NFL Combine & 40-Yard Dash Analysis

Spencer Slote - MS Practicum

University at Buffalo The State University of New York

Introduction

NFL Combine & 40-Yard Dash Background

NFL Combine & 40-Yard Dash

Every year, hundreds of NFL prospects gather for one of the most impactful talent evaluations in all of sports: the NFL Combine. Among the drills performed at this combine, none draws more attention than the 40-yard dash, which is one of the most watched and debated drills at the NFL Combine, often shaping how prospects are perceived by scouts and media. The main part of this project explores whether a player's 40 time can actually be predicted using the rest of their combine profile by using over 20 years of data and a machine learning model.

Why Predict the 40-Yard Dash?

- A player's 40-yard dash time is often seen as a snapshot of their speed, but it doesn't exist in a vacuum
- Other combine drills (vertical jump, broad jump, etc.) tend to be correlated with sprint times, revealing shared athletic traits
- Not all players run the 40 (due to injury or even strategy), which creates data gaps during evaluation
- Predicting 40 times from other combine data can help teams and scouts:
  - Estimate performance when the 40 time is missing
  - Contextualize outlier times (fast or slow)
  - Bring consistency and objectivity into evaluations
- Main question: Can we use the rest of a player's combine profile to predict their 40-yard dash time?

Data Overview

Combining & Cleaning 20+ Years of Combine Data

NFL Combine Dataset Overview

- Dataset includes over 8,000 players from NFL Combine (2000–2025)
- Covers physical measurements, drill results, position, college, and draft info
- Data from 2021 was sourced from Pro Day results due to Combine cancellation (COVID-19)
- Created new variables for some grouped positions (ex. OL, DB, Skill) to simplify analysis
- Cleaned and refined data to remove incomplete, non-comparable, or any erroneous entries
- Built multiple dataset versions depending on analysis type
  - Ex. Removed players missing 40 time for prediction model
- Data from: https://www.pro-football-reference.com/

PRO FOOTBALL REFERENCE

Data Cleaning & Preparation Decisions

- Players with missing 40-yard dash times were excluded from the main prediction model
- Used median imputation for other missing numeric drill results (shuttle, vertical, etc.)
  - Replaces missing values with the column median to reduce the influence of outliers
  - While not the most advanced method, it was still reasonable given the non-extreme amount of missing data - also didn't negatively impact model performance.
- Created both grouped and individual position variables depending on analysis goals
  - Grouped: Cleaner comparisons across roles (OL, DB, etc.)
  - Individual: More accurate for modeling & improved prediction performance
- Built custom datasets for modeling vs. exploratory analysis to avoid any bias / noise
- Addressed inconsistent data entries in excel (typos, non-standard formats) from initial dataset

