## 

## 

## 

## 

**Predicting Running Back Performance using Deep Learning and other Predictive Modelling Techniques**

## 

## 

## 

## 

## 

**By:**

**Dylan Ott**

**Ganga Surendra Basva**

**Kranthi Pedamajji**

## **Executive Report: Structure and Contents**

1. Introduction

1.1 Key Research Questions

2. Scope and Methodology

3. Data Preprocessing, Exploration, and Analysis

3.1 Data Collection

3.1.1 HTML Scrapers

3.1.2 Manual Data Integration

3.2 Data Cleaning

3.2.1 Handling Missing Values

3.2.2 Standardizing Data

3.2.3 Removing Duplicates

3.2.4 Normalization

3.2.5 Outlier Detection

3.3 Data Splitting

3.4 Data Analysis

3.4.1 Data Quality Assurance

3.4.2 First Impressions – Team

3.4.3 First Impressions – Coach

3.4.4 First Impressions – Age and Years of Service (YoS)

3.4.5 First Impressions – Games Started

3.5 Interactive Analysis Using Tableau Dashboard

3.5.1 Performance of Head Coach

3.5.2 Player Performance

3.5.3 Player Years in Service

3.5.4 Key Insight

3.5.5 Conclusion

3.6 Summary of Data Analysis

4. Modeling Methodology

4.1 Neural Net

4.2 Linear Regression

4.3 Random Forest

4.4 Deep Learning

5. Results

5.1 Results of Neural Net

5.2 Results of Linear Regression

5.3 Results of Random Forest

5.3.1 Root mean square error (RMSE)

5.3.2 Prediction Range

5.3.3 Feature Weightage

5.4 Results of Deep Learning

5.4.1 Early Stages of Development

5.4.2 Final Deep Learning Results

5.4.3 Handling Low Totals and Highlight Seasons

5.4.4 Seeing Their Differences in Predictions

6. Overfitting Assessment

6.1 Methodology

6.2 Results

6.3 Conclusion: No Overfitting

7. Discussion and Recommendations for Future Iterations

7.1 Key Findings

7.1.1 Impact of Key Features

7.1.2 Player-Based Insights

7.1.3 Concerns of Overfitting

7.2 Recommendations

7.2.1 Improving Feature Selection

7.2.2 Enhancing Team Data

7.2.3 Advanced Techniques

7.2.4 Handling Outliers

8. Conclusion

8.1 Summary of Findings

8.2 Future Directions

### 

### 

### **1. Introduction**

The modern sports world is highly data-driven, and as companies like Amazon begin to add their technology to statistical tracking and data logging, we continue to see the trend toward analytics being the driver behind numerous decisions in the professional sports world. The NFL in particular is a massive hub of data, spanning everything from player acceleration and physical attributes to their “success percentage” while on the field. This data isn’t only used by teams to determine their staffing either. As sports betting becomes more popular and accessible in the USA, companies like BetMGM, DraftKings, and Fanduel are working to create betting lines that are fair to the customer while also giving the company the best chance to turn a profit. Media outlets like ESPN and NFL Network are relying on their data analysis to generate headlines and conversations around football in order to keep viewers entertained. Much of this entertainment and discussion revolves around prediction and ranking. This became a natural springboard into the work discussed in this report.

Our team has collected and analyzed a moderate amount of data (with respect to the amount of data available) in an effort to accurately predict the performance of a given NFL player by the end of the regular season. Because there exists such a wide variety of players in the NFL, and their positions all impact the game in important but different ways, we thought it would be appropriate to narrow it down to one position on the offense - the running back - and one statistic in particular - rushing yards. This position and stat were chosen for a few reasons. First, running back rushing yards only involves one person touching the ball, rather than passing yards which would be dependent on both the quarterback and the receiver. While the rest of the offense certainly impacts a team’s rushing performance, the fact that this stat is fueled largely by the person touching the ball made it a fantastic candidate for prediction. Next, total rushing yards is a common betting line set by casinos, as well as a key stat for the NFL Offensive Player of the Year award. Finally, as stated above, the entire offense is considered to play a role in a team’s rushing offense - including the head coach and offensive coordinator’s scheme. These factors made it a great target for predictive modeling. By the end of this report, we hope to be able to answer a few common questions that fans have when trying to predict how a player will perform throughout the season

#### **1.1 Key Research Questions:**

By analyzing the data, we aim to address the following critical questions often asked by football fans, analysts, and decision-makers:

1. **Does age play a significant factor in running back performance?**
   * Older players are often believed to experience a decline in physical output. We aim to validate or challenge this assumption.
2. **Do coaches or teammates have a greater impact on a player’s performance?**
   * How much of a running back’s success can be attributed to coaching strategy or the support of teammates (e.g., offensive line strength, quarterbacks)?
3. **Is individual skill truly decisive?**
   * Are there measurable “X-factors” that distinguish elite players from average ones?

### **2. Scope and Methodology**

For this particular study, historical data has been collected from 2014 to 2023 to quantify performance indicators toward understanding the impacts of age, team performance, and coaching in predicting rushing yards using advanced preprocessing and machine learning techniques for data analysis. To validate the robustness of the predictive hypothesis, the model was trained on historical data from 2014 to 2022 and was then tested on that of the 2023 season. Also, the extended evaluation will include real-world scenarios through the prediction of rushing yards gained by current 2024 NFL season running backs through week 15 of the 2024 season (the most current data at the time of writing). With this, a field test can be carried out to observe how the model extends its predictions to an unseen, live sporting environment.

By the end of the analysis, we aim to:

* Identify the most significant factors influencing rushing performance.
* Provide actionable insights that can assist teams, analysts, and fans in evaluating running back performance.

### **3. Data Preprocessing, Exploration and Analysis**

#### **3.1 Data Collection:**

To build an accurate predictive model, we collected historical NFL player data from the **Pro Football Reference** website (<https://www.pro-football-reference.com>). This process involved the following steps:

##### **3.1.1 HTML Scrapers:**

* + We developed **two HTML scrapers** to automate data extraction:
    - **Scraper 1**: Collected player performance statistics (e.g., rushing yards, touchdowns, games played).
    - **Scraper 2**: Collected team rosters and coaching details (e.g., head coach, offensive coordinator).

