

NHL Player Contract Calculation and Player Evaluation using Advanced Metric Analysis

Stuart Scamehorn | DS 785 | stu.scamehorn@gmail.com

Abstract: This project aims to address the challenges of evaluating National Hockey League (NHL) skaters and forecasting future value in a decision environment where box-score statistics, reputation, and imperfect comparables drive negotiations. The project pipeline integrates multi-season NHL play-by-play, shift-level, and on-ice advanced statistical data with historical contract information to build a unified player-market dataset for skaters. Individual player impact is quantified via Regularized Adjusted Plus Minus (RAPM), which is translated into Goals, Wins, and Standing Points above Replacement (GAR/WAR/SPAR). Empirical age curves constructed from industry findings project each player's future value, while a market-slope ratio links WAR and SPAR to cap hit percentage before converting into dollar values based on the seasonal cap ceiling. A hybrid framework combining K-nearest-neighbors and per-term ridge regression estimates cap-hit percentages for candidate contracts, and results are displayed on an interactive Dash-based Python dashboard. This system improves contract prediction accuracy relative to simple baselines and provides actionable and clearly interpretable insights for contract negotiations.

Problem Definition: Athlete contract negotiations are a fundamental part of professional sports; assigning a dollar amount to the value a player provides both on and off the field can be contentious, especially if the player themselves does not agree with management's assessment. While some sports like Major League Baseball lack financial ceilings to limit contracts, and others like the National Football League allowing restructuring of contracts at any point during the deal, the National Hockey League features some of the strictest parameters for teams to work within; a hard salary cap ceiling for teams as well as for individual player contracts, set entry-level contract details, and a set average asking value (AAV) for each year of a set term on a contract all factor into negotiations. For team front offices, this means that overpaying players can limit the competitive window for a team, while under-offering may cause a player to test the free agent market and sign elsewhere. Another risk of contract negotiations dragging on, especially for Restricted Free Agents (RFAs), is that players are unable to play until a contract is signed. An example of this occurred in 2018 with the Toronto Maple Leafs; Up-and-coming star winger William Nylander was coming off his entry-level contract and was unable to come to an agreement with the team, leaving him unsigned after the summer free agency period. Negotiations dragged on until minutes before the RFA signing deadline on December 1st, with both sides being almost \$2 million apart as late as November.¹ This project pipeline seeks to aid the contract negotiation process by estimating a fair initial cap hit percentage for each possible term with historical comparables in addition to projecting a player's future value, estimating possible value lost or gained over the course of a contract. These findings are packaged into a user-friendly dashboard that allows front office staff to understand findings without a technical background. The results of the pipeline are then evaluated against signed contracts from the 2025-2026 offseason to determine success. The current pipeline has limitations; notably, it does not model goaltenders, and the

¹ Schlager, Brandon. "William Nylander, Maple Leafs Agree to Contract at 11th Hour of Sign-Or-Sit Standoff." *Sportingnews.com*, 2 Dec. 2018, www.sportingnews.com/us/toronto-maple-leafs/news/william-nylander-maple-leafs-reach-contract-at-11th-hour-of-sign-or-sit-standoff/147mwlw7r471f1sd7mczn9ftat.

underlying data only dates back to the 2015-2016 NHL season. Additionally, it only factors in on-ice contributions, ignoring factors like injury history, taxation rates, family considerations, contending status, and more.

Data Collection, Preparation and Cleaning: To support the contract valuation and prediction models, I assembled an integrated dataset that links on-ice performance to contract outcomes and the salary cap environment. Data for this project falls into three categories: granular, season-long, and contract. Granular shift-level and play-by-play data is sourced from the NHL Stats API alongside game schedule data and is collected through a Python-based API script. The API also set my season timeframe for this project, as the shift and play-by-play data within it is only complete back to the 2015-2016 NHL season. Season-long skater data is directly downloaded in CSV format from Natural Stat Trick², an NHL advanced statistic aggregation platform. The data is separated by situations, notably even-strength, power play, and penalty kill, and split by individual and on-ice performance. These raw per-situation files are merged together into one seasonal master file per situation and saved within the respective “Clean Data” folder for each season. Lastly, contract data is sourced primarily from CapWages³; I initially planned to utilize PuckPedia⁴ due to it having more comprehensive contract details, but changes to the site’s HTML structure required me to pivot for now, with plans to address the PuckPedia portion at a later date. This data is collected by a Python-based web-scraper that maps CapWages columns to their respective counterparts in the PuckPedia data, fills in any missing columns if possible, and saves as a unified contract master CSV file.

