

Master of Science Final Project

Forecasting Football Outcomes to Invest in Betting Markets

Business Administration

João Maria Telles Moniz Côrte-Real
2017/2018

Supervisor: Professor José Joaquim Dias Curto, PhD
Associate Professor at ISCTE Business School, Department of
Quantitative Methods for Management and Economics

Master of Science Final Project

Forecasting Football Outcomes to Invest in Betting Markets

Business Administration

João Maria Telles Moniz Côrte-Real
2017/2018

Supervisor: Professor José Joaquim Dias Curto, PhD
Associate Professor at ISCTE Business School, Department of
Quantitative Methods for Management and Economics

Acknowledgements

To Professor Dias Curto, for the patience and availability

To my father, and to my uncles and aunts from the mother's side (especially to my godparents João and Luísa), for supporting me through this journey

To my girlfriend Inês, the love of my life, for the understanding

To my aunts Isabel and Filipa, for loving me as a son

To my stepmother Domingas, for treating me as a son

To my mother, to my grandparents (Leonel, Graça, Inês and Filipe) and to my cousin Cristina, for looking after me

To my sister and niece, Gracinhas, for the laughs

To my cousin António, for the unconditional brotherhood

To all my best friends, for the good times spent together

Table of Contents

1- Abstract	I
2- Executive summary	III
3- Introduction	1
3.1- Forms of placing bets	2
3.1.1- Fixed-odds betting	2
3.1.2- Exchange betting	5
3.2- Legal context of the betting markets in Portugal	6
3.3- Purpose of the project.....	9
4- Literature review	10
4.1- Statistical models.....	10
4.1.1- Poisson regression models	10
4.1.2- Ordered probit/logit regression models	11
4.2- Fixed-odds betting markets efficiency	13
4.2.1- Arbitrage opportunities	15
4.3- Home advantage	15
4.4- Money management methods.....	19
4.4.1- Fixed-stake	19
4.4.2- Martingale	19
4.4.3- Kelly criterion	20
4.5- The importance of subjective factors	20
4.6- Conclusions about the literature review	22
5- Method	23
5.1- Poisson regression model	23
5.1.1- Counting process and Poisson process definitions.....	23
5.1.2- Poisson random variable	24
5.1.3- Dataset	25
5.1.4- Poisson regression	32
5.2- Betting strategies	39
5.2.1- Value strategy.....	40
5.2.2- Accuracy strategy	47
6- Discussion of the results.....	50
6.1- Portuguese online fixed-odds betting markets	50
6.2- Betting record	52
6.2.1- Portuguese online fixed-odds betting markets efficiency	53

Forecasting Football Outcomes to Invest in Betting Markets

6.2.2- The importance of our forecasting method in appraising the expected value	56
7- Conclusions	59
7.1. Summarizing the value strategy	59
7.1.1- Limitations and recommendations for further research	59
7.2- Summarizing the accuracy strategy.....	59
7.3- Portuguese online fixed-odds and markets efficiency.....	60
7.4- Final comment.....	61
8- References	62
8.1- Books.....	62
8.2- Periodicals	62
8.3- Unpublished works.....	64
8.4- Websites	64
Annexes	67
Annex I- Watford vs Everton (match day 28)	67
Annex II- Arsenal vs Manchester City (match day 28).....	74
Annex III- Manchester City vs Chelsea (match day 29)	80
Annex IV- Manchester United vs Liverpool (match day 30)	87
Annex V- Liverpool vs Watford (match day 31)	94
Annex VI- Bournemouth vs West Bromwich Albion (match day 31)	102
Annex VII- Everton vs Manchester City (match day 32).....	110
Annex VIII- Arsenal vs Southampton (match day 33).....	118
Annex IX- Arsenal vs West Ham (match day 35).....	125
Annex X- Tottenham vs Watford (match day 36).....	133
Annex XI- Southampton vs Manchester City (match day 38)	141

1- Abstract

Our main objective was to develop a method which can be used to obtain long-term returns on football bets. The forecasting of the matches was made using quantitative (including a Poisson regression model) and qualitative (of subjective interpretation) information. This approach was named value strategy, being able to generate a yield rate of 19,88%, and to predict 8 out of 11 outcomes. However, the number of bets made is too small to conclude that this strategy is capable of providing consistent earnings, in the long-run.

The accuracy strategy consisted in betting on favourites who played at home, and produced returns of 0,39%, not compensating for the risks assumed. If this strategy was able to create revenues, it could indicate the presence of biases in the Portuguese betting markets, and consequently, that these were semi-strong efficient. In practice, the accuracy strategy broke even, therefore not suggesting that the Portuguese markets' odds were biased. However, if we did the same bets through Pinnacle (UK-based bookmaker), the returns would be of 6,62%, hence indicating the existence of biases and semi-strong efficiency.

Surprisingly, if one removes the overround, Portuguese bookmakers would have offered better odds for the home wins than Pinnacle. This demonstrates the need of reviewing the Portuguese online gambling legal framework, which is the main reason for the poor competitiveness of Portuguese betting markets. As far as our research went, we could not find any studies about these markets, hence we consider that our work adds an important contribution to the existing literature.

Keywords: Betting Markets, Football, Forecasting, Sports Betting

JEL Classification System: Information and Market Efficiency, Sports Economics

O nosso principal objetivo foi desenvolver um método que possa ser usado para obter retornos de longo prazo em apostas de futebol. A previsão dos jogos foi feita utilizando informação quantitativa (incluindo um modelo de regressão Poisson) e qualitativa (de interpretação subjetiva). A esta abordagem chamou-se estratégia de valor, tendo sido capaz de originar retornos de 19,88%, e de prever 8 em 11 resultados. Contudo, o número de apostas feitas é insuficiente para considerar que esta estratégia consegue gerar retornos consistentemente, a longo prazo.

Forecasting Football Outcomes to Invest in Betting Markets

A estratégia de precisão consistiu em apostar nos favoritos que joguem em casa, e conseguiu retornos de 0,39%, não compensando os riscos assumidos. Se esta estratégia fosse capaz de criar retornos, podia indicar a tendenciosidade dos mercados de apostas Portugueses, e conseqüentemente, que estes possuem eficiência semiforte. Na prática, a estratégia de precisão ficou na mesma, não sugerindo, portanto, que as odds dos mercados Portugueses são tendenciosas. No entanto, se fizéssemos as mesmas apostas através do Pinnacle (casas de apostas Britânica), os retornos seriam de 6,62%, indicando a existência de tendenciosidade e de eficiência semiforte.

Surpreendentemente, se removéssemos a margem implícita, as casas de apostas Portuguesas teriam oferecido melhores odds do que o Pinnacle, para vitórias da equipa caseira, o que demonstra a necessidade de rever as leis do jogo online em Portugal, que são a principal razão para a falta de competitividade dos mercados de apostas Portugueses. Não conseguimos encontrar nenhum estudo acerca desses mercados, assim, consideramos que o nosso trabalho acrescenta um importante contributo à bibliografia existente.

Palavras-chave: Mercados de Apostas, Futebol, Projeções, Apostas Desportivas

JEL Classification System: Information and Market Efficiency, Sports Economics

2- Executive summary

A ideia de desenvolver este projeto emerge, essencialmente, do fascínio que o seu autor sempre demonstrou por temas relacionados com o futebol, nomeadamente as apostas desportivas (especialmente após a leitura do livro “Como ganhar com as apostas desportivas”, de Paulo Rebelo). Durante o período em que frequentou as aulas do mestrado em Gestão de Empresas, e apesar de não possuir grandes bases matemáticas, interessou-se pela cadeira de Métodos Quantitativos, lecionada pelo Professor Dias Curto. Daí surgiu a ambição de elaborar um método de previsão de resultados desportivos, que lhe permitisse ganhar dinheiro ao mesmo tempo que desenvolvia as suas capacidades analíticas.

A força motriz deste projeto foi a criação de um método de avaliação de jogos futuros entre equipas de futebol, com base em ocorrências passadas, que servisse de suporte a tomadas de decisão relativas a potenciais investimentos nos mercados de apostas a operar legalmente em Portugal. Considerou-se extremamente importante que este sistema fosse facilmente compreendido e aplicável (com maior ou menor sucesso) pela maioria dos apostadores, o que diferencia este trabalho da maior parte das publicações existentes, que se focam nas componentes teóricas e académicas, sendo por isso mais direcionadas a um público de estudiosos e professores dos ramos estatísticos e económicos.

Após efetuar uma análise comparativa dos mercados de apostas nacionais e internacionais, o autor apercebeu-se de limitações significativas no caso Português, tanto ao nível dos tipos de mercados existentes, como da proposta de valor implícita nas odds oferecidas, passando pelo número anormalmente pequeno de casas de apostas a operar legalmente no nosso país. Posteriormente, apurou-se que estas limitações se devem, em grande parte, à ineficiência dos decretos de lei que regulam o jogo online e as apostas desportivas em Portugal, limitações essas que são prejudiciais aos interesses dos 3 principais “players” deste mercado: os apostadores, as casas de apostas e, em última análise, o próprio Estado Português.

O passo seguinte passou por uma revisão literária extensiva, abrangendo todos os tópicos que o autor julgou relevantes para ampliar o seu grau de conhecimento acerca deste tema, facilitando assim a definição dos conteúdos a serem considerados na avaliação dos jogos. Deste modo, procurou-se perceber se já existiam modelos econométricos que estimassem probabilidades para os possíveis desfechos de um jogo de futebol, e a existirem, qual seria

o mais adequado. Analisaram-se também diversos artigos científicos, no sentido de aferir o nível de eficiência dos mercados de odds fixas, quanto à informação publicamente disponível.

Concluiu-se que o modelo de regressão Poisson era o mais indicado, por permitir estimar probabilidades não só para as hipóteses de desfecho (vitória caseira, empate, ou vitória forasteira), mas também para o número de golos a ser marcado por cada uma das equipas. Concluiu-se também que, ao longo dos anos, foram detetadas de forma sistemática odds tendenciosas nos mercados de apostas, indicando ineficiências potencialmente exploráveis por parte dos apostadores. Para além disso, nesta fase do projeto, estudaram-se igualmente outros aspetos que podem influenciar os acontecimentos das partidas, como por exemplo, a importância da análise das variáveis qualitativas (informação dificilmente quantificável), que para Rebelo, até é superior à da análise quantitativa.

Com base nas descobertas proporcionadas pela revisão literária, definiu-se um método de avaliação composto por variáveis quantitativas (nas quais se incluem, para além do modelo de Poisson, estatísticas descritivas da performance das equipas) e qualitativas (de interpretação subjetiva, mas bem definidas), no sentido de identificar os jogos em que as odds propostas pelos mercados de apostas Portugueses se enquadravam no conceito de valor esperado positivo, que é definido e explicado por Rebelo na obra acima referida. Uma vez que não foram encontrados quaisquer estudos relativos aos mercados de apostas Portugueses, tomou-se como segundo objetivo deste projeto testar a eficiência dos mesmos, e para isso, desenvolveu-se uma estratégia que consistia em apostar apenas nos claros favoritos, quando estes jogavam em casa: se esta estratégia fosse capaz de gerar retornos positivos, poderia indiciar a presença de alguma(s) das odds tendenciosas recorrentemente detetadas nas publicações anteriormente consultadas.

Foram guardados registos de todas as apostas efetuadas, para posteriormente comparar as odds oferecidas, nomeadamente em termos das odds médias para cada hipótese de desfecho, bem como da margem média cobrada por cada uma das casas de apostas (overround), com as odds oferecidas, para os mesmos jogos, por uma casa de apostas sediada no Reino Unido (Pinnacle), a fim de avaliar o nível de competitividade dos mercados de apostas Portugueses, face a uma referência internacional.

Após o final da época 2017/2018, analisaram-se os resultados, e tiraram-se as respetivas ilações:

Forecasting Football Outcomes to Invest in Betting Markets

- Através do da aplicação do método preditivo idealizado pelo autor, geraram-se retornos na casa dos 20%, face ao investimento realizado. No entanto, devido ao grau de detalhe com que os relatórios foram escritos, apenas foi possível realizar 11 apostas, o que é manifestamente pouco para se poder concluir que esse método é capaz de gerar retornos positivos de forma consistente, no longo prazo.
- A estratégia relacionada com apostar em equipas caseiras amplamente favoritas foi capaz de gerar retornos teoricamente positivos, embora praticamente nulos. Na realidade, não se pôde concluir a existência de tendenciosidades exploráveis nas odds propostas pelas casas de apostas autorizadas a operar no mercado Português.
- Em relação à competitividade dos mercados de apostas Portugueses, verificou-se que, em média, as 4 casa de apostas consideradas neste estudo cobram um overround bastante superior ao do Pinnacle (8,11% vs 2,75%). Se as mesmas apostas em equipas caseiras favoritas fossem efetuadas através do Pinnacle, seria possível obter retornos de 6,62%, que indicariam a presença de odds tendenciosas exploráveis, e a consequente ineficiência dos seus mercados. Apesar de, em média, o Pinnacle conseguir propor odds superiores para a vitória dos favoritos a jogar em casa, se retirássemos o peso do overround, os mercados de apostas Portugueses teriam fornecido, comparativamente, odds melhores. Uma vez que a pesquisa previamente conduzida identificou a legislação em vigor como a principal causa do overround excessivo nas odds portuguesas, este trabalho sugere que a mesma deveria ser rapidamente revista, sob pena que cada vez mais apostadores optem por apostar ilegalmente.

3- Introduction

“Determining the origin of sports betting is a difficult task. In its truest form, sports betting has been around for thousands of years, as people have always wagered on the outcome of an event between two competitors.” (Moody, 2017)

Although people have always been interested in betting on the outcome of sports events, more recently sports betting started to be seen as an alternative form of investment, where decision-making is based on numbers and probabilities instead of relying on gut-feeling. In fact, “the rise of sports analytics, combined with better data and computational power, has led to the creation of companies that specialize in betting on sport. (...) They pitch sports betting as an alternative asset class, bringing potential for income and diversification” (Kucharski, 2016).

“The UK betting industry (...) is leading the way internationally with the implementation of a series of significant changes, [such as the] abolition of gambling tax for punters and legitimization of online betting. In addition to the importance of this industry for the economy *per se*, betting markets have received much attention in the academic literature due to their similarities to financial markets. In a seminal paper, Thaler & Ziemba (1988) were among the first to argue that betting markets may be better suited than financial markets when testing for efficiency. The main advantage is that bets have a well-defined period of life at the end of which their value becomes certain and this makes the testing of market efficiency far less complicated” (Vlastakis *et al.*, 2009: 427).

Besides that, notorious cases of successful punters and traders working on their own also started to emerge, like Paulo Rebelo or Billy Walters, for example. In “Ganhar com as Apostas Desportivas”, Rebelo (2012) states that data analysis and statistics are useful tools to predict the outcomes of sports events, and also that it is possible to obtain positive returns on betting just by running a quantitative analysis on both teams involved in a certain match. However, he also claims that in his view, a qualitative analysis (subjective interpretation of the news and factors that are very hard to quantify numerically, such as injuries or suspensions of key players) adds far more value and, in case of conflict, should outweigh the quantitative analysis.

According to the European Gaming and Betting Association (EGBA) report on sports betting, in 2012 regulated betting accounted for 58 billion dollars, 14% of the total global

gambling yield. At the time, European countries represented the largest share of the global betting market (41%), followed by Asia (39%). The EU fixed-odds sports betting gross revenue was of 1379 million euros per year (on average), between 2010 and 2012. In 2013, European Sponsorship Association (ESA) said that commercial sponsorship of the sport sector by gambling companies had become a significant source of sponsorship funding for sports organizations, with 74 deals closed in 2012. From 2009 to 2013, regulated gambling companies invested more than 60 million pounds in English Premier League clubs, and appeared as the main sponsor in some well-known European clubs' jerseys (like Bwin for Real Madrid and AC Milan, or Betclac for Juventus).

All around the world, different sports constitute an important part of every human society, and these numbers demonstrate the growing importance of betting companies in the sport industry, over the last decade. If the sports betting companies' revenue is growing, one can conclude that there is an increasing number of people betting. One can also conclude that the clear majority of punters lose money to the bookmakers in the long term, otherwise this would not be such an income generating activity. Very likely, this happens because most punters do not base their betting decisions on a scientific study of the variables that may affect the outcome of a match, but rather on their instinct and beliefs.

3.1- Forms of placing bets

3.1.1- Fixed-odds betting

Fixed-odds betting is one of the most popular forms of sports gambling. The bookmaker determines the odds for a given future event and takes the opposite side of every transaction. “The bettor is left with a take-it-or-leave-it decision: he can either hit the market quotes or refrain from participating. This is why bookmaker markets are sometimes called quote-driven markets, by analogy to the same setting in financial markets” (Franck *et al.*, 2010: 449-450).

In 1x2 market, the punters bet in one of the three possible outcomes of a match (home win, draw or away win). The market prices of these outcomes are usually presented as ‘decimal odds’ (e.g. a certain bookmaker offers an odd of 1.45 for Arsenal beating West Ham, at home), which represent the pay-out ratio of a winning bet (in this example, the bookmaker pays 45 cents for each 1€ bet, if the home team wins). Fixed-odds can be converted into probabilities ($1/\text{odd}$, as an odd is the inverse of probability), so in this

particular case, an odd of 1.45 for an Arsenal home win *versus* West Ham can be interpreted as if the bookmaker attributes a probability of approximately 69% to Arsenal winning the game ($1/1.45 = 0.689$). However, one must always have the bookmaker's overround in consideration (Franck et al., 2010).

“The sum of the probability-odds for the possible outcomes of an event is invariably greater than one, which is the mechanism through which bookmakers collect commission. The margin by which the sum of probability-odds exceeds one is known as the overround” (Forrest *et al.*, 2008: 159). The overround represents the specific expected profit for the bookmaker, if bets are distributed in such way that identical amounts in prizes are paid to the punters, whatever the match outcome. Once that it is very unlikely that bets are distributed in such manner, the overround is simply an estimation of the bookmaker's expected return (Constantinou & Fenton, 2013). For example, let's consider the odds offered by Betclie for the Premier League match Southampton vs Tottenham, that will take place in 21st January of 2018:

- Home win- 4.90 ($1/4.90 = 20,4\%$)
- Draw- 3.90 ($1/3.90 = 25,6\%$)
- Away win- 1.60 ($1/1.6 = 62,5\%$)
- $20,4\% + 25,6\% + 62,5\% = 108,5\%$, in this specific case the bookmaker's overround is 8,5%

“In order to obtain the [bookmaker's] prediction of the outcome, we assume that the overround is equally distributed over the outcome probabilities. Therefore, we obtain the market's 'implicit probabilities' by a linear transformation”: (Franck et al., 2010: 449)

- Home win- $(1/4.90) * (1 / ((1/4.90) + (1/3.90) + (1/1.60))) = 18,8\%$
- Draw- $(1/3.90) * (1 / ((1/4.90) + (1/3.90) + (1/1.60))) = 23,6\%$
- Away win- $(1/1.60) * (1 / ((1/4.90) + (1/3.90) + (1/1.60))) = 57,6\%$
- $18,8\% + 23,6\% + 57,6\% = 100\%$

The calculations for removing the overround are automatically done by an Excel spreadsheet called “True Odds Calculator”, which is available for free download at

<http://www.football-data.co.uk/>

Rebelo (2012) explains the concept of positive expected value: when the probability of a certain outcome converted into an odd is lower than the odd offered by the bookmaker. In his opinion, the only way of consistently making a profit in the long-term is being able to estimate the probability of an event's outcome in an accurate way, and consequently identify situations where the bookmakers' mistakes of evaluation can be exploited. This author also refers that finding value is less difficult if one specializes in certain leagues or markets, because extended knowledge is essential when the objective is beating the bookmaker, who has far more resources.

Among the Big 5 leagues (English Premier League, Spanish La Liga, Italian Serie A, German Bundesliga and French Ligue 1), Premier League is the one where the money from commercial and television rights is most equitably distributed for all its participants (Macário, 2017). For example, the 2016/2017 winner Chelsea received only 1,6 times more than the last placed Sunderland, which nonetheless still got 107,5 million euros (Jornal Económico, 2017). This allows that every club is able to sign quality players and, as a consequence of that, the general level of the competition increases, as well as the unpredictability of its matches.

In the beginning of each season, there are typically 6 candidates to the top 4, which grants qualification for the groups stage of Champions League (Arsenal, Chelsea, Liverpool, Manchester City, Manchester United and Tottenham), and apart from these, nearly all other teams might finish in the bottom 3, thus being relegated to the 2nd division of English football. Having this in consideration, the outcomes of Premier League's matches would be the most difficult and challenging to predict, hence we decided to specialize in 2 different betting markets of that league: 1x2 (bet on home win, draw or away win) and under/over X goals in a match (e.g. if there will be more than 2.5 goals in a certain match).

The fixed-odds bets can be classified in 2 different types: single or multiple. In the first case, the punter bets on one outcome of a single match, and if that outcome subsequently occurs, he or she is rewarded with the return implicit in the odd (as explained above). In a multiple bet, the punter bets separately on 2 or more outcomes from 1 or more matches, and if all the bets made are accurate, he or she is rewarded as in the following example:

A punter bets 5€ on the following outcomes:

- Manchester United beats West Ham away, at 1.75
- Manchester United will score more than 1.5 goals, at 1.60
- A draw in Barcelona vs Real Madrid, at 3.20

If all these outcomes occur, the punter will be rewarded with:

- $5€ * 1.75 * 1.60 * 3.30 = 46,20€$.

In other words, the punter is leveraging the return by multiplying the risks. Rebelo (2012) also addresses this topic, saying that although many professional punters devalue the usage of multiple bets, the possibility of leveraging is good, in itself. Even if this author does not work with multiple bets himself, he understands its value in the betting markets. Nowadays, it is possible to bet on sports events in real time (live bets), but the aim of this project is to produce accurate forecasts based on past data and subjective information, so this type of betting will not be considered. In order to minimize the risks, we decided not to consider multiple bets as well.

Lastly, some bookmakers offer the option of making a single bet on 2 outcomes, in a match. For instance, that Liverpool beats Leicester at home, and scores more than 3.5 goals. Because the risk is heightened, the odd is also higher (in most cases, higher than if one makes a multiple bet on the 2 same outcomes). We believe that these markets may offer some interesting opportunities to explore and therefore, our decision-making process will take them into consideration.

3.1.2- Exchange betting

“In recent years, person-to-person exchange betting has evolved as a different betting market structure. Here, individuals contract their opposing opinions with each other. On an online platform, they can post the prices at which they are willing to place a bet either on or against a given event. The latent demand for wagers is collected and presented in the order book, which displays the most attractive odds, with the corresponding available volume, in a canonical manner. Such a market design is often referred to as an order-driven market. The bettor has the choice to either submit a limit order and wait for another participant to match his bet or submit a market order and directly match an already offered

bet. As a result, there is a continuous double auction process taking place at the online platform. If two bettors with opposing opinions agree on a price, their demands are automatically translated into a transaction. After the bets have been matched, both of the individuals hold a contract on a future cash flow. The size of the cash flow is determined by the price of the contract, while the direction of the cash flow is tied to the outcome of the underlying event. The provider of the platform charges a commission fee, which is typically lower than the bookmaker's overround, on the bettors' net profits" (Franck et al., 2010: 450).

"Online betting exchanges have experienced a fast boom. The odds analysed in this paper are from Betfair, which is one of the most prominent bet exchange platforms. With a weekly turnover of more than \$50m and over two million registered users, Betfair accounts for 90% of all exchange-based betting activity worldwide. It has been online since 2000, and claims to process five million trades a day" (Franck et al., 2010: 450).

In the same document, the authors also explain that, at least in theory, bet exchanges should be able to generate accurate forecasts, for three different reasons:

1. The betting exchange works in a way that rewards the value of information. In this way, well informed traders should be able to generate higher average returns than uninformed traders, because they are competing against each other.
2. The betting exchange encourages the truthful exposure of information. Based on the information gathered, traders put money at stake, and by doing it, are revealing their expectations of the outcome's probability.
3. Through the price mechanism, the betting exchange delivers a proficient algorithm, which is capable of gathering, storing and aggregating different information in a dynamic way.

In practice, studies have demonstrated that bet exchanges provide very accurate forecasts. In one of those studies, Croxon & Reade (2008) used high-frequency Betfair data in order to test the efficiency for the arrival of goals, and concluded that prices incorporate the relevant news swiftly and fully, indicating a high level of efficiency on Betfair odds.

3.2- Legal context of the betting markets in Portugal

The greatest difficulty anticipated is the lack of competition in the Portuguese online

Forecasting Football Outcomes to Invest in Betting Markets

betting business. The country's online betting is regulated by the decree-law n° 66/2015, which does not tax the punters winnings (unlike the financial markets), but taxes the bookmakers in 8 to 16% over the volume of bets made (won or lost, which means that the bookmakers must pay prizes and taxes for bets won by the punters). In contrast, the UK government taxes the bookmakers in 15% of their profits (which are defined as stakes received less winnings paid out), in the fixed-odds' case, and as in Portugal, does not tax the punters' winnings at all (HM Revenue & Customs, 2017).

This taxation policy forces the bookmakers to offer lower odds, otherwise they would not be able to run a profitable business. (Apostas Online, 2017). For example, on January 26 of 2018, Bet365 (UK based bookmaker) was offering an odd of 1.66 for Manchester United win or draw away vs Tottenham (to be disputed on January 31 of 2018), while Betclie.pt offered 1.48 (the difference is around 11%). Besides that, "only 39% of Portuguese punters' online betting turnover is spent with local operators, far below the desired 80-90% channeling rate that most regulated online markets strive to achieve" (Stradbroke, 2017).

Physical stores specialized in sports betting are a more traditional alternative to online fixed-odds betting (like the betting shops that exist in High Street, London). In Portugal, physical betting is regulated by the decree-law n° 67/2015, which grants the State the exclusive exploration of the physical betting. In its turn, the State concedes the monopoly of physical betting to Santa Casa da Misericórdia (a charity organization owned by the State), who introduced Placard in 2015, a physical way of placing bets in a bulletin shape, which is sold by authorized distributors, such as stationery shops and other retail stores that are not specialized in betting. The state justifies this exclusive concession of the physical betting to Santa Casa with the argument that, in this way, it is easier to assure the protection of the punters' interests (e.g. to make sure that he or she receives the prize in case of winning the bet), as well as controlling underage gambling and criminal activities related with sports betting, such as money laundering (Batista, 2017).

Placard offers the possibility of betting in eight different sports (football, tennis, basketball, handball, volleyball, rugby, ice hockey and American football) and, in spite of its undeniable popularity and commercial success (over 1 200 000 people bet on Placard, in two years), this game is very limitative to the punters, in the sense that it only covers four different markets: 1x2, 1x2 at half time, European handicap and under/over

2.5 goals in a match ¹. Betting exchanges are not even considered in the current Portuguese legal framework, and so our project is limited to fixed-odds betting markets.

Many important bookmakers (e.g. Bwin, Betfair, Bet365, Ladbrokes, etc.) consider this legal environment unattractive and prefer not to ask for the permit that would allow them to operate in the Portuguese market (Silva, 2014). Currently, Portugal has only four licensed online bookmakers (Bet.pt, Casino Portugal, ESC and Betcltic), while Denmark, considered the best regulated market in Europe, has forty (Madremedia/Lusa, 2017). This lack of competition between the bookmakers created an oligarchic market that is harmful to the punters interests, because bookmakers do not feel much pressure to reward the risks assumed by the punters with valuable odds (once that there are only 4 possible choices). In March of 2018, a new online bookmaker obtained the permit to operate in Portugal (Correio da Manhã, 2018). However, our project was already in progress when this happened, and so we decided not to consider A Nossa Aposta in this work. For the same reason, we also excluded Placard.pt, which is the online bookmaker owned by Santa Casa da Misericórdia (as mentioned above), that entered the Portuguese online betting market just before the beginning of the FIFA World Cup 2018 (Cortim, 2018).

On December 6 of 2017, the Remote Gambling Association (RGA) announced the conclusions of a report made to determine how Portugal's regulated online market (also includes casino games and poker) was doing since the first locally licensed site began its activity, in June of 2016. This report was based on a survey, which targeted a sample of Portuguese online gamblers, and its results demonstrated that 68% of the respondents were still using gambling sites not allowed to operate in Portugal. From these punters, 38% gambled solely through unlicensed websites, while 30% admitted using both licensed and unlicensed online gambling platforms. When the respondents were asked why they still use unlicensed websites, the main reason mentioned was the wish to access better odds than the ones offered by the Portugal-licensed gambling platforms. The RGA director of government relations, Pierre Tournier, claimed that the Portuguese regulated market is "clearly failing to combat the unregulated market and change is much needed to make the regulation work" (Stradbrooke, 2017).

¹ For further information, please consult <https://www.jogossantacasa.pt/web/SCNoticias/newsDetail?contentId=22534.0&headerTitle=Comunicados%20de%20Imprensa>

2017 was the first full calendar year since the regulation of the Portuguese online gambling market. The SRIJ (Serviço de Regulação e Inspeção de Jogos do Turismo de Portugal) official report on that year states that the market's total revenue was of 122,6 million euros, from which the government took 54,3 million euros (44,2%), essentially from the market's punitive tax on sports betting turnover. The total sports betting revenue hit 68,1 million euros in 2017 (55% of the overall online gambling market), and football accounted for more than 75% of the wagers, followed by tennis with 12,4%, and basketball with 8,5% (Stradbroke, 2018).

3.3- Purpose of the project

The core purpose of this project is to create an uncomplicated method, that can be understood and used by regular punters to obtain positive returns consistently, in the long-term. This system intends to forecast the outcomes of football matches with accuracy, and subsequently, to identify profitable opportunities in the Portuguese online fixed-odds betting markets (1x2 and Under/Over X goals in a match).

The forecasting of the football matches will be done through a combination of analytic methods and subjective interpretations of the aspects that can affect the outcomes of those matches, with the aim of supporting our betting decisions. The decisions will be made by comparing the probabilities calculated for a certain outcome with the odds offered by the market, for that same outcome, using Rebelo's (2000) concept of positive expected value as the main point of reference.

Finally, we will also record the odds of the bets made in the Portuguese markets and compare them with the ones provided by a British online bookmaker, that will serve as a benchmark to evaluate the efficiency and competitiveness of the Portuguese online fixed-odds betting markets.

4- Literature review

4.1- Statistical models

In the context of this project, we will not try to develop our own statistical models to predict the outcomes of football matches, but simply to find good already developed models and use them as a tool (among others) to forecast football events. Goddard (2005) mentions two methods of modelling football outcomes: the first one is modelling the probabilities for the goals scored and conceded by each team (and therefore the probabilities for the outcomes, in an indirect way), and the second one is modelling the outcomes' probabilities (home win, draw, away win) directly. In the same document, the author compared both methods and found them to yield similar performances in terms of their ability to predict match results.

