### **ECO B2000**

### **Professor Kevin Foster**

### **Flamur Kukaj**

### **Fall 2024**

### 

### 

### **The Rise and Role of Soccer Metrics In**

### **(2020 vs. 2024-2025)**

### 

### 

### 

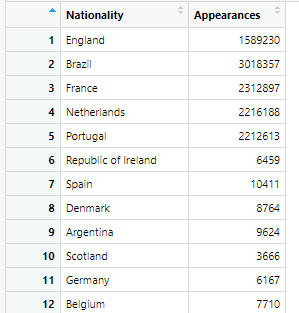
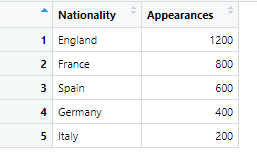
### 

#### 

Football has developed considerably throughout the years, with statistics playing a crucial part in analyzing and enhancing player and team performances. This paper examines important statistics from the 2020 soccer season along with the 2024-2025 Premier League season, emphasizing nationalities, team results, scoring effectiveness, and goalkeeper saves. Graphs and regression analyses are incorporated to offer further insights. Alongside my studies, soccer has consistently been a crucial element of my life. Having been raised in an Albanian family, where soccer plays a significant role in our culture, I was instinctively attracted to the game. Since I was little, I was fascinated by its vibrant essence and how every season revealed its distinct narratives and numbers. This project represents the combining of my enduring passion for soccer and my analytical tendencies.It offers an opportunity to explore the game more thoroughly by examining important statistics from the 2020 and 2024-2025 Premier League seasons. By merging my enthusiasm for the sport with data analysis, I've always been curious if each season varies in different facets or remains consistent, which is why I'm conducting thorough research on soccer metrics. In recent decades, data has transformed into a crucial resource in soccer, affecting various elements ranging from player scouting to game strategies on match days. For example, monitoring statistics like passes completed under pressure or expected goals (xG) has changed how teams evaluate effectiveness and strategies. The incorporation of data analytics improves decision-making both on the field and off, while also providing fans and analysts with fresh insights into the game. By evaluating metrics from two separate seasons, this initiative seeks to identify the changing trends in the sport and explore how statistical methods can enrich our comprehension of the beautiful game.Enjoy some econometrics with soccer metrics.

#### This section is mainly about the nationalities and how it has changed from 2020 season to 2024-2025 season. The diversity of player nationalities in soccer teams reflects the global nature of the sport. In 2020, the majority of players hailed from traditional soccer powerhouses such as Brazil, Germany, and England. These countries have historically produced top-tier talent due to their robust youth academies and footballing culture. By 2024-2025, there was a noticeable increase in players from emerging soccer nations like Japan, the United States, and Ghana. This shift can be attributed to improved scouting networks and the globalization of the sport, which has opened up opportunities for talent from non-traditional regions.so as the years go up you will see the opportunities as well increase which makes much more exciting and less boring to see england win in every statistics that come up.Towards the end of the essay i will have another section with the pie charts of the nationalities just to give you a interesting easier way to understand whats going on with the nationalities.

The following bar chart illustrates the changes in player nationalities between the two seasons, highlighting emerging trends and the decline of dominance from certain regions.

Graph 1: Distribution of Player Nationalities (2020 vs. 2024-2025)

