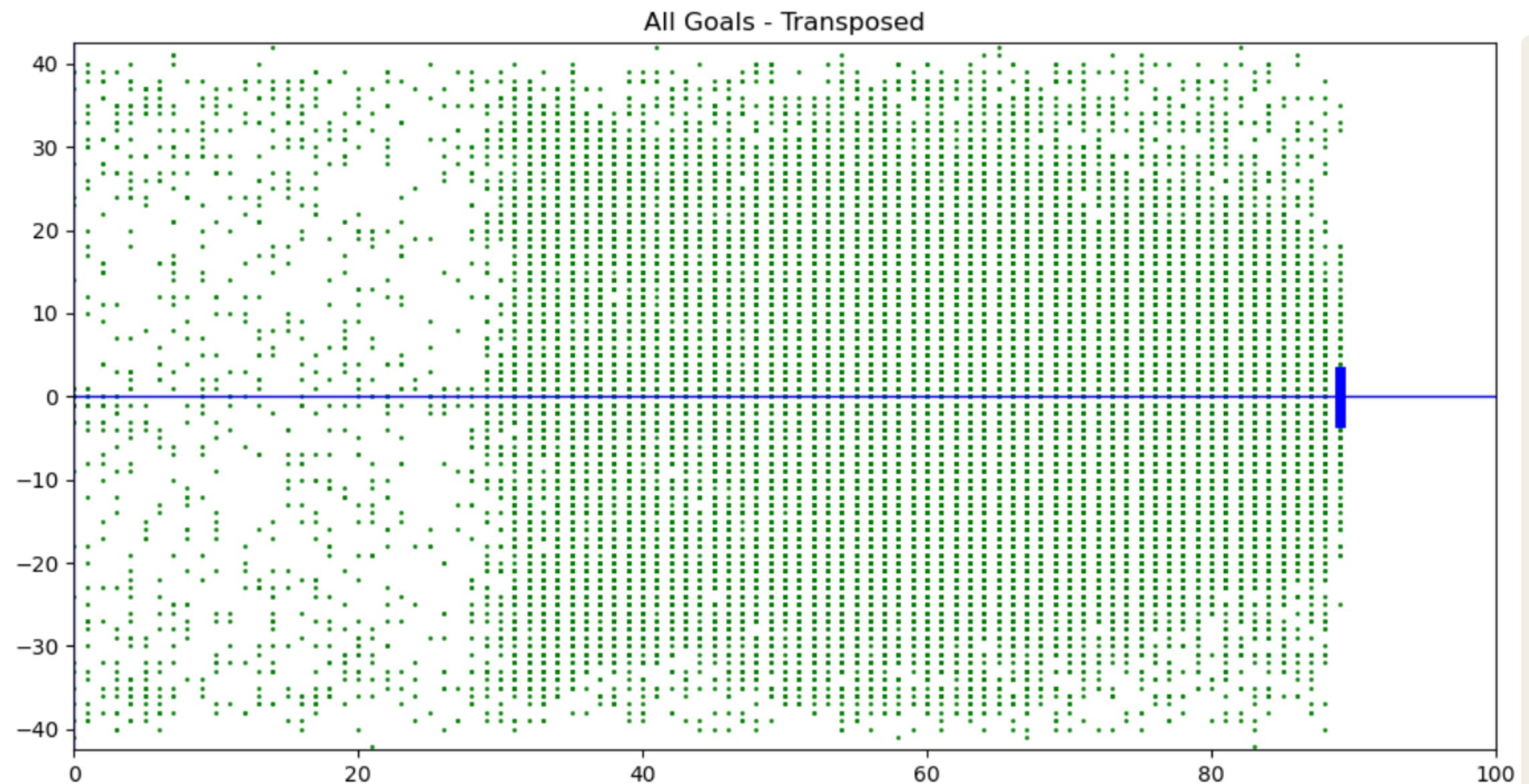


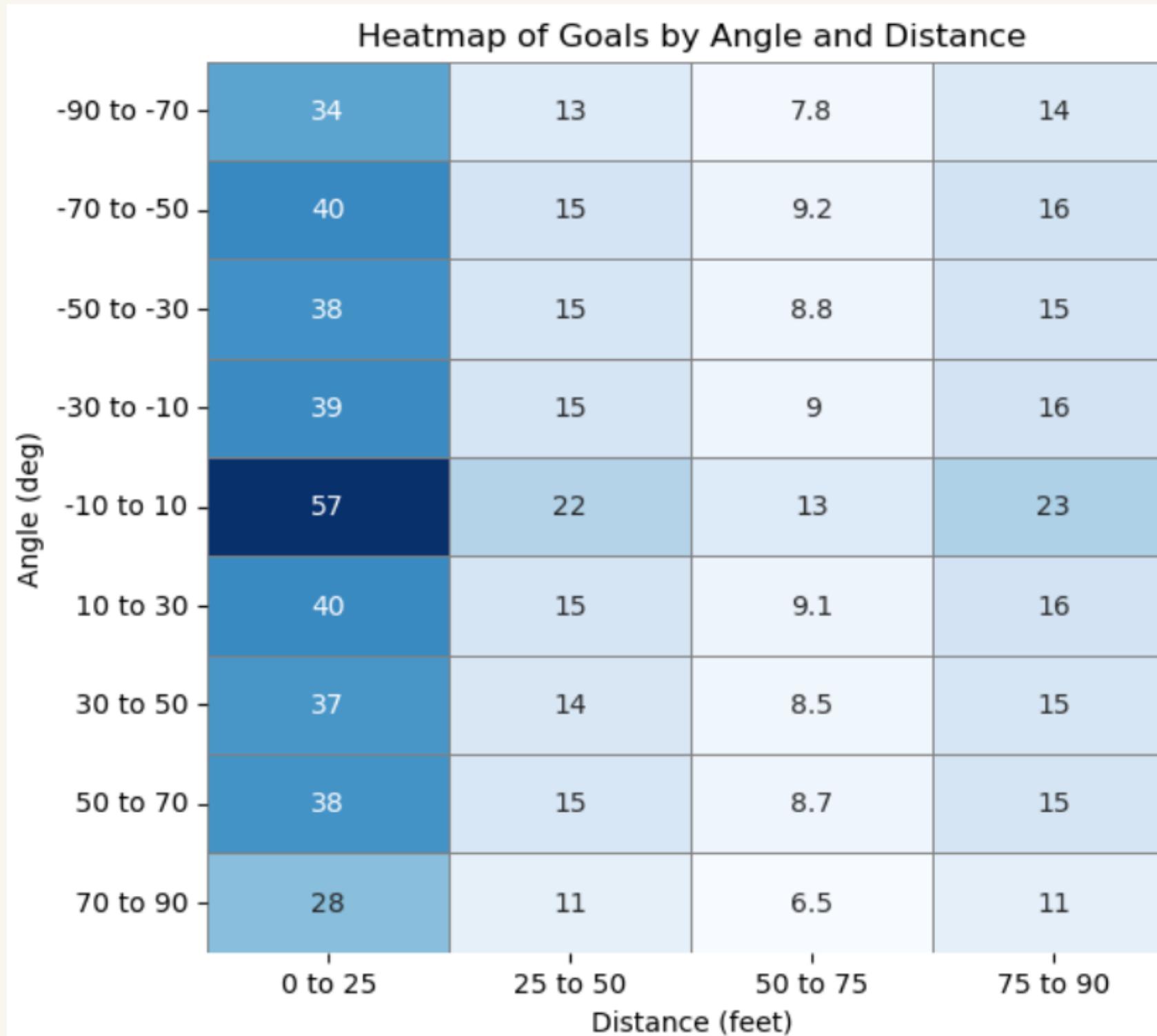
Table Transformation

game_plays Table:

- Transpose x-y to one side of the rink
- Create Goal position at x=89
- Assume all shot at goals are within x=0 to 89
