Vignette

# Problems my app is designed to solve and by whom

In the world of U.S. professional soccer, scouting a player and determining how much a player *should* be paid isn’t always straightforward. Salaries can vary wildly—even among players with similar stats—because of factors like fan appeal, media attention, or team preferences. For clubs, agents, and fans who want a more data-driven way to estimate a player's worth, there hasn’t been an easy tool to explore those relationships. That’s where this app comes in.

This app is designed to help users—particularly sports analysts, soccer fans, and team management—explore and model how different performance stats (like goals, assists, and minutes played) relate to a player’s salary. With the ability to interactively filter data, view trends, and even build simple statistical models, users can get insights into which metrics matter most when it comes to how much an offensive player is paid. Whether someone is trying to evaluate the fairness of a player's current contract or just wants to explore soccer salary data in a fun, hands-on way, this app provides a straightforward, visual way to dig in.

# Packages used by the app

This app uses the following packages:

library(shiny)

library(bslib)

library(DT)

library(ggplot2)

library(dplyr)

library(tidyverse)

library(scales)

library(itscalledsoccer)

# Data

The app uses data from the itscalledsoccer R package as its default source, which compiles detailed information on U.S. professional soccer players. The dataset is structured as a clean data frame with each row representing an individual player and includes variables like base\_salary, age, minutes\_played, goals, assists, shots\_on\_target, and key\_passes. These features allow users to explore offensive player performance and salary trends. Additional fields like player\_name and team\_name help users search for specific individuals and view their current earnings. The dataset includes several hundred observations and is well-suited for both statistical modeling and interactive exploration within the app.

# EDA Inputs, Controls, and Outputs

To begin exploring the data, a user might start by selecting the "Offensive Players" tab, where they can interactively filter players by age, minutes played, or offensive stats like goals, assists, and shots on target. Suppose a user wants to analyze how age and goals relate to salary—they can select these variables from dropdown menus and instantly view a scatterplot with a regression line to identify patterns based on the variable they choose. The app also outputs a summary of the linear model used to predict salary. The user can search for a specific player to see an estimated salary prediction based on the model and compare it with the player's actual reported salary. This helps users—whether fans, analysts, or fantasy team managers—identify undervalued or overpaid players and draw insights about what performance metrics drive pay and what player could potentially fit in their team. The search bar is the feature with a statistical analysis behind the scenes. Its simplified in the user interface. Simply search the player of interest and their actual salary is shown as well as their predictive salary. A coach might use this to scout a player and realize the player is undervalued and offer that player a contract to join his team.

# Litterature Review

He, M., Cashucho, R., & Knobbe, A. (2015). *Football player’s performance and market value*. Proceedings of the Workshop on Machine Learning and Data Mining for Sports Analytics (MLSA). <https://dtai.cs.kuleuven.be/events/MLSA15/papers/mlsa15_submission_8.pdf>

Lee, S., & Harris, J. (2012). Managing excellence in USA Major League Soccer: An analysis of the relationship between player performance and salary. *European Sport Management Quarterly, 12*(2), 149–169. https://doi.org/10.1080/13606719.2012.674389

Li, C., Kampakis, S., & Treleaven, P. (2022). *Machine learning modeling to evaluate the value of football players*. arXiv. <https://arxiv.org/pdf/2207.11361>

Several studies have explored the relationship between soccer player performance and compensation, offering valuable context for the app’s purpose. He, Cashucho, and Knobbe (2015) built a model to evaluate player market values in Spain’s La Liga, focusing specifically on forwards. They highlighted bias in data collection across different positions, reinforcing the app’s decision to emphasize offensive players. Lee and Harris (2012) analyzed salary distribution and performance in Major League Soccer (MLS), introducing the concept of the “star player effect,” where a few high-profile players earn disproportionately large salaries. This inequality underscores the need for tools like this app that can estimate salary more objectively based on measurable performance metrics. Li, Kampakis, and Treleaven (2022) applied machine learning to assess which features most strongly influence player value, noting that factors beyond the field—like fan popularity—can impact salary. Collectively, these studies support the app’s use of performance-based models while recognizing the complexity of salary prediction in professional soccer.

# Contributions/ Collaboration Assessment

Daniel completed the entire project independently.