Ishockey

Ishockey - EDA

1. Hent koden og data-sættet fra https://www.kaggle.com/code/madison88/nhl-expected-goals-model-and-heat-maps
2. Find den function der tegner banen og gem den i et script som kan sources i andre scripts
3. Hent Viktors data ind i et nyt script og plot hans data (shot_datasæt.xlsx) – tjek koordinater
4. Fra hvilken zone scores der flest mål?
5. Sammenhold vinkel og zone. Er der noget der undrer? (tjek zone 17)
6. Find formlen der beregner vinkel fra skudposition til målet

Ishockey - retrieval

1. Hent alle kampe i de efterfølgende sæsoner (én json pr kamp)
2. Erstat den første ”id” med ”_id” mhp import til mongo
3. Importer alle json-filer til mongoDB
4. Lav en query i mongo der finder all distincte event-types.

```
[
  'blocked-shot', 'delayed-penalty',
  'faceoff',      'game-end',
  'giveaway',     'goal',
  'hit',          'missed-shot',
  'penalty',      'period-end',
  'period-start', 'shootout-complete',
  'shot-on-goal', 'stoppage',
  'takeaway'
]
```


```
{
  "id": 2023020358,
  "season": 20232024,
  "gameType": 2,
  "gameDate": "2023-12-02",
  "venue": {
    "default": "Bridgestone Arena"
  },
  "startTimeUTC": "2023-12-02T21:30:00Z",
  "easternUTCOffset": "-05:00",
  "venueUTCOffset": "-06:00",
  "tvBroadcasts": [
    {
      "id": 27,
```

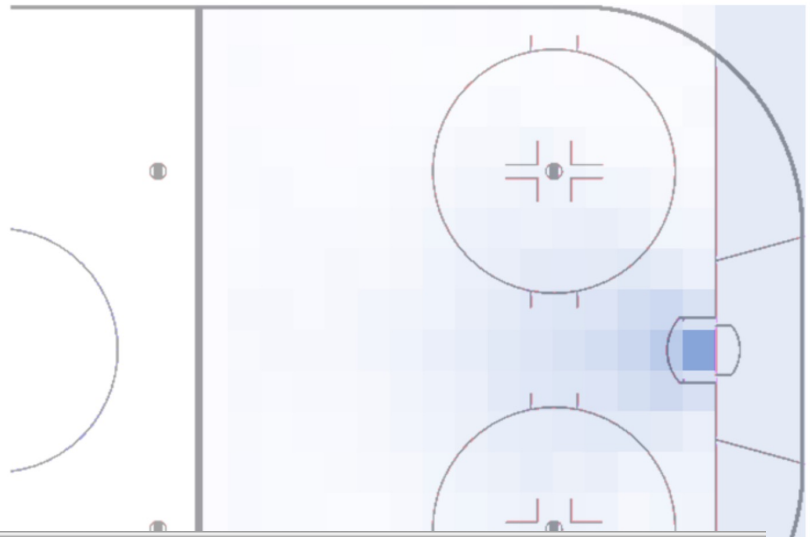

Alle kamp-id’er i sæsonen fra :
https://api-web.nhle.com/v1/club-schedule-season/WSH/20222023

Stats for hver kamp:
https://api-web.nhle.com/v1/gamecenter/2023020358/play-by-play

https://gitlab.com/dword4/nhlapi/-/blob/master/new-api.md

xG

All **All** **FSh%**



	A	B	C	D
1	Box Nr.	xG_base	x coordinate	y coordinate
2	1	0,01166	29 <= x < 33	37.5 < y < 42.5
3	2	0,01888	29 <= x < 33	32.5 < y < 37.5
4	3	0,01246	29 <= x < 33	27.5 < y < 32.5
5	4	0,01732	29 <= x < 33	22.5 < y < 27.5
6	5	0,01830	29 <= x < 33	17.5 < y < 22.5
7	6	0,01568	29 <= x < 33	12.5 < y < 17.5
254	253	0,01880	85 <= x < 89	-32.5 < y < -27.5
255	254	0,01536	85 <= x < 89	-37.5 < y < -32.5
256	255	0,02778	85 <= x < 89	-42.5 < y < -37.5
257	256	0,00856	25 <= x < 29	
258	257	0,01221	x < 25	
259	258	0,08595	x > 89	

I've decided to categorize shots in 6 strength states:

1. 5v5
2. 4v4
3. 3v3
4. PPv4 (5v4)
5. PPv3 (5v3 or 4v3)
6. SH

xG Base

1. Scraped all Play-By-Play data from 2007 to 2022 (both regular season and playoffs) using Harry Shomer's PBP scraper.
2. Removed all shot attempts that doesn't have a shot location (mostly shots from the early seasons).
3. Divided the ice into small boxes.
4. Counted 5v5 unblocked shot attempts (fenwick) and goals in each box.
5. Calculated 5v5 FSh% in each box (*FSh% = Goal/Fenwick*). This is my xG_base!

Building the 5v5 model:

In the first version of this xG model I will adjust based on the following parameters:

- Rebound shot (Shot taken within 2 seconds of the previous shot is defined as a rebound shot)
- Rush shot (Shot taken within 4 seconds of an event from the neutral or defensive zone)
- Shot type (e.g., Slap shot, wrist shot, etc.)
- Score state (e.g., trailing by 2, leading by 1, tied, etc.)
- Rink Bias (Differences from arena to arena)
- Season (Differences from season to season, for example due to smaller goalie equipment)