- Create Angle, Distance, Angle-Group/Distance-Group
- Formulate composite index for goal probability

ANALYSIS: X-Y

FINDINGS ON GOAL DISTRIBUTION



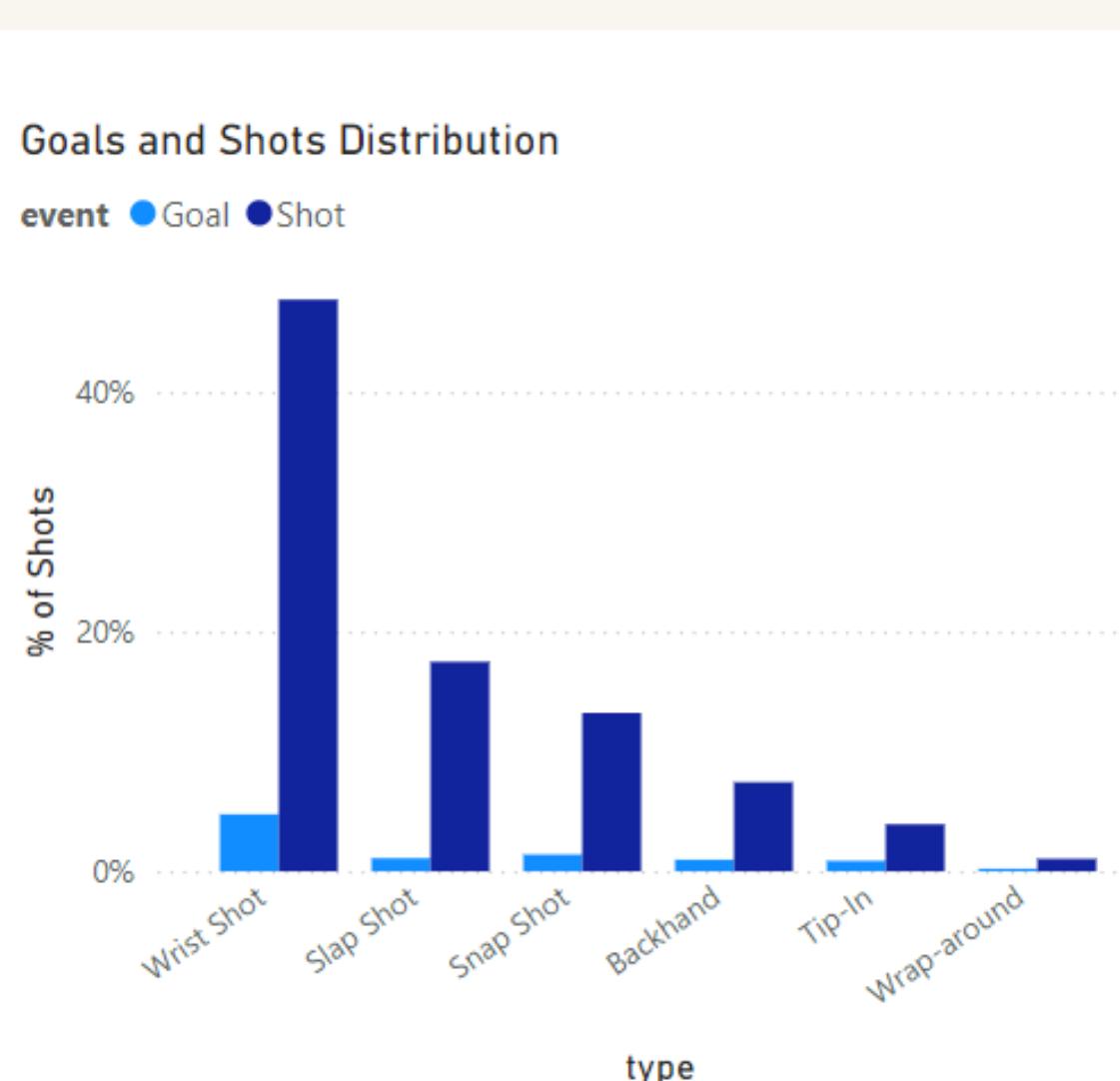
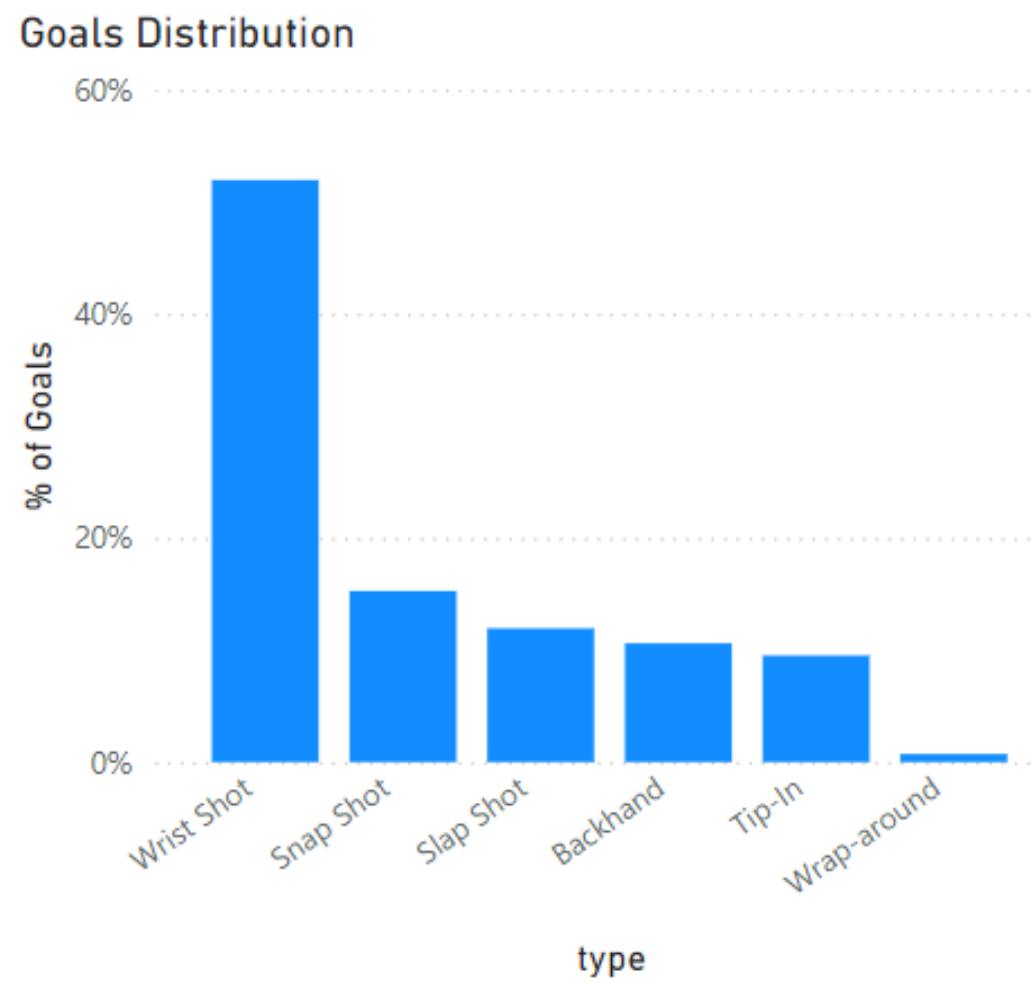
Goal Distance and Angle

