

The influence of ball possession, passes and shots on target in winning premier league football matches

A influência da posse de bola, dos passes e dos chutes ao gol na vitória das partidas de futebol da premier league

La influencia de la posesión del balón, los pases y los tiros a puerta para ganar partidos de fútbol de la liga premier

Received: 07/02/2021 | Reviewed: 07/06/2021 | Accept: 07/07/2021 | Published: 07/17/2021

Eric Matheus Rocha Lima

ORCID: <https://orcid.org/0000-0003-4254-5677>
Sao Paulo State University, Brazil
E-mail: ericmrlima@gmail.com

Ivan Wallan Tertuliano

ORCID: <https://orcid.org/0000-0001-6413-6888>
Anhembi Morumbi University, Brazil
E-mail: ivanwallan@gmail.com

Carlos Norberto Fischer

ORCID: <https://orcid.org/0000-0002-5598-6263>
Sao Paulo State University, Brazil
E-mail: carlos.fischer@unesp.br

Abstract

Football is the most popular sport in the world and involves the availability of information of footballers and teams to investigate success in this sport. In account of these aspects, the current study aims to analyze the influence of ball possession time, passes and shots on target in winning Premier League football matches. The methodology adopted consists of a quantitative, descriptive and documentary research, in which the sample is represented by all the 380 Premier League (PL) matches of the 2015/2016 season. The analysis process was made by the use of the association technique, present in the data mining process, with the support of the Weka software. The results suggest that shots on target, as well as its interaction with ball possession and passes - total passes, accuracy of passes, accuracy of passes inside the opponent's field and long passes - influenced the final result of the match. Based on the results, the conclusion is that direct play alternatives, such as presenting less ball possession time and more long passes, associated with more shots on target, showed to be more beneficial for football clubs to win football matches.

Keywords: Football; Data mining; Ball possession; Long passes; Venue; Winning football matches.

Resumo

O futebol é o esporte mais popular do mundo e envolve a disponibilidade de informações de jogadores e equipes para investigar o sucesso neste esporte. Tendo em conta estes aspectos, o presente estudo tem como objetivo analisar a influência do tempo de posse de bola, dos passes e dos chutes a gol na vitória de jogos de futebol da Premier League. A metodologia adotada consiste em uma pesquisa quantitativa, descritiva e documental, em que a amostra é representada por todos os 380 jogos da Premier League (PL) da temporada 2015/2016. O processo de análise foi feito por meio da utilização da técnica de associação, presente no processo de mineração de dados, com o apoio do software Weka. Os resultados sugerem que os chutes a gol, bem como sua interação com a posse de bola e passes - total de passes, precisão dos passes, precisão dos passes dentro do campo adversário e passes longos - influenciaram o resultado final das partidas. Com base nos resultados, conclui-se que alternativas de jogo direto, como apresentar menos tempo de posse de bola e mais passes longos, associados a mais chutes a gol, mostraram-se mais benéficas para os clubes de futebol vencerem partidas.

Palavras-chave: Futebol; Mineração de dados; Posse de bola; Passes longos; Local; Vencendo partidas de futebol.

Resumen

El fútbol es el deporte más popular del mundo e implica la disponibilidad de información de jugadores y equipos para investigar el éxito en este deporte. Teniendo en cuenta estos aspectos, este estudio tiene como objetivo analizar la influencia del tiempo de posesión de balón, pases y tiros a puerta en la victoria de los partidos de fútbol de la Premier League. La metodología adoptada consiste en una investigación cuantitativa, descriptiva y documental, en la que la muestra está representada por los 380 partidos de la Premier League (PL) de la temporada 2015/2016. El proceso de

análisis se realizó mediante la técnica de asociación, presente en el proceso de minería de datos, con el apoyo del software Weka. Los resultados sugieren que los tiros a puerta, así como su interacción con la posesión del balón y los pases (pases totales, precisión de pases, precisión de pases dentro del campo contrario y pases largos) influyeron en el resultado final de los partidos. Con base en los resultados, se puede concluir que las alternativas de juego directo, como tener menos tiempo en posesión del balón y más pases largos, asociados a más tiros a portería, fueron más beneficiosas para que los clubes de fútbol ganaran partidos.

Palabras clave: Fútbol; Procesamiento de datos; Posesión del balón; Pases largos; Lugar de eventos; Ganar partidos de fútbol.

1. Introduction

Football is the most popular sport in the world with high numbers of footballers and supporters (Britannica, 2015). It also involves the availability and the application of data from football matches, with teams' success achievement demanding the connection of the data from football matches with the theory (Wilson, 2013). Football, besides a sport, represents a cultural phenomenon that includes, in its context, cooperation and opposition relations among the footballers, involved by teams' objectives and opponents, evidencing the tactical relevance in this scenario (Garganta et al., 1997).

