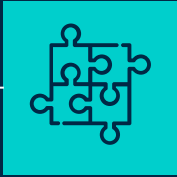# BTMA 431 Final Project: MoneyPuck

Isaac Lutzko (30026703)
Nolan Ruzicki (30132405)
Karl Specht (30120565)
Doug Strueby (30122048)

Table Of Contents

01
INTRODUCTION
Motivation, context, & the questions for our study

02
DATA COLLECTION
Data sources & how we collected the data

03
OUR FINDINGS
Summary & the importance/ implications of our findings

Project Topic & Motivation

- Final project topic: NHL player and team data from the season in 2005-06 to the current season
- As hockey fans, we were curious about what makes an NHL team successful in the regular season and in the playoffs
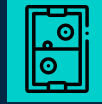- Partly inspired by the movie *Moneyball*
- Who benefits?
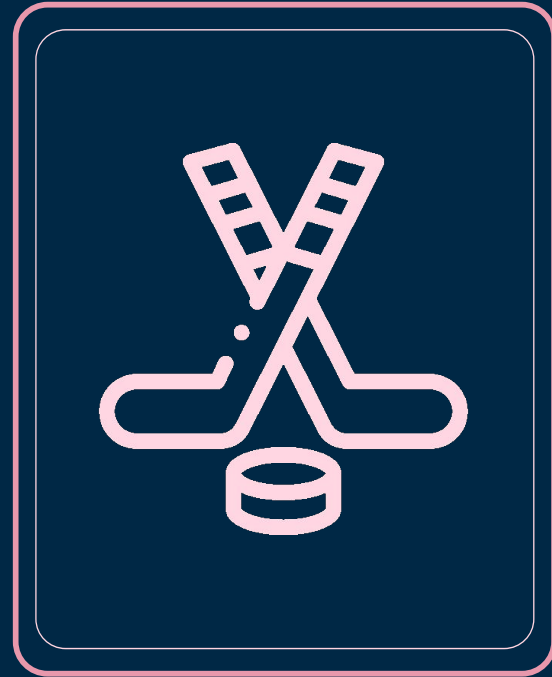
Questions for the Study

What NHL statistics are most relevant to making the playoffs?

Which NHL team has the most depth in the league during the current season (2022-23)?

What did previous Stanley Cup winning teams do that made them successful?

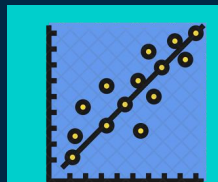What teams **should** make the playoffs?

How does our measure of depth compare to the league standings of the current season (2022-23)?

Who will win the Stanley Cup this year?

Data Source & Data Collection Process

Data Source: https://www.hockey-reference.com/

Collect player data (skaters & goalies) and team data

Web scrape with help from R packages (rvest, XML, httr, & dplyr)

