



Predicting NFL Rookie Contract Extensions

By Sarah, Drew, and Josh

Problem Framing

- How can we help predict what players will be extended
- Are there any factors that impact binary outcome?
- How is this research relevant to the current NFL?



Data Extraction and Cleaning Process

- NFL FastR Package
 - Data from contracts function and season statistics
- Finding a suitable date range
 - Narrowing down to only offensive players at skill positions
- Dealing with 49ers players collection process



Predictive Models and Initial Insights

❏ Binary Random Forest + Tuning

Confusion Matrix and Statistics

```
rf_pred_class  0  1
               0 168  68
               1  47  45
```

Accuracy : 0.6494

95% CI : (0.5951, 0.701)

No Information Rate : 0.6555

P-Value [Acc > NIR] : 0.61636

Kappa : 0.1879

Mcnemar's Test P-Value : 0.06218

Sensitivity : 0.3982

Specificity : 0.7814

Pos Pred Value : 0.4891

Neg Pred Value : 0.7119

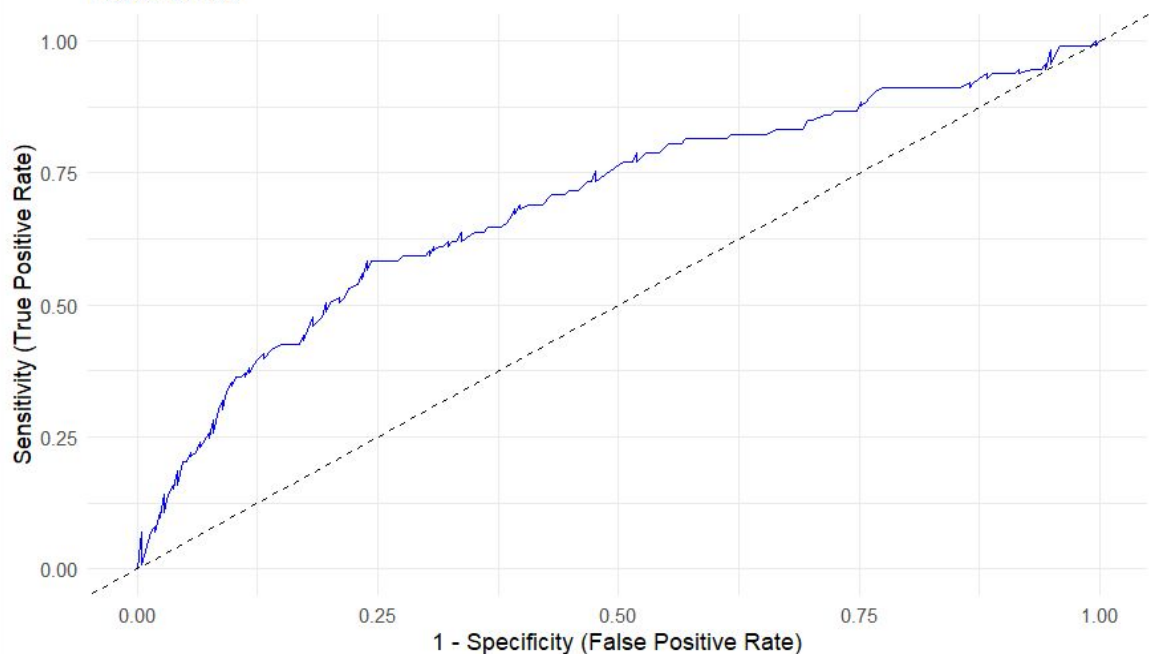
Prevalence : 0.3445

Detection Rate : 0.1372

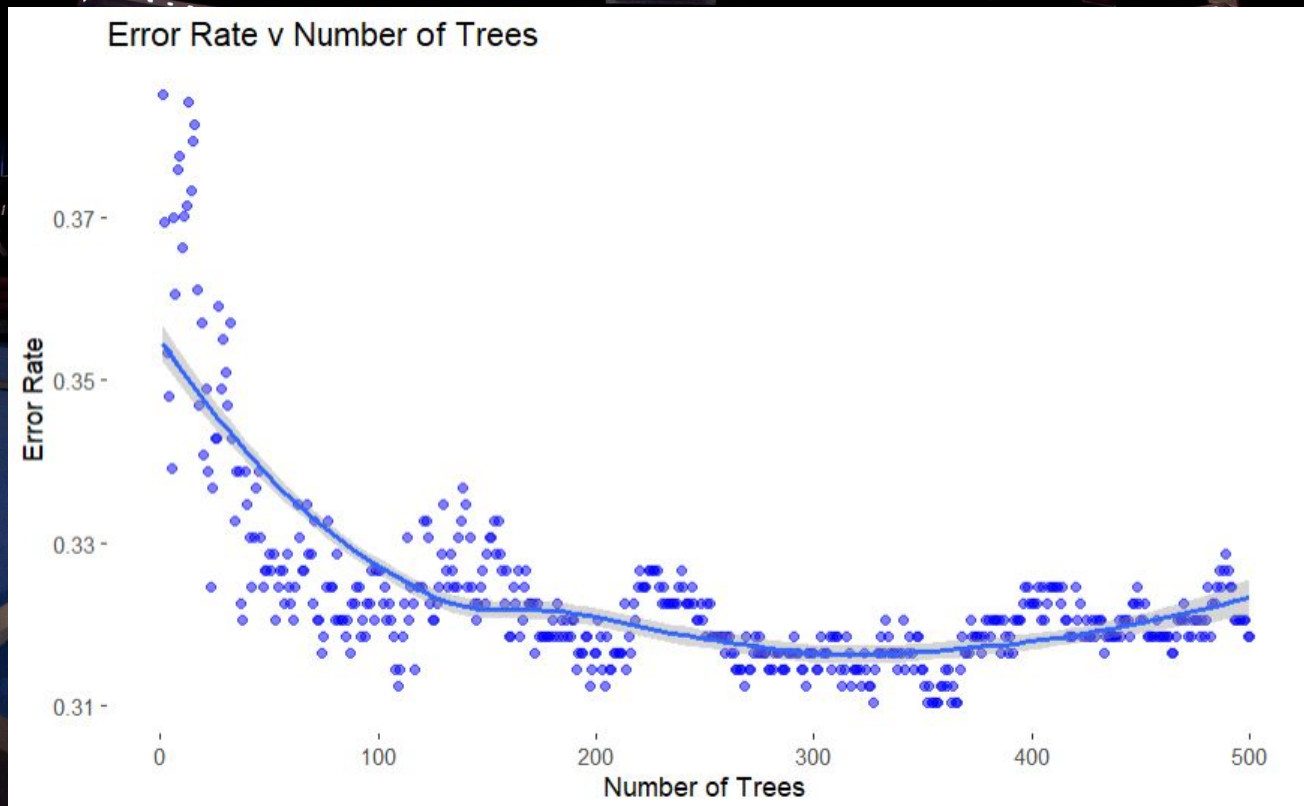
Detection Prevalence : 0.2805

Balanced Accuracy : 0.5898

ROC Curve



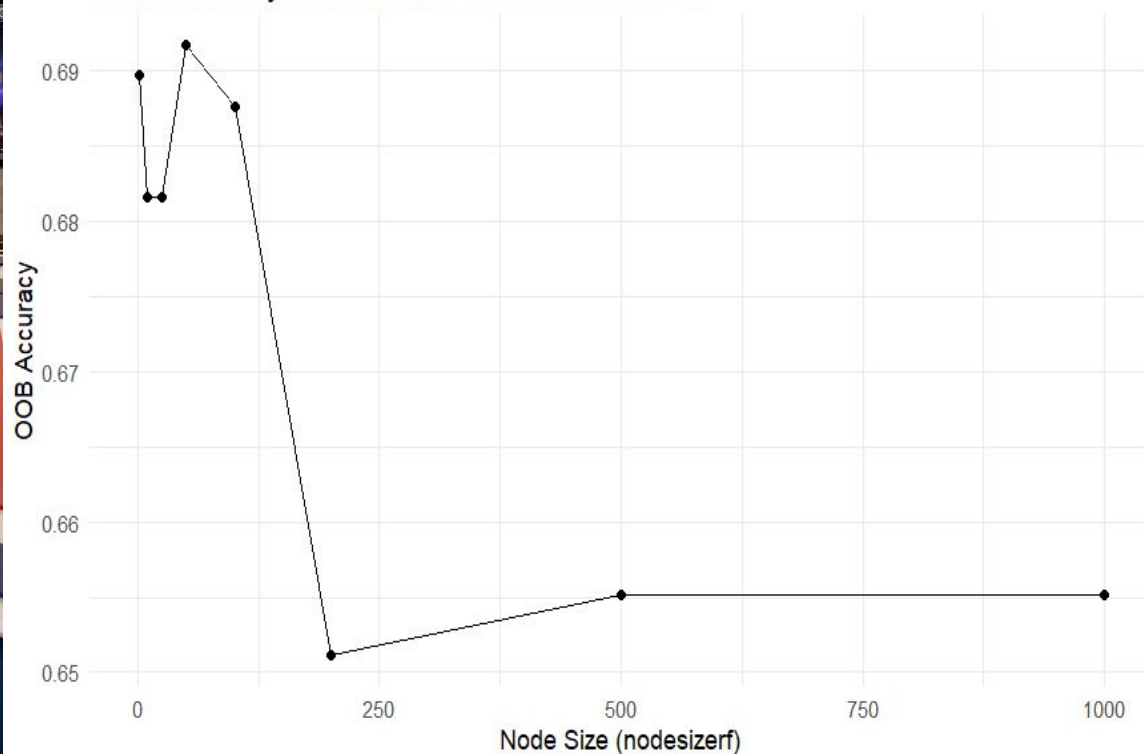
Predictive Models cont.