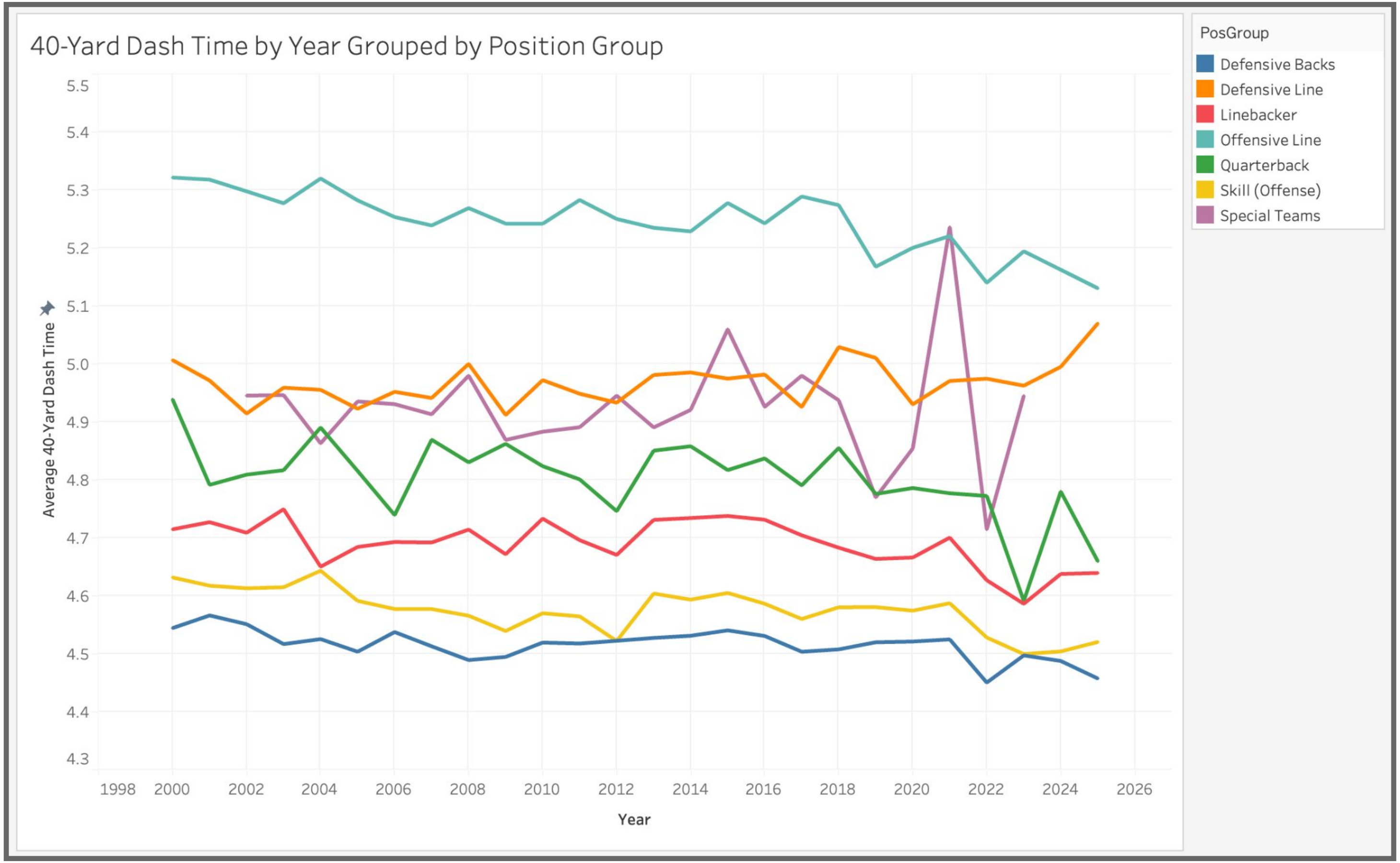
Exploratory Data Analysis

Patterns in Combine Performance Across Positions / Events

University at Buffalo The State University of New York

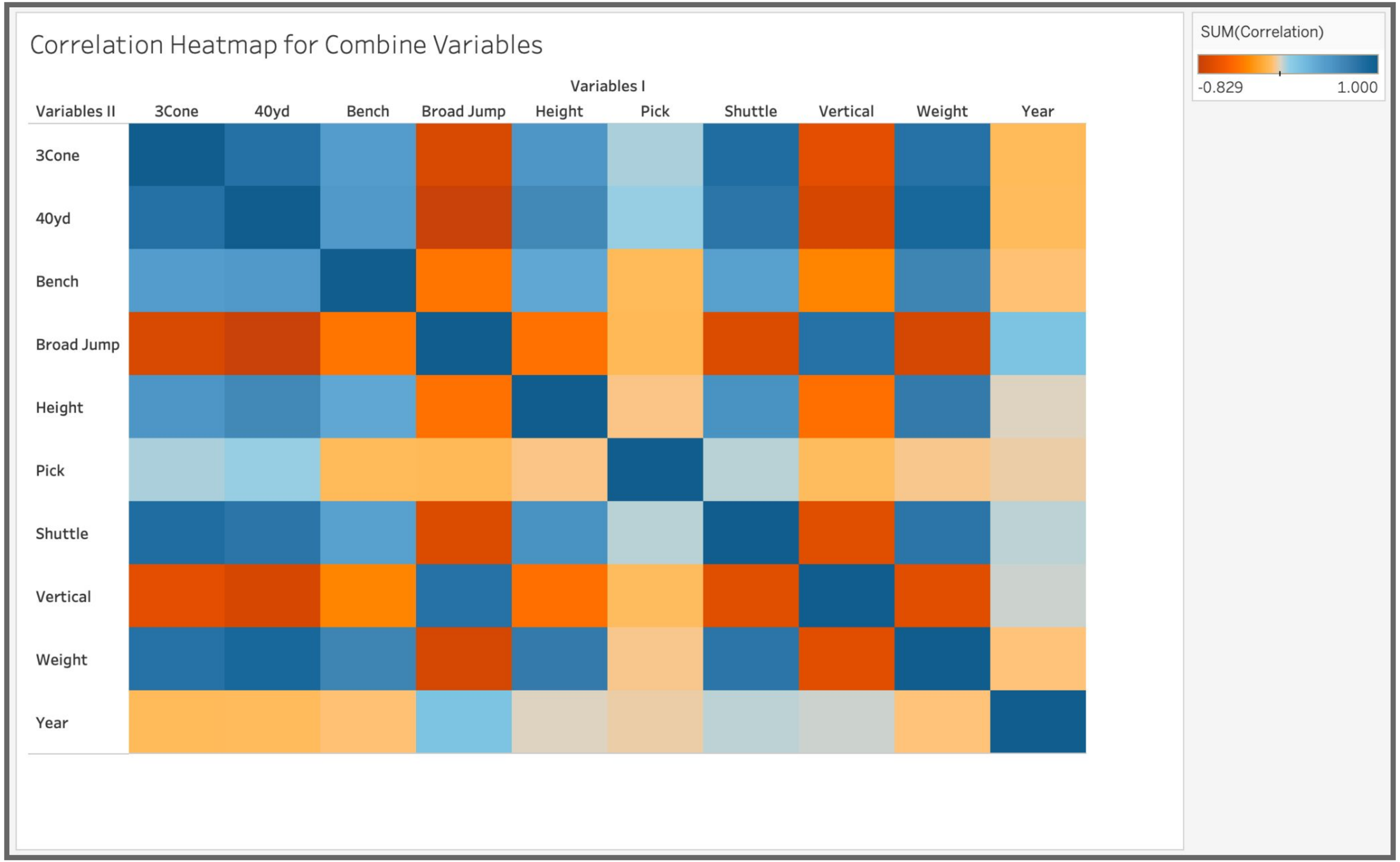
Drill Performance by Position Group

- Analyzed average 40-yard dash times across major position groups from 2000–2025
- WRs and DBs were consistently the fastest, as expected, averaging ~4.5–4.6 seconds
- OL had the slowest times on average but showed gradual improvement over the years
- Highlighted clear patterns between position group and sprint speed, though individual variation still exists
- Helped confirm that speed-based roles (WR, DB, RB) dominate this event, while the more strength-based roles trail


40-Yard Dash Time by Year Grouped by Position Group
Average 40-Yard Dash Time
5.5
5.4
5.3
5.2
5.1
5.0
4.9
4.8
4.7
4.6
4.5
4.4
4.3
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
2024
2026
Year
PosGroup
Defensive Backs
Defensive Line
Linebacker
Offensive Line
Quarterback
Skill (Offense)
Special Teams

Correlation Between Combine Factors

- Analyzed how combine measurements and drill results relate to one another
- Weight had a strong positive correlation with 40-yard dash time — heavier players tend to run slower as expected
- Broad jump and vertical jump both showed strong negative correlations with 40 time — more explosive athletes typically run faster
- 3-cone and shuttle were also highly correlated, reflecting similar agility traits
- Helped identify which metrics are linked, which are distinct, and how traits like explosiveness are related to sprint performance


Correlation Heatmap for Combine Variables
Variables I
Variables II
3Cone
40yd
Bench
Broad Jump
Height
Pick
Shuttle
Vertical
Weight
Year
SUM(Correlation)
-0.829
1.000

Correlation Matrix for Combine Variables

Variables II	3Cone	40yd	Bench	Broad Jump	Height	Pick	Shuttle	Vertical	Weight	Year
3Cone	1.000	0.824	0.466	-0.741	0.523	0.060	0.845	-0.677	0.814	-0.112
40yd	0.824	1.000	0.497	-0.829	0.623	0.082	0.783	-0.756	0.897	-0.102
Bench	0.466	0.497	1.000	-0.425	0.381	-0.109	0.427	-0.358	0.657	-0.075
Broad Jump	-0.741	-0.829	-0.425	1.000	-0.438	-0.117	-0.717	0.816	-0.759	0.193
Height	0.523	0.623	0.381	-0.438	1.000	-0.059	0.548	-0.449	0.739	-0.006
Pick	0.060	0.082	-0.109	-0.117	-0.059	1.000	0.041	-0.098	-0.052	-0.030
Shuttle	0.845	0.783	0.427	-0.717	0.548	0.041	1.000	-0.695	0.784	0.038
