



NAVAL
POSTGRADUATE
SCHOOL

Analyzing NFL Spending by Position

Ace Padilla | Andrew Tamez
Matthew Thoms | Richard Viehdorfer
OA3802 – Fall 2024



- American Football Overview
- Web Scraping
- Statistical Analysis
- Simulation Model
- Results
- Web Dashboard
- Conclusion



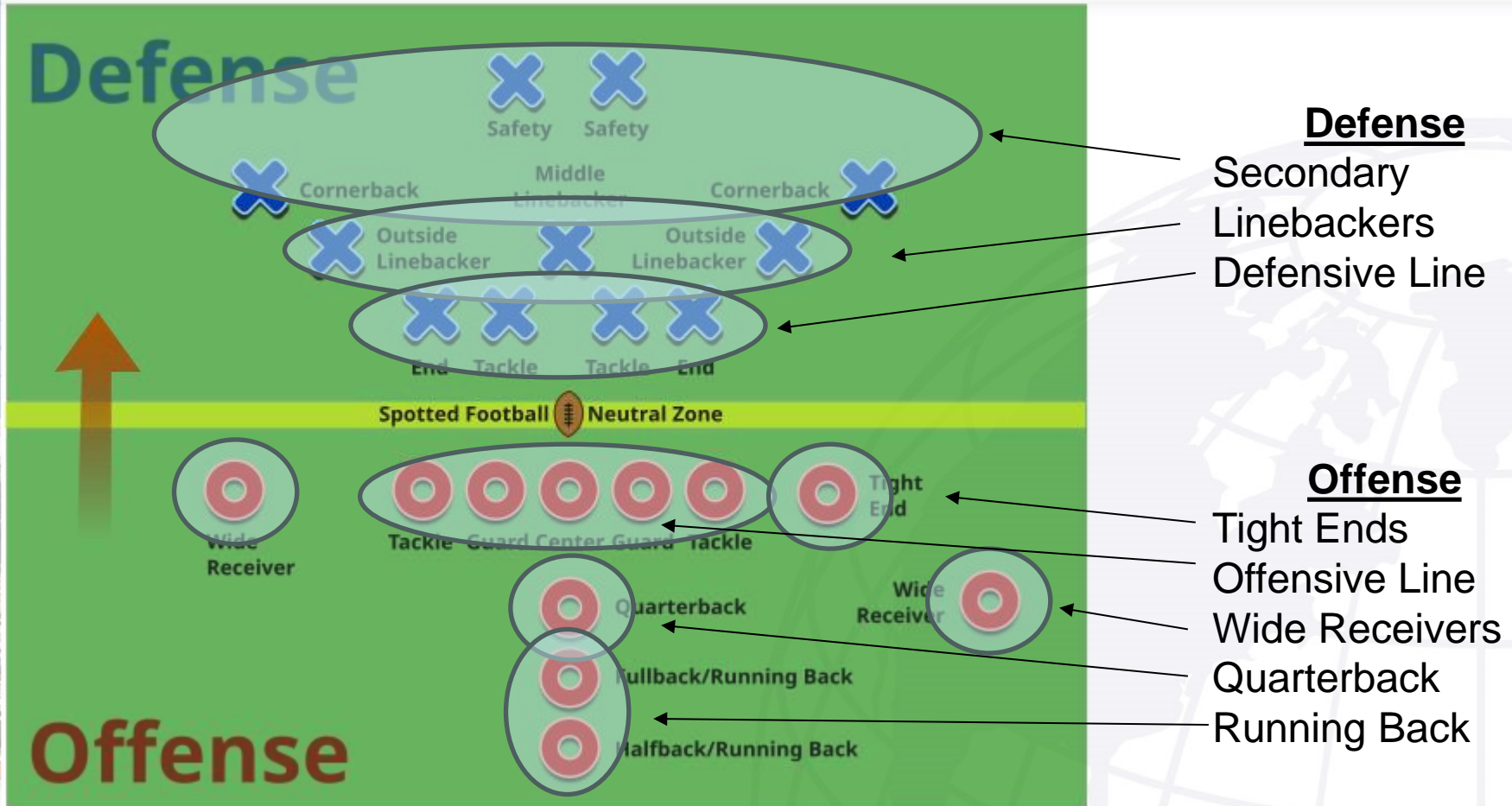
What is American Football?