A separate cleaning script finishes the data preparation process. Player identifiers and names are standardized so that records from the NHL, Natural Stat Trick, and PuckPedia align, including resolving naming formats, suffixes, and duplicate entries. Demographic data sourced from the API is matched with players and written into the contracts master file. All raw inputs are saved as season-specific CSV files, with corresponding “Clean Data” versions that keep one row per player-season-situation and enforce consistent, numeric column types. Time-on-ice (TOI), scoring, and cap figures are converted into common units, and limited missing values are filled using back calculations where needed. Outlier values, namely skaters with low TOI and high expected goals (xG), are shrunk back towards zero so they do not dominate regression. One note is that, for the 2024-2025 season, the NHL API is missing the final 57 regular-season games; this will eventually be filled once the API is updated, but for now the missing values are logged and the results are understood with that context in mind. Finally, I constructed a TeamGames file that aggregates team-level results and schedule details from the API and MoneyPuck⁵, which provide contextual data needed for later models.

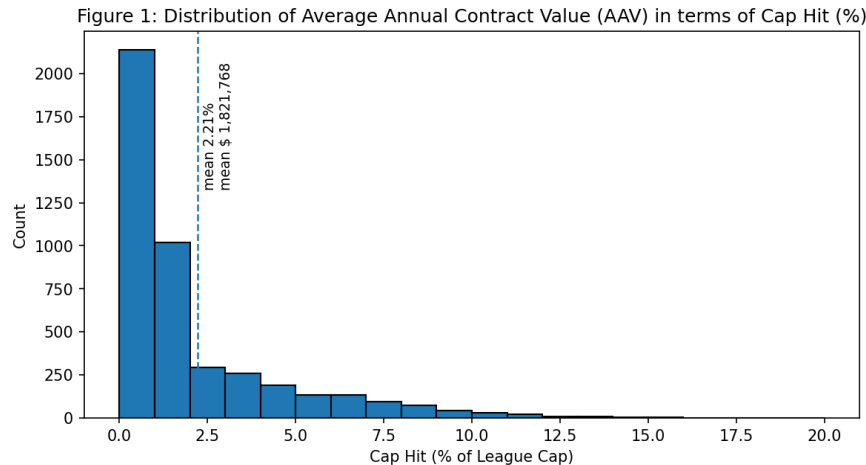
Exploratory Data Analysis and Feature Engineering: To explore possible biases within the underlying contract data, I first examined the distribution of NHL skater contract value. I analyzed cap hit percentage as opposed to raw dollar amounts due to the shifting cap ceiling year-to-year within the NHL, making cap hit percentage more interpretable across seasons.

² “Natural Stat Trick.” Naturalstattrick.com, 2024, www.naturalstattrick.com/.

³ “CapWages | Signings.” CapWages, 2025, capwages.com/signings.

⁴ “Home | Puckpedia.” Puckpedia.com, puckpedia.com/.