Vertical	-0.677	-0.756	-0.358	0.816	-0.449	-0.098	-0.695	1.000	-0.694	0.019
Weight	0.814	0.897	0.657	-0.759	0.739	-0.052	0.784	-0.694	1.000	-0.069
Year	-0.112	-0.102	-0.075	0.193	-0.006	-0.030	0.038	0.019	-0.069	1.000

SUM(Correlation)

-0.829 1.000

Key Correlation Takeaways & Coefficients

- 40-Yard Dash & Weight: **0.90** – Heavier players tended to run noticeably slower
- 40-Yard Dash & Broad Jump: **-0.83** – More explosive players generally ran faster
- 40-Yard Dash & Vertical Jump: **-0.76** – Similar pattern, reinforcing lower-body explosiveness
- 3-Cone & Shuttle: **0.84** – Strong overlap in agility-based movement
- Broad Jump and Weight: **-0.76** – Higher weight tends to reduce explosiveness

These patterns helped guide variable selection and gave early insight into which traits drive sprint speed.

Predictive Modeling Approach

Building a Model to Predict 40 Times Using Combine Results

University at Buffalo The State University of New York

Predictive Modeling Overview

- The goal of the model was to predict a player's 40-yard dash time using the rest of their NFL Combine profile
- Used a combination of physical measurements and performance in other drills as input variables
- Chose to use random forest regression due to its ability to capture complex, non-linear relationships and handle both numeric and categorical data
- Included all combine metrics and individual positions to give the model maximum relevant information because random forest can handle many input variables without overfitting
- Split data into training and testing sets (80/20) to evaluate model performance on unseen players

Preprocessing and Setup

- Removed players with missing 40-yard dash times from the dataset
- Filled in missing values for numeric combine drills using median imputation
- Converted the position variable into dummy variables so the model could interpret categorical data
    - Binary variable for each position (ex. Pos_WR = 1 for a WR or 0 for anyone else)
- Used individual positions (WR, QB, OT) instead of grouped categories in order to improve model accuracy
- Combined all preprocessing steps with the model into a single workflow to streamline training and testing
- No scaling was needed since random forest handles variables on different scales automatically
    - Ex. no need to normalize height in inches & weight in pounds