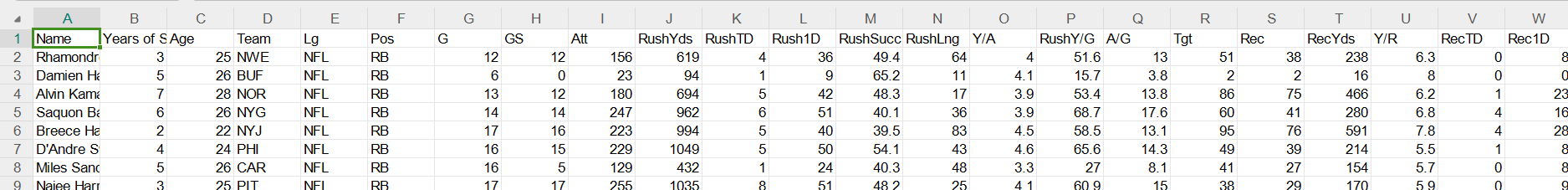
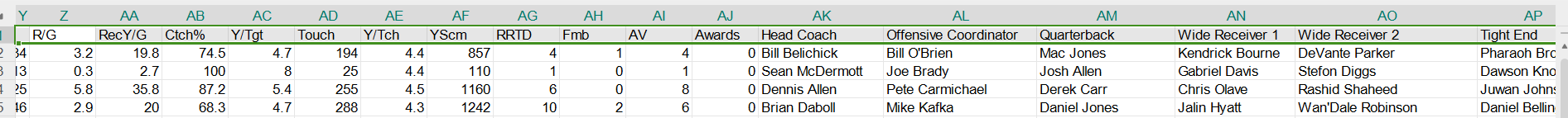
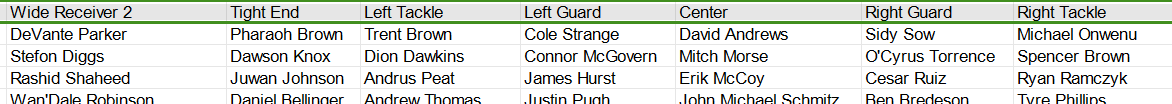
##### **3.1.2 Manual Data Integration:**

The extracted data was manually copied into two spreadsheets for further processing and analysis:

* + - **Training Data** (Data.xlsx): Contains **785 rows** and **47 columns**, covering historical player and team data from **2013 to 2022**.
    - **Testing Data** (TestData.xlsx): Includes **69 rows** and **47 columns**, focusing on player attributes and team performance for the **2023 NFL season**.

The resulting datasets included key features like:

* **Player Attributes**: Age, years of service, games played, and rushing metrics.
* **Team and Coaching Information**: Team name, head coach, offensive coordinator, and roster composition. As shown Below

#### **3.2 Data Cleaning:**

### To ensure the data was clean, consistent, and ready for analysis, we used **Google OpenRefine**, a powerful tool for data cleaning and transformation. The following steps were performed:

##### **3.2.1 Handling Missing Values:**

Identified missing values and filled them using appropriate strategies (e.g., imputation for numerical features and “N/A” for categorical features).

##### **3.2.2 Standardizing Data:**

Ensured consistency in text-based fields such as **team names**, **player positions**, and **coach names** by removing inconsistencies like extra spaces, typos, and variations.

##### **3.2.3 Removing Duplicates:**

Identified and removed duplicate rows to ensure data integrity.

##### **3.2.4 Normalization:**

We experimented with **Z-score normalization** to scale the features. However, it was observed that scaling negatively impacted model performance for Deep Learning models, so raw values were retained.

##### **3.2.5 Outlier Detection:**

Initial outliers in performance metrics like **rushing yards** and **touchdowns** were flagged for review using boxplots. These outliers were retained because standout performances often represent valuable insights in sports data and predicting these outstanding seasons is in part the goal of this model. If the model is never trained on or ‘aware of’ these outlier seasons, then it will be unlikely to predict any outstanding seasons in the future.

#### **3.3 Data Splitting:**

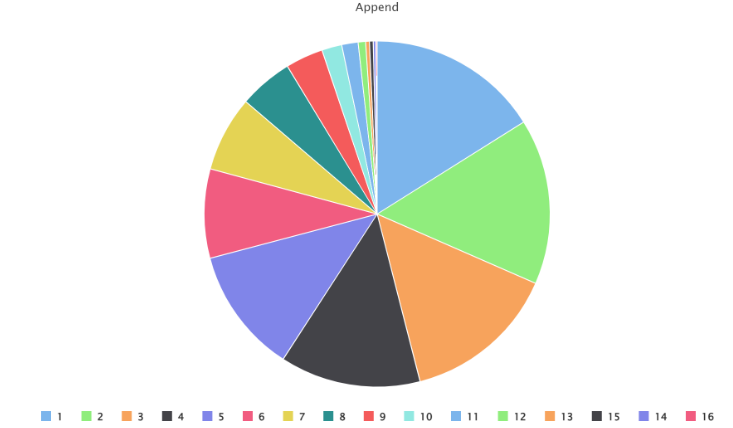
The cleaned dataset was split as follows:

* **Training Set**: Data from **2013 to 2022** (785 rows) to train the models.
* **Testing Set**: Data from **2023** (69 rows) to evaluate model performance.

#### **3.4 Data Analysis:**

##### **3.4.1 Data Quality Assurance:**

Among the different years of service (YoS) in the NFL, this pie chart shows the distribution of running backs into several segments, each representing the number of players in terms of the number of years in the league.