Gather both regular season data and playoff data
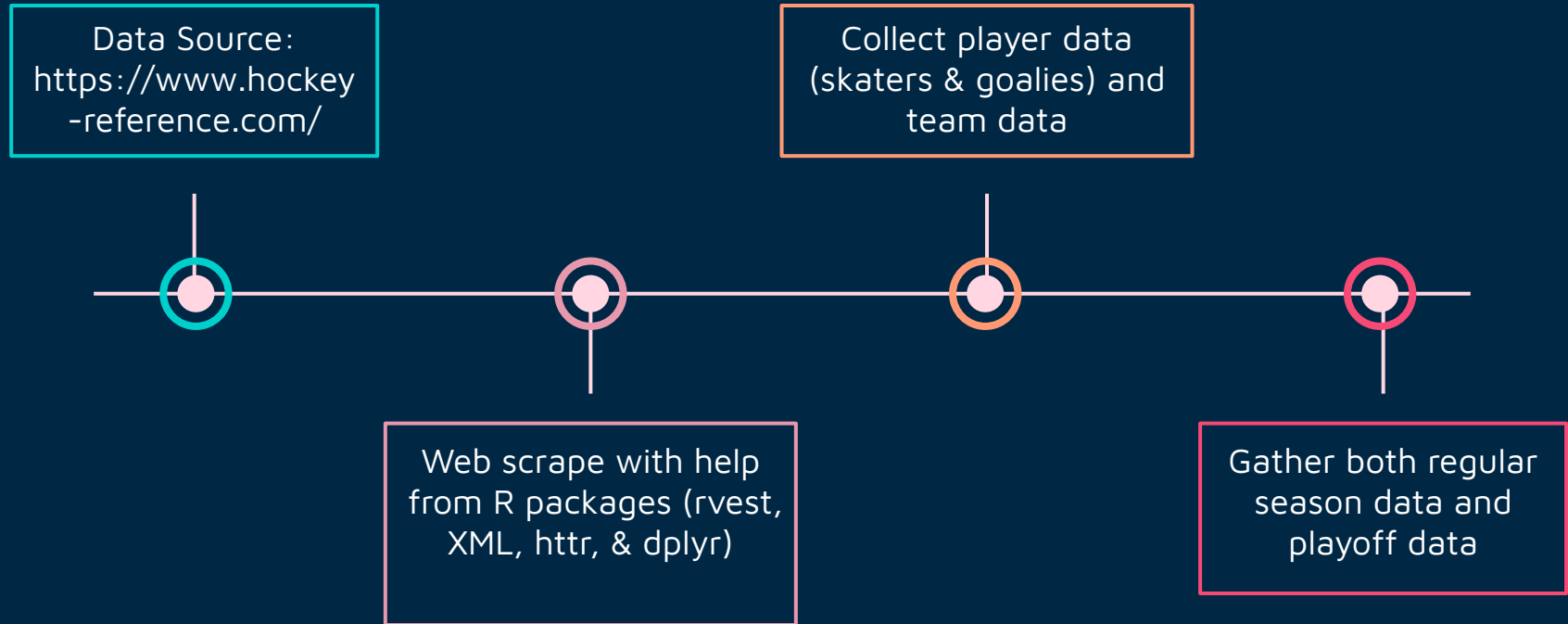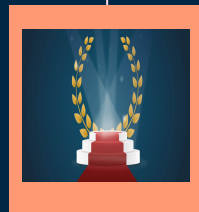
Who Makes It? And Who Will Win?

Linear regression
Trained using the data from 2006-2022

What are the Standings
The top 16 team will make the playoffs in this model

What factors we most important
The most significant stats should be a team main focus

Who Will Win the Cup
When running a simulation who will in the cup?

What Stats are most significant to making the playoffs

Removed Variables

Too obvious
W, L, OT, PT%, etc.

Changed
GF/G, GA/G, SV%,
PK%, and PP%

```
Coefficients:
              Estimate  Std. Error  t value  Pr(>|t|)
(Intercept)   547.8290    495.9348    1.105   0.26981
AvAge           1.1976      0.4651    2.575   0.01029 *
PIM.G           0.3576      0.5392    0.663   0.50751
oPIM.G          0.5280      0.5617    0.940   0.34758
SV.          -460.6209    495.0773   -0.930   0.35258
GD.G           44.7546     16.3032    2.745   0.00625 **
SD.G           -1.7357      1.5374   -1.129   0.25941
S.           -672.7012    498.1979   -1.350   0.17750
PK.           -31.0659     20.4732   -1.517   0.12976
PP.             7.7861     20.2788    0.384   0.70117
```

Predicted League Standings

| Using linear regression we were able to predict the standings	Top 16 team make the playoffs

	Team name	Points
1	Boston Bruins	106
2	New York Islanders	94
3	Minnesota Wild	93
4	Nashville Predators	92
5	Colorado Avalanche	92
6	Dallas Stars	91
7	Vegas Golden Knights*	90
8	Winnipeg Jets	89
9	New York Rangers	89
10	Toronto Maple Leafs	87
11	Tampa Bay Lightning	84
12	New Jersey Devils	84
13	Pittsburgh Penguins	84
14	Carolina Hurricanes	83
15	Washington Capitals	80
16	Edmonton Oilers	80
17	Arizona Coyotes	80
18	Florida Panthers	78
19	Anaheim Ducks	78
20	Ottawa Senators	77
21	Philadelphia Flyers	76
22	Detroit Red Wings	75
23	Chicago Blackhawks	75
24	Montreal Canadiens	74
25	Buffalo Sabres	73
26	Los Angeles Kings	72
27	Calgary Flames	71
28	St. Louis Blues	70
29	Columbus Blue Jackets	69
30	Seattle Kraken	69
31	Vancouver Canucks	65
32	San Jose Sharks	65

