



Projecting NFL Receiving Stats for Upcoming Season



★ DSI 523 Capstone ★

Rick Powell



★ TABLE OF ★ CONTENTS



- 01**
Background
- 02**
Data Analysis
- 03**
Modeling
- 04**
Conclusions

Background and Problem Statement

Projecting players statistics can be used in many different situations:

- Predicting contracts for a free agent player
- Determining which players to draft in fantasy football

This project aims to create models to predict passing yards, receptions and touchdowns for all pass-catching positions (wide receiver, tight end, and running back).



https://static.www.nfl.com/image/private/t_editorial_landscape_12_desktop/league/lhjrc0r4dfzwfb1xuar (Minnesota Vikings)

Data Acquisition and Cleaning

Data pulled from Pro-Football Reference: (<https://www.pro-football-reference.com/>)

Datasets include:

- All players who caught a pass between 2016 and 2021
- All players drafted between 2001 and 2022.
- All players who made their league debut between 2001 and 2021



Cleaning Data included:

- Filling in data for Rookies
- Filling in data for injured players
- Adding columns for All-Pro and Pro-Bowl teams

Distribution of Age

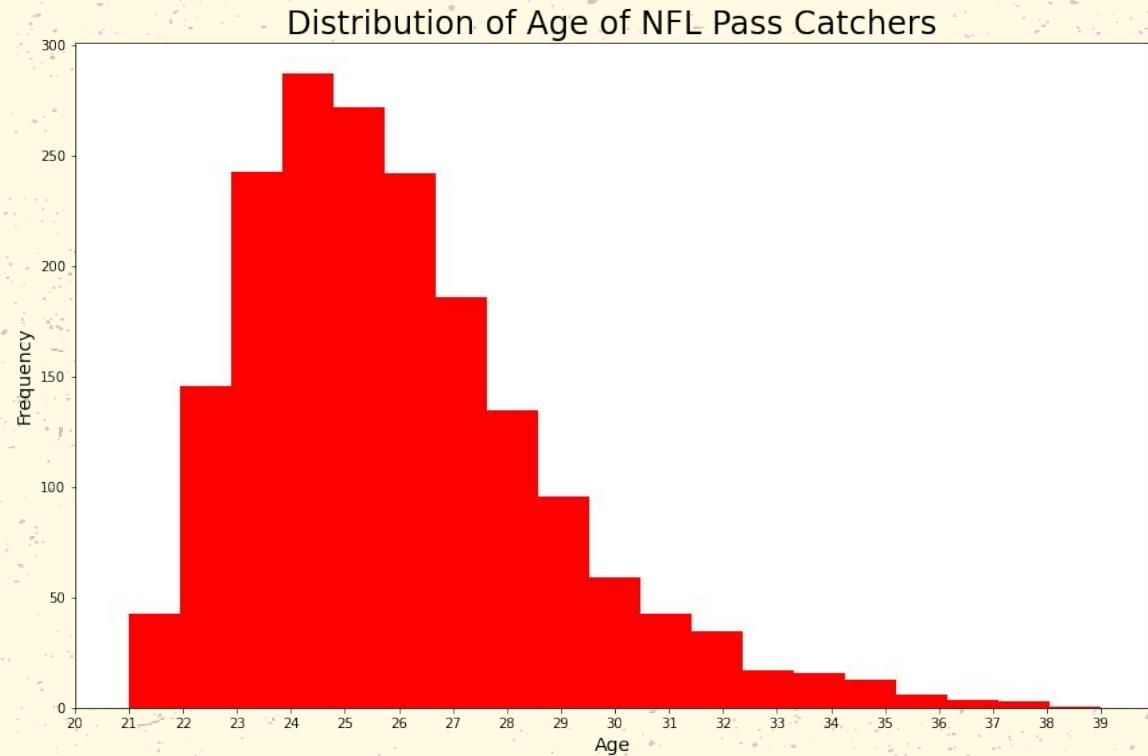
Average Player Ages:

Wide Receiver: 25.69 years

Tight End: 26.45 years

Running Back: 25.26 years

All Receiving Positions:
25.76 years



Yardage by Age

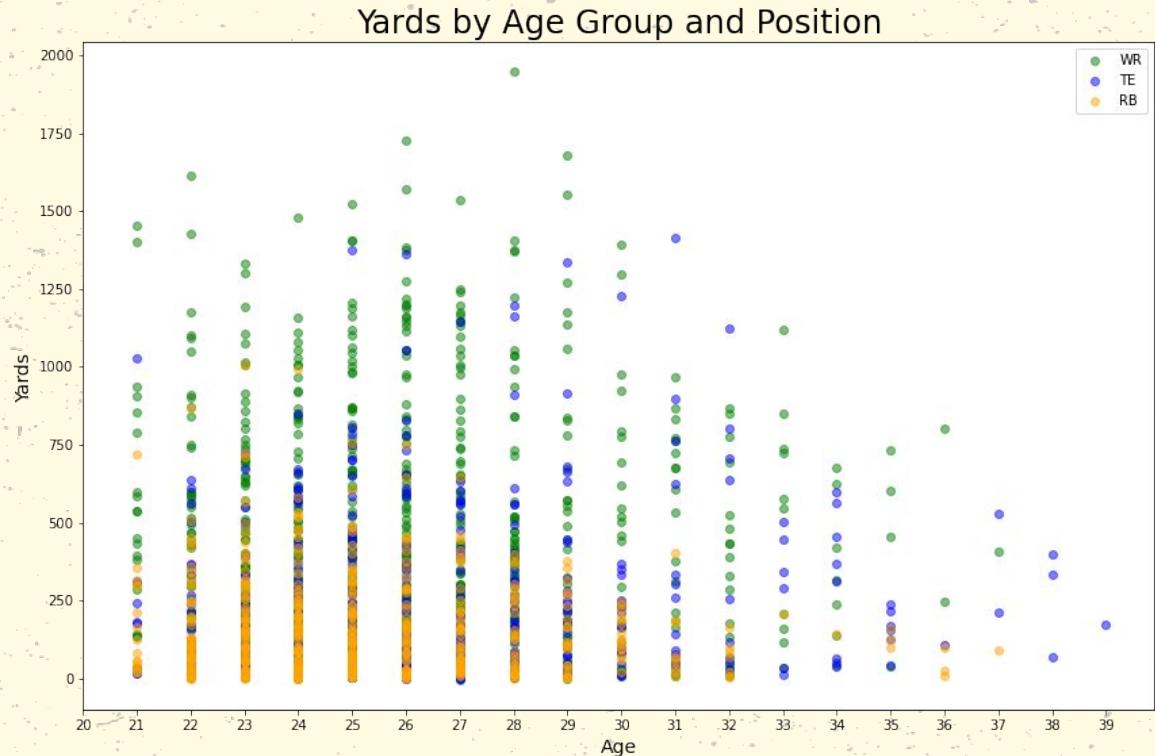
Oldest Player:

Ben Watson, TE, New England, Age 39

Most Yards:

