Jacksonville Jaguars Cap Optimization

Carson Blin, Tyler Gelinas, Gunnar Hansen

University of North Florida

Dr. Richard Shang

**Table of Contents**

[Introduction 3](#_Toc264756862)

[Data Description and Exploration 3](#_Toc481222226)

[Data Analysis and Results 11](#_Toc588137019)

[Regression, Cluster, and Machine Learning Analyses 13](#_Toc599821727)

[Cluster Analysis: 17](#_Toc274714787)

[Cluster Based Models 19](#_Toc1583493277)

[Machine Learning Models 19](#_Toc1521886297)

[Defense 21](#_Toc1574248937)

[Offense 22](#_Toc692562865)

[ML Model Applied to Jags Current Salary Cap 23](#_Toc1609728432)

[Discussions 25](#_Toc696737642)

[Key Takeaways from Regression vs. Machine Learning Models 25](#_Toc1658083969)

[Defensive Investment and Team Success 26](#_Toc1939291015)

[Offensive Investment and Team Success 26](#_Toc1947176642)

[Implications for the Jacksonville Jaguars 27](#_Toc431621566)

[Cluster Analysis Insights 27](#_Toc762485905)

[Conclusion 27](#_Toc517123656)

[Address of the Research Question 28](#_Toc591536267)

[Recommendations for the Jacksonville Jaguars Front Office 29](#_Toc881235981)

[Reallocate Spending Toward High-Impact Positions (LBs): 29](#_Toc1606009996)

[Reduce Overspending on Defensive Backs (DBs): 29](#_Toc776736366)

[Maintain a Balanced Offensive Approach: 29](#_Toc1373718355)

[Use the Model as a Strategic Planning Tool: 29](#_Toc1317887403)

[Benchmark Against Winning Archetypes: 29](#_Toc1956605769)

[Draft for Value at DB and WR: 30](#_Toc354763821)

# Introduction

The Jacksonville Jaguars have had a consistent problem of performing at a low level compared to the rest of the competition currently in the National Football League (NFL). Because of this, we asked ourselves the question of what the Jaguars could do to find more success in future seasons.

The primary goal of our research was to explore the salary cap optimization of the Jacksonville Jaguars and compare it to other teams across the NFL. Our study included a deep dive into the relationship between salary cap spending and success, based on the breakdown of dollars spent by position group, offense and defense specific spending, and whether our findings could predict better ways for the Jacksonville Jaguars to allocate their financial contracts on players to better set themselves up for success in future years.

The tools we used during our research include Excel, Python, and JMP Pro. The data analysis methods we used these tools for were describing our statistics, regression analysis, cluster analysis, and machine learning in both regression and predictive analyses (bootstrap forest/random forest).

Our process began by describing the statistics used for our primary analytical methods. From there, we performed regression analysis on the salary cap of all positions and compared its results to the same analysis with an offensive scope as well as a defensive scope to find which position groups were most significant with respect to salary cap spending. We also applied a Bootstrap Forest Machine Learning model to the offense and defense results for further context.

Once we had a better idea of the data we were working with, we built a Cluster Analysis model to segregate the NFL’s most successful teams. The goal here was to identify which teams should serve as the best examples for the Jaguars to mirror their salary cap allocations. We determined successful teams and built these clusters based on Team Wins, Offensive Rating, and Defensive Rating.

Finally, we built a predictive machine learning model with the purpose of identifying the ideal allocation percentages across each position group for future seasons and compared the Jaguars current allocation breakdown to that of the predictive model’s ideal allocation breakdown.

# Data Description and Exploration

**Table 1. Description of Data – Pittsburgh Steelers Salary Cap Summary 2015-2025**

|  |  |
| --- | --- |
| **Characteristic** | **Players, n (%)** |
| **Number of players contracted** | *n = 1,411* |
| **Cap Type** | |
| *Active* | *588 (41.67%)* |
| *Dead Money* | *650 (46.07%)* |
| *Injured Reserve* | *95 (6.74%)* |
| *Non-Football Injury* | *2 (0.14%)* |
| *Practice Squad* | *68 (4.82%)* |
| *Reserve/Did Not Report* | *2 (0.14%)* |
| *Reserve/PUP* | *5 (0.35%)* |
| *Reserve/Suspended* | *1 (0.07%)* |
| **Age group** (years) | |
| *34+* | *20 (1.42%)* |
| *27-33* | *264 (18.71%)* |
| *21-26* | *408 (28.92%)* |
| *Missing Values* | *719 (50.95%)* |
| **Cap Hit ($)** | *Mean = $1,515,478 (StD = $3,405,398)* |

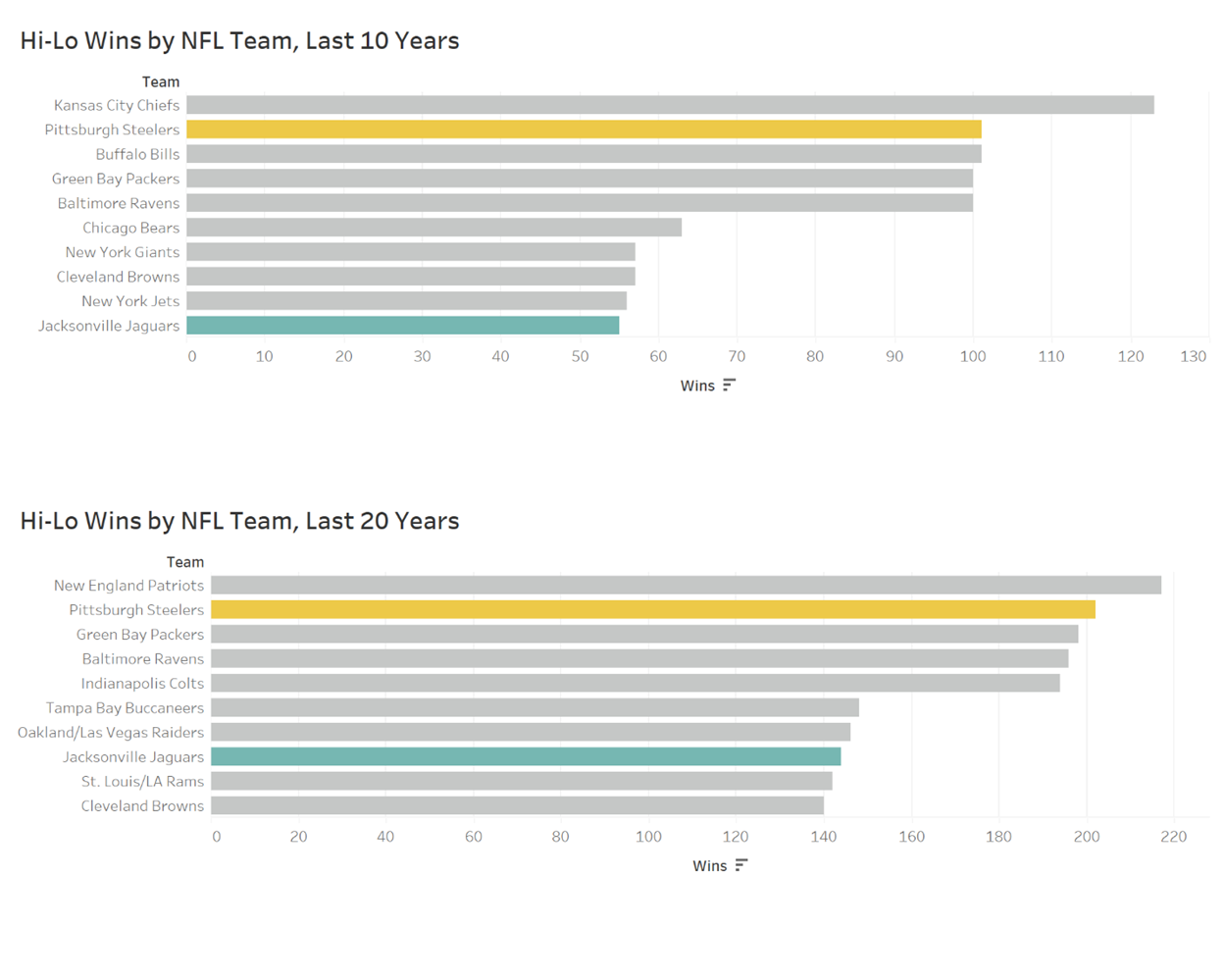
**Table 2. Description of Data – Jacksonville Jaguars Salary Cap Summary 2025**