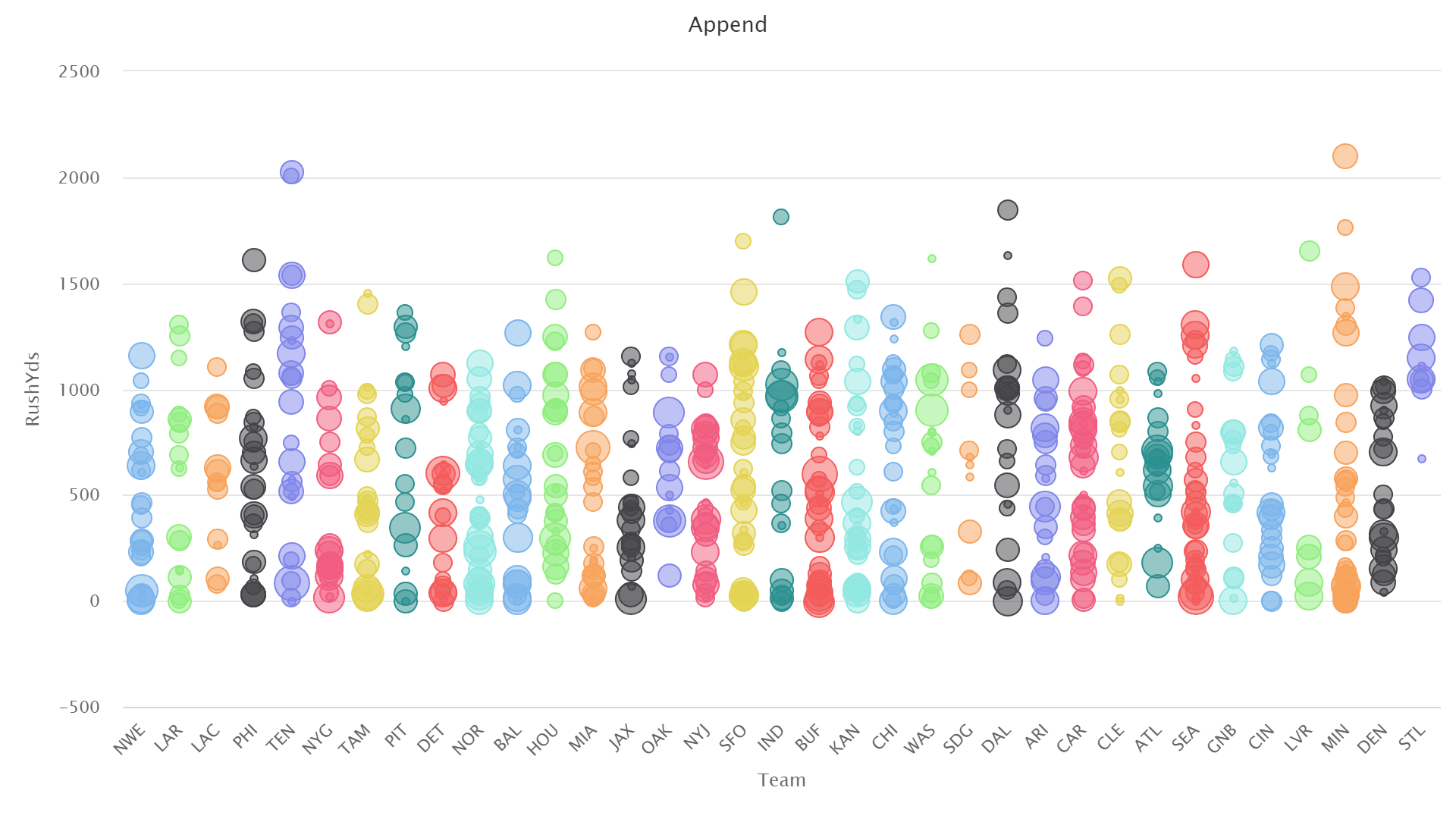
**Insights:**

* It has a nice distribution for various experience levels.
* It is mainly 1, 2, or 3 years of service players, which indicates that rookies and early-career players are most prolific in our dataset. This is beneficial, as today’s average NFL running back has a career of less than 3 years because of the way that organizations value youth at the highly physical position.
* It balances with the other end of the spectrum: those who have been in the league for less than 7 years are considered experienced; the other group in the table possibly includes mid-career players who have been in the league for somewhere between 4 and 6 years. This ensures that performances captured on a dataset will come from all stages of a player's career.

**Conclusion:**

While Years of Service is just one of our many variables, we’ve shown it here to summarize the diversity of our dataset rather than prove that every single attribute has a wide range of values. The diversity in terms of Years of Service ensures that the model does not have a bias either toward rookies or veterans. This becomes important to ensure that predictions become valid irrespective of career stage. As we continue to explore our data, we’ll begin looking at the attributes that are commonly associated with rushing performance to discover any trends in our dataset that may inform the expectations for our model’s weights and performance.

**3.4.2 First Impressions – Team (Attribute):**This bubble chart shows rushing yards (Y-axis) by team (X-axis), and each bubble represents years of service (YoS) for each player by size.

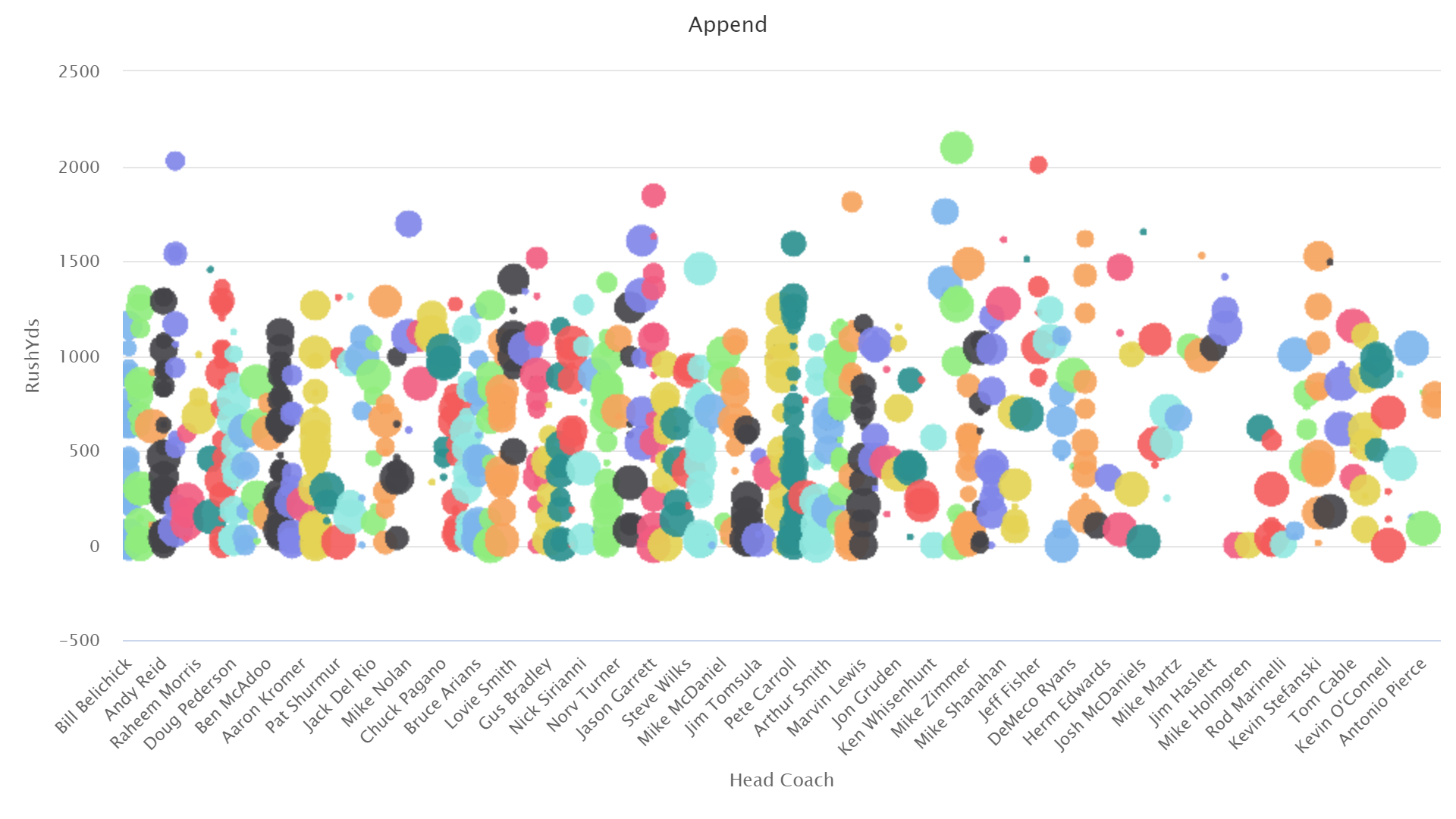
**Insights:**

* Rushing yards are pretty much distributed all over the cities, although not particularly on a team level concerning performance.
* Some of these entries have exceptionally high rushing yards (some exceed over 1800 yards, where our average is around 550), but these performances seem to come from the same subset of players, which points to the player being the key factor rather than the team.
* Bubble size indicates that both early-career players as well as veteran players are capable of performing at a high level; there is no immediate advantage or disadvantage regarding their tenure on specific teams.

**Conclusion:**

The above chart proves that attachment to a team does not directly associate itself with performance on rushing hence portraying the other impacts that player-centric property and usage effects have on them as well.

##### **3.4.3 First Impressions – Coach:**

This bubble chart displays Rushing Yards (Y-axis) with the Head Coach as the X-axis, while bubble sizes illustrate Years of Service (YoS).