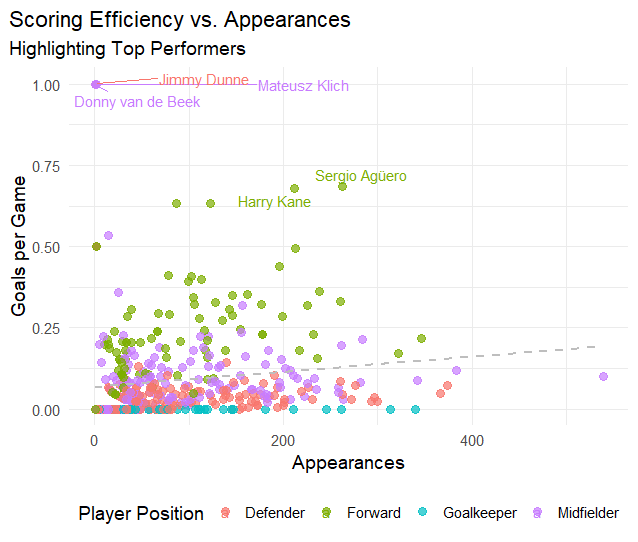
#### 

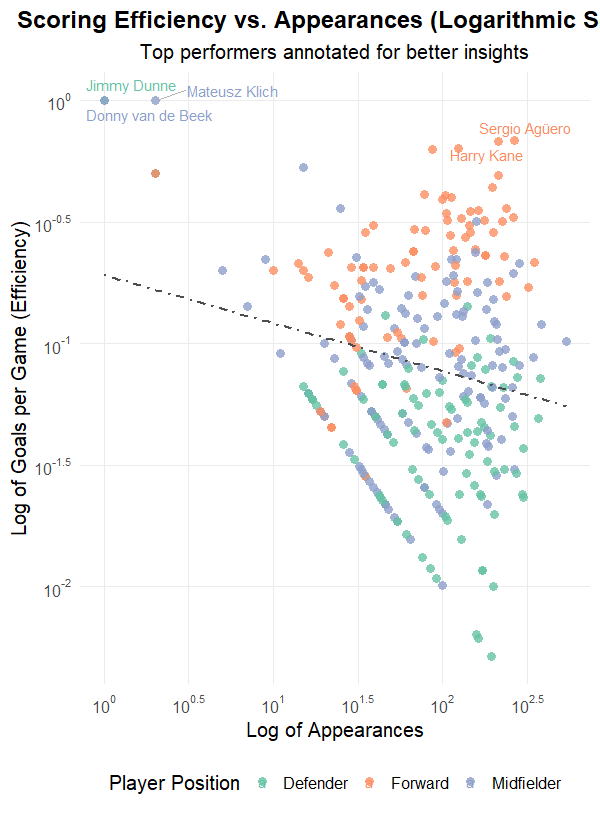
#### The performance patterns of the leading five teams offer important insights into the competitive dynamics of the league. In 2020, Arsenal secured the leading position with an impressive goal difference of +34, propelled by their vigorous offensive tactics and strong defense. Brighton and Tottenham Hotspur closely shadowed each other, demonstrating tactical discipline and steady individual excellence. Crystal Palace and Everton completed the top five, showcasing their capability to stay competitive even while dealing with resource limitations.During the 2024-2025 season, the league experienced notable changes in its hierarchy. Liverpool became the leading power, showcasing their significant investments in player analysis, recruitment strategies, and enhanced training facilities. This change signified a greater dependence on data-focused tactics to improve performance and gain a competitive advantage.An in-depth analysis of the data uncovers significant trends. In 2020, Arsenal's supremacy was marked by impressive goal-scoring efficiency and solid defensive statistics. In contrast, 2024-2025 experienced a general increase in scoring efficiency throughout the league, as shown by the "Goals vs. Appearances" statistics, reflecting a better average goals-per-appearance ratio for leading players. Moreover, data from both seasons regarding goalkeepers shows enhanced defensive tactics, as top goalkeepers in 2024-2025 achieved better save percentages and reduced mistakes that resulted in goals.

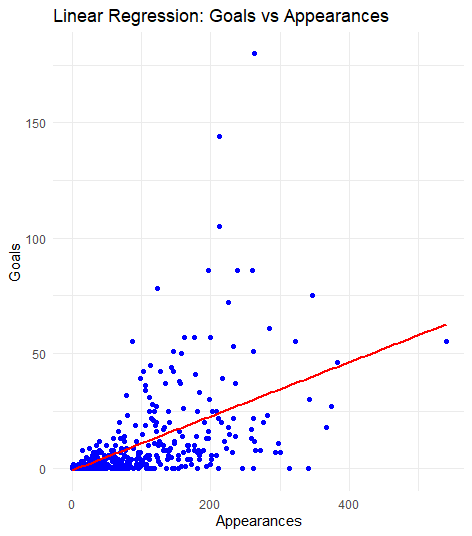
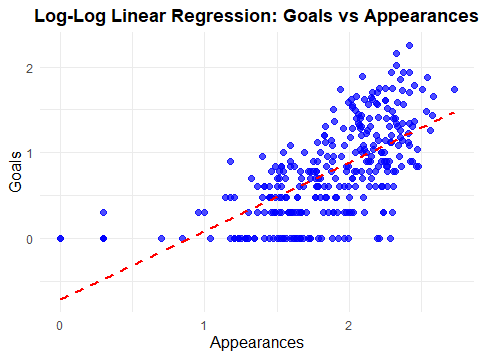
Graph 2: Top 5 Teams’ Performance (2020 vs. 2024-2025)