Random Forest Model #2 (Focus on Node Size)

nodesizerf <dbl>	oob_accuracy <dbl>
1	0.6896552
10	0.6815416
25	0.6815416
50	0.6916836
100	0.6876268
200	0.6511156
500	0.6551724
1000	0.6551724

OOB Accuracy vs Node Size in Random Forest



Random Forest #3

Confusion Matrix and Statistics

```
rf_pred_class  0   1
              0 172  69
              1  43  44
```

Accuracy : 0.6585

95% CI : (0.6044, 0.7097)

No Information Rate : 0.6555

P-Value [Acc > NIR] : 0.47924

Kappa : 0.2003

McNemar's Test P-Value : 0.01816

Sensitivity : 0.3894

Specificity : 0.8000

Pos Pred Value : 0.5057

Neg Pred Value : 0.7137

Prevalence : 0.3445

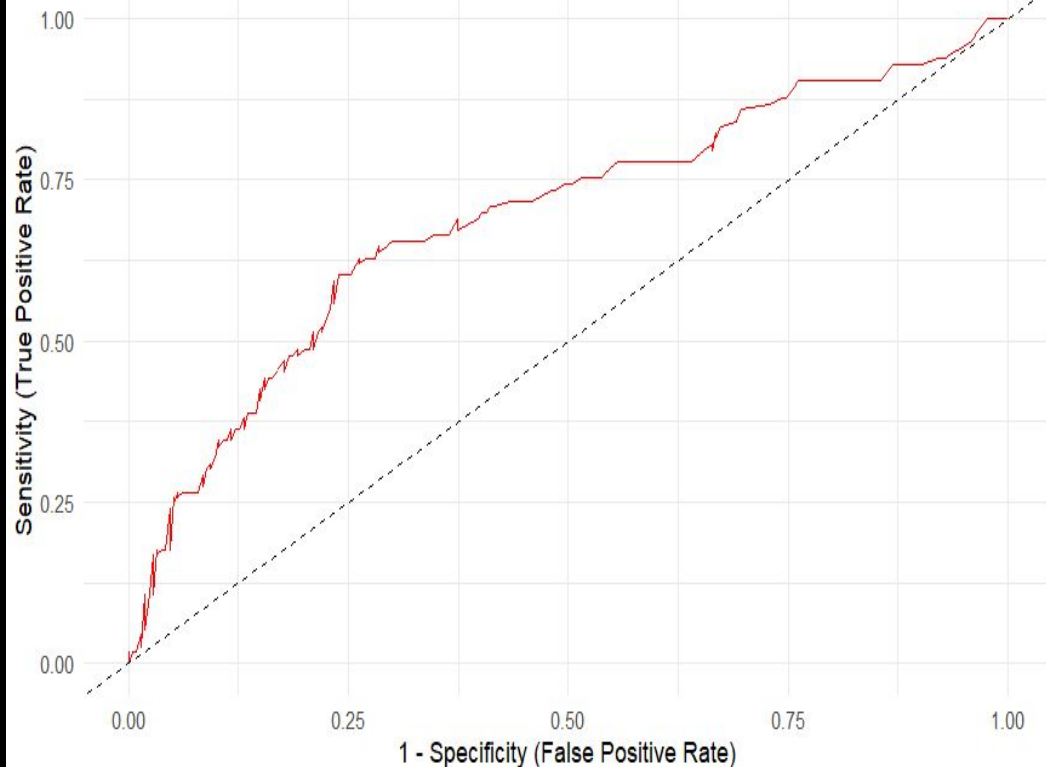
Detection Rate : 0.1341

Detection Prevalence : 0.2652

Balanced Accuracy : 0.5947

'Positive' Class : 1

ROC Curve - Random Forest Predictions Model 2



Random Forest Final Model

Confusion Matrix and Statistics

	Reference	
Prediction	0	1
0	61	27
1	6	14

Accuracy : 0.6944

95% CI : (0.5984, 0.7795)