**Insights:**

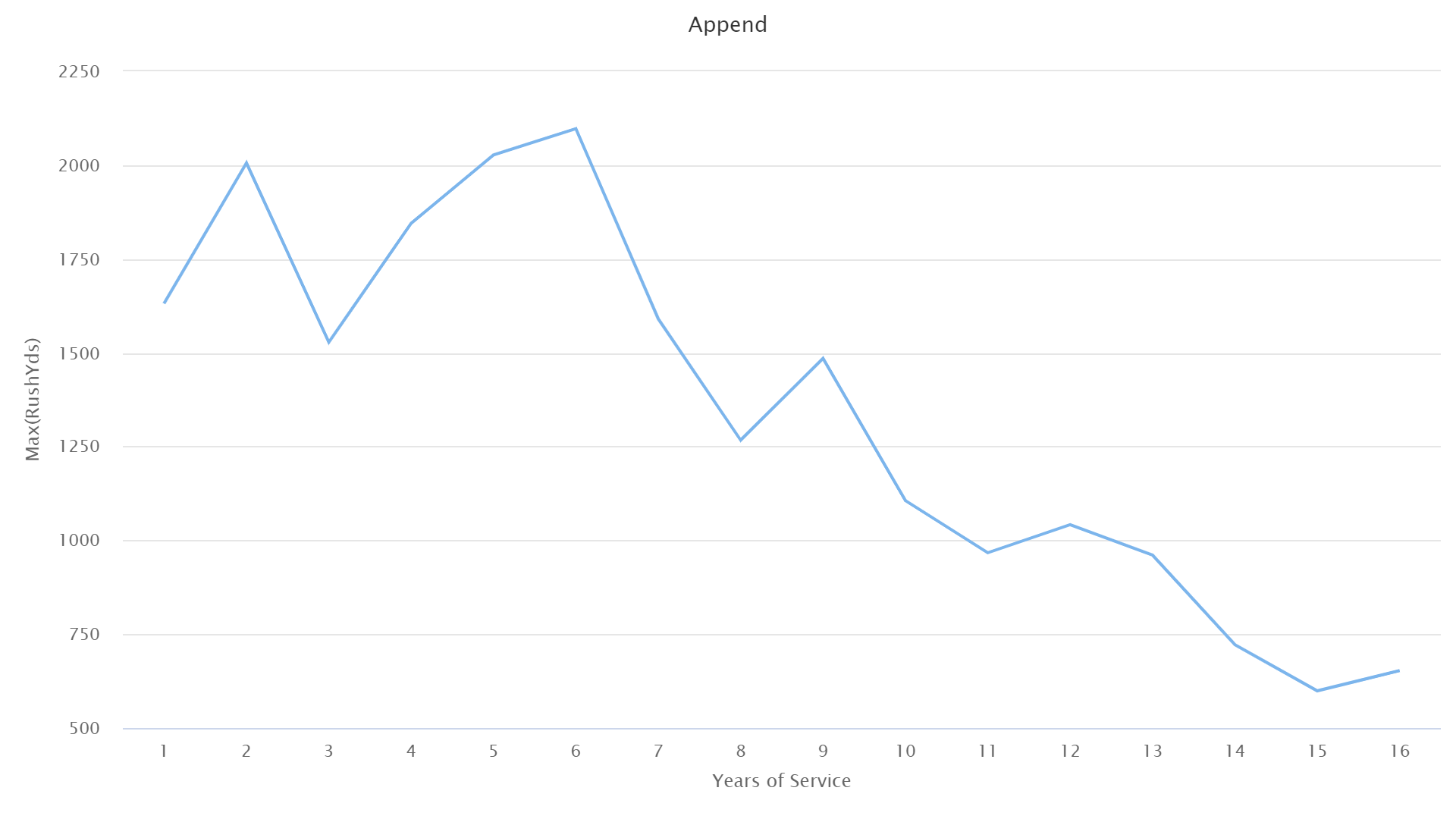
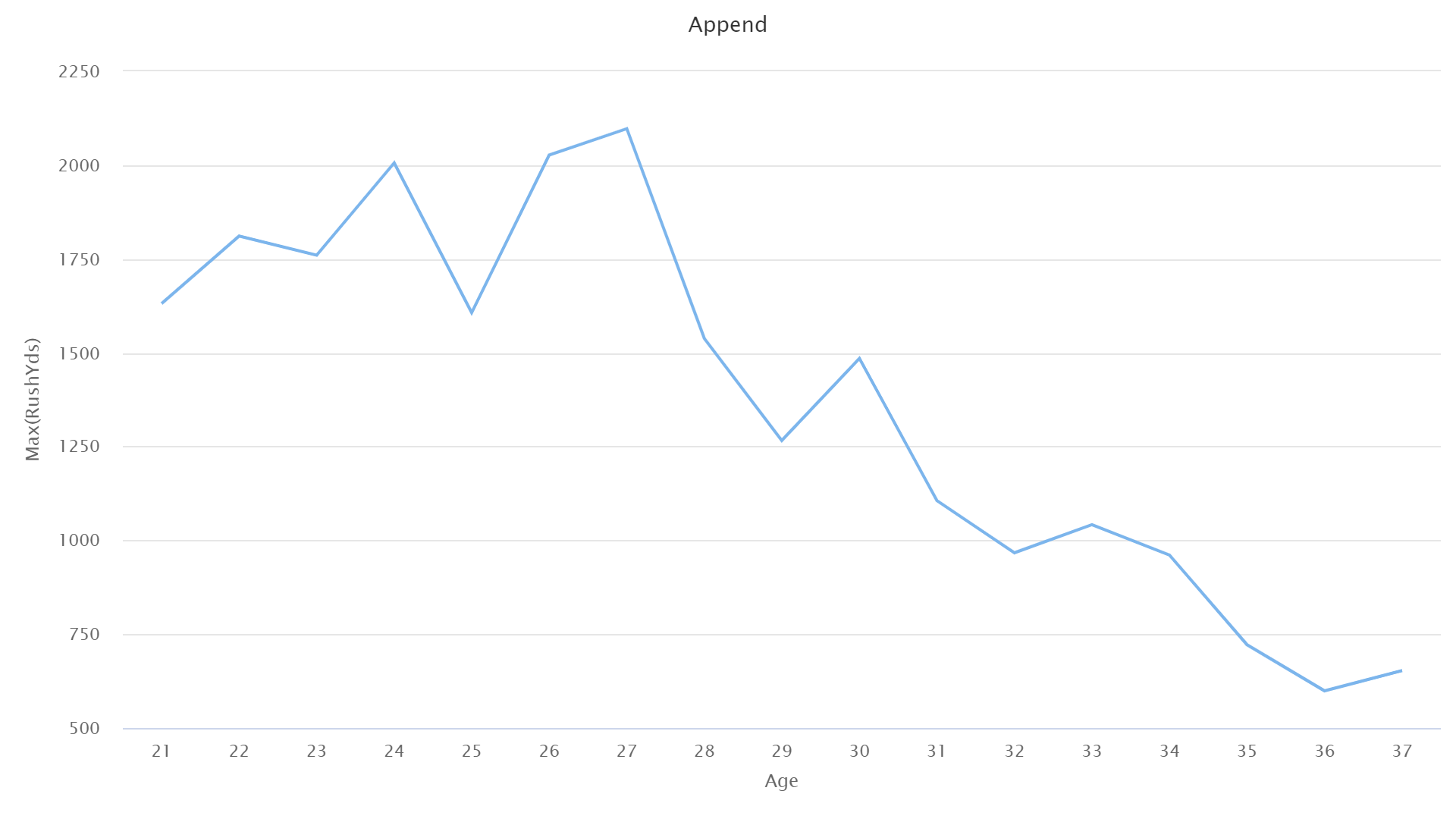
* A common theory concerning NFL player performance is that the coach and their scheme has a heavy impact on a player’s ability to contribute and perform at a high level. If that is the case, some of the coaches in our data set should have very few entries towards the bottom of our Y-axis.
* Some players have been proficient under particular coaches, but these cases do not present a pattern around which one would collect specific data. Similar to the previous findings, this would seem to indicate that the player themself - their physical attributes or capability - that plays a key role in their performance.
* As bubble sizes again show a cross-section of veteran and younger players under each coach, this may suggest that coaching effects are not always directly apparent through rushing yards averages.