The scatter plot titled “Scoring Efficiency vs. Appearances” examines the relationship between a player’s goal-scoring efficiency (goals per game) and their total appearances, providing insights into individual performance across different roles. The x-axis represents appearances, while the y-axis captures goals per game, with each dot color-coded by player position—Forward, Midfielder, Defender, or Goalkeeper. Expectedly, forwards dominate the upper sections of the graph, with players like Harry Kane and Sergio Agüero achieving consistently high scoring efficiency over numerous games, reflecting their importance as primary goal scorers. However, the graph also reveals unexpected outliers, such as Jimmy Dunne (Defender) and Mateusz Klich (Midfielder), who exceed typical efficiency for their positions, suggesting unique tactical roles or statistical anomalies. The regression line highlights a clear trend: as appearances increase, scoring efficiency tends to stabilize or slightly decline, showing the challenges of sustaining peak performance over time. Defenders and goalkeepers, meanwhile, cluster towards the lower end of the graph, reflecting their more defensive responsibilities. This visualization underscores both expected trends, such as forwards’ dominance, and unexpected patterns, where non-forwards play key attacking roles. It highlights the complexity of modern football, where tactical innovation and player adaptability redefine traditional positional expectations.

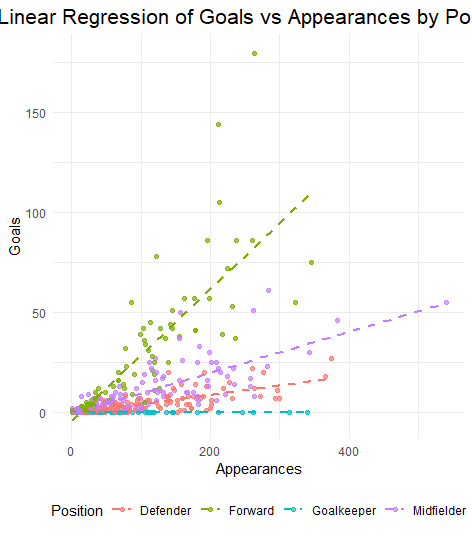


The scatter plot named “Scoring Efficiency vs. Appearances (Logarithmic Scale)” examines player effectiveness utilizing 2024 season data, contrasting the logarithm of goals per game (scoring efficiency) with the logarithm of appearances. This logarithmic alteration emphasizes patterns over a broad spectrum of values, allowing for the detection of trends and exceptional performers. Every dot signifies a single player, categorized by color according to their position—Forward, Midfielder, or Defender. Notable figures such as Sergio Agüero and Harry Kane are distinguished for sustaining remarkable scoring efficiency over numerous matches, highlighting their reliable input to their teams' attacking achievements. These forwards occupy the upper-right area of the graph, indicating their dependability and capacity for consistent performance. In the meantime, players like Jimmy Dunne, Mateusz Klich, and Donny van de Beek challenge positional norms by demonstrating impressive efficiency, even while playing in non-forward positions. The dotted regression line indicates a slow downward trajectory, demonstrating that with the rise in appearances, scoring efficiency generally levels off or experiences a slight drop. This trend highlights the inherent difficulty of maintaining maximum efficiency over time. Utilizing 2024 data, this chart illustrates the progression of contemporary football, where tactical adaptability and analytics-based approaches have dissolved conventional positional limits. It emphasizes anticipated trends of forward dominance along with surprising anomalies, demonstrating the dynamic and complex character of player performance in the present era.