- 2 Teams
- 11 Players on the field at each time
- Attempt to advance the ball down the field by passing it or running it
- If you don't advance 10 yards in 4 plays, turn the ball over to the other team
- 6 points for a touchdown (PAT, 2 Point Conversion)
- 3 points for a field goal
- 2 points for a safety





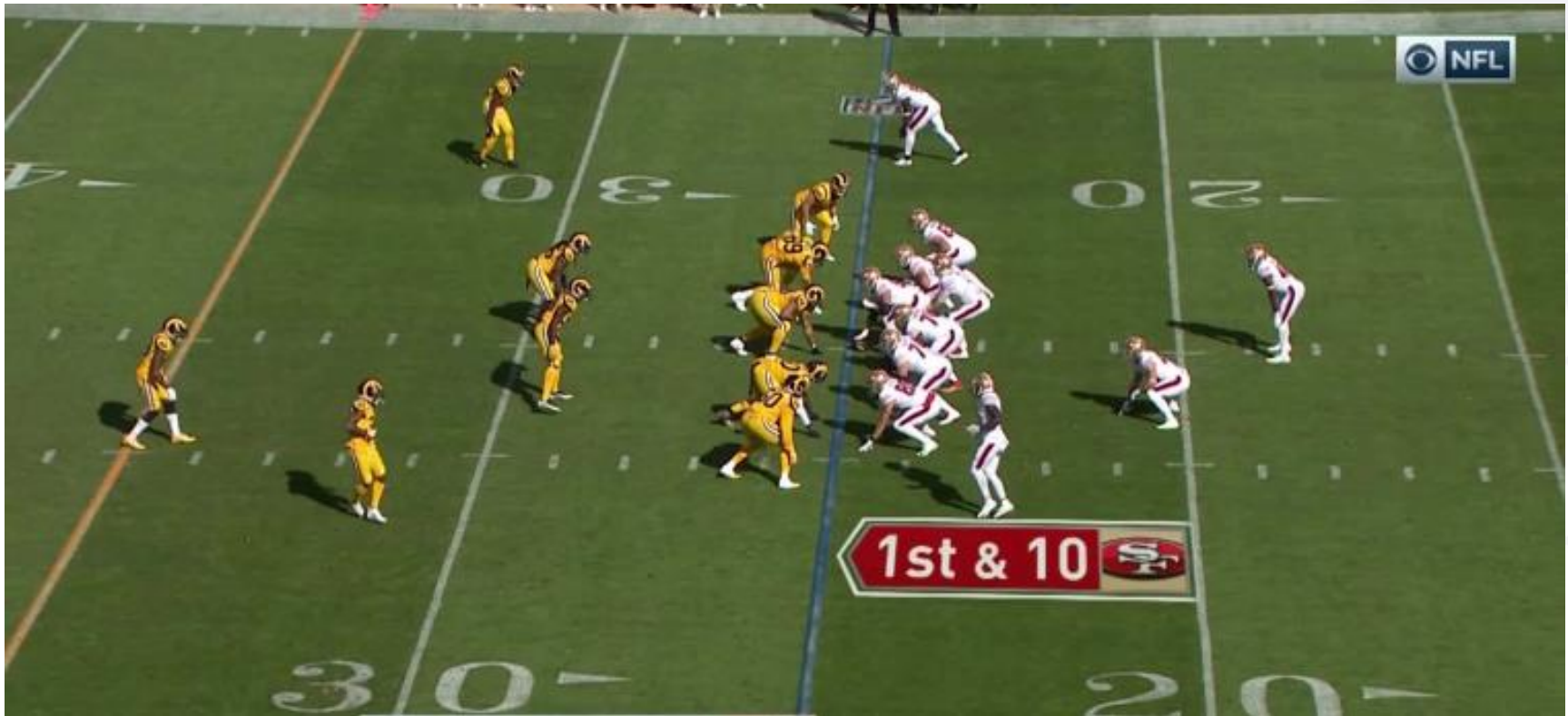
American Football Positions





NAVAL
POSTGRADUATE
SCHOOL

American Football Positions





American Football Positions



Punter

Long Snapper



Kicker



- Reflects the total amount a team can spend on its players per year
- 48% of total NFL revenue, divided by 32 teams
- Increasing salary cap results in increased player salary
- General Manager: role is to build and manage a team's roster and oversee the football operations.
 - Player signings
 - Contract Negotiations
- Working within the salary cap, how much do you pay which players to get the best chance of success?

YEAR	CAP MAXIMUM	CAP \$ +/-	CAP % +/-
2024	\$255,400,000	\$30,600,000	13.61% ↑
2023	\$224,800,000	\$16,600,000	7.97% ↑
2022	\$208,200,000	\$25,700,000	14.08% ↑
2021	\$182,500,000	\$-15,700,000	-7.92% ↓
2020	\$198,200,000	\$10,000,000	5.31% ↑
2019	\$188,200,000	\$11,000,000	6.21% ↑
2018	\$177,200,000	\$10,200,000	6.11% ↑
2017	\$167,000,000	\$11,730,000	7.55% ↑
2016	\$155,270,000	\$11,990,000	8.37% ↑
2015	\$143,280,000	\$10,280,000	7.73% ↑
2014	\$133,000,000	\$9,400,000	7.61% ↑
2013	\$123,600,000	\$3,000,000	2.49% ↑
2012	\$120,600,000	\$225,000	0.19% ↑
2011	\$120,375,000		
2010	Uncapped		
2009	\$123,000,000	\$7,000,000	6.03% ↑
2008	\$116,000,000	\$7,000,000	6.42% ↑
2007	\$109,000,000	\$7,000,000	6.86% ↑
2006	\$102,000,000	\$16,500,000	19.30% ↑
2005	\$85,500,000	\$4,918,000	6.10% ↑
2004	\$80,582,000	\$5,575,000	7.43% ↑