**Conclusion:**

That coaching strategy is important indeed. Yet, the present analysis indicates that talent and opportunities available to individual players contribute more to rushing yards than the singular head coach.

##### **3.4.4 First Impressions – Age and Years of Service (YoS):**

The first line chart displays the relationship between Years of Service (YoS) and rushing yards. The second line chart displays the relationship between age and rushing yards.



**Insights:  
Age:** Rushing yards peak around the mid-20s of the players and then taper off towards the old age. This pattern denotes that young players outperform the old veterans as they are physically much fresher.

**By Service:** Rush performance tends to be optimum with around 3-5 years of service. Beyond this point, it starts to decline, likely due to injuries and reduced efficiency.

**Conclusion:**

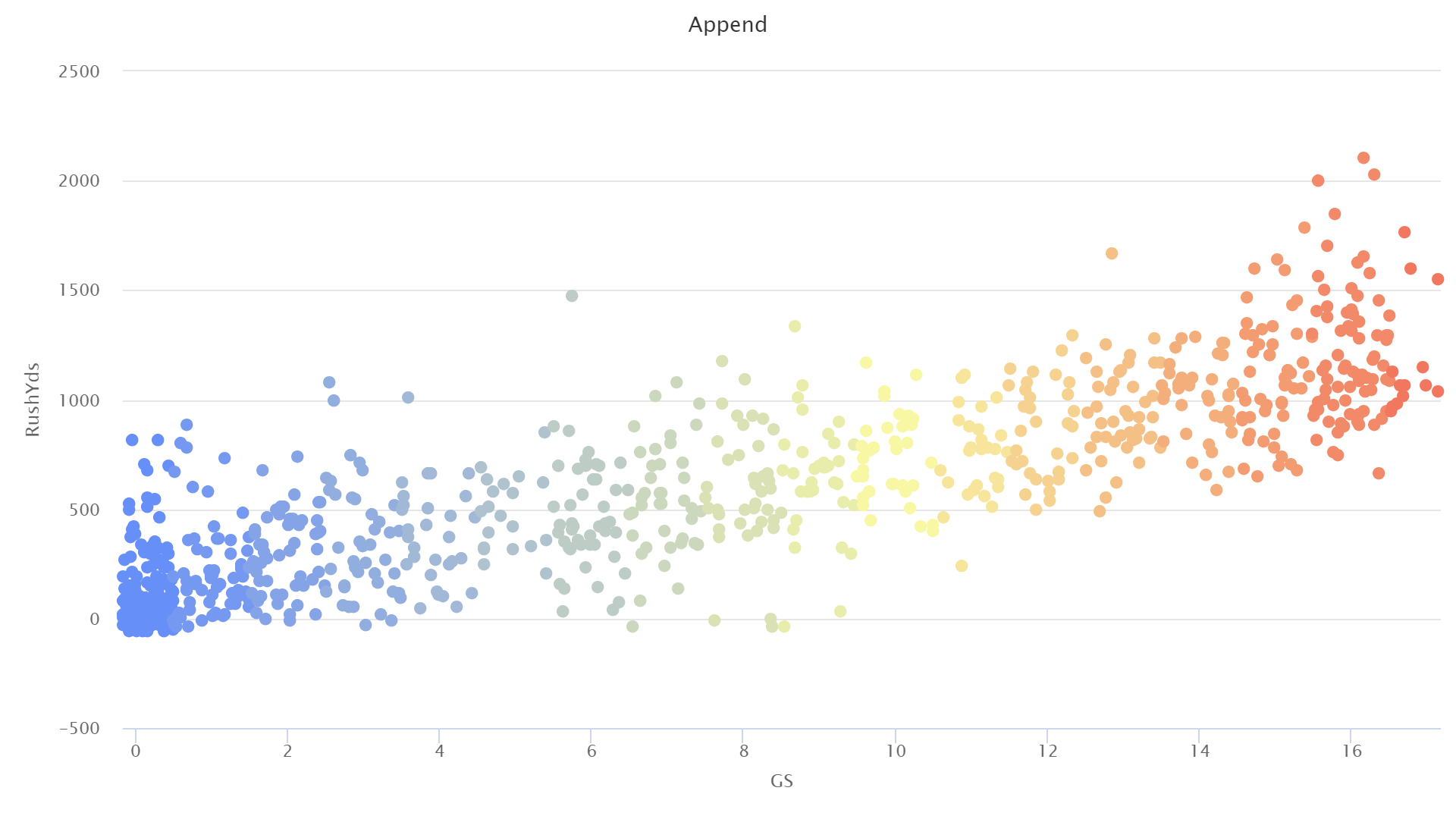
Both Age and YoS are attributes that affect rushing performance. Young and middle-career players (with 3-5 years of experience) generally outperformed the rest; hence, these are appropriate for model predictions.

##### **3.4.5 First Impressions - Games Started:**

The scatter plot shows Rushing Yards (Y-axis) against Games Started (GS) (X-axis), with a gradient color indicating increased values of GS.

##### 

**Insights:**

* Positive correlation between games started and rushing yardage: those who start more games generally have higher rushing yards.
* This chart reflects the most straightforward trend: the more games a player starts, the greater the chance he has to receive rushing attempts and the more yards he accumulates.
* This is not a hard and fast rule. There are players who have started just a few games but have managed to rush for considerable amounts of yards, which hints at effectiveness or remarkable single-game performances. Additionally, some players who have more than 6 starts record very low yardages over the course of a season, pointing maybe to injury or perhaps a shift in their role over time.

**Conclusion:**

The total number of games that have started is a good predictor of rushing performance. It mirrors the opportunity given to a player throughout a season and should probably be an important constituent in the predictive model.

#### **3.5 Interactive analysis using Tableau Dashboard**

##### **3.5.1 Performance of Head Coach:**

The first bar chart shown below depicts head coaches according to their team's average rushing yards. The efficient plays or utilization of players' talents by star players tend to always bring in higher average rushing yards as compared to normal finds by other coaches.

##### **3.5.2 Player Performance:**

The second bar chart illustrates the single high-average rushing yards measured personally. Exceptional players really contribute to performances regardless of the other features such as the coach or indeed the team to which they are affiliated.