4.1.1- Poisson regression models

For the first method, although some articles have previously rejected the goal scoring process as Poisson (e.g. Reep *et al.*, 1971), Maher (1982) assumed independence between the scores of both teams and, from the past goal-scoring data of the teams disputing the top 4 divisions of English football, estimated attacking strength (home and away) and defensive weakness (home and away) parameters for every team (4 parameters for each team), through the method of maximum likelihood.

After comparing the expected goal scoring frequencies provided by the model with the actual ones, goodness-of-fit tests concluded that, in spite of some minor but consistent discrepancies, the Poisson regression model fitted relatively well into the goal scoring process. In order to correct the effect of those small differences, which were explained by the fact that the independence assumption was not completely valid, the author proposed a bivariate Poisson regression model, with a correlation coefficient between both teams scores of about 0.2. This modification improved the fit in a significant way.

Another interesting finding was that 2 parameters for each team were enough to describe the teams' attacking and defensive behaviour, because "although home ground advantage is a highly significant factor, it applies with equal effect to all teams, and each team's inherent scoring power is diminished by a constant factor when playing away" (Maher, 1982: 113). It is important to highlight that this model was still not able to produce forecasts for the outcomes of future matches.

Starting from Maher's work, Dixon & Coles (1997) developed the Poisson regression model into one that was actually capable of predicting future match results, with some changes for enhancing the model's predictive ability. The first one was about the parameters of the model, that were calibrated to reflect the differences of quality between divisions, which allowed to estimate the outcomes of English cups (FA Cup and League Cup) matches. In the context of this project, we will not attempt to forecast cups events.

The second change is related with the fact that "in reality, a team's performance tends to be dynamic, varying from one time period to another (...) In particular, a team's performance is likely to be more closely related to their performance in recent matches than in earlier matches" (Dixon & Coles, 1997: 272). To account for this, the authors considered to formalize a stochastic development of the model's parameters, however, because this method requires that new parameters for each team would be estimated at the time of making a bet, they preferred a more simplistic approach. It was assumed that "parameters are, in a loose sense, locally constant through time and that historically information is less valuable than recent information" (Dixon & Coles, 1997: 272). Having this in consideration, the authors modified the likelihood function into a "pseudolikelihood", with a weighting function Φ which allowed for the historical data to be down-weighted in the likelihood to a degree of choice.

4.1.2- Ordered probit/logit regression models

When it comes to modelling the probabilities for the outcomes directly, Goddard & Asimakopoulos (2004) used a dataset of 10 seasons fitted into an ordered probit regression model to forecast English football results. The authors justified the emphasis on match outcomes rather than scores with simpler estimation procedures, and also with the argument that fewer parameters were required.

The long-term performance of a team was modelled by building win-loss ratios from the outcomes of the teams' past matches, which were measured from the match results data through distinct time intervals, with the objective of understanding how a team's past performance during different lengths of time explains its future performance, on average. This was achieved by using those win-loss ratios (which were built with data relative to England's top 4 divisions, from 1989 to 1998) as the explanatory variables of an ordered probit regression model. The most recent match data appeared to be the most relevant in predicting future match results. In fact, it seemed that the impact of the historical

performance in the estimation accuracy diminished as a function of time in relation to the predicted match.

Goddard & Asimakopoulos measured short-term performance levels through discrete indicators, which were then inserted into the ordered probit model. Those indicators reflected the team's results in recent matches (a home win was encoded as 1, a draw as 0.5 and an away win as 0), and were used as explanatory variables together with the long-term performance ratios. The short-term performance indicators importance to the forecast accuracy decreased in the same way as the long-term performance level indicators: the most recent matches have more influence in the forecasting of future football matches. Empirical test results also demonstrated that generally, long-term indicators contributed more to the prediction accuracy than short-term indicators.

This was the first model that quantified the predictive quality not only from data relative to past outcomes, but also from other explanatory variables: importance of the match for each team (e.g. avoiding relegation), participation in cup competitions and geographical distance between the home grounds, all of which were tested positive and significant at the 1% level.

Hvattum & Arntzen (2010) used the ELO system to rate each team's current strength. This system "was initially developed for assessing the strength of chess players, but has been widely adopted in various other sports, including association football" (Hvattum & Arntzen, 2010: 461). After computing the ratings for every team disputing the English football top 4 leagues (calculated from historical match results, they are updated at the end of each game), the rating difference between both teams involved in a match (acts as a measure of relative performance) was used to derive covariates, which were then fitted into an ordered logit regression model (works in the same way as the ordered probit, but assumes logistic distribution instead of standardized normal distribution), in order to estimate 1x2 probabilities.

"Hvattum & Arntzen compared the forecasting accuracy of the ELO model with the one of Goddard & Asimakopoulos' model. Based on the comparison with informational loss and quadratic loss, the authors concluded that the ELO model produced more accurate estimates for probability distributions. However, they were not able to determine which model would be more profitable in betting. In fact, neither of the models were able to produce abnormal returns when used in betting simulations. The authors also suggested

that the differences in forecasting accuracy between the two approaches are likely to become smaller as sample size increases. Hence, they concluded that the ELO ratings are a more efficient way of encoding past results when only short time periods are available to calibrate the model” (Heino & Sillanpää, 2013: 38).

4.2- Fixed-odds betting markets efficiency

“In general terms, the theory of efficient markets is concerned with whether prices at any point in time fully reflect available information” (Fama, 1970: 413). The author describes 3 types of information-related market efficiency: in a certain market, there is strong-form efficiency when “all public and private information is incorporated in the security prices and hence no one can consistently achieve abnormal returns in the market” (Heino & Sillanpää, 2013: 11), semi-strong-form efficiency is when “all relevant public information is quickly incorporated into security prices, and hence the opportunities for abnormal returns disappear quickly” (Heino & Sillanpää, 2013: 11), and a market has weak-form efficiency “as long as consistent abnormally large returns cannot be achieved by using historical price [in this case, odds]” (Heino & Sillanpää, 2013: 11).

Kuypers (2000) defines abnormal returns as the ones that are better than the bookmakers’ take (the overround). According to the author, if one uses all the publicly available information to estimate probabilities for the outcomes of a match and then is able to use those estimations to get abnormal betting returns, it seems logical to conclude that the markets at stake are semi-strong efficient. In the same publication, it is also argued that from a theoretical point of view, even negative returns can be interpreted as an evidence of semi-strong efficiency. For example, assuming that all the bookmaker’s overround rates are equal and have a constant value of 9%, negative betting returns of -8% would indicate that the punter has developed a better ability to estimate probabilities using publicly relevant information than the market. Constantinou & Fenton (2013: 1) explain that “a gambling market is usually described as being inefficient if there are one or more betting strategies that generate profit, at a consistent rate, as a consequence of exploiting market flaws”, and Buraimo *et al.* (2013: 182) “report striking evidence of semi-strong efficiency in the UK fixed-odds football betting market using a reputable newspaper tipster which offers probabilities of match outcomes rather than simple result indicators.”

An example of a well-documented inefficiency in the fixed-odds betting markets is the

favourite-longshot bias, which is a type of systemic error where favourites tend to win more often than what is implied by the odds (Cain *et al.*, 2000). The authors discovered that this bias existed in the 1991/1992 odds of English football leagues, and Vlastakis *et al.* (2009) found the same bias in a number of big online bookmakers. Based on the odds offered by several bookmakers over a period of 7 football seasons (from 2005/06 to 2011/12), Constantinou & Fenton (2013) demonstrated the existence of a clear home advantage bias, whereby the returns produced by betting on home wins were substantially higher in comparison with the ones obtained by placing bets on away wins. In fact, the results of this study showed that this bias is almost equally as strong as the favorite-longshot bias. By using a dataset of 79 446 matches, disputed in 21 leagues of 11 different European countries (from 2000 to 2001), Direr (2011) claims that would have been possible to generate positive rates of return of 4.45% if best odds across all the bookmakers were selected, and 2.78% if mean odds were selected, by following a strategy of backing overwhelming favourites with odds of 1.19 (odds from 1.15 to 1.24 would have generated positive returns).

Forrest *et al.* (2005) affirmed that, at the time, bookmakers set the odds approximately a week before the matches and, although they reserved the right to adjust the odds until the kick-off, they rarely did it, regardless of betting volumes and new information (e.g. an injury of a key player). This allowed informed punters to exploit those market inefficiencies, which could bring serious financial consequences for the bookmakers. More recently, Buraimo *et al.* (2013: 175) claimed that “such a position became untenable with the emergence of internet betting combined with increased global competition in betting markets”, and classified “the adjustment of betting odds on football matches by bookmakers up to kick-off” as a “relatively recent phenomenon”.

Levitt (2004) demonstrated that by intentionally setting odds that are not in consonance with their outcome probabilities estimations, the bookmakers can exploit the punters' sentiments and preferences, consequently increasing their profits. The author also explains that the bookmakers must be careful in doing so, because well-informed punters, who are able to estimate accurate probabilities for the possible outcomes of the matches, can notice the distortions and develop strategies in order to use those distortions in their benefit, generating positive returns.

4.2.1- Arbitrage opportunities

“An arbitrage opportunity is simply an opportunity whereby profit is guaranteed on the basis of a negative profit margin which results by combining the odds published by the various bookmaking firms. In particular, arbitrage opportunities depend on two factors: a) the divergence in outcome probabilities between bookmaking firms and b) the profit margin by each bookmaker. Negative profit margin is simply a scenario where a set of probabilities is found (for a single match instance) in which the sum of the probabilities within that set is < 1 . Hence, profit for the bettor can be guaranteed if the bets are placed such that the return is identical whatever the outcome” (Constantinou & Fenton, 2013: 3).

The existence of arbitrage opportunities is a classical sign of weak-form efficiency, as it is possible to exploit those opportunities in order to obtain a sure profit just from examining the historical prices, which in this case are the odds offered by the bookmakers. These opportunities were detected in several publications, namely Buraimo *et al.* (2013) Vlastakis *et al.* (2009) and Constantinou & Fenton (2013), however, as Buraimo *et al.* explain, in order to exploit arbitrage one must have the possibility of choosing between a vast number of bookmakers (as it happens in the English fixed-odds betting market). As a consequence, in this work there will be no further discussion about this issue (there are only 4 bookmakers operating legally in Portugal).

4.3- Home advantage

Home advantage is defined by Courneya & Carron (1992: 13) as “the consistent finding that home teams win over 50% of the games played under a balanced home and away schedule”, and it is a widely regarded matter in the literature related with sports events. “Home advantage in football has long been established as an important factor in determining the result of a game. Its existence is certain to affect the attitude of players, coaches, referees, fans and the media alike. Surprisingly, and despite over 25 years of research, the precise causes of home advantage and the way in which they operate are still not well understood” (Pollard, 2008: 12). In the same work, the author reviews all the hypothesis raised in order to explain this phenomenon: crowd effects, travel effects, familiarity, referee bias, territoriality, specific tactics and psychological factors.

Forecasting Football Outcomes to Invest in Betting Markets

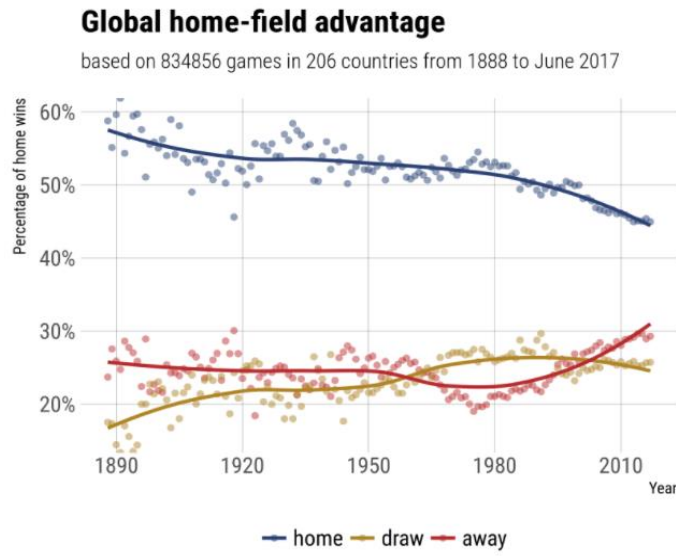


Figure 1- Percentage of home wins, draws and away wins worldwide, from 1888 to 2017 (Schoch, 2017)

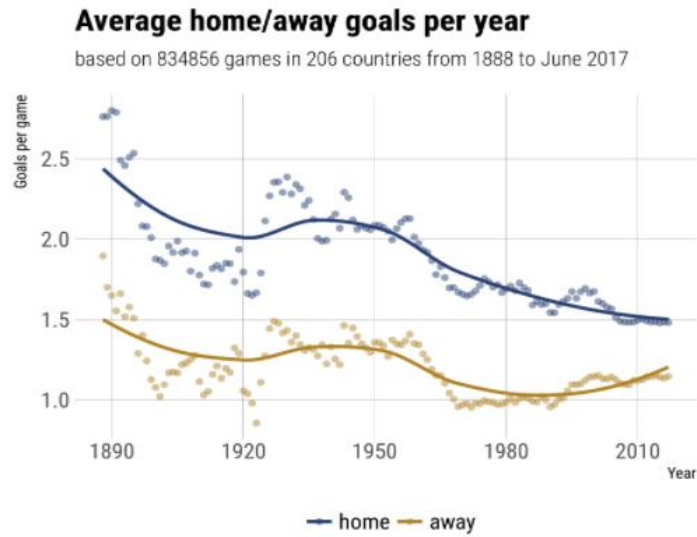


Figure 2- Average of home and away goals scored per game worldwide, from 1888 to 2017 (Schoch, 2017)

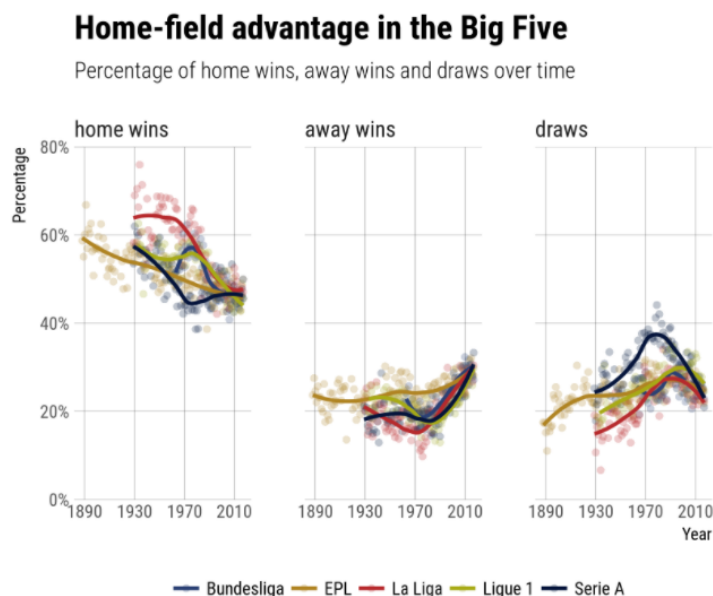


Figure 3- Percentage of home wins, draws and away wins in the big 5 leagues, from 1888 to 2017 (Schoch, 2017)

When it comes to crowd effects, the connection between home advantage and crowd size is somewhat questionable, because this advantage works even in matches with very little attendance (Pollard & Pollard, 2005). Besides that, the proximity and intensity of support can also influence what is happening on the pitch (Heuer & Rubner, 2009). It was never understood if crowd effects have more positive influence over the home team or negative influence over the away team (Pollard, 2008).

Likewise, there is no conclusive statistical evidence for a travel effect inconvenience to the away team, due to the physical and psychological fatigue caused by a trip to another team's home ground. Distance between both teams' stadiums has been researched as a factor of disadvantage for the away team (both within the same country and internationally), with paradoxical results (Pollard, 2006). Nevertheless, research consistently discovered reductions of home advantage in the case of local derbies, where no travel is required (Seckin & Pollard, 2008).

Through the analysis of the frequency of disciplinary cards and other referee decisions, investigators also gathered significant evidence of a bias related with referee decisions favouring the home team (Thomas *et al.*, 2006). However, until now researchers were unable to establish the cause for this bias.

Forecasting Football Outcomes to Invest in Betting Markets

The home team familiarity with the environment surrounding the match (e.g. stadium, locker rooms, etc.) should also be considered as an important part of the home advantage. Although familiarity is, by nature, difficult to investigate, some findings indicate it as a probable aspect of home advantage. For example, playing on a synthetic field or on fields with uncommonly large or small dimensions has been tested as an advantage for the team who is familiar with those particular conditions (Pollard, 1986). Being familiar with certain climate conditions and altitude also has a positive impact on the home team chances of winning the match (McSharry, 2007).

Humans beings react aggressively to a feeling of territory invasion (in the same manner as irrational animals do). This behaviour is called territoriality, and it seems logical to assume that this may play a role in the home advantage phenomenon. Scientists were able to demonstrate that home players increase their hormone activity before a game (Neave & Wolfson, 2003), and also that teams grounded in countries or regions isolated and/or with a history of conflict have higher levels of home advantage, probably due to an enhanced sense of territoriality (Pollard, 2006).

Very often, managers change their tactical approach to the match, depending on whether it is played at home or away. If the away team choose to approach the match in a more cautious manner, possibly they are giving a territorial and psychological advantage to the home team (Pollard, 1986). There are no solid scientific studies that relate tactics with home advantage, despite a study attesting increased home advantage in the second leg of European cup matches proposed tactical aspects as one of the explanations (Page & Page, 2007).

If managers and players are conscious of the home advantage phenomenon, it can affect their mental attitude before and throughout the match. It is possible that psychological factors are the main cause for home advantage, because even though there may be real motives to explain this advantage, they are being heightened by the players' beliefs, creating a self-perpetuating phenomenon (Pollard, 1986). Neave and Wolfson (2005) did a review of the home advantage in football, from a psychological point of view, and concluded that at the end of the day, it is what is happening in the minds of referees, managers and players that defines their actions, and consequently the outcome of the match, as well as the importance of home advantage in that outcome. The only scientific work about the psychological condition of footballers in relation to home advantage

seems to be an investigation conducted on five players (Waters & Lovell, 2002).

In his book, Rebelo (2002) claims that there is a slight correlation between the home advantage and the position of a certain team in the league table. According to this author, weaker teams are more dependent of the home advantage to win than stronger teams, once that the latter are more consistent. This is the main conclusion of a study that he did for a statistics course, in the year of 2005 (that same study is fully available at <https://www.academiadasapostas.com/blog/2011/01/o-factor-casa-no-futebol>).

“One of the problems about researching home advantage in football is the fact that the likely causes outlined above will be operating together, each interacting with the other in ways that will be difficult to investigate, isolate and quantify. (...) Thus a researcher will need to develop a strategy which either takes a multivariate approach, or which carefully controls for possible confounding variables that are not the main focus of the study” (Pollard, 2008: 13).

4.4- Money management methods

4.4.1- Fixed-stake

In this money management method, the punter defines a percentage of the initial money to invest in each bet, independently of the risk taken or the reward offered by that bet. Rebelo (2012) recommends this method due to its simplicity, and also because it assures that the punter will never lose all the initial money. Simultaneously, it allows the punter to exponentially increase that initial percentage, as his ability of betting accurately is developed. One possibility is to redefine the initial percentage at the beginning of the following month, based on the growth rate of the initial money. For example, if one begins January with 100€ and has defined a 10% stake for each bet, at the beginning of February, if the initial money grows to 110€, then the percentage at stake can be adjusted to 11% until further review, and so on.

4.4.2- Martingale

Mansuy (2009) reviewed in detail the origins and the various senses of the word martingale, with their respective etymologies, in mathematics, gambling, technology, and vernacular language. From the gamblers perspective, the word goes back to the beginning of the 18th century, when it was used to define a strategy of always betting as much as it

covers everything that was lost until that point, with the objective of recovering the all the losses in just one move. Rebelo (2012) explains that, although in theory this is a perfect method, as it virtually assures that anyone can make money in betting, he does not recommend it because, in his view, what makes a punter win is the accuracy of his bets, and not the money management method. If it is true that a good punter can lose all the initial money for not having a good money management method, a bad punter will not make money just for using a good money management method. According to this author, if a punter cannot make money betting a fixed stake, he will not be able make money consistently just by applying a martingale. On the contrary, a bad punter that falls into a negative spiral of results would be multiplying the risk many times just to break even.

4.4.3- Kelly criterion

“The Kelly criterion (Kelly, 1956) is a vital tool in the armory of both portfolio investors and gamblers. By maximizing logarithmic utility – simultaneously minimizing the risk of ruin – Kelly provided the formula that gamblers with perfect probabilistic knowledge must use to grow their bank at the largest expected rate” (O’Shaughnessy, 2011: 2).

“The basic Kelly criterion for a single option on a regular betting market gives the Kelly Bet B as:

$$B = \frac{Mp - 1}{M - 1} \quad (1)$$

where M is the team’s market price [odd] and p is the gambler’s presumed probability of the team winning. B is expressed as a percentage of the bettor’s bankroll, and a bet should be placed if $Mp > 1$. The formula is derived by maximizing \log (expected bank) with respect to the bet proportion B ” (O’Shaughnessy, 2011: 2).

4.5- The importance of subjective factors

Forrest *et al.* (2005: 562) affirm that “it seems highly likely that odds-setters’ prices are influenced by information of this kind. If the subjective information is used effectively, it may contribute positively to the odds-setters’ forecasting performance”. Constantinou *et al.* (2012: 1) presented a Bayesian framework “[graphical probabilistic belief network

that represents the conditional dependencies among uncertain variables, which can be both objective and subjective] for forecasting Association Football matches [outcomes] in which the subjective variables represent the factors that are important for prediction but which historical data fails to capture”. “Pi-football” model produces match forecasts “by considering generic factors for both the home and away team” (Constantinou et al., 2012: 5). Those factors are: 1) relative team strength, 2) recent form, 3) psychological factors and 4) fatigue.

“Components 2, 3 and 4 are predominantly dependent on subjective information. They are used to revise the forecast from component 1. The outcome of each of the components is mutually summarized in a single value (considering both teams) which we describe as ‘subjective proximity’. The subjective proximity is measured on a scale from 0 to 1. A value equal to 0.5 indicates no advantage for either of the teams; a value less than 0.5 indicates an advantage for the home team, while a value greater than 0.5 indicates an advantage for the away team” (Constantinou et al., 2012: 5).

For evaluating the performance of the model, the authors used both accuracy and profitability measurements, and concluded that the subjective information improved the forecasting ability of the model in a significant way. In terms of accuracy, they concluded that the precision of quantitative forecasts was significantly inferior to the bookmaker’s odds implicit probabilities, and also that subjective information improved the forecasts in a way that matches the bookmaker’s estimations. This suggest that the bookmakers also use information that is not captured by the standard statistical football data available to the public, as suggested by Forrest *et al.* (2005).

For the 380 Premier League matches of the 2010/2011 football season, it would be possible to obtain profitability rates from 2.87% to 9.48% (at standard discrepancy levels of 5% between the outcome’s probabilities provided by the model and the odds offered by the multiple bookmakers considered), and from 8.86% to 35.63% (at higher standard discrepancy levels, of 8% to 11%), respectively. Constantinou et al. (2012: 17) claim that, at the time, “no other published work appears to be particularly successful at beating all of the various bookmakers’ odds over a large period of time, which highlights the success of pi-football”.

4.6- Conclusions about the literature review

After carefully reviewing the existent literature about sports betting, it is possible to take some useful conclusions, which may help us in the development of a user-friendly method to achieve positive returns on the Portuguese online fixed-odds betting markets. Previously, we understood the importance of specialization to increase our chances of beating the odds in a consistent way and therefore, we decided to concentrate in the markets 1x2 and under/over X goals, of the English Premier League matches ². We are going to support our investing decisions with probabilities estimated by a regression model, and also with descriptive statistics about the Premier League teams. Since we decided to bet on both 1x2 and under/over X goals in a match, the first modelling method, estimating probabilities for the expected goals to be scored and conceded by each team (and therefore the outcomes probabilities, in an indirect way), seems to be the most appropriate. That said, we will use a Poisson regression model applied to the estimation of goal scoring probabilities.

When it comes to the efficiency of fixed-odds betting markets, we were unable to find any literature about the Portuguese context, hence we think that any disclosures about this topic would be a relevant contribution to the existing literature. In other countries, biases such as the favorite-longshot bias or the home advantage bias were consistently documented and therefore, we can attempt to develop betting strategies with the intention of identify and exploit those biases in the Portuguese online fixed-odds betting markets. If our strategies are able to generate positive returns, it can be seen as a strong evidence that the Portuguese online fixed-odds betting markets are at least semi-strong efficient, once that arbitrage opportunities (which indicate weak-form efficiency) are very unlikely, due to the short number of bookmakers functioning legally in the country.

The initial amount of money available for this investment project is relatively small (140,14€), and so we should use a fixed stake method (10% of the initial money on each bet), as it virtually assures that we will not run out of funds until the end of the football season. The importance of the qualitative analysis is highlighted by Rebelo (2012) and confirmed by Constantinou *et al.* (2012). Eventhough this analysis is mostly subjective, and consequently dependent of the interpretation and knowledge of each individual, the

² Please consult the 4th page of our work

latters defined an evaluation method that can be applied to our forecasting method.

5- Method

5.1- Poisson regression model

As mentioned previously, the aim of this work is not to build our own statistical model. From literature review, it was possible to conclude that the modelling method which estimates probabilities for the expected goals to be scored and conceded by each team, and indirectly, the probabilities for each possible outcome, was the one that made more sense to apply here. Our research also identified the Poisson regression model as the most widely used when the objective is to forecast the probabilities for the number of goals that might be scored in a certain match. Moreover, although statistical modelling is an integrating and important part of our work, it is not its core subject, but rather a supporting element to our decision-making system. In fact, this paper is more targeted to regular punters than to statisticians or academics. Consequently, this document will contain a somewhat simplistic explanation about the building process of the model and how it is able to produce accurate estimations for each possible score. If one has interest in the statistical details of this model in the full extent, we suggest the further consultation of Gardner's (2011) work, from where we retrieved the Poisson regression model used in this project.

5.1.1- Counting process and Poisson process definitions

“For each real number $t \geq 0$, let $N(t)$ denote the number of times a specified event occurs by time t . Note that $N(t)$ is a nonnegative-integer valued random variable for each $t \geq 0$. The collection of random variables $N(t): t \geq 0$ is called a counting process because it counts the number of times the specified event occurs in time” (Weiss, 2006: 686).

“A counting process $N(t): t \geq 0$ is said to be a Poisson process with rate λ if the following three conditions hold” (Weiss, 2006: 688):

- (a) $N(0) = 0$.
- (b) $N(t): t \geq 0$ has independent increments.
- (c) $N(t) - N(s) \sim \text{Poiss}(\lambda(t - s))$ for $0 \leq s < t < \infty$.

The second condition in the definition states that the increments in time must be

independent, although when it comes to football, this may not be the case. If a match is split into 2 increments, first half and second half, then it seems evident that, for several reasons (e.g. an important player gets injured during the first half), what happened in the first half might affect what will occur in the second half. It is even possible that football matches themselves are not independent, as many dependencies can affect the outcome of a match, such as differences in motivation between teams, resulting from success or underachievement in the near past, for instance. However, for the purpose of modelling football results, we will assume independence and take that a Poisson process $N(t)$ has a Poisson distribution (Gardner, 2011). A goodness-of-fit test on football scores' data will be performed later.

5.1.2- Poisson random variable

A random variable with Poisson distribution (X) can be defined as the number of occurrences in a continuous interval: timespan (in our case), volume, etc. This distribution can be used to represent different phenomena that happen in our daily lives (Curto, 2016), such as the monthly number of traffic accidents, the daily number of phone calls received by an individual or the total number of goals scored in a match.

In the context of our work, makes sense to consider a timespan of 90 minutes (plus the stoppage time given in each match), in which the goals are scored randomly by both teams (e.g. 1st goal scored by the minute 21, 2nd goal scored by the minute 43, 3rd goal scored by the minute 56, etc.). Furthermore, the values that the Poisson random variable can assume must be non-negative integers, as it is impossible to score -2 goals or 5,5 goals, for example. The probability mass function of a Poisson random variable is of the form:

$$p_X(x) = \exp(-\lambda) \frac{\lambda^x}{x!} \quad (2)$$

For values of $x = 0, 1, 2, \dots, n$ and where λ (expected value of X) is the only parameter which characterizes this distribution (Curto, 2016).

The probability mass function of a Poisson random variable demonstrates that the probability of the random variable X being a value x is equal to the right side of the equation. Since it is used to calculate probabilities for the number of times that an event

might happen, and it is clear that the sum of all the probabilities for the number of times that an event may occur is equal to 1, then (Gardner, 2011):

$$\sum_{x=1}^{\infty} \exp(-\lambda) \frac{\lambda^x}{x!} = 1 \quad (3)$$

“for any value of λ , where λ is the expected value of the variable. A sum is used here as a Poisson process and distribution uses discrete values, whereas in the continuous case, this would be an integral as the area under the curve would be needed. However, in the discrete Poisson case, it is only required that the number of occurrences at each time is recorded as the data points, which can then be viewed in a bar chart. The graphical representation of this will look different for differing values of λ . Lower values of lambda will skew the density to the left as the expected value of the random variable is low meaning that a higher percentage of low values will occur. (...) For a Poisson distribution, the value λ represents the expected number of occurrences of an event for the variable in question. A unique property of the Poisson distribution is that both the mean and the variance are equal to this value of λ ” (Gardner, 2011: 8).

5.1.3- Dataset

Rebelo (2012) claims that 3 seasons is the maximum timespan in which past data remains relevant for predicting future events, because data from beyond that period does not reflect the current strength of the teams involved in a match, due to the significant changes that occurred meanwhile, such as player transfers or managerial changes, for example. Dixon & Coles (1997) used that same data timespan in their model, and in Goddard & Asimakopoulos' (2004) case, the model measured the long time performance from the teams historical match data, which was only tested significant as a predictor of future results until two seasons beyond the current one. We find very interesting that the empirical allegations made by Rebelo correspond to the statistical testing results of Goddard & Asimakopoulos' work.