|  |  |
| --- | --- |
| **Characteristic** | **Players, n (%)** |
| **Number of players contracted** | *n = 152* |
| **Cap Type** | |
| *Active* | *136 (89.47%)* |
| *Dead Money* | *16 (10.53%)* |
| **Age group** (years) | |
| *34+* | *11 (7.24%)* |
| *27-33* | *68 (44.73%)* |
| *21-26* | *57 (37.50%)* |
| *Missing Values* | *16 (10.53%)* |
| **Cap Hit ($)** | *Mean = $6,095,593 (StD = $10,211,938)* |

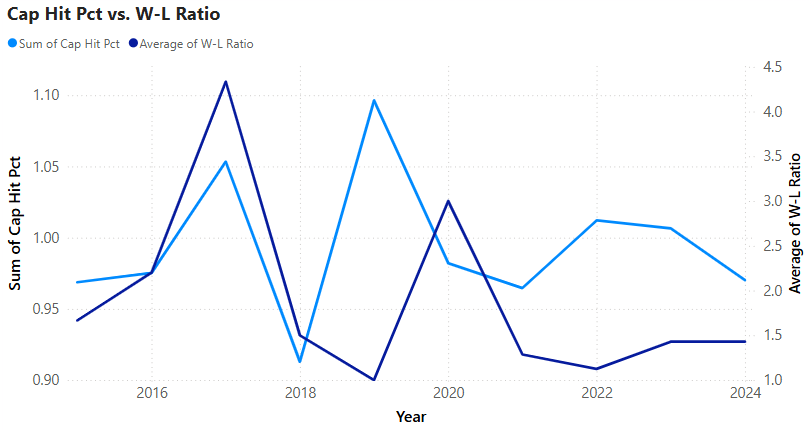
To investigate how cap spending differs between the Steelers and Jaguars, we need to dive into the data of how each team utilizes their cap space. Table 1 describes the characteristics of the primary fields assessed for the Pittsburgh Steelers salary cap over the past 10 seasons. These fields include the number of players contracted, the cap type breakdown, a breakdown of player age groups, including the discovery of missing values within that breakdown, and the mean and standard deviation of the cap hit, in dollars. Table 2 shows the same descriptive summary, under the scope of the 2025 Jacksonville Jaguars salary cap.

Key findings in the descriptive statistics begin with the difference in Dead Money spending between the two teams. Over 46% of the Steelers allocation is within Dead Money while the Jaguars only allocate a little over 10% of their cap space to it. Another important callout is the overall Cap Hit. The Steelers hit is significantly lower at $1,515,478 compared to the Jaguars at over $6 million. Additionally, the Steelers Standard Deviation on the Cap Hit is much closer to the mean. Finally, we noticed a large volume of missing values within the Age Group field. All of these missing values are only applicable to Dead Money, Practice Squad, or Injured Reserve Cap Types. This variable may still be useful when analyzing the active rosters of each team.

**Figures 1 and 2. Hi-Lo Wins by NFL Team; 10 Years vs 20 Years**

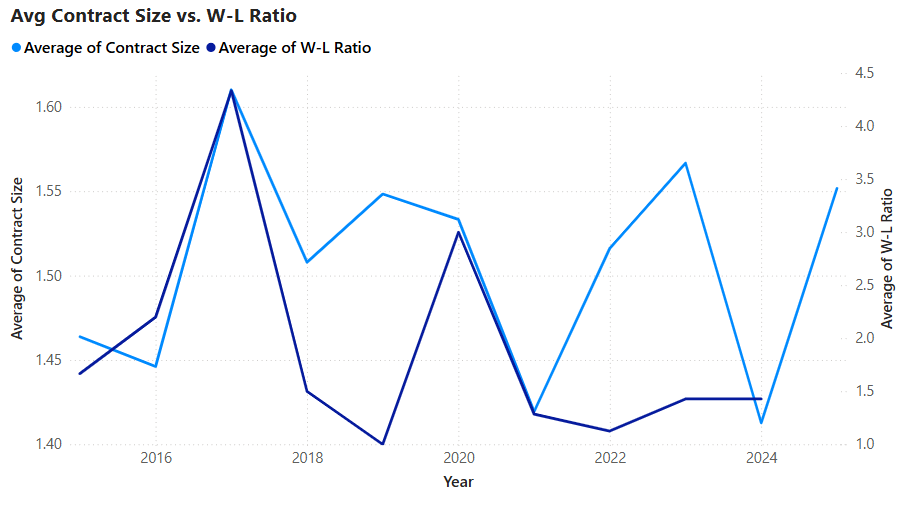
 Figures 1 and 2 represent the highest and lowest number of wins by NFL teams over the past 10 (Figure 1) years and past 20 years (Figure 2), separated by Top 5 and Bottom 5 teams. The purpose of this chart is to prove why the Pittsburgh Steelers, highlighted in yellow, are the best team for the Jacksonville Jaguars to emulate their spending for the best chances of success. Although the Steelers are shown as the number two ranked team in both charts, it’s important to note that both number one ranked teams do not appear in the Top 5 group at all in the opposite comparison. The Kansas City Chiefs for example, are the number one team over the last 10 years but are not in the Top 5 over the last 20 years. The New England Patriots comparatively are the top ranked team in the last 20 years but are not in the Top 5 group for the past 10 seasons. Only the Pittsburgh Steelers and Baltimore Ravens appear in the Top 5 in both timeframes, but the Steelers have more wins in each. Additionally, Pittsburgh has 101 wins in the last 10 years, and 202 wins in the last 20, meaning that in each decade, they have the same number of wins. There is nothing more consistent than that. Finally, we can see the Jacksonville Jaguars, highlighted in teal, are among the Bottom 5 teams in each category, and the worst over the last decade. Therefore, the Pittsburgh Steelers are the prime team to emulate, and the Jaguars have a lot of room to grow.