Cooper Kupp, WR, Los Angeles Rams, 1947 yards

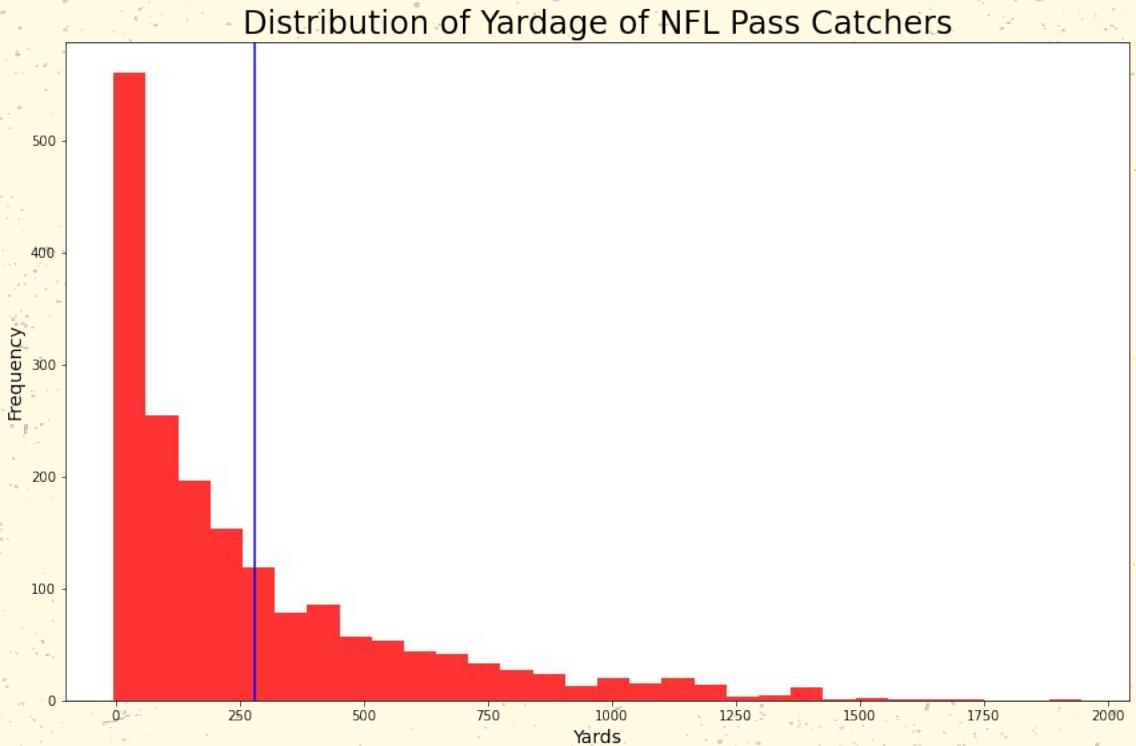
All positions decline in total yardage as they get older, and there are a lot fewer players.



Distribution of Yardage

Average Yardage:

279.65 yards



Yardage, by position

Average Yardage:

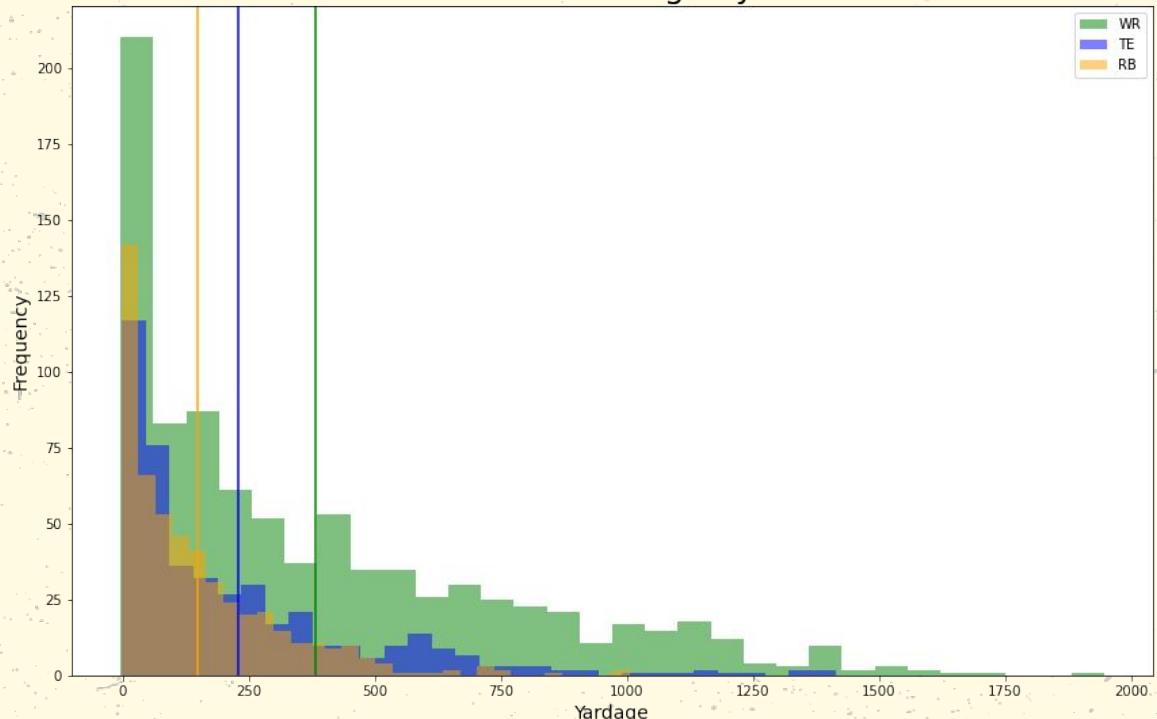
Wide Receiver: 382.64

Tight End: 229.93

Running Back: 149.14

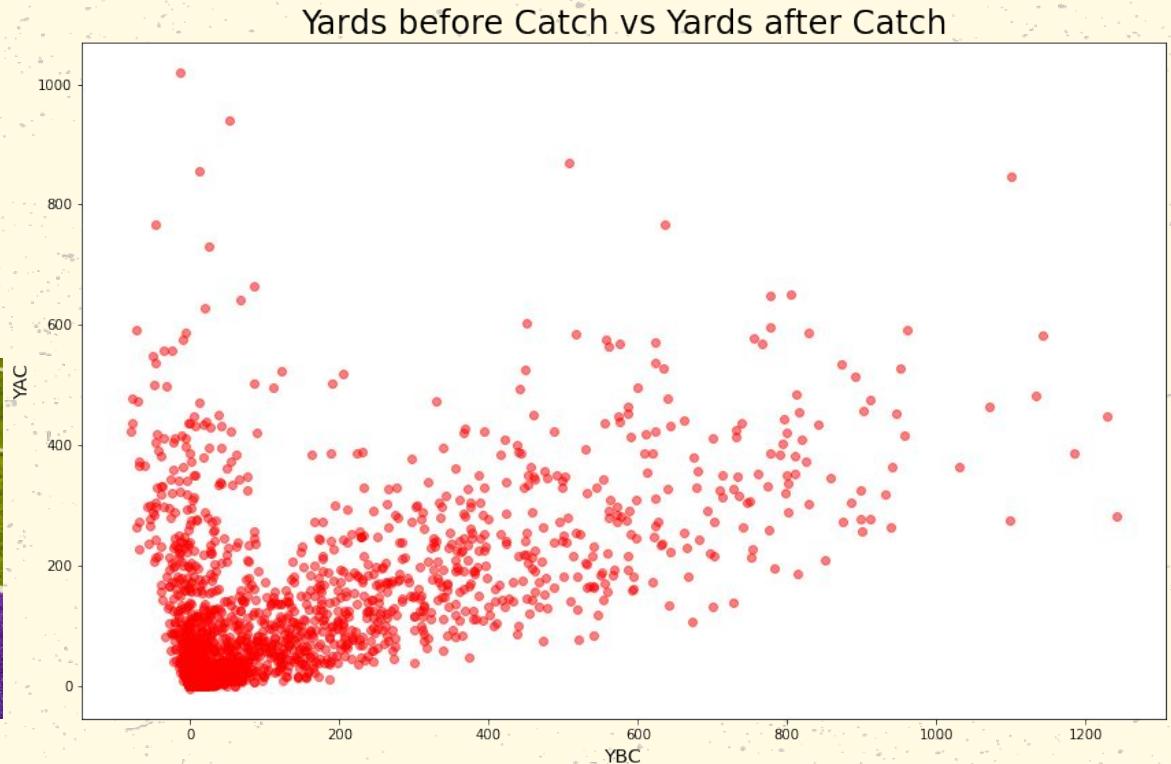


Distribution of Yardage by Position



YBC & YAC

There appears to be 2 groups with this data.



<https://www.dallascowboys.com/news/inside-ceedee-lamb-s-unbelievable-td-catch> (AP Photo/Jim Mone)