The 2015/2016 season of English Premier League was very atipic, with Leicester City being crowned champions (when they historically struggled to avoid relegation), Chelsea (champions in 2014/2015 and 2016/2017) ending the season ranked 9th and Liverpool

finishing in 8th place. On the other hand, from the end of the 2016/2017 season to the beginning of the 2017/2018 season, only 3 clubs experienced managerial changes:

- Southampton- Claude Puel left the club, Mauricio Pellegrino joined in
- Crystal Palace- Sam Allardyce was replaced by Frank de Boer (who was fired early in this season)
- Watford- Walter Mazzarri returned to Italy, and the ex-Hull City coach Marco Silva was selected as his substitute

Given these theoretical reasons, it seems more appropriate to consider in the Poisson regression model just the seasons 2016/2017 and 2017/2018, as this data timespan appears to be the most significant for measuring the teams' current attacking and defensive strengths. The historical data for both these seasons is fully available at <http://www.football-data.co.uk/>, and when it comes to the current season, the data is updated at each match day, as the season progresses. The descriptive statistics used in the forecasting reports are only relative to the 2017/2018 season, for practical reasons (the websites where those statistics are available present the information by season).

5.1.3.1- Goodness-of-fit

In order to use the Poisson regression model to estimate probabilities for the number of goals that might be scored in a given match, and indirectly, for the possible outcomes of that same match, it is imperative to assess if our dataset fits well into the Poisson distribution. To this end, we decided to generate random numbers in Excel, set to be Poisson distributed and with λ equal to the mean of our dataset (total number of goals / total number of matches). Afterwards, we compared the histogram of this distribution (left side) with another one representing the distribution of the actual number of goals scored, in a certain timespan (right side). In the histograms, the x axis represents the bins (in this case, the total number of goals scored in a certain match), while the y axis represents the number of observations for each bin. As the season 2017/2018 progressed, we needed to add new data at the end of each match day. In order to see if that new data was fitting well into the Poisson distribution, we decided to divide the data in the following pieces:

- 2016/2017 full season- $\lambda = 2,8$

Forecasting Football Outcomes to Invest in Betting Markets

- 2017/2018 full season- $\lambda = 2,67894737$
- 2016/2017 + 2017/2018 full seasons- $\lambda = 2,73947368$
- 2016/2017 + 1/4 of the 2017/2018 season- $\lambda = 2,76210526$
- 2016/2017 + 2/4 of the 2017/2018 season- $\lambda = 2,75087719$
- 2016/2017 + 3/4 of the 2017/2018 season- $\lambda = 2,75639098$

The results are presented below:

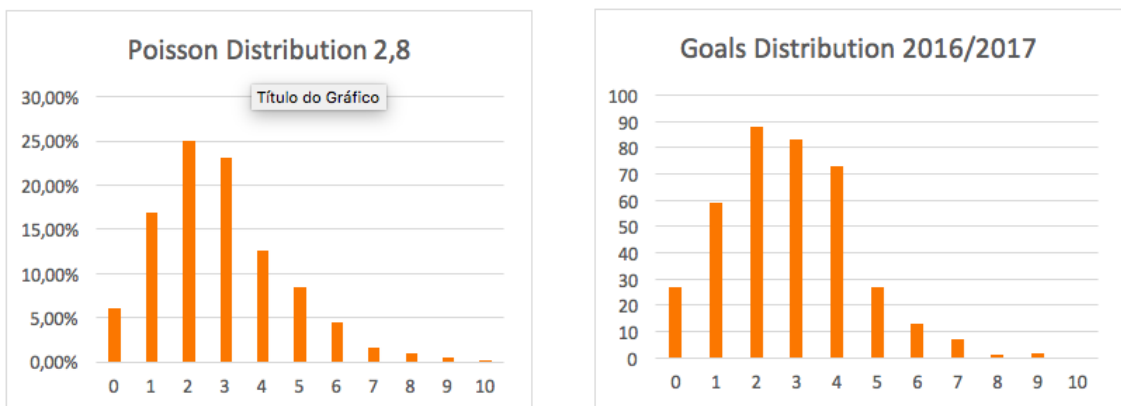


Figure 4- Random Poisson Distribution with $\lambda = 2,8$ vs Goals Distribution 2016/2017

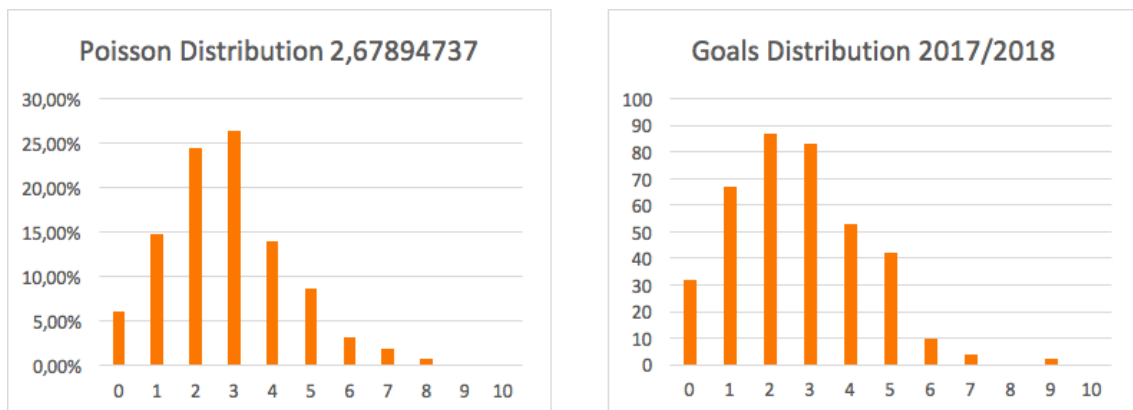


Figure 5- Random Poisson Distribution with $\lambda = 2,67894737$ vs Goals Distribution 2017/2018

Forecasting Football Outcomes to Invest in Betting Markets

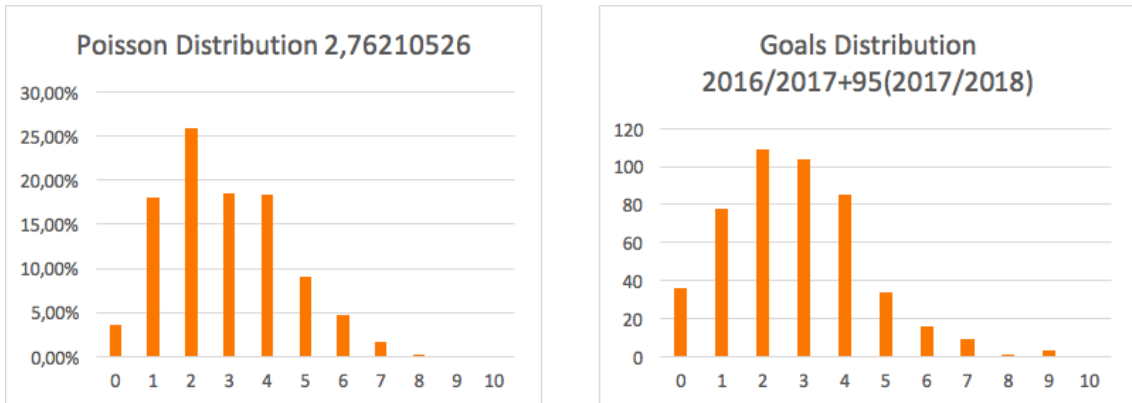


Figure 6- Random Poisson Distribution with $\lambda = 2,76210526$ vs Goals Distribution
2016/2017 + 1/4 of the 2017/2018 season

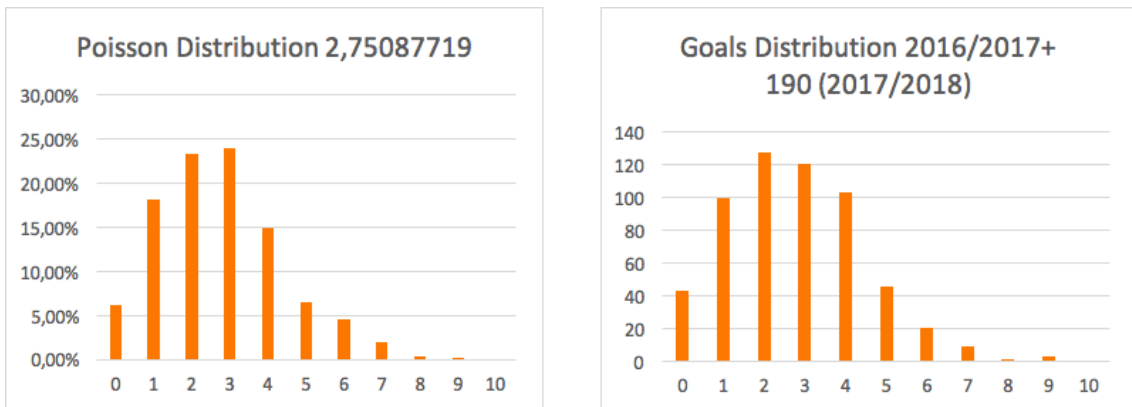


Figure 7- Random Poisson Distribution with $\lambda = 2,75087719$ vs Goals Distribution
2016/2017 + 2/4 of the 2017/2018 season

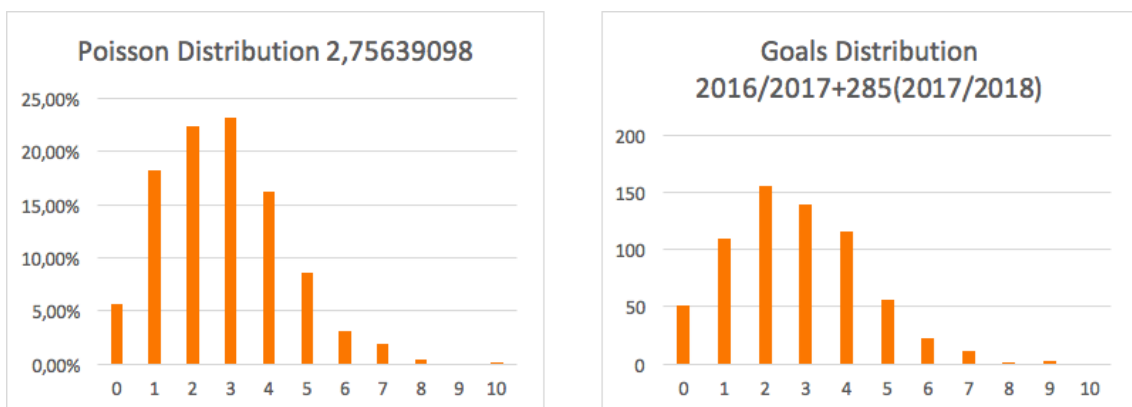


Figure 8- Random Poisson Distribution with $\lambda = 2,75639098$ vs Goals Distribution
2016/2017 + 3/4 of the 2017/2018 season

Forecasting Football Outcomes to Invest in Betting Markets

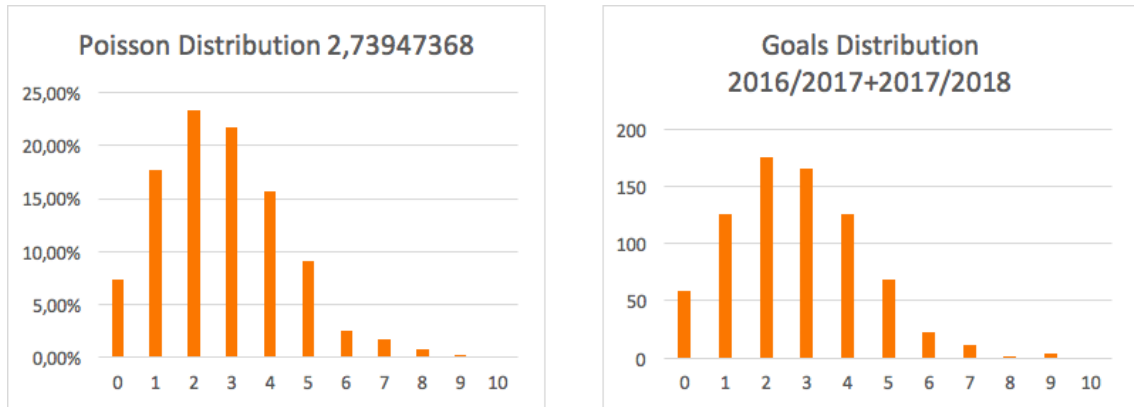


Figure 9- Random Poisson Distribution with $\lambda = 2,73947368$ vs Goals Distribution
2016/2017 + 2017/2018

As one can observe, in all cases the shapes of the distributions look very similar to each other. Another interesting observation is that the means of both seasons are also very similar to each other, which indicates that the teams' attacking and defensive abilities did not change much from one season to the other. However, in order to assure that our dataset fits into the Poisson distribution, a goodness-of-fit test must be performed.

In his work, Gardner (2011) uses the Chi-squared goodness-of fit test, as the Poisson distribution and the football data are both discrete. The aim is to test if the observed data follows a particular distribution (in this case, Poisson), and the hypothesis are formulated in this way:

- H0: The total number of goals in a match is a random variable with Poisson distribution
- H1: The total number of goals in a match has no Poisson distribution

Subsequently, the Chi-squared distribution test is calculated by this formula:

$$\chi^2_t = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (4)$$

Where:

- O_i = Observed frequencies of the bins, calculated by the Excel function

Forecasting Football Outcomes to Invest in Betting Markets

=FREQUÊNCIA

- E_i = Expected observations under the distribution in question (in this case, Poisson). The first step to obtain the values for E_i is to calculate the expected probability for each one of the bins, under the Poisson distribution, using the Excel function =DIST.POISSON. Then, it is necessary to multiply each result for the respective O_i .
- n = Number of bins (in our case, 11)

Knowing the degrees of freedom (10), it is possible to calculate the p-value for the test, using the Excel function =DIST.CHIQ.DIR, as the Chi-squared distribution is positive (skewed to the right). If the p-value > 0,05, we do not reject H_0 , and consequently, we can assume that the dataset has approximately a Poisson distribution. Once again, the results are presented below:

Bins	Frequencies (O _i)	Expected Probability Under Poisson Distribution	Expected observations (E _i)	(O _i -E _i) ²	(O _i -E _i) ² /E _i
0	27	6,08%	23,11	15,15	0,66
1	59	17,03%	64,70	32,51	0,50
2	88	23,84%	90,58	6,67	0,07
3	83	22,25%	84,54	2,38	0,03
4	73	15,57%	59,18	190,97	3,23
5	27	8,72%	33,14	37,71	1,14
6	13	4,07%	15,47	6,08	0,39
7	7	1,63%	6,19	0,66	0,11
8	1	0,57%	2,17	1,36	0,63
9	2	0,18%	0,67	1,76	2,61
10	0	0,07%	0,25	0,06	0,25
Total	380	100,00%			9,61
				P-value	0,47 > 0,05

Figure 10- Chi-squared goodness-of-fit test results for 2016/2017

Bins	Frequencies (O _i)	Expected Probability Under Poisson Distribution	Expected observations (E _i)	(O _i -E _i) ²	(O _i -E _i) ² /E _i
0	32	6,86%	26,08	35,03	1,34
1	67	18,39%	69,87	8,24	0,12
2	87	24,63%	93,59	43,43	0,46
3	83	21,99%	83,57	0,33	0,00
4	53	14,73%	55,97	8,84	0,16
5	42	7,89%	29,99	144,25	4,81
6	10	3,52%	13,39	11,49	0,86
7	4	1,35%	5,12	1,26	0,25
8	0	0,45%	1,72	2,94	1,72
9	2	0,13%	0,51	2,22	4,34
10	0	0,05%	0,18	0,03	0,18
Total	380	100,00%			14,24
				P-value	0,16 > 0,05

Figure 11- Chi-squared goodness-of-fit test results for 2017/2018

Forecasting Football Outcomes to Invest in Betting Markets

Bins	Frequencies (O _i)	Expected Probability Under Poisson Distribution	Expected observations (E _i)	(O _i -E _i) ²	(O _i -E _i) ² /E _i
0	36	6,32%	30,00	36,00	1,20
1	78	17,45%	82,86	23,66	0,29
2	109	24,09%	114,44	29,59	0,26
3	104	22,18%	105,36	1,86	0,02
4	85	15,32%	72,76	149,89	2,06
5	34	8,46%	40,19	38,35	0,95
6	16	3,90%	18,50	6,26	0,34
7	9	1,54%	7,30	2,89	0,40
8	1	0,53%	2,52	2,31	0,92
9	3	0,16%	0,77	4,96	6,41
10	0	0,06%	0,28	0,08	0,28
Total	475	100,00%			13,12
				P-value	0,22 > 0,05

Figure 12- Chi-squared goodness-of-fit test results for 2016/2017 + 1/4 of the 2017/2018 season

Bins	Frequencies (O _i)	Expected Probability Under Poisson Distribution	Expected observations (E _i)	(O _i -E _i) ²	(O _i -E _i) ² /E _i
0	43	6,39%	36,41	43,47	1,19
1	99	17,57%	100,15	1,32	0,01
2	127	24,17%	137,75	115,60	0,84
3	120	22,16%	126,31	39,85	0,32
4	103	15,24%	86,87	260,26	3,00
5	45	8,38%	47,79	7,80	0,16
6	20	3,84%	21,91	3,66	0,17
7	9	1,51%	8,61	0,15	0,02
8	1	0,52%	2,96	3,85	1,30
9	3	0,16%	0,91	4,39	4,85
10	0	0,06%	0,33	0,11	0,33
Total	570	100,00%			12,18
				P-value	0,27 > 0,05

Figure 13- Chi-squared goodness-of-fit test results for 2016/2017 + 2/4 of the 2017/2018 season

Bins	Frequencies (O _i)	Expected Probability Under Poisson Distribution	Expected observations (E _i)	(O _i -E _i) ²	(O _i -E _i) ² /E _i
0	51	6,35%	42,24	76,72	1,82
1	109	17,51%	116,43	55,25	0,47
2	155	24,13%	160,47	29,90	0,19
3	140	22,17%	147,44	55,31	0,38
4	116	15,28%	101,60	207,40	2,04
5	56	8,42%	56,01	0,00	0,00
6	23	3,87%	25,73	7,46	0,29
7	11	1,52%	10,13	0,75	0,07
8	1	0,52%	3,49	6,20	1,78
9	3	0,16%	1,07	3,73	3,49
10	0	0,06%	0,39	0,15	0,39
Total	665	100,00%			10,91
				P-value	0,36 > 0,05

Figure 14- Chi-squared goodness-of-fit test results for 2016/2017 + 3/4 of the 2017/2018 season

Forecasting Football Outcomes to Invest in Betting Markets

Bins	Frequencies (OI)	Expected Probability Under Poisson Distribution	Expected observations (EI)	(OI-EI)^2	(OI-EI)^2/EI
0	59	6,46%	49,10	98,02	2,00
1	126	17,70%	134,51	72,36	0,54
2	175	24,24%	184,24	85,34	0,46
3	166	22,14%	168,24	5,01	0,03
4	126	15,16%	115,22	116,18	1,01
5	69	8,31%	63,13	34,47	0,55
6	23	3,79%	28,82	33,91	1,18
7	11	1,48%	11,28	0,08	0,01
8	1	0,51%	3,86	8,20	2,12
9	4	0,15%	1,18	7,98	6,78
10	0	0,06%	0,43	0,18	0,43
Total	760	100,00%			15,10
				P-value	0,13 > 0,05

Figure 15- Chi-squared goodness-of-fit test results for 2016/2017 + 2017/2018

The results of the Chi-squared goodness-of-fit test are conclusive: once more, in all cases, the p-value is far above the smallest level of significance at which the null hypothesis would not be rejected (0,05). These values, in addition to the histograms, make us feel quite comfortable to assume that our dataset fits approximately into the Poisson distribution.

5.1.4- Poisson regression

After the assumption that our dataset fits into the Poisson distribution was validated by the Chi-squared goodness-of-fit test, the Poisson regression model can be used to estimate the probabilities for each possible score, in a given match.

The Poisson regression model is used as the “standard model for count data” (Cameron & Trivedi, 1998: 9), which refers to the number of times that an event occurs, and follows the same steps as a normal linear regression model ($y = X\beta + \varepsilon$), where the scalar variables y and the explanatory variables X are used to estimate the parameters β . However, it is slightly different as the values for count data need to be positive, and consequently, exponentials need to be used in the Poisson regression model’s case (Gardner, 2011).

The Poisson regression model comes directly from the Poisson distribution by using the same mass function and variables that describe the occurrences of the events. In this model, y_i are the scalar dependent variables which will correspond to the number of times that the event in question occurs, and x_i is a vector of linearly independent regressors, that is, a variable which has an effect on the values that y_i can assume (Cameron & Trivedi, 1998). As the same authors explain (1998: 21), “the standard estimator for this model is the maximum likelihood estimator”, and so the parameters β are estimated using

the method of maximum likelihood.

5.1.4.1- Applying the Poisson regression model to football data in R

As we explained in 5.1.3, the football data used in the Poisson regression model consists on the outcomes of all Premier League's matches played between 2016/2017 and 2017/2018. The raw historical data was retrieved from <http://www.football-data.co.uk/>, and when it comes to 2017/2018, was updated at the end of each match day, as the season progressed. In 5.1.3.1, we tested our data as fitting closely into the Poisson distribution and therefore, usable in a Poisson regression model.

Gardner (2011) goes further by using the Poisson regression model to simulate a whole season from the estimated parameters, and also to predict a future league table. The author utilizes the statistical program R, having developed a code that fully operationalizes all the steps necessary to the progression of his work. In the context of our research, we only need to use the Appendixes B ("Parameter Estimation") and C ("Probability Table") of his R code, and so we will not consider all the other segments that are not directly related with obtaining the probabilities for each possible score, in a given match. The parts of the R code used in this project were retrieved from <http://www1.maths.leeds.ac.uk/~voss/projects/2010-sports/Football.R>.

First, we set a working directory in R ("R_tese"), where we created an Excel workbook ("Data_MSc"), that contains a spreadsheet ("Folha4") in which we organized the data in 4 different columns:

- Home team name
- Away team name
- Home team final score (FTHG)
- Away team final score (FTAG)

The following step was to load all the relevant R packages, including "openxlsx", which allows the R to read Excel files. After that, we imported the data from "Folha4" to our R workspace.

Y corresponds to the data and, consequently, to the scores of the teams involved in our dataset's matches. As Y_i is a vector, it can be created by including all of the teams' scores

in it. Having this in consideration, if the data contains g number of matches, then the length of Y_i will be $2g$, as there are 2 bits of data for each match, the home team's score and the away team's score. It is clear that $Y_i \in N_0$ as the team's scores will be an integer of 0 or larger, with large numbers being very unlikely, and $i = 1, \dots, 2g$. Given this explanation, the vector will be of the form:

$$Y^T = (y_{i_1, j_1}^1, y_{j_1, i_1}^1, y_{i_2, j_2}^2, y_{j_2, i_2}^2, \dots, y_{i_g, j_g}^g, y_{j_g, i_g}^g) \quad (5)$$

where $y_{i,j}^g$ represents the number of goals scored by team i versus team j in game g (Gardner, 2011).

The β is a vector containing the parameters for the model, which in this case, is of length $2n$, as each of the n teams will be associated with an attacking and a defensive strength. The first half of the vector will contain the attacking strengths (α_k) of the n teams, and the second half will contain the defensive strengths (γ_k) of the n teams, resulting in a vector of the form (Gardner, 2011):

$$\beta^T = (\alpha_1, \dots, \alpha_n, \gamma_1, \dots, \gamma_n). \quad (6)$$

5.1.4.1.1- The home advantage parameter

Gardner (2011) also considers the possibility of adding a distinctive parameter between the home team and the away team, as the teams playing at home benefit from a statistically demonstrated advantage (as discussed in the subchapter 4.3 of our work). The author uses the two-sample t-test to see if the mean of the home goals is significantly different from the mean of the away goals. Testing the difference between the means of these two samples will show if there is a clear difference between them (i.e. do teams score more goals at home than away).

Although the t-test demands the variables to have normal distributions, Clarke & Cook (2004: 416) state that "If X has a very skew distribution [e.g. Poisson], we need a sample of considerable size (say, at least 200) before the distribution of X approaches reasonably closely to a normal distribution". The dataset used by Gardner (2011) includes all the matches from the Premier League 2009-2010 (380), while our dataset includes all the matches from the Premier League 2016-2017 and 2017-2018 (380 each), therefore the

Forecasting Football Outcomes to Invest in Betting Markets

variables can be assumed as normal distributed, and the t-test can be performed.

To test whether the difference between the means of the home goals and the away goals is statistically significant or not, the following hypothesis was formulated:

- $H_0: \mu_1 = \mu_2$
- $H_1: \mu_1 \neq \mu_2$

In Gardner’s (2011) case, the null hypothesis was strongly rejected at the 99% confidence level, which implies that the difference between the means is statistically significant and consequently, makes sense to add a home advantage parameter δ to the parameter vector β , as it is an extra parameter that needs to be estimated from the data. As a consequence, equation 6 becomes 7 and now has the length $2n + 1$.

We decided to test if, in our case, the difference between the means of the home goals and the away goals was also statistically significant. To do that, we used Microsoft Excel instead of the statistical program R. According to Curto (2016), the first step is to assess if the variances of both samples (home goals and away goals) differ in a statistically significant way, using the “F-Test Two-Sample for Variances” from the “Data Analysis Toolpack”. The hypothesis was:

- H_0 : The variance of the home goals is equal to the one of the away goals
- H_1 : The variance of the home goals is not equal to the one of the away goals

These were the results obtained:

F-Test Two-Sample for Variances		
	FTHG	FTAG
Mean	1,597368421	1,202631579
Variance	1,708172476	1,507644772
Observations	380	380
df	379	379
F	1,133007263	
P(F<=f) one-tail	0,112309597	
F Critical one-tail	1,270531723	

Figure 16- F-test results for 2016/2017

t-Test: Two-Sample Assuming Equal Variances		
	FTHG	FTAG
Mean	1,597368421	1,202631579
Variance	1,708172476	1,507644772
Observations	380	380
Pooled Variance	1,607908624	
Hypothesized Mean Difference	0	
df	758	
t Stat	4,290953228	
P(T<=t) one-tail	1,00476E-05	
t Critical one-tail	2,331277467	
P(T<=t) two-tail	2,00953E-05	
t Critical two-tail	2,582330918	

Figure 17- F-test results for 2017/2018

Forecasting Football Outcomes to Invest in Betting Markets

At the 1% significance level, the null hypothesis was not rejected for the season 2016/2017, and rejected in the case of the 2017/2018 season, which meant that we would need to perform a “t-Test: Two-Sample Assuming Equal Variances” to the former, and a “t-Test: Two-Sample Assuming Unequal Variances” to the latter.

F-Test Two-Sample for Variances		
	FTHG	FTAG
Mean	1,531578947	1,147368421
Variance	1,795833912	1,387196223
Observations	380	380
df	379	379
F	1,294578144	
P(F<=f) one-tail	0,006070285	
F Critical one-tail	1,270531723	

t-Test: Two-Sample Assuming Unequal Variances		
	FTHG	FTAG
Mean	1,531578947	1,147368421
Variance	1,795833912	1,387196223
Observations	380	380
Hypothesized Mean Difference	0	
df	746	
t Stat	4,197983066	
P(T<=t) one-tail	1,50898E-05	
t Critical one-tail	2,331356927	
P(T<=t) two-tail	3,01796E-05	
t Critical two-tail	2,582435754	

Figure 18- t-Test results for 2016/2017

Figure 19- t-Test results for 2017/2018

The results of our t-Test were exactly the same as the ones from the test executed by Gardner (2011), that is, the null hypothesis was strongly rejected at the 1% significance level, which confirmed the necessity of estimating a home advantage parameter.

