Exploring the Significance of the Wonderlic Test for NFL Players

# Our Team

* ***Jaiden Knecht*** ⇒ Junior Business Analytics Major | Marketing Co-Major | Working in the Audit Data Analytics Co-Op for Fifth-Third Bank next summer in Cincinnati, OH
* ***Sam Peeler*** ⇒ Junior Business Analytics Major | Currently working as a Business Analytics Intern for the Cincinnati Reds
* ***Samuel W Rogers*** ⇒ Senior Business Analytics Minor | Finance Major | Entrepreneurship Co-Major | Working as a Financial Analyst at Filtration Group in Oak Brook, IL Upon Graduation
* ***Ashley Stratton*** - Junior Business Analytics Major | Mergers and Acquisitions Consulting Intern at West Monroe next summer in New York City

# Introduction

**Problem Background**

The Wonderlic test was an IQ examination that all NFL hopeful players took at the NFL Combine from 1970 to 2022. Originating in 1932 by E.F. Wonderlic, the test measures general cognitive ability for retention, math, vocabulary, and reasoning. The test was adapted by companies in the 1940s to help employee selection, then moved to the US Army and National Football League in the 1970s. Ever since, the Wonderlic test has been a component of the NFL combine, which is the event before the NFL draft that measures potential players’ physical and mental capabilities. The Wonderlic test had been used to examine players' mental ability for 52 years and had the potential to help or tank NFL players’ draft stock.

**Significance of the Problem**

However, the NFL discontinued the Wonderlic score at the combine because of its irrelevance to players real-time performance. Determining if this was the correct decision, our team has set out to examine if and how NFL players’ Wonderlic test scores correlate with their performance in the NFL. This test was important for many years, and we are hoping to understand whether it was ever actually impactful and predictive or not. Because it was important to the NFL for so long, it is important to actually examine if the scores impacted players’ performance in the NFL.

We will be looking at salary, rankings, other scoring metrics, yards, and more to see how they all represent overall performance. This will hopefully help our team understand different measures of performance for players in the NFL, and we will get a better understanding of how Wonderlic fits in as a whole.

**Current State of Knowledge**

People have often wondered whether or not NFL on-field performance correlated Wonderlic test scores. Studies have been done in the past to try and determine some sort of correlation. One of the more prominent studies was done in 2009 by Dr. Brian Hoffman and Brian D. Lyons in collaboration with California State University (Fresno) and Towson University. In the study, NFL performance was found to have a strong correlation with the Wonderlic test among only two positions: tight ends and defensive backs. Also, these 2 position groups showed negative correlation. There have been many studies completed in many different ways. Another study focused on only quarterbacks and found that when they removed quarterbacks who never really played in the NFL (1000 passing yards or less), there was a strong positive correlation between NFL performance and Wonderlic scores. There are many different possible routes to take when examining the correlation between Wonderlic scores and performance. Our team has chosen to observe the possible correlation and trends between Wonderlic scores and various measures including Approximate Value scores, fantasy football rankings, leading quarterback statistics, and player salaries.

# Data Gathering & Quality Assessment

**Sources of Data & Approach to Gathering Data**

Our first step to collecting data was finding credible sources that are highly respected in the sports world. We hoped to find objective data by NFL experts rather than opinions on rankings by sports fans. After determining which sites were credible, such as Pro Football Reference, Fantasy NFL, ESPN, and more, we had to determine which ones were able to scrape into RStudio. We used robots txt on all of our sources, determining which data we could pull effectively. With those results, we started to get a direction for the data. Focusing mainly on player salary, fantasy rankings, and certain position rankings, we tried to get a wide variety of different performance metrics to compare to Wonderlic data.

**Data Quality and Descriptive Statistics Assessment**

Once scraped and in R, the next step had to do with cleaning. Pulling the data from a website, the data will obviously not be tidy or cleaned, which was one of our biggest hurdles. Before we could do any analysis, we had to go through and make sure the data was usable. There are many steps to cleaning data, but we first wanted to make sure the data was tidy. This included a lot of strsplit() for different columns. In our data, we had many occurrences where the player, team, and position were all in the same column. It seems simple to fix, but many names are not just two words, which made tidying the data difficult. On top of that, not all the data was formatted the same because of missing values, which complicated the cleaning process.