NFL Quarterback Salaries

Player	Team	Average Salary per Year	Super Bowl Wins	Playoff Wins
<u>Dak Prescott</u>	<u>Cowboys</u>	\$60,000,000		
<u>Joe Burrow</u>	<u>Bengals</u>	\$55,000,000		5
<u>Jordan Love</u>	<u>Packers</u>	\$55,000,000		1
<u>Trevor Lawrence</u>	<u>Jaguars</u>	\$55,000,000		1
<u>Tua Tagovailoa</u>	<u>Dolphins</u>	\$53,100,000		
<u>Jared Goff</u>	<u>Lions</u>	\$53,000,000		5
<u>Justin Herbert</u>	<u>Chargers</u>	\$52,500,000		
<u>Lamar Jackson</u>	<u>Ravens</u>	\$52,000,000		2
<u>Jalen Hurts</u>	<u>Eagles</u>	\$51,000,000		2
<u>Kyler Murray</u>	<u>Cardinals</u>	\$46,100,000		
<u>Deshaun Watson</u>	<u>Browns</u>	\$46,000,000		1
<u>Kirk Cousins</u>	<u>Falcons</u>	\$45,000,000		
<u>Patrick Mahomes</u>	<u>Chiefs</u>	\$45,000,000	3	15
<u>Josh Allen</u>	<u>Bills</u>	\$43,000,000		5
<u>Matt Stafford</u>	<u>Rams</u>	\$40,000,000	1	4
<u>Daniel Jones</u>	<u>Giants</u>	\$40,000,000		1
<u>Aaron Rodgers</u>	<u>Jets</u>	\$37,500,000	1	12
<u>Derek Carr</u>	<u>Saints</u>	\$37,500,000		
<u>Baker Mayfield</u>	<u>Buccaneers</u>	\$33,333,333		2
<u>Geno Smith</u>	<u>Seahawks</u>	\$25,000,000		
<u>Gardner Minshew</u>	<u>Raiders</u>	\$12,500,000		2

Source: overthecap.com/position/quarterback

- Key tools: requests, bs4 python packages



- Sources (no API with all the data needed):
 - Team Spending: https://www.spotracer.com/nfl/position/_/year/2024
 - Superbowl Winners (1966-Current):
<https://www.topendsports.com/events/super-bowl/winners-list.htm>
 - Conference Champions: <https://www.foxsports.com/stories/nfl/nfc-champions-complete-list-winners-year>
 - Playoff Teams: [https://en.wikipedia.org/wiki/NFL_playoff_results#All-time_playoff_records_\(NFL/AFL\)](https://en.wikipedia.org/wiki/NFL_playoff_results#All-time_playoff_records_(NFL/AFL))

```
url = 'https://www.topendsports.com/events/super-bowl/winners-list.htm'  
  
response = requests.get(url)  
soup = BeautifulSoup(response.content, 'html.parser')
```



Team	OFF	QB	RB	WR	TE
DET	0.52525	0.075417	0.060167	0.205417	0.047
PHI	0.46475	0.157667	0.015083	0.102833	0.0295
CAR	0.535833	0.045583	0.091083	0.112	0.05375
ATL	0.543833	0.135833	0.07775	0.112333	0.0805
NYG	0.478167	0.15075	0.099333	0.062833	0.012667

Total	Year	Wins	SuperBow	CC_Win	Playoffs
1.02225	2011	10	0	0	1
0.99675	2011	8	0	0	0
0.989333	2011	6	0	0	0
0.972333	2011	10	0	0	1
0.959667	2011	9	1	1	1

- **Data**

- Independent Variables: Positions
- Dependent Variables: Wins
 - Super Bowl, Conference Championships, Playoffs

- **Data Balance (0,1)**

- SB: 410 vs. 13
- CC: 397 vs. 26
- Playoffs: 268 vs. 155

- **R: Smote Package:** [SMOTE function - RDocumentation](#)

- “Artificially generate new examples of the minority class using the nearest neighbors of these cases.”
- “Majority class examples are also under-sampled, leading to a more balanced dataset.”
- Increases bias aligned with winning team particular features.

