

How to Cite:

Baid, Y., Ghosh, A., Jana, S., & Giri, A. (2022). Evaluation of the conspicuous EPL matches for sponsorships using data envelopment analysis. *International Journal of Health Sciences*, 6(S2), 8208–8217. <https://doi.org/10.53730/ijhs.v6nS2.7078>

Evaluation of the conspicuous EPL matches for sponsorships using data envelopment analysis

Yash Baid

Student of Commerce, St. Xavier's College (Autonomous), Kolkata, India

Dr. Arijit Ghosh

Assistant Professor of Mathematics, St. Xavier's College (Autonomous), Kolkata, India

Subrata Jana

Assistant Professor of Mathematics, Seacom Engineering College, Howrah, India
Corresponding author email: 87subratajana87@gmail.com

Arpita Giri

Assistant Professor of Management, Regent Education and Research Foundation, Barracpore, India

Abstract---The analysis of games data, especially football is a thrilling field. Data Envelopment Analysis (DEA) was applied to find out most conspicuous matches out of 288 matches in the EPL 2019-20 using various inputs and various outputs variables. Total Fouls, Total Yellow Cards, Total Red Cards are taken as inputs variables. Number of Goals Scored, Total Number of Shots, Total Shots on target, Total Corners, Total Plays in Opponent Final Third, Market Value of Squad are taken as outputs variables. There are six matches have been found as most conspicuous. These matches are 41st (Liverpool Vs Newcastle United), 53rd (Manchester City Vs Watford), 68th (Everton Vs Manchester City), 91st (Leicester City Vs Southampton FC), 131st (Manchester City Vs Newcastle United) and 270th (Liverpool Vs West Ham). These results not only show the most conspicuous matches but also help to make financial decision for sponsorships. Most attractive matches bring more spectators.

Keywords---EPL, performance evaluation, DEA, sponsorships.

Introduction

The Premier League or the English Premier League is the top flight league, or in other words, the most elite competition in England. It was launched in 1992 by

the FA or the Football Association in a league-style format with promotions and relegations at the end of almost a 38-week long season. The twenty best teams of the nation play against each other in a home and away format with matches being mostly played on weekends. After completion of all the 380 matches, the team with the highest number of points is crowned the Champions. The top four teams earn a spot in the UEFA Champions League while the fifth placed team earns a place in the UEFA Europa League. The bottom three teams get relegated to the EFL Championship (the 2nd division of the English football system), while the top two teams of the EFL earn a promotion and a round of fixtures among the third, fourth, fifth and sixth teams of the EFL determines the team who will take the third promotion ticket. Unlike an auction session prior to the commencement of league, the Premier League operates in an open transfer window system. This refers to a system where clubs from various nations can interact with each other to purchase players by agreeing upon a fee to be paid to the selling club and meeting the player's personal wage requirements. This window is open from June to mid-September and halfway through the season in December. Though there is no rigid upper limit on the amount spent in the transfer window, the UEFA (The European Football Association) sets up a governing body to ensure financial fair play. This is done primarily to protect the clubs from landing up in any adverse financial position, to protect the creditors and stakeholders of a team and to ensure fair competition amongst financially distant clubs.

The Premier League attracts heavy attention from football fans and sponsors across the world. With the clubs in the league having a global appeal due to a vast and effective social media connect, the global television audience of a single premier league match averages a million and a half people. In EPL matches, sponsors invest money in looking forward to handsome profit. In terms of the selection of profitable match, they go by the popularity of the players and their teams whereas it requires a wholesome measurement and meticulous observation to auger that which match is going to be a conspicuous one. In this respect they need to be circumspect enough before sponsoring any particular. Mathematical calculation can be a weapon to cut the Gordian knot here and can give a certainty.

Efficiency refers to the style of play and the tactics used by a team in a match. In other words, it a close measure of how well a team plays and its foot balling performances. There are a number of factors which constitute the efficiency of team which can be broadly categorized as attacking and defending variables.