Also, when pulling from websites, most of the positions were separated by tabs on the same website. This made the scraping process difficult, causing some processes to be repeated many different times before making a full dataframe. We ended up making large, combined dataframes for different positions and eliminating variables that were inaccurate and inconsistent. Once we had clean and tidy data, we merged each type of performance measure with Wonderlic data. This allowed us to see how individual performance indicators might correlate with Wonderlic scores. After merging data, we used the DataExplorer package in R in order to help us view the quality of the data as well as descriptive statistics and visuals:

*Quarterback Leaders Ranked x Wonderlic Scores*

Looking at the report from DataExplorer in R, we can first see that there are no missing values. For each variable, the QQ plot shows if the data stays near the trendline. The lines are pretty usual for QQ plots with curves on the ends and on the trendline in the middle. For Yds, it does vary more than the others, so we can know that this variable is less reliable and varies from normality. For the Univariate Distribution, there is a variety of normal, skewed right, and skewed left. Score and Year Start are normal distributions, but the other variables are all skewed. Since Rank and Yards are both skewed right, we can think about the amount of smaller vs larger numbers. For Yards, it is harder and less likely to get more yards in the NFL, which is why there are less players with high total yards. For year end, the data was obviously more current, making most of the year ends in the 2020s. Going into the heatmap, it is clear there is not a ton of correlation between the variables. Specifically for Wonderlic score, it does not specifically correlate to any variable. The main positive and negative correlations reside with Year Start vs Year End and Yards vs Rank.

*Fantasy Rankings x Wonderlic Scores*

For the data that was merged regarding Fantasy Rankings and Wonderlic Scores, the DataExplorer report from R gives us more detail about the data normality and correlations. Importantly, there are no missing values in this dataset. For the Univariate Distribution, the histogram for ranking is skewed to the right while the histogram for score seems fairly normal, barring any gaps in the data. The ranking was expected to be skewed because each position is ranked from 1 to 80. However, some positions are ranked 1 to 30 because there are less of them in the league, like quarterbacks. The QQ plots show the data staying on the trendlines very closely, which is good for the normality assumption. Finally, the heatmap has no glaring correlations, which may sound concerning for the rank vs score; however, this was expected based on other findings.

*Player’s Salary Data* *x Wonderlic Scores*

Using the DataExplorer package in R, we can observe a variety of things about the merged Player Salary and Wonderlic Score data. There are no missing columns or observations as well as no missing data. As we look at the QQ plot, we can see that while salary veers away from the trend line, score displays a fairly normal distribution. The univariate distribution histograms show that salary is skewed to the right, while score is normally distributed. The skewed data in salary is expected, as there are a small number of players that will achieve an above average salary. When looking at the correlation through the heatmap analysis, we see that there is not a strong correlation between score and salary. While this may seem concerning, it follows the narrative we are hoping to share that Wonderlic Scores are not in fact a relevant predictor of a player’s future success.

*AV Player Scores x Wonderlic Scores*

From using the DataExplorer package in R on the AV and Wonderlic scores data, we can first see that this data is high quality. There are no missing columns or observations in this data set, and all rows are complete. From the Missing Data Profile chart, we can see that all variables have zero data missing. As we look at the QQ Plot, we can see that all plots veer away from the trend line, suggesting that there could be some discrepancies in normality. Next, we can look at the descriptive visuals to gather early insights on the data. When looking at the Univariate Distribution histograms for AV and Wonderlic scores, we see that the Wonderlic histogram is relatively normally distributed, while the AV histogram is skewed to the right. This tells us that many players have low AV scores. Finally, when looking at the Correlation Analysis heatmap, we can see that correlation is lacking between more important variables like AV and Wonderlic scores. This could be expected, as we assumed that there would not be much correlation among the raw data without any manipulation to AV scores. Although there seems to be strong correlation between From and To, these variables are not important to the overall logic behind our insights.

After all of our individual data was clean, we merged the data sources using innerjoin and merge, making dataframes that were useful for Tableau. Once all of our data was tidy, technically correct, and consistent, we then moved on to data analysis and visualizing our data for the audience.