**Figure 3. Cap Utilization and W-L Ratio**



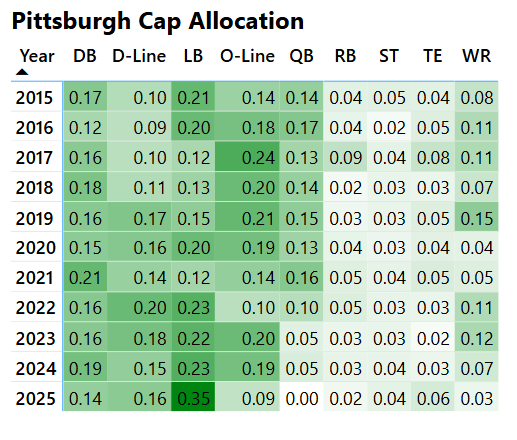
Shown above in Figure 3, is a 2-axis line graph of the Pittsburgh Steelers’ total cap hit percentage (and thus cap utilization %) and Win-Loss Ratio over time. The light blue line depicts the W-L ratio for each season. Especially successful regular seasons can be seen in 2017 and 2019. The dark blue line represents the aggregate of cap hit percentage from each contract. Interestingly there is also a spike 2017 that coincides with a successful season. In 2019 this trend is reversed with high cap utilization resulting in poor performance. These two seasons could provide important information to cap optimization especially if Jacksonville decided to utilize maximum cap space as a cornerstone of their strategy.

**Figure 4. Average Contract Size and W-L Ratio**



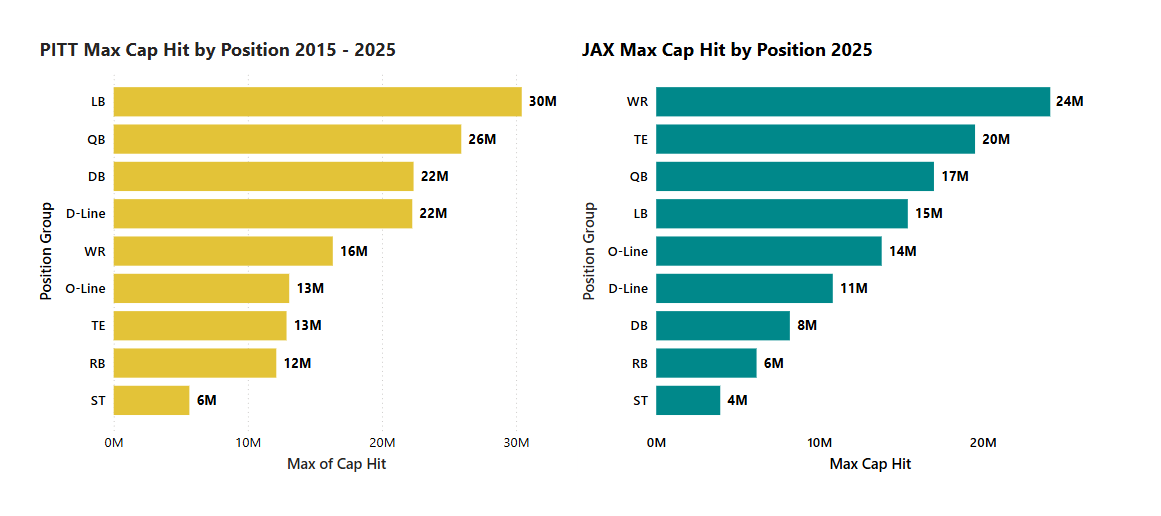
Shown above in Figure 4, is a 2-axis line graph of the Pittsburgh Steelers’ average contract size percentage and Win-Loss Ratio over time. The dark blue line depicts the W-L ratio for each season. The light blue line represents the average size of contract that is applied to Pittsburgh’s salary cap. The scale of contracts went from 1 – 6 with the following values: 1 **(<1%)**, 2 **(1-3%)**, 3 **(3 – 7%)**, 4 **(7 – 12%)**, 5 **(12 – 18%)**, and 6 **(18% or greater)**. Years where Pittsburgh had the lowest average contract scores, and a strong W-L ratio may represent the strongest spending strategies and are areas targeted for further analysis.

**Figure 5. Pittsburgh Cap Allocation Heatmap**



The above heatmap represents Pittsburgh's Salary Cap allocation by position. The darker the green the more of their cap Pitt has allocated to that position group. Consistent areas of high spend seem to DB, D-Line, LB, and O-Line and a deeper analysis into these positions could be beneficial.

**Figures 6 and 7. PITT Historic Max Spend vs. JAX Present Day Spend**



The side-by-side bar graphs represent the largest contract by position group in terms of total cap hit (in million USD). Figure 6 represents the Pittsburgh Steelers max contract by position across their previous 10 seasons. Figure 7 represents the Jacksonville Jaguars’ max contracts for 2025. It’s interesting that the Jaguars have higher spends in WR and TE than the Steelers do across 10 years.

**Figure 8. Pittsburgh Total Players Total Salary by Year**

A screenshot of a computer

AI-generated content may be incorrect., Picture

This graph visualizes the correlation between the players' total salary and the year. There was a slight decrease in the year 2021, but the total salary has been increasing at a steady rate ever since. When we bring age into this graph, we notice that salary has been increasing for all ages since 2015.

**Figure 9: Pittsburgh Total Sign on Bonus by Age**

A screenshot of a computer

AI-generated content may be incorrect., Picture

This graph considers the total sign on bonus and compares it to age. There is a significant boost going from 25 to 26 and remaining steady until age 30. After age 30, we see a drop in signs on bonuses. We factor in position into this to see if there is a specific position that gains more of a sign on compared to other positions.

**Figure 10 Pittsburgh Average Cap Hit Per Year**

A screenshot of a graph

AI-generated content may be incorrect., Picture

This graph we look at the average cap hit. We notice a slight increase in average cap hit in 2018, then see it decrease for a few years and then increase in 2021. We bring Cap Type into this to see how much money is going towards active players, injured players, etc.

# Data Analysis and Results

One of the main goals of our research was to investigate how salary cap spending differs between each team across the National Football League (NFL), and its relationship to how successful NFL teams perform. To do this, we assessed the data in two different ways. The first way was to analyze how significant each position group was to salary cap spending across entire teams themselves. The second was to break that analysis out into offense and defense specifically, to see if one side of the ball showed a better relationship between spending and success comparatively. Each of these analyses are described in Tables 1 and 2 below.

Table 1 describes the characteristics of our first assessment, the analysis of teams as a whole. We did a ten-year lookback, and since the NFL has had 32 teams for each of the past 10 seasons, there were 320 total observations. Typical NFL statistics we included in the analysis were Points For, Points Against, and Point Differential. The means of Points For and Points Against being equal, with the average Point Differential being zero make sense due to each team having to compete against one another with parity throughout every NFL season. Other fields included were Offensive Rating, Defensive Rating, and Cap Utilization. Position grouping values are expressed in Millions of Dollars. For clustering, the data was scaled using a Min-Max method with the max being 15 and the minimum being 0. For the machine learning models and regression analyses, salary data was “normalized” meaning that they were scaled to be representative of current salary cap limits in the NFL.

The key findings in Table 1 are seen in the breakdown of each position group using the “normalized” salary data. DB (Defensive Backs), DL (Defensive Line), LB (Linebackers), OL (Offensive Line), QB (Quarterbacks), RB (Running Backs), TE (Tight Ends), WR (Wide Receivers), and ST (Special Teams) were the position groups identified for our study. The amount of dollars (in millions) of salary cap spending are highest in the Offensive Line ($45.69 million), Defensive Backs ($43.48 million), and Defensive Line ($41.48 million) position groups. This is likely due to there being a higher number of roster spots designated for these positions compared to others.

