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

**Jimevwo 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 are 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:
▪ To obtain the general leading goal scoring patterns of the five leagues.
▪ To determine the relationship between the leading goal scorers and their respective final league positions.
▪ To perform comparative analysis to determine the general pattern across the five leagues.

The research will be helpful in the analysis of how the leading goal scoring patterns affect the overall competitive nature of the European leagues. The research is recommended for other sports and related activities [14-16].

## II. RESEARCH METHODS

### 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*

$H_o$: 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.

$H_1$: 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*

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

$H_1$: 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.
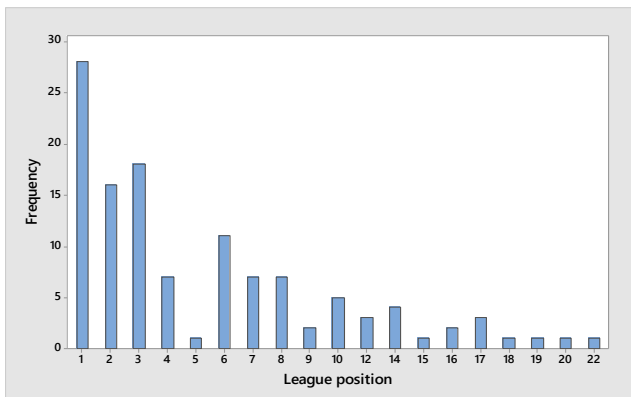
### A. England



**Figure 1: Bar chart of League positions of leading goal scorers (Premier League)**

The analysis carried out is from the year 1888 to 2018 (119 seasons) on the league positions of the leading or highest goal scorers. This is presented in **Figure 1**.

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 that 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**.
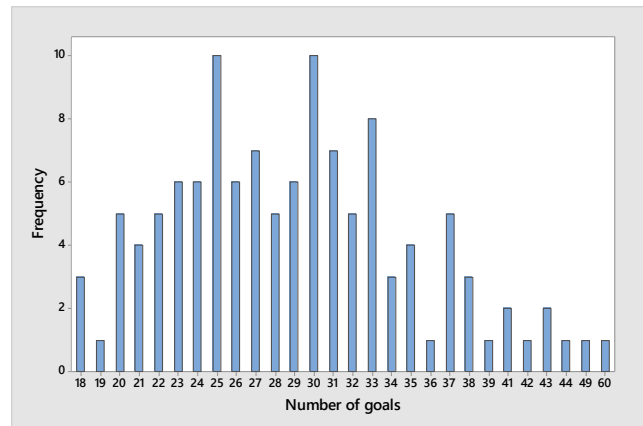


**Figure 2: Bar chart of number of goals scored by the highest goal scorers (Premier League)**

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.

### B. Spain

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**.
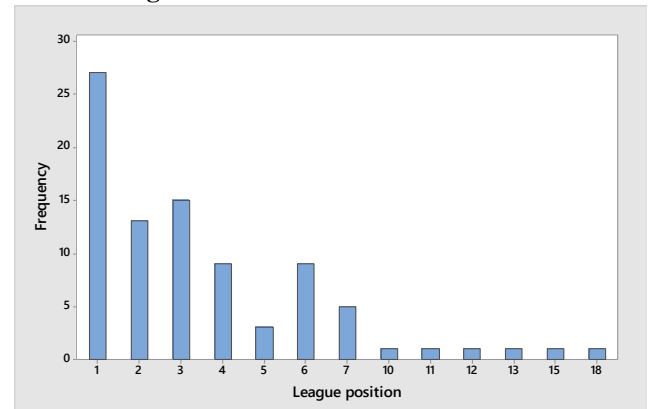