- **Emsemble Model per Independent Variable –**

- Balanced and Unbalanced – avg of (6) total
 - Ridge Regression
 - Logarithmic General Linear Model
 - Random Forest Classifier

ENSEMBLE	
Position	Overall_Weighted_Average
QB	0.133950041
TE	0.129747766
DL	0.121430187
OL	0.116630561
P	0.115268535
LB	0.08733053
SEC	0.072290829
WR	0.068452959
LS	0.063114955
K	0.057982732
RB	0.033800904

- Attempted HPC.
- Randomize Spending Percentages
- Utilize our calculated weights

$$\begin{aligned}\pi(\mathbf{X}) &= \frac{\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}{1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)} \\ &= \frac{\exp(\mathbf{X}\boldsymbol{\beta})}{1 + \exp(\mathbf{X}\boldsymbol{\beta})} \\ &= \frac{1}{1 + \exp(-\mathbf{X}\boldsymbol{\beta})},\end{aligned}$$

- Determine the probability of win using binary logistic regression
- Look at the top 10% of probabilities
- What do we get?
- Image courtesy of <https://online.stat.psu.edu/stat462/node/207/>



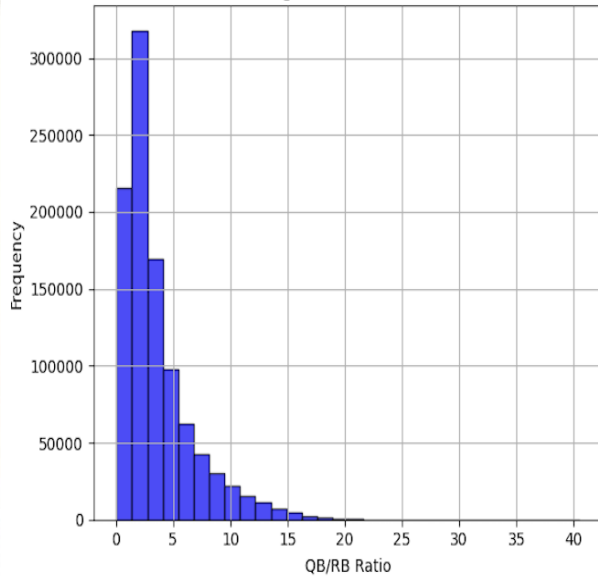
- Before we get to our results
- Over 3 million data points. Best way to determine the values we want: pandas dataframe, excel, SQL?
- With so many data points, let's go with SQL
 - Allows storage of more data within the database in a memory-efficient manner
 - Easy and efficient to query for the results we need
- Looking for the mean/max is easy:

```
query = """
SELECT
    AVG(QB) AS avg_QB,
    AVG(TE) AS avg_TE,
    AVG(DL) AS avg_DL,
    AVG(P) AS avg_P,
    AVG(OL) AS avg_OL,
    AVG(LB) AS avg_LB,
    AVG(WR) AS avg_WR,
    AVG(SEC) AS avg_SEC,
    AVG(K) AS avg_K,
    AVG(LS) AS avg_LS,
    AVG(RB) AS avg_RB,
    AVG(predicted_probability) AS avg_predicted_probability
FROM SuperBowl
WHERE predicted_probability >= 0.56221 AND predicted_probability <= 0.59018
"""
```