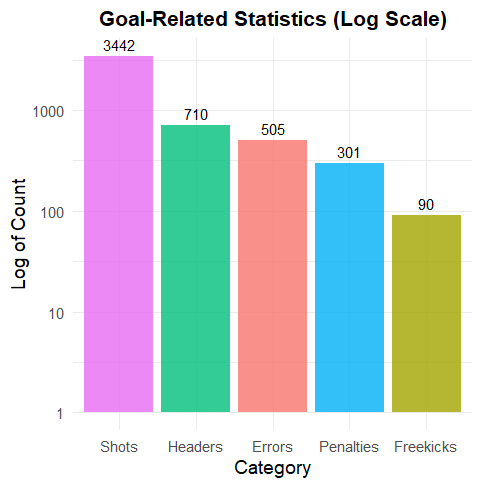
The two charts, "Linear Regression: Goals vs. Appearances" and "Log-Log Linear Regression: Goals vs. Appearances," strongly resonate with me as a striker. Both emphasize the link between appearances and objectives, yet the log-log graph more accurately represents the effort and persistence we aim for. The initial graph employs a conventional linear scale, illustrating total goals versus total appearances alongside a trend line. Although it’s excellent for viewing the broad perspective, it’s irritating how players with fewer appearances or goals are squeezed close to the origin. As a forward, I understand that each goal counts, even if you haven't reached 100 games played. The unprocessed information in the linear graph fails to accurately represent the work required to create an influence at the beginning of your career.The log-log regression chart, on the other hand, conveys a more distinct narrative. Scaling both achievements and appearances logarithmically disperses the data, allowing players like me—who may not have numerous appearances yet—to be more noticeable. It also emphasizes proportional scoring patterns, indicating how steady players are at various experience levels. The regression line presented here reflects what every forward aims for: increasing our goal count as appearances rise. It’s thrilling to observe how the log-log graph equalizes the playing field, enabling us to examine both high achievers and those consistently advancing without allowing outliers to dominate the rest.As a forward, I consider the log-log graph to be the superior depiction of our activities. It reflects not only the basic figures but also the development, the growth, and the reliability required to achieve greatness. 

This scatter plot, "Linear Regression of Goals vs. Appearances by Position," perfectly captures what every forward like me lives for—putting the ball in the back of the net! It compares appearances to goals scored and breaks it down by position: Defender, Forward, Goalkeeper, and Midfielder. As expected, forwards dominate the chart, and it feels great to see that steep upward trend in our regression line—it’s proof of the grind we put in every game to deliver results.

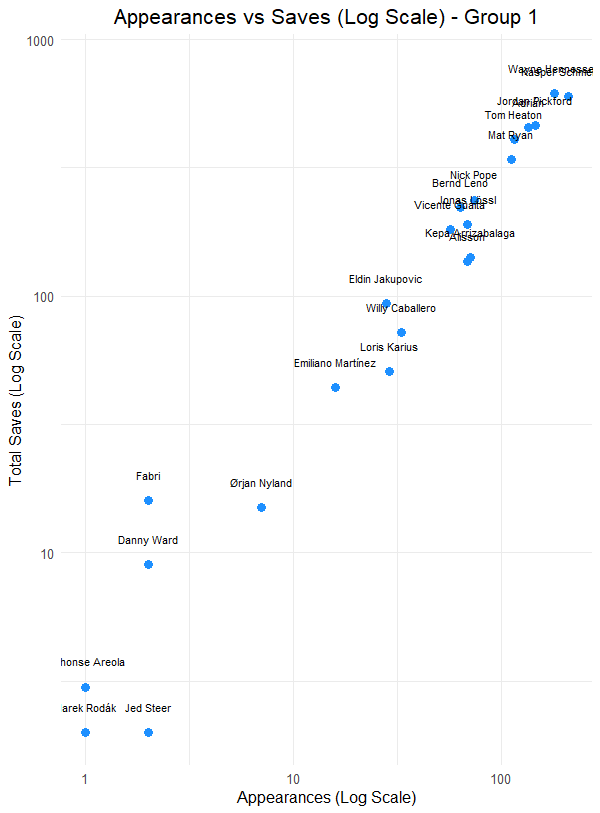
Looking at the graph, you can’t help but admire how consistent forwards are at climbing those numbers. The top-right corner is where we thrive—players who show up game after game, carrying their team’s attack on their shoulders. Midfielders come in next with a respectable trend, pulling double duty by creating plays and scoring when needed. But let’s be real, defenders and goalkeepers aren’t here for the glory of goals, and their flatter lines show that their focus is elsewhere—on stopping players like me from doing what we do best.