Higher chance of goal at closer range and narrower angles

Generally, distance to goal has more effect than angles to goal

ANALYSIS: X-Y

FINDINGS ON SHOT TYPE DISTRIBUTION



Shot Types

Wrist Shot, Snap/Slap Shot most common

Backhand and Tip-in are less common

Wrap around is the least common

ANALYSIS: X-Y

FINDINGS ON SHOT TYPE OVER DISTANCE AND ANGLE

Tip-In

Shot Type

angle / dist 0 to 25 25 to 50 50 to 75 75 to 100

a -90 to -70	2462	14	8	0
b -70 to -50	18465	105	60	0
c -50 to -30	62781	357	204	0
d -30 to -10	264665	1505	860	0
e -10 to 10	593342	3374	1928	0
f 10 to 30	260972	1484	848	0
g 30 to 50	55395	315	180	0
h 50 to 70	13541	77	44	0
i 70 to 90	1231	7	4	0

angle

- Backhand
- Slap Shot
- Snap Shot
- Tip-In
- Wrap-around
- Wrist Shot

=



distance

- Backhand
- Slap Shot
- Snap Shot
- Tip-In
- Wrap-around
- Wrist Shot

Shot Types over Distance and Angle

Heatmap of different shot types over distance and angle

- Tip In
- Slap Shot
- Snap Shot
- Backhand
- Wrap-around
- Wrist Shot

ANALYSIS: X-Y

CONCLUSION

- Dependencies of shot types over both distance and angle from goal post
 - Distance has a greater effect than angle
- Distance from goal post
 - 0 to 25 ft: Tip-In and Backhand
 - 25 to 50 ft: Snap Shot
 - 50 to 75 ft: Slap Shot
- Angles to goal post
 - -30 to 30 deg: Tip-In
 - -50 to -90 or 50 to 90 deg: Wrap Around
 - Other shot types are relatively independent of angle to goal
- Wrist Shot
 - A versatile goal conversion shot, used in wide range of distances and angles

ANALYSIS: NPC DREAM TEAM

APPROACH

Top 23 Players (under 26)

Categorized into three groups, each group is evaluated based on their recorded performance using a composite score of 10, with the following weighted metrics.