5.1.4.1.2- Y , X and β

Gardner's (2011) next step was to formulate Y and X into the R language. To achieve this, the author wrote the following code:

```
Parameters <- function(games) {
  teams <- sort(unique(c(games[,1], games[,2])), decreasing = FALSE)
  n <- length(teams)
  g <- nrow(games)
  Y <- matrix(0,2*g,1)
  for (i in 1:g) {
    Y[(2*i)-1] <- games[i,3]
    Y[(2*i)] <- games[i,4]
  }
  X <- matrix(0,2*g,(2*n)+1)
  for (i in 1:g) {
    M <- which(teams == games[i,1])
    N <- which(teams == games[i,2])
    X[(2*i)-1,M] <- 1
    X[(2*i)-1,N+n] <- -1
    X[(2*i),M] <- 1
    X[(2*i),N+n] <- -1
    X[(2*i)-1,(2*n)+1] <- 1
  }
  XX <- X[,-1]
  parameters <- glm(formula = Y ~ 0 + XX, family = poisson)
  Z <- c(0, coefficients(parameters))
  P <- data.frame(row.names=teams, Attack=Z[1:n], Defence=Z[(n+1):(2*n)])
  return(list(teams=P,home=as.numeric(Z[(2*n)+1])))
}
```

Figure 20- Gardner's (2011) code for Y and X

Forecasting Football Outcomes to Invest in Betting Markets

After that, the parameters for the attacking and defensive strengths of each team, as well as the home advantage parameter, were estimated. “As Arsenal is the first team alphabetically in 2009- 2010 Premiership season, Arsenal’s attack strength is being set to 0. (...) This actually makes sense because now the other 40 parameters (19 attack strengths, 20 defence strengths and the home advantage) can be calculated by using this as the benchmark. Since Arsenal is one of the best teams in the league, it can be expected that only Manchester United and Chelsea (who were the only two teams to come above Arsenal in the league in 2009-2010) will have strengths above 0 and the majority of other strengths will be below 0. Using data inputted into R, it is possible to estimate the parameters using a function called glm which stands for ‘generalized linear model’. The Poisson regression model is a form of generalized linear model but with the Poisson distribution taken into account (...). The glm function (...) estimates the 40 remaining parameters, using the benchmark of the [Arsenal’s] parameter (...)” (Gardner, 2011: 30).

```
parameters <- glm(formula = Y ~ 0 + XX, family = poisson)
Z <- c(0, coefficients(parameters))
P <- data.frame(row.names=teams, Attack=Z[1:n], Defence=Z[(n+1):(2*n)])
return(list(teams=P,home=as.numeric(Z[2*n+1])))
}
```

Figure 21- Poisson generalized linear model (glm) code, which was used by Gardner (2011) to estimate the parameters

```
> TeamParameters
Teams
Attack Defence
Arsenal 0.00000000 -0.46511770
Bournemouth -0.39656271 -0.73800848
Brighton -0.77118382 -0.54687244
Burnley -0.70215181 -0.41586723
Chelsea -0.03930105 -0.17091337
Crystal Palace -0.45313757 -0.65384756
Everton -0.35140968 -0.51332121
Huddersfield -0.96162411 -0.61236677
Hull -0.70430012 -0.95305727
Leicester -0.35978047 -0.70008480
Liverpool 0.06279231 -0.29834493
Man City 0.19414256 -0.11778267
Man United -0.23319767 0.06183283
Middlesbrough -1.04684774 -0.53054081
Newcastle -0.64083193 -0.41285096
Southampton -0.65786346 -0.51865188
Stoke -0.67382490 -0.69392504
Sunderland -0.95954757 -0.79675288
Swansea -0.71318275 -0.70844573
Tottenham 0.04106256 -0.04165644
Watford -0.56941077 -0.76067710
West Brom -0.70910995 -0.54514299
West Ham -0.44597078 -0.76631231

Shome
[1] 0.2862813
```

Figure 22- Estimated attack and defence parameters for each one of the Premier League’s teams, plus the home advantage parameter (using data from the seasons 2016/2017 and 2017/2018)

5.1.4.2- Simulating games

To estimate the parameters, the scores from past games, Y , were used. Now these estimates can be used to simulate future scores, by setting up the matrix X^3 for the games to be simulated and putting the set of 41 parameters into order and named as the vector β , they can be inserted into the formula $Y = \exp(X\beta)$, and the vector Y is predicted. This vector will be formed as this (Gardner, 2011):

$$Y = (\exp(\alpha_a - \gamma_b + \delta), \exp(\alpha_b - \gamma_a), \dots) \quad (7)$$

For example, considering the estimated parameters in figure 22, if we want to simulate a match between Chelsea (home) and Arsenal (away), we should replace the symbols in equation 8 for the respective numbers:

- λ Chelsea = $\exp(-0.03930105 - (-0.46511770) + 0.2862813)$
- λ Arsenal = $\exp(0 - (-0.17091337))$

As it is explained in the subchapter 5.1.2 of our work, for a Poisson distribution, the value λ represents the expected number of occurrences of an event, which in this case is the expected number of goals scored by each one of the teams. Knowing the λ for each team, it is possible to use the probability mass function of Poisson to calculate the expected probability for each team to score 1, 2, 3, etc. goals, and from there, the expected probability for each possible final score (0-0, 1-0, 1-1, 1-2, etc.). Gardner (2011) developed this R code, which does this automatically, and presents the results in a table format:

³ To understand this matrix in detail, please consult Gardner (2011: 22-28)

Forecasting Football Outcomes to Invest in Betting Markets

```

ProbTable <- function(parameters,hometeam,awayteam) {
  teams <- rownames(parameters$teams)
  P <- parameters$teams
  home <- parameters$home
  a <- which(teams == hometeam)
  b <- which(teams == awayteam)
  lambdaa <- exp(P[a,]$Attack - P[b,]$Defence + home)
  lambdab <- exp(P[b,]$Attack - P[a,]$Defence)
  A <- as.numeric()
  B <- as.numeric()
  for(i in 0:6) {
    A[(i+1)] <- dpois(i,lambdaa)
    B[(i+1)] <- dpois(i,lambdab)
  }
  A[8] <- 1 - sum(A[1:7])
  B[8] <- 1 - sum(B[1:7])
  name <- c("0","1","2","3","4","5","6","7+")
  zero <- mat.or.vec(8,1)
  C <- data.frame(row.names=name, "0"=zero, "1"=zero, "2"=zero, "3"=zero, "4"=zero,
    "5"=zero, "6"=zero, "7+"=zero)
  for(j in 1:8) {
    for(k in 1:8) {
      C[j,k] <- A[k]*B[j]
    }
  }
  colnames(C) <- name
  return(round(C*100,2))
}

Probabilities <- ProbTable(TeamParameters,
  "Chelsea",
  "Arsenal")

Probabilities

```

Figure 23- Gardner's (2011) code for each score's probability table

	0	1	2	3	4	5	6	7+
0	3.98	8.11	8.26	5.61	2.86	1.17	0.40	0.15
1	4.72	9.62	9.80	6.66	3.39	1.38	0.47	0.18
2	2.80	5.70	5.81	3.95	2.01	0.82	0.28	0.11
3	1.11	2.26	2.30	1.56	0.80	0.32	0.11	0.04
4	0.33	0.67	0.68	0.46	0.24	0.10	0.03	0.01
5	0.08	0.16	0.16	0.11	0.06	0.02	0.01	0.00
6	0.02	0.03	0.03	0.02	0.01	0.00	0.00	0.00
7+	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00

Figure 24- R probability table for the match Chelsea vs Arsenal (using data from the seasons 2016/2017 and 2017/2018)

In this example, if we sum all the probabilities for the home win (e.g. 1-0, 2-1, 2-0, etc.), for the draw (e.g. 0-0, 1-1, 2-2, etc.) and for the away win (e.g. 0-1, 1-2, 2-3, etc.), we obtain, indirectly, the probabilities for each possible outcome. Likewise, we can also sum the probabilities for the match ending with under 1.5 goals (0-0 + 1-0 + 0-1), over 2.5 goals, etc.

5.2- Betting strategies

The initial 140,14€ will be split in half and applied to the following 2 betting strategies in equal manner. At the end of the season, we will compare the performance of these strategies in terms of accuracy and return.

5.2.1- Value strategy

As we have seen in the subchapter 3.3 of this work, due to the current situation of the Portuguese fixed-odds betting markets, it may be challenging to find odds with positive expected value. However, we believe that it is possible to do it, by applying the following steps:

1. Look at the Premier League 1x2 and under/over X goals odds offered by the 4 bookmakers operating legally in Portugal, and identify the ones that offer a relatively good reward for a relatively low risk (this should be done at least once a week)
2. Run a quantitative analysis on the intended matches, using the Poisson regression model and the descriptive statistics, with the aim of finding data patterns that can support our decision-making procedure
3. Write a detailed forecasting report with the relevant results of the quantitative analysis plus all the qualitative factors (tactics of both teams, psychological moment, missing players, weather, etc.) that can influence the outcome of the event and consequently, our decision of betting or not.

5.2.1.1- Forecasting report model

Introduction:

- Name of both teams
- Competition match day
- Date and time of the match
- Location: stadium and city
- Weather at the time of the match: precipitation, humidity and wind (source: Google)
- Potential bets: markets that seem to have interesting odds

Probable line-ups: www.whoscored.com -> Previews -> Select the match from the list -> Preview

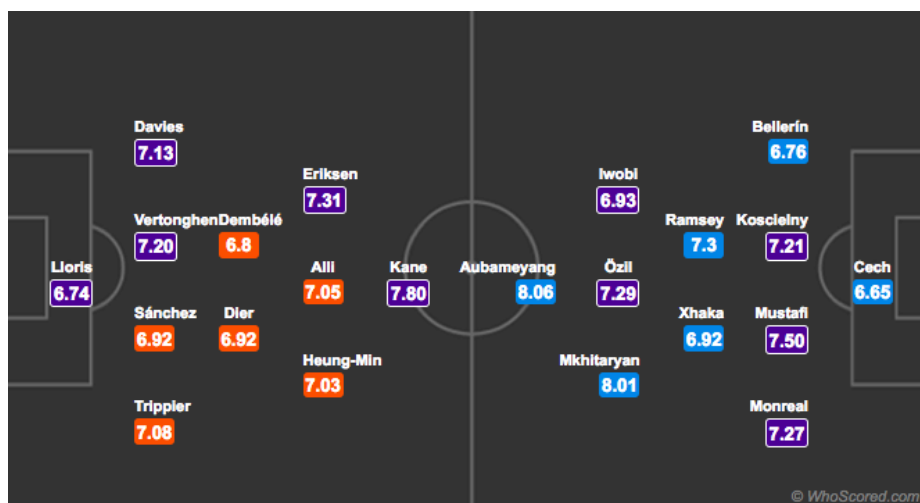


Figure 25- Probable line-up Tottenham vs Arsenal (played at 10/02/2018)

Injuries and suspensions: www.whoscored.com -> Previews -> Select the match from the list -> Preview

- List of all the players that are not available or are at risk for the match in question

Player	Reason	Status	Rating	Player	Reason	Status	Rating
Jesé	↓	Doubtful	6.48	Steve Sidwell	+	Out	N/A
Darren Fletcher	+	Doubtful	6.51	Jiri Skalak	+	Out	N/A
				Jürgen Locadia	+	Doubtful	N/A

Figure 26- Injuries and suspensions Stoke City vs Brighton (played at 10/02/2018)

Discipline table: This table is part of the tool Betting Data 2017-2018, which can be downloaded for free at <http://www.football-data.co.uk/>. This tool analyses the raw data available for download in the same website, and presents it in a way that can assist the punter's decision-making process.

- This table tells us the level of aggressiveness of both teams when playing at home, away and overall, in comparison with the Premier League's average (the right-sided table refers to the home team disciplinary data, the table in middle indicates the away team disciplinary data, and the green table shows the Premier League's average disciplinary data), measuring the average number of yellow cards, red

Forecasting Football Outcomes to Invest in Betting Markets

cards and booking points per match. In order to obtain the Premier League's average per match for a single team, one must divide the green table's numbers in half

- A yellow card worth 10 booking points and a red card corresponds to 25 booking points (there is no distinction between direct red cards and double yellows)

AVERAGE PER MATCH		YC	RC	BP	YC	RC	BP	YC	RC	BP
Match average of yellow cards (YC), red cards (RC) and booking points (BP)	Total	1,6	0,1	17,7	1,8	0,0	18,5	3,1	0,1	33,7
	Home	1,8	0,1	19,6	2,2	0,0	22,3	1,5	0,1	16,5
	Away	1,4	0,1	15,8	1,5	0,0	14,6	1,6	0,1	17,2

Figure 27- Discipline table Everton vs Crystal Palace (played at 10/02/2018)

Tables: www.academiasapostas.com -> Estadísticas -> England flag -> select match from the list

- Head-to-head over the last 3 years table: allows to understand if one team has an historical advantage over the other, maybe due to having a style of play that is suited to exploit the other team's weaknesses, or because of psychological factors

2017-11-18	🇬🇧 17/18	Burnley	2-0	Swansea
2017-03-04	🇬🇧 16/17	Swansea	3-2	Burnley
2016-08-13	🇬🇧 16/17	Burnley	0-1	Swansea
2015-02-28	🇬🇧 14/15	Burnley	0-1	Swansea
2014-08-23	🇬🇧 14/15	Swansea	1-0	Burnley

Figure 28- Head-to-head over the last 3 years Swansea vs Burnley (played at 10/02/2018)

Forecasting Football Outcomes to Invest in Betting Markets

- Premier league tables overall, home and away: allows to understand the comparative strength of both teams, and the importance of home advantage for each one of them

Total							Casa							Fora									
#	Equipa	P	J	V	E	D	Golos	#	Equipa	P	J	V	E	D	Golos	#	Equipa	P	J	V	E	D	Golos
1	Man City	69	26	22	3	1	74:19	1	Man City	37	13	12	1	0	45:9	1	Man City	32	13	10	2	1	29:10
2	Man Utd	56	26	17	5	4	51:18	2	Man Utd	32	13	10	2	1	29:5	2	Chelsea	24	13	7	3	3	25:13
3	Liverpool	51	26	14	9	3	59:31	3	Arsenal	32	13	10	2	1	36:14	3	Liverpool	24	13	7	3	3	32:22
4	Chelsea	50	26	15	5	6	46:23	4	Tottenham	28	13	8	4	1	28:9	4	Man Utd	24	13	7	3	3	22:13
5	Tottenham	49	26	14	7	5	51:24	5	Liverpool	27	13	7	6	0	27:9	5	Tottenham	21	13	6	3	4	23:15
6	Arsenal	45	26	13	6	7	51:35	6	Chelsea	26	13	8	2	3	21:10	6	Burnley	18	13	4	6	3	12:13
7	Burnley	36	26	9	9	8	21:23	7	Everton	23	13	7	2	4	20:17	7	Leicester	14	13	3	5	5	19:20
8	Leicester	35	26	9	8	9	38:35	8	Leicester	21	13	6	3	4	19:15	8	Watford	14	13	4	2	7	17:19
9	Bournemouth	31	26	8	7	11	30:37	9	Bournemouth	18	13	5	3	5	18:19	9	Arsenal	13	13	3	4	6	15:21
10	Everton	31	26	8	7	11	29:45	10	Burnley	18	13	5	3	5	9:10	10	Bournemouth	13	13	3	4	6	12:18
11	Watford	30	26	8	6	12	37:45	11	Brighton	18	13	4	6	3	15:19	11	Newcastle	12	13	3	3	7	13:21
12	West Ham	27	26	6	9	11	32:46	12	Stoke	18	13	5	3	5	15:20	12	Southampton	11	12	2	5	5	12:19
13	Brighton	27	26	6	9	11	21:35	13	Crystal Palace	17	13	4	5	4	16:19	13	West Ham	11	14	2	5	7	18:29
14	Crystal Palace	27	26	6	9	11	24:39	14	West Ham	16	12	4	4	4	14:17	14	Swansea	10	13	2	4	7	9:19
15	Southampton	26	26	5	11	10	28:38	15	Watford	16	13	4	4	5	20:26	15	Crystal Palace	10	13	2	4	7	8:20
16	Newcastle	25	26	6	7	13	24:36	16	Huddersfield	16	13	4	4	5	11:19	16	Brighton	9	13	2	3	8	6:16
17	Swansea	24	26	6	6	14	19:37	17	Southampton	15	14	3	6	5	16:19	17	Everton	8	13	1	5	7	9:28
18	Stoke	24	26	6	6	14	26:52	18	Swansea	14	13	4	2	7	10:18	18	Huddersfield	8	13	2	2	9	8:27
19	Huddersfield	24	26	6	6	14	19:46	19	West Bromwich	13	13	2	7	4	14:18	19	West Bromwich	7	13	1	4	8	7:19
20	West Bromwich	20	26	3	11	12	21:37	20	Newcastle	13	13	3	4	6	11:15	20	Stoke	6	13	1	3	9	11:32

Figure 29- Premier League tables Swansea vs Burnley (played at 10/02/2018)

SoccerSTATS data: <http://www.soccerstats.com/latest.asp?league=england> -> select match -> H2H

- Home/away performance: this table describes the percentage of points, goals scored and goals conceded by each team at home and away

Home	Away	HOME / AWAY PERFORMANCE	Home	Away
59%	41%	% Points	53%	47%
44%	56%	% Goals scored	54%	46%
37%	63%	% Goals conceded	58%	42%

Figure 30- Home/away performance West Ham vs Watford (played at 10/02/2018)

- Points per game table: this table describes the relative form of each team, which corresponds to the difference between the average points per game in the last 8

Forecasting Football Outcomes to Invest in Betting Markets

matches and the average points per game in all matches played so far in the Premier League. If positive, it means that the team is performing better in the last 8 matches than in the overall season, if negative it means the opposite

Points Per Game (all matches)	1.04
Points Per Game (last 8 matches)	1.25
Relative Form (PPG last 8 - PPG all matches)	+0.21

Figure 31- Points per game West Ham (at 10/02/2018)

- Scoring table: This table provides all the information that is relevant to bet in the goals related markets

SCORING	Home	Away	Total
Goals scored	14	18	32
Goals conceded	17	29	46
Goals scored per match	1.17	1.29	1.23
Goals conceded per match	1.42	2.07	1.77
Scored+conc. per match	2.58	3.36	3.00
Matches over 1.5 goals	75%	93%	85%
Matches over 2.5 goals	42%	71%	58%
Matches over 3.5 goals	25%	50%	38%
Matches over 4.5 goals	25%	21%	23%
Over 2.5 goals at half-time	0%	7%	4%

Figure 32- Scoring table of West Ham (at 10/02/2018)

Style of play data: www.whoscored.com -> Previews -> Select the match from the list -> Head to Head -> Team Statistics

- Team characteristics: provides a detailed description of every team's strengths, weaknesses and style of play. It allows to understand if any of each team's strengths match with the other team's weaknesses, and vice-versa

Forecasting Football Outcomes to Invest in Betting Markets

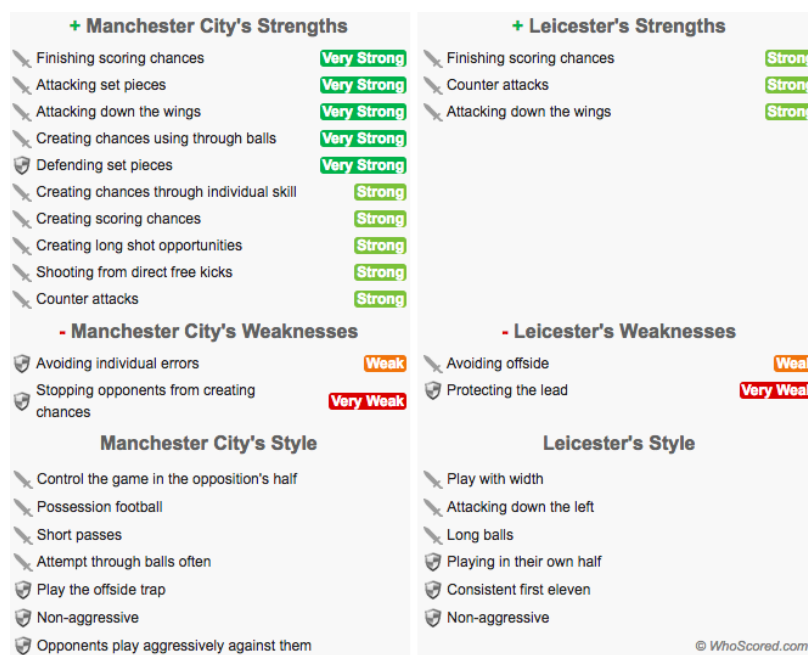


Figure 33- Team characteristics Manchester City vs Leicester (played at 10/02/2018)

Subjective factors:

As mentioned earlier, these factors are very important, and having them in consideration is essential when trying to produce accurate forecasts of what is likely to happen on the pitch. However, they are hard to quantify and therefore left to subjective interpretation. Constantinou *et al.* (2012) attribute key importance to three in particular: form, psychological moment and fatigue. We believe that some of the descriptive statistics present in this report describe relatively well the factor form, and so we will focus on the factors psychological moment and fatigue. In the same work, the authors propose a system for assessing these last two factors that we will adopt in our forecasting reports.

Table A.3. Team Psychology (as presented in Figure 4)

ID	Variable (node)	Description	Subjective Scenarios
I.	Team spirit and motivation	Expert indication regarding the team's level of motivation and team spirit	[Very High, High, Normal, Low, Very Low]
II.	Confidence	Expert indication regarding its confidence about his input in (I).	[Very High, High, Medium, Low, Very Low]
III.	Managerial impact	Expert indication regarding the impact of the current managerial situation.	[Very High, High, Normal, Low, Very Low]
IV.	Head-to-Head bias	Expert indication regarding potential biases in a head-to-head encounter between the two teams.	[High advantage for home team, Advantage for home team, No bias, Advantage for away team, High advantage for away team]

Figure 34- Evaluation method for the psychological moment of a certain team
(Constantinou et al., 2012: 19)

The point IV. of the Team Psychology evaluation will not be considered, as it was previously discussed in the Head-to-head over the last 3 years table ⁴

Table A.4. Team Fatigue (as presented in Figure 5)

ID	Variable (node)	Description	Subjective Scenarios
I.	Toughness of previous match	Expert indication regarding the toughness of previous match.	[Lowest, Very Low, Low, Medium, High, Very High, Highest]
II.	First team players rested during last match	Expert indication regarding the first team players rested during last match.	[1-2, 3, 4, 5, 6+]
III.	National team participation	Expert indication regarding the level of international participation by the first team players.	[None, Few, Half team, Many, All]

Figure 35- Evaluation method for the fatigue of a certain team (Constantinou *et al.*, 2012: 19)

Poisson regression model:

The results of the Poisson regression model are presented in this Excel table, including the probabilities for each possible score, the probabilities for 1x2, the probabilities for under/over X goals and the respective odd conversion.

	0	1	2	3	4	5	6	7+	Home
0	5,16%	7,02%	4,77%	2,16%	0,74%	0,20%	0,05%	0,01%	
1	8,27%	11,26%	7,66%	3,47%	1,18%	0,32%	0,07%	0,02%	
2	6,64%	9,03%	6,14%	2,79%	0,95%	0,26%	0,06%	0,01%	
3	3,55%	4,83%	3,28%	1,49%	0,51%	0,14%	0,03%	0,01%	
4	1,42%	1,94%	1,32%	0,60%	0,20%	0,06%	0,01%	0,00%	
5	0,46%	0,62%	0,42%	0,19%	0,07%	0,02%	0,00%	0,00%	
6	0,12%	0,17%	0,11%	0,05%	0,02%	0,00%	0,00%	0,00%	
7+	0,03%	0,05%	0,03%	0,01%	0,00%	0,00%	0,00%	0,00%	
Away									
	5,16%	11,26%	6,14%	1,49%	0,20%	0,02%	0,00%	0,00%	
Home win	32,50%		Odd	3,08					
Draw	24,27%		Odd	4,12					
Away win	43,23%		Odd	2,31					
Total	100,00%								
Under 2.5	43,12%		Odd	2,32					
Over 2.5	56,88%		Odd	1,76					
Total	100,00%								

⁴ Please consult the 41st page of our work

Figure 36- Poisson regression model results on Watford vs Everton (played at 24/02/2018)

Betting decision (based on the odds):

Here, the report's content is compared with the potential bets, and afterwards a decision is made about which markets worth a bet (if any).

5.2.2- Accuracy strategy

According to the document "The Football Betting Mastermind" (Betinfo24: 2010), it is very unlikely that an individual punter is able to find value and beat the bookmakers in the long term (in major leagues), because these have top-level statisticians and football experts working for them, backed up by the best technology that money can buy (consequently, the odds are pretty accurate). So, the only way of making consistent profits is being able to identify the teams that are expected to win almost every game, and betting high stakes on those teams (even if the odds are low).

The section 1 of that document explains how to identify the teams that have a very high percentage of wins, dividing them in "hot teams" and "strong teams". "Hot teams" are defined as the ones that are very likely to win every match in the leagues where they compete (except when playing against each other). By its turn, "strong teams" are defined as the ones that are maintaining good performances as well, but are not as consistent as the "hot teams", especially when playing away (usually these teams are ranked between 3rd and 7th in the league table). As the "strong teams" do not perform as regularly as the "hot teams", bookmakers tend to offer better odds for their wins. However, the objective of this particular strategy is to place bets as "certain" as possible, and therefore we will rule the "strong teams" out.

The rules for discarding "hot teams" to bet on at each match day are also stated in "The Football Betting Mastermind":

- Discard "hot teams" that play against other "hot teams"
- Discard "hot teams" when playing away against a "strong team" (this rule is not applicable to our case, as we decided to exclude the "strong teams")

Forecasting Football Outcomes to Invest in Betting Markets

- Discard “hot teams” when odds offered for the win are < 1.15 (Betclie and Bet.pt do not allow single bets with odds lower than 1.20)

Our research indicates that the home advantage is a factor that increases the chances of winning considerably, so we decided to add another rule:

- Discard “hot teams” when playing away

In the major leagues of England, Spain, Germany, Italy, France and Portugal (the ones that we know better), we identified these as “hot teams” (at 14/02/2018):

England:

- Manchester City- 1st place, 14 matches played at home, 13 wins, 1 draw, 0 losses -> home win % of 92,86 (equivalent to an odd of 1.08), 50 goals scored and 10 conceded

Spain:

- Barcelona- 1st place, 11 matches played at home, 9 wins, 2 draws, 0 losses -> home win % of 81,82 (equivalent to an odd of 1.22), 31 goals scored and 5 conceded

Germany:

- Bayern Munich- 1st place, 11 matches played at home, 10 wins, 1 draw, 0 losses -> home win % of 90,91 (equivalent to an odd of 1.10), 34 goals scored and 9 conceded

Italy:

- Napoli- 1st place, 12 matches played at home, 9 wins, 2 draws, 1 loss -> home win % of 75 (equivalent to 1.33 odd), 29 goals scored and 8 conceded
- Juventus- 2nd place, 12 matches played at home, 10 wins, 1 draws, 1 loss -> home win % of 83,33 (equivalent to an odd of 1.20), 30 goals scored and 4 conceded

France:

- PSG- 1st place, 12 matches played at home, 12 wins, 0 draws, 0 losses -> home win % of 100, 46 goals scored and 7 conceded

Forecasting Football Outcomes to Invest in Betting Markets

Portugal:

- Porto- 1st place, 11 matches played at home, 10 wins, 1 draw, 0 losses -> home win % of 90,91 (equivalent to an odd of 1.10), 34 goals scored and 7 conceded
- Benfica- 2nd place, (10 matches played at home, 9 wins, 1 draw, 0 losses -> home win % of 90 (equivalent to an odd of 1.11), 31 goals scored and 5 conceded
- Sporting- 3rd place, 12 matches played at home, 10 wins, 2 draws, 0 losses -> home win % of 83,33 (equivalent to an odd of 1.20), 26 goals scored and 4 conceded

Average win %: 87,57 (equivalent to an odd of 1.14)

In conclusion, this strategy means that before each match day, we bet on every “hot team” that plays at home against other teams that are not considered “hot”, if the reward offered by the bookmakers in case of success is higher than 15% of the stake placed.

6- Discussion of the results

6.1- Portuguese online fixed-odds betting markets

	Bookmaker's odds				Equal margin (no bias)			
	Home	Draw	Away	Margin	Home	Draw	Away	Margin
Bayern vs Hertha ESC	1,17	7,50	19,00	4,07%	1,22	7,80	19,77	
Barcelona vs Girona ESC	1,18	7,50	15,00	4,75%	1,24	7,86	15,71	
PSG vs Marseille Bet.pt	1,23	5,75	9,00	9,80%	1,35	6,31	9,88	
Sporting vs Moreirense CP	1,17	6,00	13,00	9,83%	1,29	6,59	14,28	
Benfica vs Marítimo ESC	1,17	7,50	17,00	4,69%	1,22	7,85	17,80	
Napoli vs Roma Betcltic	1,55	3,70	5,25	10,59%	1,71	4,09	5,81	
Man City vs Chelsea CP	1,57	3,80	5,00	10,01%	1,73	4,18	5,50	
Barcelona vs Atletico Bet.pt	1,57	3,70	5,20	9,95%	1,73	4,07	5,72	
Juventus vs Udinese Bet.pt	1,24	5,20	11,50	8,57%	1,35	5,65	12,49	
Juventus vs Atalanta CP	1,40	4,00	7,50	9,76%	1,54	4,39	8,23	
Napoli vs Genoa CP	1,18	5,50	13,00	10,62%	1,31	6,08	14,38	
Barcelona vs Athletic ESC	1,18	7,00	17,00	4,91%	1,24	7,34	17,84	
Sporting vs Rio Ave Betcltic	1,25	5,25	11,50	7,74%	1,35	5,66	12,39	
Bayern vs Dortmund CP	1,36	4,50	6,50	11,14%	1,51	5,00	7,22	
Benfica vs Guimarães ESC	1,17	7,50	17,00	4,69%	1,22	7,85	17,80	
Juventus vs Milan CP	1,60	3,40	5,50	10,09%	1,76	3,74	6,06	
Man City vs Man United CP	1,73	3,50	4,33	9,47%	1,89	3,83	4,74	
Sporting vs Paços ESC	1,18	7,50	15,00	4,75%	1,24	7,86	15,71	
Napoli vs Chievo ESC	1,15	8,00	19,00	4,72%	1,20	8,38	19,90	
Bayern vs M'gladbach CP	1,25	5,50	8,50	9,95%	1,37	6,05	9,35	
Barcelona vs Valencia Betcltic	1,25	5,50	9,00	9,29%	1,37	6,01	9,84	
Juventus vs Sampdoria ESC	1,18	7,00	17,00	4,91%	1,24	7,34	17,84	
PSG vs Monaco CP	1,25	5,50	8,50	9,95%	1,37	6,05	9,35	
Napoli vs Udinese Bet.pt	1,19	5,85	15,25	7,69%	1,28	6,30	16,42	
Sporting vs Boavista Bet.pt	1,21	5,75	15,25	6,59%	1,29	6,13	16,26	
Bayern vs Frankfurt Betcltic	1,45	4,25	5,50	10,68%	1,60	4,70	6,09	
Barcelona vs Real Madrid Betcltic	1,65	4,10	4,20	8,81%	1,80	4,46	4,57	
Napoli vs Torino CP	1,22	5,50	10,00	10,15%	1,34	6,06	11,01	
Barcelona vs Villarreal Bet.pt	1,33	4,70	7,25	10,26%	1,47	5,18	7,99	
Manchester City vs Brighton ESC	1,19	7,00	13,00	6,01%	1,26	7,42	13,78	
Bayern vs Stuttgart ESC	1,19	7,40	11,00	6,64%	1,27	7,89	11,73	
PSG vs Rennes CP	1,22	6,00	9,00	9,74%	1,34	6,58	9,88	
Napoli vs Crotone CP	1,44	4,50	5,50	9,85%	1,58	4,94	6,04	
Barcelona vs Real Sociedad Bet.pt	1,28	6,35	9,00	4,98%	1,34	6,67	9,45	
Average overround	8,11%							
Average overround without ESC	9,40%							
Average overround ESC	5,01%							
Average overround Bet.pt	8,26%							
Average overround Betcltic	9,42%							
Average overround CP	10,05%							
	Percentage				Percentage			
Average home win odd	1,31	76,62%			1,41	70,81%		
Average draw odd	5,64	17,74%			6,07	16,48%		
Average away win odd	10,71	9,33%			11,49	8,70%		
Total		103,69%				95,99%		

	Bookmaker's odds				Equal margin (no bias)			
	Home	Draw	Away	Margin	Home	Draw	Away	Margin
Bayern vs Hertha	1,17	8,44	19,88	2,35%	1,20	8,64	20,35	
Barcelona vs Girona	1,19	7,95	15,70	2,98%	1,23	8,19	16,17	
PSG vs Marseille	1,26	7,04	10,99	2,67%	1,29	7,23	11,28	
Sporting vs Moreirense	1,21	7,05	16,75	2,80%	1,24	7,25	17,22	
Benfica vs Marítimo	1,17	8,52	20,23	2,15%	1,20	8,70	20,67	
Napoli vs Roma	1,57	4,24	6,42	2,86%	1,61	4,36	6,60	
Man City vs Chelsea	1,62	4,22	5,72	2,91%	1,67	4,34	5,89	
Barcelona vs Atletico	1,66	4,07	5,70	2,35%	1,70	4,17	5,83	
Juventus vs Udinese	1,27	5,96	14,40	2,46%	1,30	6,11	14,75	
Juventus vs Atalanta	1,46	4,46	8,67	2,45%	1,50	4,57	8,88	
Napoli vs Genoa	1,22	6,61	18,93	2,38%	1,25	6,77	19,38	
Barcelona vs Athletic	1,22	6,93	16,58	2,43%	1,25	7,10	16,98	
Sporting vs Rio Ave	1,27	6,16	12,83	2,77%	1,31	6,33	13,19	
Bayern vs Dortmund	1,41	5,19	8,00	2,69%	1,45	5,33	8,22	
Benfica vs Guimarães	1,16	9,10	19,27	2,39%	1,19	9,32	19,73	
Juventus vs Milan	1,65	3,81	6,44	2,38%	1,69	3,90	6,59	
Man City vs Man United	1,79	3,89	4,90	1,98%	1,83	3,97	5,00	
Sporting vs Paços	1,19	7,73	16,78	2,93%	1,22	7,96	17,27	
Napoli vs Chievo	1,15	8,10	24,91	3,32%	1,19	8,37	25,74	
Bayern vs M'gladbach	1,32	5,94	10,06	2,53%	1,35	6,09	10,31	
Barcelona vs Valencia	1,30	6,24	10,37	2,59%	1,33	6,40	10,64	
Juventus vs Sampdoria	1,19	7,45	19,25	2,65%	1,22	7,65	19,76	
PSG vs Monaco	1,28	6,38	11,13	2,78%	1,32	6,56	11,44	
Napoli vs Udinese	1,22	6,75	16,50	2,84%	1,25	6,94	16,97	
Sporting vs Boavista	1,24	6,64	15,08	2,34%	1,27	6,80	15,43	
Bayern vs Frankfurt	1,51	4,66	6,68	2,65%	1,55	4,78	6,86	
Barcelona vs Real Madrid	1,70	4,50	4,75	2,10%	1,74	4,59	4,85	
Napoli vs Torino	1,24	6,79	14,23	2,40%	1,27	6,95	14,57	
Barcelona vs Villarreal	1,37	5,25	8,30	4,09%	1,43	5,46	8,64	
Manchester City vs Brighton	1,19	7,84	15,47	3,25%	1,23	8,10	15,97	
Bayern vs Stuttgart	1,17	8,71	12,71	4,82%	1,23	9,13	13,32	
PSG vs Rennes	1,25	6,85	12,31	2,72%	1,28	7,04	12,65	
Napoli vs Crotone	1,49	5,3	6,07	2,46%	1,53	5,43	6,22	
Barcelona vs Real Sociedad	1,30	6,56	8,50	3,93%	1,35	6,82	8,83	
	Percentage				Percentage			
Average overround	2,75%							
Average home win odd	1,34	74,87%			1,37	72,89%		
Average draw odd	6,33	15,79%			6,51	15,36%		
Average away win odd	12,49	8,01%			12,83	7,79%		
Total		98,67%				96,04%		

Figure 37- Bookmakers' odds from the Portuguese online betting markets (right side) vs Pinnacle's (British bookmaker) odds (left side)

If we compare the odds proposed by 4 bookmakers operating in the Portuguese online betting markets with the ones offered by Pinnacle, the first thing to notice is the substantial difference between the average overrounds: 8,11% practiced by the Portuguese bookmakers, against the 2,75% taken by Pinnacle. This means that the bookmakers' commission ⁵ is 5,36 percentage points (pp) higher in the Portuguese online betting markets' case, and consequently, there should be considerably less opportunities for the punters to exploit. Furthermore, Casino Portugal (CP) was the bookmaker that

⁵ Please consult the 3rd page of our work

most often offered the better odds (12 times), followed by Estoril Sol Casinos, or ESC, (10 times), Bet.pt (7 times) and Betclic (5 times).