As it is a sport inserted in collective sports games (Barreira et al., 2011), football teams operate dynamically since they face contextual situations that can be unpredictable and each game variation relates to a match moment (Machado et al., 2011; Casanova et al., 2012). In account of these aspects, the present article can help treating variables to be explored in the football matches context.

Football teams are susceptible to suffer perturbations in frequently inter-team synchrony and for short periods of times, since previous investigations have shown that football consists in a sport which teams tend to move synchronously in all pitch directions (Frencken et al., 2012; Gonçalves et al., 2014). On this way, football components as ball possession, as short and long passes are included on footballers and teams' data and have been already investigated by several authors as well, contributing to new researches' development, as contributing to this research development as well, which brings the differential of analyzing these mentioned components (ball possession and passing variables) associated with shots on target.

Ball possession represents a component that consists in delivering the ball to a teammate and making passes without losing it to the opponent team (Anderson & Sally, 2013). According to the literature, ball possession's time obtained by football teams in a match is directly associated with success and failure (Lago-Peñas et al., 2010), passes frequency and precision (Anderson & Sally, 2013; Bradley et al., 2014; Collet, 2013), venue (Bradley et al., 2014; Lago-Peñas & Dellal, 2010), scoreboard (Bradley et al., 2014; Lago-Peñas & Dellal, 2010), teams' quality (Bradley et al., 2014; Lago-Peñas & Dellal, 2010), defensive movements (Vogelbein et al., 2014), goals scored (Yi et al., 2019) and also distance covered with and without the ball control (Bradley et al., 2014).

In addition to ball possession, in order to share contributions to football teams in their respective contexts, to help them not only to score goals, but also to win football matches and trophies, long passes, an alternative present in the direct play, have also been investigated by many researchers, as the direct play itself as well (Anderson & Sally, 2013; Bate, 1988).

Direct play is described as a possibility that involves set pieces, counter attacks, attacking movement of at least one long pass, a maximum of two passes and just a few ball touches, making fast transitions passing through the midfield (Tenga & Larsen, 2003), with its alternatives being seen as a good option to achieve success (Anderson & Sally, 2013; Bate, 1988; Garganta et al., 1997; Hughes, 1990; Stanhope, 2001).

Since both alternatives – ball possession and long passes – are considered by football teams to build their playing model to achieve their objectives, data analysis, in general, is gaining space inside the sport, sharing patterns to be explored by the clubs, in order to try to guarantee advantages against their opponents.

The data analysis area in football has significantly grown since 2010 (Sarmiento et al., 2014) and its main intention

consists in identifying strengths and weaknesses of football teams (Lago-Peñas et al., 2010). Data analysis also has shown itself to be a useful instrument, capable of describing, monitoring and modeling patterns, connecting them with performance in competitions (Lago-Peñas & Dellal, 2010; Lago-Peñas et al., 2010) and showing that football is not a sport isolated of innovations that appear to contribute for victories achievement (Drust & Green, 2013). Thus, because of those aspects, the current study aimed to analyze the influence of ball possession, - total passes, accuracy of passes, accuracy of passes inside the opponent's field and long passes - and shots on target to win Premier League football matches.

2. Methodology

The present research is characterized as a quantitative, descriptive and documentary research. The approaches related to this method involve a strict variable control with the use of precise measurements, focusing on objective analysis of a large database components. The descriptive aspect aims to describe certain phenomenon characteristics or obtained relations among investigated variables (Johnson & Wichern, 2007).

To carry out this research, data from the 380 Premier League (PL) matches, the first division of the English professional football, of the 38 matchweeks played at 2015/2016 season were analyzed. From these 380 matches, 157 were won by the home teams (41%), 107 ended in a draw (28%) and 116 matches were won by the away teams (31%).

In the first instance, it emphasizes that only statistical data from the matches played in the PL at 2015/2016 season are needed, so it is not necessary to submit this research to an ethics and research committee approval. To obtain the desired data, accesses to the Daily Mail website (<https://www.dailymail.co.uk/sport/football/premier-league/matchzone.html>) were made, specifically in the 'sport' topic, which is possible to access the information about the PL, to extract the desired indicators – ball possession, passes, shots on target, among others – to be analyzed.

At first instance, a spreadsheet program was used to facilitate the data visualization and make preliminary analyzes of the relevant indicators of those matches, allowing to generate additional information about these data and adequate the stored numbers. After that, the generated set of data was analyzed using the data mining software Weka (Weka, 2020), in particular, considering the association technique.

In this work, some data “classifiers” were proposed and used to make easier and faster subsequent analyses. These classifiers were created in function of the extracted indicators (Hughes, 1990). So, the match results classifier received 3 possible concepts: A, home team victory; B, draw; and C, away team victory.

The equivalent was used for the classifiers of indicators of ball possession, types of passes and shots on target, which also received 3 possible concepts: A, home team presented higher value for the specific indicator than the away team; B, home and away teams with the same value for the indicator; and C, away team had higher value for the specific indicator than the home team.