##### **3.5.3 Player Years in Service:**

The table establishes a relationship between Years of Service to performance.

##### **3.5.4 Key Insight:**

The greatest average rushing yards numbers are recorded by relatively inexperienced players (3-5 years). Thus, we would expect our model to predict higher performances for players entering this stretch of their careers.

##### **3.5.5 Conclusion:**

This Tableau Dashboard expertly consolidates the insights in terms of the influence of coaching, player performance, and experience, and offers the best analysis extraction points for predictive modeling using interactivity.

#### **3.6 Summary of Data Analysis:**

1. **Years of Service** and **Age** reveal expected trends in performance decline with time.
2. **Games Started** is a strong predictor of rushing yards due to increased opportunities.
3. Neither **team affiliation** nor **head coach** shows a strong standalone correlation with performance, emphasizing individual player attributes.
4. The **Tableau dashboard** validates these findings and highlights notable coaches and players.

### **4. Modelling Methodology:**

When beginning to develop predictive models, we took into consideration what inputs from our dataset would result in the most usable and beneficial model in a real-world setting. Our data contains every relevant stat, including total rushing attempts, yards per game, and other statistics that are calculated by their overall stats. Immediately, we knew that using these as inputs to the model would be detrimental to its usability, regardless of how much better it performed. If a general manager needed to know a player’s end-of-season stats to predict their end-of-season performance, the model would be useless. As such, the model was trained on and given as input the following variables from our dataset:

* The running back’s name, age, and years of service,
* The name of the running back’s team, head coach, offensive coordinator, wide receivers, and offensive linemen, and
* The running back’s total rushing yards (the target variable).

#### **4.1 Neural Net:**

Given that this problem involves several non-linear relationships that are hard to visualize and describe, a neural net seemed like the natural choice. However, neural nets are not well-suited to polynominal data such as names or abbreviations, so these fields had to be converted to numerical equivalents. Because several of these fields could have the same value in different variables that would represent the same data point (e.g., an offensive coordinator who became a head coach), all 1440 points of our nominal data had to be mapped to unique integers in one large map. This map was used when training and applying the model to allow the neural net to run on our data. Afterwards, the output had to be unmapped using a reverse process and the same map.

The net we used featured 3 fully connected layers of 10 nodes. The number of nodes was auto-populated by Altair Studio RapidMiner based on the number of inputs. This net was trained over 500 cycles at a learning rate of 0.02 and a momentum of 0.9 to avoid local maxima and minima. Decay, data shuffling, and normalization were also employed to improve the model’s performance. Adding even just one extra layer resulted in the model making the same prediction across all entries, which was roughly the average of the data set.

#### **4.2 Linear Regression:**

The design for our linear regression model followed that of our neural net, due to this node also being unsuited to polynomial data. After setup was complete, we had a linear regression model using the M5 Prime function with a minimum tolerance of 0.05.

#### **4.3 Random Forest:**

While not as well-suited to numerical prediction as it is classification, the flexibility of a Random Forest model encouraged us to give it a chance. The first encouraging sign was the model’s acceptance of our raw data, rather than needing to normalize or map any values to work.

The final design of our forest featured 425 trees using the least square criterion for deciding how to split their decisions. Pre-pruning was also enabled in this model, requiring 5 pre-pruning alternatives. Additionally, tree depth was not limited, so as to allow them to gain as much information as possible, no matter how small.

When testing this model, we had a disappointing RMSE of 213. Confused by the drop in performance, we investigated the weights and discovered a key difference - the random forest model chose Years of Service and Age as its highest weights (behind games started, of course). This divergence from using head coach and quarterback as its key weights is what we believe caused the degraded performance when compared to other models.

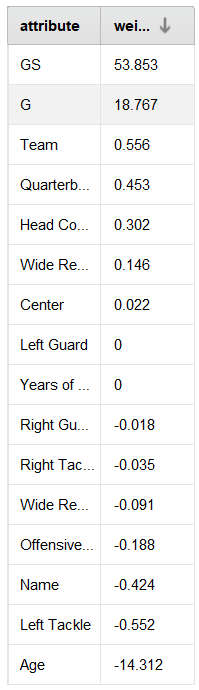
#### **4.4 Deep Learning:**

Seeing that the neural net had the best performance overall thus far, the natural next step in attempting to improve it was to step up to a Deep Learning model. Once again this model was simple to work with, as it consumed our data without the need for any mapping or normalization.

Our final deep learner had nine layers of 100 nodes each using a ReLU activation function over 28 training epochs. The model was set to train for 30 epochs, but it would stop early at 28 due to not meeting its learning threshold anymore.

### **5. Results:**

#### **5.1 Results of Neural Net:**

After running this model, we return a root mean squared error (RMSE) of 164.710, which is to say that our prediction was 164 yards off on average. This is adequate performance for the number of inputs given, and we were happy with this result, but felt the overall error could be lowered by a different model that involved a less cumbersome process. Looking at the resulting model, it seemed to heavily weigh games started, quarterback, and head coach when making its predictions. 

#### **5.2 Results of Linear Regression:**

Once again, we achieved a satisfactory RMSE of 165.94, slightly worse than our neural net but not significantly so. As we explored the weights that came out of this model, we once again saw that games started was the most significant factor by far, followed by team, quarterback, and head coach (although at much lower values). All other factors had slightly negative weights that were loosely bounded by -0.5 aside from age, which had a strong negative weight of -14. This is interesting, as it directly opposes the common theory that running back age is a determining factor in their performance.

#### **5.3 Results of Random Forest:**

The performance of the Random Forest model was average, but it did have certain drawbacks:

##### **5.3.1 Root mean square error (RMSE):** 213.655 was the RMSE of the model, which can be improved, especially more in the case of outliers.RF_Results.png

##### **5.3.2 Prediction Range:**

* The model almost never predicted less than 250 yards or over more than 1000 yards, although these performances are quite common in the NFL.
* For example, under-predicting signature players like Christian McCaffrey showed a lack of sensitivity in the prediction of some phenomenal seasons.

##### **5.3.3 Feature Weights:**

* Games Started can be defined as the most important feature with a weight of 53.853; followed closely would be Games Played (G) at 18.767.
* Other important attributes like Team, Coaches, and Offensive Line did not have much weight which might have resulted in these attributes not being able to capture team dynamics too well within the model.

**Key Insight:**

Although the Random Forest model learned patterns from the data, it performed poorly on extreme predictions and portrayed a disinterest in using team-related features, which de facto reduced its scope of accuracy concerning those standout performances.

#### **5.4 Results of Deep Learning:**

Deep learning showed superior performance over Random Forest with less RMSE at 161.481, as against 213.655, which indicates better accuracy. The weights from this model aren’t as straightforward to read, because it weights each attribute and its possible value individually. Still, it’s evident from our investigation games started, that Name, Team, and HC are all important, with certain players having a bigger impact than others (none as great an impact as the RB themself, though). Interestingly, this is the only model to have the player’s name consistently show up as a key factor. We believe that this is what drives the RMSE lower than the other models, even if only slightly.

