

# Analysis of NHL Stats

Jase Estrin

July 25, 2024

## 1 Introduction

### 1.1 Analysis Overview

The purpose of this analysis is to determine the statistical similarities between the past three Stanley Cup champions in the NHL. This analysis compares winning teams to each other to find correlations in statistics that would come across as unexpected to one who watches the NHL.

### 1.2 Teams

The three primary teams used in this analysis are the three Stanley Cup Champions for the 2021-22, 2022-23, and 2023-24 seasons respectively. The teams are listed below:

- 2021-22 Stanley Cup Champions: Colorado Avalanche [COL]
- 2022-23 Stanley Cup Champions: Vegas Golden Knights [VGK]
- 2023-24 Stanley Cup Champions: Florida Panthers [FLA]

### 1.3 Advanced Stats

Within this analysis are three stats that one must know the meaning of. These are `xGoals%`, `corsi%`, and `fenwick%`.

### 1.3.1 xGoals%

The xGoals% stat is used to measure the quality of a team's scoring chances, otherwise known as expected goals. Simply put, the higher the number is, the greater the chance that a scoring opportunity will result in a goal.

### 1.3.2 corsi%

The corsi% stat is a measure of the quantity of shots. It is calculated by taking the number of shot attempts a team has and dividing that by the sum of shot attempts for and against. This includes blocked shots. A corsi percentage over 50% implies that a team creates more chances than it gives up.

### 1.3.3 fenwick%

The fenwick% stat quantifies how much a team controls the play. It is calculated with the same formula as corsi, though fenwick disregards blocked shots. Just like corsi, a percentage of over 50% shows that a team is controlling the play.

## 2 Stat Correlation

This portion of the analysis will focus on more simple stats, such as the total goals a team scored all season. There are three categories of heatmaps created: offense, defense, and miscellaneous. The measured statistics for each category are listed below:

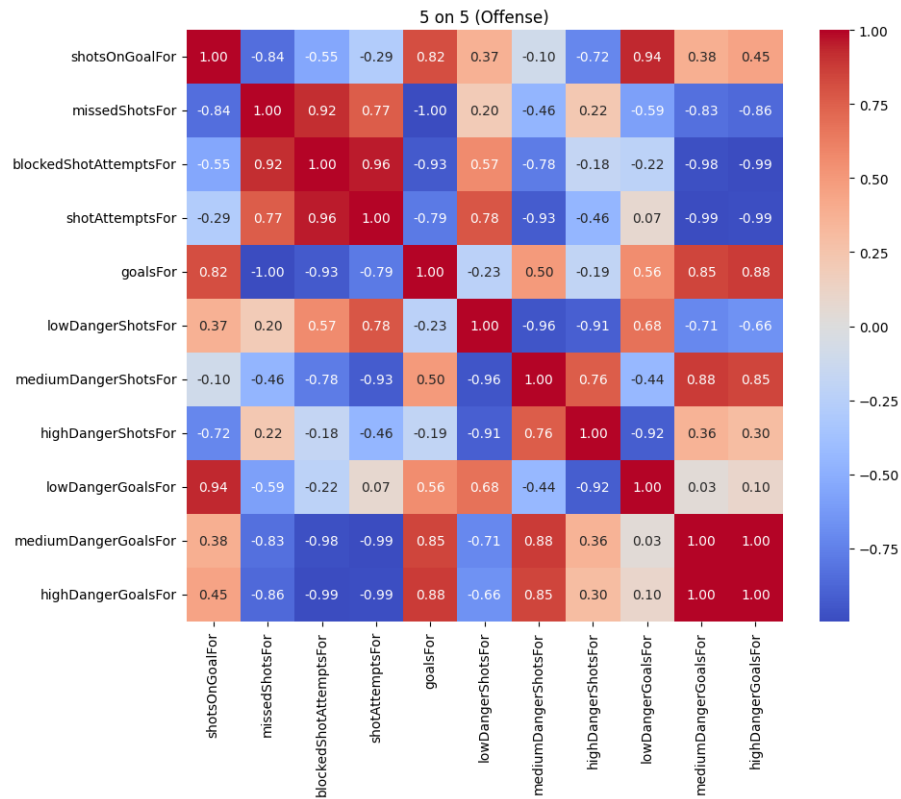
- Offense: xGoalsFor, shotsOnGoalFor, missedShotsFor, blockedShotAttemptsFor, shotAttemptsFor, goalsFor, lowDangerShotsFor, mediumDangerShotsFor, highDangerShotsFor, lowDangerGoalsFor, mediumDangerGoalsFor, highDangerGoalsFor
- Defense: shotsOnGoalAgainst, missedShotsAgainst, blockedShotAttemptsAgainst, shotAttemptsAgainst, goalsAgainst, savedShotsOnGoalAgainst, lowDangerShotsAgainst, mediumDangerShotsAgainst, highDangerShotsAgainst, lowDangerGoalsAgainst, mediumDangerGoalsAgainst, highDangerGoalsAgainst
- Miscellaneous: penaltiesFor, penaltyMinutesFor, faceOffsWonFor, hitsFor, takeawaysFor, giveawaysFor, penaltiesAgainst, penaltyMinutesAgainst, faceOffsWonAgainst, hitsAgainst, faceOffsWonAgainst, hitsAgainst, takeawaysAgainst, giveawaysAgainst

