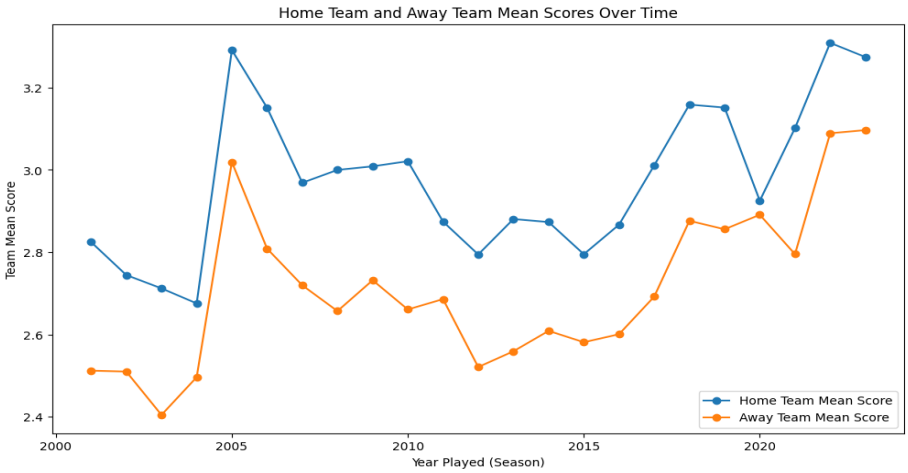
NAME – Perumalla Vijay (22096689)

DATASET LINK – https://github.com/fivethirtyeight/data/tree/master/nhl-forecasts

GITHUB LINK – https://github.com/vjaychandra/assignments

1. Mean Score Line Plot



Graph Explanation:

The average scores of home and away teams in NHL games between 2001 and 2023 are displayed on the mean score line graphic. The years are represented by the x-axis, while the mean scores are displayed on the y-axis. The evolution of average scores over time is represented visually by two lines, one for each of the home and away teams.

Interpretation:

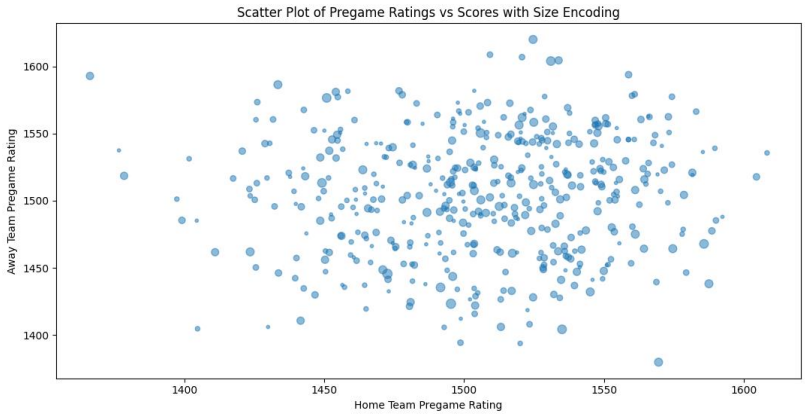
Parallel patterns can be seen in the Mean Home Team Score and Mean Away Team Score, with the former constantly being 0.2 points higher. Both ratings decrease in the first years, reaching a low point in 2004 and around 2.87 (home) and 2.5 (away) in 2001. The year 2005 had a significant peak of 3.4 for the Mean Home Team Score and about 3.0 for the Mean Away Team Score.

Both scores show a downward trend from 2005 to 2012, with 2.6 (away) and 2.8 (home) being the lowest in that year. After then, there is a rising tendency that experiences sporadic highs and lows. The mean home team score (3.4) and mean away team score (4.0) reach their major peaks in 2021 after a notable increase that starts in 2021. This rising trend from 2021 might indicate a shift in the dynamics of the game or enhanced offensive capability. Further investigation into variables like player moves, rule modifications, or team tactics throughout this time period may offer a more profound understanding of the tendencies that have been noted.