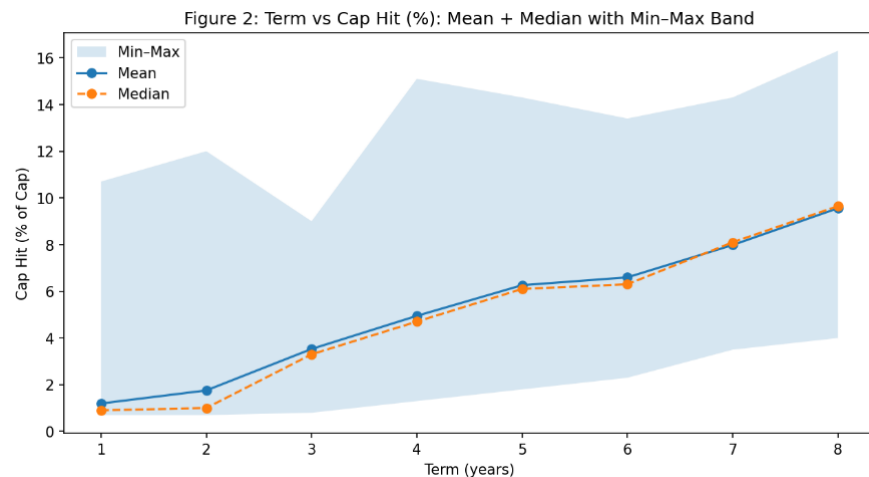
⁵ “MoneyPuck.com -Download Datasets.” Moneypuck.com, moneypuck.com/data.htm.



As shown in Figure 1, cap hit percentage is strongly right-skewed, even without the presence of ELC contracts, with the majority of player contracts being near the league minimum and a decreasing amount of contracts as the cap hit percentage increases. This trend was confirmed after splitting contracts by forwards and defensemen;

forwards had a slightly higher mean contract value at 2.24%, while defenders lagged slightly with 2.15%. However, defenders had a tighter mid-range of contracts, meaning that while forwards have a higher average value, defenders tend to be more consistent in the contract value they earn, staying fairly steady in the 3-7% range. This suggests that position carries meaningful structure in the market and should be included as a core predictor. Additionally, I analyzed how contract value shifts with each allowable term under the current NHL Collective Bargaining Agreement (CBA). These restrictions include a maximum term length of 8 years, and a maximum individual cap hit percentage of 20%.

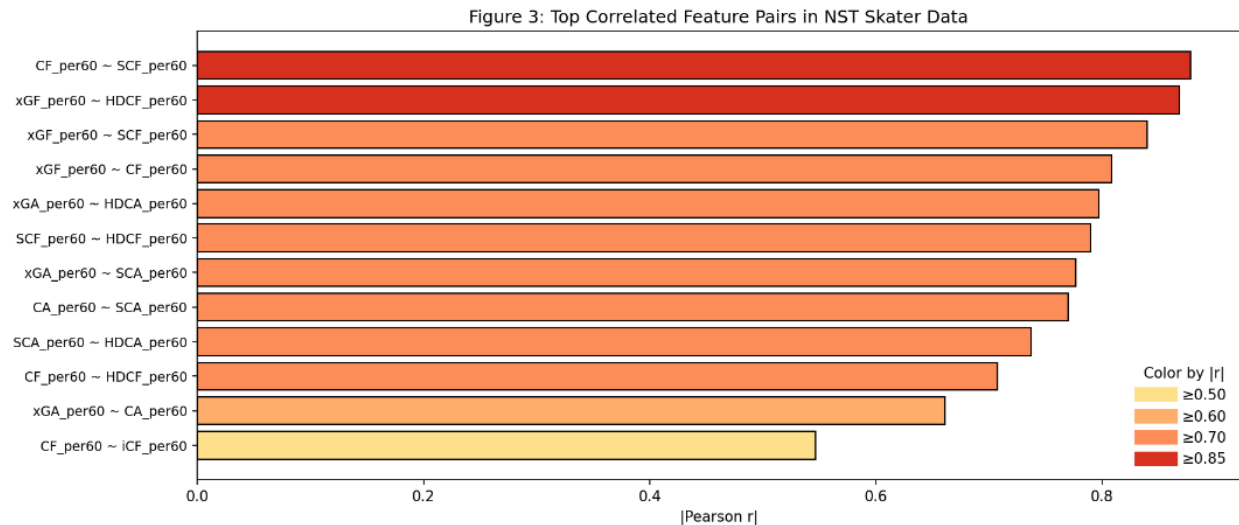
Figure 2 shows that the min-max band widens considerably past the 3-year mark, with mean and median contract value rising gradually as term increases. This indicates that multi-year deals exhibit more dispersion and risk than short-term contracts. Together, these patterns motivated a modeling approach that treats term and position as primary structural features and uses the historical cap hit min-max band as a sanity check for predicted values.