**Table 3. Description of Data – NFL Full Roster Salary Cap Summary 2015-2024**

|  |  |
| --- | --- |
| **Characteristic** | **Values** |
| **Number of data observations (32 teams x 10 years)** | *n = 320* |
| **Teams** | *n = 32* |
| **Years** | *n = 10* |
| **Points For** | *Mean = 373.29 (StD = 69.83)* |
| **Points Against** | *Mean = 373.29 (StD = 54.26)* |
| **Point Differential** | *Mean = 0.00 (StD = 98.60)* |
| **Wins** | *Mean = 8.17 (StD = 3.13)* |
| **DB (Defensive Backs)** *- Millions of Dollars* | *Mean = $43.38 (StD = 12.10)* |
| **DL (Defensive Line)** *- Millions of Dollars* | *Mean = $41.48 (StD = 15.13)* |
| **LB (Linebackers)** *- Millions of Dollars* | *Mean = $32.09 (StD = 13.94)* |
| **OL (Offensive Line)** *- Millions of Dollars* | *Mean = $45.69 (StD = 10.60)* |
| **QB (Quarterbacks)** *- Millions of Dollars* | *Mean = $27.15 (StD = 12.36)* |
| **RB (Running Backs)** *- Millions of Dollars* | *Mean = $12.76 (StD = 5.49)* |
| **TE (Tight Ends)** *- Millions of Dollars* | *Mean = $12.82 (StD = 5.46)* |
| **WR (Wide Receivers)** *- Millions of Dollars* | *Mean = $29.70 (StD = 9.96)* |
| **ST (Special Teams)** *- Millions of Dollars* | *Mean = $7.39 (StD = 2.60)* |
| **Offensive Rating** | *Mean = 0.00 (StD = 4.16)* |
| **Defensive Rating** | *Mean = 0.00 (StD = 3.16)* |
| **Cap Utilization** | *Mean = 98.85 (StD = 5.94)* |

Table 3 describes the characteristics of our second assessment, the analysis of team spending based on the offensive and defensive sides of the ball separately. The data for this analysis only differentiates from that of Table 1 in that position groups are broken down as a percentage of overall team salary cap spending instead of a normalized dollar amount. We still have a ten-year lookback period, using all 32 teams with 320 total observations.

The key findings from Table 3, similar to Table 1, are that the highest percentages of salary cap spending are seen in the Offensive Line (18%), Defensive Backs (17%), and Defensive Line (16%) position groups. This is again likely due to there being a higher number of roster spots designated for these positions compared to others. Quarterbacks for example, who may get paid more as individuals, only represent about 11% of total salary cap spending across the NFL.

**Table 4. Description of Data – NFL Offense v Defense Salary Cap Summary 2015-2024**

|  |  |
| --- | --- |
| **Characteristic** | **Values** |
| **Number of data observations (32 teams x 10 years)** | *n = 320* |
| **Teams** | *n = 32* |
| **Years** | *n = 10* |
| **Points For** | *Mean = 373.29 (StD = 69.83)* |
| **Points Against** | *Mean = 373.29 (StD = 54.26)* |
| **Point Differential** | *Mean = 0.00 (StD = 98.60)* |
| **Wins** | *Mean = 8.17 (StD = 3.13)* |
| **DB (Defensive Backs)** | *Mean = 17.19% (StD = 4.70)* |
| **DL (Defensive Line)** | *Mean = 16.14% (StD = 5.81)* |
| **LB (Linebackers)** | *Mean = 12.74% (StD = 5.51)* |
| **OL (Offensive Line)** | *Mean = 18.09% (StD = 4.01)* |
| **QB (Quarterbacks)** | *Mean = 10.76% (StD = 4.85)* |
| **RB (Running Backs)** | *Mean = 5.05% (StD = 2.13)* |
| **TE (Tight Ends)** | *Mean = 5.08% (StD = 2.17)* |
| **WR (Wide Receivers)** | *Mean = 11.75% (StD = 3.85)* |
| **ST (Special Teams)** | *Mean = 2.94% (StD = 1.05)* |
| **Offensive Rating** | *Mean = -0.02 (StD = 4.24)* |
| **Defensive Rating** | *Mean = 0.00 (StD = 3.22)* |

## Regression, Cluster, and Machine Learning Analyses

Table 4 below shows a regression analysis using Wins as the dependent variable and each Position Group as independent variables.  The coefficients represent the number of wins expected for each dollar spent on the applicable position group.  Out of the 9 position groups, 6 show significance based on their p-value.  The 6 significant position groups are Defensive Back (DB), Defensive Line (DL), Linebackers (LB), Quarterback (QB), Tight End (TE), and Wide Receiver (WR).

**Table 5: Salary Cap of All Positions**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Variables* | *Coefficients* | *P-value* | *Multiple R* | *R Square* |
| Intercept | -3.08621985 | 0.29021226 | 0.2509997 | 0.063000849 |
| DB | 4.34869E-08 | 0.011237017 | - | - |
| DL | 5.35464E-08 | 0.000901993 | - | - |
| LB | 3.61337E-08 | 0.036544216 | - | - |
| OL | 2.89624E-08 | 0.108116867 | - | - |
| QB | 6.55975E-08 | 0.000186782 | - | - |
| RB | 2.56536E-08 | 0.432763012 | - | - |
| ST | 5.74458E-08 | 0.402937636 | - | - |
| TE | 6.81164E-08 | 0.038571733 | - | - |
| WR | 4.23691E-08 | 0.029189918 | - | - |

Shown in Figure 1, the random forest model for overall cap allocation and its effect on team performance demonstrates an R² of 0.654, meaning it explains 65.4% of the variance in the target variable. This indicates a stronger predictive power compared to the defensive-only model. The Root Average Squared Error (RASE) is 1.836, reflecting the average prediction error. The model was trained on 320 observations using 100 decision trees. The out-of-bag RASE (1.810849) is closer to the in-bag error (1.349477), suggesting a better generalization capability compared to the defensive-focused model. Defensive Line (DL) is also a very significant feature, accounting for 14.56% of the model’s importance, followed by Quarterback (QB) at 13.13%. Other significant contributors include Defensive Backs (DB) and Running Backs (RB), with portions of 11.67% and 11.59%, respectively. Wide Receiver (WR) and Linebacker (LB) features also play a notable role, while Tight End (TE), Offensive Line (OL), and Special Teams (ST) have the lowest impact, though they still contribute to the model's predictive capability.

A main difference between the models is the importance of the RB position which was not significant in the regression model but was in the ML model. The ML model accounted for much more of the variance when compared to the regression model, improving the r-squared value by a factor of ten going from 6% to 60%. Because of this the ML model is more accurate, suggesting that the relationships in the data are likely nonlinear and/or involve complex interactions between variables.

**Figure 11: All Players Bootstrap Forest Output**

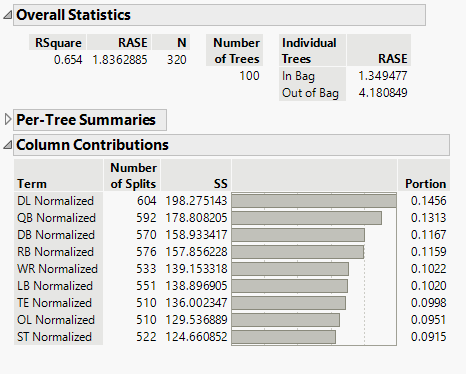


Table 6 below shows a regression analysis using Defensive Rating as the dependent variable and each Defensive Position Group as independent variables.  This model uses a min/max scaling weight and shows Defensive Back (DB) as the only significant position group.  The key inference to note here is that most teams who have success in defensive efficiency are spending a statistically significant amount of their salary cap on at least one premium defensive back.