Model Training and Tuning

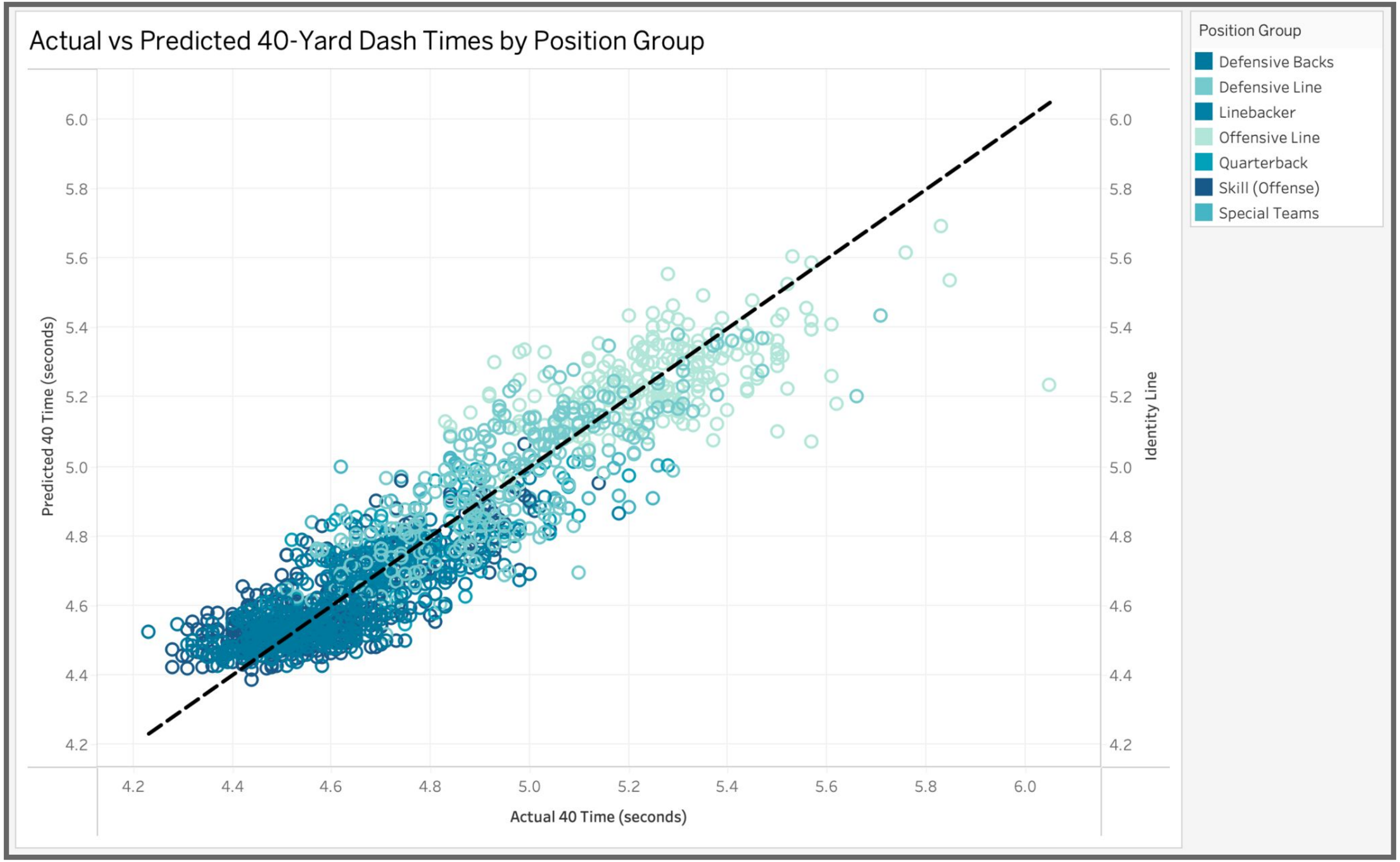
- Tuned two key hyperparameters:
  - mtry: number of variables considered at each split
  - min_n: minimum number of observations required to split a node
- Used 5-fold cross-validation to reliably evaluate different parameter combinations
  - The training data is split into 5 parts – the model trains on 4, and tests on the 5th, rotating until all parts are used
- Tested 10 combinations of mtry and min_n to find the best-performing setup
  - More combinations were possible, but the improvements would've been minimal compared to the extra time and complexity it would have resulted in
- Selected the model with the lowest RMSE to prioritize prediction accuracy

Model Performance

- Final model evaluated on the 20% test set using actual 40-yard dash times
- Achieved strong performance across all key metrics:
  - R^2 = **0.855** – the model explains 85.5% of the variation in 40 times
  - RMSE = **0.114** seconds – on average, predictions were off by about a tenth of a second
  - MAE = **0.088** seconds – most predictions were within roughly 0.09 seconds of the actual time
- Most 40-yard dash times fall between 4.2 and 5.5 seconds, so the prediction error was low
- Results show the model explained over 85% of the variation in 40-yard dash times
- Confirms that combine traits can accurately estimate sprint speed, even when the 40-yard dash is missing