No Information Rate : 0.6204

P-Value [Acc > NIR] : 0.0669761

Kappa : 0.2797

McNemar's Test P-Value : 0.0004985

Sensitivity : 0.9104

Specificity : 0.3415

Pos Pred Value : 0.6932

Neg Pred Value : 0.7000

Prevalence : 0.6204

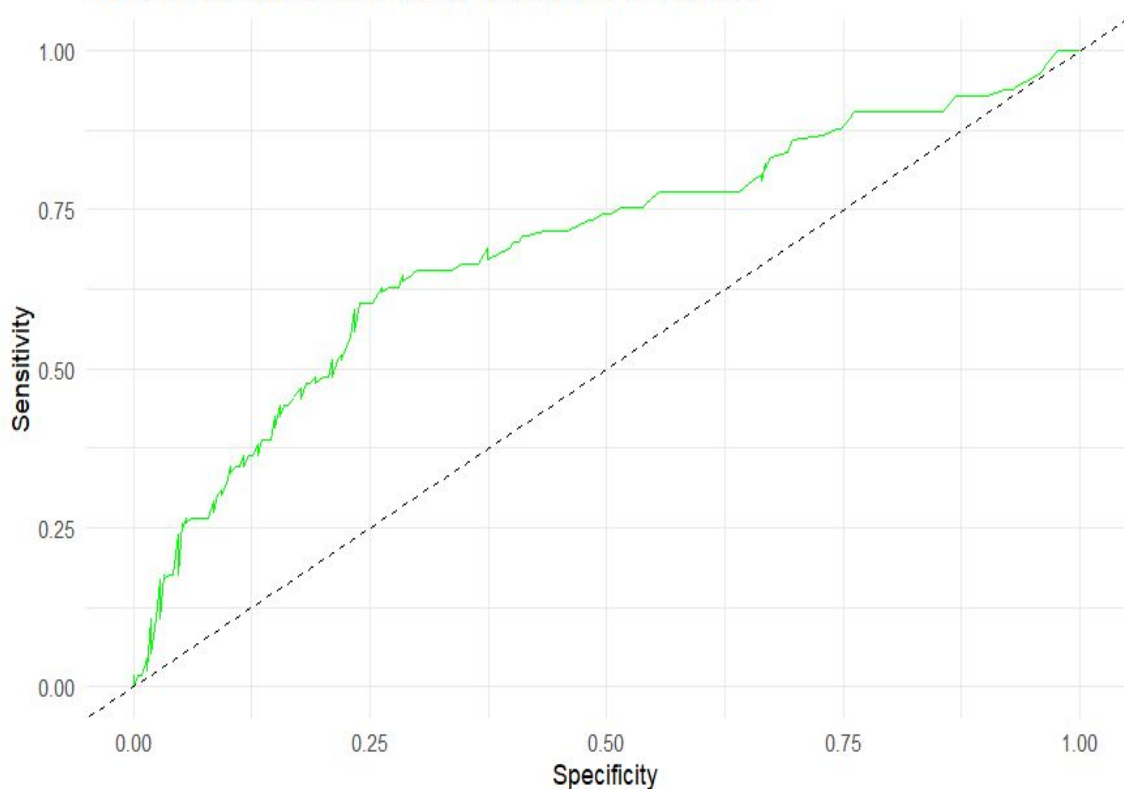
Detection Rate : 0.5648

Detection Prevalence : 0.8148

Balanced Accuracy : 0.6260

'Positive' Class : 0

ROC Curve - Random Forest Predictions Final Model



Rookie Data Models by Position

Quarterbacks

- Carries, rushing fumbles, rushing first downs, rushing EPA, rushing 2 point conversions