**Figure 3: Bar chart of League positions of leading goal scorers (La Liga)**

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.**
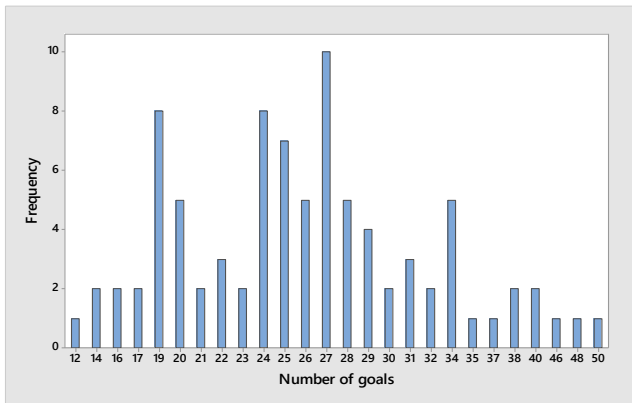
**Figure 4: Bar chart of number of goals scored by the highest goal scorers (La Liga)**

From **Figure 4**, the number of goals scored by the highest goal scorers has been mostly 27 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**.
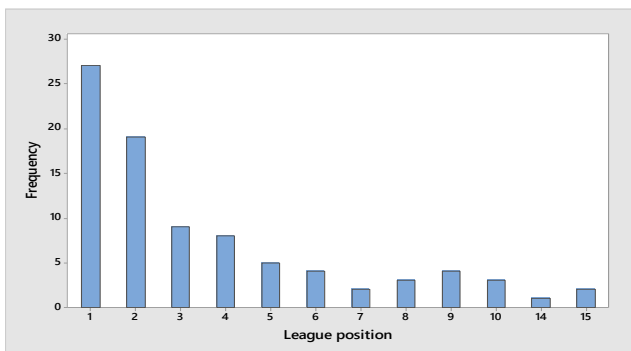


**Figure 5: Bar chart of League positions of leading goal scorers (Serie A)**

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 to the making of the leading or highest goal scores in Italian Serie A is presented in **Figure 6.**
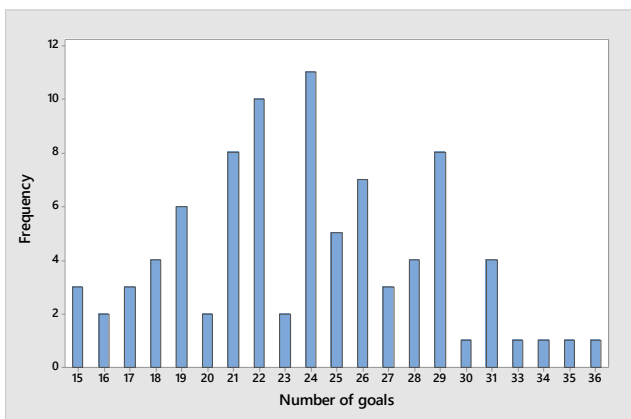


**Figure 6: Bar chart of number of goals scored by the highest goal scorers (Serie A)**

From **Figure 4.6**, the number of goals scored by the highest goal scorers have been mostly 24 whereas the least goals scored are 30, 33, 34, 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**.
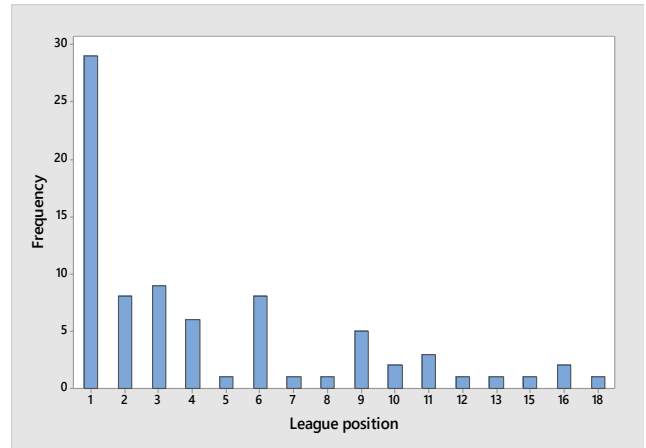


**Figure 7: Bar chart of League positions of leading goal scorers (Ligue 1)**

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**.



**Figure 8: Bar chart of number of goals scored by the highest goal scorers (Ligue 1)**

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**.