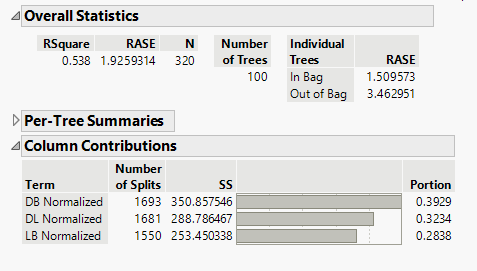
**Table 6: Salary Cap of Defensive Players**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Variable* | *Coefficients* | *P-value* | *Multiple R* | *R Square* |
| Intercept | 7.08266 | 2.05E-16 | 0.158773 | 0.025209 |
| DB | 0.183792 | 0.005837 | - | - |
| DL | 0.002497 | 0.973121 | - | - |
| LB | 0.020766 | 0.748756 | - | - |

Shown in figure 12, the random forest model for defensive cap allocation and its effect on defensive rating demonstrates an R² of 0.538, meaning it explains 53.8% of the variance in the target variable. While this suggests a moderate level of predictive power, the Root Average Squared Error (RASE) is 1.925, indicating the average number of prediction errors. The model was trained on 320 observations using 100 decision trees. The out-of-bag RASE (3.462951) is significantly higher than the in-bag error (1.509573), suggesting the model may struggle to generalize beyond the training data. Defensive Backs (DB) are the most significant feature, accounting for 32.9% of the model’s importance, followed by Defensive Line (DL) at 32.3%. Linebackers (LB) also play a significant role, contributing 28.3% to the model's predictive power.

Like in the previous analysis, the ML model accounted for much more of the variance when compared to the regression model, improving the r-squared value by a factor of twenty going from 2.5% to 54%. Because of this the ML model is more accurate, suggesting that the relationships in the data are likely nonlinear and/or involve complex interactions between variables.

**Figure 12: Defensive Player Bootstrap Forest Output**



Below, Table 7 shows the same regression analysis as Table 6 but from an offensive perspective, using Offensive Rating as the dependent variable and each Offensive Position Group as independent variables.  This model also uses a min/max scaling weight and shows Quarterback (QB) as the only significant position group.  The key inference to note here is that most teams who have success in offensive efficiency are spending a statistically significant amount of their salary cap on at least one premium quarterback.

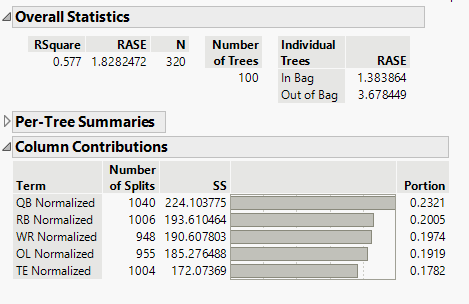
**Table 7: Salary Cap of Offensive Players**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Variables* | *Coefficients* | *P-value* | *Multiple R* | *R Square* |
| Intercept | 4.20217049 | 7.88807E-06 | 0.158385469 | 0.025085957 |
| OL | 0.068436042 | 0.232051645 | - | - |
| QB | 0.151542051 | 0.014163923 | - | - |
| RB | 0.042648169 | 0.433367201 | - | - |
| TE | 0.051312845 | 0.458207873 | - | - |
| WR | 0.086612442 | 0.120190095 | - | - |

Shown below in Figure 13, the output of a random forest model for offensive cap allocation and its effect on offensive performance demonstrates an R² of 0.577, meaning it explains 57.7% of the variance in the target variable. While this suggests a moderate-to-strong level of predictive power, the Root Average Squared Error (RASE) is 1.828, indicating the average magnitude of prediction errors. The model was trained on 320 observations using 100 decision trees. The out-of-bag RASE (3.678449) is significantly higher than the in-bag error (1.383864), suggesting that the model may struggle to generalize beyond the training data. Quarterback (QB) is the most significant feature, accounting for 23.2% of the model’s importance, followed by Running Back (RB) at 20.1%. Wide Receiver (WR) also plays a significant role, contributing 19.7% to the model’s predictive power, while Offensive Line (OL) and Tight End (TE) contribute 19.1% and 17.8%, respectively.

The strong importance of quarterback spending aligns with the regression model's output. Once again, the ML model accounted for much more of the variance when compared to the regression model, improving the r-squared value by a factor of around twenty going from 2.5% to 57%. Because of this the ML model is more accurate, suggesting that the relationships in the data are likely nonlinear and/or involve complex interactions between variables.

**Figure 13: Offensive Player Bootstrap Forest Output**



## Cluster Analysis:

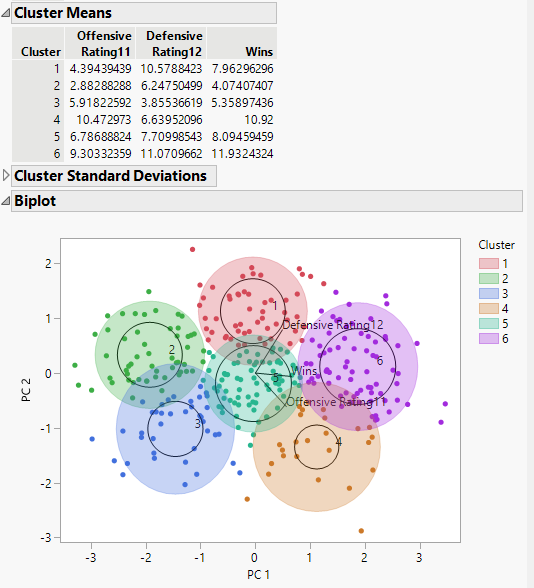
  Depicted in the figure below are the results of our cluster analysis on offensive rating, defensive rating, and win total for all NFL teams over a 10-year period. The K-Means clustering analysis identified a six-cluster solution (K=6) as optimal for grouping teams based on their performance metrics.

Examining the cluster means, teams in Cluster 6 demonstrate the highest offensive rating (9.30), defensive rating (11.07), and win total (11.93), indicating elite teams with strong performance on both sides of the ball. Because of this, Cluster 6 would likely target to extract insightful findings from further analysis. Meanwhile, teams in Cluster 2 exhibit the lowest ratings across all three metrics, averaging just 4.07 wins per season, suggesting struggling teams with weaker overall performance.

Clusters 3, 4, and 5 represent teams with varying strengths and weaknesses, where Cluster 4 teams have a balanced approach with an above-average offensive rating (10.47) and a solid win total (10.92), while Cluster 5 teams show moderate performance levels but lower win totals (5.36).

The biplot visualization highlights how the clusters are distributed in a two-dimensional space using principal component analysis (PCA). The arrows indicate that wins are positively correlated with both offensive and defensive ratings, reinforcing the idea that teams excelling in these areas tend to win more games. The spatial separation between clusters suggests clear distinctions in team archetypes, from dominant, well-rounded teams to those that struggle on both ends of the field.

**Figure 14: Preliminary Cluster Analysis of Team Wins, Offensive Ratings, and Defensive Ratings**



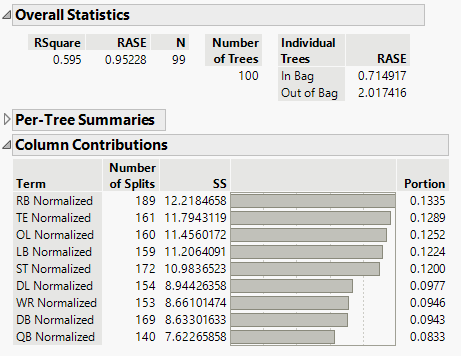
## Cluster Based Models

We decided to run the machine learning models on the data contained in clusters 4 and 6 as they are the ideal place to be for the Jacksonville Jaguars given the highest wins per season and would give us insightful information about these performing teams. Any other cluster would not give us meaningful information and would not lead to a good concluding statement for the Jacksonville Jaguars. We decided to run machine learning models on the offensive positions, defensive positions, and all positions for these respective performing teams and the results are shown below using Machine Learning Models.