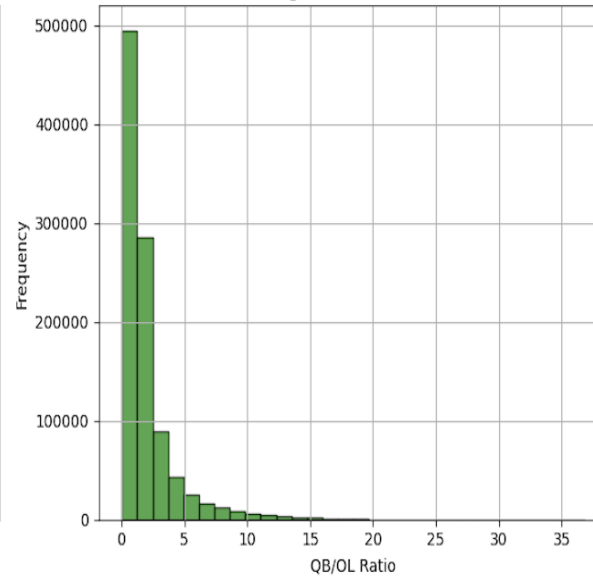


Results: SuperBowl

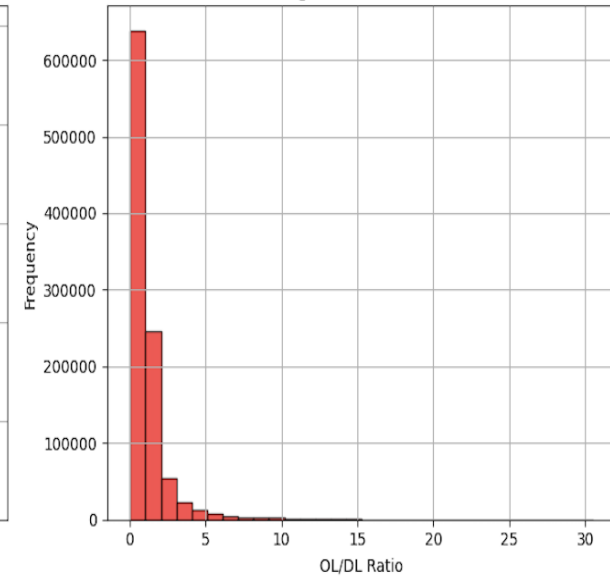
Histogram of QB/RB Ratio



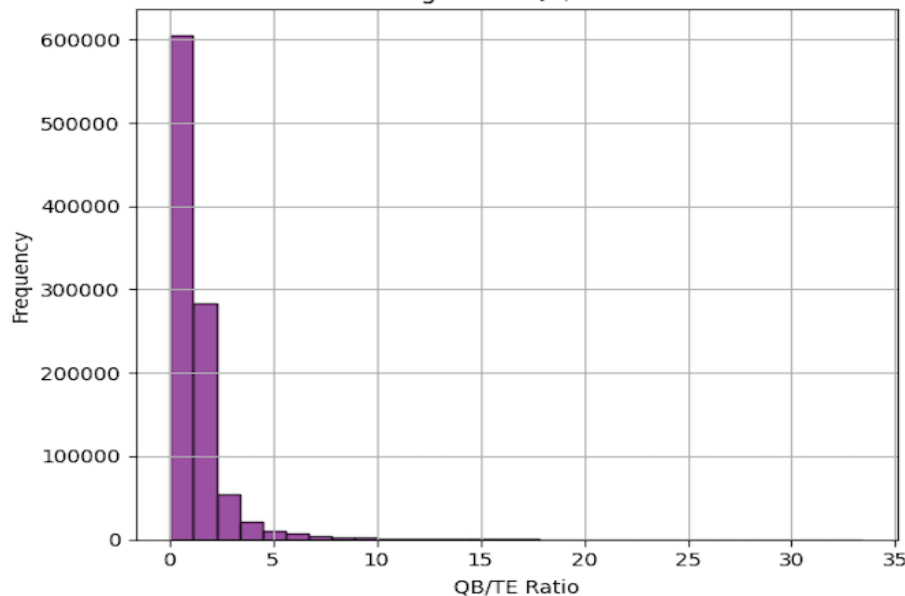
Histogram of QB/OL Ratio



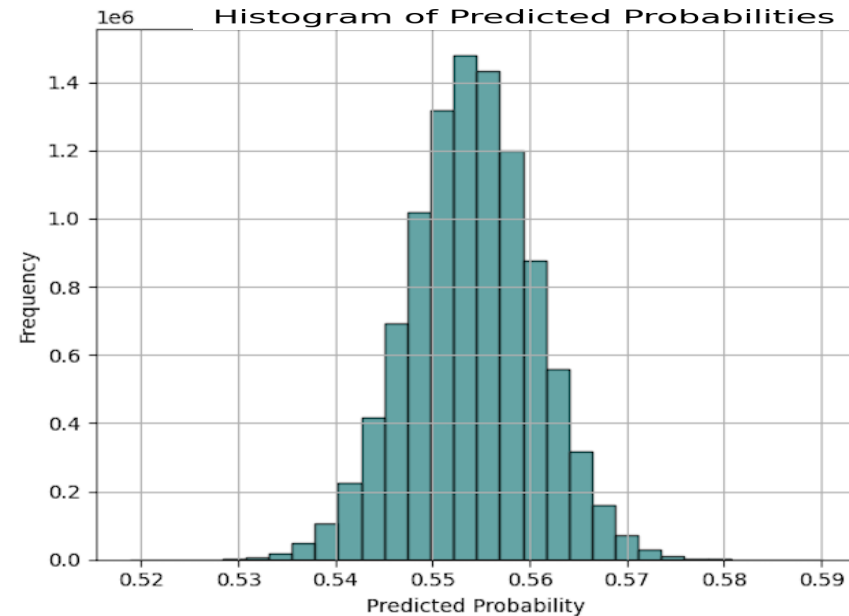
Histogram of OL/DL Ratio



Histogram of QB/TE Ratio



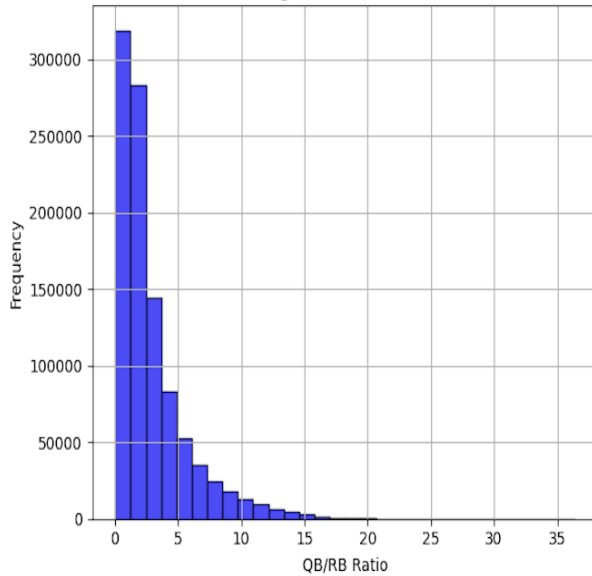
Histogram of Predicted Probabilities