**Figure 9: Bar chart of League positions of leading goal scorers (Bundesliga)**

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.**
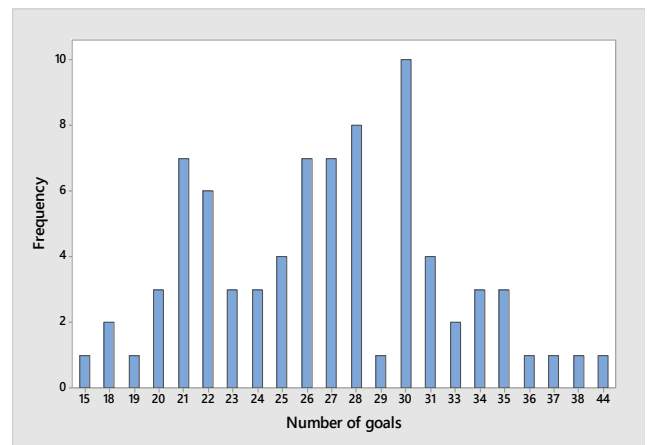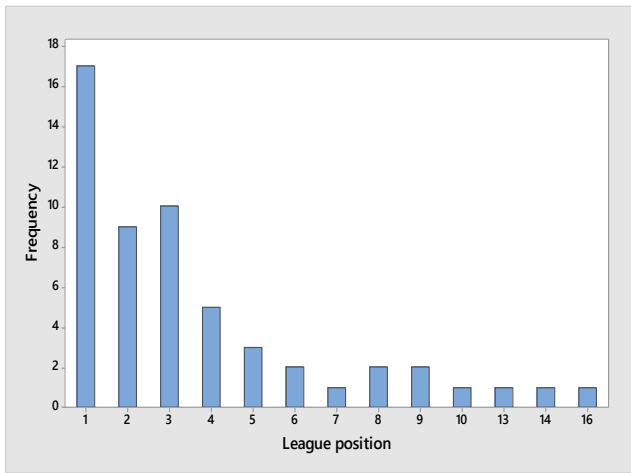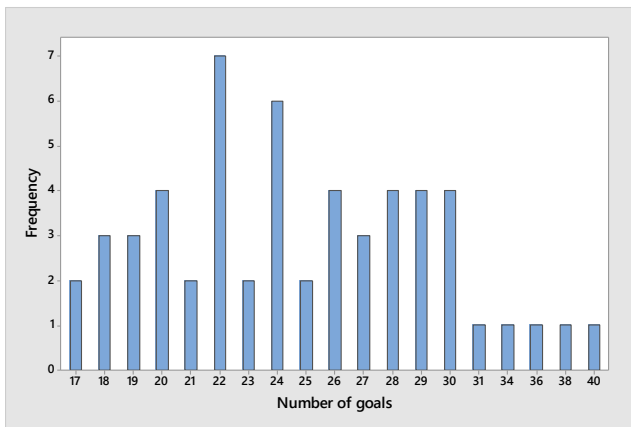


**Figure 10: Bar chart of number of goals scored by the highest goal scorers (Bundesliga)**

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**

|          | England | Spain | Italy | France | Germany |
|----------|---------|-------|-------|--------|---------|
| Mean     | 6       | 4     | 4     | 4      | 4       |
| Median   | 3       | 3     | 2     | 3      | 3       |
| SD       | 5       | 3     | 3     | 4      | 3       |
| Skewness | 1.33    | 2.09  | 1.64  | 1.36   | 1.83    |

**Table 2: Summary of the number of goals scored by the highest goal scorers across the respective leagues**

|          | England | Spain | Italy | France | Germany |
|----------|---------|-------|-------|--------|---------|
| Mean     | 29      | 26    | 24    | 27     | 25      |
| Median   | 29      | 26    | 24    | 27     | 24      |
| SD       | 9       | 7     | 5     | 5      | 5       |
| Skewness | 1,10    | 0.89  | 0.25  | 0.43   | 0.75    |

### 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**

|         | England | Spain  | Italy  | France | Germany |
|---------|---------|--------|--------|--------|---------|
| R       | -0.067  | -0.265 | -0.062 | -0.154 | -0.276  |
| p value | 0.467   | 0.013  | 0.569  | 0.176  | 0.042   |

This is the answer to the first hypothesis statement.