The attacking variables comprise of namely:

- a) Number of Shots Taken, b) Number of Shots on Target, c) Number of Goals Scored, d) Number of Crosses, e) Number of Corner Kicks, f) Number of Free Kicks, g) Number of Counter attacks, etc

The defending variables comprise of namely:

- a) Number of Tackles, b) Number of Interceptions, c) Number of duels won, d) Number of shots saved, e) Number of fouls versus the number of tackles attempted, etc.

In this research, we have implemented Output-oriented DEA in order to find out the most attractive matches of the tournament. We would further extend our study to investigate whether there are any particular teams whose matches turn out to be the most attractive ones and comment on other variables which show significant correlation with attractiveness. The grounds on which we have classified the variables as inputs and outputs is that, factors whose minimisation is favourable for the match being attractive have been considered to be inputs. Factors whose maximisation enables a match to be more alluring have been taken as outputs.

Literature Review

Rezania, K. et al (2013) employed data envelopment analysis to evaluate the efficiency of 48 sport association boards of Isfahan region regarding championship. Factors like selecting to matches, holding matches, the number of players and etc. are conducted. As per the constant return on scale calculation is done. Evaluating the effectiveness disclose that out of fifty present boards in Isfahan region, twenty four boards in the group of men & twenty two boards in the group of women have been competent in year 1990. In the first place in the group of women, blind and weak-sighted board, deaf board and martial art board have been identified after making rank. In men's group amidst fifty active association, blind and weak-sighted board holds the top position. At last sensitivity analysis of input data exhibits that sending to matches affects the ability of association boards most significantly. Djordjević, D. P. et al (2015) used the non-parametric data envelopment analysis approach to analyse the technical efficacy of football teams who were a part of the qualification stages of 2010 FIFA World Cup. Pyatunin, A. V. et al (2016) conducted a study on the overall efficiency of the top European football clubs by taking into account the factors attributing to operational or economic efficiency as well as factors pertaining to technical efficiency. The results from this study proved to be more holistic as economic factors (or factors influencing the financial position of a club), such as the prize money of a competition, revenue generated by sale of tickets, etc were also taken into account. Solntsev, I. et al (2016) devised a rating model for rating the sports development efficiency with respect to the investments made and how well were these resources put to use. Villa, G., & Lozano, S. (2016) presented a new way to measure the ability of scoring of football teams. It used the defensive and attacking variables of the 2 teams, their respective percentage of possession of the ball & the economic value of teams, a parallel process network DEA model taken into account. The projected way gives estimates of the number of goals each team could have scored in each match conducted. The team scoring efficiencies in the season can be calculated by averaging. For the season 2013/14, the projected way has been implemented to Spanish First Division teams & a virtual final league table has been evaluated. Freitas, M. M. D. et al (2017) used of DEA and Tobit Model, conducted an analysis of the efficiency of Brazilian football clubs in generating revenues. The Tobit model being primarily used to comprehend the influence of winning titles and the elite status of club having any financial impact. Oukil, A., & Govindaluri, S. M. (2017) used DEA and ordered weighted averaging operators to devise a ranking system for various football players on the basis of their desired and undesired skill sets, such as shot accuracy, fouls conceded, etc. Zelenkov, Y., & Solntsev, I. (2017) applied a quantitative analysis methodology for