The data used to create these graphs consists of stats from each of the three aforementioned teams within a single combined dataset. the graphs measure the correlation between different statistics on a scale from -1 to 1. The higher the value, the higher the correlation between the two statistics.

Each graph will be divided by scenarios such as even strength, power play, and penalty kill. Additionally, graphs that measure the values from all scenarios combined, as well as all other scenarios not mentioned (5 on 3, 3 on 5, 4 on 4, 3 on 3, etc.).

## 2.1 Offense

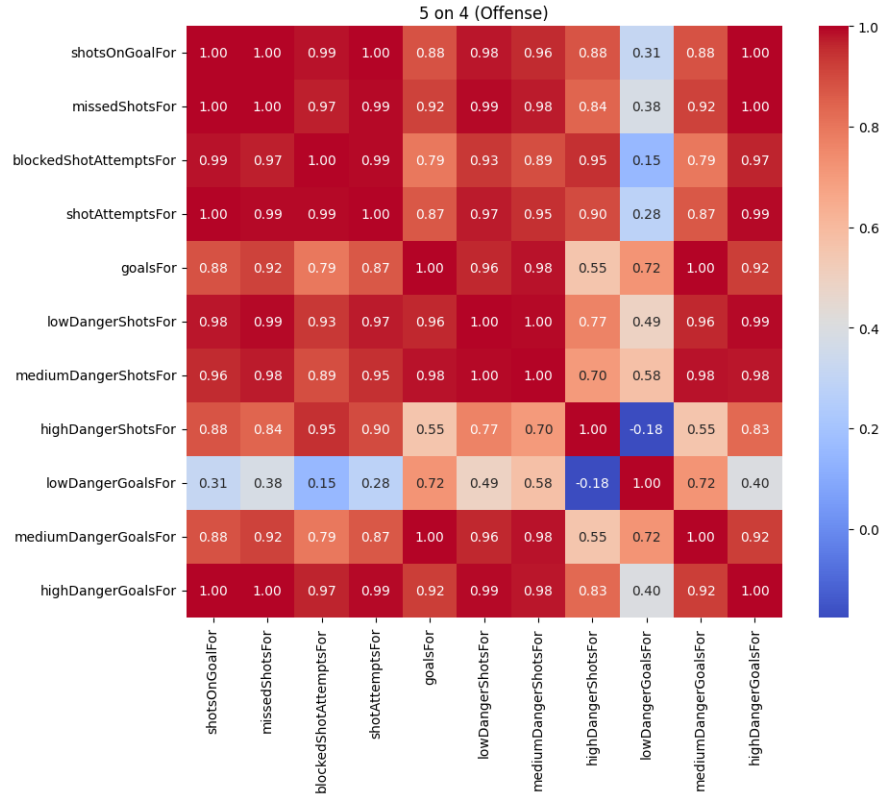
### 2.1.1 Even Strength



Notable correlations:

- missedShotsFor & shotAttemptsFor (positive correlation): Greater volume in shots that miss the net implies a greater volume in overall scoring attempts.
- highDangerShotsFor & shotAttemptsFor (negative correlation): Implies that teams that take more high-danger shots end up with less shot attempts overall. This proves that teams who focus on making the perfect play end up with less to show for it.