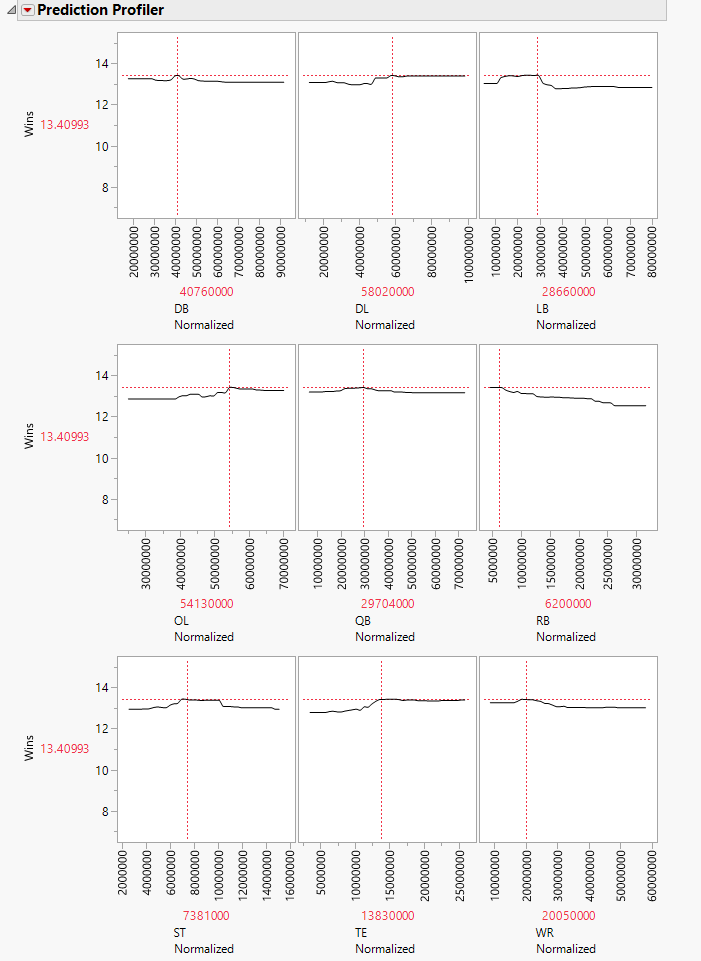
## Machine Learning Models

Below are the overall statistics on all types of positions of the NFL teams in clusters 4 and 6 from the Machine Learning output. As shown in the figure below, there is a greater emphasis on Running Backs (RB), Tight End (TE), Offensive Line (OL), Linebacker (LB), and Special Teams (ST). Additionally, we see a slight drop-off in the other positions compared to the other positions listed. In Figure 6, we can identify that some positions like the running backs are not performing compared to the other positions and show a slight decrease in wins per season.

**Figure 15: Overall Statistics for All positions for NFL Teams in Clusters 4 and 6**



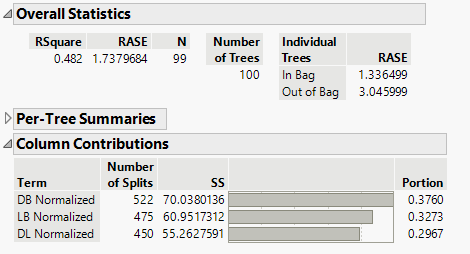
**Figure 16: Prediction Profiler for All positions for NFL Teams in Clusters 4 and 6**



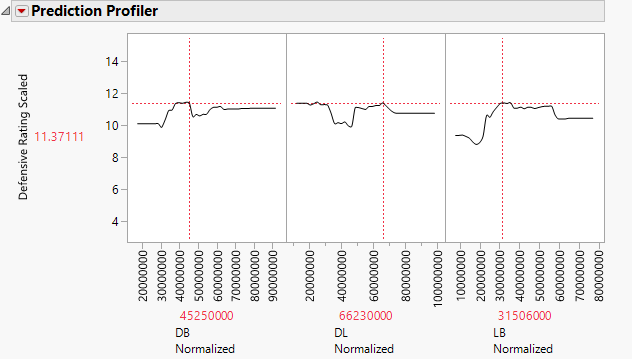
## Defense

Below are the results of the Machine Learning model for Defensive positions. In Figure 17, we can see that given the limited number of defensive positions, the Defensive Back (DB) is shown to be the most significant in this machine learning output as compared to the rest of the positions. Additionally, we see the R² value being on the weaker side compared to the other machine learning model outputs. When analyzing Figure 18, we see that the Linebacker has the most noticeable trend where it starts low, rapidly increases, remains steady, and then starts to decrease slightly.

**Figure 17: Overall Statistics for Defensive Positions for NFL Teams in Clusters 4 and 6**



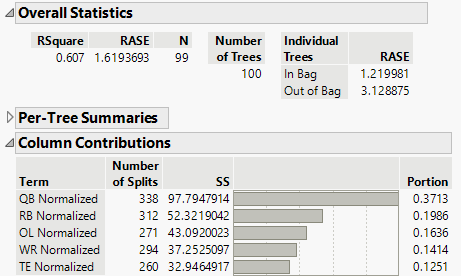
**Figure 18: Prediction Profiler for Defensive Positions for NFL Teams in Clusters 4 and 6**



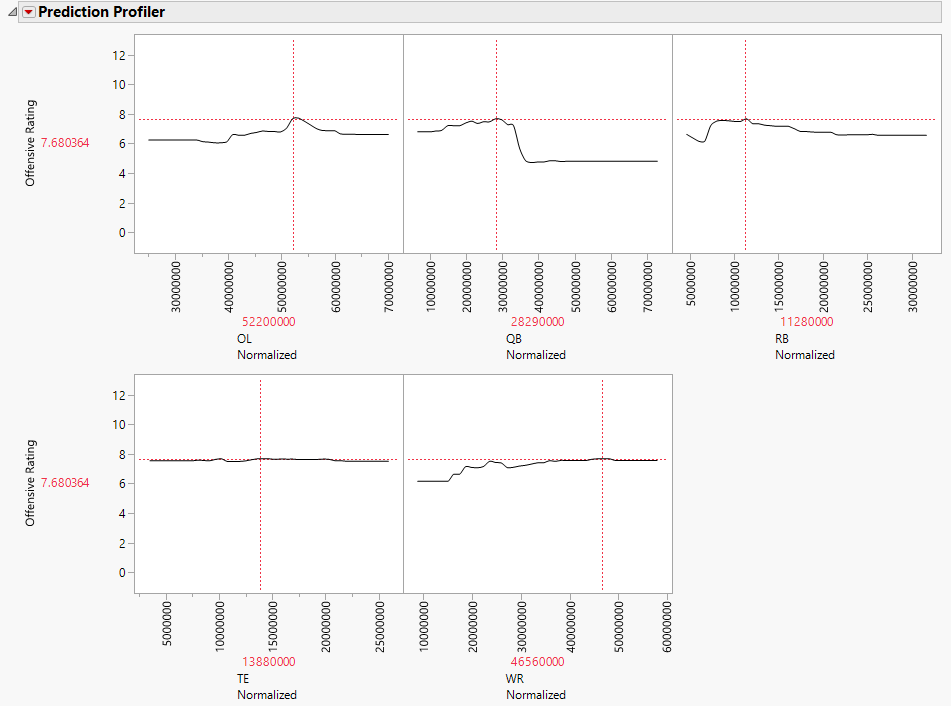
## Offense

Below are the results of the Machine Learning models in the context of offensive positions. In Figure 18, we can see that the quarterback (QB) is by far the most significant position and takes up most of the portion compared to the rest of the positions. Additionally, this output has the largest R² value compared to the defensive positions output, which means there is a greater emphasis on offense than defense when it comes to these performing teams. In Figure 18, we can see there is a major dip in offensive rating for the quarterback and then stays consistent. The Tight End (TE) is the only position that has stayed the most consistent in terms of performance.