Running Backs

- Rushing yards, rushing fumbles, receiving air yards, receiving EPA

Wide Receivers

- Receiving air yards, receiving first downs, receiving EPA

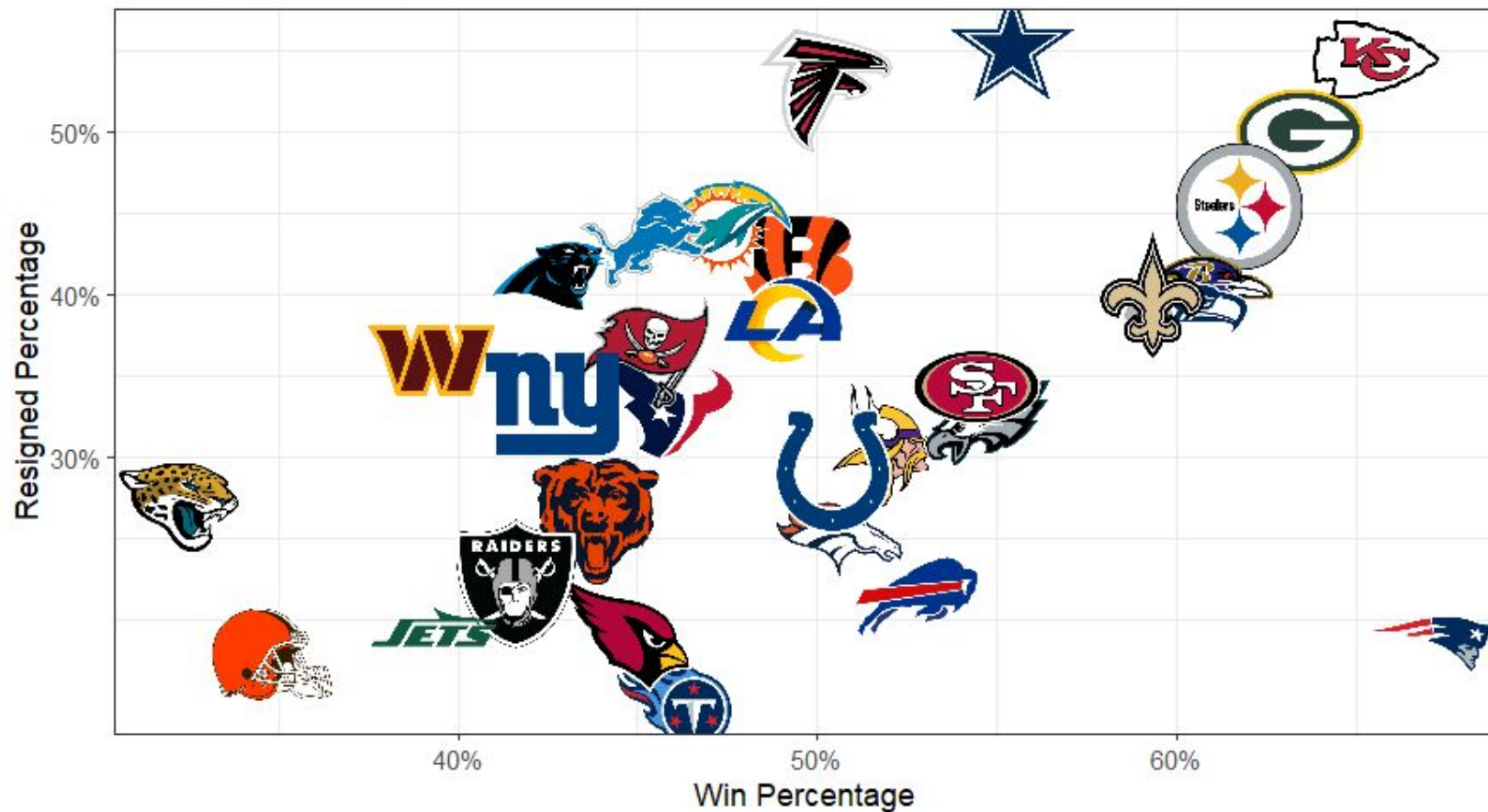
Tight Ends

