

Data Visualization Of Nfl Offensive Player Stats,1999-2013 Dataset

Author: Gajabheenkar Roshini

Scholar ,Dept. CSE(AI&DS)

Methodist College Of Engineering and Technology Hyderabad,India

roshinigajabheenkar@gmail.com

Guided By:

Dr.Lavanya Pamulaparty

HOD of CSE Department

Methodist College Of Engineering and Technology, Hyderabad,India

Abstract- This research focuses on analyzing NFL offensive player statistics from 1999–2013 using Tableau visualization techniques. The dataset contains player demographics combine performance, draft information, and offensive statistics. Various visualizations such as bar charts, heat maps, dashboards, and highlight tables are used to identify player performance trends and statistical patterns. The analysis helps understand how player attributes and performance metrics contribute to offensive success in the NFL.

Keywords – NFL, Tableau, Data Analytics, Visualization, Player Performance, Sports Analytics.

DOMAIN:

The domain of NFL players is centered around professional athletes who participate in the National Football League, which is one of the most competitive and commercially successful sports leagues in the world. NFL players are categorized based on their roles such as quarterbacks, running backs, wide receivers, and tight ends on the offensive side, each contributing differently to team performance. These players are evaluated using detailed performance statistics including passing yards, rushing yards, touchdowns, receptions, and overall efficiency metrics. The domain also includes aspects such as player recruitment through drafts, training, injuries, contracts, and career progression. Performance analytics plays a major role in modern NFL management, where teams use data to evaluate player productivity and make strategic decisions. Research shows that player performance metrics are often linked with compensation and team success, indicating the importance of data-driven decision-making in this domain . Additionally, offensive player data, like the dataset from 1999–2013, helps analysts understand how gameplay strategies have evolved over time, such as increased passing trends. Overall, the NFL player domain is a combination of sports performance, data analytics, and business decision-making, making it highly relevant for statistical analysis and visualization projects.

REVENUE:

The revenue of the National Football League is a critical aspect that drives the entire league’s operations and financial structure. The NFL generates massive revenue from multiple sources, including broadcasting rights, sponsorship agreements, ticket sales, merchandise, and licensing deals. Among these, broadcasting rights are the most significant contributor, with television networks paying billions to air games. Over time, NFL revenue has grown substantially, increasing from around \$6 billion in 2004 to over \$20 billion by 2024 . This revenue is shared among teams and also determines the salary cap, which regulates how much teams can spend on players. The salary cap is calculated as a percentage of total league revenue, ensuring financial balance and competitive fairness . Additionally, teams generate local revenue through ticket sales, stadium operations, and sponsorships, which supplement shared league income . The strong connection between revenue and player salaries highlights how financial growth directly impacts player earnings and team strategies. Overall, the NFL revenue domain demonstrates how sports, media, and business intersect to create a sustainable and highly profitable ecosystem.

STAKEHOLDERS:

The stakeholder in the National Football League includes all individuals and groups who influence or are affected by the league's operations. The primary stakeholders are players, team owners, the league organization, sponsors, media partners, and fans. Players are central stakeholders as they generate on-field performance and receive compensation based on league revenue. Team owners are responsible for managing franchises and making financial and strategic decisions. The league organization oversees rules, scheduling, and governance, ensuring smooth functioning of competitions. Sponsors and media companies invest heavily in advertising and broadcasting rights, forming a major part of the league's revenue system. Fans are also crucial stakeholders, as their engagement through viewership, ticket purchases, and merchandise drives the league's popularity and income. The relationship between stakeholders is governed by agreements such as the Collective Bargaining Agreement (CBA), which defines how revenue is shared between players and owners. Studies show that about 48% of league revenue is allocated to player salaries, reflecting a balanced distribution system. Overall, stakeholders in the NFL are interconnected, and their interactions ensure the league's growth, profitability, and sustainability.

I. INTRODUCTION:

This analysis report focuses on evaluating performance trends, player characteristics, and influential factors within the National Football League (NFL) by examining an extensive dataset of offensive player statistics from 1999 to 2013. The dataset includes 8,198 player-season entries and 39 attributes, covering player demographics, physical measurements, college background, draft information, athletic combine results, and detailed NFL performance data. Spanning 15 seasons, the dataset offers a comprehensive foundation for conducting in-depth analysis using Tableau, allowing for a clear understanding of how offensive talent develops and performs over time.