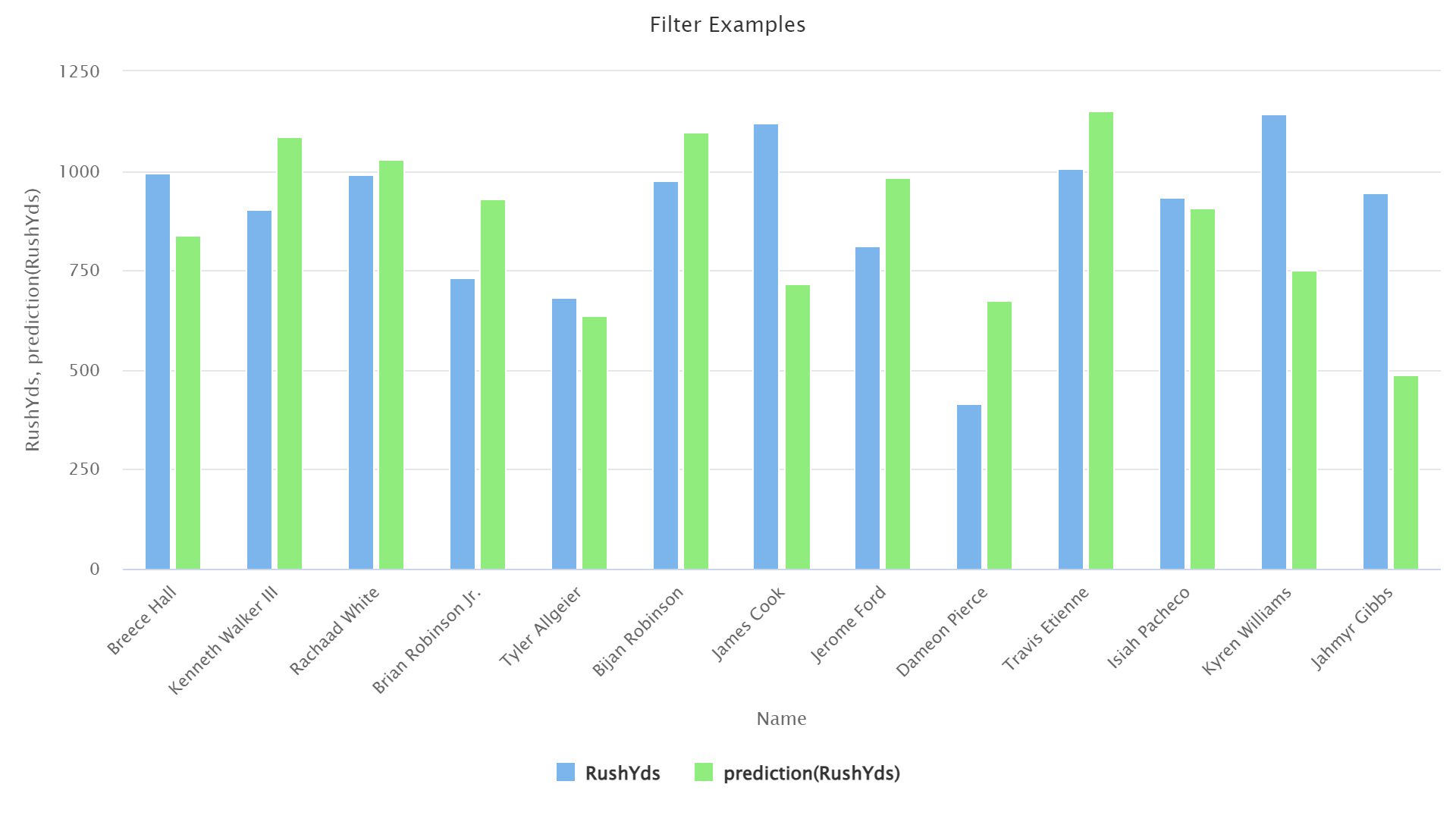
**Low Totals:** This model effectively predicted low totals of rushing yards underestimating the amounts as seen in Random Forest.

**Highlight Seasons:** A key aspect of our model’s performance that we mentioned earlier was the inclusion of outliers to assist our model in predicting standout seasons from players. The deep learner showed that it was capable of this, predicting the highest rushing value for the 2023 season to be 1,333 yards, which was roughly our RMSE off of the actual value, of 1,459 yards. It attributed this to the wrong player, naming Derrick Henry as the rushing leader instead of Christian McCaffrey, but this is still evidence that the model is capable of predicting the sort of season that can be expected from the input data.

**Rookie Outcome:**

Another key aspect of our model is assessing rookie data, as it would have no prior knowledge of these players. Looking at the subset of players with less than 2 years of service in our test set, it managed the rookies and early professionals quite well, with predictions for such players being deduced accurately irrespective of sparse historical data.

##### **5.4.4 Seeing Their Differences in Predictions:**

This bar chart visualization of predicted rushing yards (in green) along with actual rushing yards in blue brings forth the model's accuracy for the individual players.

For the prediction aspect, for almost each of the incoming players, Bijan Robinson and James Cook, the predicted values came out very close to the actual rushing yards.

However, for some standout players like the returning rookie Jahmyr Gibbs, rather than performing above standard expectations by a huge mile during the season, minor over-predictions popped up.

**Key Insight:**

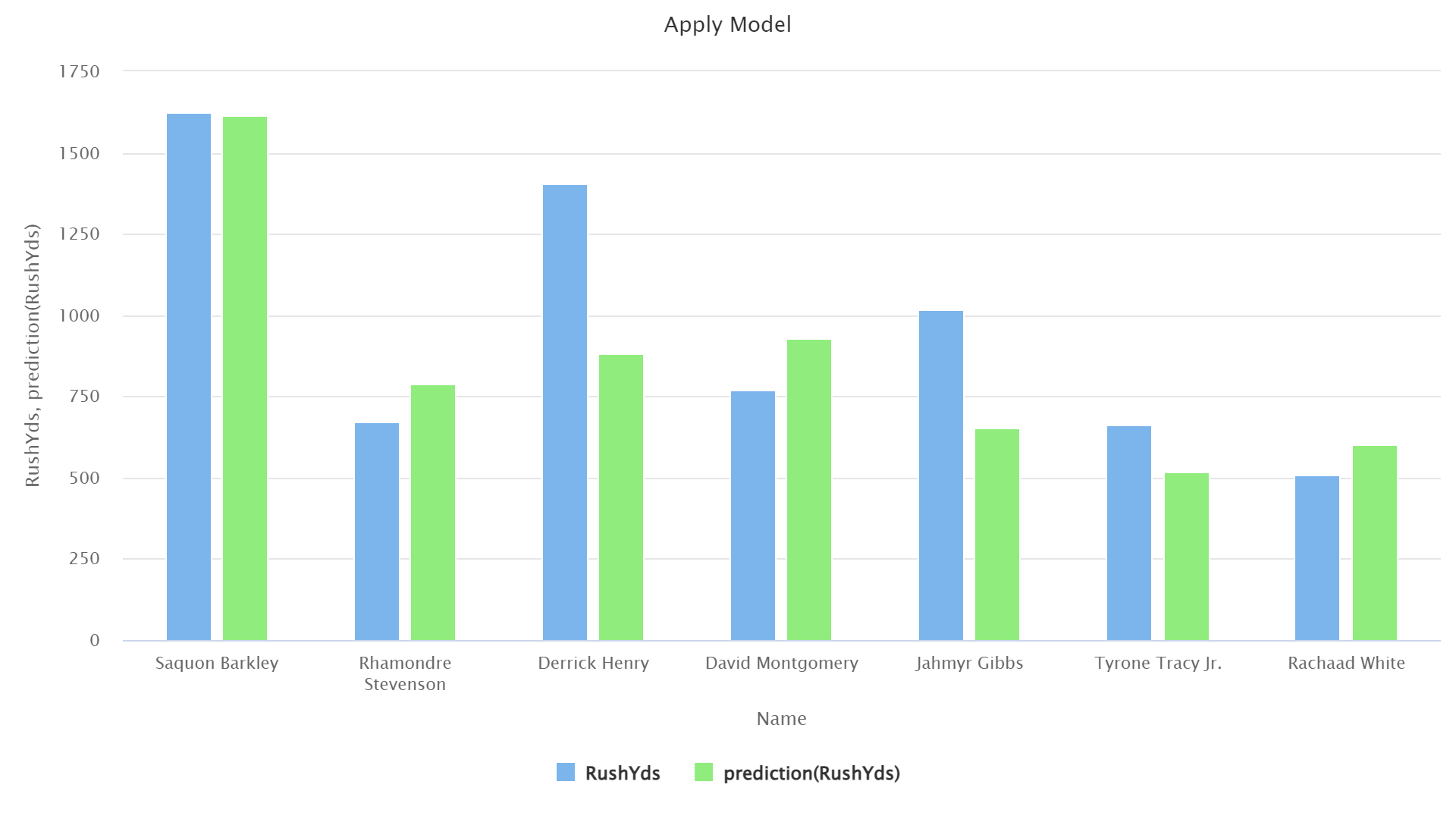
The Deep Learning model produced solid predictions for every category of player, making it a far better choice for real-life predictions.