NHL Playoff bracket
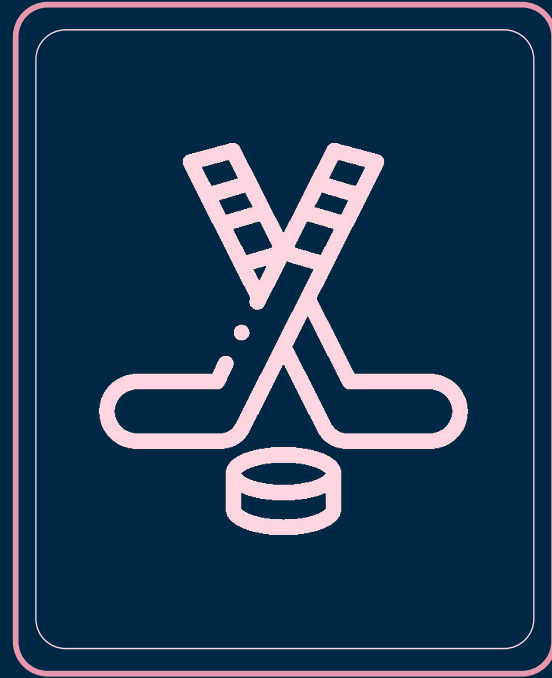
Using simulation

Based on our Model the Champs are the Winnipeg jets

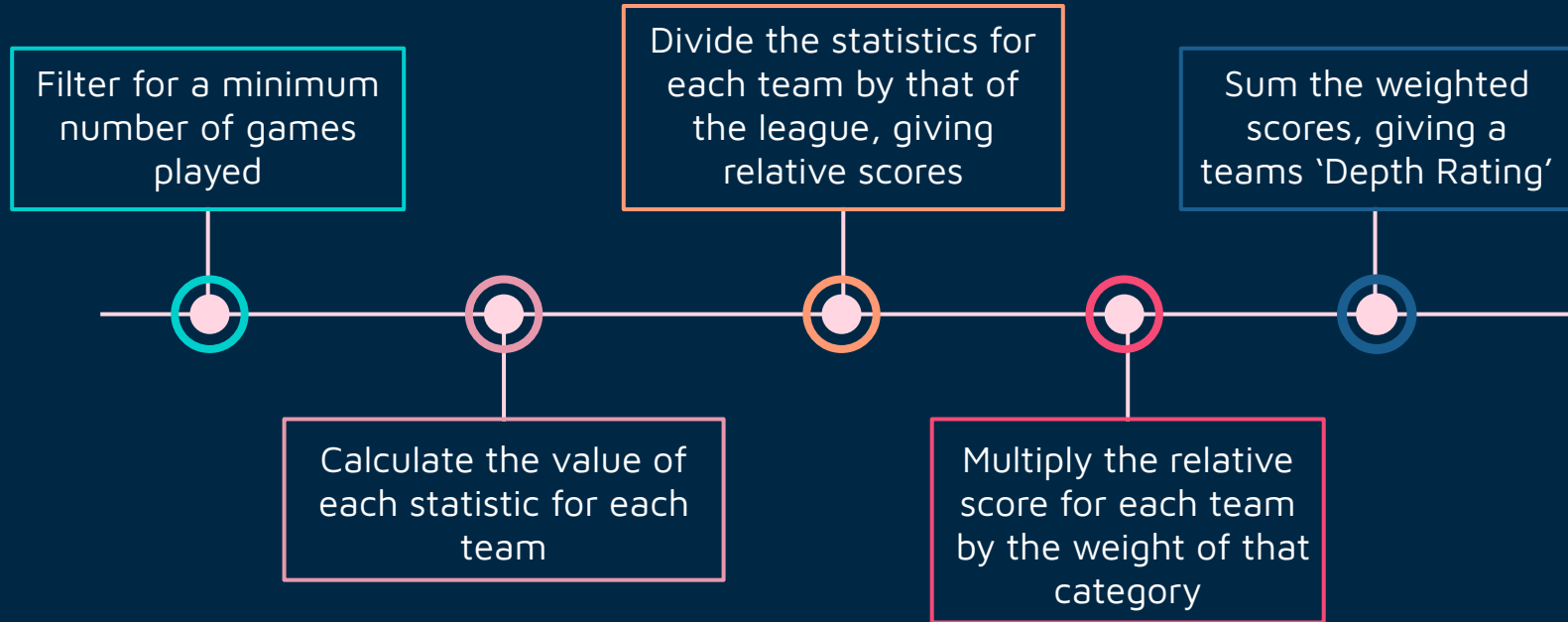Seven game series first to four wins moves on
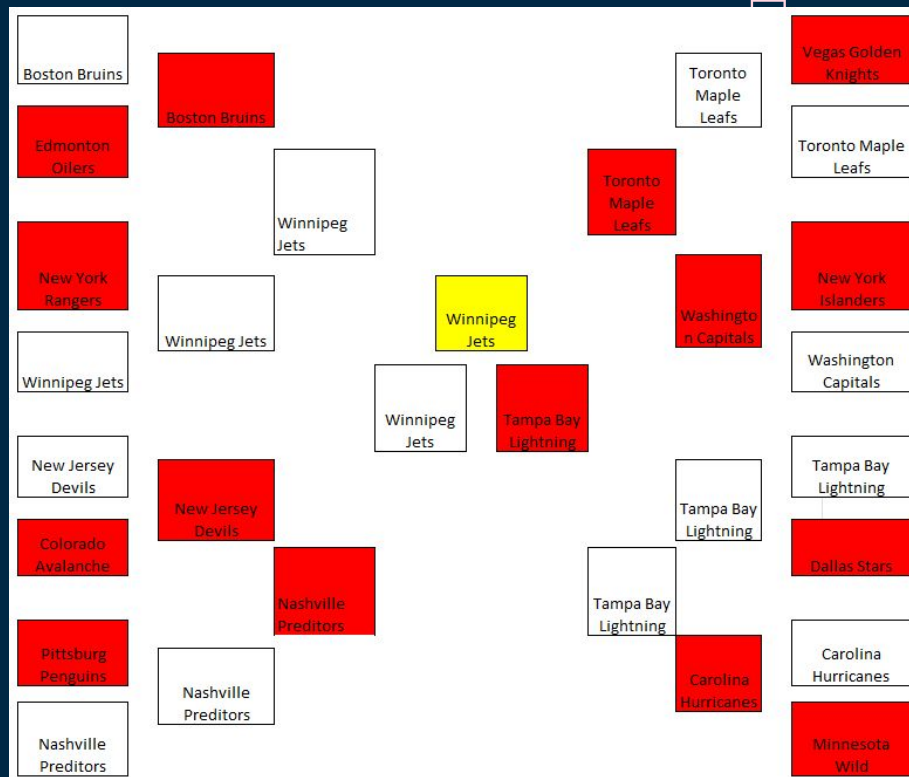
Defining Depth