**Figure 19: Overall Statistics for Offensive Positions for NFL Teams in Clusters 4 and 6**



**Figure 20: Prediction Profiler for Offensive Positions for NFL Teams in Clusters 4 and 6**



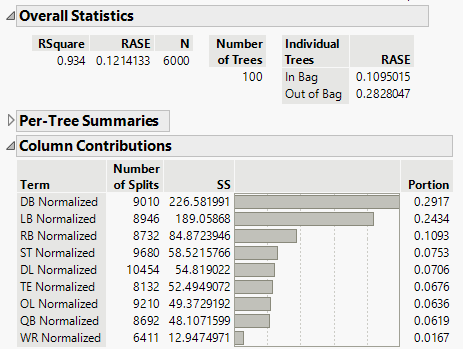
## ML Model Applied to Jags Current Salary Cap

6,000 lines of synthetic Jacksonville Jaguars salary cap allocations were generated using a Python script (Appendix). The dataset incorporates variations in salary distributions, including 10% extreme shifts, where allocations fluctuate by up to 70% within each position group. Additionally, 30% of the data features moderate shifts, with changes reaching up to 49% in certain position groups, while the remaining data consists of small shifts ranging from 5% to 19% across different position groups.

The original machine learning model, trained on historical salary cap data, was applied to the synthetic dataset to evaluate ideal spending strategies by position group for the Jacksonville Jaguars. A new ML model was then run on the synthetic data to analyze how different spending strategies influenced predicted wins the output is shown below in Figure 11.

The R-Square (0.934) of the model indicates 93.4% of the variance in the data is explained, suggesting a strong fit. The Root Average Squared Error (RASE) (0.121413) measures prediction error, with lower values indicating better performance. The RASE values for individual trees are 0.1095 (In Bag) and 0.2828 (Out of Bag), where the lower in-bag RASE suggests the model performs well on training data, while the higher out-of-bag RASE reflects generalization error. Among key variables, DBs was the most influential, with followed by LBs indicating that defensive spending plays a critical role. RBs, ST and DL had moderate impacts. Positions such as TE, OL, and QBs (6.19%) played a smaller role, which is surprising given the traditional importance of QBs. Finally, Wide Receivers had the least impact, contributing only 1.67%.

**Figure 21: Machine Learning Output of Potential Jags Spending Allocations**



The prediction profile depicted below in Figure 22 shows sharp increases in wins when more money is allocated to the Linebacker position and when less money is allocated towards defensive backs given the possible scenarios in the data.

**Figure 22: DB & LB Prediction Profile**

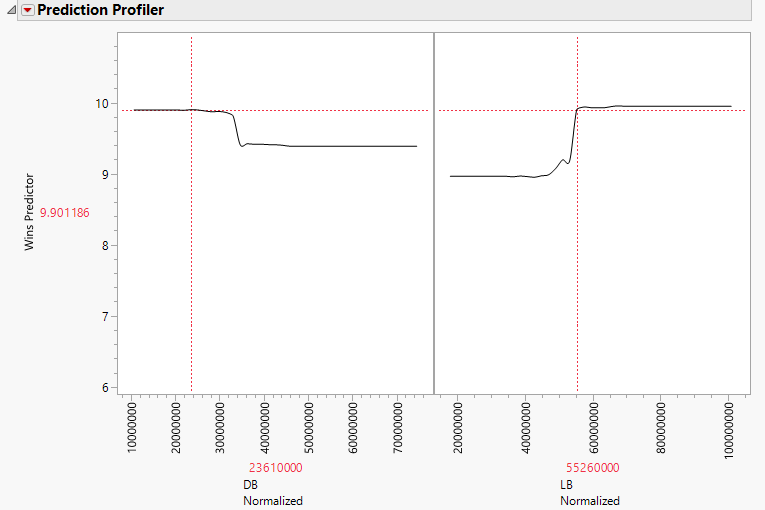


Table 8 below shows the salary allocation that led to the highest number of predicted wins. From Jacksonville's current cap allocation, it involved decreases in DB and DL spending with increases across all other positions. This is feasible due to the increasing of the NFL salary cap for 2025.

**Table 8: Current Allocation and for Maximum Success**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Year** | **DB** | **DL** | **LB** | **OL** | **QB** | **RB** | **ST** | **TE** | **WR** |
| Ideal | 23.61 | 22.84 | 56.12 | 61.12 | 18.32 | 9.11 | 6.96 | 24.87 | 49.54 |
| Actual | 32.82 | 31.08 | 47.07 | 46.90 | 17.00 | 8.65 | 6.24 | 22.98 | 46.43 |
| %Dif | -28 | -27 | 19 | 30 | 8 | 5 | 11 | 08 | 06 |

# Discussions

The findings from this analysis highlight the complexities of NFL salary cap allocation and its impact on team performance. The regression models provided initial insights into which position groups are statistically significant in predicting wins, offensive ratings, and defensive ratings. However, the machine learning (ML) models demonstrated a substantially greater ability to explain variance in team success, suggesting that salary cap allocation strategies involve nonlinear relationships and intricate interactions between variables.

## Key Takeaways from Regression vs. Machine Learning Models

One of the most striking differences between the regression and ML models was the variance explained (R²). The regression models had relatively low explanatory power, with R² values ranging from 2.5% to 6%, while the ML models significantly improved upon this, explaining up to 65.4% of the variance in team performance. This suggests that traditional linear regression may not fully capture the complexities of salary cap allocation strategies, whereas ensemble learning techniques like random forests are better suited for such analyses.

A key discrepancy between the models was the importance of the Running Back (RB) position. In the regression model, RB spending was not statistically significant, yet in the ML model, it emerged as a major contributor to wins. This suggests that RB contributions may depend on interactions with other variables rather than having a direct linear impact on team success. Additionally, Quarterback (QB) spending was statistically significant in both models, reinforcing the long-standing notion that a franchise QB is crucial to offensive performance. However, its overall importance was lower than expected, particularly in the ML analysis, where Defensive Backs (DBs) and Linebackers (LBs) played a more significant role in predicting wins.

## Defensive Investment and Team Success

The results from the defensive-specific models reinforce the importance of Defensive Backs (DBs) in constructing a successful defense. Both regression and ML models identified DBs as the most significant contributors to defensive efficiency, suggesting that teams prioritizing elite DB talent tend to perform better defensively. Additionally, Defensive Line (DL) and Linebackers (LBs) played meaningful roles in the ML models, indicating that a well-rounded defense requires investment across multiple position groups.

