Soccer Report 18/19

Alejandro Gomez, Santiago Torralva, Kay Maletzki

Professor Perez

CCIS 546 – Data Visualization

St. Thomas University

05/05/2025

**Abstract**

Using a FIFA player dataset, this research investigates the variables affecting the performance and market value of professional football players. Our main research topics investigate whether a player's birth month affects their chances of becoming a professional, how player positions correspond to skill sets, and which elements have the strongest correlations with a player's market value. We produced interactive dashboards with heatmaps, scatter plots, and bar charts using Tableau Public for data visualization. Key findings showed that FIFA rating has a moderate, but nonlinear, correlation with value, players born in the first few months of life are slightly overrepresented in the professional ranks, and physical and technical traits differ significantly by position. Other investigations looked at performance across age groups, compared wages to market worth, and investigated national variations in player evaluations. Sports management experts, youth academies, and scouting departments can all benefit from this information by better comprehending player development and valuation trends. We conclude that although certain characteristics (such position-specific skills) exhibit distinct patterns, others—including wage fairness or nationality—show greater variance and deserve deeper exploration.

**Introduction**

Player development and football scouting are multimillion-dollar businesses. For both club administration and youth development initiatives, it is essential to comprehend what makes a player successful or valued. This initiative investigates a number of important questions:

* *What positional differences exist in player skills?*
* *Does the possibility of becoming a professional depend on one's birth month?*
* *What elements affect salary disparities and market value?*
* *What is the relationship between performance ratings, nationality, and age?*

Our objective is to identify trends in skill development, valuation, and demographic distribution by utilizing data from FIFA ratings, which are a commonly used indicator of player performance. When assessing athletes, teams, agents, and analysts may find these results useful in making better decisions.

**Methodology**

Dataset

We used the FIFA player dataset sourced from Kaggle. It includes data on over 17,000 players, detailing attributes such as position, skill ratings (e.g., dribbling, passing), age, nationality, overall and potential ratings, market value, wage, and birth month.

* **Source:** FIFA 19 Player Dataset on Kaggle

Cleaning & Preprocessing

* Removed null or extreme outlier values (e.g., zero market value or wage).
* Converted categorical position data into grouped categories (Defender, Midfielder, Forward, Goalkeeper) for clarity.
* Pivoted skill rating columns to analyze position vs. abilities heatmap.
* Extracted birth months and grouped player ages to enable demographic analysis.

Tool Used

We used **Tableau Public** due to its intuitive interface and strong visualization capabilities. Tableau allows for interactive dashboards and seamless filtering, ideal for exploring multiple variables simultaneously. The pivoting and calculated field functions made it well-suited for our skill-position heatmap and demographic analyses.

**Analysis & Results**

1. Heatmap - Position vs. Abilities

In this visualization, we will study the abilities by position. After pivoting on the various abilities we considered most decisive for a football player, we can select other options and group them based on our needs. Following the pivoting process, we can categorize our abilities and examine their values across different grouped positions. We created a calculated field for this categorical variable since there are too many positions, and this approach enhances clarity. By centering the color range, we can observe different positions and skills, and analyze areas for improvement in each case. The color range was centered around the average, making it easier to spot above- or below-average traits for each position group. Defenders excel in strength, aggression, interceptions, and marking, while midfielders and attackers score higher in sprint speed, positioning, dribbling, ball control, and balance.

A screenshot of a computer

Description automatically generated

2. Line Chart - Birth Month vs. Player Count (Relative Age Effect)

The next two visualizations are gonna study whether a professional soccer player’s birth month is important,

But in this first visualization, we focus on the quantity of players born in each month and we can see a fairly similar repeated pattern regardless of the position group they are included in, as well as their nationality or age groups, which include under 23, prime, and senior. In this case, we can see that the approximation of the trend line p-value is approaching 0, which is a good signal. Therefore, we can be sure that there is a relative age effect, where the month has a significant impact. For example, younger children born earlier in the year may gain early advantages in youth academies.

A screenshot of a computer

Description automatically generated

3. Line Chart - Birth Month vs. Market Value

In this view, we are studying the potential value of the market relationship based on the month of birth, using various filters. Surprisingly, market value showed no consistent trend across birth months. This suggests that while early birth months may increase **chances** of going pro, they don’t directly influence a player’s **value** once professional.A screen shot of a graph

Description automatically generated

*4. Line Chart -* FIFA Rating vs. Market Value

In this visualization, we are examining the relationship between value and rating, categorized by player position. We have grouped the different positions since we are working with averages to avoid potential outliers due to unequal counts per position. It is essential to highlight the patterns that emerge and the significance of player positions in this analysis. The correlation was visible but non-linear—some highly rated players had moderate values, while others were overvalued due to market dynamics, fame, or potential.

A screenshot of a graph

AI-generated content may be incorrect.

5. Scatter Plot -Age vs. Market Value (by positions)

This graphic, which is broken down by position groups, looks at how a football player's market value varies with age. The graph reveals how market views differ between roles on the field by charting average value against age and using color to identify job groups. A trend line, which shows the overall course of value development over time, provides clarification. This kind of analysis can assist teams in making strategic decisions about transfers and squad building by highlighting the times when players in certain positions are typically highly valued. Using average wage for size and position groups for color, we find that while goalkeepers, defenders, and midfielder typically reach their prime in their mid-to-late twenties. Strikers on the other hand hold their value longer, being in their prime with around 31 years. Clubs can use this view to find undervalued talent based on role, potential, and salary as well as optimum age ranges by position.

A screenshot of a graph

Description automatically generated

6. Bar Chart - Nationality Breakdown – Top 20 Market Value (by positions)

By looking at the average overall rating (or market value) of football players from each nation, this graphic investigates how talent is distributed among nationalities. We restricted the research to the top 20 nations by player representation in order to maintain a clear and understandable display. Based on FIFA ratings, this chart assists in determining which nations are continuously creating players of a greater caliber. On the other hand, by employing average market value rather than rating, we may reveal how the global market views talent from different geographical areas, emphasizing variations in reputation, visibility, or development systems among countries. A bar chart comparing average player ratings by nationality showed expected dominance by countries like Brazil, Germany, and France. However, some nations with fewer players had surprisingly high averages—possibly due to a few elite stars boosting the mean. A filter was used to show only countries with a minimum player count to ensure fairness.

A screenshot of a graph

Description automatically generated

**7.** Scatter Plot - Wage vs. Market Value by Position

We examine the connection between a football player's salary and market worth for various positions in this graphic. By graphing wage on the x-axis and market value on the y-axis, we hope to uncover potential outliers—players who may be overpaid or underpaid relative to their valuation. We eliminated players who had zero or missing data in either field to prevent clutter. To see broad patterns throughout the field, we grouped the players by position. To show player quality, we may also choose to size the bubbles according to the players' overall rating. This gives us information about club spending plans and player negotiations by enabling us to determine whether particular positions are more vulnerable to salary inflation or undervaluation. We observed certain discrepancies: goalkeepers and defenders were often undervalued in wage compared to value, while some attacking players appeared overpaid. This suggests wage does not always reflect market value accurately.

**A screenshot of a computer

Description automatically generated**

**Presentation**

https://www.canva.com/design/DAGmibnMvrA/cKfNX7HNvgJ5Kmd8rdkiUg/edit?utm\_content=DAGmibnMvrA&utm\_campaign=designshare&utm\_medium=link2&utm\_source=sharebutton

**Thank you!**