NFL Hindsight: the search for NFL prospects

F. Baumgarten[1,2]

[1]University of Southern Denmark, Odense, Denmark
[2]Technical University of Denmark, Risø Campus, Denmark

Abstract

I here present "NFL Hindsight": a dashboard-styled Shiny app designed to explore how NFL Combine performance correlates with Overall player-rating scores achieved in the yearly updated NFL video game: Madden. All drafted players from 2013 to and including 2016 are evaluated (n = 1019). The visual tools contained within include the output of Principal Component Analysis (PCA) and Gaussian Kernel Density comparisons on NFL Combine data. To illustrate potential uses of this app, I work through an example analysis centered around the questions: Are NFL teams able to scout prospects? And are long arms paramount for "very good" Defensive Lineman? 66% of players achieving an overall Madden-rating of 89 in year four or five were selected in the first two rounds of the NFL Draft, proving that NFL scouting is not random. Finally, Defensive Linemen reaching Madden-ratings of 89 or above are not characterized by long arms.

1. Introduction

The NFL Combine and NFL Draft are two major stepping-stones for a player embarking on a quest to achieve a successful NFL career. Having - almost always - completed successful college careers, players are invited to the NFL Combine for an opportunity to bolster their draft "stock" before entering the NFL Draft, where NFL teams take turns selecting players to join their rosters. At the Combine, prospects participate in organized athletic (and other) disciplines to showcase their athletic ability as well as their physical profile and interview skills.

In this work I present a Shiny [CCA*22] visualization app I developed to explore Combine performance through the lens of career success. Quantifying player success is not a new idea. In fact, a very competitive market of player-evaluation has developed over the years, the perhaps most notable of which, anno 2022, is PFF [PFF22]. For this work, I have opted to use Overall Madden-ratings as the principal measure of success, a choice which is likely to cause some degree of eye-rolling among some NFL scouts. Madden is a NFL video game which has a new installment released every year, just before the NFL season starts. Madden-ratings range from 1 to 99, with 99 being the highest Overall rating that a player can achieve. Most metrics suffer from the same issues, highlighted by inherent subjectivity and complexity of evaluation and the black box nature of the underlying parameters to determine scores. Having said this, I will contend that Overall Madden-ratings do carry utility for evaluating groups of players. While I am sure it is possible to point to individual players or player-to-player comparisons with high disagreement among scouts and experts, when it comes to groups of players, I would posit that ratings correlate with general expert opinion. Further debate surrounding the utility of player evaluation metrics is certainly interesting, but beyond the scope of this paper.

1.1. Example analysis

In what follows I will present NFL Hindsight and its features while working through an example of how this tool might be used to explore questions pertaining to the relevance of NFL Combine metrics and success of NFL teams in the NFL Draft. For this example analysis I am looking for players that reach a Madden-rating of 89 or above in year four or five. The average starting player in the 2014 NFL season had a Madden-rating of $81 \pm 8(sd, n = 680)$ in the Madden game that came out prior to the start of that season. If we define a player who reach the mean rating of 81 as "good" then, a rating of 89 could be viewed as "really good". Much is made every year at the Combine of Defensive Lineman and the length of their arms. In fact, in the 2022 NFL Draft, one Combine highlight was the "short arms" of Aidan Hutchinson. It has been speculated to have been one of the major reasons why Hutchinson was drafted number two overall instead of number one overall. For this example analysis I therefore pose two questions: (i) are NFL teams able to scout prospects? (ii) is there evidence that arm length is important for NFL Defensive Lineman?

2. Related Work

The question of Combine performance vs future success has been investigated in the past with some reporting a general lack of an ability of the Combine to predict NFL success [KA08, CRSR20], although most are able to connect at least one metric with success in at least one position [VBH18, TCW15, LM20, KA08, CRSR20]. One example is the consistent finding that Running Back success and speed tests are linked, with faster 40-yard dash times correlating with success [KA08, VBH18, TCW15].

My work with NFL Hindsight is apart from the research mentioned above in that NFL Hindsight is an interactive environment meant for exploration of not only athletic disciplines but physical traits as well. Users are also able to survey NFL Draft outcomes with draft classes generally more recent than found in the literature cited above (most make use of classes from before 2011).