Conclusions:

By examining the patterns in average scores throughout time, we may understand the general effectiveness of both home and away teams. Variations in the lines might represent intervals of more scoring or superior defensive play. Recognizing these patterns may be related to noteworthy occurrences or modifications to the league's rules, and it can aid in comprehending how the NHL games are evolving.

2. Scatter Plot of Pregame Ratings vs Scores



Graph Explanation:

The scatter plot looks at the difference in home and away teams' pregame ratings. Every data point represents a game, and the size of the markers indicates the home team's score. The x-axis displays the home team's pregame rating, while the y-axis displays the away team's pregame rating.

Interpretation:

Most data points in the scatter plot are concentrated around the coordinates (1500, 1500), which represent the home and away teams' pregame ratings being equal. Interestingly, the scores are spread out throughout a range of 1450 to 1550 for both axes, indicating a wider variation in pregame ratings.

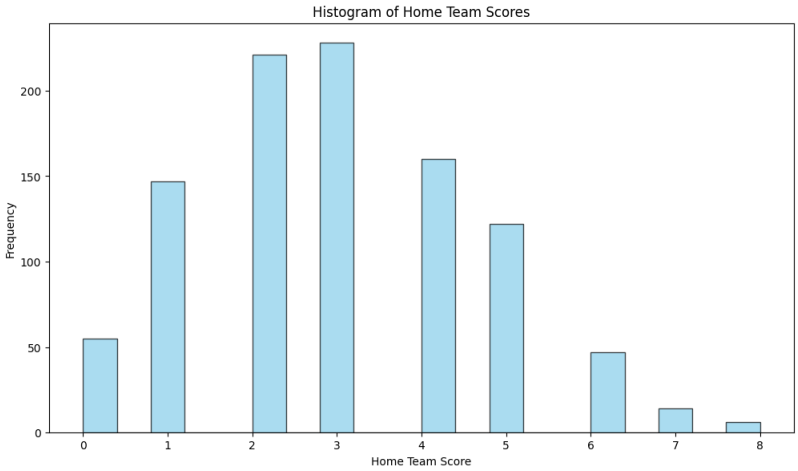
Points that leave the center beyond this central cluster are called outliers. These outliers, which are dispersed from the main concentration, show when home and away teams' pregame ratings diverge significantly.

Conclusions:

We can identify trends and connections between pregame ratings and actual game scores by examining the scatter plot. The home team's score is represented by the size of markers, which provides information about the relationship between improved team performance and higher pregame ratings.

Marker deviations from the center might indicate abnormalities in team performance or unanticipated events. One useful method for evaluating the predictive ability of beforehand assessments is the scatter plot. We can learn a lot about how well rating systems predict a team's performance by analyzing how pregame ratings match up with real game results. Bigger markers may be associated with higher scoring games, which would lead to more research into the elements influencing such results.

3. Histogram of Home Team Scores



Graph Explanation:

The distribution of home team scoring, taken from NHL games played between 2001 and 2023, is displayed in the histogram. Score ranges are shown on the x-axis, and the frequency of occurrence within each designated range is shown on the y-axis. This histogram shows the variety and common score ranges in the performances of the home side.

Interpretation:

Within different score ranges, the Home Team Scores Histogram shows different frequency patterns:

- A score of 0 occurs 50 times.
- A score of 1 is observed 150 times.
- A score of 2 is prevalent with a frequency of 220.
- A score of 3 has a frequency of 240.
- A score of 4 is recorded 160 times.

- A score of 5 has a frequency of 130.
- A score of 6 occurs 40 times.
- A score of 7 has a frequency of 20.
- A score of 8 has a frequency of 9 (Approx.)

Conclusion:

Analyzing the histogram yields valuable information about how home teams scoring is distributed throughout NHL games. The histogram's peaks indicate score ranges that regularly occur and provide insight into typical results. The histogram's spread and form highlight the variances in scoring patterns and the diversity of game outcomes.

We obtained a thorough grasp of the distribution of home team performances by locating recurrent score ranges and evaluating the spread. Peaks might represent instances of hotly contested games, while wider distributions would point to a variety of results, such as low- and high-scoring contests.