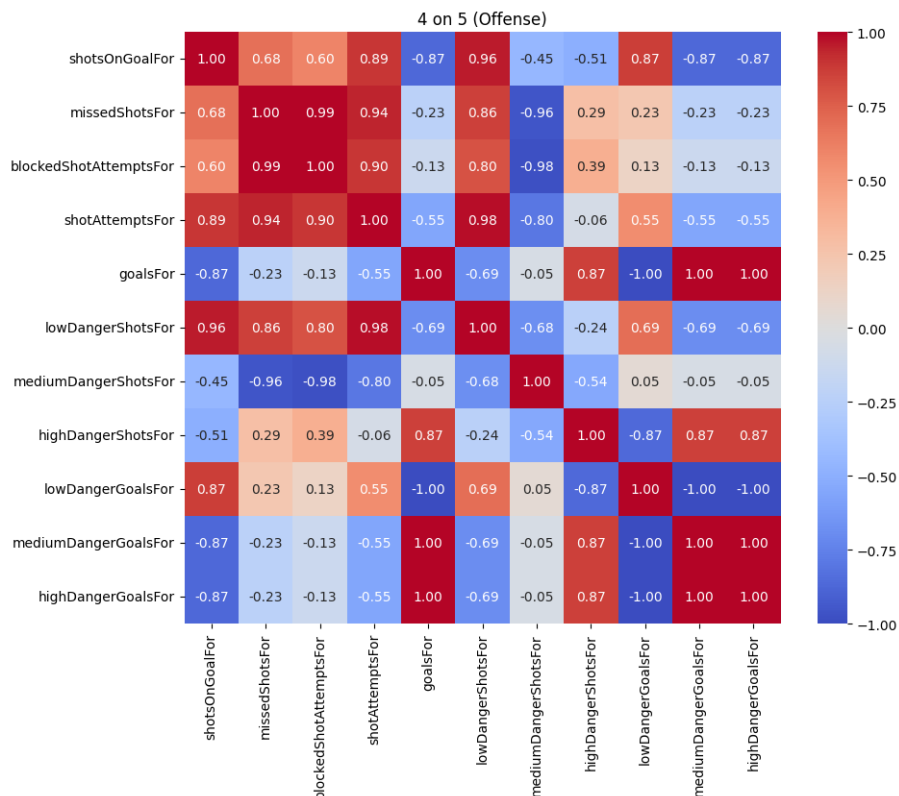
### 2.1.2 Power Play



Notable correlations:

- missedShotsFor & goalsFor (positive correlation): The correlation between goals scored and missed shots implies that on the power play, every opportunity to get the puck to the net should be taken as it provides the chance of a goal. This further proves that teams that move the puck and wait for the perfect chance will not see the same levels of success that teams that shoot first do.