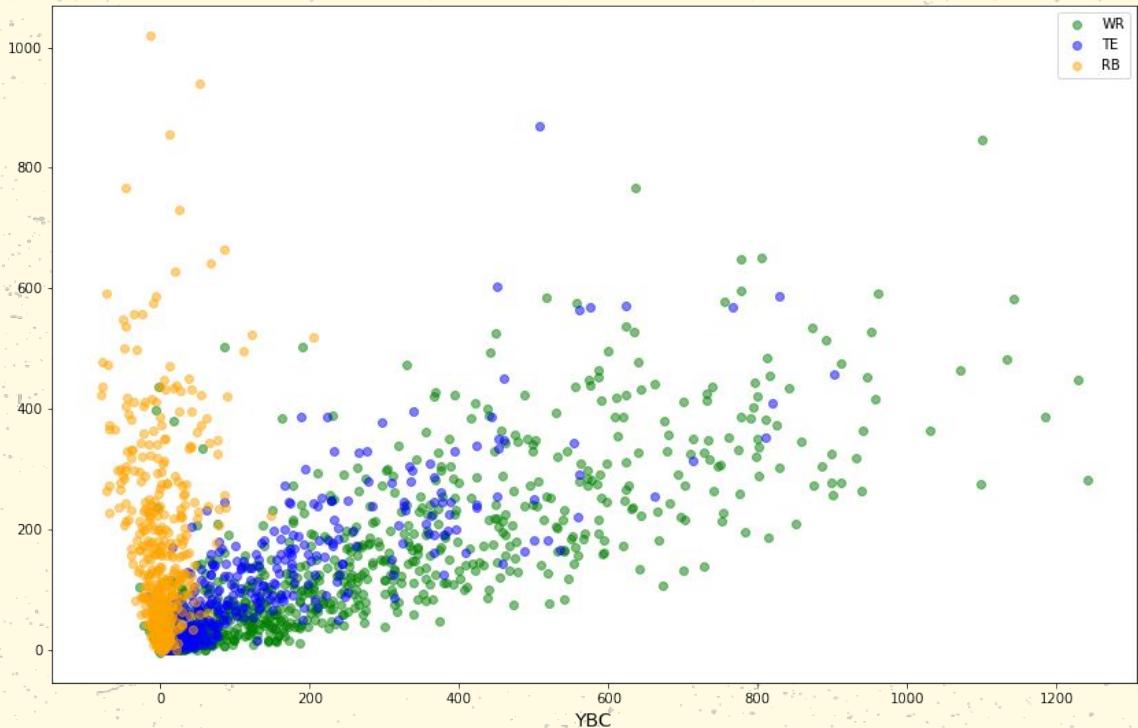
YBC & YAC, by position

When splitting the data by position, the difference between the 2 groups becomes much clearer.



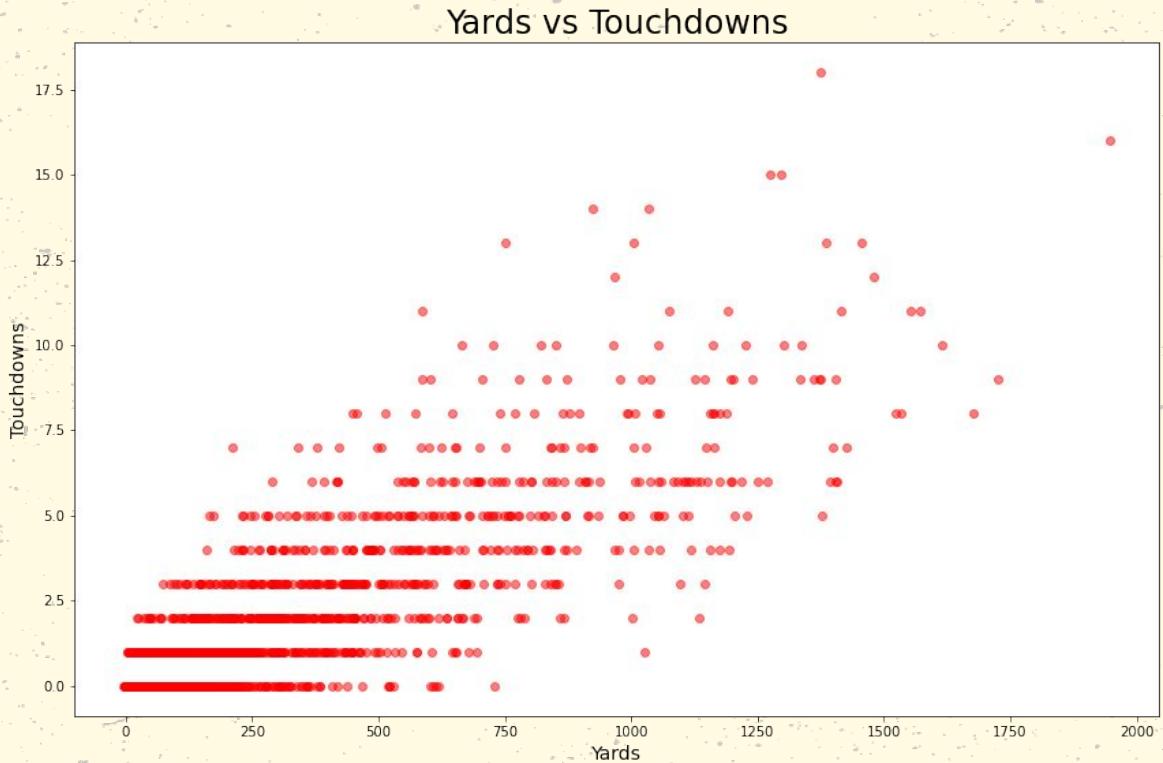
(Dan Powers/USA TODAY NETWORK-Wisconsin / USA TODAY NETWORK)