the ability of football clubs comprising steps such as (i) in football theoretical analysis of production processes (ii) its explorative proof on the basis of PLS-SEM approach; (iii) Using DEA for calculating the efficacy of transformation multiple inputs into multiple outputs. The paper considered four seasons of RFPL (2012-13 to 2015-16). Proposed way can supports a football club for identifying the impotence and spotlight on efficacy-increasing strategies. Arabzad, S. M. et al (2018) evaluate the excellence of WC 2014 football matches. The quality criteria of football matches are evaluated for this purpose. The DEA-CWA is applied on account of evaluating and ranking the excellence of competitions. In Brazil's 2014 World Cup first round, the match between the national teams of Argentina and Nigeria was selected as the top quality match. Solntsev, I., & Osokin, N. (2018) applied a performance measurement system gives a frame to coordinate regional networks on the basis of different football expansion patterns in Russia. 5 performances relating to the Russian context were selected such as elite sport, player development, grassroots, infrastructure, development and promotion activities. This paper is also introduced a mapping procedure which supports for visualizing the potency of football development networks. The results of this paper showed that there are several areas which are directly connected in staging FIFA 2018 WC matches. Badmus, S. O. (2019) assessed the efficiency of the EPL football clubs ranking. The paper searched for identifying the Football Clubs that could be considered as most skilled by the help of a mixed methodological approach combining DEA with Naturalistic Approach. As per the suggestion of Substance (2010), the research approach is fixed in both qualitative and quantitative methods for including the case study research. DEA with NA used NVivo text analyses from data obtained through semi-structured questionnaire and mini-interview of football stakeholders for assessing the efficiency of English Premier League football clubs for twelve seasons from 2004-05 to 2015-16. 240 clubs were identified with the sum of 4,560 games. Miragaia, D. et al (2019) worked upon exploring a sustainable investment approach in European football clubs with the increasing volatility in the inflow and outflow of funds in the football industry. They decided to study the interdependence of financial efficiency and sports performance. Ghosh, A. et al (2021) looked into the efficacy of batsmen, bowlers and all-rounder, participated in the IPL 2019. At first, input data and output data were verified by using SEM and DEA was applied to evaluate the efficacy of the batsmen, bowlers and all-rounder by making their ranks based on DEA Scores. To evaluate the value for money spent on each player, the Auction Price was taken as a parameter. Cricketing parameters together with Auction price describes the ranking from the point of view of money. The outcomes helps the decision makers to figure out the players position based on cricketing performances and also cricketing return on the basis of investment of money upon them. Jana, S. et al (2021) looked into the efficacy of 8 teams in the IPL 2019 by the use of DEA & SEM to estimate the efficiency.

Methodology

DEA is a mathematical technique which is seen to have an extension over the basic linear programming methods. It draws substantial interest from the field of economics due to links with the production theory. The basic function of a DEA model is to comprehend the efficiency of various heterogeneous decision making units. It lays on the three core concepts of efficiency, effectiveness and

productivity. Efficiency refers to the ability of generating favourable output from a given set of resources. Effectiveness, on the other hand refers to the application of right techniques in order to achieve the desired results. Productivity can be termed as a function of both efficiency and effectiveness or the juncture at which both efficiency and effectiveness successfully integrate.

DEA technique is familiar in most efficiency evaluation models as it overcomes the limitation of one input and one output models. It is capable to compute efficiency in a tractable way even after considering multiple inputs and outputs. We employ an output-oriented model under which we seek to maximise the observed output while holding the inputs constant.

Min θ_j

Subject to : $\sum_{j=1}^p \lambda_j x_{mj} \leq \theta x_{jm}; (m = 1,2,3, \dots 1), (j = 1,2,3, \dots p)$

$\sum_{j=1}^{\infty} \lambda_j y_{mj} \leq y_{jm}; (m = 1,2,3, \dots k), (j = 1,2,3, \dots n)$

$\sum_{j=1}^p \lambda_j = 1;$

$\lambda_1 \geq 0 (j = 1,2,3, \dots, p)$

We assume that the DEA model has 'p' number of Decision-Making Units. For the 'l' inputs considered in the model, we see that every DMU produces 'k' outputs. The efficiency of each DMU can be calculated iteratively by using linear programming.