### 2.1.3 Penalty Kill

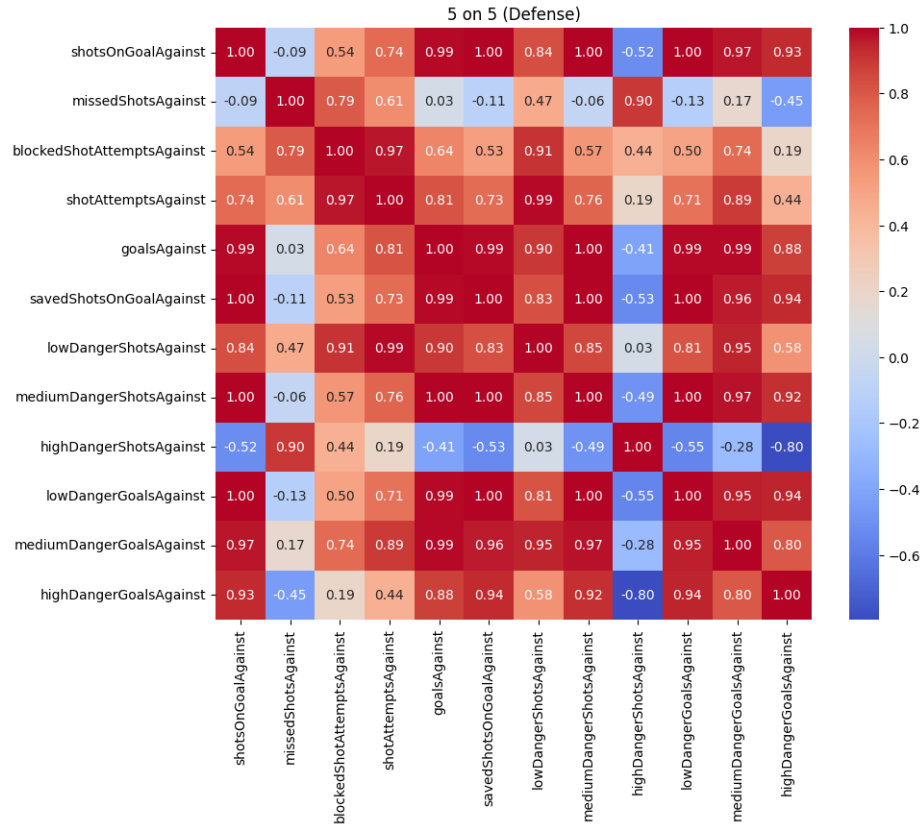


Notable correlations:

- shotAttemptsFor & goalsFor (negative correlation): The negative correlation between these two stats tells a different story from even strength and power play. On the penalty kill, a team that takes more shots does not greatly improve their chances of scoring a goal.
- highDangerShotsFor & goalsFor (positive correlation): This statistic further proves the importance of being selective with scoring chances on a penalty kill. High-danger shots have a direct association with goals, but when comparing this association to that of low and medium-danger shots to goals, it is seen that high-danger shots are the only ones that work.

## 2.2 Defense

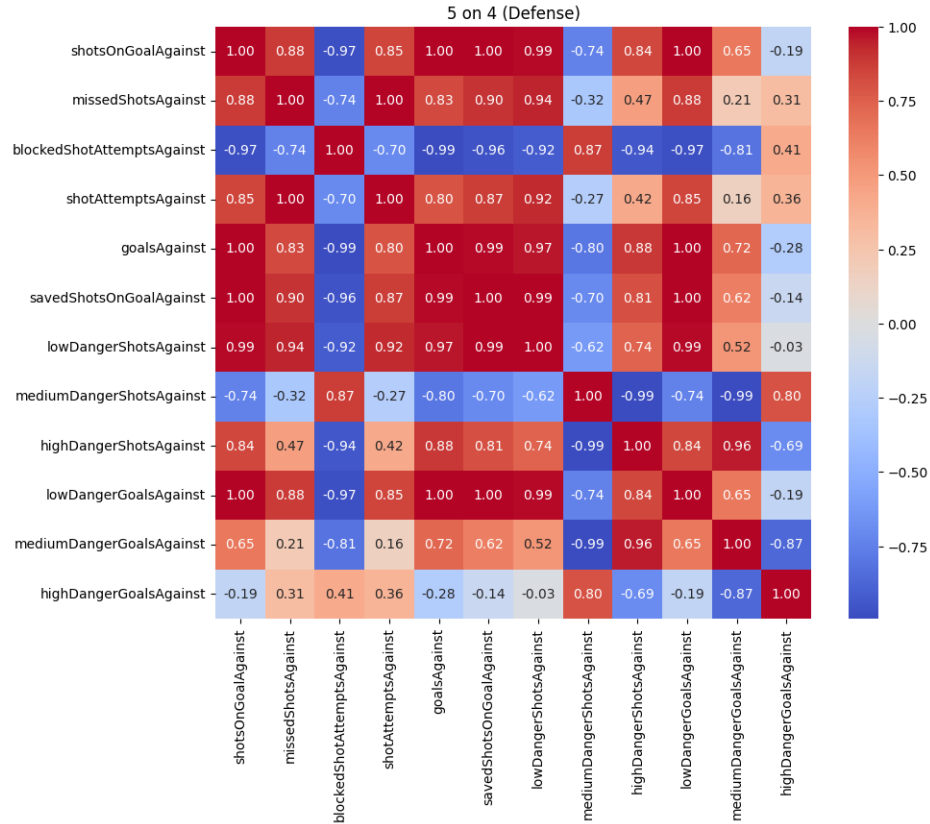
### 2.2.1 Even Strength



Notable correlations:

- goalsAgainst & lowDangerShotsAgainst (positive correlation): This correlation shows the importance of limiting an opponent's shots on goal. As previously mentioned, the more shots on goal a team gets blatantly correlates to the amount of goals they score.