12 Forwards

- Goals: 3 pts
- Shots: 3 pts
- Assists: 1 pt
- Time on Ice: 3 pts

8 Defensemen

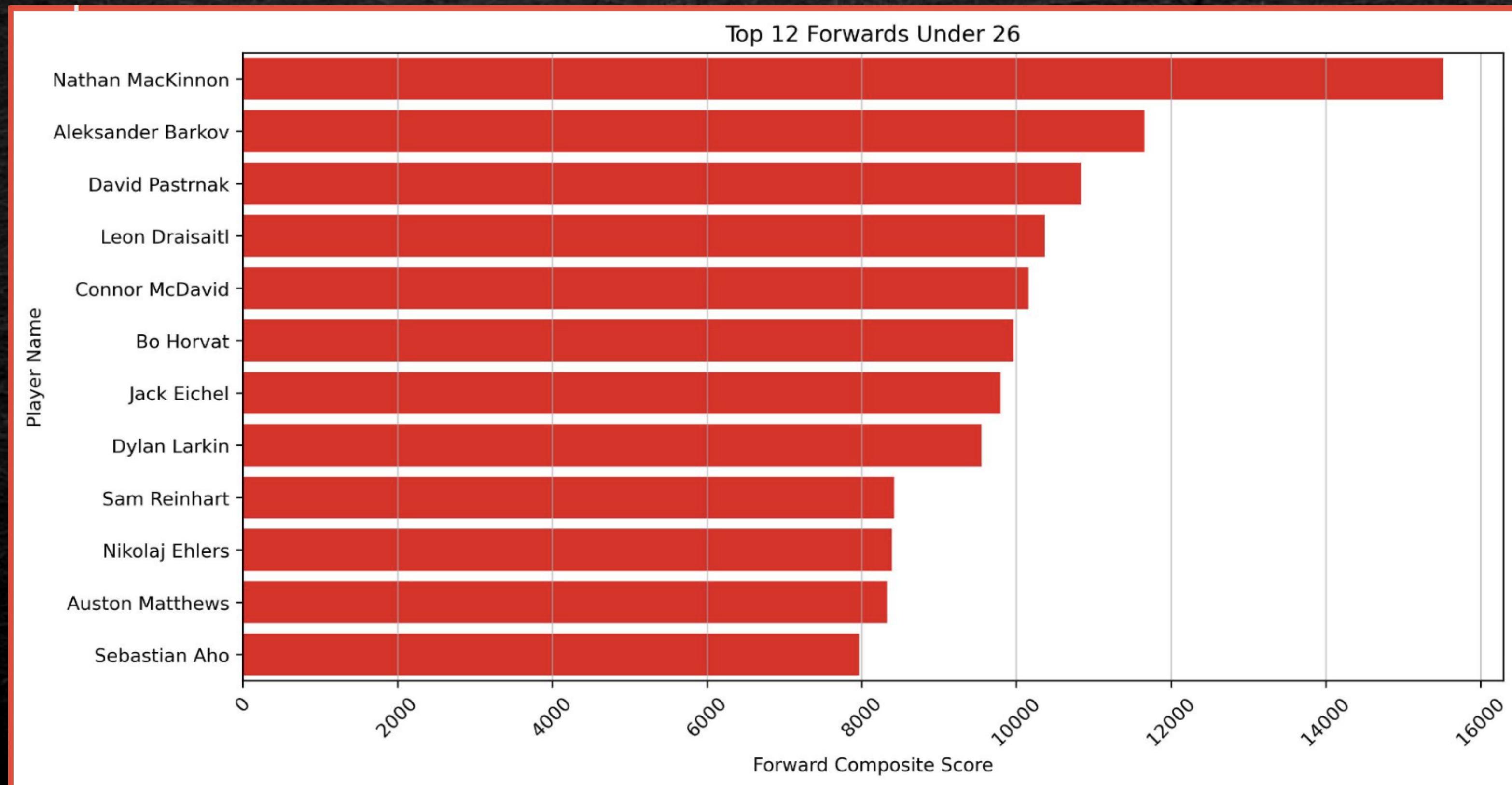
- Takeaways: 3 pts
- Blocked: 3 pts
- Assists: 1 pt
- Time on Ice: 3 pts