Data

The data on which we have based our analysis is of 288 matches of the English Premier League 2019-20 from 8th August 2019 to 9th March 2020. This data has been obtained from the website [GitHub.io](https://github.com) and has manually filtered and cleaned to emphasis more specifically on the variables of our interest. We have also incorporated the market value of each squad to this data. This statistic was obtained from the www.transfermarkt.co.uk

A key element of any research is the degree of accuracy with which the analysis has been conducted before any conclusion is drawn upon. As we all know that in our attempts to successfully conclude on certain issues pertaining the domain chosen, we are restricted by the fact that we only have access to a limited number of resources. Therefore, after careful consideration of the factors, we decided to proceed with our research by including the following variables:

Inputs:

- i. Total Fouls
- ii. Total Yellow Cards
- iii. Total Red Cards

Outputs:

- i. Number of Goals Scored
- ii. Total Number of Shots
- iii. Total Shots on target

- iv. Total Corners
- v. Total Plays in Opponent Final Third
- vi. Market Value of Squad

Data Envelopment using R Programming

R Programming was developed by Ross Ihaka and Robert Gentleman. The R language is considered to be an extension and improvement over the S and S Plus Language. It is a very interactive and user-friendly tool for statistical computations. It is regarded to be interactive due to the fact that R developers can publish and improvise on the functionality of R programming by introducing various packages which are available for all users across the globe. It is at the discretion of the user to enable such packages by downloading them as and when required or felt necessary. This makes the tool an ever developing and a self-sustaining one. Hence, we have chosen the R Programming for further analysis.

R Script for DEA

```
getwd()
setwd("C:/Users/Yaarav/Desktop/Sem6 Project")
#OPENING LIBRARIES
library(lpSolveAPI)
library(ucminf)
library(Benchmarking)
PremierLeague <- read.csv("EplModified.csv", header=TRUE)
x1 <- PremierLeague$TF
x2 <- PremierLeague$TYC
x3 <- PremierLeague$TRC
y1 <- PremierLeague$TG
y2 <- PremierLeague$TPP
y3 <- PremierLeague$Market.Value
y4 <- PremierLeague$TS
y5 <- PremierLeague$TSOT
y6 <- PremierLeague$TC
#Combining inputs and outputs
inputs <- data.frame(x1,x2,x3)
output <- data.frame(y1,y2,y3,y4,y5,y6)
#model
EffMeasure <- dea(inputs, output, RTS="crs", ORIENTATION="out",SLACK=TRUE)
```

Results and Analysis

After successfully conducting the data envelopment analysis using R, we obtain a number of statistics or values which enable us to have an insightful look at the optimization problem in hand. The “EFF” variable denotes the efficiency measurement of each decision-making unit, which in our case is each individual football match played. The “LAMBDA” is the matrix of values for constraint coefficients in the analysis and so on. With our primary analysis, we base our inferences on the efficiency measure of each match.

In the entire dataset consisting of 288 matches, we find that 6 matches have an efficiency score of exactly 1. These matches are 41,53,68,91,131 and 270. We take a brief look at the statistics of these matches:

Match 41 - This match was played between Liverpool and Newcastle United. The final score of the match was 3-1 in favour of Liverpool. On the offensive end of things, we see that there was a total of 29 shots taken with 10 of them being on target. A total of 11 corners and 29 plays in the final third happened in the entire game. Defensively, there were 9 fouls and no bookings. The individual squad value of Liverpool football club was around 966.95million Euros while that of Newcastle was 246.18 million Euros.

Match 53 - This match took place between Manchester City and Watford. Manchester City had the most expensive squad in the entire premier league, valued at around 1020 million Euros while Watford had a squad value of 181.64 million Euros. The match had a total of 8 goals, all of which were scored by Manchester City. There was a total of 4 yellow cards and 14 fouls. Offensively, we witnessed 27 shots with 10 of them being on target. There were 11 corners and a total of 16 plays in the final third.

Match 68 - The 68th match of Premier League 2019-2020 saw Everton take on Manchester City. Manchester City were positioned 2nd in the league and with the most expensive squad in the entire premier league, valued at around 1020 million Euros, had already attracted a huge crowd towards their style of play and the quality players who played for them. Everton had a squad 407.85 million Euros. Manchester City had won this match with the final score line being 3-1. A total of 32 shots were taken with 17 of them being on target. There were 7 corners, 19 plays in the final third composite, 4 yellow cards and 12 fouls.