- High average player

- Low variation

- We defined based on the 6 factors of Goals, Assists, Shots on Goal, Average time on Ice, Hits, and Blocks

Calculating Depth

Filter for a minimum number of games played

Calculate the value of each statistic for each team

Divide the statistics for each team by that of the league, giving relative scores

Multiply the relative score for each team by the weight of that category

Sum the weighted scores, giving a teams 'Depth Rating'

Different Methods Used

Mean

- Simple
- Does not account for high variance

Mean Minus Standard Deviation

- Heavily accounts for variance
- In practice, very inaccurate

Median

- Less straightforward than mean
- Better at accounting for variance

Mean Divided by Standard Deviation

- Heavily accounts for variance
- Much more accurate than Mean Minus SD

Depth Rating vs. Actual Rating Across All Methods
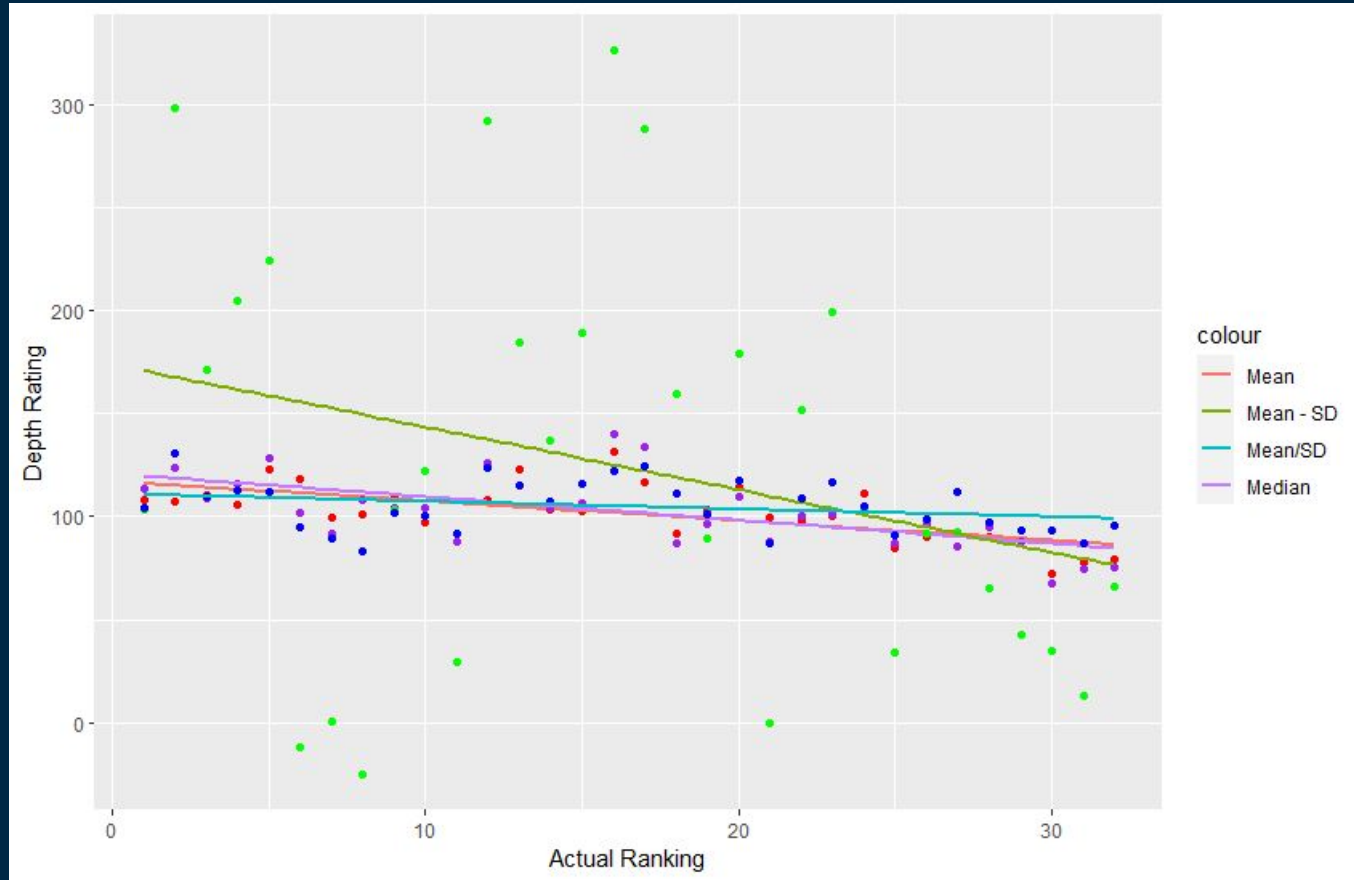
Depth Rating vs. Actual Rating Across the Top 3 Methods
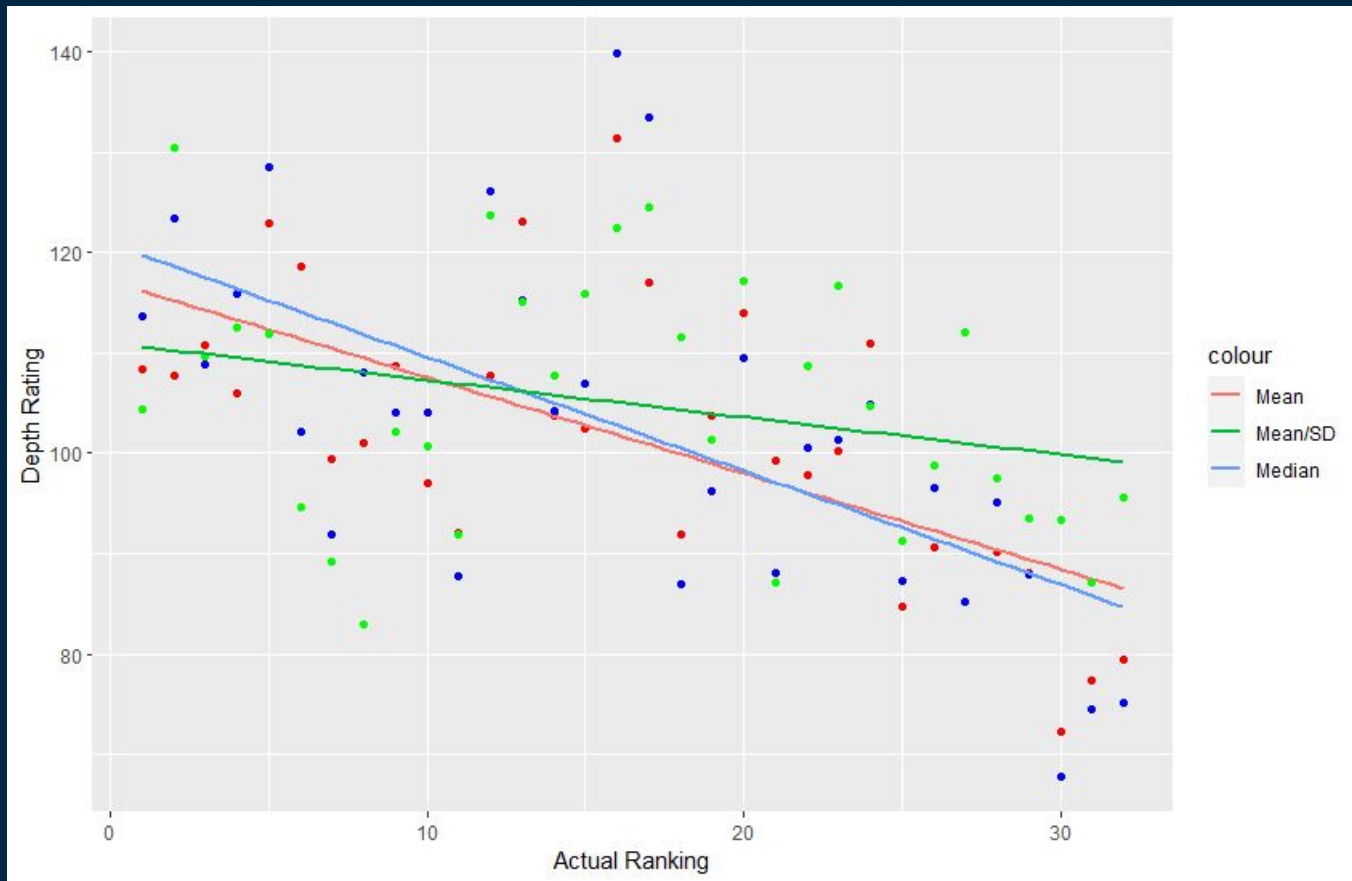
Defining Data Frames