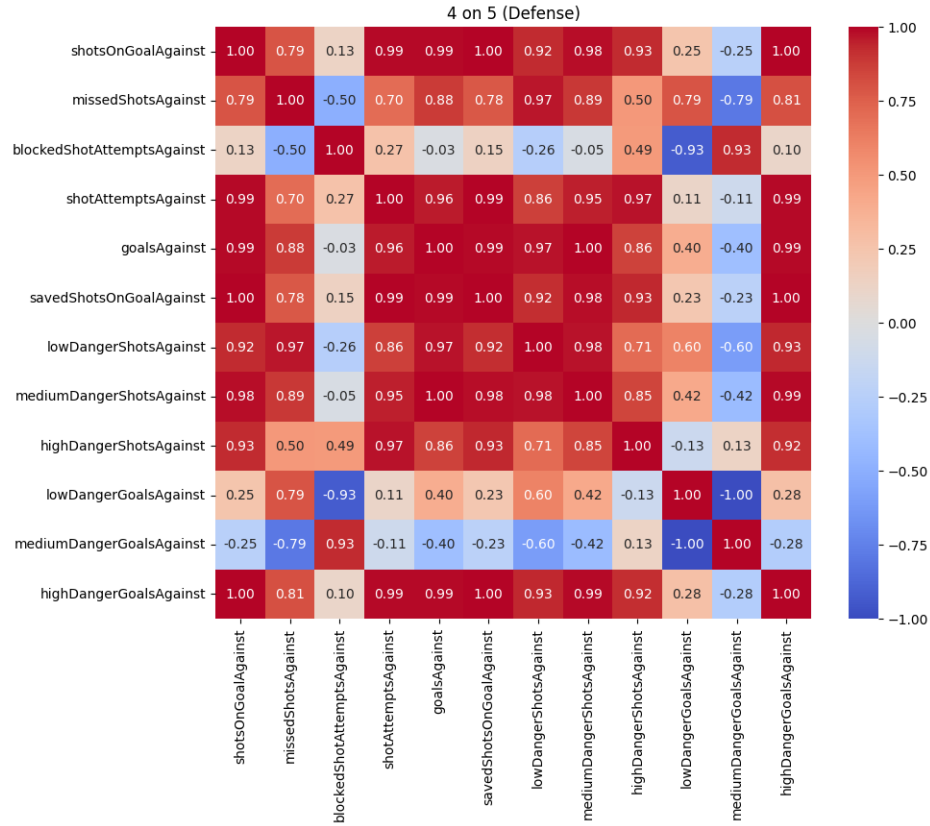
## 2.2.2 Power Play



Notable correlations:

- goalsAgainst & blockedShotAttemptsAgainst (negative correlation): The negative correlation here proves the importance of blocking shots in the defensive zone. The power play is not a situation where a team should face many shots on their goalie, but when they happen, it is important and worth it to get in the way of them.

### 2.2.3 Penalty Kill



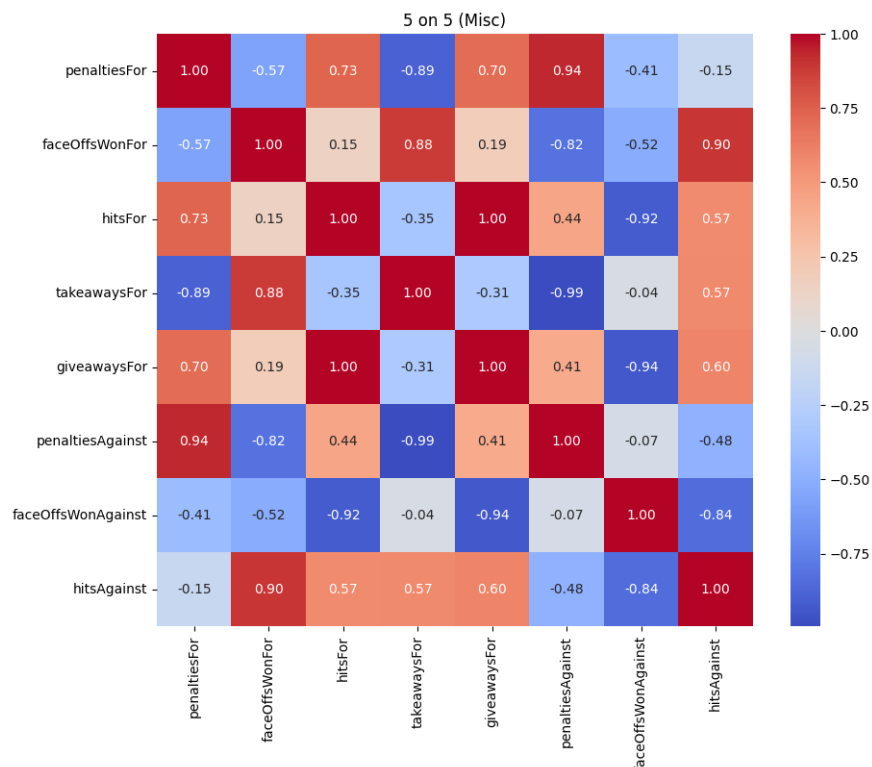
Notable correlations:

- missedShotsAgainst & highDangerGoalsAgainst (positive correlation): The correlation between missed shots against and high-danger goals against continues to reinforce the fact that all shots on goal are dangerous.



## 2.3 Miscellaneous

### 2.3.1 Even Strength



Notable correlations:

- giveawaysFor & hitsFor (positive correlation): The correlation between these two stats shows the importance of playing a physical game. More hits mean more giveaways, which then leads to more possession.

## 2.4 Code Used

```
1 def plot_heatmap(df, title):
2     plt.figure(figsize = (10, 8))
3     sns.heatmap(df[["listOfCategoriesIncluded"]].corr(),
4                 annot = True, cmap = "coolwarm", fmt = ".2f")
5     plt.title(title)
6     plt.show()
```

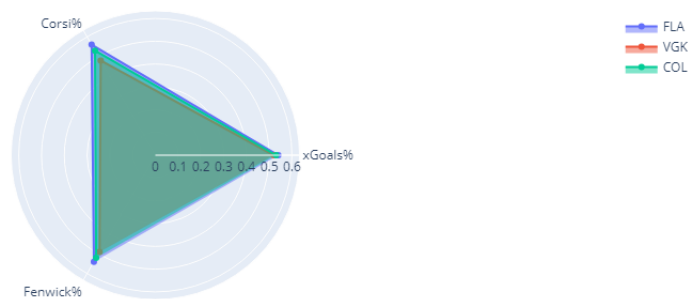
### 3 xGoals%, corsi%, and fenwick%

These three stats are easily the three most telling and important in hockey. This section will compare the three champion teams' stats. It will also compare the stats of the 2021 champions, the Colorado Avalanche, to the best team in the league to not make the playoffs that year, as well as the team that placed at the bottom of the league standings that year.

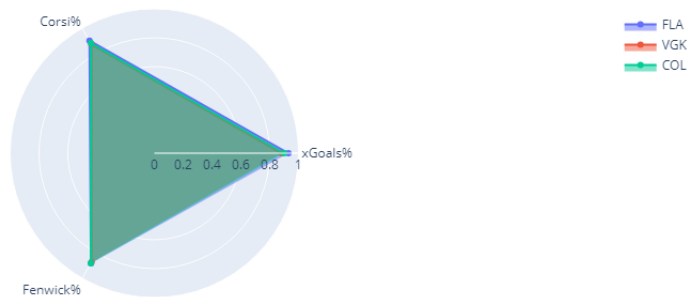
It is important to understand that between every team, the percentages will differ by a small visual amount. However, this difference in values still creates a huge impact. To make the contrast easier to spot, the graphs have been scaled accordingly.

