Statistical Analysis of Leading Goal Scoring Pattern of Europe’s Top Five Football Leagues

Jimewo G. Oghonyon, Hilary I. Okagbue, Emmanuella D. Amanie, Aderemi A. Atayero

Abstract: Previous researchers have discussed goal-scoring patterns of different football leagues, but little has been done on leading or top goal scorer. In this study, the focus is on the relationship between leading goal scoring and final league position for the top five European leagues (Premier League, La Liga, Serie A, Bundesliga and Ligue 1). The five leagues are the current top leagues in Europe as ranked by the Union of European Football Association (UEFA). The data were obtained from the respective leagues’ football association websites. The data were entered into Microsoft Excel spreadsheet to create frequency counts. Thereafter correlation and analysis of variance (ANOVA) was done using Minitab 18.0 and SPSS version 23. From the result, it was observed that there is a significant relationship between leading goal scoring and league positions for La Liga at a p value equals 0.013 and Germany at a p value equals 0.042, but no relationship for the remaining three leagues. Choosing the last 20 seasons only, the analysis of variance (ANOVA) revealed that the final league positions of the clubs that produced the leading goal scorers are the same across the leagues at p value = 0.349. However, the opposite is the case of the goal-scoring pattern and p value < 0.005 were obtained. Further investigation using the Post Hoc analysis revealed that the goal-scoring pattern of leading goal scores in Spanish La Liga is significantly different from others.

Keywords: Leading Goal Scoring, Highest Goal Scoring, Football, UEFA, Sports Analytics, Statistics.

I. INTRODUCTION

The leading or top goal scorer is a player among the top scorers that scores the most goals for the team in a season. There is an award given to the leading goal scorer, which is known as the Golden Boot. The leading goal scorer is determined at the end of a season of the various football leagues and commonly referred to as the highest goal scorer.

The paper examines the leading goal scoring pattern across the top five (5) football association leagues in Europe which are: Premier League (England and Wales), La Liga (Spain), Serie A (Italy), Bundesliga (Germany) and Ligue 1 (France). The leagues are occupying the most prestigious top five positions in Europe, according to the Union of European Football Association (UEFA) ranking.

A goal is scored if the whole ball passes over the specified goal line between goal posts. The teams competing try to score at one end of the football pitch while also stopping the other from scoring. If a ball enters a goal post knowingly or unknowingly by a defender of that goal post, then an own goal is said to have occurred. If a team wins, the team is awarded three points, if the team draws, just one point is awarded and if the team loses, no point is awarded. The cumulative of those points determines the order of which the teams finish at the end of the season. The champion is the team that tops the league and the last three teams are consigned to the lower league at the end of the season.

Goal scoring is the most important part of football. Goals decide outcomes and win matches [1]. The only way a team can be victorious is to outscore its opponent. Goal scoring determines if a team is to be promoted or relegated. Without goal scoring, football will be boring and tactics, individual efforts, strategies and team work often results to goals [2-5]. Goal scoring is purely probabilistic and always needed in the computation and finalization of the top teams at the end of a season [6]. Goal scoring is another important metric in determining the competitive nature of the football league [7]. Goal scoring helps in football prediction and in sports betting.

The pattern and comparison of leading goal scoring in European football leagues are rarely discussed. What is available is the goal scoring patterns of different football leagues and international or continental football competitions [8-13].

The aim of this research work is to explore the Leading goal scoring patterns of the top five European leagues and their relationship with final league position. The objectives of this study are:

A. Data Source

The data were obtained from Wikipedia and the respective football association websites.
B. Statement of Hypothesis

The following are the hypotheses formulated to answer the question raised in the objectives.

Hypothesis Statement 1

H0: There is no significance in the relationship between the leading goal scorer and the position of the player’s team in the final league table.

H1: There is significance in the relationship between the leading goal scorer and the position of the player’s team in the final league table.

Hypothesis Statement 2

H0: There is no significance in the leading goal scoring patterns across the five leagues.

H1: There is significance in the leading goal scoring patterns across the five leagues.

C. Inclusion Criteria

For this study, we are considering the top five football association leagues in Europe; Premier League, Bundesliga, La Liga, Serie A and Ligue 1. We have chosen these top leagues because they are the most competitive according to UEFA indices and ranking. Secondly, undoubtedly, these leagues are the ones with the largest fan base across the world.

D. Statistical Analysis

The data were entered into Microsoft Excel spreadsheet to create frequency counts. Thereafter, correlation, analysis of variance (ANOVA) were done using Minitab 18.0 and SPSS version 23.

III. RESULTS

The aforementioned statistical methods were applied to the data and the results are presented. The first analysis is to determine the frequency of final league positions of the clubs that produced the leading goal scorers and the number of goals scored by the scorers. The results are presented using the bar charts.

A. England

From Figure 1, it is observed that the highest goal scorer has mostly come from the team that finished first, then, the team that finished third and second respectively. We also have a situation where a team that finished 22nd produced the highest goal scorer and that is because we considered from the year 1888, whereas now only 20 teams compete.

The number of goals that resulting in the making of the leading or highest goal scores in English premier league is presented in Figure 2.

B. Spain

From Figure 2, the number of goals scored by the highest goal scorers has been mostly 25 and 31 whereas the least goals scored are 19, 36, 39, 42, 44, 49 and 60 in a season.

The analysis carried out is from the 87 seasons on the league positions of the leading or highest goal scorers. This is presented in Figure 3.

C. Germany

From Figure 3, it is observed that the highest goal scorers are often players that finished in the top four positions of the league.

The number of goals that resulting in the making of the leading or highest goal scores in Spanish La Liga is presented in Figure 4.
The number of goals scored by the highest goal scorers has been mostly 24 whereas the least goals scored are 12, 35, 37, 46, 48, and 50 in a season.