3. Design

NFL hindsight is an app that uses recent NFL Combine and NFL Draft data to produce exploratory visualizations of NFL Combine metrics' effect on the future Overall Madden-rating of players. I developed the app in R [R C22] using the Shiny [CCA*22] R-package. I made all figures using various combinations of R-packages: "cowplot" [Wil20], "ggplot2" [Wic16] and "ggridges" [Wil22]. For the full app design I used R-packages: "bslib" [SC22], "DT" [XCT22], "imager" [Bar22], "plyr" [Wic11], "png" [Urb22], "shinydashboard" [CB21] and "summaryBox" [per22]. The data, visualizations and explorative techniques are described further in the following subsections.

3.1. Data

The data contains 1019 observations from four years of NFL Combines and NFL Drafts spanning the years 2013 - 2016. A few examples of what the data contains are the round in the Draft a player was drafted along with NFL Combine metrics such as 40-yard dash times, height and hand-size. The list of NFL Combine metrics used in the analysis are shown in Table 1, and the list of player positions included are shown in Table 2. Note that I did not include all Combine metrics, because some have very low participation from player prospects. I also recorded each player's Madden-rating from the first iteration in which they appear up to and including the fifth year. All in all, the data consists of 43 columns. I sourced the data from "NFL Combine Results" [NFL22] and cross-checked with "Bleacherreport" [ble22] and "PRO-FOOTBALL-REFERENCE" [PRO22].

Table 1: *NFL Combine metrics with variable names in parenthesis.*

NFL Combine metrics (variable name)
Arm length (arm)
Broad jump (broad$_{jump}$)
Hand size (hand)
Height (height)
Vertical jump (vertical)
Weight (lbs) (wt)
225 lbs bench press (bench)
10-yard split (s) (X10ydsplit)
40-yard dash (s) (X40combine)

Table 2: *Player positions included in NFL Hindsight*

Player positions (abbreviation)
Quarterback (QB)
Runningback (RB)
Wide Receiver (WR)
Tight End (TE)
Offensive Line (OL)
Defensive Tackle (DT)
Defensive End (DE)
Defensive Back (DB)

3.2. Visual design and user interaction

NFL Hindsight is a dashboard-styled app with three main visualization tabs named "DRAFT OUTCOME", "PCA" and "PERCENTILES". Each tab provides an unique look into aspects of the data. Any setting available to the user on any tab can be changed anytime and will dynamically update the plots shown. I have made some effort to make plot-titles reflect at least part of the current settings being used for clarity's sake. The app also includes an "INTRO" tab and a "DATA" tab in which users can survey the data-set themselves.

3.2.1. Tab: DRAFT OUTCOMES

Provides a bar-chart look at the proportion of players meeting a user-set Madden-rating threshold in year four or five of their NFL career held against the round of the Draft in which they were drafted. One bar-chart (left) will always display all data from all positions, while the user is able to subset the data according to Draft class and player position(s) on the second bar-chart (right).

For the example analysis, these figures could be used to see if NFL Scouts were generally able to detect "really good" future players, because if they were, one would expect more "hits" in round 1 of the Draft, less "hits" in round 2 and so on.

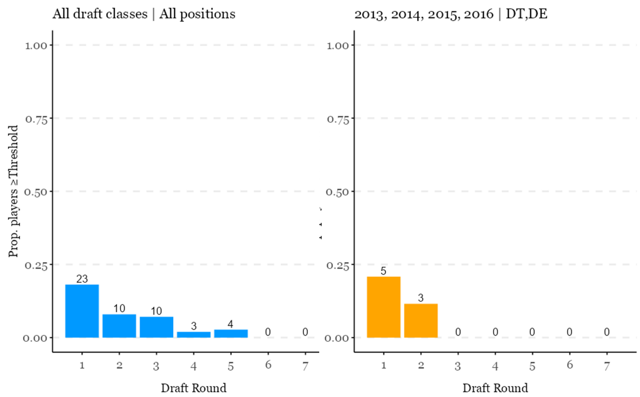

Figure 1: *The left-hand barplot shows the proportion of players meeting the Madden-rating threshold of 89 for all draft classess and positions. The right-hand barplot shows the same, except only for the "DE" and "DT" - positions.*