Visualizing Model Accuracy: Actual vs. Predicted

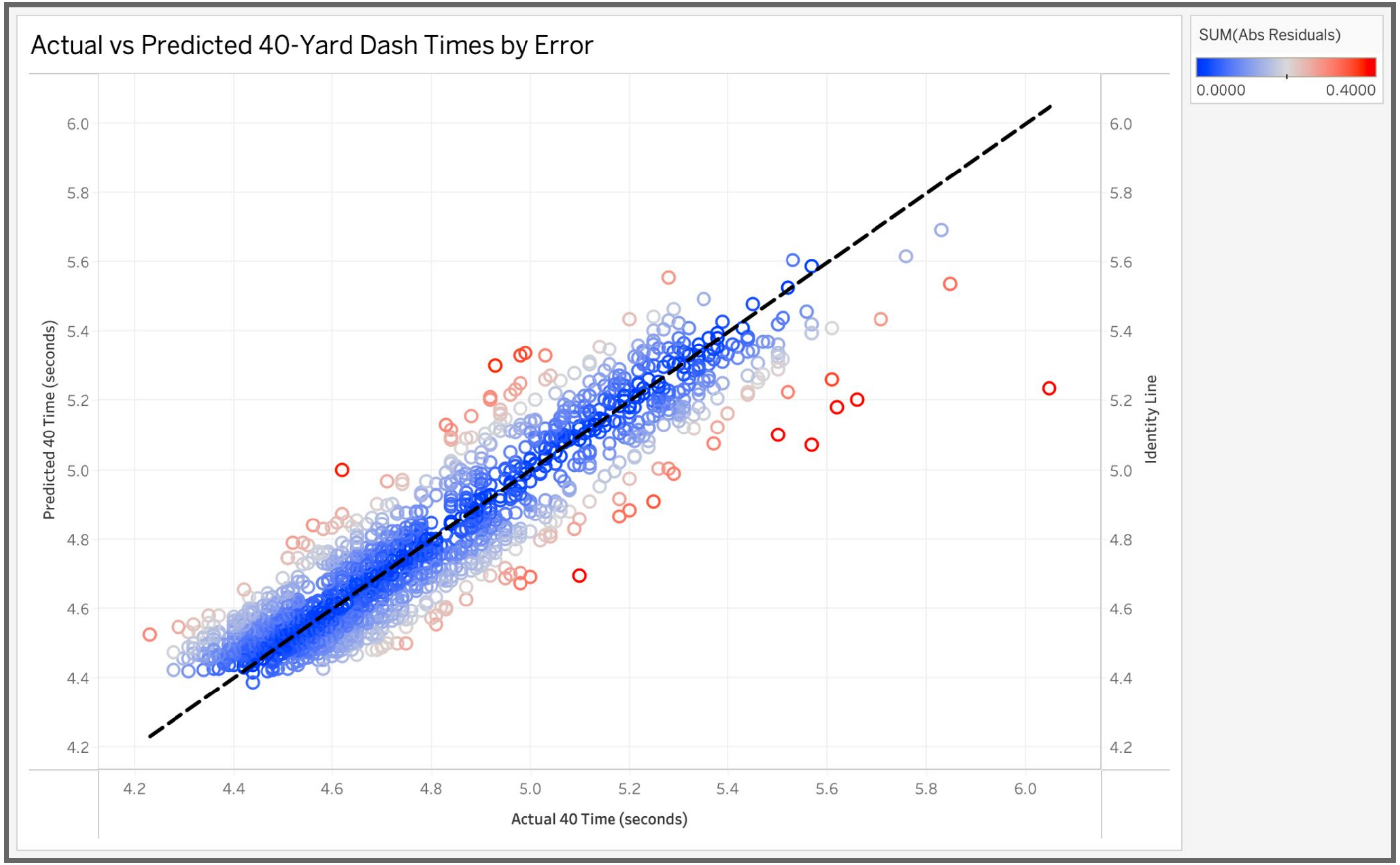
- Scatterplot shows predicted 40-yard dash time vs. actual time for each player
  - Grouped / colored by position groups
- Dashed diagonal line represents perfect predictions – points closer to the line are more accurate
- Most points clustered tightly around the line, showing strong consistency
- Some larger errors appeared in slower position groups (e.g., OL, DL), suggesting more variation in those predictions, since 40 times are larger and have a greater spread
- Confirms the model is reliable across a wide range of positions, with only a few outliers



Actual vs Predicted 40-Yard Dash Times by Position Group
Predicted 40 Time (seconds)
Actual 40 Time (seconds)
Identity Line
6.0
5.8
5.6
5.4
5.2
5.0
4.8
4.6
4.4
4.2
4.2
4.4
4.6
4.8
5.0
5.2
5.4
5.6
5.8
6.0
Position Group
Defensive Backs
Defensive Line
Linebacker
Offensive Line
Quarterback
Skill (Offense)
Special Teams