I then shifted focus to exploring the on-ice performance data. Using the NHL API shift and play-by-play data, I constructed stints as continuous segments of play where the sets of skaters on the ice remain constant, anchored on faceoffs, stoppages, and any line change from either team. These stints are then contextualized with the scoring state, zone starts, manpower situation, and more. For assigning xG values, I tracked Fenwick events, defined as “any unblocked shot attempt (goals, shots on net and misses) outside of the shootout.”⁶ This is opposed to Corsi events, which is all Fenwick events in addition to blocked shots.⁶ Then, all the xG are summed for and against within each stint window. This stint structure provides a consistent unit of analysis for measuring

⁶ “Natural Stat Trick - Glossary.” *Naturalstattrick.com*, 2025, www.naturalstattrick.com/glossary.php?players.

how different skater combinations drive xGF (expected goals for) and xGA (expected goals against), and it avoids double-counting events across overlapping shifts. The resulting stint-level xG numbers are then converted into per-60-minute rates; while these rates are right-skewed, especially in special teams situations, that behavior is expected and mainly reinforces the need to handle very low TOI cases carefully. Within the NST skater data, exploring correlations between candidate metrics shows that many intuitive pairs are highly connected.



As shown in Figure 3, pairings such as Corsi For per 60 and Scoring Chances For per 60 or xGF per 60 and High-Danger Corsi For per 60 are very highly correlated with a Pearson R value greater than 0.85. To address collinearity within my data, the Regularized Adjusted Plus Minus model features L2 regularization to prune redundant variables by keeping a single representative from each correlated group. The final predictor set features stint-based, situation-specific per-60 metrics (even-strength, power play, and penalty kill), along with age, position, term, and selected team-context variables.

Model Development and Evaluation: For the modelling stage, I began by fitting a Regularized Adjusted Plus Minus (RAPM) model inspired by the work of Patrick Bacon⁷, a renowned data scientist in the world of hockey analytics. RAPM, ported to hockey from basketball, addresses the issue of players inheriting team impacts that unregularized regression and raw plus-minus suffer from. The RAPM model is an L2 regularized ridge-regression model using the computed stints as the unit of analysis. For each stint, I modelled xGF/60 and xGA/60 using indicator variables for home skaters, away skaters, and goaltenders, along with context dummies for zone starts, score state, manpower situation, home-ice advantage, and pulled-goalie states. The model was fit with an L2 penalty on the player and context coefficients, so longer stints receive more weight and the

⁷ Bacon, Patrick. "A Comprehensive Guide to Using Regression (RAPM) to Evaluate NHL Skaters (with Source Code)." Medium, TDS Archive, 25 Dec. 2020, medium.com/data-science/a-comprehensive-guide-to-using-regression-rapm-to-evaluate-nhl-skaters-with-source-code-ad76d1e1b8da. Accessed 26 Nov. 2025.

Situation	Model	Lambda	RMSE/Avg. Stint Length	P-Value
Even Strength	Ridge	1.4	0.0431 xG	0.0002
Even Strength	OLS	0	0.0433 xG	
Power Play	Ridge	1.1	0.0811 xG	0.0034
Power Play	OLS	0	0.0849 xG	
Penalty Kill	Ridge	1.1	0.0861 xG	0.0013
Penalty Kill	OLS	0	0.0899 xG	

TABLE 1: RMSE COMPARISON FOR RAPM VS OLS MODELS WITH P-VALUE

Additionally, the model lambda (λ) value was tuned with a grid between 1 and 1.5, applying a $2*\lambda$ penalty to prevent xGA results being applied too heavily to defenders. To evaluate the RAPM model, I compared the root mean squared error (RMSE) per average stint length against the results from an unregularized OLS model with the same design matrix. As shown in Table 1, the RAPM model achieves improved RMSE per average stint length in every play situation. Using a paired t-test with a standard α value of 0.05, these improvements were verified to be statistically significant, indicating that the regularized RAPM fit both stabilizes player coefficients and delivers meaningfully better out-of-sample performance than the baseline.