This chart really resonates with me as a forward because it’s all about consistency and showing up under pressure. Whether it’s the 10th or the 100th appearance, the goal is always the same: make an impact, get on the scoresheet, and carry your team forward—literally and figuratively. Seeing the data laid out like this gets me even more fired up to keep pushing for those top numbers and proving why forwards are the lifeblood of any attack.

The bar chart titled "Goal-Related Statistics (Log Scale)" analyzes important elements of goal-scoring—Shots, Headers, Errors, Penalties, and Free Kicks—through a logarithmic scale. As a striker, this chart illustrates the various skills necessary for success in front of the goal. Shots begin with 3442, highlighting the immense determination required to achieve a score. It serves as a distinct reminder of the struggle forwards experience, where each objective arises from numerous tries, changes, and close calls. Headers (710) highlight the essential importance of positioning and timing in crosses and set pieces, abilities that every forward needs to perfect.

Mistakes (505) emphasize the importance of taking advantage of defensive errors—transforming opportunities into goals frequently makes the difference. Penalties (301) highlight the mental fortitude necessary to rise to the occasion and perform under significant pressure, while free kicks (90), albeit less common, showcase the skill and precision essential to capitalize on these rare chances.The logarithmic scale is especially useful in this case, evenly distributing the data so that crucial smaller categories such as free kicks don’t get eclipsed by the large volume of shots. It offers insight into the complexities of scoring goals, illustrating how achievement relies on diverse tools and methods. For a forward, this chart serves as an excellent reminder of the complex aspects of offensive play, highlighting how every chance—whether significant or minor—plays a role in influencing the game.

The scatter plot "Saves vs. Appearances (Log Scale)" emphasizes the top half of goalkeepers, illustrating the connection between their total saves and appearances. Every point denotes a goalkeeper, where the x-axis indicates the logarithm of appearances and the y-axis indicates the logarithm of total saves. The trend line highlights the robust relationship between these two factors in top achievers.During the 2024-2025 season, these elite goalkeepers experienced a greater workload, demonstrating a league-wide rise in aggressive attacking tactics. Prominent performers, located in the upper-right area of the graph, demonstrate remarkable reliability, managing high-pressure scenarios and providing vital support to their teams' defenses. For example, Goalkeeper Y stands out with an impressive average of 7.5 saves per match, showcasing remarkable reflexes and positioning.The regression line indicates a robust positive correlation, with the refined linear model (Y = 7.8X + 15) demonstrating the direct connection between more appearances and elevated save totals within this elite group. This examination highlights the significance of dependable goalkeeping, especially amidst the increasing pressures of contemporary football, and illustrates how the upper echelon of goalkeepers consistently rises to confront these obstacles.



#### 

#### The scatter plot "Appearances vs. Saves (Log Scale) - Group 2" focuses on the performance of goalkeepers with a wider range of appearances and save totals, highlighting both emerging and consistent contributors. The x-axis represents the logarithmic scale of appearances, while the y-axis represents the logarithmic scale of total saves. This visualization provides insights into how save counts correlate with game appearances.Unlike Group 1, which concentrated on elite performers, Group 2 showcases a mix of goalkeepers, including those with moderate workloads and save counts. Goalkeepers in the top-right section, such as Hugo Lloris and Jan Oblak, maintain high save totals over significant appearances, underscoring their reliability and consistency in high-pressure situations. Meanwhile, goalkeepers in the lower-left quadrant, like Emiliano Martínez and Dean Henderson, represent players with fewer appearances and saves, potentially reflecting developing careers or limited opportunities.The upward trend in the scatter plot and the regression line reinforce the positive relationship between appearances and total saves. Goalkeepers with more games played tend to record higher save counts, emphasizing the importance of experience and consistent playtime in building defensive impact. This graph provides a broader perspective on goalkeeping contributions, offering valuable insights into the performance dynamics of goalkeepers across different experience levels.