After classifying all relevant collected data of the 380 PL matches, the next step consisted in analyzing these data in order to identify relations between variables and the teams' success in the PL matches. The used software, Weka, generated association rules considering the variables and classifiers analyzed. An association rule represents an implication hypothesis in the format “A → B with probability P” (Agrawal & Srikant, 1994; Yang, 2005), describing the occurrence of B, dependent on A, with probability P for this occurrence, with support and confidence as the most used measures (Weiss & Zhang, 2003) – the first measures the frequency of the relations and the second the strength of the observed relations.

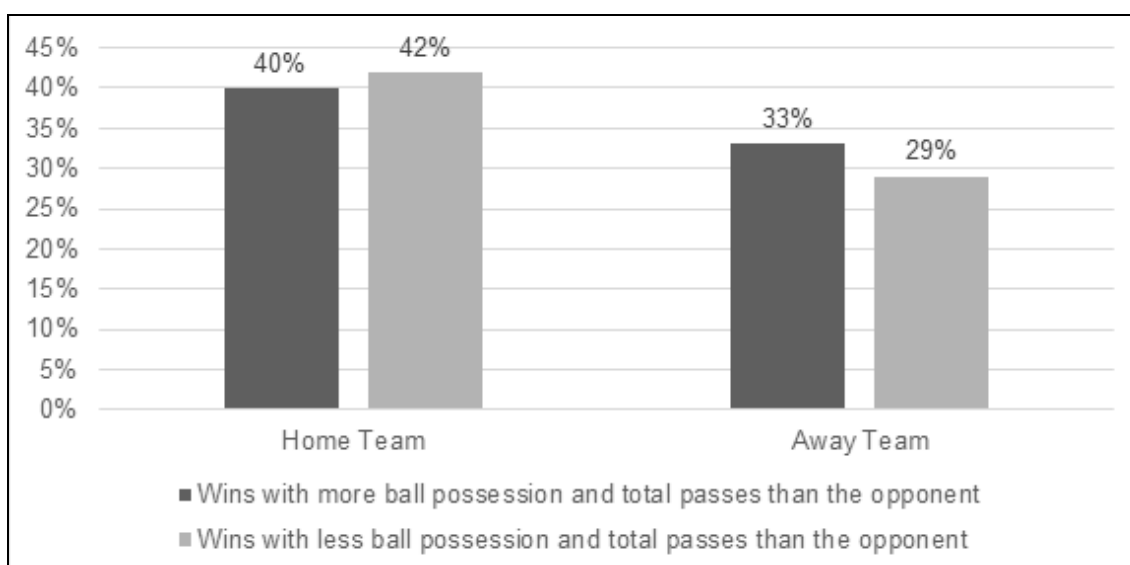
Weka produced a large set of association rules that was analyzed aimed to verify the contributions to the objectives of this research, characterizing the data exploitation process, that prioritizes making the information valid, confronting them with existent appointments and possibly getting to broad and effective generalizations (Weiss & Zhang, 2003).

3. Results

Results produced by Weka revealed that, at PL's 2015/2016 season, from a total of 199 matches in which the home teams had more ball possession time and performed more passes than the away teams, the home teams won 40% of these matches. Regarding the away teams, in 166 matches that they had more ball possession and also more passes, they obtained only 33% of wins (Figure 1). According to these results, having more ball possession time and also making more passes than the opposing teams were not enough to guarantee a high occurrence of victories at PL.

According to these results, Figure 1. is important to observe because it demonstrated that having more ball possession time and also making more passes than the opposing teams were not enough to guarantee a high occurrence of victories at PL and indicated that staffs should include other variables in their gameplay strategies in order to help their football teams to increase winning chances.

Figure 1. Wins occurrence connected to ball possession and total passes classifiers.



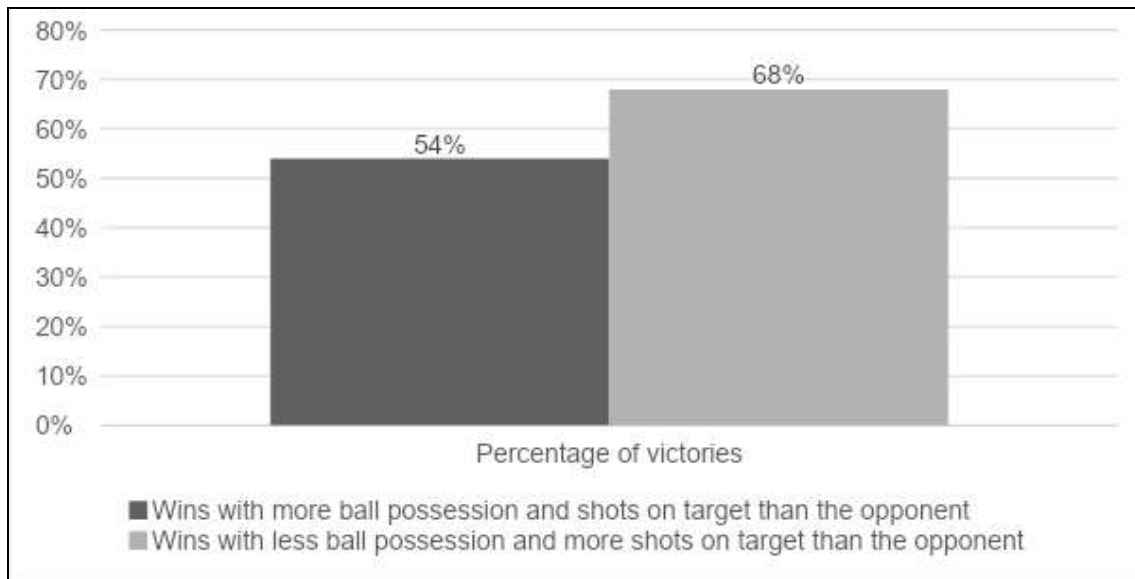
Source: Authors.