Championships

Playoffs - Mean per game
Regular Season - Mean of all games

Rest of the League

Playoffs - Mean per game
Regular Season - Mean of all games

Differences Between Champions and Rest of the League

Higher Plus Minus
+0.25 higher per game in the playoffs

More Shots per Game
Average of 6.7 more shots in the regular season

More Efficient Power Plays
Average of 0.16 more powerplay goals in the regular season

Less Penalty Minutes per Game
0.5 less penalty minutes on average in the regular season
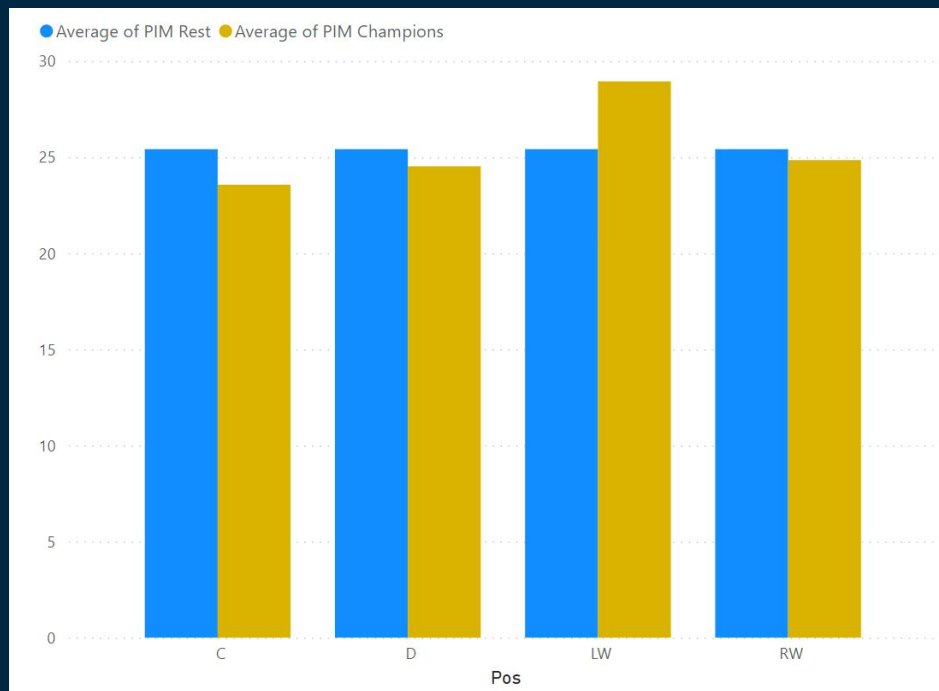
Average Games by Shots