### 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 scores 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**

| Variable        | F    | P-value |
|-----------------|------|---------|
| League Position | 1.12 | 0.349   |
| Number of Goals | 7.04 | 0.000   |

**Table 5: Tukey Pairwise Comparison for the Number of Goals**

| Factor | N | Mean | Grouping* |
|---|---|---|---|
| La Liga | 20 | 32.30 | A |
| Serie A | 20 | 26.15 | B |
| Premier League | 20 | 26.00 | B |
| Ligue 1 | 20 | 24.45 | B |
| Bundesliga | 20 | 24.25 | B |

*Means that do not share a letter are significantly different.

## IV. CONCLUSION

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 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.

## ACKNOWLEDGMENT AND CONFLICT OF INTEREST STATEMENTS

## REFERENCES

1. Sarmento H, Bradley P, Anguera MT, Polido T, Resende R, Campaniço J. Quantifying the offensive sequences that result in goals in elite futsal matches. J. Sports Sci. 2016; 34(7): 621-629.
2. Machado JC, Barreira D, Garganta J. The influence of match status on attacking patterns of play in elite soccer teams. Revista Brasileira de Cineantropometria e Desempenho Humano. 2014; 16(5): 545-554.
3. Mara JK, Wheeler KW, Lyons K. Attacking strategies that lead to goal scoring opportunities in high level women's football. Int. J. of Sports Sci. Coach. 2012; 7(3): 565-577.
4. Wright C, Atkins S, Polman R, Jones B, Sargeson L. Factors associated with goals and goal scoring opportunities in professional soccer. Int. J. Perform. Analysis in Sport. 2011; 11(3): 438-449.
5. Ridgewell A. Passing patterns before and after scoring in the 2010 FIFA World Cup. Int. J. Perform. Analysis in Sport. 2011; 11(3): 562-574.
6. Evangelos B, Gioldasis A, Ioannis G, Georgia A. Relationship between time and goal scoring of European soccer teams with different league ranking. Journal of Human Sport and Exercise. 2018; 13(3): 518-529.
7. Vilar L, Araújo D, Travassos B, Davids K. Coordination tendencies are shaped by attacker and defender interactions with the goal and the ball in futsal. Human Movement Science. 2014; 33(1): 14-24.
8. Garganta J, Maia J, Basto F. Analysis of goal-scoring patterns in European top level soccer teams. Science and football III, 246-250, 1997.
9. Yiannakos A, Armatas V. Evaluation of the goal scoring patterns in European Championship in Portugal 2004. Int. J. Perform. Analysis in Sport. 2006; 6(1): 178-188.
10. Alberti G, Iaia FM, Arcelli E, Cavaggioni L, Rampinini E. Goal scoring patterns in major European soccer leagues. Sport Sciences for Health. 2013; 9(3): 151-153.
11. Njororai WWS. Analysis of goals scored in the 2010 world cup soccer tournament held in South Africa. Journal of Physical Education and Sport. 2013;13(1): 6-13.
12. Lastella M, Lovell GP, Rampinini E. Evaluation of goal scoring patterns between the 2016 Copa America and the 2016 European championship. Asian J. Sports Med. 2018; 9(1): Article number e59960.
13. Razali N, Mustapha A, Clemente FM, Ahmad MF, Salamat MA. Pattern analysis of goals scored in Malaysia super league 2015. Indonesian J. Elect. Engine. Comp. Sci. 2018; 11(2): 718-724.
14. Okagbue HI, Erondu EC, Atayero AA, Oguntunde PE, Opanuga AA, Olawande TI, Ijezie OA, Eze GA. Statistical analysis of frequencies of opponents' eliminations in Royal Rumble wrestling matches, 1988–2018. *Data in Brief*. 2018; 19: 1458–1465.
15. Okagbue HI, Adamu MO, Oguntunde PE, Opanuga AA, Rastogi MK. Exploration of UK Lotto results classified into two periods. *Data in Brief*. 2017; 14: 213-219.
16. Okagbue HI, Adamu MO, Atayero AA, Opanuga AA, Omogbadegun ZO, Obasi ECM. Regional Popularity and Gender Differences in Solving Sudoku Game among Some Sampled Secondary School Students in Lagos, Nigeria. *International Journal of Soft Computing*.2015; 10(6): 405-407.

## 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.