In addition, results produced by Weka showed that, from the total matches in which home teams performed more shots on target and had more ball possession time than the opponents (118 matches), they won 54% of these matches. However, in 63 matches in which home teams made more shots on target but registered less ball possession, the home teams won 68% of them (Figure 2). Weka's results also suggest that, in particular for the home teams, more shots on target associated with less ball possession time represent an important combination to be explored in matches.

Besides, Figure 2. is relevant to be observed because it indicated the importance of performing more shots on target than opponents in order to increase winning chances, being this component essential in gameplay strategies.

Moreover, Figure 2. is also important to be considered since it showed that teams which preferred to play controlling ball possession can succeed if they win the shots on target duel against their opponents, but not that much if they have preferred to play having less ball possession time and have kept winning the shots on target duel. Thus, Figure 2. also indicated that direct play alternatives are more recommended than indirect alternatives in order to guarantee a higher winning chance.

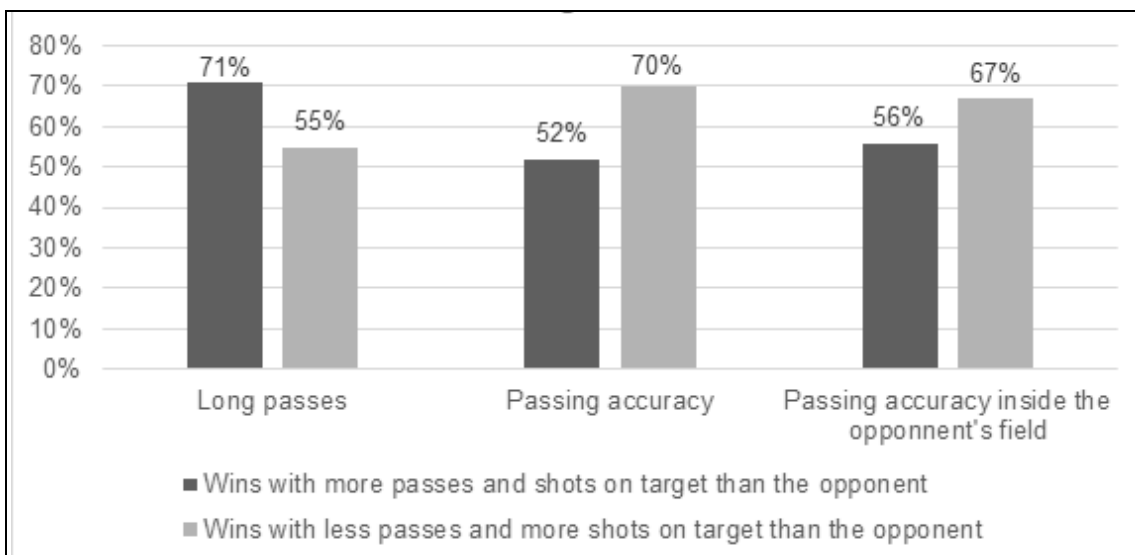
Figure 2. Wins occurrence connected to ball possession and shots on target classifiers.



Source: Authors.

Complementarily, Weka also generated other interesting association rules that show possible connections between the match results and the shots on target associated with the numbers of passes. Three different classifiers for passes are treated: long passes, passing accuracy and passing accuracy inside the opponent's field. In all situations, a particular case is considered: the home teams performed more shots on target; Figure 3 shows the resulting percentages for these cases. When the home teams had a higher amount of long passes than the opponents, they obtained 71% of wins in 56 matches; when they had less long passes, the percentage of wins was 55% in 131 corresponding matches. Related to the passing accuracy classifier, in 115 matches in which the home teams presented more passing accuracy, they won 52%; when they had less passing accuracy than the away teams, they won 70% of the 71 corresponding matches. Finally, when the home teams had more passing accuracy inside the opponent's field, in 126 matches, they won 56% of them; however, from a total of 61 matches in which the home teams presented less passing accuracy inside the opponent's field, they registered 67% of wins.