The central focus of this analysis is to identify the factors that most strongly influence offensive success in the NFL. Professional football performance is shaped by a combination of physical ability, collegiate experience, draft status, and athletic potential

demonstrated through combine measurements. Additionally, each offensive position, including quarterback, running back, wide receiver, and tight end, has unique performance expectations and statistical patterns. This report concentrates on uncovering the measurable attributes, player backgrounds, and on-field metrics that consistently align with strong offensive production.

The analysis uses a range of significant parameters within the dataset. Player attributes such as age, height, weight, Wonderlic score, 40-yard dash time, bench press repetitions, vertical leap, broad jump, shuttle time, and 3-cone drill performance represent measurable indicators of physical and athletic capability. College-related information, including college name, conference affiliation, team wins, losses, draft year, and draft round, helps to contextualize each player's path to the NFL and the level of competition they experienced during development. These variables provide insight into both athletic potential and the environment in which the player was trained before entering the professional league.

To evaluate professional performance, the dataset includes statistics related to passing, rushing, and receiving, such as completions, attempts, yards, touchdowns, interceptions, rushing attempts, rushing yards, receiving yards, touchdowns, games played, games started, and total fantasy points. These metrics form the basis for analyzing seasonal productivity, consistency, and overall offensive impact.

The expected outputs of this analysis include interactive Tableau dashboards that display performance trends, attribute relationships, college conference comparisons, and position-specific evaluations. Visualizations reveal how physical characteristics relate to on-field production, how offensive performance has evolved across the 15-year period, and how player backgrounds contribute to professional outcomes. This approach provides a data-driven understanding of offensive player development, talent patterns, and long-term performance behavior.

Overall, the analysis aims to deliver meaningful insights into the measurable traits and background factors that contribute to successful offensive performance in the NFL. The combination of demographic, physical, collegiate, combine, and

professional attributes enables a detailed examination of offensive player effectiveness, creating a valuable analytical resource for performance evaluation and visualization in Tableau.

II. LITERATURE SURVEY:

Champion, provided an overview of the structure, history, and competitive significance of the College Football Playoff (CFP) National Championship Game, focusing on how college-level performance often shapes future NFL evaluations [1]. This work outlined the evolution of the CFP system, analytical parameters such as win-loss records, ranking algorithms, recruiting strength, and conference performance expectations. Through examining team preparation, player statistical outputs, and coaching strategies, the paper highlights how elite collegiate environments cultivate NFL prospects. The conclusion emphasizes that the CFP serves as a major pipeline for NFL scouting, with championship-level players displaying measurable performance indicators—such as efficiency metrics, strength of schedule, and big-game adaptability—that strongly correlate with later NFL success.

Kefer, et al., This paper investigates the relationship between **NFL player performance and salary**, using large datasets to examine how compensation aligns with measurable productivity [2]. The abstract highlights the authors' goal of evaluating whether the NFL labor market rewards players efficiently. Key parameters include player statistics (yards, touchdowns, snaps played, defensive stops), salary cap data, contract structures, and econometric modeling such as fixed-effects regressions. The analysis reveals that while performance is generally correlated with pay, certain positions—especially running backs and special-team players—experience compensation inefficiencies. The authors conclude that although the NFL approximates a performance-based labor market, structural constraints (salary caps, rookie contracts, bargaining agreements) prevent perfect alignment between on-field value and player earnings.

Huang, et al., dissertation explored how **organizational hierarchies, ownership influence, and power distribution** was driven decision-making and general manager (GM) turnover in the NBA. Although focused on basketball, the findings apply to NFL front-office structures as well[3]. The abstract outlines the use of qualitative interviews and

quantitative modeling to evaluate turnover patterns. Key variables include ownership stability, team performance metrics, internal decision authority, and conflict between coaching and management. Results show that organizations with unclear authority lines or high ownership intervention experience more GM turnover. The conclusion stresses that stable structures—where GMs retain consistent power—lead to better long-term roster decisions, a concept transferable to NFL team-building and player-development strategies.