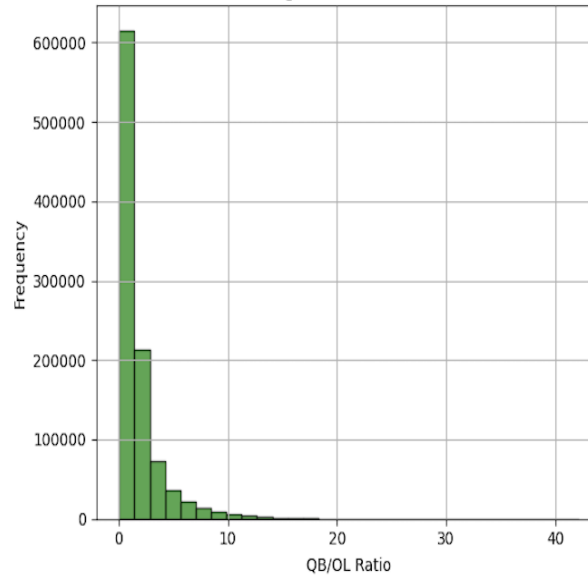


Results: Conference Champ. Wins

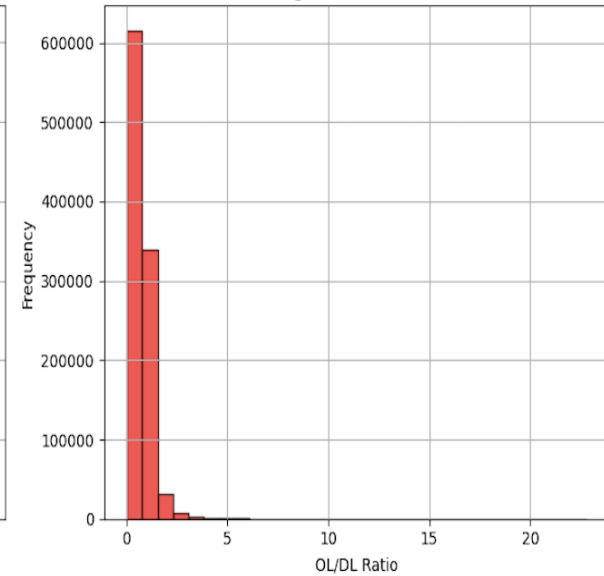
Histogram of QB/RB Ratio



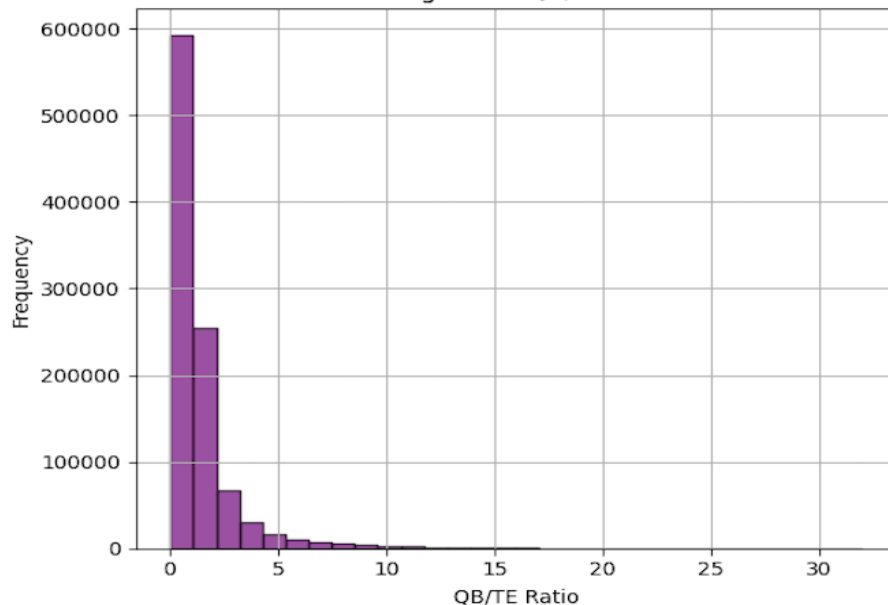
Histogram of QB/OL Ratio



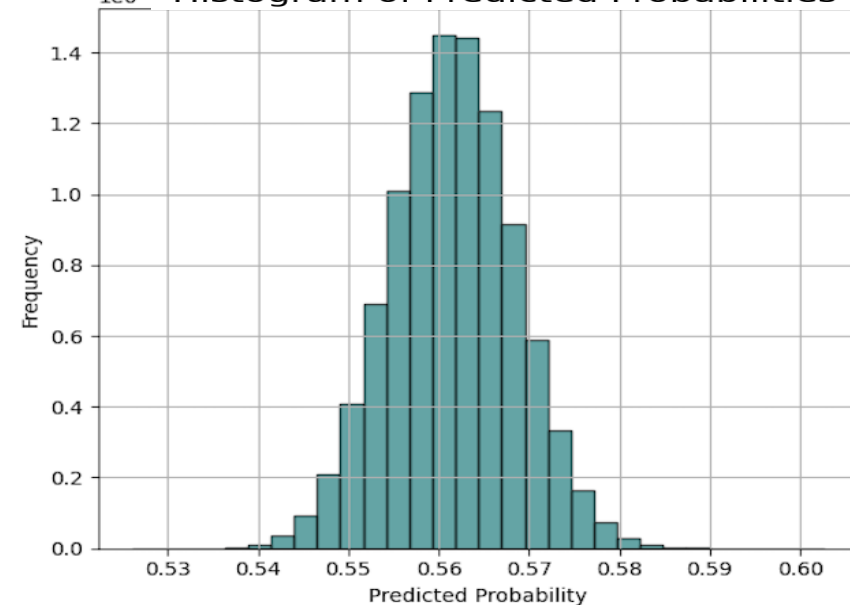
Histogram of OL/DL Ratio



Histogram of QB/TE Ratio



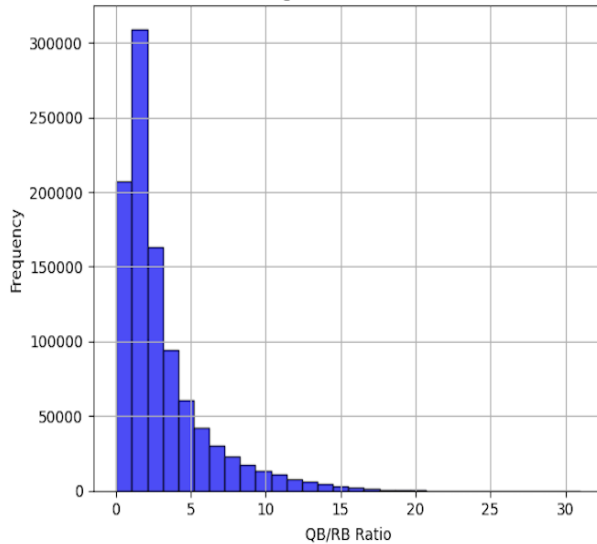
Histogram of Predicted Probabilities