Figure 3. Wins occurrence connected to passes and shots on target classifiers.



Source: Authors.

By the end, Figure 3. is relevant to be considered because it reinforces Figure 2. and the argument that is possible to win if teams play controlling ball possession time during their matches, but that the winning chance considerably increases through fast and direct attacks, mainly associating the performance of more precise long passes and shots on target their opponents. Figure 3. also indicated that efficiency is what matters and it seems to be better achieved through direct alternatives instead of indirect approaches.

4. Discussion

Since the objective of this research is to analyze the influence of ball possession time, passes - total passes, accuracy of passes, accuracy of passes inside the opponent's field and long passes - in the victory of Premier League football matches, the results described above provide relevant contributions to professional football and to the context of the league in question.

Winning occurrence related to ball possession time and total passes

About the data revealed in Figure 1, at first instance about the ball possession and the total passes aspect, literature presents arguments that teams capable of dominating ball possession reveal high offensive efficiency, with the use of passes being an important aspect to the teams that are controlling the ball (Collet, 2013), and also authors that defend ball possession as a success enhancer indicator in tournaments as the World Cup (Lago-Peñas et al., 2010), the Spanish football league (Lago-Peñas & Dellal, 2010) and to win more matches (Anderson & Sally, 2013).

However, there are authors that alert that ball possession might not represent a relevant component to achieve success in the football scenario (Bate, 1988; Stanhope, 2001), meeting the results presented in this research, which can also be justified by rigid defensive formations, that might reduce available spaces for attacking movements and the chance to score goals, as concede counter-attacks hard to defend as well.

The venue aspect, which revealed itself as a relevant component to the winning percentages presented at Figure 1, has already been investigated into the sport context (Legaz-Arrese et al., 2013), being directly associated with ball possession, since home teams tend to play more in the offensive areas of the pitch, while defensive movements in deep areas were mostly evidenced in away teams (Lago et al., 2016; Taylor et al., 2010).

Match venue is considered an important factor for teams' defensive and offensive performance (Lago-Peñas & Lago-Ballesteros, 2011; Mackenzie & Cushion, 2012; Sarmiento et al., 2014). Generally, the best teams playing at home defend in a more consistent way, not depending much on goalkeepers interventions. Besides, performance of football teams is improved due to a variety of aspects, as the supporters, for example, that might increase the footballers' aggressive answer, rising the defensive actions efficiency (Almeida et al., 2014; Lago-Peñas & Lago-Ballesteros, 2011).

In addition, authors also argue that: (a) playing at home generates a winning occurrence of 42%, against 27% when playing away (Anderson & Sally, 2013), (b) venue associated with opponent's quality explain 48% of teams ball possession's variation (Lago-Peñas & Dellal, 2010), (c) the best teams of PL, Bundesliga and La Liga are expected to have approximately 54% of ball possession in their respective matches (Collet, 2013) and (d) playing away from home presents the tendency of drop in the ball possession time between 2% (Lago-Peñas & Dellal, 2010) and 3% (Bradley et al., 2014).

Wins related to shots on target and ball possession

Figure 2 shows that having more ball possession and more shots on target than the opponents presents a high victory occurrence, as supported by the literature. Authors point that associating more ball possession with less balls lost takes teams to higher positions in the classification table (Anderson & Sally, 2013; Lago-Peñas & Dellal, 2010), as it seems to exist a

tendency that teams with higher revenues reveal a higher ball possession percentage and more final third entrances, what indicates higher chance of controlling the matches actions (Lago-Peñas & Gómez-López, 2014).

However, the fact of having more shots on target and less ball possession than the opponent teams, revealing itself as an even better variable to be explored by football teams, is also supported by the literature. Since it is argued that 85% of goals scored comes from plays of 5 or less passes in a row, while less than 3% of the goals scored comes from plays of 10 or more passes in a row and the idea of the ball possession model produces less shots and goals (Hughes, 1990).

Matches won according to passes classifiers and shots on goal

Figure 3 reinforces, as Figure 2 does as well, that shots on target are more determinant for the success than having more ball possession and passes made. Authors defend that teams with higher passing and ball touches volume and precision are considered as well succeeded in finishing, scoring goals and victories occurrence (Collet, 2013; Grund, 2012).

Furthermore, they are capable of presenting more stable attacking movements (Liu et al., 2016) and game proposals, independent of the scoreboard condition (Lago-Peñas & Dellal, 2010). In addition, it is argued that teams with higher ball possession time than their opponents have won 39.4% of their matches, reducing the defeat quantity in nearly 7.6% and increasing the attacking production, registering, then, 7.8% more victories than teams with lower ball possession time (Anderson & Sally, 2013).

