Summary of Project in Progress:
Assessing NHL Players' Offensive Talent
Using a Linemate-Adjusted Metric

As an avid hockey fan and fantasy hockey enthusiast, I pursued a project in Python to scrape NHL data and implement a regression model to uncover undervalued NHL players. The motivation behind my analysis was that many publicly available advanced hockey statistics (e.g. Corsi) evaluate a player's performance without accounting for the skill levels of the teammates and opposition players that player usually plays on the ice with. Therefore, I constructed a model that quantifies players' offensive talent using a metric based on shots that controls for the impact of other players.

In order to conduct this project, I needed a dataset at the shift-level, containing all the shot attempts that occurred during each shift from every NHL game in a season. I was able to leverage an NHL API in order to retrieve game data on shot attempts, however the dataset did not show all the players who were on the ice when each shot attempt happened. Therefore, I scraped data on individual game summaries available online, which contained information on the start and end time of every shift for each player. I combined this data with the NHL game data for my analysis.

For my model, I first ran OLS regression, but then realized I had issues with multicollinearity and overfitting. Therefore, I decided to implement regularization and switched to a ridge regression model, which helped to reduce the variance of my estimates. I am working on a more detailed description of the data preparation required and the regression model used to derive the player metric.

The results of my analysis based on data up until the 2021-2022 season helped me identify players such as Jason Robertson that were likely more offensively talented than other publicly available statistics suggested. Before publishing results on my website, I plan to update this analysis to use data from the 2022-2023 season, tune my model, and prepare visuals to more easily compare players' offensive talent based on this linemate-adjusted metric.