These results are then translated into value metrics by moving from RAPM to GAR, WAR, SPAR, and finally dollars. To define my replacement-level baseline, I took players outside the top 12 forwards and 8 defenders at even strength, top 5 forwards and 2 defenders on power play, and top 4 forwards and 4 defenders on penalty kill for each team by TOI, and calculated replacement-level xG baselines using the selected players, splitting forwards and defenders. Then, I calculated offensive GAR components for each player p in situation s using the formula:

$$GAR_{p,s}^{off} = \left(\frac{xGF}{60}_{p,s} - \frac{xGF}{60}_{repl,s} \right) * \left(\frac{TOI_{p,s}}{60} \right)$$

For defense, the formula is flipped:

$$GAR_{p,s}^{def} = \left(\frac{xGF}{60}_{repl,s} - \frac{xGF}{60}_{p,s} \right) * \left(\frac{TOI_{p,s}}{60} \right)$$

Finally, total GAR is calculated as:

$$GAR_p = \sum_s (GAR_{p,s}^{off} + GAR_{p,s}^{def})$$

GAR is then translated to WAR and SPAR using Goals per Win (GpW) and Goals per Standing Point (GpSP); these values are sourced from Evolving Hockey⁸ and filled for the later seasons. Conversion for each follows the same format for a player p and season t :

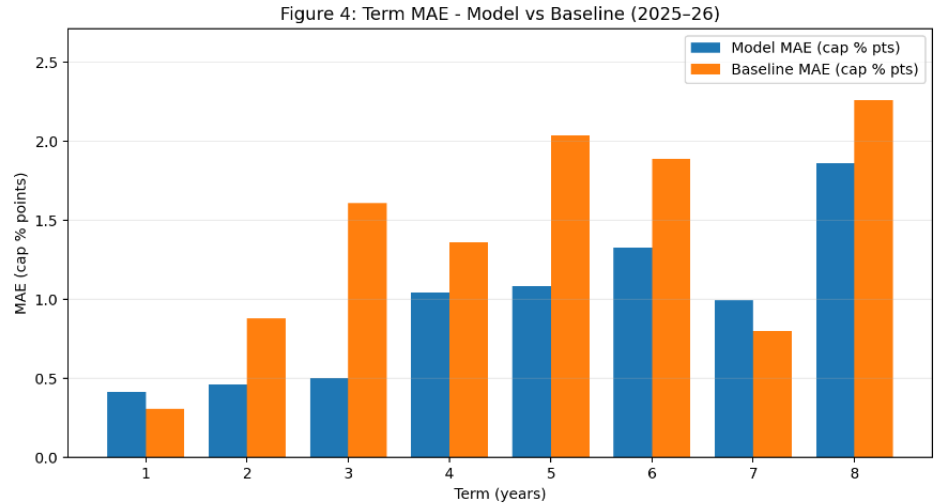
$$WAR/SPAR_{p,t} = \frac{GAR_{p,t}}{GpW/GpSP_t}$$

To stabilize season-to-season fluctuations within the WAR and SPAR values, I blended each on a three-year weighted average then applied a compression function that shrinks outlier values towards the center before pricing. To convert from raw WAR and SPAR values to dollars, I estimated a market rate by taking the median ratio of cap-hit percentage to WAR (or SPAR) among non-ELC contracts with positive value in each season and position bucket. After trimming outliers, I then multiplied that cap-percentage-per-WAR ratio by the current salary cap to get dollars per WAR (or SPAR). Each player's WAR\$ and SPAR\$ were then the compressed WAR/SPAR

regularization term shrinks coefficients toward zero to stabilize estimates for players with limited TOI but large xG values. I tuned the ridge penalty over a small grid using 5-fold GroupKFold cross-validation grouped by game ID to avoid leaking information across folds.