Match 91 - Match 91 saw the heaviest defeat in the Premier League history in 24 years. Leicester City defeated Southampton FC by 9 goals to nil. There were a cumulative of 31 shots and 18 of them were on target, 9 corners and 10 press plays in the final third. On the defensive side, there were a total of 9 challenges made and only 1 yellow card. The market value of the Leicester squad was 403.4 million Euros and that of Southampton was 182.2 million Euros.

Match 131 - This match took place between Manchester City and Newcastle United. The match statistics show a total of 30 shots and 12 being on target with 11 corners and total of 4 goals. Manchester City had the most expensive squad in the entire premier league, valued at around 1020 million Euros while Newcastle United had a squad value of 246.18 million Euros. The final score line was 2-2 and both the teams took a single point. Defensively, 10 fouls were made and 3 yellow cards.

Match 270 - The match between Liverpool and West Ham saw 5 goals with Liverpool winning the match by 3 goals to 2. The estimated squad value of Liverpool was 966.95 million, the second most expensive squad in the league and West ham had a squad valued at 287.25 million Euros. The match had a total of 32 shots taken with 11 being on target. There were 24 corners and a total of 15

plays in the final third. Defensively, there were 14 fouls made and 3 yellow cards shown.

We should also have a quick look at the average statistics of every match:

Average Goals	Average Shots	Average Shots on Target	Average Corners	Average Cards	Average Plays in the final third	Average Fouls
2.7708	25.476	8.9063	10.903	3.5729	13.347	21.156

Out of the 6 matches, we see that there are three matches in which Manchester City are involved. One of the matches saw 8 goals being scored by them while the other two had a total of 4 goals in each of them. When compared to the average goals per match, the three games above when considered together against it have approximately 92% more goals. The squad also has big players such as Kevin De Bruyne, Sergio Aguero, Raheem Sterling. While considering the market value, we have taken the average of squad value at the start and end of season. This has been done with an aim to incorporate the rise in valuation which arises of the fact that the players have been performing well. So, the team was not only the most expensive one but was among the top two teams throughout the year-long season. These facts when combined with our analysis, present a very logical output of them being involved in the predicted attractive matches.

Liverpool FC finished as the champions in the Premier League Season 2019-2020. Out of the efficient matches, we find that two matches have Liverpool FC as one of the teams playing them. The club had made huge money signing of Virgil Van Dijk in the season prior to this one and went to the European Cup Final as well. This had already led to a huge hype among the crowd with respect to their performances. With the consistency in their performances in the 2019-2020 as well as good goal scoring form, their matches had to be one of the most interesting ones.

The match between Leicester City and Southampton was among the interesting ones because of the nine goals in them. It was a goal scoring delight with the biggest victory in 24 years. The addition of Market Value to other technical variables of the football match has resulted in a combination of teams which performed well with respect to the strength of their squad as well as the outcome they secured against very tough opponents such as Manchester City and Liverpool. Newcastle have also appeared twice on the list of most interesting games, when intuitively explaining this, we see that they drew against Manchester City with a squad worth 5 times lesser than Manchester City and displayed a very attacking performance against a Liverpool team which had a squad 4 times as expensive as theirs.

These results not only show the most interesting matches but also have financial implications. Since, it is evident that these teams were involved in the most interesting matches, we can draw various inferences about how effectively they have been utilising their resources. Improved performance of the club enables the club to attract finances from across the world with ease. These finances can be

invested further in the squad development and the infrastructure of the club such as the stadium, other training facilities and youth academies.

Conclusion

In this research paper, we have applied DEA, a simplistic approach to football matches and measuring the attractiveness which is very helpful for sponsors by taking efficiency as its indicator. Along with the inclusion of parameters which directly tell us the technical aspects of the game, we have included an additional parameter in the form of Market Value of Squad.

Limitations

This research paper is limited to EPL database of 2019-20 & limited number of inputs & outputs variables.