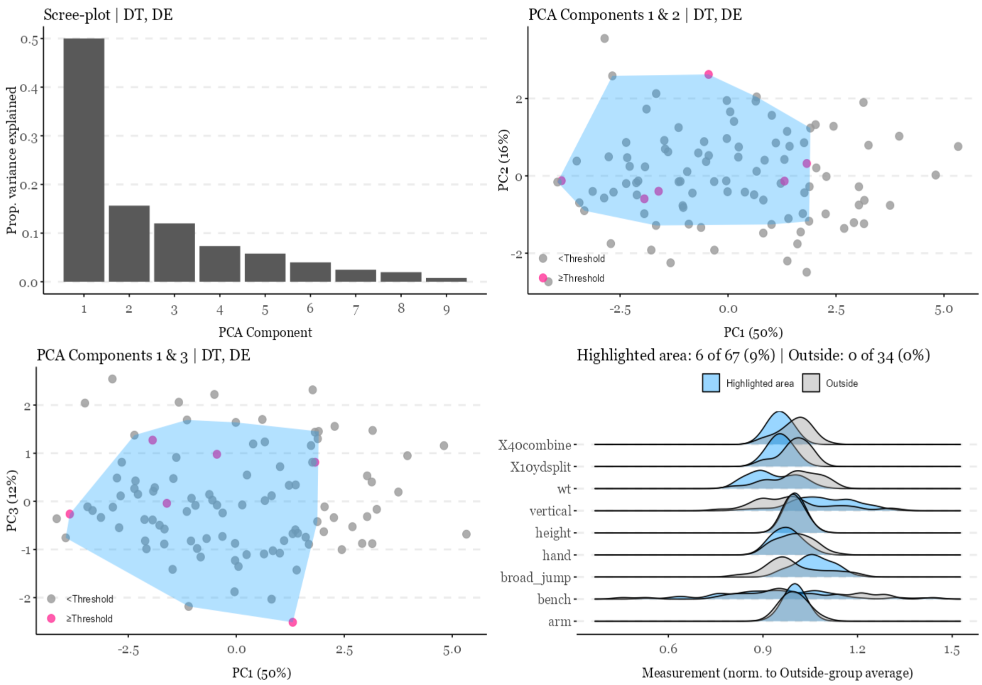

Figure 2: *Results of the Principal Component Analysis (PCA) for the "DE" and "DT"-positions, including a Scree-plot (top-left), PCA components 1 and 2 scatter-plot (top-right), PCA components 1 and 3 scatter-plot (bottom-left) and a Ridgeline-plot showing Gaussian Kernel Density-plots for NFL Combine metrics. The Madden-rating threshold was set at 89.*