- Receiving yards, receiving TDs , Receiving air yards



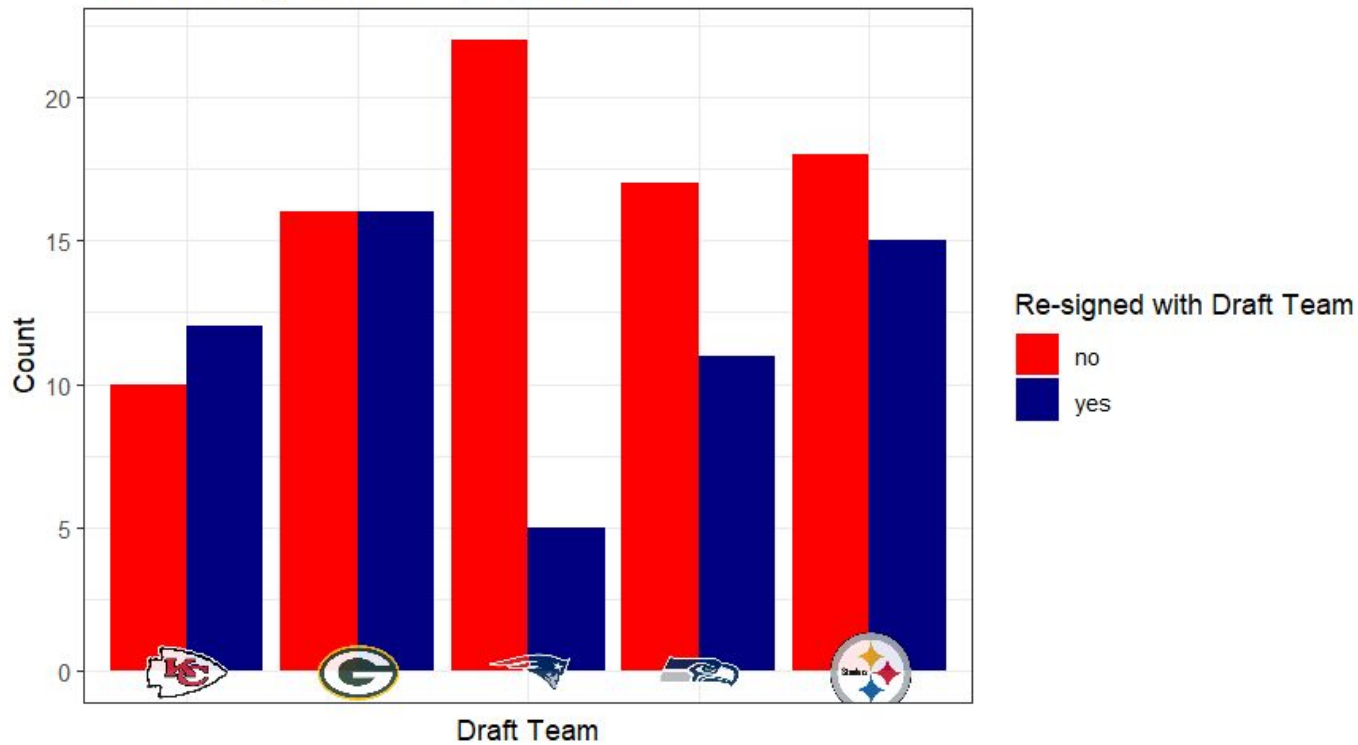
Correlation Between Win % and Resign %

Data Since 2010 Years



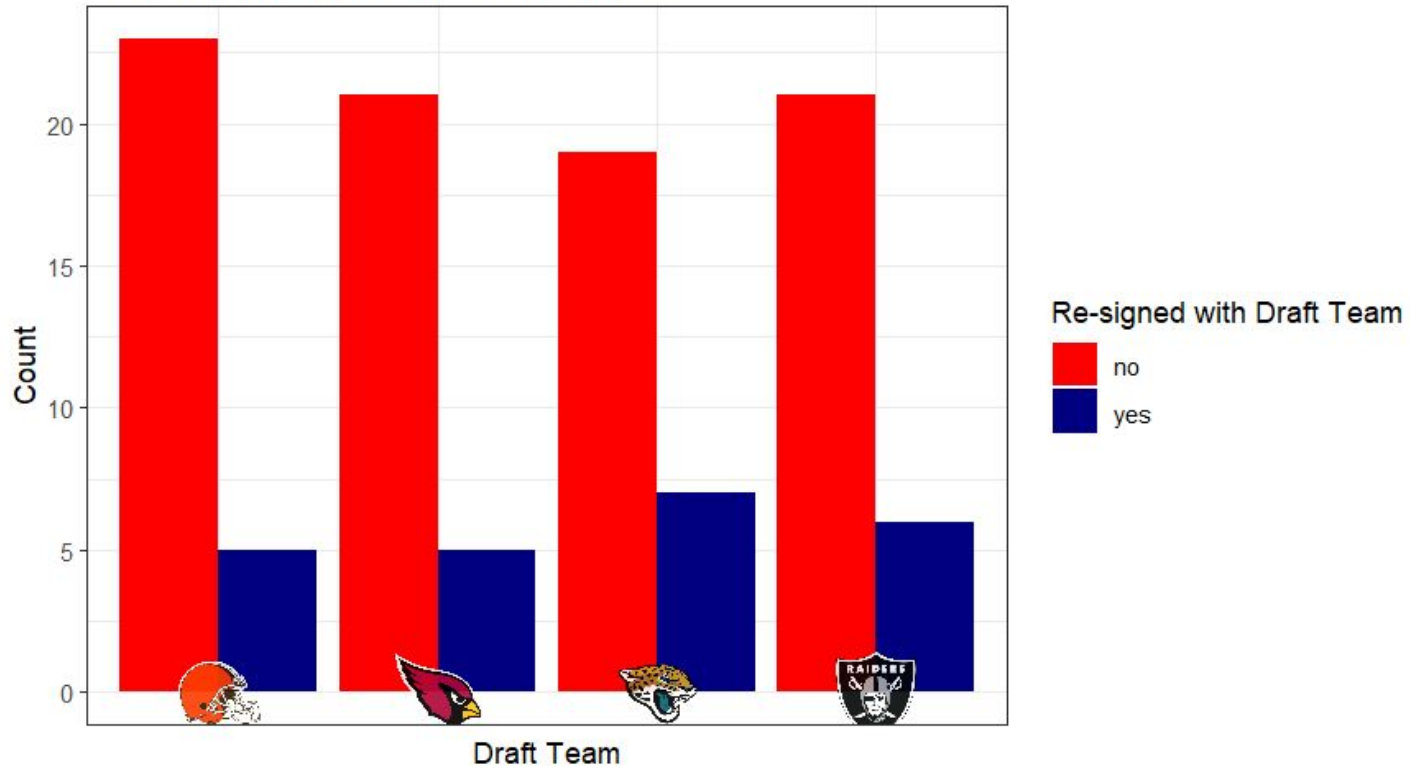
Re-signed with Draft Team by Team