C. Italy

The analysis carried out is from the (1929 to 2018) 87 seasons on the league positions of the leading or highest goal scorers. This is presented in Figure 5.

From Figure 5, it is observed that the highest goal scorers are often players that finished in the top four positions of the league.

The number of goals that resulting in the making of the leading or highest goal scores in Italian Serie A is presented in Figure 6.

From Figure 6, the number of goals scored by the highest goal scorers have been mostly 24 whereas the least goals scored are 30, 33, 35, and 36 in a season.

D. France

The analysis carried out is from the 79 seasons on the league positions of the leading or highest goal scorers. This is presented in Figure 7.

From Figure 7, it is observed that the highest goal scorers are often players that finished in the top four positions of the league.

The number of goals that resulting in the making of the leading or highest goal scores in French Ligue 1 is presented in Figure 8.

From Figure 8, the number of goals scored by the highest goal scorers has been mostly 30 whereas the least goals scored are 15, 19, 29, 33, 36, etc. in a season.

E. Germany

The analysis carried out is from the (1963 to 2018) 55 seasons on the league positions of the leading or highest goal scorers. This is presented in Figure 9.
From Figure 9, it is observed that the highest goal scorers are often players that finished in the top four positions of the league.

The number of goals that resulting in the making of the leading or highest goal scores in German Bundesliga is presented in Figure 4.10.

From Figure 10, the number of goals scored by the highest goal scorers has been mostly 22 whereas the least goals scored are 31, 34, 36, 38 and 40 in a season.

F. Summary of the League Positions and Number of Goals

The mean of the respective league positions is higher than the median causing the skewness to be positive. This is presented in Tables 1 and 2 respectively. On the average, teams that finished in the top four often produce the leading or highest goal scorers.

Table 1: Summary of the Final League positions of the clubs that produced the highest goal scorers

<table>
<thead>
<tr>
<th></th>
<th>England</th>
<th>Spain</th>
<th>Italy</th>
<th>France</th>
<th>Germany</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Median</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>SD</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Skewness</td>
<td>1.33</td>
<td>2.09</td>
<td>1.64</td>
<td>1.36</td>
<td>1.83</td>
</tr>
</tbody>
</table>

G. Correlation Analysis

Correlation analysis between the league positions and the number of goals scored by the highest goal scorers showed a significant negative relationship at p-value = 0.05 in Spain and Germany and not significant elsewhere. Details are presented in Table 3.

Table 3: Summary of the result of the correlation analysis between the league positions and the number of goals scored by the highest goal scorers

<table>
<thead>
<tr>
<th></th>
<th>England</th>
<th>Spain</th>
<th>Italy</th>
<th>France</th>
<th>Germany</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>-0.067</td>
<td>-0.265</td>
<td>-0.062</td>
<td>-0.154</td>
<td>-0.276</td>
</tr>
<tr>
<td>p value</td>
<td>0.467</td>
<td>0.013</td>
<td>0.569</td>
<td>0.176</td>
<td>0.042</td>
</tr>
</tbody>
</table>

H. Analysis of Variance

This was done as a response to the second hypothesis which is to investigate whether the goal scoring and League positions are the same or different across the five leagues.

In order to ensure uniformity since all the leagues started at different years, only the last 20 years were considered for all the leagues. That is from 1998 to 2018. In addition, the analysis was done to reflect the current trend of football in Europe.

In the two results, the following applies;
- Null hypothesis All means are equal.
- Alternative hypothesis At least one mean is different.
- Significance level $\alpha = 0.05$.

Equal variances were assumed for the analysis.

The analysis of variance (ANOVA) revealed that the final league positions of the clubs that produced the leading goal scorers are the same across the leagues. However, the opposite is the case of the goal-scoring pattern. Further investigation using the Post Hoc analysis revealed that the goal-scoring pattern of leading goal scorers in Spanish La Liga is significantly different from the other leagues. The ANOVA and Post Hoc analyses are presented in Tables 4 and 5 respectively.

Table 4: Summary of the Analysis of Variance

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>League Position</td>
<td>1.12</td>
<td>0.349</td>
</tr>
<tr>
<td>Number of Goals</td>
<td>7.04</td>
<td>0.000</td>
</tr>
</tbody>
</table>
The football clubs of the leading goal scorers across the UEFA top five leagues finish between first and sixth position at the end of the season. The leading goal scorers score between 24 and 29 goals per season. Considering only the last 20 years, it was found that the average final league positions of the football clubs where the leading goal scorers emerges are the same across the five leagues. In addition, it was found that the average number of goals scored by the leading goal scorers across the leagues are different. Post Hoc analysis revealed that the average number of goals scored by the leading goal scorers of the leading goal scorers of the leading goal scorers of the leading goal scorers of the leading goal scorers across the five leagues. In addition, it was found that the average number of goals scored by the leading goal scorers across the leagues are different. Post Hoc analysis revealed that the mean number of goals scored in Spanish La Liga is higher than the other four football leagues. Lower leagues, continental and other football related competitions can be analyzed using the methods advocated in this work.

IV. CONCLUSION

ACKNOWLEDGMENT AND CONFLICT OF INTEREST STATEMENTS

There is absence of conflict of interest among the authors. The paper was sponsored by Covenant University, Ota, Nigeria.

REFERENCES


AUTHORS PROFILE

Jimevwo G. Oghonyon is of the Department of Mathematics, Covenant University, Ota, Nigeria.

Hilary I. Okagbue is of the Department of Mathematics, Covenant University, Ota, Nigeria.

Emmanuella D. Amanie is of the Department of Mathematics, Covenant University, Ota, Nigeria.

Aderemi A. Atayero is of the Department of Electrical and Information Engineering, Covenant University, Ota, Nigeria.