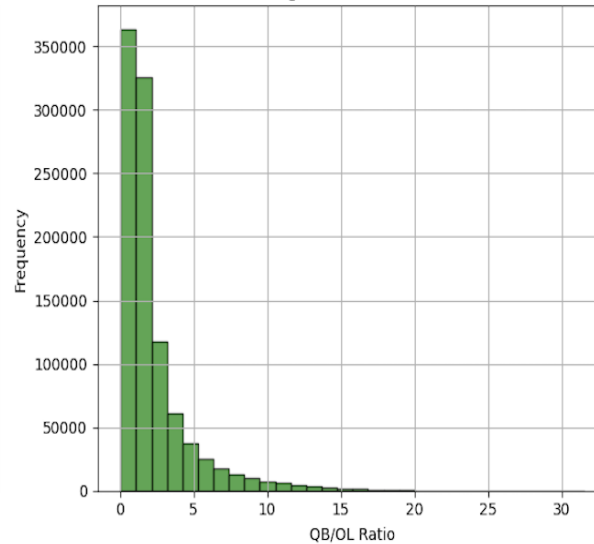


Results: Make the Playoffs

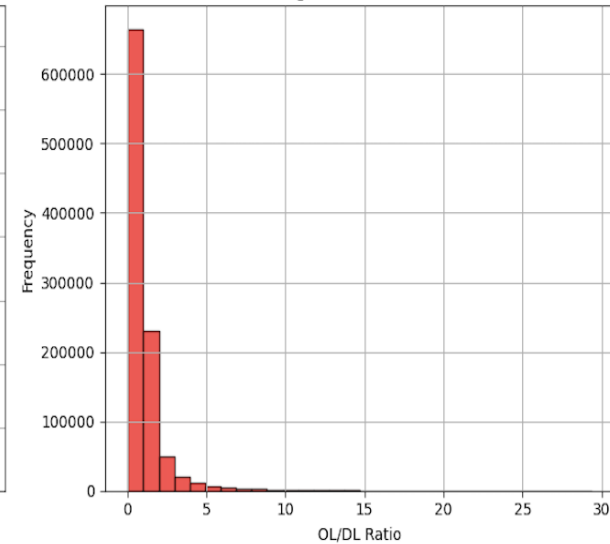
Histogram of QB/RB Ratio



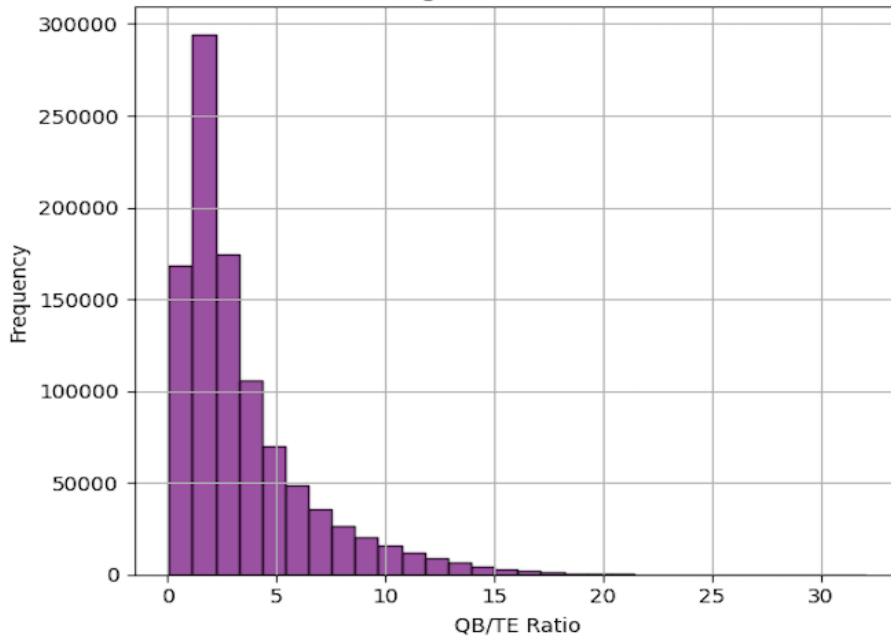
Histogram of QB/OL Ratio



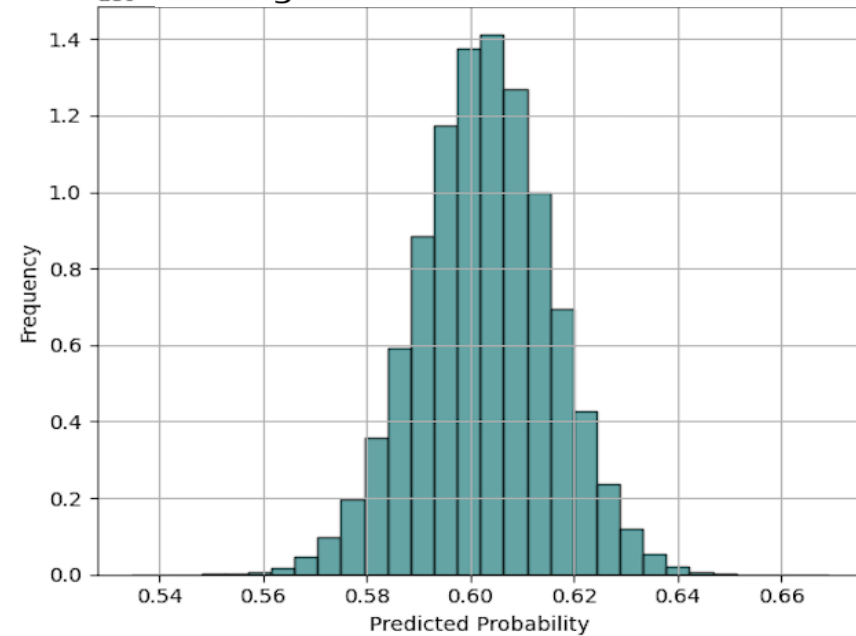
Histogram of OL/DL Ratio



Histogram of QB/TE Ratio



1e6 Histogram of Predicted Probabilities





NAVAL
POSTGRADUATE
SCHOOL

Dashboard

- Recommendations: Pay QB twice as much as RB, OL = DL spending, QB = TE spending
- Investing most of your money on a specific position doesn't result to wins; you need to balance the ratio between positions.
- Future research:
 - Salary versus specific players stats that contribute toward wins
 - Analyze coach spending





NAVAL
POSTGRADUATE
SCHOOL

Questions?