W200   
Section 1

**Project 2 Proposal**

**Team Members:** Robert Bergan, Brent Biseda, Michael Hardy

**Team GitHub Repository:** https://github.com/UCB-INFO-PYTHON/project\_2\_Bergan\_Biseda\_Hardy

**Dataset Overview:** National Hockey League (“NHL”) game data that represents all of the official metrics for each game in the NHL over the past 6 seasons. Below is a chart of the table relationships of the data. <https://www.kaggle.com/martinellis/nhl-game-data>

**Proposal Summary:**

Our team intends to use the official NHL stat data set to analyze home ice advantage in professional hockey.

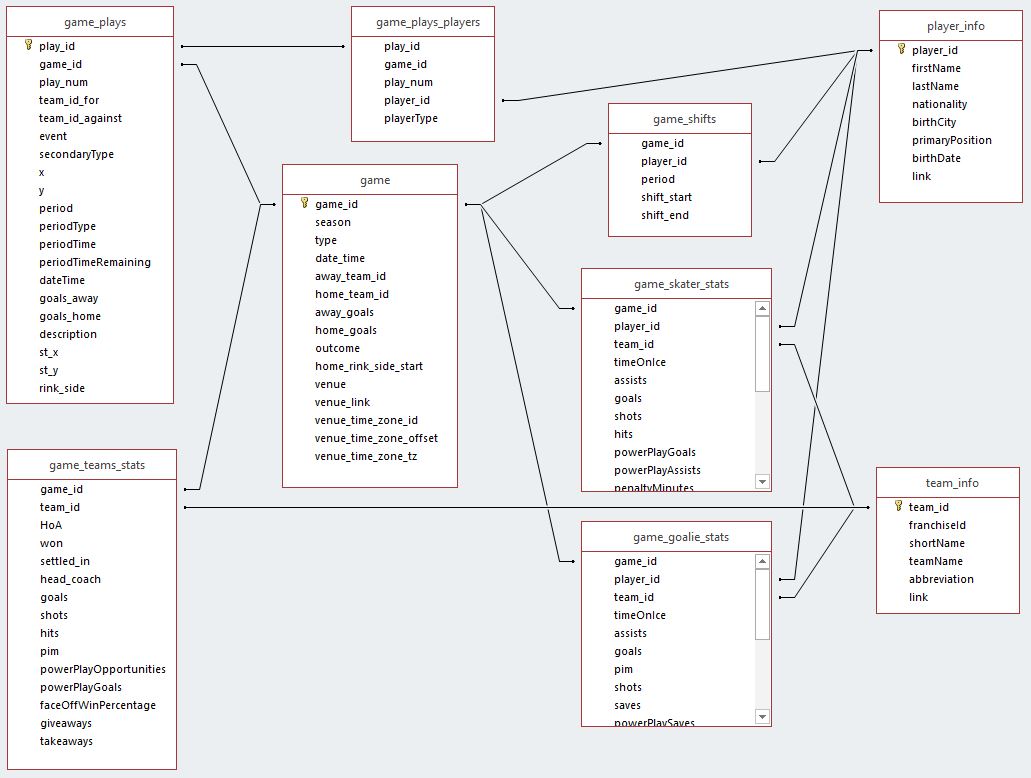
In all sports there is typically a supposed and oft-discussed home team advantage that is assumed to be derived from factors ranging from crowd noise to players’ comfort at home. Our report intends to quantify the actual home ice advantage in the NHL and to determine the associated drivers and factors.

The specific questions we intend to ask of the data include:

* The home team advantage is defined as the improvement in percentage of games won by the home team vs. the road team. Can the home ice advantage be quantified? And in what ways does it manifest itself?
  + Has this advantage varied by season?
  + How does the advantage vary by team or by “good” teams vs. “bad” teams?
  + Does the home team win percentage vary by month of the season?
  + Does the advantage vary in the regular season vs. the playoffs?
* Are there underlying drivers of home ice advantage that we could determine?
  + Does home ice advantage have a greater impact on goals-for or goals-against?
  + Are goals scored in different areas of the ice by the home team vs. the away team? This could potentially indicate a different style played by the home team vs. the away team.
    - Does this occur during regular season vs. playoff season?
  + Are goals scored at different time periods during the game by the home team vs. the away team?
  + Are goals by the home team correlated with face-off wins? In hockey the home team is in a slightly better position (away player has to put his stick down first) so this may be one of the factors in any home-ice advantage.
  + Is there a difference in the number of penalties that the home team has compared to the away-team? Is there a difference in the number of goals scored by the home team when there is a penalty?
  + Is there a difference in the number of goals scored home vs. away by each team’s top 3 players (as determined by total points)? Similarly, does the time played by the top 3 players vary home vs. away? The home team gets more opportunity to determine match ups throughout the course of the game so this could be another driver of home-ice advantage.
  + Is there a time-zone factor or geographic factor to home ice advantage caused by greater travel distance?

**Dataset Detail:**

The NHL dataset includes detailed gameplay data for each of the past 6 years in the tables outlined in the below table relationship diagram.



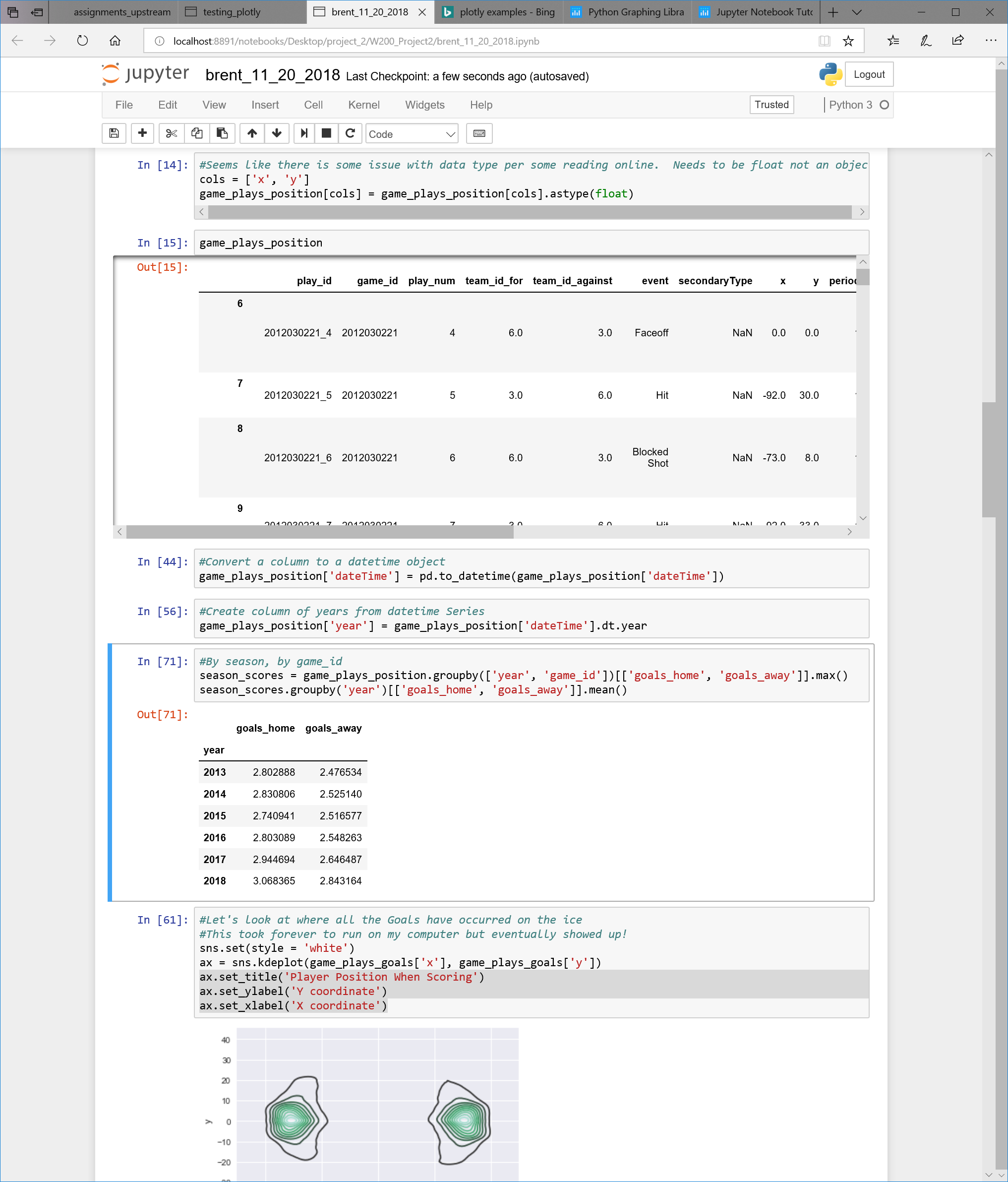
The primary tables we will focus on in our analysis include:

* **Game:** includes the aggregated info for each game in the data set including home team, away team, number of goals scored by each team and outcome (winning team and score)
* **Game\_plays:** includes each play of each game and the on-ice location of each play
* **Game\_teams\_stats:** includes more detailed aggregation data for each game such as face-offs won and the number of penalties

**Initial Exploratory Analysis:**

When grouping data into calendar years and evaluating average scores for home team vs away team, it can be seen that the home team outperforms.

Table – Tabulated Average Scores by Calendar Year



We can see that the distribution of goals for home team differs from the distribution of goals for the away team.

A close up of a plant

Description automatically generated

Figure - Home Advantage by Season (measured in goals)

A picture containing text, map

Description automatically generated

Figure - For Home Advantage by Regular vs Playoff (measured in goals)

We can investigate if team strategies have changed by season and during the play of each period.

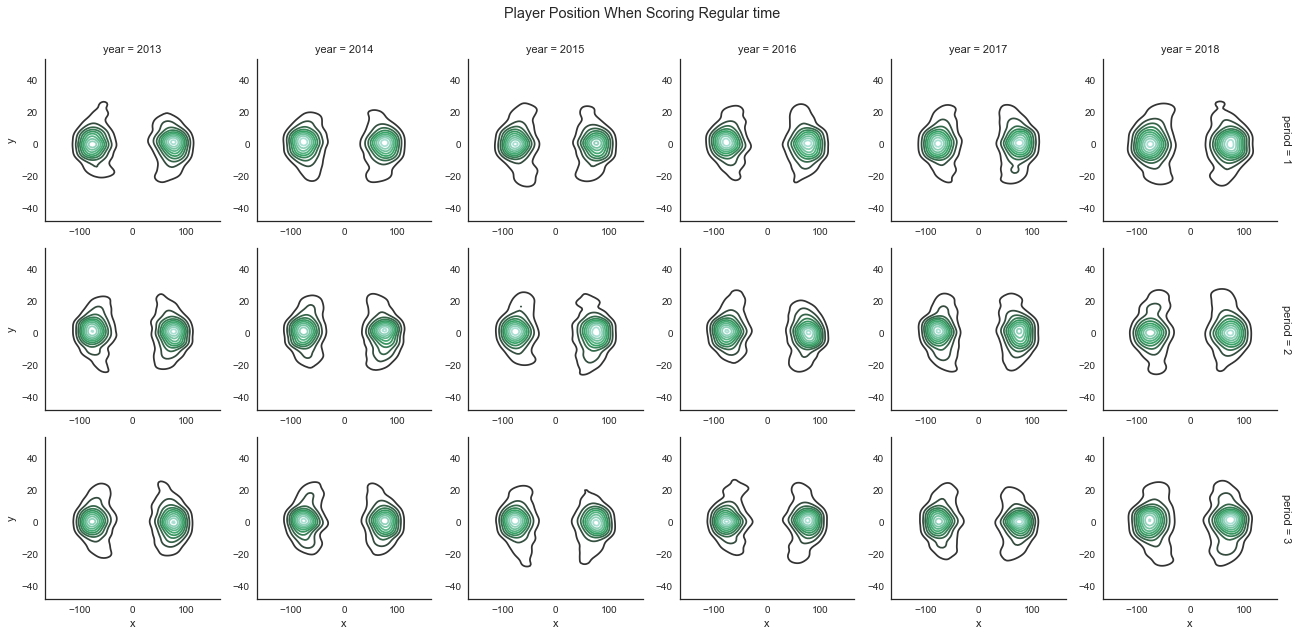


Figure – Positions of Scoring Players During Regular Play (Period 1 – 3)