#### 3.1 Comparison of Three Champion Teams

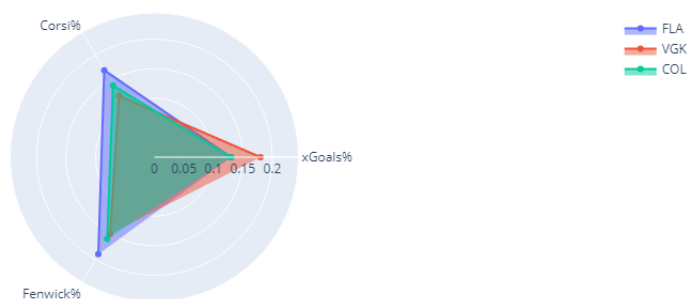
Corsi, Fenwick, and xGoals Percentages (Even Strength)



Corsi, Fenwick, and xGoals Percentages (Power Play)



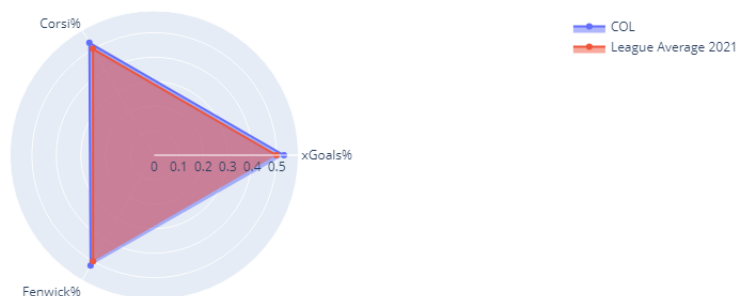
Corsi, Fenwick, and xGoals Percentages (Penalty Kill)



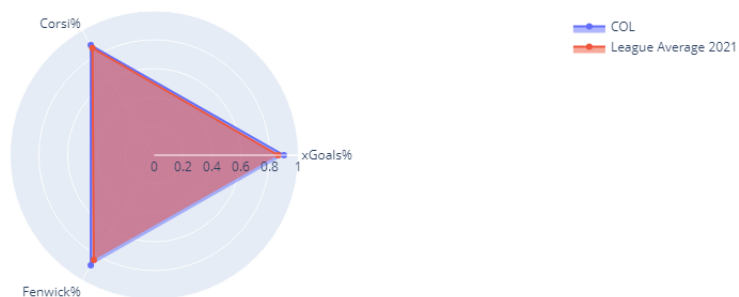
As seen with these three graphs, the three measurements are extremely close for each winning team, showing a consistency in xGoals%, corsi%, and fenwick%.

### 3.2 2021 Champions vs. 2021 League Average

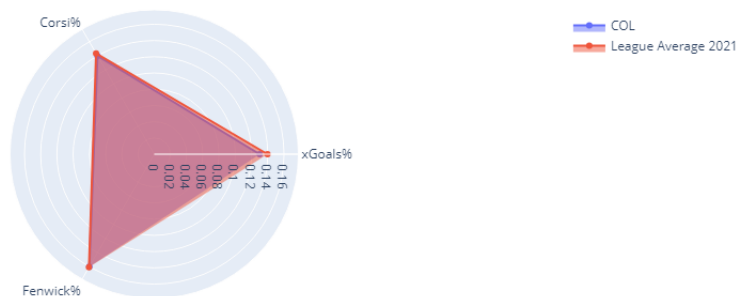
Colorado Avalanche vs. League Average (2021) (Even Strength)



Colorado Avalanche vs. League Average (2021) (Power Play)



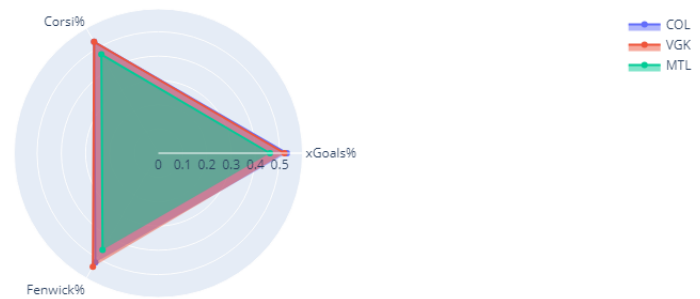
Colorado Avalanche vs. League Average (2021) (Penalty Kill)



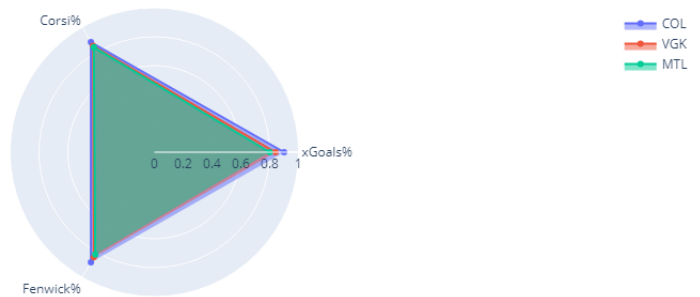
These three graphs show how the 2021 Colorado Avalanche are above league average in almost all statistics, except for xGoals% when they are on the penalty kill. Despite this anomaly, these graphs further illustrate the importance of having a good xGoals%, corsi%, and fenwick%.