3 Goalies

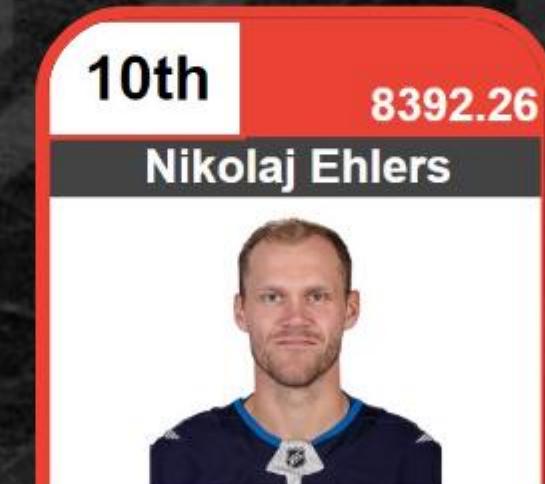
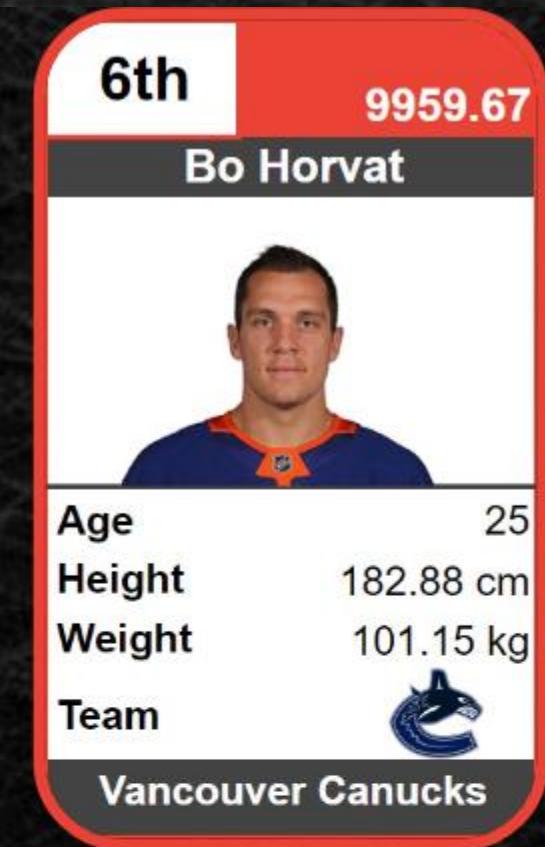
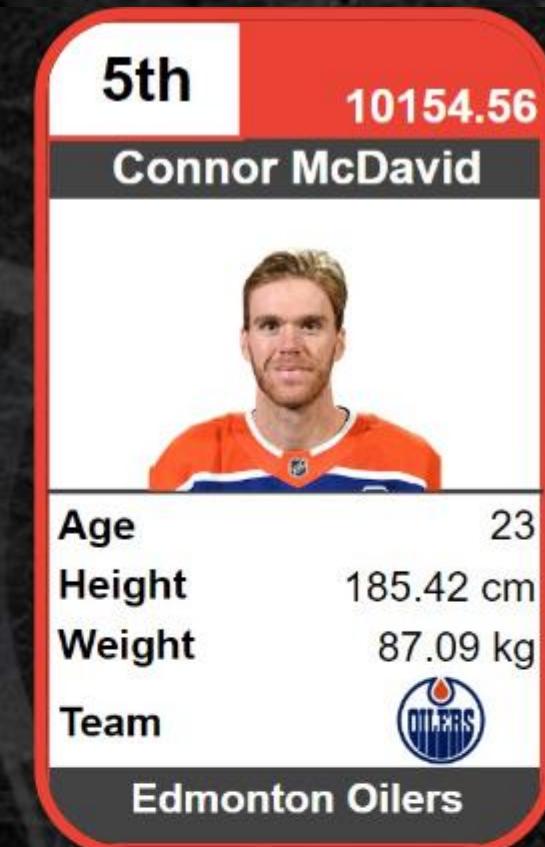
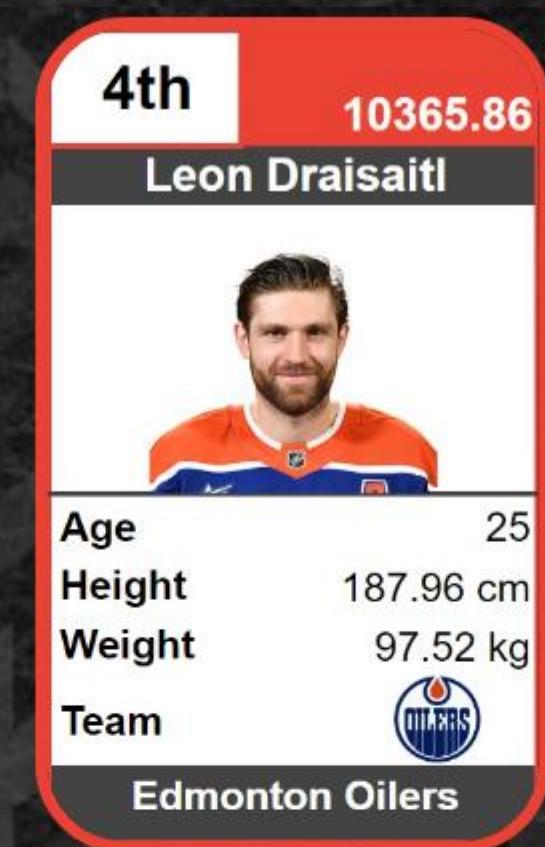
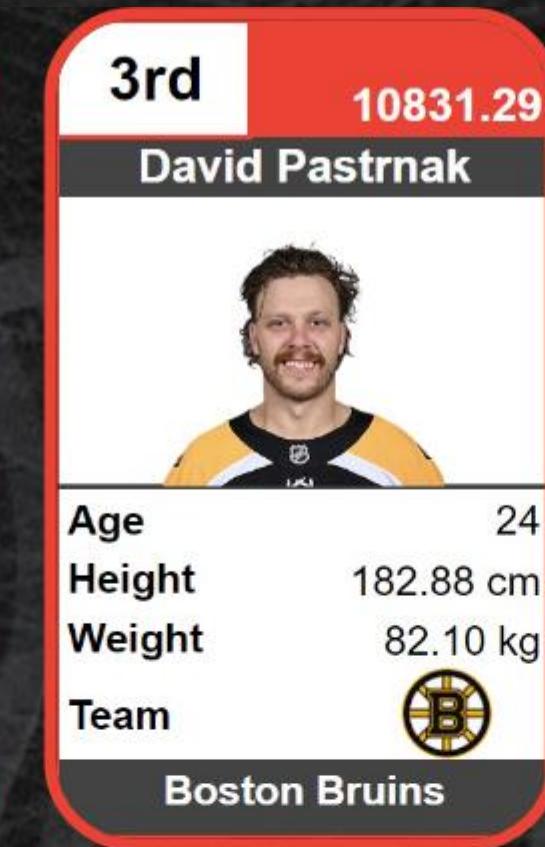
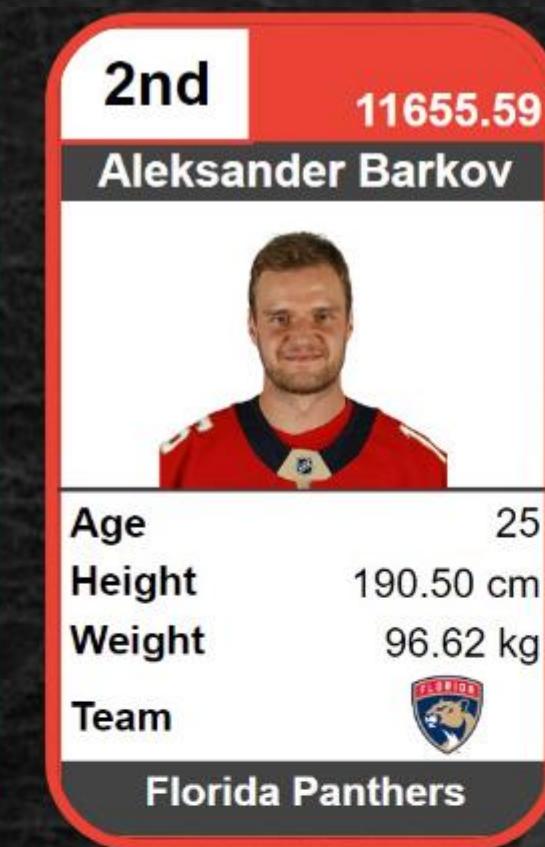
- Saves: 7 pts
- Time on Ice: 3 pts