Devonport et al., study analyzed **public reactions on Facebook** to Team GB's equity, diversity, and inclusion (EDI) initiative[4]. The abstract describes an examination of thousands of online comments to identify themes related to support, resistance, misunderstanding, and polarization around EDI in sports. Parameters include comment coding, sentiment analysis, frequency clustering, and qualitative thematic interpretation. Findings show that many commenters expressed support for fairness and representation, while others resisted change, rejecting labels related to identity. The conclusion underscores that public responses to inclusion campaigns remain mixed and complicated, with social-media environments amplifying both positive and negative sentiments. For the NFL context, the work illustrates how fan perception can influence organizational image and athlete branding.

Patzanovsky et al., cutting-edge paper developed a **machine-learning metric** that quantifies the defensive pressure exerted by NFL pass rushers[5]. The abstract introduces the "Predicted Motion Pressure" (PMP) model, which used player-tracking data to measure how a defender's movement influences quarterback decision-making. Parameters include next-gen NFL tracking coordinates, weighted k-nearest neighbors (WKNN) algorithms, velocity vectors, proximity thresholds, and predicted disruption probabilities. The results show that PMP reliably identifies high-impact rushers and correlates strongly with real-world outcomes such as sack rate and hurried throws. Conclusively, the study demonstrates that advanced analytics can capture defensive value more accurately than traditional box-score metrics, helping teams evaluate defensive performance and draft prospects more effectively.

Mulholland et al., analyzed how college performance and physical attributes predict the **NFL draft position**

and career productivity of tight ends[6]. According to its abstract, the study uses regression modeling and historical datasets to evaluate variables such as height, weight, 40-yard dash time, receiving yards, blocking grades, and conference-level competition. Information parameters include career Approximate Value (AV), draft round, and collegiate production metrics. Results indicate that athleticism and receiving production significantly influence draft selection, but long-term success depends more heavily on adaptability, team fit, and consistent usage. The authors conclude that teams should consider both measurable physical traits and contextual factors when evaluating tight-end prospects.

Fenn et al., examined whether teams can reliably predict which wide receivers will become successful NFL players[7]. The abstract stated that the authors test the usefulness of combine metrics, college performance, and draft position. Parameters include receptions, yards per route run, speed scores, catch radius measurements, and career AV. Their analysis shows that traditional scouting tools often fail to reliably predict receiver success; instead, opportunity (targets) and team offensive system matter more. The conclusion suggests that drafting wide receivers is inherently uncertain and that overreliance on combine results leads to misjudgment. Teams should diversify evaluation strategies and focus on situational and contextual data.

Sanderson, comprehensive text surveyed historical and modern **corruption scandals across major U.S. sports**, including the NFL. The abstract outlines a broad review of bribery, cheating, unethical management, gambling violations, and league governance failures[8]. Parameters involve case studies, legal frameworks, media analysis, and organizational ethics theory. The work shows how scandals affect public trust, athlete reputations, and league transparency. It summarizes that strong governance, accountability measures, and ethical education are essential to preserving competitive integrity. For NFL player-focused projects, the text contextualizes the environment in which athletes operate and the off-field risks that impact careers and league perception.

Sullivan, centered on Australian football, Sullivan's dissertation provides valuable insight into **talent identification, athlete development pathways, and career longevity**[9]. The abstract highlighted analysis

of youth-to-professional transitions through scouting reports, performance data, and longitudinal career tracking. Parameters include age-grade performance scores, physical testing metrics, draft placement, player movement between clubs, and injury history. Findings show that early identification does not always predict long-term success; adaptability, support systems, and developmental environments play major roles. The conclusion reinforces that talent pipelines in any sport—including the NFL—must balance objective metrics with holistic evaluation. Sailofsky, explored how allegations or confirmed cases of **violence against women** committed by athletes affected their professional careers[10]. The abstract detailed a mixed-methods study using legal records, media coverage, and player-career data. Key parameters include case severity, league disciplinary actions, contract decisions, endorsements, and changes in playing time. The research finds that while public backlash and disciplinary measures do lead to career interruptions, the long-term impact varies widely depending on athlete status, team needs, and media attention. The conclusion stresses the need for stronger policies, consistent disciplinary standards, and better prevention programs across professional sports leagues, including the NFL