Interestingly, the ML model improved R² from 2.5% (regression) to 54%, reinforcing that defensive spending strategies are likely nonlinear. The model also struggled with generalization, as evidenced by the higher out-of-bag RASE (3.46) compared to the in-bag RASE (1.50), suggesting potential overfitting when applied to new data. This indicates that while defensive spending strategies can be optimized, they may not always translate directly into improved defensive ratings due to other external factors such as injuries, coaching schemes, or in-game adjustments.

## Offensive Investment and Team Success

On the offensive side, Quarterback (QB) spending was the only statistically significant variable in the regression model, reinforcing its role as the most critical position in football. The ML model supported this, with QB spending accounting for 23.2% of the model’s importance. However, RBs (20.1%) and Wide Receivers (WRs, 19.7%) also emerged as significant contributors in the ML analysis, suggesting that skill position players play a larger role in offensive success than linear models indicate.

Additionally, Tight Ends (TEs) and Offensive Line (OL) were more influential than initially expected, contributing 19.1% and 17.8%, respectively. These findings suggest that a balanced offensive approach—rather than one solely focused on QB investment—may be more effective in optimizing offensive performance. The ML model improved R² from 2.5% (regression) to 57.7%, further supporting the idea that spending strategies involve complex, nonlinear relationships.

## Implications for the Jacksonville Jaguars

By applying the ML model to 6,000 lines of synthetic salary cap allocations for the Jacksonville Jaguars, we identified optimal spending strategies for maximizing wins. The results showed that Linebacker (LB) spending had the most substantial positive impact on predicted wins, whereas over-investment in Defensive Backs (DBs) led to diminishing returns. This suggests that reallocating resources from DBs to LBs could be a viable strategy for improving team performance.

Additionally, the highest win total scenario involved reductions in Defensive Back (DB) and Defensive Line (DL) spending while increasing investment in all other position groups. Given the expected increase in the NFL salary cap for 2025, such adjustments appear feasible. Notably, Quarterbacks (QBs) played a smaller-than-expected role in win prediction (6.19%), suggesting that Jacksonville should prioritize a well-balanced approach across multiple position groups rather than relying solely on QB performance.

## Cluster Analysis Insights

The cluster analysis identified six distinct team archetypes, with Cluster 6 representing elite teams that excel both offensively and defensively. Jacksonville should aim to align with Cluster 4 or 6, which were associated with the highest win totals. The ML models trained specifically on these clusters revealed a strong emphasis on RBs, TEs, OL, LBs, and Special Teams (ST), while

positions like WR and DB showed more volatility. This supports the idea that a diversified salary cap strategy—rather than an over-reliance on a single position group—is key to sustained success.

# Conclusion

Our analysis investigated the relationship between NFL team salary cap spending and team success, with a focus on formulating optimal strategies for the Jacksonville Jaguars' 2025 roster, given their $279 million salary cap. Using data from the past ten seasons (2015–2024), we examined how allocations to specific position groups influenced overall team outcomes. The core objective was to identify spending patterns that could set the Jaguars up for sustained success in future seasons and avoid the inconsistency they have faced over the past several years.

To do this, we employed a mix of analytical techniques—including multiple linear regression models, K-means cluster analyses, and machine learning (ML) models, particularly bootstrap forest (random forest) models—to assess the relationship between positional spending and key performance indicators such as win totals, offensive ratings, and defensive ratings.

Ultimately, the machine learning models played the most critical role in deriving actionable insights. The ML model was trained on historical salary data (over the past ten seasons) and then applied to over 6,000 simulated salary cap scenarios for the Jaguars. These scenarios included realistic variations in how much the team might allocate to each position group and were produced using python. The model then predicted the expected number of wins, defensive rating, and offensive rating for each spending strategy, allowing us to identify the salary configurations most closely associated with high team performance.

By learning from patterns across the league, the ML model uncovered which positional investments had the strongest association with success—both individually and in combination with others. For example, it showed how strategic increases in Linebacker spending and reductions in Defensive Back investment could result in better predicted outcomes, even when conventional wisdom might suggest prioritizing other positions. This approach enabled us to recommend a more balanced and data-driven allocation strategy tailored to Jacksonville’s specific needs.

However, it’s important to include a disclaimer: machine learning models are based on identifying statistical associations in historical data, not establishing causation. While the model was highly predictive—with the best scenario explaining up to 93.4% of the variance in expected wins—it cannot account for uncontrollable real-world factors such as player injuries, coaching decisions, team culture, or in-game variability. As such, the insights generated should be seen as strategic guidance—not definitive forecasts or causal relationships.

## Address of the Research Question

Our analysis effectively addressed the research objective by combining robust data sources with advanced data methods to generate actionable insights. In addition to descriptive and correlational statistics, the study also leveraged predictive modeling and cluster analysis to test a wide range of hypothetical spending scenarios tailored to the Jaguars. This approach provided a practical, forward-looking answer to the core question: **how should the Jacksonville Jaguars allocate its’ salary cap to improve future performance?**

The use of synthetic data and cluster-based modeling allowed the analysis to simulate real-world decision-making environments, accounting for variation in team strategies and positional depth, making results more relevant to the question. By focusing on high-performing team clusters, the model offered benchmarks grounded in success—not just league averages—making the results comparable to the ideal state of Jags Football.

While the study does not claim to offer a perfect forecast, it presents a scalable, data-informed framework for decision-making. It gives the Jaguars a toolset to evaluate not just where they spend, but how much flexibility they have within each position group to experiment with alternative configurations. This adaptability is a key strength of the analysis.

In evaluating the research objective, the findings offer strong support: the models generated clear recommendations, aligned with observed patterns from successful franchises, and did so in a way that is repeatable, explainable, and practical for a front office to interpret. The approach also invites continuous improvement as new seasons and data become available.

## Recommendations for the Jacksonville Jaguars Front Office

Based on the machine learning insights and cluster-based benchmarking, several actionable recommendations emerge for the Jacksonville Jaguars’ front office as they prepare for the 2025 season:

## Reallocate Spending Toward High-Impact Positions (LBs):

The model identified Linebackers (LBs) as having the strongest positive influence on predicted win totals. The Jaguars should prioritize acquiring experienced, impact LBs—either through free agency or trade—and allocate additional cap space to this group.

## Reduce Overspending on Defensive Backs (DBs):

Over-investment in DBs showed diminishing returns across both league-wide and Jaguars-specific models. While elite DBs remain valuable, the Jaguars should seek cost-efficient depth in the secondary, possibly through the draft or short-term contracts.

## Maintain a Balanced Offensive Approach:

Rather than overcommitting cap space to any single offensive position (like QB or WR), the team should aim for balanced investment across OL, TE, and RB groups. These roles showed strong cumulative contributions in high-performing teams, and a diversified offense is more resilient and efficient.

## Use the Model as a Strategic Planning Tool:

The machine learning framework can be repurposed each offseason to test different roster configurations and simulate expected outcomes. This makes it a powerful decision-support system, not just a one-time analysis. Future cap scenarios, player contract negotiations, or draft class depth can all be incorporated to refine recommendations.

## Benchmark Against Winning Archetypes:

The cluster analysis highlighted characteristics of successful franchises. The Jaguars should align their roster and spending strategy with these team profiles—balancing both offensive and defensive strength rather than focusing narrowly on one side of the ball.

## Draft for Value at DB and WR:

The findings suggest that drafting for depth in DB and WR, while (as mentioned above) spending strategically on premium LBs and OL, is a cost-effective strategy. This approach aligns financial investment with positional return on wins.