### **6. Overfitting Assessment:**

It is a very common issue that arises with the deep learning models with larger architectures: Overfitting. To evaluate if the model generalized well beyond the training data, new performance data were collected from selected players, which were used for validating the model.

#### **6.1 Methodology:**

Model predictions were tested against a cohort of modern players, including:

* **Saquon Barkley:** A luminous star now shining above across a new level, on pace to break rushing yards records. At the time of running the test, he had a league-leading 1,623 yards - our model predicted he would have 1,615.
* **Derrick Henry:** Henry is known for high-yardage seasons, but his new team (the Baltimore Ravens) has had running back struggles recently. Our model still predicts him to have an above-average year, but this lean-on historical data from Baltimore is likely to blame for his low prediction.
* **Jahmyr Gibbs & David Montgomery:** A competent running back duo powering through one of the league's strongest offensive lines. This two-headed attack is likely why Jahmyr Gibbs is under-predicted because David Montgomery is the player getting more ‘starts’. If we were to artificially inflate Gibbs’ games started, we believe we’d see a much more accurate prediction.
* **Tyrone Tracy Jr.:** Despite not having prior knowledge of Tracy, the model still accurately predicts the one bright spot on a historically failing offense.
* **Rhamondre Stevenson and Rachaad White:** These two players are included as ‘control’ variables, as they did not have much change around them this season.

#### **6.2 Results:**

The model exhibited overall excellent predictability, being able to predict values for almost all the players that very closely mimic the actual performance trends in the real world.

The greatest example is the prediction that mentioned Saquon Barkley's breakout season; it indicates the agility of the model to change conditions both in players and teams.

#### **6.3 Conclusion: No Overfitting**

Testing indicated the Deep Learning model against overfitting; which was very impressive considering its nine-layer architecture and high-fangled training parameters. The model also generalized fairly well on new data, and that makes it all the more credible and robust for real-world application.

### **7. Discussion and Recommendations for Future Iterations:**

#### **7.1 Key Findings:**

This deep learning model surpassed the other models tested in the area of estimate prediction for low to mid-range rushing yard totals and generalization of new players and unseen data. The following key findings emerged from the study:

##### **7.1.1 Impact of Key Features:**

Games Started (GS), Games Played (G), and Name were important predictors of rushing yard performance. Although it had Team and Coach data, it contributed less than that and hence needs improvement in weightage or adding more team-related metrics.

##### **7.1.2 Player-Based Insights:**

* **Rookies:** The model performed well but showed slight exaggerations in predicting outstanding rookie performers like Jahmyr Gibbs, who outperformed expectations.
* **Exceptional Seasons:** Predicted outliers where very few predictions were above average but not quite to the accuracy that is better than Random Forest.

##### **7.1.3 Concerns of Overfitting:**

The Deep Learning model has all the potential to avoid overfitting, given 9 layers, 100 nodes in each, and the performance analysis of the model produced consistently without change to new data.

#### **7.2 Recommendations:**

The following suggestions are recommended for improving model performance and mitigating concerns surrounding limitations in the model:

##### **7.2.1 Improving feature selection:**

* This means adding more player stats like weight, height, and fitness scores, which may give more clues to performance.
* Add the pre-season statistics to aid early-season prediction where there are training camp statistics and injury reports.

##### **7.2.2 Enhancing Team Data:**

* Improved weighting of team variables like Head coach and Offensive coordinator Dynamics on Offensive Line performance variables.
* Support it with some advanced team-level stats like offensive scheme type or frequency of rushing play from the team.

##### **7.2.3 Advanced Techniques:**

* Stacking methods Exploration: Combining outputs between Random Forest and Deep Learning into a meta-model will enhance the accuracy of meta-model testing.
* Trying other advanced models like Gradient Boosting or XGBoost for performance comparison.

##### **7.2.4 Handling Outliers:**

* Deep dive for very spectacular performances as to why they happened in the case of that individual- player skill, team performance, or opponent.
* Add some outlier-handling-specific methods like weighted loss functions to improve model sensitivity to extremes.

### **8. Conclusion:**

#### **8.1 Summary of Findings:**

The project has successfully created models for predicting NFL running back rushing yards using historical player performances and demographic factors. Some of the models put to the test include the following:

1. **Deep Learning** performed better than **Random Forest** with a lower Root Mean Square Error (RMSE) of 161.481 vs. 213.655.
2. **Key Features:**

* Games Started (GS) and Years of Service - were the most important predictors.
* Age had a negative effect on the expected declining performance effect with age.
* Team metrics such as Head Coach and Offensive Line had net minor potentials in terms of changes that would be in scope for improvement in these features.

That says what potential unsolicited-from-data insights into sports have, whether helping coaching to decide player ranking or Fantasy Football vision. Again, these models may help set betting odds in sportsbooks, as well as improve player evaluation systems.

#### **8.2 Future Directions:**

1. **Enhanced Feature Set:** More physical metrics, for example, height, weight, and injury history.
2. **More Sophisticated Techniques:** Stacking methods or an algorithm such as XGBoost may be applied here for better accuracy.
3. **Outlier Analysis:** Investigate stellar seasons or rookie performances to enhance the predictions of high-value outcomes.

This project has built a good base for further investigation and the establishment of links between statistical modeling and real-world performance analysis in the NFL.