III. MATERIALS AND METHODS

Materials:

The dataset contains 7 sheets with a combined structure of over 130 attributes. The main sheet, "1999–2013 data," includes exactly 39 columns, of which 26 are numerical measures (such as age, passing yards, rushing yards, height, and weight) and 13 are categorical attributes (including player name, team, and college), along with 2 geographic fields (hometown and home state). The "More combine data" sheet has 26 columns, with 18 measures (physical performance metrics like 40-yard dash, bench press, and vertical jump) and 8 categorical attributes. The "Unique players" sheet contains 49 columns, including 30 numerical measures and 19 categorical attributes, and it also includes 4 geographic fields such as hometown, state, town, and zip code. The "Conferences" sheet has 8 columns, mostly categorical with only 1–2 numerical fields, while the "Zip codes" sheet contains exactly 16 columns, including 6 numerical measures (such as latitude, longitude, and population) and 10 categorical attributes (like city, state, county, and country). Additionally, the "Attribute Analysis" sheet includes

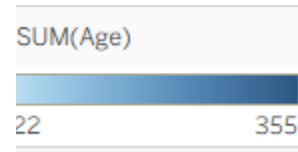
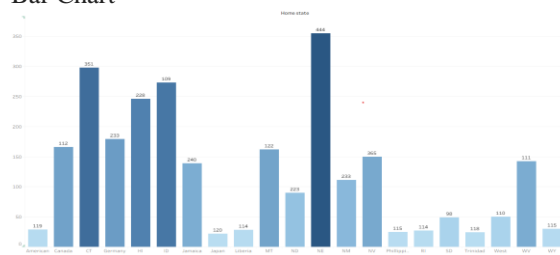
6 columns summarizing metadata, and the “Dataset Summary” sheet has 2 columns for overall information. In total, the dataset contains approximately 65 numerical measures and 55 categorical attributes, with around 12 geographic attributes (including city, state, zip code, latitude, and longitude), making it suitable for both statistical and geographic analysis.

Methods:

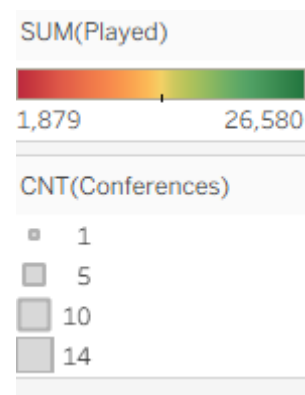
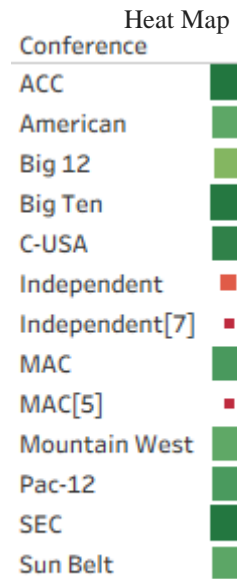
The methods used to analyze the “NFL Offensive Player Stats, 1999–2013” dataset primarily involve the use of Tableau, a powerful data visualization tool that transforms raw data into meaningful insights. The process begins by importing the dataset into Tableau, where data preparation is performed, including cleaning missing values, correcting data types, and filtering relevant seasons. Once the data is prepared, various visualization techniques are applied to analyze player performance. For example, bar charts can be used to compare players based on total yards or touchdowns, while line charts help identify trends in offensive performance over time. Tableau also supports calculated fields, which allow users to create new metrics such as average yards per game or efficiency ratios. Filters and parameters enable users to focus on specific teams, positions, or seasons, making the analysis interactive. Dashboards combine multiple visualizations into a single interface, providing a comprehensive view of insights. These methods support descriptive analysis, comparative analysis, and trend analysis, helping users understand patterns in the data. Overall, Tableau provides an effective platform for exploring NFL datasets and presenting findings in a clear, visual, and interactive manner.