Teams with Highest Win % in Last 10 Years

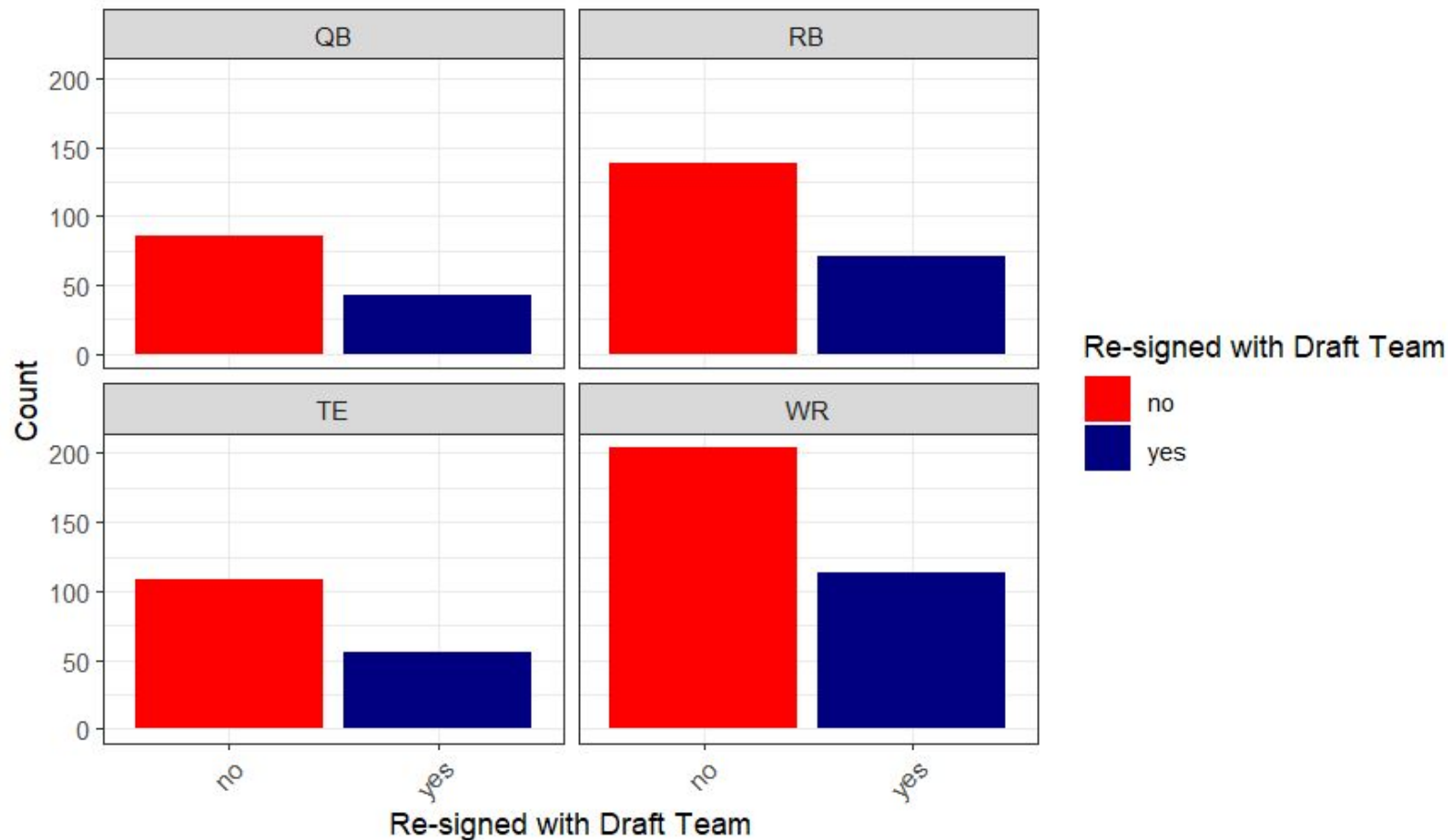


Re-signed with Draft Team by Team

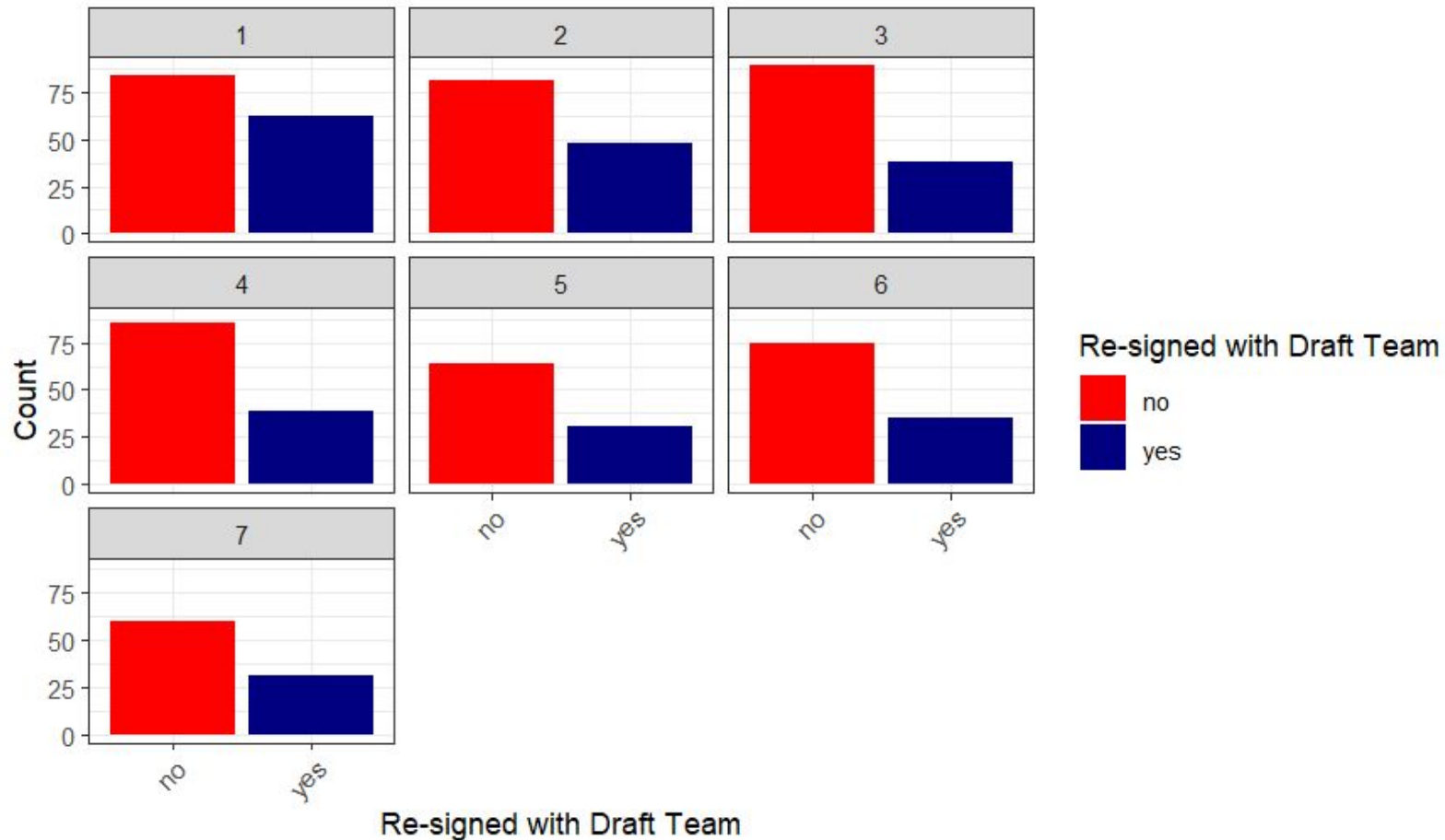
Teams with Lowest Win % in Last 10 Years