ANALYSIS: NPC DREAM TEAM

FINDINGS : TOP 12 FORWARDS

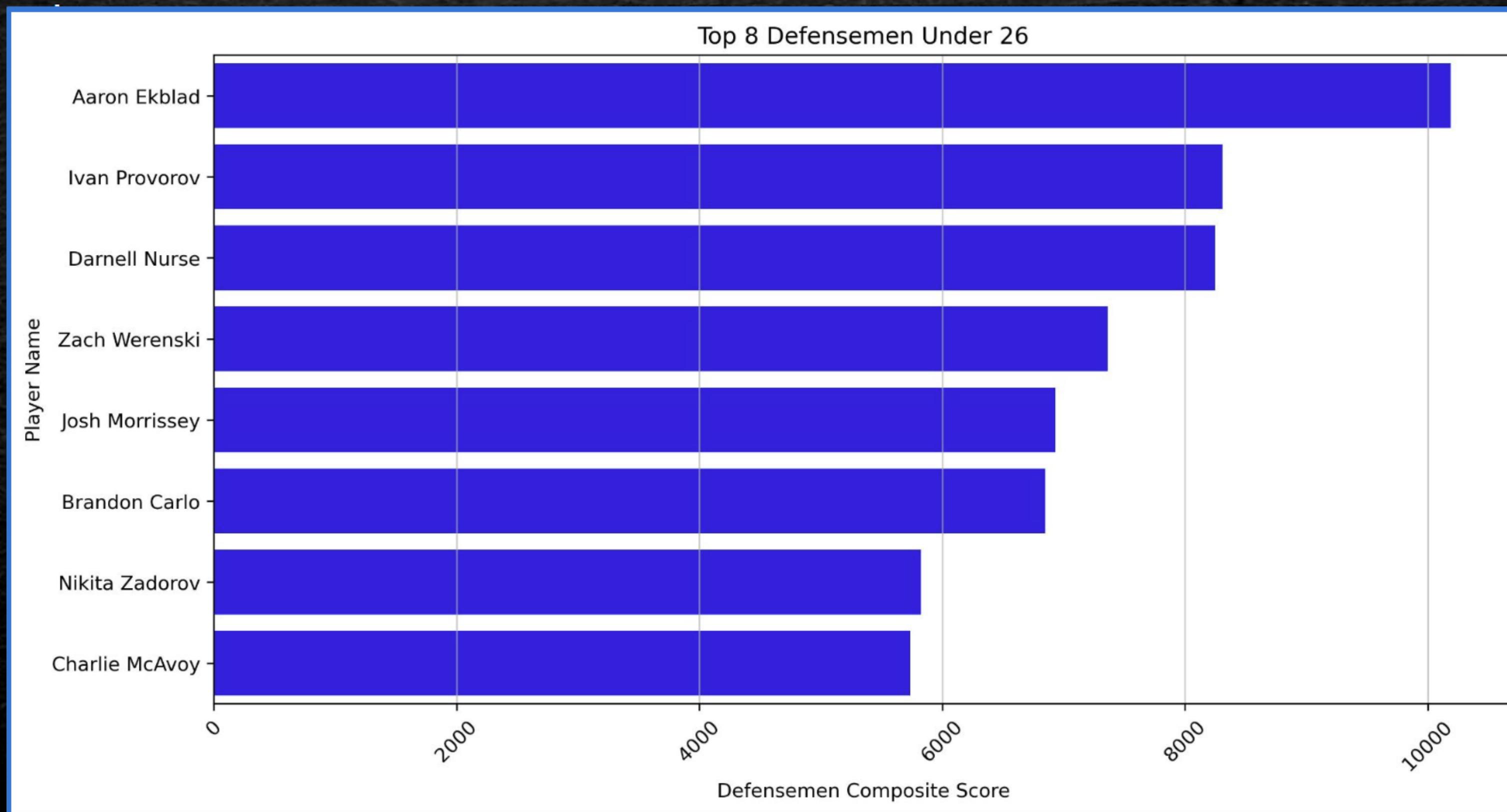


ANALYSIS: NPC DREAM TEAM

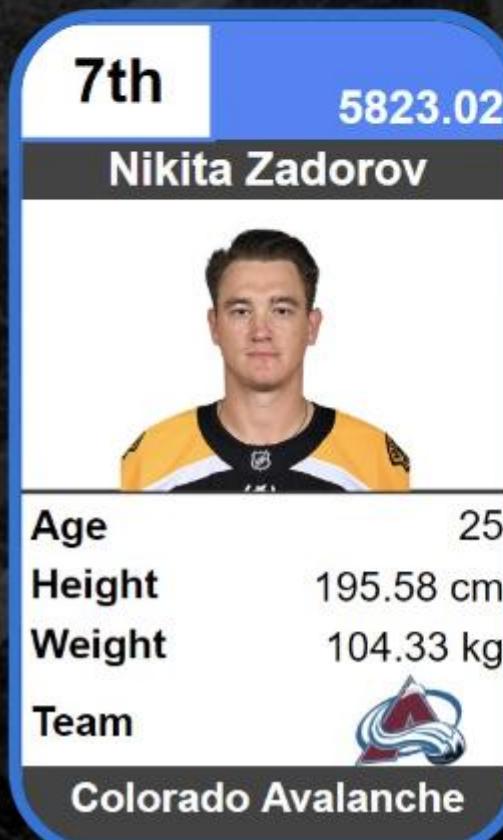
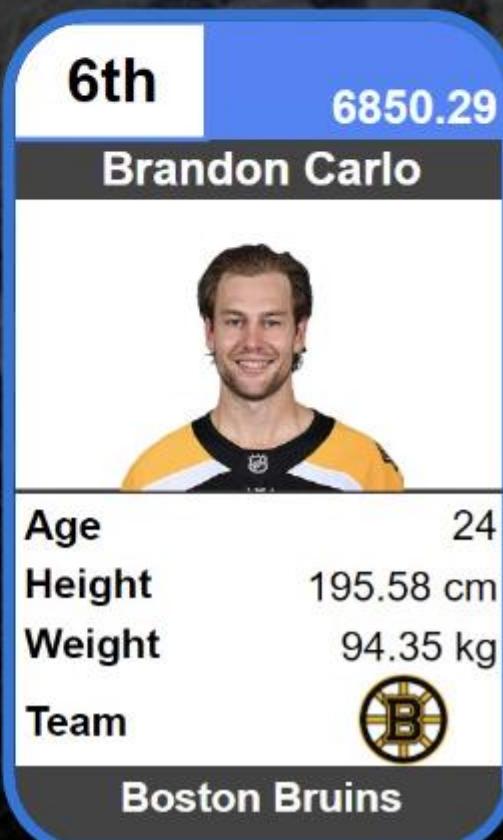