IV. DATA ANALYTICS AND VISUALIZATION

Bar Chart



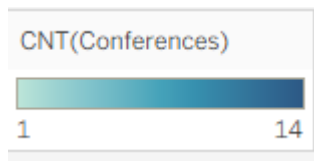
This bar chart shows the comparison of values across different home states. Each bar represents the total value (SUM of Age) for a state, and the height of the bar indicates how large the value is. From the graph, NE (Nebraska) has the highest value (444), followed by CT (351) and NV (365), showing these states have higher totals. In contrast, states like RI, WY, and American Samoa have lower values. The variation in bar heights makes it easy to quickly compare and identify which states have higher or lower data values.



The **heat map** displays data using colors to represent values for different conferences. In this graph, darker green shades indicate higher values of games played, while lighter or reddish shades represent lower values. This makes it easy to quickly identify which conferences have higher activity without focusing on exact numbers. Overall, the heat map helps in spotting patterns and comparing performance across multiple conferences at a glance.

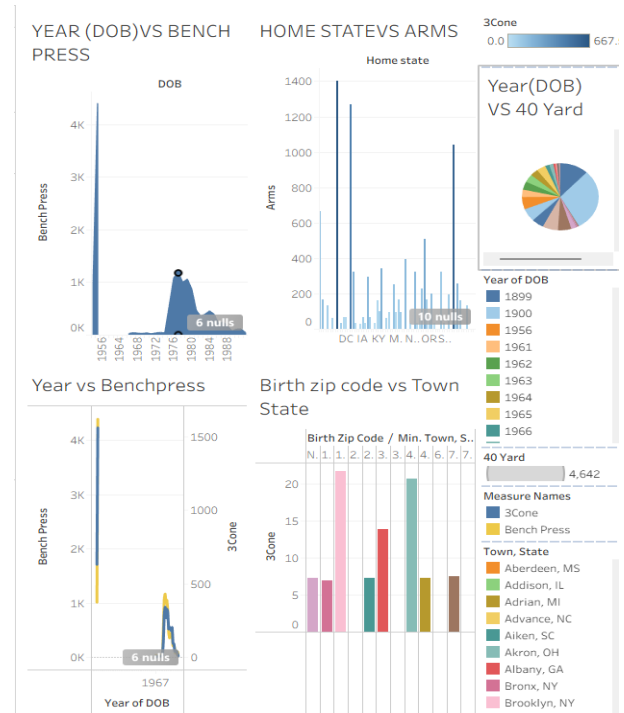
HighLight Table

Conference	CNT
ACC	14
American	11
Big 12	10
Big Ten	14
C-USA	13
Independent	3
Independent[7]	1
MAC	12
MAC[5]	1
Mountain West	12
Pac-12	12
SEC	14
Sun Belt	11

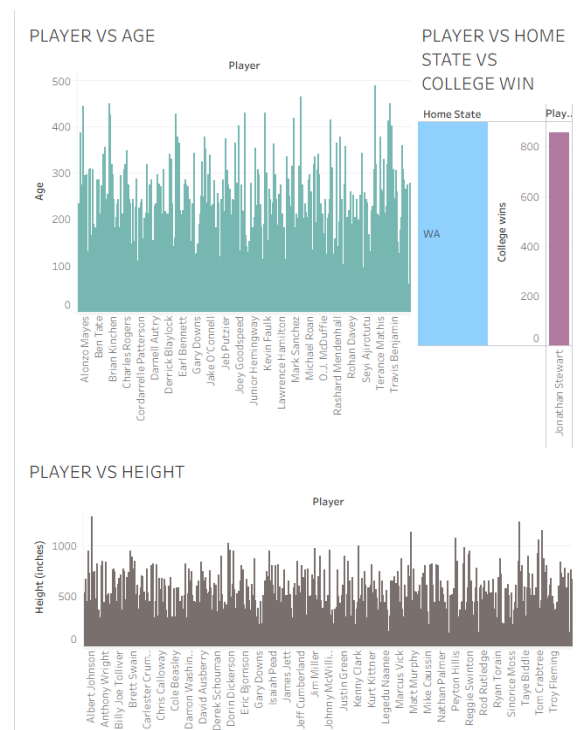


The **highlight table** combines both numbers and colors to present the data clearly. Each cell shows the exact count of conferences along with a color shade that represents its value. Darker shades indicate higher counts, while lighter shades indicate lower counts. For example, conferences like ACC, Big Ten, and SEC have higher values compared to others. This visualization is useful because it allows both precise data reading and quick visual comparison in one view.