In addition, results of this research corroborate with the literature (Anderson & Sally, 2013; Grund, 2012). For example, shots on target were considered relevant not only in this research to achieve triumphs but also in a research on La Liga (Lago-Peñas et al., 2010). Complementarily, the shots on target and passing association is also supported by the literature that emphasizes that the ability to make shorts and long passes added to footballers' skills might influence the effectiveness in both game models, as influence in the defensive strategy adopted by the opponents as well (Fernandez-Navarro et al., 2016).

Still treating Figure 3, the association between long passes and shots on target showed itself not only as the most interesting variable to be explored but also evidenced the strength of direct play components at PL 2015/2016 season as well, since long passes are included in this game model. Literature reinforces that direct play can bring benefits to football teams in moments in which the game models based on more ball possession time and more total passes than the opponents do not reveal themselves relevant to obtain triumphs (Fernandez-Navarro et al., 2016; Tenga & Larsen, 2003), specifically at PL, since PL is a league based on attacks and counter attacks, which the midfield represents an area to pass through, not to control the game (Perarnau, 2016).

Although there are authors who have not found significant differences in goals scored through both having more ball possession or having more long passes (Tenga et al., 2010) and others who pointed that ball possession approaches produce more goals than long passes approaches (Hughes & Franks, 2005), literature brings data that support long passes effectiveness, as it was also possible to check that 61.8% of attacking actions of another research consisted of four or less passes as well (Mitrotasios & Armatas, 2014).

Literature also brings some football teams as examples of direct play effectiveness, like Newcastle United, that changed their game model after 3 FA Cup final defeats, in which they have adopted a game model based on having more ball possession (Wilson, 2013), as Chelsea and Real Madrid coached by José Mourinho as well, which have beaten Barcelona with less than 30% of ball possession time in three occasions, being the first two matches by the English team and the remaining one by the Spanish team (Anderson & Sally, 2013).

In addition, it is reinforced that direct play approaches increase the number of finishing opportunities, since the higher the time to make an attacking movement, the higher the time to the opponent team to defensively reorganize (Hughes, 1990)

and the lower the finishing conversion ratio (Anderson & Sally, 2013). Even if teams with game models based on long passes have fewer scoring opportunities, goals and tend to fight against relegation, there are exceptions in format of teams that have found game models to help maximizing the existent resources and ambitions (Anderson & Sally, 2013), being one of these exceptions the Leicester City, PL champions at the 2015/2016 season.

This mentioned efficiency of direct play resources is comprehensible, since attacking with more space and less opponents to pass through helps on finishing attacking movements and putting speed at footballers displacements. Besides, it possibly takes advantage of opponents' failures, which can also provide higher success chances, since it is easier to overcome a disorganized defense than a solid block of footballers already in the right positions to defend

5. Conclusion

According to the analysis carried out in this work, in order to achieve the objective of this research, which consisted in analyzing the influence of ball possession and passes - total passes, accuracy of passes, accuracy of passes inside the opponent's field and long passes - and shots on target in the victory of Premier League football matches, the results show that having more ball possession time and making more passes do not lead to a high percentage of victories. However, making more shots on target associated with either a greater volume of long passes or accuracy passes (on the whole field or on the opponent's field) increases the occurrence of wins. On the other hand, it seems that home teams have more advantage over the visiting ones when they have more shots on target and less ball possession or less accuracy passes, suggesting that these indicators should be further explored by the teams.

Thus, shots on target appeared as a strong indicator to be considered and maximized by football teams, being more relevant than ball possession and passes for success. In addition, this research reveals that it is more beneficial for football clubs to obtain victories when adopting direct game alternatives, such as presenting less time in possession of the ball and more long passes associated with more shots on target.

Ultimately, this study, like any study, has some limitations that are worth pointing out. Among the limitations are: the use of only one season; the monitoring of a single competitive european league. Thus, it is important that these limitations are investigated in future studies, as well as include classifiers referring to the quality of the football league teams in the analyses, with all available data of the games.

References

- Agrawal, R., & Srikant, R. (1994). Fast Algorithms for Mining Association Rules. *Proceedings of the 20th International Conference on Very Large Data Bases*, 1–13. VLDB Conference. <https://rakesh.agrawal-family.com/papers/vldb94apriori.pdf>
- Almeida, C. H., Ferreira, A. P., & Volossovitch, A. (2014). Effects of match location, match status and quality of opposition on regaining possession in UEFA champions league. *Journal of Human Kinetics*, 41(1), 203–214. <https://doi.org/10.2478/hukin-2014-0048>
- Anderson, C., & Sally, D. (2013). *The Numbers Game: Why Everything You Know about Soccer is Wrong*. Penguin Books.
- Barreira, D., Garganta, J., & Anguera, T. (2010). In search of nexus between attacking game-patterns, match status and type of ball recovery in European Soccer Championship 2008. In M. Hughes, H. Dancs, K. Nagyvaradi, T. Polgar, N. James, G. Sporis, G. Vuckovic, & M. Jovanovic (Eds.), *Research Methods and Performance Analysis* (pp. 226–237). University of West Hungary.
- Bate, R. (1988). Football chance: tactics and strategy. In T. Reilly, A. Lees, K. Davids, & W. Murphy (Eds.), *Science and football* (pp. 293–301). E & F.N. Spon.
- Besemann, C., Denton, A., & Yekkirala, A. (2004). Differential association rule mining for the study of protein-protein interaction networks. In M. J. Zaki, S. Morishita, & I. Rigoutsos (Eds.), *BIOKDD'04: Proceedings of the 4th International Conference on Data Mining in Bioinformatics* (pp. 72–80). Springer-Verlag. <https://doi.org/10.5555/3000580.3000590>
- Bradley, P. S., Lago-Peñas, C., Rey, E., & Sampaio, J. (2014). The influence of situational variables on ball possession in the English Premier League. *Journal of Sports Sciences*, 32(20), 1867–1873. <https://doi.org/10.1080/02640414.2014.887850>