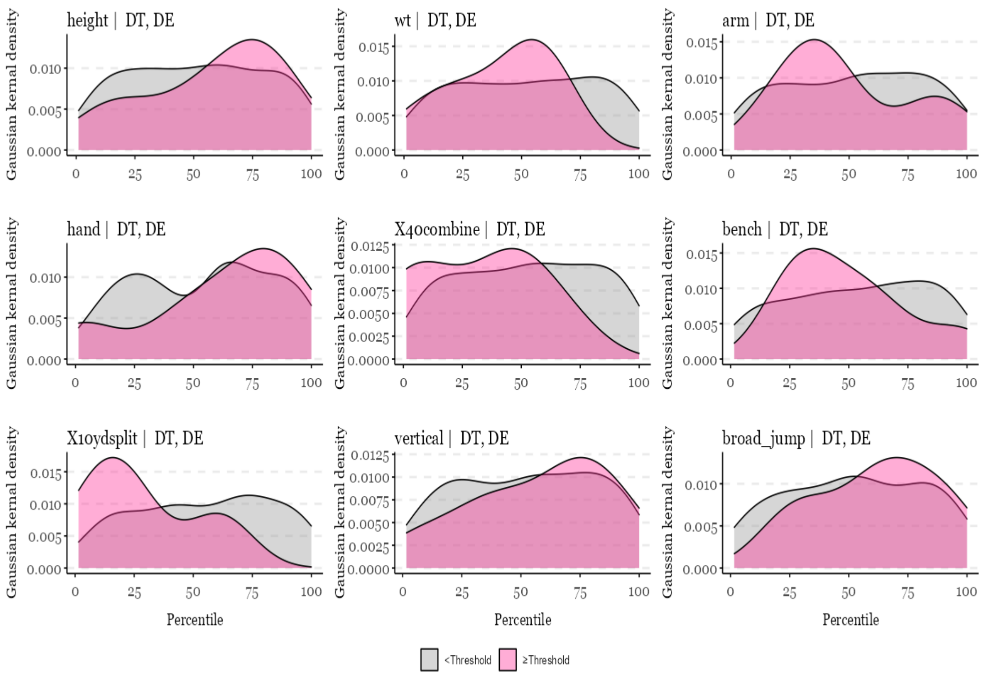

Figure 3: *Gaussian Kernel Density-plots of NFL Combine metrics for the "DE" and "DT" positions. Each metric shows the densities for the group of players that ended up meeting a Madden-rating threshold of 89 and of those that did not. Note that it is the density of player percentile scores for a metric compared only within the same position as the player.*

3.2.2. Tab: PCA

This tab shows the outcome of a Principal Component Analysis (PCA) performed on NFL Combine metrics chosen by the user (all included metrics are on by default). The user also sets which player positions are shown. The plot window includes a Scree-plot and a two scatter-plot view of the first three principal components. The user again sets a Madden-rating threshold, and all players equal to that rating or above in year four or five are marked. The user is able to set min-max sliders for each of the three components and highlight observations in the specified principal component space. A Ridgeline Gaussian density plot then compares distributions for each chosen metric between observations in the highlighted area and those outside it.

Returning to my example analysis, we could see if a clustering pattern emerges in PCA space. Do "really good" players find themselves in the same neighborhood in PCA space? If so, how dense is that space with lower-rated players? and how does this areas' Combine metrics stack up to those outside it?

3.2.3. Tab: PERCENTILES

In the final tab, each players' scores in each NFL Combine metric is compared with all other scores of players playing the same position, and the percentile that a player belongs to in that metric is produced. For example, Aidan Hutchinson's arms of 32.125 inches would belong to the 5th percentile among DEs in the years investigated here. Two groups of players are then produced: the group of players that achieved a Madden-rating equal to or surpassing a user-set Madden-rating threshold and the group of players that did not. Gaussian density plots are then compared between each group. The user is able to set any combination of player positions to view aggregated results from.

For the final consideration in my example analysis, the percentile-plots could help illustrate differences in individual Combine metrics such as "Arm length" between "really good" players and lower-rated players.

4. Results

Results from Figure 1 show that not many players ended up with Madden-ratings of 89 or above. Of 1019 players, only 50 met this threshold (5%). Nevertheless, it can be said that NFL teams chose these players primarily in the first round, then the second round and so on.

For the example analysis, my focus is on Defensive Lineman (DTs and DEs). The Scree-plot (Figure 2, top-left) shows that the first component is encompassing much of the variance (information) in the data (50%). And right after PC1, a sharp drop occurs which is then followed by an immediate-level off in the rate of change. This forms a break known as the "elbow-joint", and one interpretation is that I would gain little relative information including more components than PC1. By evaluation PCA-scores only on PC1, it is not obvious that any clear pattern exists for where Defensive Linemen trend to for Combine metrics looking at PCA scores only. It does not seem that any distinct clusters are formed.

It may be more useful to look directly at Arm-length distribution comparisons between "really good" Defensive Linemen and those that did not meet the "really-good" standard (Figure 3). Comparisons show that "really good" Defensive Linemen are not generally long-armed.

5. Conclusion

In conclusion of my example analysis: I originally posed two questions:
(i) are NFL teams able to scout prospects? 66% of the 50 "really good" players were selected in the first two rounds compared with only 33% in the subsequent five rounds. This shows with some clarity that NFL teams are indeed able to scout prospects.
(ii) is there evidence that arm length is important for NFL Defensive Lineman? Percentile density distribution comparisons show that "really good" players from the draft classes included here do not possess particularly long arms.

In this paper, I have presented the core functionalities of NFL Hindsight, the data that undergirds it, and shown one example analysis using it.