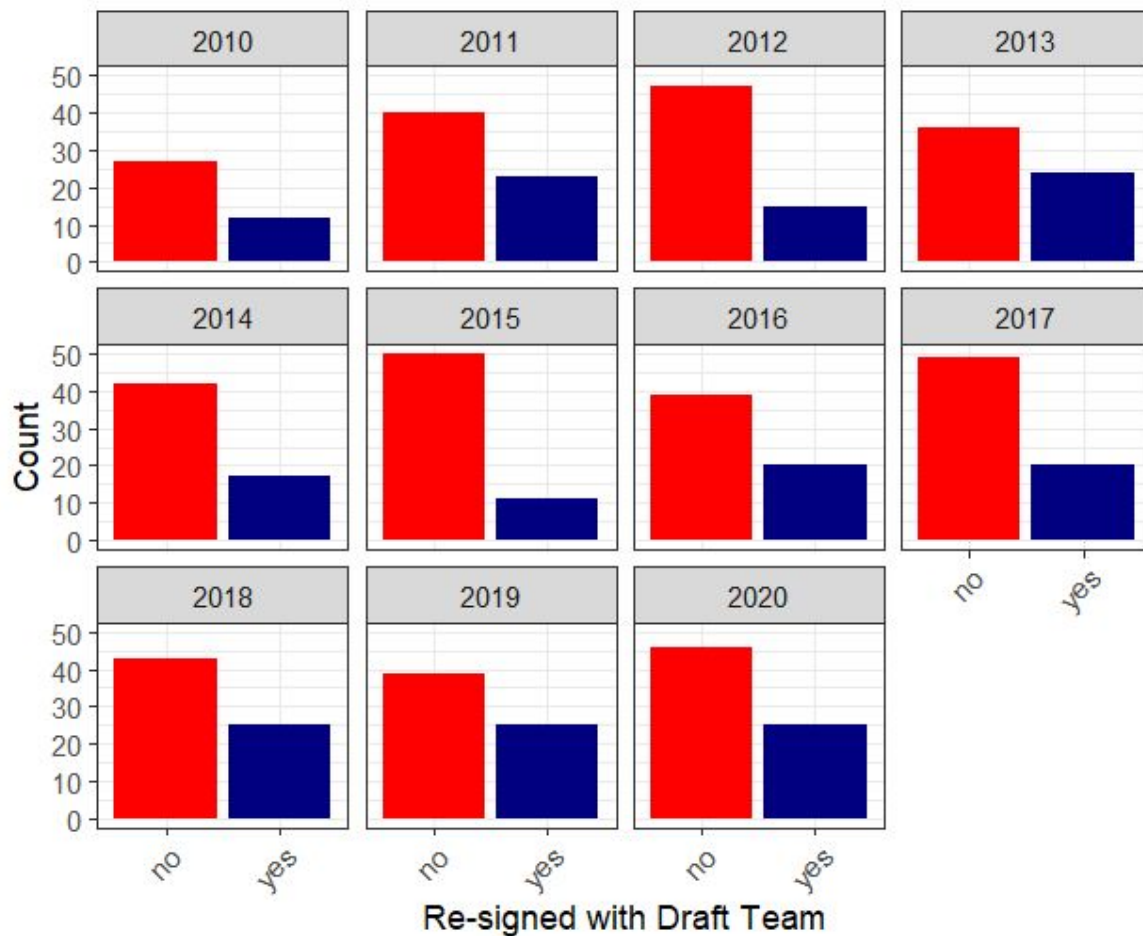
Positional Breakdown of Draft Picks Resigning



Draft Round Breakdown of Draft Picks Resigning



Year by Year Breakdown of Draft Picks Resigning



Re-signed with Draft Team



Final Thoughts and End Summary

Outside factors impact contract extensions

- Cap space allocation
- Draft picks
- Internal roster changes

What is the scope of what we've analyzed

- Team Specific?
- Player Specific?
- Position Specific?