- Britannica, E. (2015). Football. www.global.britannica.com/sports/football-soccer
- Collet, C. (2013). The possession game? A comparative analysis of ball retention and team success in European and international football, 2007-2010. *Journal of Sports Sciences*, 31(2), 123–136. <https://doi.org/10.1080/02640414.2012.727455>
- Casanova, F., Garganta, J., & Oliveira, J. (2012). Representativeness of offensive scenarios to evaluate perceptual-cognitive expertise of soccer players. *The open sports sciences journal*, 5, 161-166. <https://doi.org/10.2174/1875399x01205010161>
- Drust, B., & Green, M. (2013). Science and football: evaluating the influence of science on performance. *Journal of Sports Sciences*, 31(13), 1377–1382. <https://doi.org/10.1080/02640414.2013.828544>
- Fernandez-Navarro, J., Fradua, L., Zubillaga, A., Ford, P. R., & McRobert, A. P. (2016). Attacking and defensive styles of play in soccer: analysis of Spanish and English elite teams. *Journal of Sports Sciences*, 34(24), 2195–2204. <https://doi.org/10.1080/02640414.2016.1169309>
- Frencken, W., de Poel, H., Visscher, C., & Lemmink, K. (2012). Variability of inter-team distances associated with match events in elite-standard soccer. *Journal of Sports Sciences*, 30(12), 1207–1213. <https://doi.org/10.1080/02640414.2012.703783>
- Garganta, J., Maia, J., & Basto, F. (1997). Analysis of goal- scoring patterns in European top level soccer teams. In J. Bangsbo, M. Hugles, & T. Reilly (Eds.), *Science and football III. Proceedings of the Third World Congress of Science and Football* (pp. 246–250). E & F.N. Spon.
- Gonçalves, B. V., Figueira, B. E., Maças, V., & Sampaio, J. (2014). Effect of player position on movement behaviour, physical and physiological performances during an 11-a-side football game. *Journal of Sports Sciences*, 32(2), 191–199. <https://doi.org/10.1080/02640414.2013.816761>
- Greene, W. H. (2012). Econometric Analysis. In *Money Laundering in the Real Estate Sector* (7th ed.), Pearson education.
- Grund, T. U. (2012). Network structure and team performance: The case of English Premier League soccer teams. *Social Networks*, 34(4), 682–690. <https://doi.org/10.1016/j.socnet.2012.08.004>
- Hughes, C. (1990). *The Winning Formula*. London William Collins & Co.
- Hughes, M., & Franks, I. M. (2005). *Notational Analysis of Sport* (2nd ed.). Routledge.
- Johnson, R. A., & Wichern, D. W. (2007) *Applied multivariate statistical analysis* (6nd ed.). Pearson Prentice Hall.
- Lago-Peñas, C., & Dellal, A. (2010). Ball possession strategies in elite soccer according to the evolution of the match-score: The influence of situational variables. *Journal of Human Kinetics*, 25(1), 93–100. <https://doi.org/10.2478/v10078-010-0036-z>
- Lago-Peñas, C., & Gómez-López, M. (2014). How important is it to score a goal? the influence of the scoreline on match performance in elite soccer. *Perceptual and Motor Skills*, 119(3), 774–784. <https://doi.org/10.2466/23.27.PMS.119c32z1>
- Lago-Peñas, C., & Lago-Ballesteros, J. (2011). Game location and team quality effects on performance profiles in professional soccer. *Journal of Sports Science and Medicine*, 10(3), 465–471.
- Lago-Peñas, C., Lago-Ballesteros, J., Dellal, A., & Gómez, M. (2010). Game-related statistics that discriminated winning, drawing and losing teams from the Spanish soccer league. *Journal of Sports Science and Medicine*, 9(2), 288–293.
- Lago, I., Lago-Peñas, C., & Lago-Peñas, S. (2016). Democracy and football. *Social Science Quarterly*, 97(5), 1282–1294. <https://doi.org/10.1111/ssqu.12281>
- Legaz-Arrese, A., Moliner-Urdiales, D., & Munguía-Izquierdo, D. (2013). Home advantage and sports performance: Evidence, causes and psychological implications. *Universitas Psychologica*, 12(3), 933–943. <https://doi.org/10.11144/Javeriana.UPSY12-3.hasp>
- Liu, H., Gómez, M.-A., Gonçalves, B., & Sampaio, J. (2016). Technical performance and match-to-match variation in elite football teams. *Journal of Sports Sciences*, 34(6), 509–518. <https://doi.org/10.1080/02640414.2015.1117121>
- Machado, J., Barreira, D., & Garganta, J. (2010). Attacking game-patterns in soccer. a sequential analysis of the winner team of World Cup 2010. In M. Hughes, H. Dancs, K. Nagyvaradi, T. Polgar, N. James, G. Sporis, G. Vuckovic, & M. Jovanovic (Eds.), *Research Methods and Performance Analysis* (pp. 238–245). Szombathely: University of West Hungary.
- Mackenzie, R., & Cushion, C. (2012). Performance analysis in football: A critical review and implications for future research. *Journal of Sports Sciences*, 31(6), 1–38. <https://doi.org/10.1080/02640414.2012.746720>
- Menard, S. (2002). *Applied logistic regression analysis* (2nd ed.). Sage Publications. <https://doi.org/10.4135/9781412983433>
- Mitrotasios, M., & Armatas, V. (2014). Analysis of goal scoring patterns in the 2012 european football championship. *The Sport Journal*, 1–11. <https://doi.org/10.1057/9781137455062>
- Oliveira, S. R. M., Zaiane, O. R., & Saygin, Y. (2004). Secure association rule sharing. *8th Pacific-Asia Conference, PAKDD 2004*, 74–85. Sydney: Pacific-Asia Conference on Knowledge Discovery and Data Mining. https://doi.org/10.1007/978-3-540-24775-3_10
- Perarnau, M. (2016). *Pep Guardiola: La Metamorfosis*. Córner.
- Sarmento, H., Marcelino, R., Anguera, M. T., Campaniço, J., Matos, N., & Leitão, J. C. (2014). Match analysis in football: a systematic review. *Journal of Sports Sciences*, 32(20), 1831–1843. <https://doi.org/10.1080/02640414.2014.898852>
- Stanhope, J. (2001). An investigation into possession with respect to time in the Soccer World Cup 1994. In Mike Hughes (Ed.), *Notational analysis of sport III* (pp. 155–162). Cardiff: Centre for Performance Analysis, UWIC.