ESC can be considered a very *sui generis* bookmaker, because it sets odds that are, in most cases, the highest in the Portuguese online betting markets. Its average overround is equal to 5,01%, a lower value in comparison with the 8,26%, 9,42% and 10,05%, charged by Bet.pt, Betclic and CP, respectively. However, if one is betting on a single outcome, and the odd is higher than 1.20, it automatically drops to values worse than the ones offered by the other 3 bookmakers. Because of this, we believe that makes sense to compare the 3 Portuguese bookmakers' and Pinnacle's overrounds without considering ESC. In that scenario, the difference is even bigger (6,65%), as the average overround of Bet.pt, Betclic and CP is equal to 9,40%.

As it is explained in chapter 3.3, the lack of competitiveness in the Portuguese betting markets created an oligarchy “that is harmful to the punters interests, because bookmakers do not feel much pressure to reward the risks taken by the punters with valuable odds”. On the contrary, the British markets have so much offer that in order to stay competitive, the bookmakers have to provide valuable odds, and therefore, to reduce drastically their commissions, that is, the overrounds. Besides that, if we compare the average odds for the 3 possible outcomes, in all cases, Pinnacle pays more than the 4 Portuguese bookmakers (home win: 1.34 vs 1.31; draw: 6.33 vs 5.64 and away win: 12.49 vs 10.71).

Another explanation for the higher overrounds in the Portuguese bookmakers' case is the punitive tax of 8 to 16% over the volume of bets made, which means that they must pay prizes and taxes for bets won by the punters, while the UK government taxes the bookmakers in 15% of their profits (HM Revenue & Customs, 2017). This legal framework “forces” the Portuguese bookmakers to increase the overround, and decrease the odds, reducing the value offered to the market.

It is known that competitive markets deliver higher value to its customers than uncompetitive markets, however, our investigation could not find any records related with the Portuguese markets' odds. Thus, we believe that this document adds a valid contribution to the existing literature, as it provides numbers which support the idea that Portuguese betting markets are, in fact, less competitive (at least in comparison with the ones from Great Britain), and as a consequence, it seems reasonable to conclude that it is more difficult for its customers (punters) to consistently profit, in the long-run.

Forecasting Football Outcomes to Invest in Betting Markets

In a survey that served as the basis for the Remote Gambling Association report (Stradbroke, 2017), which was made to determine how Portugal’s regulated online market was doing since its regulation, “68% of the respondents were still using gambling sites not allowed to operate in Portugal. From these punters, 38% gambled solely through unlicensed websites, while 30% admitted using both licensed and unlicensed online gambling platforms. When the respondents were asked why they still use unlicensed websites, the main reason mentioned was the wish to access better odds than the ones offered by the Portugal-licensed gambling platforms”⁶. These results suggest that punters feel that Portuguese odds do not reward the risks taken in a fair way, and many prefer to use unlicensed bookmakers, breaking the law and diminishing the government’s returns on the sports betting business.

The findings of our study appear to coincide with the conclusions drawn in the Remote Gambling Association’s reports (2017 and 2018), which claim that although the tax on sports betting turnover is a significant source of income for the Portuguese government, the current Portuguese betting regulation is harmful to both the punters’ and the bookmakers’ interests, and in the words of RGA director Pierre Tournier, is “clearly failing to combat the unregulated market, (...) change is much needed to make the regulation work” (Stradbroke, 2017). Hence, our research indicates that, from a long-term strategic perspective, reviewing the decrees of Law n° 66 and 67/2015 would benefit all 3 parts involved: the punters, the bookmakers and the Portuguese government itself.

6.2- Betting record

Match Date	Betting Date	Money	Home Team	Away Team	Competition	Market	Bet	Bookmaker	Stake (10%)	Odds	Revenues	Profit	ROI	Final Score
24/02/18	24/02/18	70,07 €	Watford	Everton	Premier League	2.5	Over	Bet.pt	7,01 €	1,86	- 7,01 €	- 7,01 €	-100,00%	1-0
01/03/18	01/03/18	63,06 €	Arsenal	Man City	Premier League	1x2	1	Betcllc	6,31 €	1,75	11,04 €	4,73 €	75,00%	0-3
04/03/18	02/03/18	67,79 €	Man City	Chelsea	Premier League	1x2	1	CP	6,78 €	1,57	10,64 €	3,86 €	57,00%	1-0
10/03/18	07/03/18	71,65 €	Man United	Liverpool	Premier League	1x2, +/- X	1x, + 1.5	Betcllc	7,17 €	1,80	12,90 €	5,73 €	80,00%	2-1
17/03/18	14/03/18	77,39 €	Bournemouth	WBA	Premier League	1x2	1	Betcllc	7,74 €	1,82	14,08 €	6,35 €	82,00%	2-1
17/03/18	14/03/18	77,39 €	Liverpool	Watford	Premier League	1x2, +/- X	1, -4.5	Betcllc	7,74 €	1,58	- 7,74 €	- 7,74 €	-100,00%	5-0
31/03/18	28/03/18	75,99 €	Everton	Man City	Premier League	1x2, +/- X	2, +1.5	Betcllc	7,60 €	1,45	11,02 €	3,42 €	45,00%	1-3
08/04/18	04/04/18	79,41 €	Arsenal	Southampton	Premier League	1x2	1	Betcllc	7,94 €	1,50	11,91 €	3,97 €	50,00%	3-2
22/04/18	21/04/18	83,38 €	Arsenal	West Ham	Premier League	1x2, +/- X	1, +1.5	Betcllc	8,34 €	1,58	13,17 €	4,84 €	58,00%	4-1
30/04/18	25/04/18	88,22 €	Tottenham	Watford	Premier League	3.5	Under	Bet.pt	8,82 €	1,58	13,94 €	5,12 €	58,00%	2-0
13/05/18	10/05/18	93,34 €	Southampton	Man City	Premier League	1x2, +/- X	2, +1.5	Betcllc	9,33 €	1,55	- 9,33 €	- 9,33 €	-100,00%	0-1
Final Money	84,00 €			Value-Accuracy M	13,65 €		Value-PIN M	9,29 €			Std DV	6,05 €	77,09%	
Final ROI	19,88%			Value-Accuracy ROI	19,49%		Value-PIN ROI	13,26%			Mean	1,27 €	18,64%	
Betting Accuracy	72,73%			Value-Accuracy A%	-6,68%		Value-PIN A%	-6,68%			CV	477,48%	413,66%	
Final Profit	13,93 €													
Average odd	1,64													

Figure 38- Betting record of the value strategy

⁶ Please consult the subchapter 3.3 of our work

Forecasting Football Outcomes to Invest in Betting Markets

Match Date	Betting Date	Home Team	Away Team	Competition	Market	Bet	Bookmaker	Money	Stake (10%)	Odd	Revenues	Profit	ROI	Money Pinnacle	Stake Pinnacle (10%)	Odd Pinnacle	Revenues Pinnacle	Profit Pinnacle	ROI Pinnacle	Final Score	Total Scored	Total Conceded	Goal Difference
24/02/18	23/02/18	Bayern	Hertha	Bundesliga	1x2	1	ESC	70,07 €	7,01 €	1,17	-	-7,01 €	-100,00%	70,07 €	7,01 €	1,17	-	-7,01 €	-100,00%	0-0	0	0	0
24/02/18	24/02/18	Barcelona	Girona	La Liga	1x2	1	ESC	70,07 €	7,01 €	1,18	8,27 €	1,26 €	18,00%	70,07 €	7,01 €	1,19	8,34 €	1,33 €	19,00%	6-1	6	1	5
25/06/18	23/02/18	PSG	Marseille	Ligue 1	1x2	1	Bet.pt	70,07 €	7,01 €	1,23	8,62 €	1,61 €	23,00%	70,07 €	7,01 €	1,26	8,83 €	1,82 €	26,00%	3-0	9	1	8
26/02/18	23/02/18	Sporting	Moreirense	Liga NOS	1x2	1	CP	70,07 €	7,01 €	1,17	8,20 €	1,19 €	17,00%	70,07 €	7,01 €	1,21	8,48 €	1,47 €	21,00%	1-0	10	1	9
03/03/18	01/03/18	Benfica	Maritimo	Liga NOS	1x2	1	ESC	67,12 €	6,71 €	1,17	7,85 €	1,14 €	17,00%	67,69 €	6,77 €	1,17	7,85 €	1,14 €	16,02%	5-0	15	1	14
03/03/18	01/03/18	Napoli	Roma	Serie A	1x2	1	Betcllc	67,12 €	6,71 €	1,55	-	-6,71 €	-100,00%	67,69 €	6,77 €	1,57	-	-6,77 €	-100,00%	2-4	17	5	12
04/03/18	02/03/18	Man City	Chelsea	Premier League	1x2	1	CP	67,12 €	6,71 €	1,57	10,54 €	3,83 €	57,00%	67,69 €	6,77 €	1,62	10,87 €	4,16 €	60,64%	1-0	18	5	13
04/03/18	02/03/18	Barcelona	Atletico	La Liga	1x2	1	Bet.pt	67,12 €	6,71 €	1,57	10,54 €	3,83 €	57,00%	67,69 €	6,77 €	1,66	11,14 €	4,43 €	64,61%	1-0	19	5	14
11/03/18	06/03/18	Juventus	Udinese	Serie A	1x2	1	Bet.pt	69,20 €	6,92 €	1,24	8,58 €	1,66 €	24,00%	70,65 €	7,07 €	1,27	8,79 €	1,87 €	24,39%	2-0	21	5	16
14/03/18	12/03/18	Juventus	Atalanta	Serie A	1x2	1	CP	70,86 €	7,09 €	1,40	9,52 €	2,83 €	40,00%	72,52 €	7,25 €	1,46	10,35 €	3,26 €	42,66%	2-0	23	5	18
18/03/18	12/03/18	Napoli	Genoa	Serie A	1x2	1	CP	70,86 €	7,09 €	1,18	8,36 €	1,28 €	18,00%	72,52 €	7,25 €	1,22	8,65 €	1,56 €	19,21%	1-0	24	5	19
18/03/18	12/03/18	Barcelona	Athletic	La Liga	1x2	1	ESC	70,86 €	7,09 €	1,18	8,36 €	1,28 €	18,00%	72,52 €	7,25 €	1,22	8,65 €	1,56 €	19,21%	2-0	26	5	21
18/03/18	12/03/18	Sporting	Rio Ave	Liga NOS	1x2	1	Betcllc	70,86 €	7,09 €	1,25	8,86 €	1,77 €	25,00%	72,52 €	7,25 €	1,27	9,00 €	1,91 €	24,10%	2-0	28	5	23
31/03/18	27/03/18	Bayern	Dortmund	Bundesliga	1x2	1	CP	78,02 €	7,80 €	1,36	10,61 €	2,81 €	36,00%	80,81 €	8,08 €	1,41	11,00 €	3,20 €	36,13%	6-0	34	5	29
31/03/18	27/03/18	Benfica	Guimaraes	Liga NOS	1x2	1	ESC	78,02 €	7,80 €	1,17	9,13 €	1,33 €	17,00%	80,81 €	8,08 €	1,16	9,05 €	1,25 €	11,99%	2-0	36	5	31
31/03/18	27/03/18	Juventus	Milan	Serie A	1x2	1	CP	78,02 €	7,80 €	1,60	12,48 €	4,68 €	60,00%	80,81 €	8,08 €	1,65	12,87 €	5,07 €	59,30%	3-1	39	6	33
07/04/18	04/04/18	Man City	Man United	Premier League	1x2	1	CP	86,83 €	8,68 €	1,73	-	-8,68 €	-100,00%	90,33 €	9,03 €	1,79	-	-9,03 €	-100,00%	2-3	41	9	32
08/04/18	04/04/18	Sporting	Paços	Liga NOS	1x2	1	ESC	86,83 €	8,68 €	1,18	10,25 €	1,56 €	18,00%	90,33 €	9,03 €	1,19	10,33 €	1,65 €	14,40%	2-0	43	9	34
08/04/18	07/04/18	Napoli	Chievo	Serie A	1x2	1	ESC	86,83 €	8,68 €	1,15	9,99 €	1,30 €	15,00%	90,33 €	9,03 €	1,15	9,99 €	1,30 €	10,55%	2-1	45	10	35
14/04/18	11/04/18	Bayern	M'gladbach	Bundesliga	1x2	1	CP	81,02 €	8,10 €	1,25	10,13 €	2,03 €	25,00%	84,25 €	8,42 €	1,32	10,69 €	2,59 €	26,94%	5-1	50	11	39
14/04/18	11/04/18	Barcelona	Valencia	La Liga	1x2	1	Betcllc	81,02 €	8,10 €	1,25	10,13 €	2,03 €	25,00%	84,25 €	8,42 €	1,30	10,53 €	2,43 €	25,01%	2-1	52	12	40
15/04/18	11/04/18	Juventus	Sampdoria	Serie A	1x2	1	ESC	81,02 €	8,10 €	1,18	9,56 €	1,46 €	18,00%	84,25 €	8,42 €	1,19	9,64 €	1,54 €	14,44%	3-0	55	12	43
15/04/18	11/04/18	PSG	Monaco	Ligue 1	1x2	1	CP	81,02 €	8,10 €	1,25	10,13 €	2,03 €	25,00%	84,25 €	8,42 €	1,28	10,37 €	2,27 €	23,09%	7-1	62	13	49
18/04/18	17/04/18	Napoli	Udinese	Serie A	1x2	1	Bet.pt	88,55 €	8,86 €	1,19	10,54 €	1,68 €	19,00%	93,08 €	9,31 €	1,22	10,80 €	1,95 €	16,07%	4-2	66	15	51
22/04/18	20/04/18	Sporting	Boavista	Liga NOS	1x2	1	Bet.pt	90,23 €	9,02 €	1,21	10,92 €	1,89 €	21,00%	95,03 €	9,50 €	1,24	11,19 €	2,17 €	17,75%	1-0	67	15	52
28/04/18	25/04/18	Bayern	Frankfurt	Bundesliga	1x2	1	Betcllc	92,13 €	9,21 €	1,45	13,36 €	4,15 €	45,00%	97,19 €	9,72 €	1,51	13,91 €	4,70 €	43,13%	4-1	71	16	55
06/05/18	02/05/18	Barcelona	Real Madrid	La Liga	1x2	1	Betcllc	96,27 €	9,63 €	1,65	-	-9,63 €	-100,00%	101,89 €	10,19 €	1,70	-	-10,19 €	-100,00%	2-2	73	18	55
06/05/18	02/05/18	Napoli	Torino	Serie A	1x2	1	CP	96,27 €	9,63 €	1,22	-	-9,63 €	-100,00%	101,89 €	10,19 €	1,22	-	-10,19 €	-100,00%	2-2	75	20	55
09/05/18	09/05/18	Barcelona	Villarreal	La Liga	1x2	1	Bet.pt	77,02 €	7,70 €	1,33	10,24 €	2,54 €	33,00%	81,51 €	8,15 €	1,37	10,55 €	2,85 €	26,45%	5-1	80	21	59
09/05/18	09/05/18	Man City	Brighton	Premier League	1x2	1	ESC	77,02 €	7,70 €	1,19	9,17 €	1,46 €	19,00%	81,51 €	8,15 €	1,19	9,17 €	1,46 €	12,44%	3-1	83	22	61
12/05/18	09/05/18	Bayern	Stuttgart	Bundesliga	1x2	1	ESC	77,02 €	7,70 €	1,19	-	-7,70 €	-100,00%	81,51 €	8,15 €	1,17	-	-8,15 €	-100,00%	1-4	84	26	58
12/05/18	09/05/18	PSG	Reims	Ligue 1	1x2	1	CP	77,02 €	7,70 €	1,22	-	-7,70 €	-100,00%	81,51 €	8,15 €	1,25	-	-8,15 €	-100,00%	0-2	84	28	56
25/05/18	19/05/18	Napoli	Crotone	Serie A	1x2	1	CP	65,62 €	6,56 €	1,44	9,45 €	2,89 €	44,00%	69,52 €	6,95 €	1,49	9,78 €	3,22 €	40,64%	2-1	86	29	57
25/05/18	19/05/18	Barcelona	Real Sociedad	La Liga	1x2	1	Bet.pt	65,62 €	6,56 €	1,28	8,40 €	1,84 €	28,00%	69,52 €	6,95 €	1,30	8,53 €	1,97 €	22,70%	1-0	87	29	58
Final Money	70,35 €													74,71 €						PIN-Accuracy M	4,36 €		
Final ROI	0,39%													6,62%						PIN-Accuracy ROI	6,23%		
Betting Accuracy	79,41%													79,41%						PIN-Accuracy AS%	0,00%		
Final Profit	0,28 €													4,64 €									

Figure 39- Betting record of the accuracy strategy (4 Portuguese bookmakers vs Pinnacle)

As explained in 5.2, we started with an initial amount of 70,07€ for each one of the strategies adopted. With the value strategy, we were able to obtain 13,93€, which corresponds to a Return on Investment (ROI) of 19,88%. The betting accuracy rate is equal to 72,73%, as we accurately predicted 8 out of 11 outcomes. In the Portuguese online fixed-odds betting markets, the accuracy strategy was able to produce 0,28€, which is equivalent to a ROI of 0,39%. This was accomplished by correctly forecasting 27 out of 34 results (accuracy rate of 79,41%). At last, if we have used Pinnacle’s website to execute the accuracy strategy, our money would have grown by 6,62%, that is, more 4,71€ than we began with. For convenience, from this point on, we will refer to this as if it was another different strategy, simply called Pinnacle.

6.2.1- Portuguese online fixed-odds betting markets efficiency

“In general terms, the theory of efficient markets is concerned with whether prices at any point in time fully reflect available information” (Fama, 1970: 413). The author describes 3 types of information-related market efficiency: strong-form efficiency, semi-strong-form efficiency and weak-form efficiency (these concepts are defined and explained in

the beginning of subchapter 4.2). Kuypers (2000) defines abnormal returns as the ones that are better than the bookmakers' overround): if one uses all the publicly available information to estimate probabilities for the outcomes of a match, and then is able to get abnormal returns from it, makes sense to conclude that the markets at stake are semi-strong efficient.

Theoretically, even negative returns can suggest semi-strong efficiency. For example, assuming that all the bookmaker's overround rates are equal and have a constant value of 9%, negative returns of -8% would indicate that the punter has developed a better aptitude to estimate outcome probabilities than the market. In another publication, Constantinou & Fenton (2013: 1) explain that "a gambling market is usually described as being inefficient if there are one or more betting strategies that generate profit, at a consistent rate, as a consequence of exploiting market flaws"⁷. According to these theories, either the accuracy strategy, the value strategy and Pinnacle were able to produce abnormal returns, once that all the yield rates produced by following these strategies were positive.

In reality, the bookmakers' overrounds are not equal and constant, thus one can use the average overround as a benchmark to evaluate the returns obtained. The accuracy strategy and Pinnacle would have to generate returns of at least - 8,10% and - 2,74%, respectively, to fit in the Kuypers (2000) definition of "abnormal returns". The former produced 0,39% and the latter 6,62%, which are values clearly above the threshold set by this author, and that, as he claims, indicate the existence of some form of semi-strong efficiency among these 2 markets. We did not record the odds offered for each outcome of the matches used in the value strategy, and as a consequence of that, we were unable to calculate the average overround taken by the bookmakers, in this case. However, as this strategy generated a ROI of 19,88%, it would certainly fit in the Kuypers (2000) notion of "abnormal returns".

As it is stated in the subchapter 4.2 of our work, "an example of a well-documented inefficiency in the fixed-odds betting markets is the favourite-longshot bias, which is a type of systemic error where favourites tend to win more often than what is implied by the odds (Cain *et al.*, 2000). Vlastakis *et al.* (2009) found this bias in a number of big online bookmakers. Based on the odds offered by several bookmakers over a period of 7 football seasons (from 2005/06 to 2011/12), Constantinou & Fenton (2013) demonstrated

⁷ Please consult the subchapter 4.2 of our work

the existence of a clear home advantage bias, whereby the returns produced by betting on home wins were substantially higher in comparison with the ones obtained by placing bets on away wins. In fact, the results of this study showed that this bias is almost equally as strong as the favorite-longshot bias. By using a dataset of 79 446 matches, disputed in 21 leagues of 11 different European countries (from 2000 to 2001), Direr (2011) claims that would have been possible to generate positive rates of return of 4.45% if best odds across all the bookmakers were selected, and 2.78% if mean odds were selected, by following a strategy of backing overwhelming favourites with odds of 1.19 (odds from 1.15 to 1.24 would have generated positive returns)”.

The accuracy strategy is about betting on home team favorites when the bookmakers offer odds equal or higher than 1.15 ⁸. If this strategy was able to produce positive returns, it could suggest the presence of at least one of these biases (or both at the same time) in the Portuguese online fixed-odds betting markets. As mentioned above, the accuracy strategy was able to get 0,28€, which in theory can be seen as a positive return. However, being the return so close to 0, one cannot say that such a low-income generation compensates the degree of risk taken, and therefore, one cannot conclude that the Portuguese betting markets have some form of inefficiency resulting from these biases, which can be exploited by the punters to profit consistently, in the long-term.

When it comes to the British online bookmaker, the situation is different. Betting in the same matches using Pinnacle’s platform would have provided 4,71€, which is equivalent to a ROI of 6,62%, hence better than the one obtained by Direr (2011), if best odds were chosen (although our number of observations is incomparably smaller). These numbers are clearly above the 0 threshold and consequently, it is possible to interpret them as a sign that at least one of these biases may exist in the odds offered by this bookmaker.

Figure 37 compares the odds of 4 Portuguese bookmakers vs Pinnacle, with and without the overround. If one looks at odds without the overround (right side of the tables), it is possible to perceive that Pinnacle offered better average odds for draws (6.51 vs 6.07) and away wins (12.83 vs 11.49), while the Portuguese bookmakers offered better average odds for home wins (1.41 vs 1.37). In contrast, if one considers the overround, Pinnacle offered better average odds for the home wins than the 4 Portuguese bookmakers (1.34

⁸ Please consult the subchapter 5.2.2 of our work

vs 1.31).

This suggests that, if the favorite-longshot bias and/or the home advantage bias exist in Pinnacle's odds, they also exist in the odds offered by the 4 Portuguese bookmakers. The problem is that the average overround is much higher in the latter's case (8,11% vs 2,75%), nullifying the possibility to take advantage of those biases, in a way that benefits the punters' interests. This is a clear example of the issues associated with the lack of competitiveness in the Portuguese fixed-odds betting markets, which is largely related with the country's current online gambling laws (as discussed in 6.1).

6.2.2- The importance of our forecasting method in appraising the expected value

The importance of doing a proper qualitative analysis to all the determinants of a match that are not easily quantifiable has been discussed over the course of our paper. In Rebelo's (2012) opinion, this type of evaluation is even more important than the quantitative analysis, and in case of conflict, should prevail over the latter. Forrest *et al.* (2005) claim that it is very probable that bookmakers consider subjective factors (determinants that are important for prediction but which historical data fails to capture) when setting the odds, and Constantinou *et al.* (2012) developed a Bayesian framework, which took into account both objective and subjective variables. The model was evaluated through accuracy and profitability measurements, and it was concluded that the subjective information improved significantly its performance. At the date, these authors even stated that, in terms of profitability, and over such a long period of time (one full season), no other published work appeared to be as successful as their own⁹.

In the value strategy, we did not follow some set of pre-determined rules to decide which bets to make. Instead, our focus was to maximize the expected value of each bet, through a combination of analytic methods and subjective interpretations of the information that is not expressed in numbers¹⁰. The bets made were individually selected after going through a detailed analysis, with well-defined criteria. That analysis is presented in the form of a report, whose contents are described in the subchapter 5.2.1.1. By following this strategy, we were able to obtain a ROI of 19,88% (as mentioned above), therefore 19,49 pp and 13,26 pp higher than the ones from the accuracy strategy and Pinnacle,

⁹ Please consult the subchapter 4.5 of our work

¹⁰ Please consult the subchapter 5.2.1 of our work

respectively, although the betting accuracy rate was 6,68% lower in the value strategy's case (72,73% vs 79,41%). Furthermore, the latter's average odd is considerably high (1.64) in comparison with the ones from other 2 strategies (1.31 and 1.34).

Starting from Rebelo's (2000) concept of positive expected value ¹¹, the average odd of the value strategy implies substantially more risk exposure than the accuracy strategy or Pinnacle's mean odds, at least according to the bookmakers' assessment. On the other hand, it also entails higher returns when the bets made are accurate. The average profit obtained by the accuracy strategy corresponds to 0,01€ (virtually none), which means that its expected returns, based on past trades, is approximately null. The standard deviation is 4,34€, that is, the observations tend to deviate from the mean about that value, up or down. Investors can use the standard deviation as a measure of risk: the more unpredictable the price action and the wider the range, the greater the risk. The coefficient of variation characterizes the relation between the standard deviation and the mean, determining how much risk is assumed in comparison with the amount of return expected: lower values indicate a better risk-return trade-off. Once that the average is so close to 0, in this case we should not consider the coefficient of variation, as it is too high for the interpretation to be meaningful.

Pinnacle would have generated an average profit of 0,14€, with a standard deviation equal to 4,61€. The value strategy, by its turn, produced, on average, 1,27€ for each bet placed, and its standard deviation is 6,05€, thus suggesting a higher risk than the one from Pinnacle (and consequently, also than the one from the accuracy strategy).

These results seem to indicate that, through a combination of analytic methods and subjective interpretations of all the aspects that can affect the outcomes of the matches selected, it is possible to beat the bookmakers which operate in the Portuguese online fixed-odds betting markets. Despite having a lower accuracy rate in comparison to the accuracy strategy (and Pinnacle), the value strategy produced much higher returns. Therefore, our forecasting method appears to increase the average returns while helping to better appraise and control the degree of risk assumed, hence maximizing the expected value of the bets made. However, the subjectivity of this strategy requires knowledge and specialization to predict the events accurately, and as a consequence of that, obtain

¹¹ Please consult the 4th page of our work

positive returns.

Nonetheless, it is important to highlight that we only have 11 observations for the value strategy, which is less than 3 times the number of observations for the accuracy strategy (and Pinnacle). Having this in consideration, we cannot take definitive conclusions about the contribution added by our forecasting method, and thus about the value strategy's ability to generate profits in the long-term. The reason for such a small number of observations is that each forecasting report, being so extensive, took several days to write, which constitutes a problem to the user-friendliness of our method. In order to clearly demonstrate to the readers how to perform this analysis, we had to do it in this way, however, the punters can use our reports as a model to do their own assessment, without needing to write such a detailed evaluation of all the situations that may condition the outcome of a football game.

7- Conclusions

7.1. Summarizing the value strategy

The main objective of this project was to develop an uncomplicated method, which can be used by ordinary punters to consistently obtain long-term positive returns. To achieve that, our method would have to be able to accurately predict the outcomes of matches with potentially valuable odds, within the Portuguese online fixed-odds betting markets. The forecasting of the football games was made combining quantitative information (including a Poisson regression model) with a qualitative analysis (thus mostly subjective) of the chosen matches. This approach was named value strategy, being able to generate a Return on Investment of nearly 20%, and to correctly predict 8 out of 11 results (which corresponds to an accuracy rate of 72,73%).

7.1.1- Limitations and recommendations for further research

Although the value strategy was able to produce significant yields, the number of bets made is too small to conclude that it is, in fact, capable of providing consistent earnings in the long-run. Besides that, due to the subjective nature of decision-making, it would be interesting to perceive if our forecasting method works for other punters and competitions (once that we focused on the Premier League's matches). Therefore, we will divulgate this paper in betting forums, and as a suggestion for further research, we challenge all the interested punters to use our forecasting report as a benchmark to make betting decisions, and subsequently, to give feedback about the returns obtained over a period of, at least, 1 entire season.