⁸ "Evolving-Hockey." Evolving-Hockey.com, 2021, evolving-hockey.com/glossary/goals-above-replacement/.

multiplied by these rates, yielding a dollar representation of a player’s on-ice impact. To model future player value, I constructed blended aging curves using published work from Dom Luszczyszyn⁹, Josh and Luke Younggren¹⁰, and Eric Tulsy¹¹, each of whom independently showed consistent



peak-performance ages (~25 for forwards, ~26 for defensemen) followed by predictable decline rates. Due to my dataset containing relatively few skaters in the 35+ range, fitting a purely data-driven curve would produce unstable late-career behavior, so the blended curve provides a more reliable and domain-aligned shape while providing simulated uncertainty from the differences in findings from each. I applied these position-specific curves multiplicatively while projected WAR remains above a small threshold, and transition to a linear decay once the curve nears zero to avoid unrealistic negative values caused by compounding decline.

To predict cap hit percentages for each possible contract term, I fit a hybrid model that blends a K-nearest-neighbors (KNN) comparables layer with a ridge regression component inspired by EvolvingHockey¹² to bolster terms where the KNN layer comes up short. Neighbors were identified using a Mahalanobis distance over grouped feature blocks (usage, impact, value, age, demographics, and context). I tuned the model by looping over different block-weight configurations and blend weights to minimize the mean absolute error (MAE) in cap-hit percentage on each player’s actual signed term. To evaluate the model, I compared it against a simple baseline that always predicts the historical median cap-hit percentage for the relevant term, and Figure 4 summarizes this comparison by plotting MAE by term for both the baseline and the tuned model. In all but 2 term cases, the blended model produced a lower MAE than the median baseline, especially on mid-length contracts where the pool of historical comparables is largest. In the case of 1-year contracts, the median contract value is the league minimum, so minimal error is expected in that case. For 7-year deals, the blended model lacks sufficient comparables to learn what differentiates a 7-year deal from a 6- or 8-year one value-wise. This indicates that the combination of Mahalanobis-based neighbor matching and value-aware features captures meaningful structure in how teams convert projected player value into contract dollars, especially on an expanded contract data set.

⁹ Dom Luszczyszyn. “As NHL Free Agency Nears, How Will Players Age on Their New Contracts?” The New York Times, 5 June 2025, www.nytimes.com/athletic/6398802/2025/06/05/nhl-free-agency-mitch-marner-age/.

¹⁰ EvolvingWild. “A New Look at Aging Curves for NHL Skaters (Part 1).” *Hockey Graphs*, 23 Mar. 2017, hockey-graphs.com/2017/03/23/a-new-look-at-aging-curves-for-nhl-skaters-part-1/.

¹¹ Tulsy, Eric. “How NHL Scoring Rates Change with Age.” *SBNation.com*, 13 Mar. 2014, www.sbnation.com/nhl/2014/3/13/5500522/nhl-scoring-stats-rates-age-analysis.

¹² “Evolving-Hockey.” Evolving-Hockey.com, 2023, evolving-hockey.com/glossary/contract-projections/. Accessed 26 Nov. 2025.

Term	# Contracts	# Pass (1*MAE)	Pass Rate (1*MAE)	# Pass (2*MAE)	Pass Rate (2*MAE)
1	115	97	84.3%	107	93.0%
2	76	61	80.3%	72	94.7%
3	19	12	63.2%	17	89.5%
4	15	5	33.3%	9	60.0%
5	8	4	50.0%	6	75.0%
6	11	2	18.2%	7	63.6%
7	8	2	25.0%	6	75.0%
8	18	3	16.7%	7	38.9%
Total	270	186	68.8%	231	85.5%

TABLE 2: PASS RATES WITHIN MAE THRESHOLDS

model is to get as close as possible, allowing for the difference to be explained using situational context. On the 270 contracts signed for the 2025-2026 season, the blended model produces an overall mean absolute error of 0.639 cap-percentage points. Using this overall MAE as a tolerance band, 186 out of 270 contract predictions (68.9%) have absolute errors less than 1*MAE, and 231 out of 270 (85.6%) fall within 2*MAE. These summary statistics, broken out by term in Table 2, show how often the model's predictions land reasonably close to the actual cap hit percentage when measured in terms of the model's own typical error. For short deals, the model is relatively strong, with 84.3% of 1-year contracts and 80.3% of 2-year contracts falling within 1*MAE, and over 93-95% of these short deals are within 2*MAE. However, the Pass Rate drops off as the number of contracts decreases per term. This, paired with the results shown in Figure 4, showed that the blended model generally reduced MAE on most term lengths, particularly in short-range deals, and faced greater volatility on very long contracts due to a lack of comparables to train on within the contract data set. One example of a player whose prediction fails the 1*MAE test but is explainable by situational context is Brock Faber of the Minnesota Wild.