Yards before Catch vs Yards after Catch



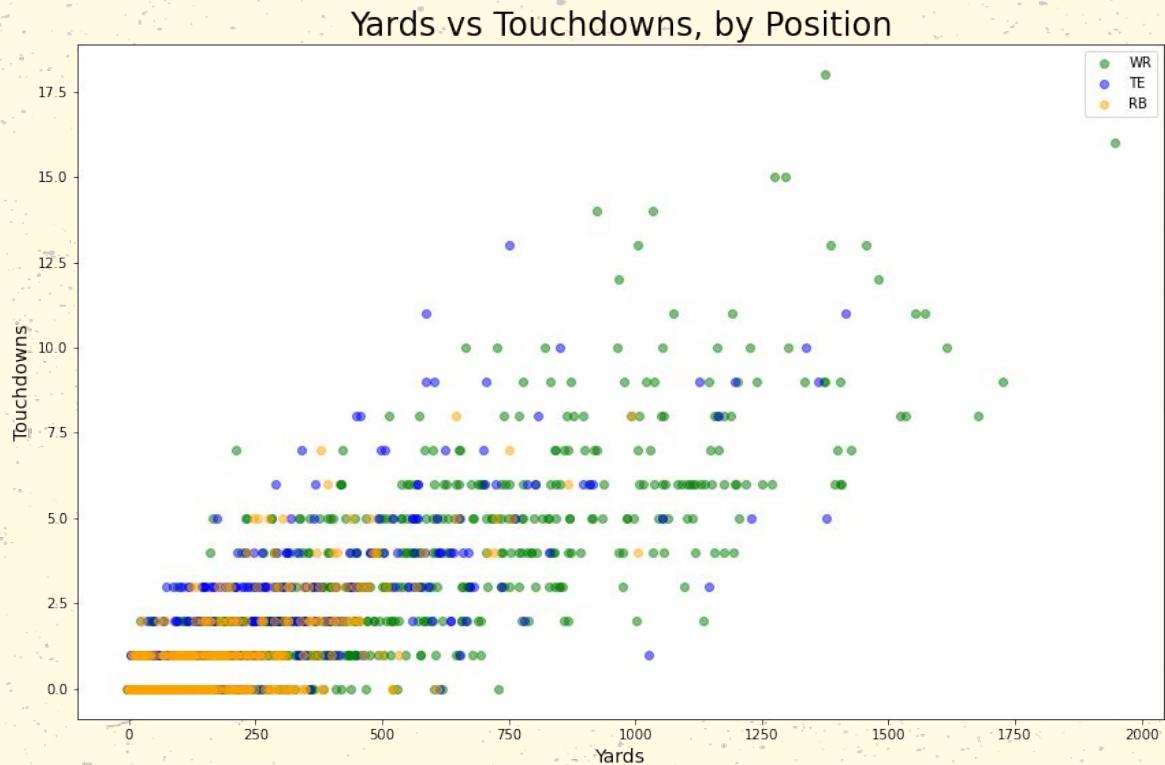
Yards vs Touchdowns

As Yardage increases, the number of Touchdowns scored tends to increase as well.



Yards vs Touchdowns, by position

Similar to the YBC vs YAC chart, when we split up the data by position, we get a better picture of what's going on.



Modeling

To predict Yardage, Receptions, and Touchdowns, I ran multiple models for each metrics.