Future Scopes

The analysis could be further extended if we had the access to the number of viewers per match. This could help these sponsors or channels to maximise their revenues and allocate resources accordingly. We could conduct the same analysis on a much larger dataset. An analysis on the leagues of various countries during a season could give us international level information on the most established leagues and the leagues which are being run the most effectively. Conducting sensitivity analysis on the study could prove to be a significant extension and an upgrade on the study. This could enable us to vary the various factors and assess the responsiveness of the outputs with respect to the input parameters considered. Use of Multiple-Criteria Decision Analysis can enable us to evaluate multiple conflicting criteria in decision making. Such techniques could be incorporated to make the study more holistic and give a better insight on various aspects of football leagues.

References

- Arabzad, S. M., Berjis, N., & Shirouyehzad, H. (2018). Ranking World Cup 2014 Football Matches by Data Envelopment Analysis Models with Common Weights. *Iranian Journal of Optimization*, 10(2), 139-150.
- Badmus, S. O. (2019). *An Assessment of Performance and Efficiency of Team-Based Organisations: Empirical Evidence of English Premier League (EPL) using Data Envelopment Analysis (DEA) and Naturalistic Approach (NA) 2005 to 2016* (Doctoral dissertation, DEPARTMENT OF ACCOUNTING AND FINANCE A Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy, School of Business and Law, London Metropolitan University, London, United Kingdom).
- Djordjević, D. P., Vujošević, M., & Martić, M. (2015). Measuring efficiency of football teams by multi-stage DEA model. *Technical Gazette*, 22(3), 763-770.
- Freitas, M. M. D., Farias, R. A. S., & Flach, L. (2017). Efficiency determinants in Brazilian football clubs. *Freitas, MM*, 1-23.

- Ghosh, A., Dev, M., Guha, B., Jana, S., & Sarkar, A. (2021). Performance evaluation & rankings of players in IPL 2019 by DEA & SEM. *Malaya Journal of Matematik*, Vol. S, (1), 46-56.
- Jana, S., Ghosh, A., & Guha, B. (2021). IPL 2019: Evaluating the performance of teams by DEA & SEM. *Malaya Journal of Matematik* Vol. S, (1), 41-45.
- Miragaia, D., Ferreira, J., Carvalho, A., & Ratten, V. (2019). Interactions between financial efficiency and sports performance: Data for a sustainable entrepreneurial approach of European professional football clubs. *Journal of Entrepreneurship and Public Policy*.
- Oukil, A., & Govindaluri, S. M. (2017). A systematic approach for ranking football players within an integrated DEA-OWA framework. *Managerial and Decision Economics*, 38(8), 1125-1136.
- Pyatunin, A. V., Vishnyakova, A. B., Sherstneva, N. L., Mironova, S. P., Dneprov, S. A., & Grabozdin, Y. P. (2016). The Economic Efficiency of European Football Clubs--Data Envelopment Analysis (DEA) Approach. *International Journal of Environmental and Science Education*, 11(15), 7515-7534.
- Rezania, K., Mokhatab Rafiei, F., & Shirouyehzad, H. (2013). Performance Evaluation of sport association board of Isfahan Province through DEA and a championship approach. *International Journal of Data Envelopment Analysis*, 1(4), 227-246.
- Solntsev, I., & Osokin, N. (2018). Designing a performance measurement framework for regional networks of national sports organizations: evidence from Russian football. *Managing Sport and Leisure*, 23(1-2), 7-27.
- Solntsev, I., Vorobyev, A., Irmatova, E., & Osokin, N. (2016). Rating evaluation of sports development efficiency using statistical analysis: evidence from Russian football. *arXiv preprint arXiv:1612.07543*.
- Villa, G., & Lozano, S. (2016). Assessing the scoring efficiency of a football match. *European Journal of Operational Research*, 255(2), 559-569.
- Zelenkov, Y., & Solntsev, I. (2017). Measuring the efficiency of Russian Football Premier League clubs. *Electronic Journal of Applied Statistical Analysis*, 10(3), 773-789.