Prediction Error by Player: Residuals Plot II

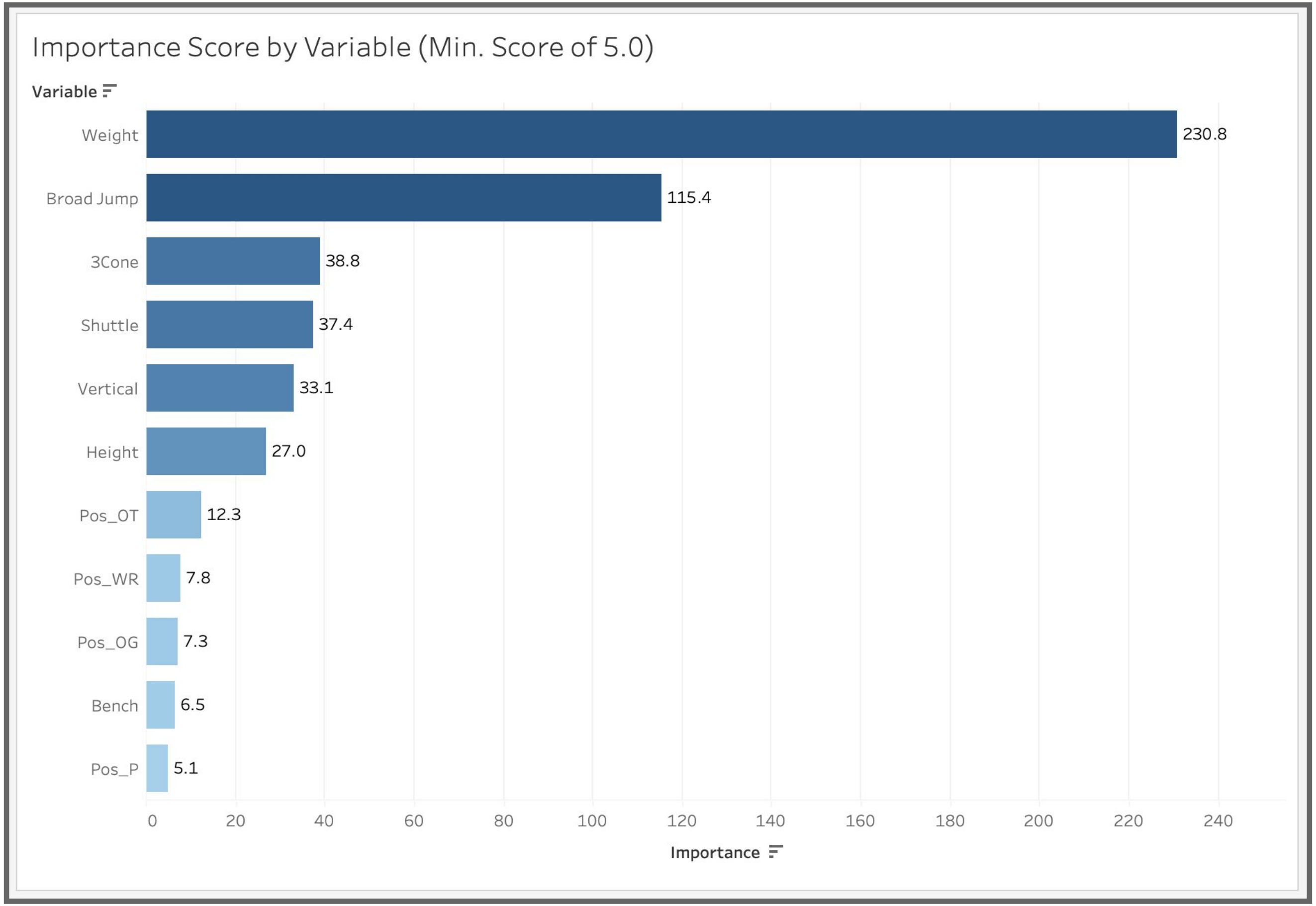
- Each point still represents a player, now colored by size of prediction error (residual)
  - Blue points = accurate predictions
  - Red points = larger misses
- Majority of players had small residuals, again indicating strong prediction accuracy
- Larger errors tended to come from players with unusual profiles or outlier performances
- Plot reinforces that the model is generally consistent, with only a handful of notable mispredictions



Actual vs Predicted 40-Yard Dash Times by Error
SUM(Abs Residuals)
0.0000
0.4000
Predicted 40 Time (seconds)
Identity Line
Actual 40 Time (seconds)
6.0
5.8
5.6
5.4
5.2
5.0
4.8
4.6
4.4
4.2

Variable Importance

- Used random forest's built-in feature for importance scores to identify top predictors
- **Weight** was by far the most important variable having the strongest link to 40-yard dash time
- **Broad jump**, **3-cone**, **shuttle**, and **vertical jump** were also high-importance, which makes sense, as they are all tied to explosiveness / agility
- **Height** had moderate importance, but much lower than weight
- Some position-specific dummy variables (WR, OT, OG) showed up with moderate influence
- Bench press and several other position indicators had low predictive value, suggesting limited relevance for straight-line speed


Importance Score by Variable (Min. Score of 5.0)
Variable
Weight
230.8
Broad Jump
115.4
3Cone
38.8
Shuttle
37.4
Vertical
33.1
Height
27.0
Pos_OT
12.3
Pos_WR
7.8
Pos_OG
7.3
Bench
6.5
Pos_P
5.1
0
20
40
60
80
100
120
140
160
180
200
220
240
Importance