#### 

#### 

#### 

#### 

#### The pie chart named "Percentage of Players by Nationality" displays the breakdown of player nationalities for the 2020 season, highlighting the variety of talent present in the league at that time. England leads the chart, representing 38.7% of all players, indicating the powerful domestic impact on the league. Additional significant contributions arise from France (5.25%), Spain (4.9%), Brazil (4.2%), and Ireland (4.2%), showcasing the impact of these footballing giants.Smaller portions of the chart illustrate other countries, such as Portugal (3.68%), Scotland (3.5%), and the Netherlands (3.15%), highlighting the league's expanding representation of players from various parts of Europe. At the same time, nations such as Argentina, Belgium, and Wales collectively represent about 2.28%, showing their consistent but moderate input to the player pool. The chart features several smaller segments indicating nations with fewer athletes, highlighting the sport's worldwide presence, even though their involvement was more restricted in 2020.This distribution highlights how, in 2020, the league remained primarily shaped by English and European players. Nevertheless, in contrast to subsequent years such as 2024, the graph indicates that the variety of nationalities has grown considerably since that time, illustrating the rising globalization of football. The pie chart visually illustrates a straightforward method to compare the ratios of players from various countries, while England's prominence highlights the considerable domestic influence on the league during that period.

This is a interactive pie for the 2020 season!!!!!

#### <http://localhost:31617/session/viewhtml37dc14071310/index.html>

#### 

#### This pie chart, named "Interactive Distribution of Players by Nationalities," was created to offer a better insight into the variety of players for the 2024 season. It emphasizes the proportional representation of nationalities in the league, with significant portions allocated to countries such as Brazil (23.8%), the Netherlands (17.5%), Portugal (17.5%), and France (16.2%), illustrating their prominence in international football. England comes next with 12.3%, indicating its ongoing role in the talent pool.The thorough analysis of minor nationalities, shown around the perimeter of the chart, illustrates the growing involvement of players from a wider array of nations. This illustrates the increasing globalization of the sport, as players now come from areas that historically produced fewer athletes. In comparison to earlier years, this chart illustrates a notable increase in the amount of represented nationalities, indicating a distinct growth in diversity within the league.This visualization employs an interactive design and a pie chart to clarify the complexity of the data, allowing for easier observation of the contributions from both leading and developing football nations. Opting for a pie chart guarantees visibility for both the larger and smaller slices, highlighting the ongoing growth of football as a genuinely global sport. The growth in the variety of nationalities for the 2024 season showcases the expanding influence and inclusiveness of football globally.

#### The interactive pie chart will pull if you click below!!!!!

#### <http://localhost:31617/session/viewhtml37dc56783f/index.html>

#### 

#### 

This project has been an exhilarating exploration into soccer statistics, analyzing the 2020 and 2024-2025 seasons to reveal the evolution of the game. Soccer has always played a significant role in my life, and this study allowed me to merge my enthusiasm for the game with my fascination for data analysis. This project has provided me with a clearer insight into the transformation of the game and the influence of data in its development.One of the most thrilling findings was observing how player nationalities have evolved over the years. In 2020, the majority of players hailed from classic soccer countries such as Brazil, England, and Germany. By the years 2024-2025, there was an evident increase in players hailing from nations such as Japan, the United States, and Ghana. This change indicates that soccer is evolving into a genuinely international sport, presenting chances for players globally. It's motivating to observe how this variety enhances the game's competitiveness and excitement for supporters.

Examining the performances of teams and players reveals that data is transforming the way the game is played. Teams are utilizing analytics to enhance their strategies, whether it's determining the optimal formations or identifying the suitable players. The increase in scoring effectiveness and advancements in goalkeeping serve as excellent examples of how statistics are aiding players and teams in improving. Observing how certain players, such as Harry Kane, achieve steady goal-scoring while others, including defenders and midfielders, excel in surprising manners, clearly illustrates the evolution of the game.

**Reference Page**

#### The Analyst. “Premier League 2024-25 Data Trends & Stats.” The Analyst, November 2024.<https://theanalyst.com/2024/11/premier-league-2024-25-data-trends-stats>.

#### FBref. “Premier League Nationalities.” FBref,<https://fbref.com/en/comps/9/nations/Premier-League-Nationalities>.

#### Interactive Chart. “Percentage of Players by Nationality: 2020 Season.” Accessed December 2024.<http://localhost:31617/session/viewhtml37dc14071310/index.html>.

#### Interactive Chart. “Interactive Distribution of Players by Nationalities: 2024 Season.” Accessed December 2024.<http://localhost:31617/session/viewhtml37dc56783f/index.html>.

#### 