- Taylor, B. J., Mellalieu, D. S., James, N., & Barter, P. (2010). Situation variable effects and tactical performance in professional association football. *International Journal of Performance Analysis in Sport*, 10(3), 255–269. <https://doi.org/10.1080/24748668.2010.11868520>
- Tenga, A., Holme, I., Ronglan, L. T., & Bahr, R. (2010). Effect of playing tactics on achieving score-box possessions in a random series of team possessions from Norwegian professional soccer matches. *Journal of Sports Sciences*, 28(3), 245–255. <https://doi.org/10.1080/02640410903502766>
- Tenga, A., & Larsen, Ø. (2003). Testing the validity of match analysis to describe playing styles in football. *International Journal of Performance Analysis in Sport*, 3(2), 90–102. <https://doi.org/10.1080/24748668.2003.11868280>
- Vogelbein, M., Nopp, S., & Hökelmann, A. (2014). Defensive transition in soccer – are prompt possession regains a measure of success? A quantitative analysis of German Fußball-Bundesliga 2010/2011. *Journal of Sports Sciences*, 32(11), 1076–1083. <https://doi.org/10.1080/02640414.2013.879671>
- Weiss, S. M., & Zhang, T. (2003). Performance Analysis and Evaluation. In N. Ye (Ed.), *The handbook of data mining* (pp. 425–440). Lawrence Erlbaum Associates.
- Weka. (2020). *Documentação*. <http://www.cs.waikato.ac.nz/ml/weka.org/>
- Wilson, J. (2013). *Inverting the Pyramid: The History of Football Tactics* (2nd ed.). Orion.
- Yang, L. (2005). Pruning and visualizing generalized association rules in parallel coordinates. *IEEE Transactions on Knowledge and Data Engineering*, 17(1), 60–70. <https://doi.org/10.1109/TKDE.2005.14>
- Yi, Q., Gómez, M. A., Wang, L., Huang, G., Zhang, H., & Liu, H. (2019). Technical and physical match performance of teams in the 2018 FIFA World Cup: Effects of two different playing styles. *Journal of Sports Sciences*, 37(22), 2569–2577. <https://doi.org/10.1080/02640414.2019.1648120>
- Zheng, Z., Kohavi, R., & Mason, L. (2001). Real world performance of association rule algorithms. *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 401–406. ACM. <https://doi.org/10.1145/502512>