Predictive Modeling Takeaways

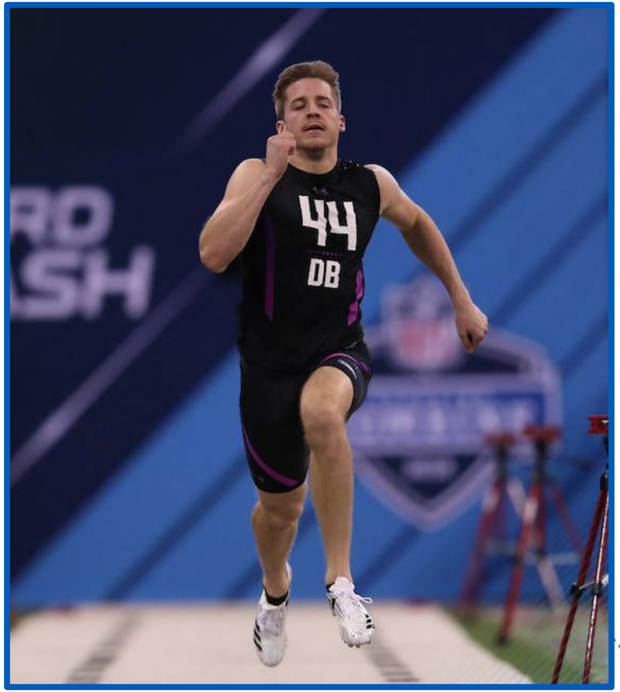
- The model was able to consistently and accurately predict 40-yard dash times using the rest of a player's combine profile
- Strong performance across all metrics (R^2 = 0.855, MAE = 0.088) helped to confirm the model's reliability
- Strongest predictors were weight as well as lower-body explosiveness and agility-based drills
- Showed that position matters, but more as a supporting factor than a primary driver
- Model provides a consistent, data-driven way to estimate 40-yard dash times, which is especially useful when results are missing, skewed, or taken out of context

Additional Analysis

Elite Athletes, Drill Trends, & Positional Standouts

Beyond the 40: Additional Combine Insights

- Explored the full combine dataset to uncover broader patterns in athletic performance
- Analyzed how position groups compare when results are normalized across events
- Looked at standout players with elite performances across multiple drills
- These analyses helped put the 40-yard dash into a bigger context and highlighted what rare athleticism really looks like



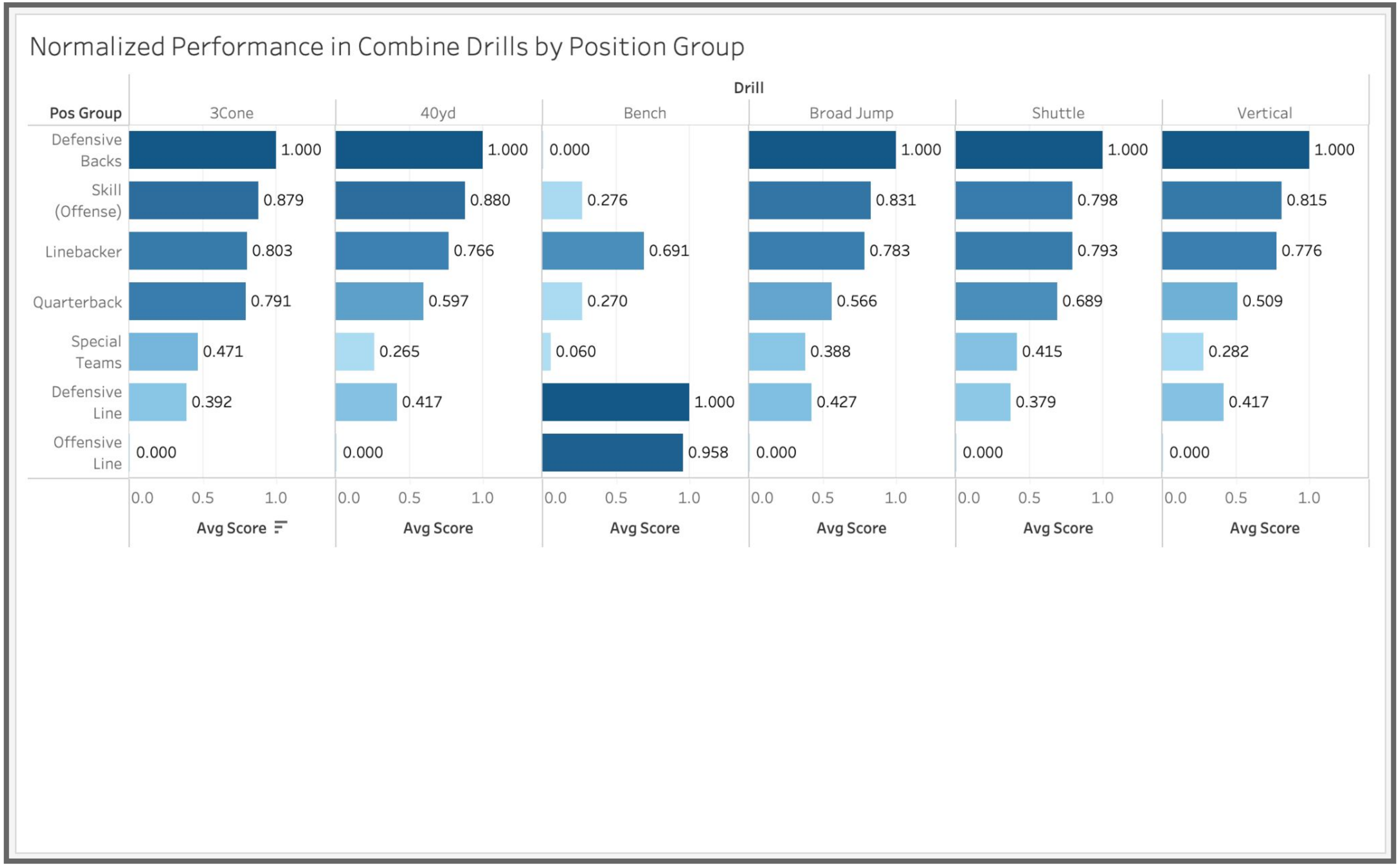
Beyond the 40: Additional Combine Insights