# [Link to Storyboard](https://public.tableau.com/app/profile/samuel.rogers7112/viz/MAINWonderlicProjectISA401/ExploringtheSignificanceoftheWonderlicTestforNFLPlayers?publish=yes)

# Data Explanation & Analysis

**Data Sources and Exploration**

*Wonderlic Scores*

This data provides 406 observations throughout the history of the NFL that gives a glimpse into where certain players and positions Wonderlic scores landed. This allows us to see where position groups ranked on average with their Wonderlic scores. We were able to pull from two different sources of historic Wonderlic data so that we could increase our sample size. What was limiting about our dataset was how many different generations it covered and how few were sampled from today’s athletes. This data was at the center of a lot of our analysis and it was important that we pull from both data sources.

*Quarterback Leaders Ranked*

This data, pulled from Pro Football Reference, gives 250 observations about past and current NFL quarterbacks. The data is composed of 7 columns: Rank, Player, Yards, Team(s), Year Start, Year End, and Hall of Famer. These metrics allow us to see the best quarterbacks, their overall yards, and whether they are in the Hall of Fame. When this data was pulled, it needed some cleaning due to multiple metrics sharing the player column. That is where the Hall of Fame column comes from. Also, for the Rank, that is a metric that was created by Pro Football Reference, which may have some subjectivity to the quarterback rankings. We use this data for many different functions, comparing Wonderlic scores to hall of famers, yards, ranking, and years in the NFL.

*Fantasy Rankings*

This data pulls the top 262 individuals in football based on their Fantasy Rankings. Fantasy.NFL.com has players separated by position then ranked based on other metrics that are not included. This data gives those final rankings with the player, position, and team for 2023. For cleaning purposes, the news article links were eliminated and the name, position, and team were separated. This table can help us compare different positions and their Wonderlic scores while also analyzing the top versus bottom ranked players.

*Player’s Salary Data*

This data includes over 1,700 observations of players from all current NFL team rosters and their salaries. The data is composed of team abbreviations, player salary, position, player first name, player last name, and the state and city where the team they play for resides. This data was useful in investigating the relationship between player’s wonderlic scores and their salary. Additionally, we used this data to compare average salaries between all positions.

*AV Player Scores*

This data pulls 22,123 NFL players from ProFootballReference.com based on Approximate Value scores. Approximate Value (AV) scores are a single number that measures the value of any given NFL player’s career. It is important to note that this is more of a career measure than just a performance measure, as AV scores have a strong positive correlation with career length. AV scores do a great job of assigning a number to a player based on how successful their career was. The data itself was pulled from tables on ProFootballReference.com. Each position group had its own table with hundreds of players listed. The variables that were pulled for each position group were player name, position, total AV score, career start year, and career end year.

**Data Analysis**

After all of our data processing and visualizations, we were able to produce some insightful analyses. We found that different approaches, like trying to separate by position or look into different metrics such as fantasy and AV, were beneficial when analyzing NFL performance and Wonderlic. For specific positions versus average scores, there is not much defining the Wonderlic scores to different positions, and even though punter is the highest scoring average, there are only 2 observations for that group, making the punters skewed. As for career length with the wonderlic metric, there is not a strong linear correlation between the two based on the tableau dashboard, but there is strong positive correlation for career length and AV. This is when we started to realize that Wonderlic may not be the best predictor of success in the NFL for players.

For AV analysis, we originally discovered that there was no correlation between AV and Wonderlic scores for all players. However, when we remove players with an AV score less than 50, we can start to see a slight positive correlation. Filtering out players that never got much playing time in the NFL also worked for one of the studies that was explained in the current state of knowledge section above. Continuing with viewing data where AV score is 50 or greater, we broke down average AV versus average Wonderlic score by position. Here, although not perfect, we saw a positive correlation with defensive backs on the low end and quarterbacks and offensive linemen on the high end. Defensive and offensive linemen deviated slightly from the trend line due to an imbalance between higher wonderlic scores and slightly lower AV scores. Finally, we looked at the correlation between AV and Wonderlic scores for each position individually, containing all players in each respective position. For this, we continued to use AV scores of 50 and greater. Here, we saw that defensive backs and quarterbacks both showed positive correlation, while wide receivers and running backs both showed negative correlation. These findings tell us that for defensive backs and quarterbacks, the higher the Wonderlic score, the higher the AV. And, for wide receivers and running backs, the lower the Wonderlic score, the higher the AV. In plain English, this means that if given playing opportunities, defensive backs and quarterbacks with higher Wonderlic scores have a greater chance to have better careers. And, if given playing opportunities, wide receivers and running backs with lower Wonderlic scores have a greater chance to have better careers.