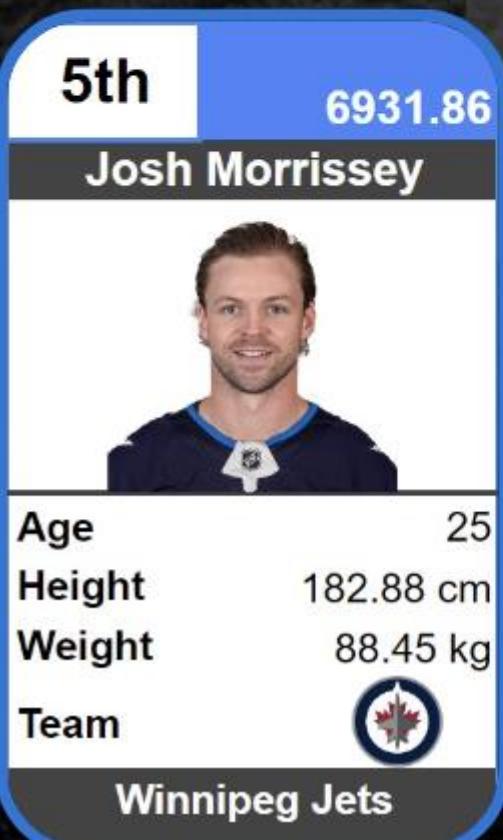
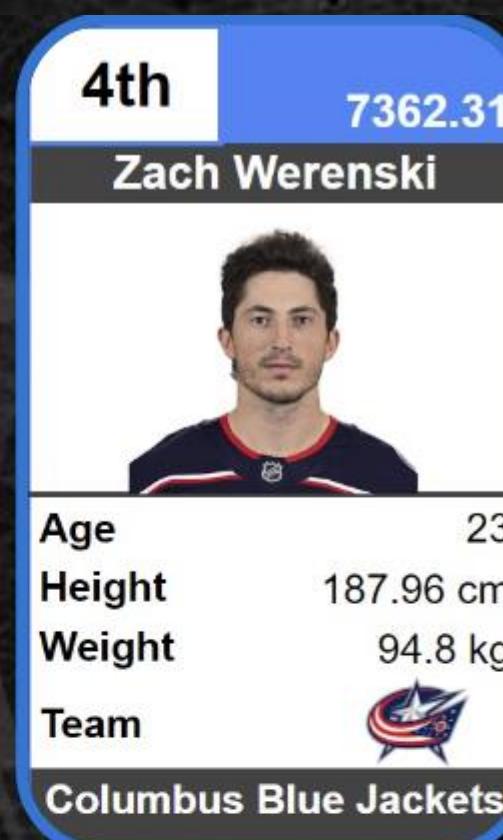
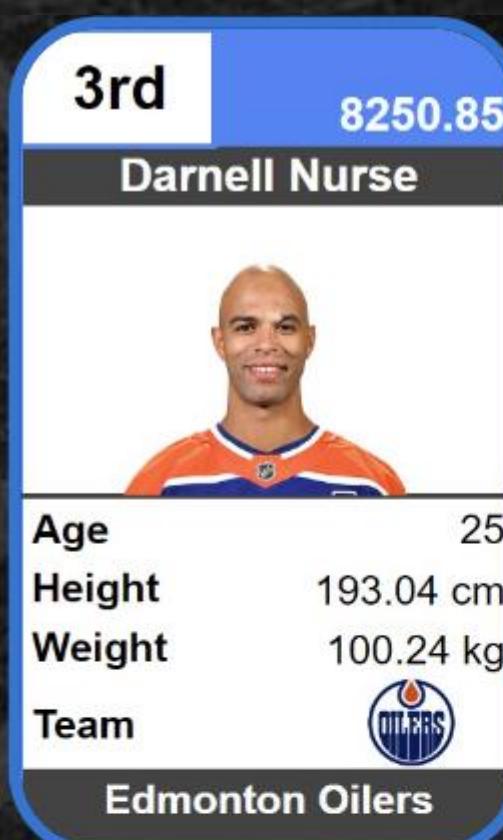
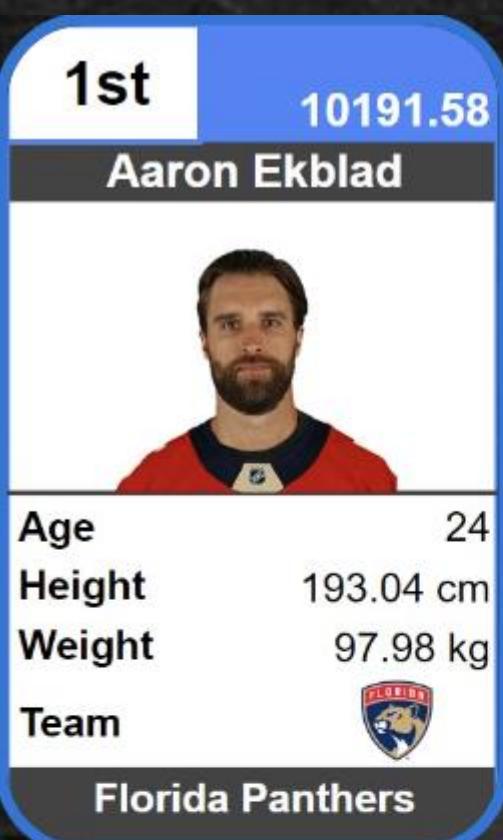