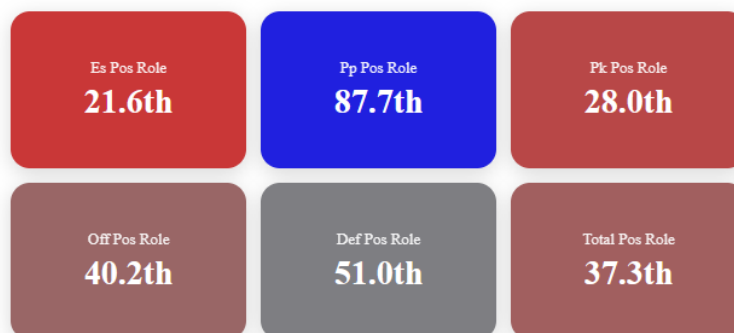
Brock Faber

Pos: D | Shot: R | Team: MIN

Age at signing: 22.0 | H: 87.0 in | W: 200.0 lb

Role: 1st Pair

FIGURE 5: BROCK FABER DEMOGRAPHICS FROM DASHBOARD



Selected term: 8 - Cap%: 9.68 - AAV (anchored to 2025-26 cap): \$9,242,383

Name	Pos	Year	Term	Cap%	Similarity
Zach Werenski	D	2019	3	6.1	56
Miro Heiskanen	D	2022	8	10.2	47
Mikhail Sergachev	D	2020	3	5.9	45
Ryan Murray	D	2016	2	3.9	42
Jacob Trouba	D	2016	2	4.5	39

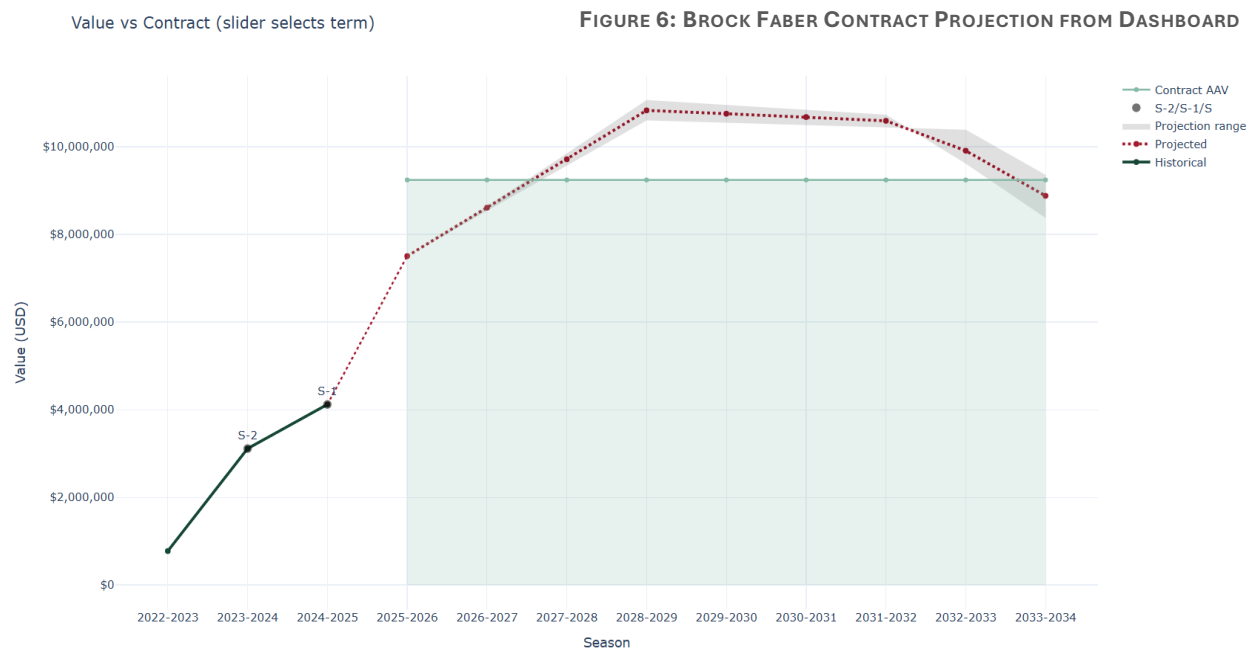
Results, Insights and Recommendations:

Though the blended model shows promising results, it does not model several important factors that drive NHL contract negotiations, including player injury history, taxation rates, family considerations, contending status of teams, contract clauses, and more.

As such, the goal of the

are percentile blocks

highlighting a player's contribution in even-strength, power play, penalty kill, offense, defense, and total. These tiles are separated by position as well as TOI-based role for each player. For example, Faber is a 22-year-old 1st Pairing defenseman who ranks in the 87th percentile on Power Play and 37th Percentile of 1st Pairing defenseman. His projected contract cap hit percentage for an 8-year deal (mirroring his actual signed deal for the 2025-2026 season) is 9.68% for the initial season, coming out to roughly \$9,250,000 a year, with comps to other high-profile defenseman such as Zach Werenski and Miro Heiskanen. This projection is also favorable based on his projected age curve shown in Figure 6, the main plot from the dashboard showing historical production (bold line), projected production for the next 9 years (dotted line), and the predicted contract for the given term in the lightly shaded region. However, on an 8-year term, the model overpredicts Faber's actual signing number of an initial cap hit percentage of 8.9%, yielding an error of 0.778%. However, there are two important considerations in Faber's case: first, Faber has a modified No Move Clause and 15-team No Trade Clause in the final 3 years of his deal, which drives the value down slightly. Additionally, Faber is from Minnesota; he grew up watching the Wild and attended at the University of Minnesota. As such, him accepting a team-friendly discount to play for his hometown team makes more sense.



Conclusion: This project aims to give NHL front offices a more data-driven starting point for skater contract negotiations by tying together on-ice value, projecting ageing, and historical market behavior. The pipeline combines shift and play-by-play data from the NHL API, situation-specific advanced stats from Natural Stat Trick, and contract data from CapWages to create a unified player-market dataset. RAPM with L2 regularization provides a more stable estimation of individual on-ice impact compared to an unregularized baseline, with the results being translated into GAR, WAR, SPAR, WAR\$, and SPAR\$ using team-TOI-based replacement baselines and season-specific goals per win and standing point rates. Those values are projected forward using a hybrid age curve and fed into a blended KNN and ridge model that predicts cap hit percentages for each possible term. This blended model outperformed a median-cap-hit baseline in 6 out of 8 terms, with 68.9% of predictions falling within 1*MAE and 85.6% falling within 2*MAE. Taken

together, these results indicate that the pipeline is capturing meaningful structure in how teams convert projected player value into contract dollars, and that the dashboard can serve as a practical tool for sanity-checking offers, exploring alternative term structures, and identifying over- or under-valued skaters. However, the model does not currently account for off-ice factors and certain contract mechanics like clauses and statuses, meaning that predictions must be interpreted with appropriate context.

Looking ahead, there are several avenues for continued improvement of the pipeline. Widening the data scope to include more seasons would give the prediction model more contracts to learn from, increasing the accuracy of predictions. It would also give the KNN layer more comparables to pull from, particularly for elite players that lack adequate comparables within the current dataset. Additionally, factoring in additional contract granularity, such as contract statuses, clauses, etc., would give the model a more nuanced understanding of some additional drivers behind contract negotiations, further improving results. Lastly, I would like to expand the model to include goalies beyond the RAPM step, as well as patch stand-ins for the age curves and demographic estimations as more data becomes available. Together, these enhancements would deepen the model's connection to the realities of NHL roster building while preserving the core goal of the project: giving decision-makers a transparent, interpretable, and data-driven foundation for negotiations that can have a significant impact on the future of the organization.