7.2- Summarizing the accuracy strategy

The accuracy strategy, which is based on "The Football Betting Mastermind" (Betinfo24: 2010), consisted in betting on the teams which were expected to win almost every match, once that according to this document, it is very unlikely that an individual punter is able to find enough value, in the major leagues' odds, to consistently beat the bookmakers. Summing up, this strategy is driven by accuracy, and not by finding positive expected value. The value strategy provided returns of 0,39%, thus not compensating for the level of risk assumed.

In spite of having an accuracy rate of 79,41%, that is higher than the one from the value

strategy, the returns obtained by the former were incomparably smaller, which suggests that making an individual appreciation of the matches confers protection against risk exposure, while increasing the expected returns. Comparatively, the value strategy's average odd implies much higher risks, while its returns imply a much higher expected value, in relation to the accuracy strategy. Perhaps if we have used the pre-determined rules of the latter to identify the matches with betting potential and then, from those, carefully selected the less risky alternatives, we would have been closer to the accuracy rate of 90%, which according to "The Football Betting Mastermind" (Betinfo24: 2010), is the benchmark to get significantly positive returns.

7.3- Portuguese online fixed-odds and markets efficiency

As we decided to bet only on overwhelming favourites who played at home, if the accuracy strategy was able to create positive returns, it would indicate the existence of the well-documented favourite-longshot and/or home advantage biases in the Portuguese online fixed-odds betting markets, and consequently, that these markets were semi-strong efficient. In practice, the accuracy strategy broke even, therefore not suggesting exploitable biases in the odds proposed by the 4 Portuguese bookmakers. However, if we have made the same bets through an UK-based online bookmaker, the rate of return would be equal to 6,62%, hence indicating the presence of this biases in the odds set by Pinnacle.

Amazingly, if we remove the weight of the overround from the odds, the Portuguese bookmakers would have offered better odds for the home wins than Pinnacle. This example clearly expresses the necessity of reviewing the Portuguese gambling legal framework (as has been claimed by punters and betting agencies), which was identified as the main reason for the poor competitiveness of Portuguese fixed-odds betting markets, in comparison with other countries. As far as our research went, we could not find any literature, studies nor even a record of past odds, about the Portuguese fixed-odds betting markets. Therefore, we consider that our work adds an important contribution, in the sense that it provides information about the odds offered by the Portuguese online betting markets in general, and also about each one of the 4 Portuguese bookmakers considered in this paper. Moreover, it offers a comparative analysis with the odds proposed by an UK-based online bookmaker, which highlights the lack of competitiveness of the Portuguese betting context.

7.4- Final comment

In spite of the difficulties encountered, if one is willing to invest a considerable amount of time and effort, we believe that it is indeed possible to find profitable opportunities within the Portuguese online fixed-odds. Rebelo (2012) states that the ability to accurately forecast football events continuously improves with practice, although not everyone is able to do it, being necessary to possess an innate gift, as it happens with professional sports players, for example. However, the number of licenced online bookmakers has been growing in Portugal, and as a consequence of that, the value implied by the odds is expected to increase, as well, creating new opportunities to exploit.

8- References

8.1- Books

- Cameron, A. C., & Trivedi, P. K. 1998. *Regression analysis of count data*. Cambridge (UK): Cambridge University Press.
- Clarke, G. M., & Cook, D. 2004. *A basic course in statistics*. London: Arnold.
- Curto, J. 2016. *Estatística, muitas aplicações em Excel e poucas fórmulas*. Lisbon: GUIDE- Artes Gráficas, Lda.
- Neave, N. & Wolfson, S. 2005. The Home Advantage: Psychological and Physiological Factors in Soccer. In D. Lavallee, J. Thatcher & M. Jones (Eds.), *Coping and Emotion in Sport*: 131-148. Hauppauge (NY): Nova Science Publishers.
- Rebelo, P. 2012. *Ganhar com as apostas desportivas*. Lisbon: Marcador.
- Weiss, N. A. 2006. *A course in probability*. Boston: Pearson Education Inc.

8.2- Periodicals

- Buraimo, B., Peel, D., & Simmons, R. 2013. Systematic Positive Expected Returns in the UK Fixed Odds Betting Market: An Analysis of the Fink Tank Predictions. *International Journal of Financial Studies*, 1(4): 168–182.
- Cain, M., David, L., & Peel, D. 2000. The favourite-longshot bias and market efficiency in UK football betting. *Scottish Journal of Political Economy*, 47(1): 25–36.
- Constantinou, A. C., & Fenton, N. E. 2013. Profiting from arbitrage and odds biases of the European football gambling market. *The Journal of Gambling Business and Economics*, 7(2): 41–70.
- Constantinou, A. C., Fenton, N. E., & Neil, M. 2012. Pi-football: A Bayesian network model for forecasting Association Football match outcomes. *Knowledge-Based Systems*, 36: 322–339.
- Courneya, K. S. & Carron, A. V. 1992. The Home Advantage in Sport Competitions: A Literature Review. *Journal of Sport & Exercise Psychology*, 14: 13-27.
- Direr, A. 2011. Are Betting Markets Efficient? Evidence from European Football Championships. *Applied Economics*, 45(3): 343–356.
- Dixon, M. J., & Coles, S. G. 1997. Modelling Association Football Scores and Inefficiencies in the Football Betting Market. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 46(2): 265–280.
- Fama, E. F. (1970). Efficient Capital Markets- A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2): 383–417.
- Forrest, D., Goddard, J., & Simmons, R. 2005. Odds-setters as forecasters: The case

of English football. *International Journal of Forecasting*, 21(3): 551–564.

- Forrest, D., McHale, I., & McAuley, K. 2008. “Say It Ain’t So”: Betting-Related Malpractice in Sport. *International Journal of Sport Finance*, 3(3): 156–166.
- Franck, E., Verbeek, E., & Nüesch, S. 2010. Prediction accuracy of different market structures- bookmakers versus a betting exchange. *International Journal of Forecasting*, 26(3): 448–459.
- Goddard, J. 2005. Regression models for forecasting goals and match results in association football. *International Journal of Forecasting*, 21(2): 331–340.
- Goddard, J., & Asimakopoulos, I. 2004. Forecasting Football Results and the Efficiency of Fixed-odds Betting. *Journal of Forecasting*, 23(1): 51–66.
- Heuer, A., & Rubner, O. 2009. Fitness, chance, and myths: An objective view on soccer results. *European Physical Journal B*, 67(3): 445–458.
- Hvattum, L. M., & Arntzen, H. 2010. Using ELO ratings for match result prediction in association football. *International Journal of Forecasting*, 26(3): 460–470.
- Kelly, J. L. 1956. A New Interpretation of Information Rate. *Bell System Technical Journal*, 35(4): 917–926.
- Kuypers, T. 2000. Information and efficiency: An empirical study of a fixed odds betting market. *Applied Economics*, 32(11): 1353–1363.
- Levitt, S. D. 2004. Why are gambling markets organized so differently from financial markets? *Economic Journal*, 114(3): 223–246.
- Maher, M. J. 1982. Modelling association football scores. *Statistica Neerlandica*, 36(3): 109–118.
- Mansuy, R. 2009. The origins of the word “martingale.” *Electronic Journal for History of Probability and Statistics*, 5(169): 105–113.
- McSharry, P. E. 2007. Altitude and athletic performance: Statistical analysis using football results. *British Medical Journal*, 335(7633): 1278–1281.
- Neave, N., & Wolfson, S. 2003. Testosterone, territoriality, and the “home advantage.” *Physiology and Behavior*, 78(2): 269–275.
- Page, L., & Page, K. 2007. The second leg home advantage: Evidence from European football cup competitions. *Journal of Sports Sciences*, 25(14): 1547–1556.
- Pollard, R. 1986. Home advantage in soccer: A retrospective analysis. *Journal of Sports Sciences*, 4(3): 237–248.
- Pollard, R. 2006. Worldwide regional variations in home advantage in association football. *Journal of Sports Sciences*, 24(3): 231–240.
- Pollard, R. 2008. Home Advantage in Football: A Current Review of an Unsolved Puzzle. *The Open Sports Sciences Journal*, 1(1):
- Pollard, R., & Pollard, G. 2005. Long-term trends in home advantage in professional

team sports in North America and England. *Journal of Sports Sciences*, 23(4): 337–350.

- Reep, C., Pollard, R., & Benjamin, B. 1971. Skill and Chance in Ball Games. *Journal of the Royal Statistical Society*, 134(4): 623–629.
- Seckin, N., & Pollard, R. 2008. Home Advantage in Turkish Professional Soccer. *Perceptual and Motor Skills*, 107: 51–54.
- Thaler, R. H., & Ziemba, W. T. 1988. Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries. *Journal of Economic Perspectives*, 2(2): 161–174.
- Thomas, S., Reeves, C., & Smith, A. 2006. English Soccer Teams Aggressive Behaviour When Playing Away From Home. *Perceptual and Motor Skills*, 1023: 17–320.
- Vlastakis, N., Dotsis, G., & Markellos, R. N. 2009. How efficient is the European football betting market? Evidence from arbitrage and trading strategies. *Journal of Forecasting*, 28(5): 426–444.
- Waters, A., & Lovell, G. 2002. An Examination of the Homefield Advantage in a Professional English Soccer Team from a Psychological Standpoint. *Football Studies*, 5(1): 46–59.

8.3- Unpublished works

- Gardner, J. A. 2011. *Modeling and Simulating Football Results*. Unpublished assignment in mathematics, University of Leeds School of Mathematics, UK. Retrieved from <https://www1.maths.leeds.ac.uk/~voss/projects/2010-sports/JamesGardner.pdf>
- Heino, O., & Sillanpää, V. 2013. *Forecasting football match results - A study on modeling principles and efficiency of fixed-odds betting markets in football*. Unpublished Master of Science dissertation, Aalto University School of Business, Helsinki. Retrieved from http://epub.lib.aalto.fi/fi/ethesis/pdf/13445/hse_ethesis_13445.pdf

8.4- Websites

- Apostas Online. (2017). Porque razão há tão poucas casas de apostas em Portugal? Retrieved November 18, 2017, from <https://apostasonline.pt/porque-razao-ha-tao-poucas-casas-de-apostas-em-portugal/>
- Batista, Rodrigo. (2017). *Apostas desportivas: O estado actual*. Retrieved January 21, 2018, from <https://blogvisaodmercado.pt/2017/12/apostas-desportivas-o-estado-actual/>
- Betinfo24. (2010). The Football Betting Mastermind. Retrieved November 16, 2017, from <http://betinfo24.co.uk/FootballMastermindB24.pdf>
- Correio da Manhã. (2018). Site “A Nossa Aposta” vai ter apostas desportivas. Retrieved May 27, 2018, from <http://www.cmjornal.pt/tv-media/detalhe/site-a-nossa->

[aposta-vai-ter-apostas-desportivas](#)

- Cotrim, António. (2018). Site de apostas “placard.pt” já abriu e vai canalizar lucros para causas sociais. Retrieved June 14, 2018, from <https://observador.pt/2018/06/11/site-de-apostas-placard-pt-ja-abriu-e-vai-canalizar-lucros-para-causas-sociais/>
- European Gaming and Betting Association. Sports Betting: Commercial and Integrity Issues. Retrieved November 22, 2017, from <http://www.egba.eu/media/Sports-Betting-Report-FINAL.pdf>
- HM Revenue & Customs. (2017). General Betting Duty, Pool Betting Duty and Remote Gambling Duty. Retrieved September 28, 2018, from <https://www.gov.uk/guidance/general-betting-duty-pool-betting-duty-and-remote-gaming-duty>
- Jornal Económico. (2017). Direitos de TV na Premier League: último classificado encaixou mais do que o Benfica. Retrieved October 24, 2018, from <https://jornaleconomico.sapo.pt/noticias/direitos-de-tv-na-premier-league-ultimo-classificado-encaixou-mais-do-que-benfica-166640>
- Kucharski, Adam. (2016). Betting and investment both require skill and luck. Retrieved November 15, 2017, from <https://www.ft.com/content/a472722e-0e25-11e6-b41f-0beb7e589515>
- Macário, José. (2017). Direitos televisivos: como são distribuídos nas principais ligas? Retrieved October 24, 2018, from <https://jornaleconomico.sapo.pt/noticias/direitos-televisivos-como-sao-distribuidos-nas-principais-ligas-193225>
- Madremedia/Lusa. (2017). Jogo 'online': Associação internacional pede "revisão urgente" do regulamento português. Retrieved November 18, 2017, from <http://24.sapo.pt/atualidade/artigos/jogo-online-associacao-internacional-pede-revisao-urgente-do-regulamento-portugues>
- Moody, Allen. (2017). History of Sports Betting. Retrieved November 15, 2017, from <https://www.thoughtco.com/history-of-sports-betting-3116857>
- O'Shaughnessy, D. 2011. Optimal Exchange Betting Strategy for Win-Draw-Loss Markets. Retrieved December 12, 2017, from <http://www.rankingsoftware.com/research/OptimalBettingStrategyWinDrawLoss.pdf>
- Schoch, David. (2017). Home-field Advantage. Retrieved January 24, 2018, from http://soccerverse.com/articles/home_field_advantage
- Silva, Ana Rute. (2014). Operadores dizem que impostos sobre jogo online tornam “inviável” acesso ao mercado. Retrieved November 18, 2017, from <https://www.publico.pt/2014/07/03/economia/noticia/operadores-dizem-que-impostos-sobre-jogo-online-tornam-inviavel-acesso-ao-mercado-1661529>
- Stradbroke, Steven. (2017). RGA: Portugal's regulated online gambling market a failure. Retrieved July 5, 2018, from

<https://calvinayre.com/2017/12/06/business/portugal-online-gambling-market-failure/>

- Stradbroke, Steven. (2018). Portugal's online gambling market closed out 2017 on a high. Retrieved July 5, 2018, from <https://calvinayre.com/2018/02/16/business/portugal-online-gambling-record-revenue/>

Annexes

Annex I- Watford vs Everton (match day 28)

- 24/02/2018 at 5:30 pm, Vicarage Road Stadium (Watford)
- Weather: 3°C; precipitation probability of 0%; humidity percentage of 58% and wind speed of 21 km/h
- **Analysis:** Weather as perfect as it gets to play football in the UK, at this time of the year. It allows for a fast circulation of the ball, and consequently for an open match.

Potential bets:

- Over 2.5 goals in the match- 1.88
- Both teams score- 1.71 (yes) and 1.80 (no)
- Watford or draw and over 1.5 goals in the match- 1.70
- Watford or draw and both teams score- 2.20

Probable line-ups:

- **Watford-** 3-4-2-1: Gomes; Mariappa, Prodl and Holebas; Femenía, Capoue, Doucouré and Zeegelaar; Deulofeu and Richarlison; Deeney
- **Everton-** 4-2-3-1: Pickford; Coleman, Keane, Williams and Martina; Gueye and Schneiderlin; Walcott, Davies and Sigurdsson; Tosun

Injuries and suspensions:

- **Watford-** Out: Hoban, Catchcart and Chalobah; doubtful: Kabasele and Cleverly (all injured)
- **Everton-** Out: Baines, McCarthy, Stekelenburg, Mangala; doubtful: Coleman, Gueye and Jagielka (all injured)
- **Analysis-** No key players missing on both teams, as many of these are virtually chronically injured.

Forecasting Football Outcomes to Invest in Betting Markets

Discipline table:

- **Watford** (average per match at home)- Yellow cards: 1.9, red cards: 0.2, booking points: 23.1
- **Everton** (average per match away)- Yellow cards: 1.4, red cards: 0.1, booking points: 15.8
- **Premier League** (team average per match overall)- Yellow cards: 1.6, red cards: 0.1, booking points: 16.7
- **Analysis-** At home, Watford is more aggressive than the average of Premier League teams, while Everton is below the average in terms of aggressiveness, when playing away. This means that Watford may win more duels and recover the ball more swiftly, but on the other hand, The Hornets are probably going to be more exposed to disciplinary conditioning through yellow or red cards.

Head-to-head over the last 3 years:

- **5 matches-** 2 wins for Everton (40%), 2 draws (40%) and 1 win for Watford (20%); 4 times over 1.5 (80%) and 1 time under (20%); 3 times over 2.5 (60%) and 2 times under (40%); neither of the teams won away
- **Analysis-** There is not a significant bias favouring any of the teams.

Premier league tables overall, home and away:

- **Overall-** Watford: 11th place, 30 points, 37 goals scored and 47 conceded; Everton: 9th place, 34 points, 32 goals scored and 46 conceded
- **Home-** Watford: 18th place, 16 points, 20 goals scored and 26 conceded
- **Away-** Everton: 17th place, 8 points, 9 goals scored and 28 conceded
- **Analysis-** Both teams performed in a very similar way this season, with a big difference: while Watford does not rely much on the home advantage (won 16 points at home and 14 when playing away), Everton is hugely dependent on this factor (only 1 win away from Goodison Park). This can influence the match in Watford's favour.

Forecasting Football Outcomes to Invest in Betting Markets

Home/away performance tables:

- **Watford-** won 53% of its points at home and 47% when playing away, scored 54% of its goals at home and 46% when playing away, conceding 55% of its goals at home and 45% away
- **Everton-** won 76% of its points at home and 24% when playing away, scored 72% of its goals at home and 28% when playing away, conceding 39% of its goals at home and 61% away
- **Premier League-** On average, 57% of the goals were scored at home and 43% were scored away
- **Analysis-** This table provides numerical evidence that Everton is very reliant on the home advantage, even in comparison with the Premier League's average.

Points per game tables:

- Watford has a relative form of -0.11, while Everton scores -0.26. The obvious conclusion is that none of the teams have performed good recently, but Everton has been even less consistent than Watford

Scoring tables:

- **Watford** scores, on average, 1.45 goals per match at home, while Everton concedes 1.70 goals per match when playing away. If we average these numbers, we obtain Watford's "scoring strength", which corresponds to 1.62 goals. Besides that, Watford's matches at home ended with over 2.5 goals 69% of the times
- **Everton** scores, on average, 1.19 goals per match when playing away, while Watford concedes 2.00 goals per match at home. If we average these numbers, we obtain Everton's "scoring strength", which corresponds to 1.60 goals. Besides that, Everton's away matches ended with over 2.5 goals 46% of the times
- **Analysis-** If we sum Watford and Everton scoring strengths, we obtain a total of 3.22 goals expected in the forthcoming match. Besides that, if we average both teams' percentage of matches that ended with over 2.5 goals, we get a 57.5% probability that this match in particular ends with over 2.5 goals, which can be converted into an odd of 1.74.

Forecasting Football Outcomes to Invest in Betting Markets

Team characteristics:

- Both teams are strong at stealing the ball from the opposition and coming back from losing positions (Everton is even stronger in the latter)
- Watford is very weak at protecting the lead and defending counter-attacks. This team uses an offside trap defensive strategy, which means that the defensive line must be positioned high in the field. As Everton has very fast players (e.g. Walcott), and attacks mostly through long balls, there might be some space to exploit behind Watford's defensive line
- Everton is very weak at defending against skilful players. Watford's front line is rich in players with such characteristic (e.g. Deulofeu and Richarlison)
- Both teams are weak at avoiding fouling in dangerous areas, so this match might be prone to harmful individual mistakes from both sides
- Watford likes to attack down the left, while Everton likes to attack down the right, so we can expect most of the action to be developed on that wing. However, Everton is very weak when it comes to defend attacks down the flanks

Team psychology:

- **Team spirit and motivation-** As this is a direct confrontation between two teams with the same ambitions in terms of classification, and with just 4 points separating Watford from Everton, both teams should be highly motivated to win the match
- **Confidence-** In spite of being highly motivated for this match in particular, both teams are very inconsistent, as they easily go from an incredible victory to a surprising defeat in the following match day. Besides that, both teams have problems in their defensive processes, hence we believe that their confidence levels should be classified as low
- **Managerial impact-** Both clubs have experienced managerial sackings over the course of the season. Everton sacked Ronald Koeman after the defeat at home against Arsenal (2-5), on the 22nd October, that left the Liverpool based side in relegation zone (18th place). On the other hand, Marco Silva was fired on the 20th

Forecasting Football Outcomes to Invest in Betting Markets

January, after suffering a categorical defeat against Leicester (2-0 away). In spite of being criticized by many for his outdated style of play, Sam Allardyce was able to put Everton in 9th place, with 5 wins, 4 draws and 4 losses. Because of this clear improvement in Everton's results, we consider that he had a high impact on the team's performance. New Watford's boss Javi Gracia started by drawing with no goals away against Stoke, achieved an incredible win at home against Chelsea (4-1), but then lost away to direct competitor West Ham (2-0). So far, we consider that his managerial impact on Watford has been normal

Team fatigue:

- **Toughness of previous match-** Watford loss 2-0 away against West Ham, while Everton won 3-1 at home against Crystal Palace. At match day 26, Everton had 31 points, Watford had 30, West Ham and Crystal Palace both had 27, and none was safe in terms of avoiding relegation. The toughness of these matches is considered high because the teams involved were direct competitors for the safety of the mid-table places
- **Resting-** Both teams are very rested, as the last Premier League's match day finished at 12/02/2018
- **National team participation-** The last round of international matches was in November of 2017, and so this will not be considered

Forecasting Football Outcomes to Invest in Betting Markets

Poisson regression model:

	0	1	2	3	4	5	6	7+	Home
0	5,16%	7,02%	4,77%	2,16%	0,74%	0,20%	0,05%	0,01%	
1	8,27%	11,26%	7,66%	3,47%	1,18%	0,32%	0,07%	0,02%	
2	6,64%	9,03%	6,14%	2,79%	0,95%	0,26%	0,06%	0,01%	
3	3,55%	4,83%	3,28%	1,49%	0,51%	0,14%	0,03%	0,01%	
4	1,42%	1,94%	1,32%	0,60%	0,20%	0,06%	0,01%	0,00%	
5	0,46%	0,62%	0,42%	0,19%	0,07%	0,02%	0,00%	0,00%	
6	0,12%	0,17%	0,11%	0,05%	0,02%	0,00%	0,00%	0,00%	
7+	0,03%	0,05%	0,03%	0,01%	0,00%	0,00%	0,00%	0,00%	
Away									
	5,16%	11,26%	6,14%	1,49%	0,20%	0,02%	0,00%	0,00%	
Home win	32,50%		Odd	3,08					
Draw	24,27%		Odd	4,12					
Away win	43,23%		Odd	2,31					
Total	100,00%								
Under 2.5	43,12%		Odd	2,32					
Over 2.5	56,88%		Odd	1,76					
Total	100,00%								

- Analysis-** Our Poisson regression model grants more chances to an away win, in contrast with the bookmaker's odds, and the descriptive statistics as well. This is probably because in 2016/2017, Everton performed very good, finishing in 7th place, while Watford ended 17th, just 6 points above the relegation zone. It is possible to detect a data pattern that is relevant for betting on over 2.5 goals, for which the model provides a probability that is very similar to the one provided by the descriptive statistics (56,88% and 57,50%, respectively).

Betting decision:

- Our analysis indicates that we can expect an open match, as both teams should be very motivated to win, in spite of the low confidence levels that are a consequence of a very inconsistent defensive process. Usually, Everton has troubles when playing away, thus we believe in a home win or draw. Watford and Everton have attacking players with high quality, that can be decisive against most of the Premier League's teams. Our descriptive scoring analysis determines a probability of 57.5% to the match ending with over 2.5 goals, and the Poisson regression model determines a 56.88% probability to the same outcome. According to our

Forecasting Football Outcomes to Invest in Betting Markets

judgement, this means that an offered odd above 1.76 on over 2.5 goals would have positive expected value. Therefore, we think that betting on over 2.5 goals at 1.88 is the best possible choice.

Annex II- Arsenal vs Manchester City (match day 28)

- 01/03/2018 at 7:45 pm, Emirates Stadium (London)
- Weather: -2°C; precipitation probability of 48%; humidity percentage of 64% and wind speed of 32 km/h
- **Analysis:** Acceptable weather conditions to play football in the UK, at this time of the year. The match was at risk due to the eastern cold wave that painted in white the Emirates Stadium surrounding area, however, it is equipped with undersoil heating, so there is no threat of the pitch being frozen and, as a consequence of that, unplayable.

Potential bets:

- Manchester City- 1.75

Probable line-ups:

- **Arsenal**- 4-3-3: Cech; Bellerín, Koscielny, Mustafi and Kolassinac; Ramsey, Elneny and Wilshere; Ozil, Aubameyang and Mkhitarjan
- **Manchester City**- 4-3-3: Ederson; Walker, Stones, Laporte and Danilo; De Bruyne, Gundogan and D. Silva; B. Silva, Aguero and Sané

Injuries and suspensions:

- **Arsenal**- Out: Cazorla, Lacazette and Monreal (all injured)
- **Manchester City**- Out: Mendy and Fernandinho (injured), Delph (suspended); doubtful: Sterling (injured)
- **Analysis**- 2 important players injured on each side: Lacazette and Monreal for Arsenal, Fernandinho and Sterling for Manchester City.

Discipline table:

- **Arsenal** (average per match at home)- Yellow cards: 1.2, red cards: 0.0, booking points: 12.3
- **Manchester City** (average per match away)- Yellow cards: 1.9, red cards: 0.1,

Forecasting Football Outcomes to Invest in Betting Markets

booking points: 21.2

- **Premier League** (team average per match overall)- Yellow cards: 1.6, red cards: 0.1, booking points: 16.6
- **Analysis-** At home, Arsenal is less aggressive than most Premier League's teams, while Manchester City is way above the average in terms of aggressiveness, when playing away. This means that The Citizens may win more duels and recover the ball more swiftly, but on the other hand, they are probably going to be more exposed to disciplinary conditioning through yellow or red cards.

Head-to-head over the last 3 years:

- **11 matches-** 4 wins for Arsenal (36,36%), 4 draws (36,36%) and 3 wins for Manchester City (27,28%)
- **Analysis-** There is not a significant bias favouring any of the teams. On the 25th of February, Manchester City won the Carabao Cup by defeating Arsenal 3-0 at Wembley, in a match where The Citizens clearly overpowered The Gunners.

Premier league tables overall, home and away:

- **Overall-** Arsenal: 6th place, 45 points, 51 goals scored and 36 conceded; Manchester City: 1st place, 72 points, 79 goals scored and 20 conceded
- **Home-** Arsenal: 3rd place, 32 points, 36 goals scored and 14 conceded
- **Away-** Manchester City: 1st place, 32 points, 29 goals scored and 10 conceded
- **Analysis-** Manchester City has been, by far, the strongest team in the Premier League this season, placing 1st with the best goal difference at home and away. On the other hand, Arsenal has been very consistent at home, being able to match Manchester City's away performances. However, Arsenal relies too much on the home advantage, as around 71% of its Premier League points were earned when playing at the Emirates Stadium.

Home/away performance tables:

- **Arsenal-** won 71% of its points at home and 29% when playing away, scoring also 71% of its goals at home and 29% when playing away, conceding 39% of its

Forecasting Football Outcomes to Invest in Betting Markets

goals at home and 61% away

- **Manchester City**- won 56% of its points at home and 44% when playing away, scored 63% of its goals at home and 37% when playing away, conceding 50% of its goals at home and 50% away
- **Premier League**- On average, 57% of the goals were scored at home and 43% were scored away
- **Analysis**- This table provides numerical evidence that Arsenal is very reliant on the home advantage, even in comparison with the Premier League's average. Manchester City achieved very good results away, and even better results at home, hence we can assume that the home advantage has its influence on the team's performance, but The Citizens do not depend on it to obtain good results.

Points per game tables:

- Arsenal has a relative form of -0.29, while Manchester City scores -0.54. At first glance, it seems weird that Arsenal has a less negative relative form in comparison with City. This happens because Manchester City has been very consistent over the course of the season, only losing against Liverpool, and drawing against Everton, Crystal Palace and Burnley. All these matches are included in the last 8 matches, with the exception of Manchester City vs Everton

Scoring tables:

- **Arsenal** scores, on average, 2.77 goals per match at home, while Manchester City concedes 0.77 goals per match when playing away. If we average these numbers, we obtain Arsenal's "scoring strength", which corresponds to 1.77 goals
- **Manchester City** scores, on average, 2.33 goals per match when playing away, while Arsenal concedes 1.08 goals per match at home. If we average these numbers, we obtain Manchester City's "scoring strength", which corresponds to 1.71 goals

Team characteristics:

- Both teams like to play a possession-based football, using short passes to control the game in the opposition's half, but usually Wenger adapts Arsenal style of play

Forecasting Football Outcomes to Invest in Betting Markets

when facing Guardiola's teams. Therefore, we should expect Manchester City to dominate this match, against a more speculative Arsenal. City defends high on the pitch, using the offside trap strategy. Although Manchester City has this process very well automatized, the team is prone to individual mistakes (like Arsenal), so there might be some space to exploit behind The Citizens' defensive line, as Arsenal has very fast offensive players (e.g. Aubameyang and Mkhitarjan)

- Arsenal creates a lot of scoring chances, and Manchester City is very weak when it comes to stopping opponents from creating chances (the opposite is also truth). Arsenal is strong at finishing scoring chances, while Manchester City is very strong
- Manchester City and Arsenal are very strong at creating chances using through balls. However, Arsenal is weak at defending against through balls
- City is strong in counter-attacks, while Arsenal is weak at defending them
- Arsenal is weak at protecting the lead, but is strong at coming back from losing positions

Team psychology:

- **Team spirit and motivation-** Pep Guardiola promotes a squad rotation policy, so that every player feels useful and integrated. There are no cases (at least publicly known) of indiscipline or internal issues between the players and the manager. The way the team is performing (and winning) provides strong evidence of a positive team culture. Apart from that, Manchester City should be highly motivated to win the Premier League as soon as possible, in order to then focus its full attention on the latter stages of the Champions League. On the contrary, Arsenal is 8 points above the 7th placed Burnley and 10 points below the Champions League spots, so the team's faith is more or less defined. Arsène Wenger is a spent manager, that can no longer produce good results and inculcate a positive spirit among the players. Nonetheless, Arsenal is hurt in its pride after the humiliating defeat in the League Cup final, and surely will try to clean up the image left in that match. For those reasons, we classify Arsenal's team spirit and motivation as normal

Forecasting Football Outcomes to Invest in Betting Markets

- **Confidence-** In 2017/2018, Manchester City has been perhaps the best team in Europe, only losing to Shakhtar Donetsk (Champions League), Wigan (FA Cup) and Liverpool (Premier League), and tying against Everton, Crystal Palace and Burnley (all for Premier League). On the last 25th of February, City won 3-0 to Arsenal, conquering the League Cup. On these circumstances, we believe it is fair to say that The Citizens' confidence levels must be very high. Arsenal, by its turn, is performing below the expectations, ranking 6th in the Premier League, and coming from traumatic losses in the Northern London derby (vs Tottenham) and in the League Cup final (vs Manchester City). Therefore, Arsenal's confidence levels should be low at this point
- **Managerial impact-** Although Arsenal do not win the Premier League since 2003/2004, Arsène Wenger has been in charge for over 22 years. For better or worse, his impact on the club has to be considered very high. Guardiola arrived at Manchester and after a mediocre first season, has been able to turn Manchester City into one of the most complete teams in Europe. The financial investment in the squad was huge, but nowadays City plays in the image of its coach, and is definitely one of the best sides that he built in the entire career. That being said, we have to consider Guardiola's impact on the club as very high, as well

Team fatigue:

- **Toughness of previous match-** Arsenal loss the final of the League Cup 3-0 against Manchester City. Considering this, we classify the toughness of Arsenal's last match as the highest, and the toughness of Manchester City's last match as low
- **Resting-** Both teams last match was against each other, for the League Cup final (4 days ago), as mentioned above. From today's probable line-up, the only players that rested were the goalkeeper Ederson, and defenders John Stones and Aymeric Laporte. Bernardo Silva started on the bench, but replaced the injured Fernandinho by the minute 52. Arsenal rested Petr Cech, plus the defender Kolasinac (who entered by the minute 26, replacing the injured Monreal) and the midfielder Elneny. Aaron Ramsey was replaced by Iwobi, on the 73rd minute
- **National team participation-** The last round of international matches was in

Forecasting Football Outcomes to Invest in Betting Markets

November of 2017, and so this will not be considered

Poisson regression model:

	0	1	2	3	4	5	6	7+	Home
0	3,32%	5,19%	4,06%	2,12%	0,83%	0,26%	0,07%	0,02%	
1	6,11%	9,56%	7,48%	3,90%	1,53%	0,48%	0,12%	0,03%	
2	5,62%	8,80%	6,88%	3,59%	1,40%	0,44%	0,11%	0,03%	
3	3,45%	5,40%	4,22%	2,20%	0,86%	0,27%	0,07%	0,02%	
4	1,59%	2,48%	1,94%	1,01%	0,40%	0,12%	0,03%	0,01%	
5	0,58%	0,91%	0,72%	0,37%	0,15%	0,05%	0,01%	0,00%	
6	0,18%	0,28%	0,22%	0,11%	0,04%	0,01%	0,00%	0,00%	
7+	0,06%	0,09%	0,07%	0,04%	0,02%	0,00%	0,00%	0,00%	
Away									
	3,32%	9,56%	6,88%	2,20%	0,40%	0,05%	0,00%	0,00%	
Home win	33,05%		Odd	3,03					
Draw	22,41%		Odd	4,46					
Away win	44,47%		Odd	2,25					
Total	100%								

- Analysis-** Our Poisson regression model grants more chances to an away win, which is expectable, as Manchester City performed considerably better than Arsenal over the last 2 seasons. Nevertheless, the model says that betting on a City's win would only have positive expected value if the bookmakers offer an odd over 2.25, which is considerably higher than the best odd offered by the Portuguese bookmakers (1.75).