### 3.3 2021 Champions vs. Best Team to Miss Playoffs vs. Bottom of League Standings

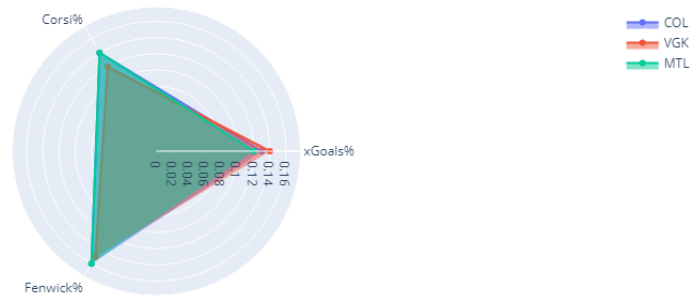
2021 Champion vs. top team to not make playoffs vs. bottom team in league standings (Even Strength)



2021 Champion vs. top team to not make playoffs vs. bottom team in league standings (Power Play)



2021 Champion vs. top team to not make playoffs vs. bottom team in league standings (Penalty Kill)



These graphs show the comparison between the 2021 champion (Colorado Avalanche), the best team to not make the playoffs in 2021 (Vegas Golden Knights), and the team that finished at the bottom of the league standings in 2021 (Montreal Canadiens). The progression in these graphs continues to solidify the importance of these three stats. The team that struggled the most had by far worse stats than the team that ended up winning the Cup.

### 3.4 Code Used

```

1 def plot_radar_chart_compare2(team1, team2, team1_name,
2   team2_name, situation, title):
3
4     team1_situation = filter_situation(team1, situation)
5     team2_situation = filter_situation(team2, situation)
6
7     stats = ["xGoalsPercentage", "corsiPercentage", "
8             fenwickPercentage"]
9
10    team1_stats = team1_situation[stats].mean().values.
11      flatten().tolist()
12    team2_stats = team2_situation[stats].mean().values.
13      flatten().tolist()
14
15    categories = ["xGoals%", "Corsi%", "Fenwick%"]
16
17    fig = go.Figure()
18
19    fig.add_trace(go.Scatterpolar(
20      r = team1_stats,
21      theta = categories,
22      fill = "toself",
23      name = team1_name

```

```

20     ))
21     fig.add_trace(go.Scatterpolar(
22         r = team2_stats,
23         theta = categories,
24         fill = "toself",
25         name = team2_name
26     ))
27
28     fig.update_layout(
29         polar = dict(
30             radialaxis = dict(
31                 visible = True,
32                 range = [0, 1]
33             ),
34             title = title
35         )
36
37     fig.show()

```

## 4 Conclusions

### 4.1 Heat Maps Analysis

The heat map analysis provided valuable insights into the offensive and defensive metrics that correlate with team success in the NHL. By examining the correlation between different statistics, we identified key patterns, such as the positive correlation between missed shots and shot attempts, as well as the negative correlation between high-danger shots and overall shot attempts. These findings suggest that teams focusing on quality scoring opportunities may take fewer overall shots but achieve better outcomes. Defensively, the analysis highlighted the importance of limiting low-danger shots and blocking shot attempts, particularly during power play situations. This underscores the need for a comprehensive defensive strategy that minimizes scoring chances against.

### 4.2 Radar Chart Analysis

The radar chart analysis compared the advanced stats of the three most recent Stanley Cup champions, the league average, and teams at varying performance levels. The charts revealed that the champions consistently outperformed the league average and struggling teams in xGoals%, corsi%, and fenwick%. These metrics are crucial indicators of a team's ability to generate scoring opportunities, control possession, and ultimately, achieve success. The comparison also emphasized that even small differences in these percentages can significantly impact a team's overall performance. This analysis reaffirms the importance of maintaining high levels in these key stats, as they are closely linked to a team's chances of winning games and championships.