ANALYSIS: NPC DREAM TEAM

FINDINGS : TOP 8 DEFENSEMEN

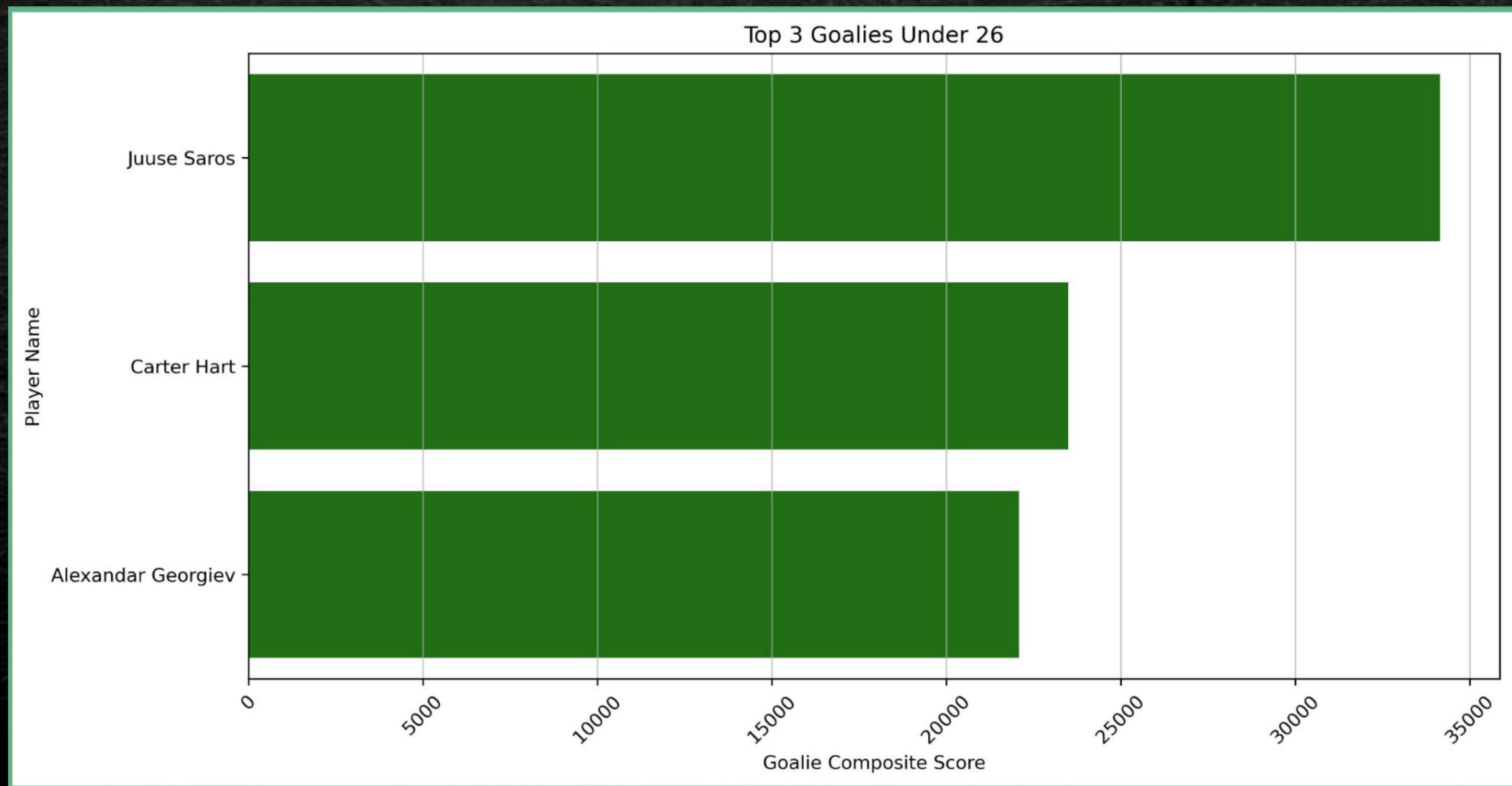


ANALYSIS: NPC DREAM TEAM



ANALYSIS: NPC DREAM TEAM

FINDINGS : TOP 3 GOALIES

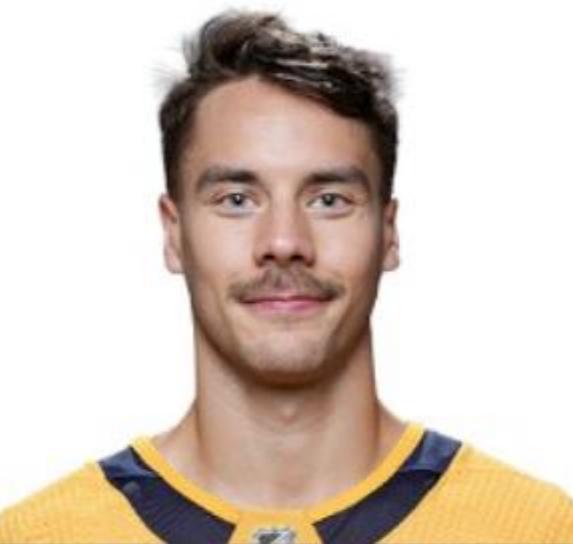


ANALYSIS: NPC DREAM TEAM

1st

34155.22

Juuse Saros



Age

25

Height

180.34 cm

Weight

81.65 kg

Team



Nashville Predators

2nd

23491.58

Carter Hart



Age

22

Height

187.96 cm

Weight

82.1 kg

Team



Philadelphia Flyers

3rd

22080.00

Alexandar Georgiev



Age

24

Height

185.42 cm

Weight

81.65 kg

Team



NY Rangers

ANALYSIS: NPC DREAM TEAM

CONCLUSION

The analysis identifies young talents in key positions with the highest potential to excel in the rankings and establishes a solid foundation for creating engaging characters in game development.

►► “Fast Forward” to today ►►

Statistics from 2021 to 2023:

- 20 of our proposed 23 players were ranked within the Top 100 in their respective fields.
- Accuracy estimated at **87%**.

CHALLENGES ... SOLUTIONS

Unfamiliar with Azure and NHL

.....

Resource Restrictions

Tight Timelines

Limited knowledge on ETL pipeline implementation in Azure

Limited knowledge of hockey rules, difficult to contextualize data and select relevant metrics

Faced budget and resource constraints due to licensing limits

Short delivery timelines

Collaborative Learning

Studied and selected appropriate services

Spent time learning NHL rules to align data.
Team shared findings to accelerate learning in Azure and NHL data.

Cost Management

Monitored cloud usage
Deactivated resources when not in use

Project Management

Defined roles, parallel task execution, and milestone tracking ensured timely project completion.

CONCLUSION

The project provided the team with important key learnings:

- 1 Building ETL Pipelines using Microsoft Azure
- 2 Data-driven Sports Analytics and Importance of domain knowledge
- 3 Collaborative Learning and Adaptability





Thank You !

Special Thanks to Christine and :

Generation
SINGAPORE