Betting decision:

- Overall, our analysis suggests that Manchester City is, these days, a much more reliable, confident and motivated team than Arsenal (as the League Cup match confirms), and in spite of the latter's good performances at home, City should win this match, with more or less difficulties. If The Citizens score first, Arsenal will be forced to risk more, exposing themselves to counter attacks. The level of fatigue is the same for both teams, but Manchester City defensive balance can be affected by the absence of Fernandinho. Although our Poisson regression model indicates that an odd of 1.75 is very distant from the positive expected value threshold (2.25), we still believe that betting on a Manchester City's win offers a relatively good reward for the level of risk assumed.

Annex III- Manchester City vs Chelsea (match day 29)

- 04/03/2018 at 4:00 pm, Etihad Stadium (Manchester)
- Weather: 5°C; precipitation probability of 9%; humidity percentage of 70% and wind speed of 25 km/h
- **Analysis:** Weather as perfect as it gets to play football in the UK, at this time of the year. It allows for a fast circulation of the ball, and consequently for an open match.

Potential bets:

- Manchester City- 1.57

Probable line-ups:

- **Manchester City**- 4-3-3: Ederson; Walker, Kompany, Otamendi and Danilo; De Bruyne, Gundogan and D. Silva; Sterling, Aguero and Sané
- **Chelsea**- 3-4-2-1: Courtois; Azpilicueta, Christensen and Rudiger; Moses, Kanté, Fàbregas and Alonso; Willian and Hazard; Morata

Injuries and suspensions:

- **Manchester City**- Out: Mendy (injured) and Delph (suspended); doubtful: Sterling, Walker and Fernandinho (all injured)
- **Chelsea**- Doubtful: Bakayoko, Barkley and D.Luiz (all injured)
- **Analysis**- Manchester City has 3 important players in doubt, while Chelsea has 3 medium-importance players in doubt

Discipline table:

- **Manchester City** (average per match at home)- Yellow cards: 1.3, red cards: 0.1, booking points: 14.6
- **Chelsea** (average per match away)- Yellow cards: 1.1, red cards: 0.1, booking points: 13.2
- **Premier League** (average per match)- Yellow cards: 1.6, red cards: 0.1, booking

Forecasting Football Outcomes to Invest in Betting Markets

points: 16.6

- **Analysis-** Both teams are considerably below the Premier League average in terms of aggressiveness. It seems unlikely that cards or injuries might condition the outcome of this match.

Head-to-head over the last 3 years:

- **8 matches-** 3 wins for Manchester City (37,5%), 3 wins for Chelsea (37,5%) and 2 draws (25%)
- **Analysis-** There is not a significant bias favouring any of the teams.

Premier league tables overall, home and away:

- **Overall-** Manchester City: 1st place, 75 points, 82 goals scored and 20 conceded, Chelsea: 5th place, 53 points, 50 goals scored and 25 conceded
- **Home-** Manchester City: 1st place, 40 points, 50 goals scored and 10 conceded
- **Away-** Chelsea: 3rd place, 24 points, 26 goals scored and 15 conceded
- **Analysis-** Manchester City has been, by far, the strongest team in the Premier League this season, placing 1st with the best goal difference, at home and away. On the other hand, Chelsea has been performing more or less in equal manner at home and away, so it is fair to assume that none of the teams is relying hugely on the home advantage factor.

Home/away performance tables:

- **Manchester City-** won 53% of its points at home and 47% when playing away, scored 61% of its goals at home and 39% when playing away, conceding 50% of its goals at home and 50% away
- **Chelsea-** won 55% of its points at home and 45% when playing away, scoring also 48% of its goals at home and 52% when playing away, conceding 40% of its goals at home and 60% away
- **Premier League-** On average, 57% of the goals were scored at home and 43% were scored away

Forecasting Football Outcomes to Invest in Betting Markets

- **Analysis-** This table provides numerical evidence indicating that neither of the teams is very reliant on home advantage. Manchester City achieved very good results away, and even better results at home, hence we can conclude that the home advantage has its influence on the team's performance (specially on the number of goals scored), but City do not depend on it to obtain good results. Regarding Chelsea's performance, it is possible to say exactly the same, with the difference that the overall results obtained are not nearly as good as the ones achieved by Manchester City. In this case, the influence of the home advantage on Chelsea's performance is mostly related with suffering less goals at home.

Points per game tables:

- Manchester City has a relative form of -0.55, while Chelsea scores -0.51. At first glance, it seems weird that Chelsea has a less negative relative form in comparison with City. This happens because Manchester City has been very consistent over the course of the season, only losing against Liverpool, and drawing against Everton, Crystal Palace and Burnley. All these matches are included in the last 8 matches, with the exception of Manchester City vs Everton

Scoring tables:

- **Manchester City** scores, on average, 3.57 goals per match at home, while Chelsea concedes 1.07 goals per match when playing away. If we average these numbers, we obtain Manchester City's "scoring strength", which corresponds to 2.32 goals
- **Chelsea** scores, on average, 1.86 goals per match when playing away, while Manchester City concedes 0.71 goals per match at home. If we average these numbers, we obtain Chelsea's "scoring strength", which corresponds to 1.29 goals

Team characteristics:

- Both teams like to play a possession-based football, using short passes to control the game in the opposition's half, but usually Conte adapts Chelsea style of play when facing Guardiola's teams. Therefore, we should expect Manchester City to dominate possession in this match, against a more speculative Chelsea. City defends high on the pitch, using the offside trap strategy. Although Manchester City has this process very well automatized, the team is prone to individual

Forecasting Football Outcomes to Invest in Betting Markets

mistakes, so there might be some space to exploit behind The Citizens defensive line. However, Chelsea is very weak in avoiding offside.

- Manchester City is strong at attacking set pieces, but Chelsea is equally strong at defending them
- Chelsea likes to attack down the left, while Manchester City prefers to attack down the right, so we can expect most of the action to be developed on that wing
- Both teams are weak at disputing aerial duels
- Chelsea is strong at creating scoring chances, while The Citizens are very weak at stopping opponents from creating scoring chances

Team psychology:

- **Team spirit and motivation-** Pep Guardiola promotes a squad rotation policy, so that every player feels useful and integrated. There are no cases (at least publicly known) of indiscipline or internal issues between the players and the manager. The way the team is performing (and winning) provides strong evidence of a positive team culture. Besides that, Manchester City should be highly motivated to win the Premier League as soon as possible, in order to then focus its full attention on the latter stages of the Champions League. Chelsea is the current Premier League champion but the results have been disappointing this season. The possible causes for this decline are, most likely, related with the selling of top players, such as Matic or Diego Costa, which were not replaced by new players with, at least, the same level in terms of quality. In addition, Antonio Conte is having problems with the board, and simultaneously, with key players like Courtois, Morata, David Luiz or Hazard. Due to these internal issues, plus the results below the expectations, we have to classify Chelsea's team spirit and motivation as low
- **Confidence-** In 2017/2018, Manchester City has been perhaps the best team in Europe, only losing to Shakhtar Donetsk (Champions League), Wigan (FA Cup) and Liverpool (Premier League), and tying against Everton, Crystal Palace and Burnley (all for Premier League). On these circumstances, we believe it is fair to say that The Citizens' confidence levels must be very high. Chelsea, on the other

hand, is going through a complicated period, dropping down to 5th in the Premier League, out of the 2018/2019 Champions League qualifying places. On top of that, probably Chelsea will be eliminated from the 2017/2018 Champions League, after an upsetting 1-1 at home against Barcelona. The team will play the FA Cup quarter-finals, still having a chance to win that competition, which is clearly not enough for a club with Chelsea's ambitions. Given the overall context of the team's behaviour in this season, we think that its current confidence level should be classified as low

- **Managerial impact-** Guardiola arrived at Manchester and, after a mediocre first season, has been able to turn Manchester City into one of the most complete teams in Europe. The financial investment in the squad was huge, but nowadays City plays in the image of its coach, and is definitely one of the best sides that he built in the entire career. That being said, we have to consider Guardiola's impact on the club as very high. Antonio Conte achieved an astonishing Premier League winning in his first season as Chelsea's manager, 7 points ahead of the 2nd ranked Tottenham Hotspurs. Conte introduced the 3-4-2-1 tactic in the English top division, which is commonly used in Italy and was subsequently adopted by Arsenal, Watford and occasionally by Manchester rivals, Tottenham and Liverpool, for example. Although this season is not being nearly as good as the past one, we still have to consider the impact of Antonio Conte on Chelsea as high, at least

Team fatigue:

- **Toughness of previous match-** Manchester City won comfortably 3-0 vs Arsenal at Emirates, while Chelsea lost 2-1 against direct competitors Manchester United, at Old Trafford, in a match where The Blues scored first. Having this scores in consideration, we decided to rate the toughness of previous matches as medium for Manchester City, and as the highest for Chelsea
- **Resting-** In City's last match, from the probable line-up for today, the only player that rested was Sterling, because he was injured. In this match, Sterling is expected to replace Bernardo Silva in the starting 11. During the match, Guardiola substituted Walker for Zinchenko (72nd minute), Aguero for Touré (82nd minute) and D.Silva for Jesus (87th minute). Against Manchester United, Antonio Conte

Forecasting Football Outcomes to Invest in Betting Markets

only rested Fàbergas, although the latter was called to replace Drinkwater on the 81st minute. Apart from that, Conte replaced Hazard for Pedro (73rd minute) and Moses for Giroud (78th minute)

- **National team participation-** The last round of international matches was in November of 2017, and so this will not be considered

Poisson regression model:

	0	1	2	3	4	5	6	7+	Home
0	5,19%	9,31%	8,35%	5,00%	2,24%	0,81%	0,24%	0,08%	
1	6,04%	10,84%	9,73%	5,82%	2,61%	0,94%	0,28%	0,09%	
2	3,51%	6,31%	5,66%	3,39%	1,52%	0,55%	0,15%	0,05%	
3	1,36%	2,45%	2,20%	1,31%	0,59%	0,21%	0,06%	0,02%	
4	0,40%	0,71%	0,64%	0,38%	0,17%	0,06%	0,02%	0,01%	
5	0,09%	0,17%	0,15%	0,09%	0,04%	0,01%	0,00%	0,00%	
6	0,02%	0,03%	0,03%	0,02%	0,01%	0,00%	0,00%	0,00%	
7+	0,00%	0,01%	0,01%	0,00%	0,00%	0,00%	0,00%	0,00%	
Away									
	5,19%	10,84%	5,66%	1,31%	0,17%	0,01%	0,00%	0,00%	
Home win	52,14%		Odd	1,92					
Draw	23,18%		Odd	4,31					
Away win	24,67%		Odd	4,05					
Total	100%								

- **Analysis-** Our Poisson regression model grants more than 50% chances to a home win. Nevertheless, the model says that betting on a Manchester City win would only have positive expected value if the bookmakers would offer an odd over 1.92, which is considerably higher than the best odd offered by the Portuguese bookmakers (1.57). It is important to understand that Chelsea's chances might be overestimated, as they performed considerably better in 2016/2017 than in this season, while City's chances, on the contrary, are probably underestimated (as its performance improved a lot in the current season).

Betting decision:

- The 0-3 away win against Arsenal just added more substance to our view that, these days, Manchester City is the best team in the Premier League (and quite possibly, in the world) by far. The difference to the other teams is enormous, as The Citizens have the best (and more expensive) squad (in terms of quality and quantity), the best manager, the most balanced and refined style of play (best

Forecasting Football Outcomes to Invest in Betting Markets

attack and defence, both at home and away). Manchester City is also an extremely confident team, that never renounces on its playing principles, regardless of the opponent's name or status. With the exception of the away match against Liverpool, every single team that played against Manchester City just assumed its superiority and gave away the control of the match, trying not to suffer a heavy loss. From our standpoint, every City match on which the bookmakers are offering odds of around 1.5 for its win, delivers a great deal in terms of risk-reward trade-off.

Annex IV- Manchester United vs Liverpool (match day 30)

- 10/03/2018 at 12:30 pm, Old Trafford (Manchester)
- Weather: 9°C; precipitation probability of 1%; humidity percentage of 71% and wind speed of 24 km/h
- **Analysis:** Weather as perfect as it gets to play football in the UK, at this time of the year. It allows for a fast circulation of the ball, and consequently for an open match.

Potential bets:

- Manchester United or Draw- 1.40
- Manchester United or Draw and over 1.5 goals- 1.80

Probable line-ups:

- **Manchester United-** 4-2-3-1: De Gea; Valencia, Bailly, Smalling and Young; Matic and McTominay; Mata, Lingard and Sánchez; Lukaku
- **Liverpool-** 4-3-3: Karius; Arnold, Matip, Van Dijk and Robertson; Chamberlain, Henderson and Can; Salah, Firmino and Mané

Injuries and suspensions:

- **Manchester United-** Out: Herrera, Blind, Jones, Rojo, Martial and Pogba (all injured)
- **Liverpool-** Doubtful: Robertson and Henderson (all injured)
- **Analysis-** Manchester United has 6 important players injured (especially Martial and Pogba), while Liverpool has 2 potential starters in doubt.

Discipline table:

- **Manchester United** (average per match at home)- Yellow cards: 1.6, red cards: 0.0, booking points: 15.7
- **Liverpool** (average per match away)- Yellow cards: 1.4, red cards: 0.1, booking points: 16.1

Forecasting Football Outcomes to Invest in Betting Markets

- **Premier League** (team average per match overall)- Yellow cards: 1.6, red cards: 0.1, booking points: 16.6
- **Analysis-** Both teams are slightly below the Premier League average in terms of aggressiveness (at home, Manchester United is equal to the average in terms of yellow cards, but has not seen any red cards so far). It seems unlikely that cards or injuries might condition the outcome of this match.

Head-to-head over the last 3 years:

- **10 matches-** 5 wins for Manchester United (50%), 4 draws (40%) and 1 win for Liverpool (10%); 7 times over 1.5 (70%) and 3 time under (30%); Liverpool never won away, while United did it 3 times. The last 4 matches between these 2 teams ended up drawn
- **Analysis-** Given these numbers, it is fair to conclude that there is a certain level of bias favouring Manchester United, in a match that represents arguably the biggest historical rivalry in the English football.

Premier league tables overall, home and away:

- **Overall-** Manchester United: 2nd place, 62 points, 56 goals scored and 22 conceded, Liverpool: 3rd place, 60 points, 67 goals scored and 32 conceded
- **Home-** Manchester United: 2nd place, 35 points, 31 goals scored and 6 conceded
- **Away-** Liverpool: 2nd place, 27 points, 34 goals scored and 22 conceded
- **Analysis-** Both teams have been performing very consistently at home and away. United defends better than Liverpool, and Liverpool attacks better than United, but the overall goal difference is virtually the same (+34 goals for Manchester and +35 goals for Liverpool).

Home/away performance tables:

- **Manchester United-** won 57% of its points at home and 43% when playing away, scored 55% of its goals at home and 45% when playing away, conceding 27% of its goals at home and 73% away
- **Liverpool-** won 55% of its points at home and 45% when playing away, scoring

Forecasting Football Outcomes to Invest in Betting Markets

also 49% of its goals at home and 51% when playing away, conceding 31% of its goals at home and 69% away

- **Premier League-** On average, 57% of the goals were scored at home and 43% were scored away
- **Analysis-** This data tells us that both teams won around 55% of their points at home, and 45% when playing away (10% more when playing at home). Liverpool's ability to score goals is nearly the same at home and away, while Manchester United is near the Premier League average, but both teams suffer considerably more goals when playing away (around 30%-70%).

Points per game tables:

- Manchester United has a relative form of +0.11, while Liverpool scores +0.31. Both teams are experiencing good levels of recent form, but Liverpool even more than Manchester United

Scoring tables:

- **Manchester United** scores, on average, 2.20 goals per match at home, while Liverpool concedes 1.60 goals per match when playing away. If we average these numbers, we obtain Manchester United's "scoring strength", which corresponds to 1.90 goals. Besides that, United's matches at home ended with over 1.5 goals 73% of the times
- **Liverpool** scores, on average, 2.33 goals per match when playing away, while Manchester United concedes 0.47 goals per match at home. If we average these numbers, we obtain Liverpool's "scoring strength", which corresponds to 1.40 goals. Besides that, Liverpool's away matches ended with over 1.5 goals 93% of the times
- **Analysis-** If we sum United and Liverpool scoring strengths, we obtain a total of 3.30 goals expected in the forthcoming match. Besides that, if we average both teams' percentage of matches that ended with over 1.5 goals, we get an 83% probability that this match in particular ends with over 1.5 goals, which can be converted into an odd of 1.21.

Team characteristics:

- Both teams like to use short passes to control the game in the opposition's half.
- Liverpool uses an offside trap defensive strategy, which means that the defensive line must be positioned high in the field. As Manchester United has very fast and skilful players (e.g. Sánchez, Lingard, Lukaku or Rashford), and is good in counter-attacking, there might be some space to exploit behind Liverpool's defensive line
- Both teams often attempt through balls
- Liverpool is weak at protecting the lead, while United is strong at coming back from losing positions
- Manchester United is weak at stopping opponents from creating chances, while Liverpool is very strong at creating scoring chances
- Both teams have many similarities, as United is good and Liverpool even better at shooting from direct free kicks, finishing scoring chances, creating long shot opportunities, creating chances through individual skill and counter attacking

Team psychology:

- **Team spirit and motivation-** Both teams' spirit and motivation have to be very high, as they are only separated by 2 points, and this is historically the most electrifying rivalry in the Premier League
- **Confidence-** Although Manchester United is not being able to compete with Manchester City for the Premier League title, the team has been performing consistently this season, being in 2nd place (the only loss at home was against Manchester City), and still in the FA Cup quarter-finals, in which United will play against Brighton, at Old Trafford. In the Champions League, United passed the group stage for the first time since Alex Ferguson's departure (2012/2013), facing Sevilla at home for a place in the quarter-finals (0-0 away). Liverpool has been performing nearly as good as Manchester United, placing 3rd in the league, 2 points behind The Red Devils, and already with a guaranteed spot among the best 8 teams of the 2017/2018 Champions League, after defeating Porto 0-5 at Estádio

do Dragão. The Reds were eliminated in the 4th round of the FA Cup by West Brom (2-3 at home), but on the other hand, were the only team in the Premier League that was able to defeat Manchester City (4-3 at home). Given this context, we believe that it is fair to classify both teams' confidence levels as high

- **Managerial impact-** Klopp arrived at Liverpool in October of 2015, in replacement of Brendan Rodgers, and in his first season (incomplete), although The Reds finished 8th in the Premier League, he was able to reach the Europa League final, which was lost to Sevilla (1-3). About 2 and a half years later, Klopp's "rock 'n roll" style of football is fully implemented at Merseyside: Liverpool defends high, pressing a lot, and attacks at the speed of light, as Borussia did when Klopp was in charge. To achieve that, some very high investments in new players had to be made (e.g. Salah, Mané, Wijnaldum or Van Dijk), but these days, there is no doubt in considering Klopp's managerial impact on Liverpool as high, not being very high just because he did not win any silverware, so far. José Mourinho came to Manchester United with the club immersed in a deep crisis, after 2 years of heavy investments without returns (poor results), with Van Gaal at the helm. Mourinho brought a more pragmatic approach to the club, focusing on a solid and compact defensive process, then trying to exploit the opponent's mistakes over the course of the match. In 2 years, the Portuguese spent around 340 million euros in new players, such as Pogba (105 million euros, at the time the world transfer fee record), Alexis Sánchez, Lukaku, Ibrahimovic or Matic, and won the FA Community Shield, the League Cup and the Europa League in the first season. This season has been a disappointing one, as the team was expected to fight for the Premier League title, but is just in 2nd place, 16 points below Manchester City. The Red Devils have always been known (at least in Ferguson's era) for playing fast, attacking football (similar to Klopp's style), and José Mourinho's cautious and cynical style of play does not convince neither the fans or the club's historical players. Because Mourinho has been able to implement his ideas (either people like them or not) and to win some titles, we consider that he is having a high managerial impact in the club, as well

Team fatigue:

- **Toughness of previous match-** Liverpool won comfortably 2-0 vs Newcastle at

Forecasting Football Outcomes to Invest in Betting Markets

Anfield Road, while Manchester United won 2-3 against Crystal Palace away, in a match where Palace was winning 2-0 at half-time, and United only scored the winning goal after the 90th minute. Having this in consideration, we consider the toughness of previous matches as low for Liverpool, and as the highest for Manchester United

- Resting-** In United's last match, the only player that fully rested was Bailly. McTominay was replaced at half-time by Rashford, and on the 67th minute, Mourinho decided to change Young for Mata, and Valencia for Luke Shaw. Against Newcastle, none of the players that compose the Liverpool's probable line-up for Old Trafford rested entirely, but Matip only entered by the minute 88, in replacement of Firmino. Besides that, Mané was replaced by Lallana (74th minute), and Chamberlain was replaced by Milner, also on the 88th minute. Overall, both teams rested a central defender, and United was able to rest a midfielder for 45 minutes, and both full-backs for about 20 minutes, while Liverpool rested a winger for 15 minutes
- National team participation-** The last round of international matches was in November of 2017, and so this will not be considered

Poisson regression model:

	0	1	2	3	4	5	6	7+	Home
0	7,67%	11,74%	8,99%	4,58%	1,75%	0,54%	0,14%	0,04%	
1	7,95%	12,18%	9,32%	4,76%	1,82%	0,56%	0,14%	0,04%	
2	4,13%	6,31%	4,83%	2,47%	0,94%	0,29%	0,07%	0,02%	
3	1,43%	2,18%	1,67%	0,85%	0,33%	0,10%	0,03%	0,01%	
4	0,37%	0,57%	0,43%	0,22%	0,08%	0,03%	0,01%	0,00%	
5	0,08%	0,12%	0,09%	0,05%	0,02%	0,01%	0,00%	0,00%	
6	0,01%	0,02%	0,02%	0,01%	0,00%	0,00%	0,00%	0,00%	
7+	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	
Away									
	7,67%	12,18%	4,83%	0,85%	0,08%	0,01%	0,00%	0,00%	
Home win	48,72%	Odd	2,05		Home or Draw		Odd		
					74,34%		1,35		
Draw	25,62%	Odd	3,90		Home or Away		Odd		
Away win	25,68%	Odd	3,89		74,40%		1,34		
Total	100%				Draw or Away		Odd		
					51,30%		1,95		
Under 1.5	27,36%	Odd	3,65						
Over 1.5	72,64%	Odd	1,38						

Forecasting Football Outcomes to Invest in Betting Markets

- **Analysis-** Our Poisson regression model grants 74,34% chances to a home win or draw, and 74,40% to a home or away win, values that can be converted into an odd of around 1.35. The model also grants a chance of 72,64% for the match to end with over 1.5 goals, which can be converted into an odd of 1.38. We believe that, in this case, the results given by the model should be pretty accurate, as both teams are performing considerably better this season than in the last one, and have not changed their managers or styles of play.

Betting decision:

- Our analysis suggests that, over the last 3 years, there has been a head-to-head bias favouring Manchester United in the direct confrontations against Liverpool, although in the last match between these teams (0-0 at Anfield), Liverpool completely overpowered United, missing several opportunities to score. In 10 matches, Liverpool never won away, the last 4 ended up drawn, and 7 had over 1.5 goals. In 2017/2018, The Red Devils have been very strong at home, only losing to Manchester City, and defeating top 6 rivals Chelsea and Tottenham. However, injuries have been affecting United severely this season, with key French players Martial and Pogba unavailable for this match. The scoring table says that there is a probability of 83% that the match ends with over 1.5 goals (odd of 1.21), and the Poisson regression determines a probability of 72,64% (odd of 1.38) for the same outcome. The model also attributed a likelihood of 74,34% (odd of 1.35) to United's win or draw. About the match, our expectation is that neither of the teams will risk too much (at least in the beginning), because they are only separated by 2 points, and so a draw can be an interesting outcome for both (especially for United). Liverpool uses a risky style of play, with high pressure, but with the exception of Van Dijk, has not enough individual quality on the defence to do so, and can be exposed to United's counter attacks. Mourinho knows this, and probably will adopt the usual cautious tactics, waiting for the right moment to strike. The Red Devils must be careful though, because they are weak in stopping opponents from creating scoring chances, and Liverpool's front line is absolutely devastating. Having all this in consideration, we believe that an offered odd of 1.80 for a Manchester United win or draw, plus over 1.5 goals in the match delivers positive expected value, therefore worthing a bet.

Annex V- Liverpool vs Watford (match day 31)

- 17/03/2018 at 5:30 pm, Anfield Road (Liverpool)
- Weather: 2°C; precipitation probability of 0%; humidity percentage of 55% and wind speed of 35 km/h
- **Analysis:** Good weather to play football in the UK, at this time of the year. The precipitation probability of 0% allows for a fast circulation of the ball, and consequently for an open match.

Potential bets:

- Liverpool and over 2.5 goals- 1.51
- Liverpool and under 4.5 goals- 1.58

Probable line-ups:

- **Liverpool-** 4-3-3: Karius; Arnold, Matip, Van Dijk and Robertson; Wijnaldum, Henderson and Can; Salah, Firmino and Mané
- **Watford-** 4-2-3-1: Karnezis; Janmaat, Prodl, Mariappa and Britos; Doucouré and Capoue; Femenía, Pereyra and Richarlison; Deeney

Injuries and suspensions:

- **Liverpool-** None
- **Watford-** Out: Hoban, Catchcart, Chalobah, Kabasele, Cleverley and Deulofeu (all injured)
- **Analysis-** Watford has 5 important players injured (especially Kabasele and Cleverley), and also a key attacking player injured, Gerard Deulofeu.

Discipline table:

- **Liverpool** (average per match at home)- Yellow cards: 1.0, red cards: 0.0, booking points: 10.0
- **Watford** (average per match away)- Yellow cards: 1.3, red cards: 0.1, booking points: 15.5

Forecasting Football Outcomes to Invest in Betting Markets

- **Premier League** (team average per match overall)- Yellow cards: 1.5, red cards: 0.1, booking points: 16.5
- **Analysis-** Both teams are below the Premier League average in terms of aggressiveness, therefore, it seems unlikely that cards or injuries might condition the outcome of this match.

Head-to-head over the last 3 years:

- **5 matches-** 3 wins for Liverpool (60%), 1 draw (20%) and 1 win for Watford (20%); 3 times over 2.5 (60%) and 2 times under (40%); 2 times over 4.5 (40%) and 3 times under (60%)
- **Analysis-** Given these numbers, it is fair to conclude that there is a certain level of bias favouring Liverpool, however, the last match for Premier League ended up in a draw, with Watford scoring the equalizing goal at the stoppage time.

Premier league tables overall, home and away:

- **Overall-** Liverpool: 4th place, 60 points, 68 goals scored and 34 conceded, Watford: 10th place, 36 points, 39 goals scored and 50 conceded
- **Home-** Liverpool: 5th place, 33 points, 33 goals scored and 10 conceded
- **Away-** Watford: 8th place, 14 points, 17 goals scored and 24 conceded
- **Analysis-** This table provides support to the idea that Liverpool is a much stronger team than Watford, investing a lot more to be placed among the Champions League places (top 4), while Watford is just a team that fights to achieve its stability in the Premier League. Overall, Liverpool has more 24 points, 29 goals scored and less 16 conceded than The Hornets, which is a very considerable difference.

