

Detail Of Project

Introduction

In the world of football, strikers play a pivotal role in deciding the fate of matches and championships. Identifying the best strikers among a pool of talent involves a comprehensive analysis of various factors ranging from performance metrics to personal attributes. In this project, titled "Segmenting and Classifying the Best Strikers," we delve into a dataset containing information on 500 strikers, aiming to uncover patterns, insights, and classifications that distinguish top-performing strikers from the rest.

Project Description

The project involves utilizing data analytics techniques to explore and understand the characteristics and performance metrics of strikers. By employing descriptive statistics, data visualization, feature engineering, and machine learning algorithms, we aim to identify the key attributes that contribute to a striker's success on the field and classify them into different categories based on their performance.

Purpose

The primary purpose of this project is to provide a systematic framework for analyzing and categorizing strikers based on their performance metrics and personal attributes. By doing so, coaches, scouts, and football analysts can gain valuable insights into the characteristics of top-performing strikers and make informed decisions in team selection, recruitment, and strategic planning.

Dataset Description

The dataset comprises various variables related to 500 strikers, encompassing both demographic information and performance metrics. Key variables include nationality, footedness, marital status, goals scored, assists, shot accuracy, dribbling success, and many more, providing a comprehensive overview of each striker's profile and on-field performance.

- **Striker ID:** Unique identifiers assigned to each striker.
- **Nationality:** The country of origin for each striker.
- **Footedness:** Indicates whether the striker is right or left-footed.
- **Marital Status:** Indicates whether the striker is married (yes) or unmarried (no).
- **Goals Scored:** The total number of goals scored by the striker, a fundamental performance metric.
- **Assists:** The number of assists provided by the striker, indicating their ability to create goal-scoring opportunities for teammates.

- **Shots on Target:** The number of shots taken by the striker that hit the target, reflecting their ability to create scoring opportunities and test the goalkeeper.
- **Shot Accuracy:** The percentage of shots on target out of total shots taken, showing the striker's precision and effectiveness.
- **Conversion Rate:** The percentage of shots that result in goals, revealing the striker's efficiency in front of goal.
- **Dribbling Success:** A metric indicating the striker's ability to bypass defenders and create goal-scoring opportunities through individual skill.
- **Movement off the Ball:** Reflects how actively the striker moves to find space and create opportunities for themselves and teammates.
- **Hold-up Play:** Measures the striker's ability to retain possession and bring teammates into play with passes or layoffs.
- **Aerial Duels Won:** The number of aerial duels won by the striker, important for strikers strong in the air as it can create scoring chances.
- **Defensive Contribution:** Reflects the striker's defensive efforts such as tracking back, pressing opponents, and making interceptions.
- **Big Game Performance:** Indicates the striker's performance in important matches, which can elevate their reputation.
- **Consistency:** Reflects how regularly the striker performs at a high level over the course of a season or multiple seasons.
- **Versatility:** Measures the striker's ability to adapt to different tactical systems and roles within the team.
- **Penalty Success Rate:** The efficiency of the striker from the penalty spot, crucial in tight matches.
- **Impact on Team Performance:** Reflects how the team's results and overall attacking play are influenced by the striker's presence.
- **Off-field Conduct:** Measures the striker's professionalism, leadership, and behavior, which can impact their overall performance and value to the team.
- **Questions for this assignment**
 - What is the maximum goal scored by an individual striker?
 - What is the portion of Right-footed strikers within the dataset?
 - Which nationality strikers have the highest average number of goals scored?
 - What is the average conversion rate for left-footed player?

- How many left footed players are from France?
- What is the correlation co-efficient between hold up play and consistency score?
- What is the p-value for the shapiro wilk test of consistency score? Is it normally distributed?
- What is the p-value for the levene's test of ANOVA analysis? Is the heteroscedasticity assumed?
- Is there any significant correlation between strikers' Hold-up play and consistency rate?
- Describe the beta value of Hold-up Play you have found in your regression analysis.
- What is the average Total contribution score you get for the best strikers?
- What is the accuracy score of your LGR model? How many regular strikers your model predicted correctly? How many best strikers your model predicted incorrectly?