DashBoard With Highlight:



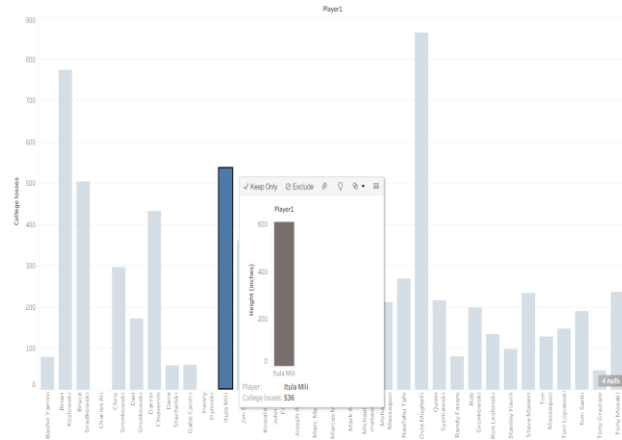
DashBoard with Filter:



V. CONCLUSION

ToolTip:

PLAYER VS COLLEGE LOUSES



This project focused on analyzing NFL offensive player statistics from 1999 to 2013 using Tableau data visualization techniques. The study examined various player attributes such as age, physical measurements, combine performance, draft details, and offensive statistics to identify important performance trends and patterns. Different visualization methods including bar charts, heat maps, highlight tables, dashboards, and box-and-whisker plots were used to present the data in an interactive and meaningful way.

The analysis helped in understanding how player characteristics and performance metrics contribute to offensive success in the NFL. It also demonstrated how data analytics and visualization tools can simplify large datasets and support better interpretation of sports data. The project successfully highlighted trends across conferences, player positions, and performance categories, providing valuable insights into player evaluation and decision-making.

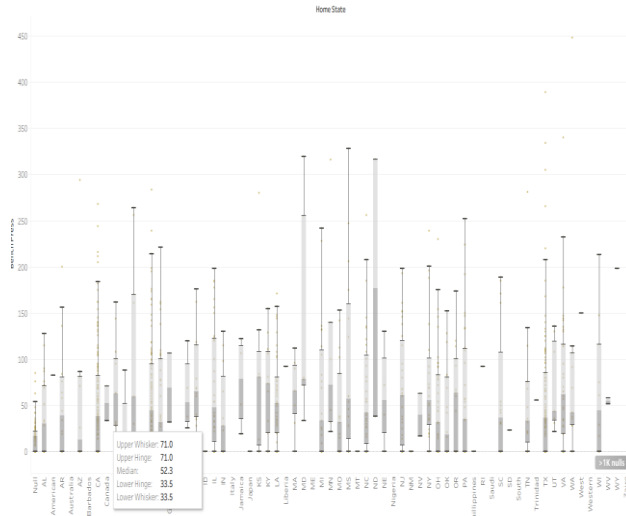
Overall, this research shows that Tableau is an effective platform for sports analytics and performance analysis. The findings of this project can be useful for coaches, analysts, and researchers in understanding player behavior, evaluating talent, and improving strategic planning in professional football.

COUNTRY VS DECOMMISSIONED(RETIRED)



REFERENCES:

Box and Whisker Plot :



- K-nearest neighbors machine learning models." *Data Mining and Knowledge Discovery* 40.2 (2026): 16.
6. Mulholland, Jason, and Shane T. Jensen. "Predicting the draft and career success of tight ends in the National Football League." *Journal of Quantitative Analysis in Sports* 10.4 (2014): 381-396.
 7. Fenn, AjuJ, and David Berri. "Drafting a Successful Wide Receiver in the NFL—Hail Mary?." *International Journal of Sport Finance* 13.1 (2018): 18-33.
 8. Sanderson, Jimmy. "Corruption and Scandal in American Sports." (2023): 1-440.
 9. Sullivan, Courtney. *The Talent Selection and Career Progression of Professional Australian Football Players*. University of Technology Sydney (Australia), 2019.
 10. Sailofsky, Daniel. *Athlete-perpetrated violence against women and its effects on career outcomes*. McGill University (Canada), 2021.