References

[Bar22] BARTHELME S.: *imager: Image Processing Library Based on 'CImg'*, 2022. R package version 0.42.14. URL: https://CRAN.R-project.org/package=imager. 2

[ble22] bleacherreport, 2022. http://www.bleacherreport.com/, Last accessed on 2022-12-22. 2

[CB21] CHANG W., BORGES RIBEIRO B.: *shinydashboard: Create Dashboards with 'Shiny'*, 2021. R package version 0.7.2. URL: https://CRAN.R-project.org/package=shinydashboard. 2

[CCA*22] CHANG W., CHENG J., ALLAIRE J., SIEVERT C., SCHLOERKE B., XIE Y., ALLEN J., MCPHERSON J., DIPERT A., BORGES B.: shiny: Web application framework for r. R package version 1.7.4.9000. URL: https://shiny.rstudio.com/. 1, 2

[CRSR20] COOK J., RYAN G. A., SNARR R. L., ROSSI S.: The relationship between the national football league scouting combine and game performance over a 5-year period. *Journal of strength and conditioning research 34*, 9 (September 2020), 2492—2499. URL: https://doi.org/10.1519/JSC.0000000000003676, doi:10.1519/jsc.0000000000003676. 1, 2

[KA08] KUZMITS F., ADAMS A.: The nfl combine: Does it predict performance in the national football league? *Journal of strength and conditioning research / National Strength Conditioning Association 22* (11 2008), 1721–7. doi:10.1519/JSC.0b013e318185f09d. 1, 2

[LM20] LAPLACA D., MCCULLICK B.: National football league scouting combine tests correlated to national football league player performance. *Journal of Strength and Conditioning Research 34* (01 2020), 1. doi:10.1519/JSC.0000000000003479. 2

[NFL22] NFL Combine Results, 2022. http://www.nflcombineresults.com/, Last accessed on 2022-12-22. 2

[per22] PERSON C. C.: *summaryBox: Value and Info Boxes in Shiny Apps and Rmarkdown*, 2022. R package version 0.1.0. URL: https://github.com/deepanshu88/SummaryBox/. 2

[PFF22] PFF: Pro Football Focus, 2022. http://www.PFF.com/, Last accessed on 2022-12-22. 1

[PRO22] PRO FOOTBALL REFERENCE, 2022. http://www.pro-football-reference.com/, Last accessed on 2022-12-22. 2

[R C22] R CORE TEAM: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2022. URL: https://www.R-project.org/. 2

[SC22] SIEVERT C., CHENG J.: *bslib: Custom 'Bootstrap' 'Sass' Themes for 'shiny' and 'rmarkdown'*, 2022. R package version 0.4.1. URL: https://CRAN.R-project.org/package=bslib. 2

[TCW15] TERAMOTO M., CROSS C., WILLICK S.: Predictive value of nfl scouting combine on future performance of running backs and wide receivers. *Journal of Strength and Conditioning Research 30* (09 2015), 1. doi:10.1519/JSC.0000000000001202. 2

[Urb22] URBANEK S.: *png: Read and write PNG images*, 2022. R package version 0.1-8. URL: https://CRAN.R-project.org/package=png. 2

[VBH18] VINCENT L., BLISSMER B., HATFIELD D.: National scouting combine scores as performance predictors in the national football league. *Journal of Strength and Conditioning Research 33* (10 2018), 1. doi:10.1519/JSC.0000000000002937. 2

[Wic11] WICKHAM H.: The split-apply-combine strategy for data analysis. *Journal of Statistical Software 40*, 1 (2011), 1–29. URL: https://www.jstatsoft.org/v40/i01/. 2

[Wic16] WICKHAM H.: *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. URL: https://ggplot2.tidyverse.org. 2

[Wil20] WILKE C. O.: *cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'*, 2020. R package version 1.1.1. URL: https://CRAN.R-project.org/package=cowplot. 2

[Wil22] WILKE C. O.: *ggridges: Ridgeline Plots in 'ggplot2'*, 2022. R package version 0.5.4. URL: https://CRAN.R-project.org/package=ggridges. 2

[XCT22] XIE Y., CHENG J., TAN X.: *DT: A Wrapper of the JavaScript Library 'DataTables'*, 2022. R package version 0.26. URL: https://CRAN.R-project.org/package=DT. 2