Models included:

- Linear Regression
- RandomForest
- GradientBoost
- AdaBoost
- Ridge
- Lasso
- ElasticNet
- PCA using Linear Regression
- and a Stacking Model using a variety of models listed above



<https://www.boston.com/sports/new-england-patriots/2021/10/24/tom-brady-600th-touchdown-ball-mike-evans/> (AP Photo/Mark LoMoglio)



Features Used for Modeling



Games Played/Started



Player Age and Seasons played



Targets, Receptions, Yards, Touchdowns



Yards before & after catch



Dropped Passes, Broken Tackles



Stats from Previous Two Years



All-Pro and Pro Bowl Teams



Draft Position

Yardage Model

Linear Regression Model
Training Score: .579
Testing Score: .533

Why I chose this model?

I chose to use the Linear Regression model because it was not as overfit as most of our other models, but it still had a decent testing score.



<https://media2.miaminewtimes.com/mia/imager/u/magnum/12931599/jaylen-waddle.jpg?cb=1642608815> (Michael Reaves/Getty Images)

Receptions Model

Linear Regression Model
Training Score: .542
Testing Score: .479

Why I chose this model?

I chose this model due to it's relatively high testing score in comparison with other models, but also did not appear as overfit as some better scoring models.



https://static.www.nfl.com/image/private/t_new_photo_album_2x/f_auto/league/lhwciphxcrq8hpwchsp.jpg (Ben Liebenberg / NFL)

Touchdowns Model

Stacking Model

(RandomForest, GradientBoost, AdaBoost, and Lasso, results then fed into Linear Regression)

Training Score: .332

Testing Score: .314

Why I chose this model?

All of our touchdown models had relatively low testing scores, but I chose this model because of it's close training and testing scores, and it was one of the higher scoring models.



<https://www.raiders.com/photos/top-shots-te-darren-waller> (Tony Gonzales / Las Vegas Raiders)

Conclusions

Best Models for...

Yards: Linear Regression

Receptions: Linear Regression

Touchdowns: Stacking Model

This is a great start for trying to predict all three stats,
but the R^2 scores for all 3 models can be improved...



★ Next Steps ★

Improving the model:

- How does QB play impacts stats?
- How defenses impact stats?
- Develop a better method of filling in rookie stats / injuries
- Get more Data! PFR started keeping track of Advanced Receiving Metrics in 2018



2022 Predictions



<https://www.latimes.com/sports/rams/story/2022-02-13/rams-receiver-cooper-kupp-mvp-super-bowl-lvi>

(Robert Gauthier / Los Angeles Times)

Player	Receptions	Yards	Touchdowns
Cooper Kupp	116	1365	9
Davante Adams	107	1240	8
Justin Jefferson	100	1359	8
Stefon Diggs	85	1089	7
Chris Godwin	84	972	5
Tyler Lockett	84	1035	7
Mike Williams	82	1111	5
Deebo Samuel	81	1016	5
Diontae Johnson	79	893	6
Tyreek Hill	79	969	7
D.K. Metcalf	79	1062	7
Ja'Marr Chase	77	1010	6
Amari Cooper	75	975	6
Mark Andrews	75	876	7
Travis Kelce	74	889	6
Terry McLaurin	73	968	5
Mike Evans	73	959	7
A.J. Brown	73	977	6
Tee Higgins	71	933	5
Hunter Renfrow	71	765	5

My Top 5 vs ESPN's Top 5

Player	Receptions	Yards	Touchdowns
Cooper Kupp	116	1365	9
Davante Adams	107	1240	8
Justin Jefferson	100	1359	8
Stefon Diggs	85	1089	7
Chris Godwin	84	972	5

ESPN Player	Receptions	Yards	Touchdowns
Cooper Kupp	114	1348	10
Justin Jefferson	98	1367	8
Davante Adams	96	1156	7
Keenan Allen	93	1028	7
Stefon Diggs	93	1147	8

Both had Cooper Kupp as the #1 Receiver, and 4 of the 5 players were on both lists!

★ Questions? ★

CREDITS: This presentation template was created by [Slidesgo](#),
including icons by [Flaticon](#), infographics & images by [Freepik](#)

Please keep this slide for attribution