Now looking at the Fantasy rankings for NFL players, we found that there were differing correlations for each position. For one position, it is a positive correlation of Wonderlic, while others were not highly correlated or negative. The fantasy rankings for players is based off of recent performance and expected performance, so it is a more present indicator of success, while Wonderlic is a test from the NFL combine that is supposed to measure a player’s mental capacity for their whole career. This difference in time periods can play a part in the variety of correlations. For other rankings, such as all-time quarterback rankings, there is a similar result. Wonderlic does not necessarily correlate, but another measure of success, specifically total career yards, the quarterback rankings correlates. This furthers the discovery that Wonderlic does not necessarily correlate with career success.

Finally, the last metric we focused on for player success was salary. The tableau shows different graphs with Wonderlic and salary encoded by size, and as shown across the league, salary shows a bigger difference in big market, successful teams while Wonderlic average stays similar throughout. This shows that salary is a better reflection of success for a player and team than Wonderlic scores. When looking into different positions, it is clear that there is an inequality between the salaries for players in certain positions. From all of this, there is a clear lack of correlation for Wonderlic scores and player performance in the NFL. The other metrics: AV, ranking, yards, and salary, were an overall better representation of player success for an NFL career.

# Learning Outcomes & Conclusion

**Data Insights and Extraction Process**

After our data visualization and analysis processes, we took away several key insights on player performance. The main hypothesis that we wanted to solve was whether or not the NFL could justify a test that has been used for decades in the league. Our visualizations prove that at the base level, the measure of player intelligence was not an indicator about how they would perform. An underlying assumption our team had was a heavy correlation between Wonderlic scores and quarterback performance due to that position requiring intelligence and quick decision making. Much to our surprise, the Wonderlic data did not prove this assumption correct unless quarterbacks who never really got a chance in the NFL were removed from the AV data.

Through our data analysis journey, we found that looking towards other metrics to determine player performance was worthwhile. Utilizing metrics such as all-time fantasy rankings, AV scores, and compensation proved effective in measuring player ability. We shifted our project to focus more on comparing Wonderlic and the other metrics for their correlation, and we ended up finding that the Wonderlic score was one of the worst determinants for success. With that, we came to the final conclusion that the NFL was correct in getting rid of the Wonderlic test at the combine last year, saving players from possible bad scores that are arguably insignificant.

Throughout this whole process, our team took what we knew from ISA 401 to analyze a meaningful and relevant topic. We found that the hardest part was the process of transforming the data from numbers to actual insights. After we extracted, transformed, and loaded the data into Tableau, we felt stuck on where to take the project. Once we started to find that Wonderlic score is not a great indicator of success, we started to go back and draw out potential data visuals that could be beneficial. From there, the insights from the visuals became clear that the NFL was correct in discontinuing the Wonderlic score. Overall, this whole project taught us more about building insightful visuals that can help create concrete answers from data, which we saw in our Wonderlic data.

# Sources

* [Huge List of NFL Wonderlic Scores by Position](https://wonderlictestpractice.com/nfl-wonderlic-scores/)
* [NFL Passing Yards Career Leaders - Pro Football Reference](https://www.pro-football-reference.com/leaders/pass_yds_career.htm)
* [Player Rankings - NFL Fantasy](https://fantasy.nfl.com/research/rankings?leagueId=0&position=QB&statSeason=2023&statType=seasonStats&week=10)
* [The Story of Wonderlic](https://wonderlic.com/about/history/)
* [What Does the Wonderlic Test Really Mean? - Bleacher Repor](https://bleacherreport.com/articles/2431408-what-does-the-wonderlic-test-really-mean#:~:text=%22The%20issue%20isn%27t%20whether,Wonderlic%20results%20and%20future%20performance)t
* [What You Need To Know About the Wonderlic Test and the NFL Draft](https://www.playerprofiler.com/article/wonderlic-test-nfl-draft-advanced-stats-metrics-analytics-profile/)
* [Wonderlic Score Database](https://footballiqscore.com/wonderlic-score-database)
* [AV Score Description](https://www.pro-football-reference.com/about/approximate_value.htm)

[GitHub Repository](https://github.com/sampeeler/isa401_project)