- Identified players who scored in the top 5% in at least four out of six core combine drills
- Only 12 players met this threshold out of over 8,000, showing how rare this type of all around athleticism is
- They were all WRs or DBs, which aligns with the speed and explosiveness focus of the majority of NFL combine drills
- Troy Apke was the top overall performer, having an elite performance in 5 out of 6 drills, more than any other player
- Even among elite performers however, few became major NFL names, showing that combine dominance doesn't guarantee career success


Elite Combine Performers
Drills
Player
elite 3Cone
Elite 40Yd
elite Bench
elite Broad Jump
elite Shuttle
elite Vertical
Troy Apke
Bob Sanders
Chris Chambers
Ja'Marr Chase
Jacob Harris
Josh Robinson
Marvell Tell
Parris Campbell
Scott Starks
Terence Newman
Tye Hill
Zyon McCollum
Pivot Field V...
Null
False
True

Position Group Comparison – Normalized Scores

- Compared performance across all six core combine drills using normalized scores (0 to 1 scale)
- Defensive backs posted the highest overall athletic profiles, scoring 1.0 in every drill except bench press
- Skill positions (WR, RB, TE) also performed well, especially in speed and agility drills like the 40 and 3-cone
- Offensive linemen ranked lowest in most drills, but scored near the top in bench press (0.96)
- Defensive linemen were below average overall but scored the highest in bench press (1.0)
- Linebackers showed the most balanced performance across all drills
- Results aligned with positional expectations – speed and explosiveness for skill roles, strength for linemen


Normalized Performance in Combine Drills by Position Group
Drill
Pos Group
3Cone
40yd
Bench
Broad Jump
Shuttle
Vertical
Defensive Backs
Skill (Offense)
Linebacker
Quarterback
Special Teams
Defensive Line
Offensive Line
1.000
1.000
0.000
1.000
1.000
1.000
0.879
0.880
0.276
0.831
0.798
0.815
0.803
0.766
0.691
0.783
0.793
0.776
0.791
0.597
0.270
0.566
0.689
0.509
0.471
0.265
0.060
0.388
0.415
0.282
0.392
0.417
1.000
0.427
0.379
0.417
0.000
0.000
0.958
0.000
0.000
0.000
0.0
0.5
1.0
Avg Score

What This Tells Us About Combine Performance

- Athletic expectations vary widely by position, and the combine reflects those role-specific demands
- All-around elite performances are extremely rare, even among thousands of top prospects
- Speed and explosiveness dominate the combine, especially for WRs and DBs
- Strength-based positions (OL, DL) tend to stand out only in bench press (only strength drill)
- Strong combine numbers don't guarantee NFL success, but they can highlight the physical traits that can set players apart
- These insights helped frame the predictive model and showed how combine data captures more than just raw speed

Conclusion

Key Takeaways From Modeling Sprint Speed & Combine Data

Project Summary

- Built a predictive model to estimate 40-yard dash times using the rest of a player's NFL Combine results
- Collected and cleaned data from over two decades of Combine performances (2000–2025)
- Used a random forest regression model, which achieved strong results:
  - $R^2 = 0.855$, RMSE = 0.114, MAE = 0.088
  - Identified key traits tied to sprint speed, especially weight, broad jump, 3-cone, and shuttle
- Model performed consistently across positions and could estimate 40 times when they are either missing or potentially unreliable

Takeaways and Future Impact

- Combine metrics are deeply connected, especially speed, explosiveness, and agility-based drills
- All-around elite performances are rare, and even the top combine performers don't always succeed in the NFL
- Predictive modeling adds much needed context to raw numbers – it can improve scouting decisions and reduce the over reliance on one drill
- This project showed how data science can translate athletic testing into actionable insights
- Opens the door to future work, such as predicting draft outcomes, career success, or analyzing bias in player evaluations

Thank You!

SCOUTING
COMBINE