Non-Stanley Cup Champions
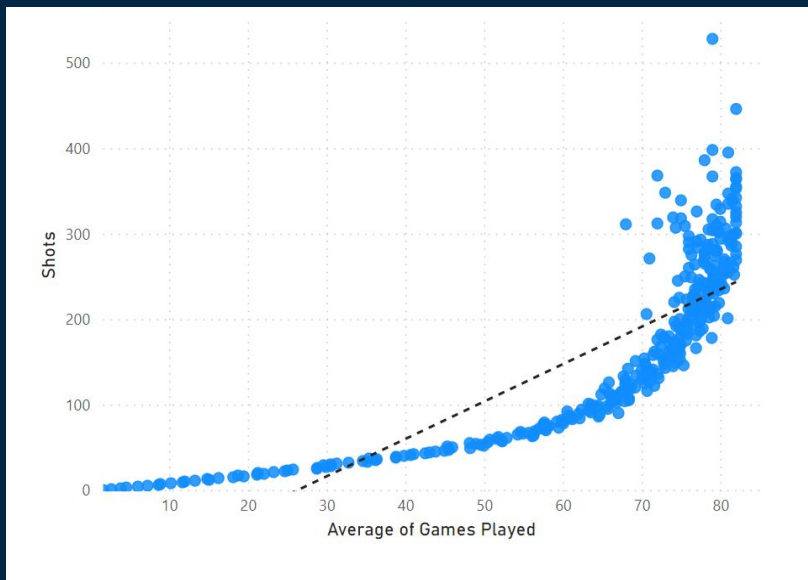
Stanley Cup Champions
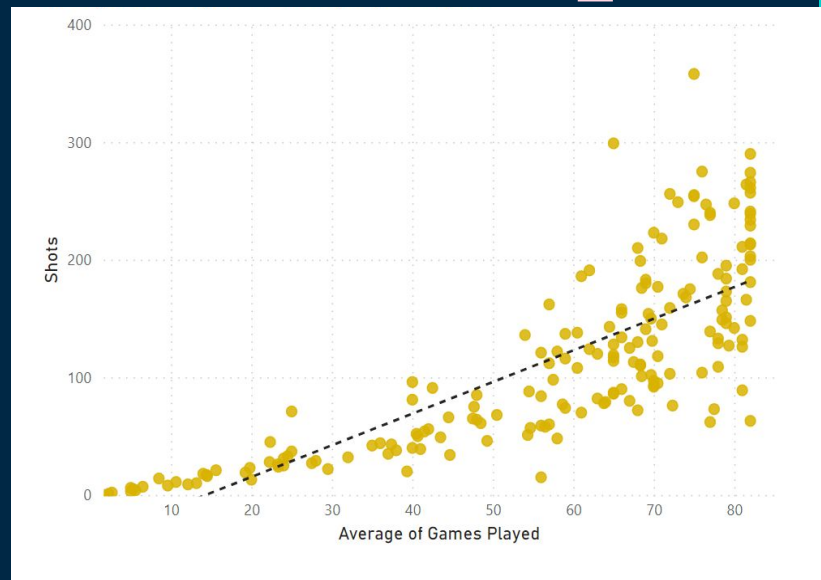
Average Penalty Minutes by Position

QUESTIONS?

Do you have any questions?

youremail@freepik.com
+91 620 421 83
yourcompany.com

THANKS