Home/away performance tables:

- **Liverpool-** won 55% of its points at home and 45% when playing away, scoring also 49% of its goals at home and 51% when playing away, conceding 29% of its goals at home and 71% away
- **Watford-** won 61% of its points at home and 39% when playing away, scoring

Forecasting Football Outcomes to Invest in Betting Markets

also 56% of its goals at home and 44% when playing away, conceding 52% of its goals at home and 48% away

- **Premier League-** On average, 57% of the goals were scored at home and 43% were scored away
- **Analysis-** Liverpool won 10% more points at home than away, and Watford 21%, which means that Watford is 11% more dependent on the home advantage factor than Liverpool. In terms of goals scored, it is fair to say that Watford is consonant with the Premier League average, while Liverpool scores virtually the same number of goals at home and away (49%-51%), maybe due to the immense strength of its offensive line. However, Liverpool concedes a lot more goals when playing away from Anfield (29%-71%), probably because its defensive line is not very reliable, and opponents tend to attack more when playing at their own grounds. Watford, by its turn, concedes approximately the same number of goals at home and away (52%-48%).

Points per game tables:

- The Reds have a relative form of +0.00, while Watford scores +0.18, which means that Liverpool's level of recent form is neutral, and Watford's is positive

Scoring tables:

- **Liverpool** scores, on average, 2.20 goals per match at home, while Watford concedes 1.60 goals per match when playing away. If we average these numbers, we obtain Liverpool's "scoring strength", which corresponds to 1.90 goals. Besides that, Liverpool's matches at home ended with over 2.5 goals 53% of the times, while 80% of those matches ended with less than 4.5 goals
- **Watford** scores, on average, 1.13 goals per match when playing away, while Liverpool concedes 0.67 goals per match at home. If we average these numbers, we obtain Watford's "scoring strength", which corresponds to 0.90 goals. Besides that, Watford's away matches ended with over 2.5 goals 53% of the times, while 87% of those matches ended with less than 4.5 goals
- **Analysis-** If we sum Liverpool and Watford scoring strengths, we obtain a total of 2.80 goals expected in the forthcoming match. Besides that, if we average both

Forecasting Football Outcomes to Invest in Betting Markets

teams' percentage of matches that ended with over 2.5 goals, we get an 53% probability that this match in particular ends with over 2.5 goals, which can be converted into an odd of 1.87. When it comes to the probability that this match in particular ends with less than 4.5 goals, the average obtained is 83,5%, which corresponds to an odd of 1.20.

Team characteristics:

- Liverpool likes to play a possession-based football, using short passes to control the game in the opposition's half
- Both teams are strong at stealing the ball from the opposition, but have different styles of doing it. Watford can be considered an aggressive team, while Liverpool is non-aggressive. Besides that, Watford is strong in aerial duels, while Liverpool is weak
- Liverpool is weak at protecting the lead, while Watford is very weak. However, the latter is strong at coming back from losing positions
- Liverpool is strong in attacking set pieces, while Watford is weak at defending them
- Watford uses an offside trap defensive strategy, which means that the defensive line must be positioned high in the field. As Liverpool has a very fast and skilful offensive triangle (Mané, Firmino and Salah), and is very good in counter-attacking, there might be some space to exploit behind Watford's defensive line, especially because The Hornets are very weak at defending counter attacks

Team psychology:

- **Team spirit and motivation-** Watford suffered upsetting early eliminations from the domestic cups, so the only focus now is to avoid relegation from the Premier League as soon as possible, preferably placing within the first half of the table. The team seems united enough around Gracia, although some important players expressed, at the time, their dissatisfaction about Silva's dismissal (e.g. Richarlison). On the other hand, Liverpool needs to guarantee a Champions League qualification place (among the top 4), ideally as close as possible to Manchester City. Klopp is a charismatic coach, highly passionate about football

and a great motivator, as well. He has the support of the team, and will not allow any kind of relaxation until the end of the season. Both teams are in a process of establishing themselves at a higher level (Liverpool as serious title contenders, Watford as a mid-table team), consequently, they need to consolidate their growth by winning as often as possible. That being said, we consider that the team spirit and motivation of Liverpool and Watford must be high

- **Confidence-** Liverpool has been performing consistently this season, placing 4th in the Premier League, and already with a guaranteed spot among the best 8 teams of the 2017/2018 Champions League, after defeating Porto 0-5 at Estádio do Dragão. The Reds were eliminated in the 4th round of the FA Cup by West Brom (2-3 at home), but on the other hand, were the only team in the Premier League that was able to defeat Manchester City (4-3 at home). Given this context, we believe that it is fair to classify Liverpool's level of confidence as high. The Hornets have been performing much better in 2017/2018 than in the last season, where they ranked 17th, the last place above relegation zone. Currently, Watford ranks 10th, 8 points above relegation, when there are still 7 matches left to play. In the national cups, Watford was surprisingly eliminated from the League Cup by Bristol (2-3 at home), in the 3rd round, and defeated by Southampton (1-0 away) in the 4th round of FA Cup. Considering this, we think that Watford's level of confidence should be classified as medium
- **Managerial impact-** Klopp arrived at Liverpool in October of 2015, in replacement of Brendan Rodgers, and in his first season (incomplete), although The Reds finished 8th in the Premier League, he was able to reach the Europa League final, which was lost to Sevilla (1-3). About 2 and a half years later, Klopp's "rock' n roll" style of football is fully implemented at Merseyside: Liverpool defends high, pressing a lot, and attacks at the speed of light, as Borussia did when Klopp was in charge. To achieve that, some very high investments in new players had to be made (e.g. Salah, Mané, Wijnaldum or Van Dijk), but these days, there is no doubt in considering Klopp's managerial impact on Liverpool as high, not being very high just because he did not win any silverware, so far. Marco Silva was fired on the 20th January, after suffering a categorical defeat against Leicester (2-0 away), being replaced by Javi Gracia,

who started by changing the usual formation from 3-4-2-1 to 4-2-3-1. Besides that, some players that were important under Silva's management, like Hughes, Carrillo or Gray, lost relevance to others, such as Roberto Pereyra or Troy Deeney. The Spanish coach began his spell by drawing away, against Stoke (0-0). After that, he achieved an incredible home win against Chelsea (4-1), but then lost away to direct competitor West Ham (2-0). In the last 3 matches, Watford was able to beat Everton and West Brom (both 1-0) at home, but lost 3-0 away to Arsenal in the last match day. Since Javi Gracia took the helm, Watford has been stable in the 10th place, 4 points below 9th ranked Everton, therefore we think that, so far, his managerial impact at Vicarage Road has been normal

Team fatigue:

- **Toughness of previous match-** Liverpool lost the North-West Derby vs Manchester United (2-1), at Old Trafford. It was an even and intense match, that must have risen the fatigue levels of both teams (United lost to Sevilla the chance of qualifying for the Champions League quarter-finals, 3 days later, at home). Arsenal gave no possibilities to Watford, winning 3-0 at Emirates, although the outcome does not reflect what really happened in the game, as The Hornets disputed the match, had several scoring chances, but were not effective when it came to finish those occasions. Having this in consideration, we classify the toughness of previous matches as high for The Hornets, and as the highest for Liverpool
- **Resting-** In Liverpool's last match, from the probable line-up for today, the only player that fully rested was Matip. Arnold was replaced on the 80th minute by Wijnaldum, and 4 minutes later, Klopp risked even more for the equalizing goal, changing Robertson for Solanke. Against Arsenal, none of the players that compose Watford's probable line-up for Anfield Road rested entirely, but Britos only entered by the minute 80, in replacement of Janmaat. Besides that, Femenía was replaced by Hughes (62nd minute), and Pereyra was replaced by Okaka, 5 minutes later
- **National team participation-** The last round of international matches was in November of 2017, and so this will not be considered

Forecasting Football Outcomes to Invest in Betting Markets

Poisson regression model:

	0	1	2	3	4	5	6	7+	Home
0	1,98%	6,10%	9,41%	9,67%	7,46%	4,60%	2,37%	1,64%	
1	1,66%	5,11%	7,89%	8,11%	6,26%	3,86%	1,99%	1,38%	
2	0,70%	2,14%	3,31%	3,40%	2,62%	1,62%	0,83%	0,58%	
3	0,19%	0,60%	0,93%	0,95%	0,73%	0,45%	0,23%	0,16%	
4	0,04%	0,13%	0,19%	0,20%	0,15%	0,09%	0,05%	0,03%	
5	0,01%	0,02%	0,03%	0,03%	0,03%	0,02%	0,01%	0,01%	
6	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	
7+	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	
Away									
	1,98%	5,11%	3,31%	0,95%	0,15%	0,02%	0,00%	0,00%	
Home win	81,55%		Odd	1,23					
Draw	11,52%		Odd	8,68					
Away win	6,90%		Odd	14,49					
Total	100%								
Under 2.5	24,96%		Odd	4,01					
Over 2.5	75,04%		Odd	1,33					
Under 3.5	45%		Odd	2,23					
Over 3.5	55%		Odd	1,82					
Under 4.5	64,37%		Odd	1,55					
Over 4.5	35,63%		Odd	2,81					

- Analysis-** Our Poisson regression model grants 81,55% chances to a home win, which can be converted into an odd of 1.23. The model also grants a chance of 75,04% to the match ending with over 2.5 goals, which corresponds to an odd of 1.33. The probability estimated for the match having less than 4.5 goals is 64,37%, which is equivalent to an odd of 1.55. We believe that, in this case, the results given by the model should be relatively accurate, as both teams are performing considerably better this season than in the last one.

Betting decision:

- Liverpool will approach this match in its full strength, while Watford has 6 players currently injured. Even if Watford had no injuries, Liverpool would still have a squad with much more individual quality. Furthermore, The Reds play at home, and so they are expected to dominate and win this match, as The Hornets have defensive flaws that can be exploited. In fact, both teams are strong in recovering the ball, but Watford has problems in controlling the space behind its defensive

Forecasting Football Outcomes to Invest in Betting Markets

line, and Liverpool is very strong in the offensive transitions. Besides that, our analysis indicates that there is a bias in the head-to-head confrontations between these teams, in Liverpool's favour. The Poisson regression model attributes 81,55% chances to a home win, which is equivalent to an odd of 1.23. The model also grants probabilities of 75,04% to this match having more than 2.5 goals (odd of 1.33), and of 64,37% to this match ending with less than 4.5 goals (odd of 1.55). Additionally, the sum of both teams' "scoring strengths" is equal to 2.80 goals. Although, at this stage, it seems more logical to bet on Liverpool's win and over 2.5 goals, 80% of Liverpool's matches at home, and 87% of Watford's matches away ended with less than 4.5 goals. Considering this, we believe that betting on Liverpool and under 4.5 goals in the match, has a positive expected value when the offered odd is of 1.58.

Annex VI- Bournemouth vs West Bromwich Albion (match day 31)

- 17/03/2018 at 3:00 pm, Dean Court Stadium (Bournemouth)
- Weather: 2°C; precipitation probability of 0%; humidity percentage of 46% and wind speed of 37 km/h
- **Analysis:** Weather as perfect as it gets to play football in the UK, at this time of the year. It allows for a fast circulation of the ball, and consequently for an open match.

Potential bets:

- Bournemouth- 1.82

Probable line-ups:

- **Bournemouth-** 4-4-1-1: Begovic; Francis, S.Cook, Aké and Daniels; Smith, L.Cook, Gosling and Stanislas; King; Wilson
- **West Bromwich-** 4-4-2: Foster; Dawson, Hegazi, Evans and Gibbs; Phillips, Krychowiak, Livermore and Brunt; Rodriguez and Rondón

Injuries and suspensions:

- **Bournemouth-** Out: Mings; doubtful: Pugh (all injured)
- **West Bromwich-** Out: Chadli, Morrison, Sturridge, and Barry; doubtful: Evans (all injured)
- **Analysis-** Bournemouth has a mid-importance player in doubt (Pugh), while West Bromwich will have to deal with the absences of Nacer Chadli, Gareth Barry and probably also Johnny Evans (very important players). Besides that, the skilful but “chronically” injured Daniel Sturridge will not be fit to play this match, as well.

Discipline table:

- **Bournemouth** (average per match at home)- Yellow cards: 1.6, red cards: 0.1, booking points: 17.7
- **West Bromwich** (average per match away)- Yellow cards: 1.9, red cards: 0.1,

Forecasting Football Outcomes to Invest in Betting Markets

booking points: 20.3

- **Premier League** (average per match)- Yellow cards: 1.6, red cards: 0.1, booking points: 16.5
- **Analysis-** Both teams are above the Premier League's average in terms of aggressiveness, especially West Bromwich. Therefore, it seems likely that at least a high number of yellow cards can be shown during the 90 minutes. Furthermore, red cards or injuries might occur, conditioning the outcome of the match. As WBA has a very physical and combative style of play (the squad was made at the image of Tony Pullis, the coach that began the season), we believe that its excess of impetuosity can work as an advantage for Bournemouth.

Head-to-head over the last 3 years:

- **6 matches-** 3 wins for Bournemouth (50%), 2 wins for West Bromwich (33,3%) and 1 draw (16,7%)
- **Analysis-** Over the last 3 years, Bournemouth won 50% of the matches between these teams, however, the last 2 were won by WBA, who was never able to win away. Considering this, we think that the equation is balanced and consequently, there is not a significant bias favouring any of the teams.

Premier league tables overall, home and away:

- **Overall-** Bournemouth: 12nd place, 33 points, 35 goals scored and 48 conceded, West Bromwich: 20th place, 20 points, 23 goals scored and 47 conceded
- **Home-** Bournemouth: 15th place, 19 points, 21 goals scored and 25 conceded
- **Away-** West Bromwich: 20th place, 7 points, 7 goals scored and 23 conceded
- **Analysis-** This table allows to understand that, in one hand, Bournemouth is performing as a mid-table team over the course of the 2017/2018 season, currently placing 6 points above relegation zone, and 4 points below the first-half of the table. On the other hand, West Bromwich's season has been disappointing to say the least, as The Baggies have a reasonably good squad, and were expected to finish around mid-table. The reality, however, is that West Brom has been the worst team in the Premier League so far, losing the last 6 matches in a row. The

Forecasting Football Outcomes to Invest in Betting Markets

West Bromwich Albion's current situation is very concerning, as the club ranks in the bottom place of the division, already 8 points below the non-relegation zone, when there are only 8 matches left to play.

Home/away performance tables:

- **Bournemouth-** won 58% of its points at home and 42% when playing away, scored 60% of its goals at home and 40% when playing away, conceding 52% of its goals at home and 48% away
- **West Bromwich-** won 65% of its points at home and 35% when playing away, scoring also 70% of its goals at home and 30% when playing away, conceding 51% of its goals at home and 49% away
- **Premier League-** On average, 57% of the goals were scored at home and 43% were scored away
- **Analysis-** By analysing the numerical data provided in this table, two main conclusions can be made: the first one is that both teams depend on the home advantage to win points and score goals, which is normal in low-to-mid table teams, but WBA's dependence on this factor is huge. Even so, West Bromwich results at home have been very poor, which anticipates serious difficulties in the forthcoming away matches. The second conclusion is that both Bournemouth and West Brom are conceding virtually the same number of goals at home and away, which can be interpreted as a sign of weakness in their defensive processes, as each one of these teams are suffering goals in a consistent way, regardless of the situation.

Points per game tables:

- Bournemouth has a relative form of +0.40, while WBA scores -0.17. In fact, Bournemouth did not begin well the season, but is experiencing good recent form, only losing twice in the last 11 matches. West Bromwich, by its turn, is in a dreadful moment, with 6 losses and a draw in the last 7 matches. Having this in consideration, the fact that the score for its relative form is only of -0.17, is a striking evidence of how terrible The Baggies season has been, until this moment

Scoring tables:

Forecasting Football Outcomes to Invest in Betting Markets

- **Bournemouth** scores, on average, 1.40 goals per match at home, while West Brom concedes 1.53 goals per match when playing away. If we average these numbers, we obtain Bournemouth's "scoring strength", which corresponds to 1.47 goals
- **West Bromwich** scores, on average, 0.47 goals per match when playing away, while Bournemouth concedes 1.67 goals per match at home. If we average these numbers, we obtain WBA's "scoring strength", which corresponds to 1.07 goals

Team characteristics:

- Bournemouth is strong at coming back from losing positions, while West Bromwich is very weak at protecting the lead
- Bournemouth take long shots often, while West Bromwich is weak at defending them
- Bournemouth uses a non-aggressive approach, as it is weak in aerial duels, while West Bromwich Albion uses an aggressive approach, and it is strong in aerial duels, hence we expect West Brom to win most of the disputed balls
- Bournemouth likes to attack down the left, while WBA likes to attack down the right, so we can expect most of the action to be developed on that wing. However, West Brom is very weak when it comes to defend attacks down the flanks

Team psychology:

- **Team spirit and motivation-** Bournemouth's motivation to win this match should be considered very high, as this is a great opportunity to win 3 points at home, against a weakened opponent. Those points would be a valuable contribution to the achievement of the club's objectives for this season: the main one is to avoid relegation, and after that is guaranteed, it would be nice to assure a place in the first half of the table. The team is going through a good moment, and at least to our knowledge, there are no public cases of indiscipline or instability at Dean Court. On the other hand, The Baggies desperately need points in order to escape from relegation. West Brom is under overwhelming pressure, and that is surely affecting the players' minds, as some of the recent lost matches were not as negative as the final scores suggest. This season, everything seems to

go wrong with West Bromwich and as we know, the team spirit cannot be good when results do not show up for such a long period of time. Regarding this, the WBA's team spirit and motivation has to be classified as very low

- **Confidence-** The 2017/2018 season began terribly for Bournemouth, with 4 defeats in 4 matches, however the team's performance improved a lot after the defeat vs Manchester City (4-0 away), including incredible wins against Arsenal (2-1 at home) and Chelsea (0-3 away). In the last 10 matches for the Premier League, Bournemouth only lost twice. When it comes to its performance in the domestic cups, Bournemouth was surprisingly defeated by Championship teams: in the 3rd round of the League Cup, losing 2-3 at home vs Preston, and also in the 3rd round of the FA Cup (3-0 away vs Millwall). WBA's performance in the Premier League has been a complete disaster, as they were only able to win 3 matches, so far: 1-0 to Bournemouth in the 1st match day (at home), 0-1 to Burnley in the 2nd match day (away), and the only win under Pardew's management was against Brighton, at home (2-0). In the FA Cup, The Baggies eliminated Liverpool in the 4th round (2-3 at Anfield Road), but loss in the next round against Southampton (1-2 at home). West Bromwich adventure in the League Cup ended after losing naturally vs Manchester City (1-2 at home). Given these results, we consider Bournemouth's level of confidence for this match as high, and West Brom's as very low
- **Managerial impact-** Eddie Howe is a manager that is fully identified with Bournemouth's culture. Howe began and finished his career as a player at Dean Court, and after that, the Englishman also began his career as a manager in the club. After a brief spell at Burnley, he returned to Bournemouth, achieving promotions from Sky Bet League 1 to Championship (2012/2013), and from there to the Premier League, in 2014/2015. In the English main division, Howe lead the team to a 16th place in 2015/2016, when few expected Bournemouth to avoid relegation, and then to a surprising 9th place in 2016/2017. This season, Bournemouth is in 10th place, playing an entertaining fast pace football that mixes British and Dutch football philosophies. Today, Eddie Howe is widely considered the most promising English manager (has been linked with a move to Arsenal, in replacement of long-time manager Arsène Wenger), and of course, his managerial

Forecasting Football Outcomes to Invest in Betting Markets

impact on Bournemouth must be considered very high. West Bromwich began the season with Tony Pullis at the helm, which is a manager that is known for promoting a very physical and direct style of play, relying more on muscle than skill. Pullis built the team in his image, hiring several players that match the characteristics mentioned above (for example, Krychowiak or Jay Rodriguez). Although Tony Pullis is criticized for his outdated style of play, the truth is that he has been able to achieve relatively good results, specially during his tenure in Stoke City. However, this season the results were not so good, and Pullis was sacked after being defeated by Chelsea (0-4 at home), leaving the club just one point above the relegation zone. In his replacement, the board nominated Alan Pardew, a man who had successful experiences as a manager in the Premier League, being able to save Newcastle and Crystal Palace from relegation. However, Pardew is not being able to implement his ideas for WBA's football, probably because these are quite different from Pullis vision, and The Baggies went down to the last place in the League, now already 8 points below the "salvation zone". Considering this, we believe that Pardew's managerial impact should be evaluated as low, not being very low just because things were already not so bright when he arrived at The Hawthorns Stadium

Team fatigue:

- **Toughness of previous match-** In the last match day, both teams lost 1-4 at home: Bournemouth against top 4 candidates Tottenham, and West Bromwich against Leicester City. Given these outcomes, we have to consider the toughness of previous matches as the highest
- **Resting-** In Bournemouth vs Tottenham, from the probable line-up for today, the only player that began as a substitute was King, although he replaced Mousset at the 67th minute. On the same minute, Eddie Howe also replaced Daniels for Ibe. Finally, Stanislas was replaced for Defore (minute 76). Against Leicester, Nyom replaced the injured Evans, and Burke was preferred over Rodriguez to begin the match. Alan Pardew unsuccessfully tried to change the course of the events by replacing Krychowiak for Field (minute 59), Phillips for Rodriguez (minute 70) and Burke for Robson-Kanu (minute 83). After observing the starting elevens and substitutions made by each one of the teams in the last match day, it seems

Forecasting Football Outcomes to Invest in Betting Markets

reasonable to conclude that they have more or less equal resting levels

- **National team participation-** The last round of international matches was in November of 2017, and so this will not be considered

Poisson regression model:

	0	1	2	3	4	5	6	7+	Home
0	7,52%	11,88%	9,39%	4,94%	1,95%	0,62%	0,16%	0,05%	
1	7,58%	11,97%	9,46%	4,98%	1,97%	0,62%	0,16%	0,05%	
2	3,82%	6,03%	4,76%	2,51%	0,99%	0,31%	0,08%	0,02%	
3	1,28%	2,03%	1,60%	0,84%	0,33%	0,11%	0,03%	0,01%	
4	0,32%	0,51%	0,40%	0,21%	0,08%	0,03%	0,01%	0,00%	
5	0,07%	0,10%	0,08%	0,04%	0,02%	0,01%	0,00%	0,00%	
6	0,01%	0,02%	0,01%	0,01%	0,00%	0,00%	0,00%	0,00%	
7+	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	
Away									
	7,52%	11,97%	4,76%	0,84%	0,08%	0,01%	0,00%	0,00%	
Home win	50,66%		Odd	1,97					
Draw	25,18%		Odd	3,97					
Away win	24,14%		Odd	4,14					
Total	100%								

- **Analysis-** Our Poisson regression model grants more than 50% chances to a home win. Nevertheless, the model says that betting on a Bournemouth's win would only have positive expected value if the bookmakers would be willing to offer an odd over 1.97, which is higher than the best odd offered by the Portuguese bookmakers (1.82). However, one must consider the possibility that the model is overestimating WBA's chances, as they performed considerably better in 2016/2017 than in the current season.

Betting decision:

- The information contained in this report suggests that West Bromwich Albion is going through the worst moment of the season, and that avoiding relegation would be considered almost a miracle, as the 6 losses in the last 6 matches attest. As things were not bad enough already, The Baggies have some important absences confirmed, such as the cases of Sturridge, Gareth Barry or Chadli, (plus the probable absence of Johnny Evans) while Bournemouth has not significant absences for this match. Besides that, West Brom is a very aggressive team, and it is likely that a high number of cards will be shown, conditioning the performance of the team. Over the last 3 years, WBA was never able to beat

Forecasting Football Outcomes to Invest in Betting Markets

Bournemouth away from home, although The Baggies have won the last 2 matches between these teams. Bournemouth plays at Dean Court and is in a good moment of form, so we believe that betting on a home win at 1.82 has positive expected value.

Annex VII- Everton vs Manchester City (match day 32)

- 31/03/2018 at 5:30 pm, Goodison Park (Liverpool)
- Weather: 8°C; precipitation probability of 3%; humidity percentage of 59% and wind speed of 14 km/h
- **Analysis:** Weather as perfect as it gets to play football in the UK, at this time of the year. It allows for a fast circulation of the ball, and consequently for an open match.

Potential bets:

- Manchester City and over 1.5 goals- 1.45
- Manchester City and over 2.5 goals- 1.78
- Manchester City and below 4.5 goals- 1.60

Probable line-ups:

- **Everton-** 4-2-3-1: Pickford; Coleman, Keane, Jagielka and Baines; Gueye and Rooney; Walcott, Davies and Bolasie; Tosun
- **Manchester City-** 4-3-3: Ederson; Walker, Kompany, Otamendi and Delph; De Bruyne, Fernandinho and D.Silva; Sterling, Jesus and Sané

Injuries and suspensions:

- **Everton-** Out: McCarthy, Stekelenburg, Sigurdsson, Mangala (all injured) and Williams (suspended); doubtful: Holgate (injured)
- **Manchester City-** Doubtful: Delph and Aguero (all injured)
- **Analysis-** Besides the chronically injured McCarthy, Everton has 2 important players out for this match (Sigurdsson and Williams), and one in doubt (Holgate). Manchester City, by its turn, has 2 players in doubt, being one of them very important to the team's attacking dynamics (Sergio Aguero).

Discipline table:

- **Everton** (average per match at home)- Yellow cards: 1.6, red cards: 0.1, booking

Forecasting Football Outcomes to Invest in Betting Markets

points: 17.7

- **Manchester City** (average per match away)- Yellow cards: 1.7, red cards: 0.1, booking points: 19.0
- **Premier League** (team average per match overall)- Yellow cards: 1.6, red cards: 0.1, booking points: 16.5
- **Analysis-** Both teams are above the Premier League average in terms of aggressiveness (especially Manchester City), therefore, it seems possible that cards or injuries might condition the outcome of this match.

Head-to-head over the last 3 years:

- **5 matches-** 2 wins for Everton (22%), 4 draws (45%) and 3 wins for Manchester City (33%); 7 times over 1.5 (78%) and 2 times under (22%); 3 times over 2.5 (33%) and 6 times under (67%); 9 times under 4.5 (100%)
- **Analysis-** The head-to-head record seems pretty even, nevertheless one must not forget that in the summer of 2017, both teams did the biggest investments of their history. As a consequence of that, Manchester City is performing like never before, and Everton is having an average season. That being said, we think that the head-to-head data over the last 3 years does not truthfully reflect the difference of strength between these teams (Manchester City is incomparably stronger).

Premier league tables overall, home and away:

- **Overall-** Everton: 9th place, 40 points, 37 goals scored and 50 conceded, Manchester City: 1st place, 81 points, 85 goals scored and 20 conceded
- **Home-** Everton: 7th place, 29 points, 25 goals scored and 18 conceded
- **Away-** Manchester City: 1st place, 38 points, 34 goals scored and 10 conceded
- **Analysis-** These numbers support the fact that Manchester City is the best team in the Premier League, and of course, a much stronger team than Everton. Overall, The Citizens have more 41 points, 50 goals scored and less 30 conceded than Everton, which is a huge difference.

Forecasting Football Outcomes to Invest in Betting Markets

Home/away performance tables:

- **Everton-** won 72% of its points at home and 28% when playing away, scoring also 68% of its goals at home and 32% when playing away, conceding 36% of its goals at home and 64% away
- **Manchester City-** won 53% of its points at home and 47% when playing away, scoring also 60% of its goals at home and 40% when playing away, conceding 50% of its goals at home and 50% away
- **Premier League-** On average, 57% of the goals were scored at home and 43% were scored away
- **Analysis-** Manchester City gained 6% more points at home than away, in contrast with Everton, that won 44% more points at home than away from Goodison Park. This data allows to understand that Everton depends hugely on the home advantage factor to win, while Manchester City does not. Furthermore, Everton depends on this factor not only to win, but also to score and prevent goals. On the contrary, Manchester City conceded exactly the same number of goals at home and away, which in this case, can be interpreted as a sign of a well-implemented defensive process, as The Citizens have the best defence of all the Premier League's teams, and play exactly the same way independently of being at home or away. They score 20% more when playing at Etihad Stadium, which can be explained by the "fear" that Guardiola's pupils instil to the vast majority of their opponents (including rivals such as Manchester United), apart from the home advantage effect on its own.

Points per game tables:

- The Toffees have a relative form of +0.23, while City scores -0.32. At first glance, it may seem weird that Everton has a better level of relative form than Manchester City. This is because City has been very consistent over the course of the season, only losing against Liverpool, and drawing against Everton, Crystal Palace and Burnley. The defeat and one of these draws happened over the last 8 matches. Everton began the season performing very poorly, under Ronald Koeman's guidance. After the Dutchman was dismissed, Sam Allardyce took charge and the results improved significantly, as Everton went from relegation zone to mid-table.

Forecasting Football Outcomes to Invest in Betting Markets

From the last 8 matches, Everton won 4, lost 3 and tied 1

Scoring tables:

- **Everton** scores, on average, 1.67 goals per match at home, while Manchester City concedes 0.67 goals per match when playing away. If we average these numbers, we obtain Everton's "scoring strength", which corresponds to 1.17 goals. Besides that, Everton's matches at home ended with more than 1.5 goals 80% of the times, with over 2.5 goals 53% of the times, and with less than 4.5 goals 87% of the times
- **Manchester City** scores, on average, 2.27 goals per match when playing away, while Everton concedes 1.20 goals per match at home. If we average these numbers, we obtain City's "scoring strength", which corresponds to 1.74 goals. Besides that, The Citizens' away matches ended with more than 1.5 goals 80% of the times, with over 2.5 goals 53% of the times, and with less than 4.5 goals 80% of the times
- **Analysis-** If we sum Everton and Manchester City scoring strengths, we obtain a total of 2.91 goals expected in the forthcoming match. Besides that, Everton and City matches finished with more than 1.5 goals 80% of the times, which is equivalent to an odd of 1.25. If we average both teams' percentage of matches that ended with over 2.5 goals, we get an 53% probability that this match in particular ends with over 2.5 goals, which can be converted into an odd of 1.87. When it comes to the probability of this match in particular ends with less than 4.5 goals, the average obtained is 83,5%, which corresponds to an odd of 1.20.

Team characteristics:

- Manchester City is very strong at attacking down the wings, while Everton is very weak at defending against these type of attacks
- Everton is very weak at defending against skilful players, while The Citizens are strong at creating chances through individual skill, as they have very technical players, such as De Bruyne, David Silva or Leroy Sané
- Manchester City is strong at attacking set pieces, while Everton is weak at defending them

Forecasting Football Outcomes to Invest in Betting Markets

- Both teams are weak in aerial duels, and also at avoiding individual errors
- Manchester City style is all about keeping possession of the ball, using short passes to control the game in the opposition's half. The Toffees like to play in their own half, are weak at keeping possession of the ball, but strong at stealing the ball from the opposition
- City likes to attack down the left, while Everton likes to attack down the right, so we can expect most of the action to be developed on that wing
- City defends high on the pitch, using the offside trap strategy. As Everton likes to attack using long balls, and has some fast players in its offensive line (e.g. Walcott, Bolasie or Niasse), there might be some space to exploit behind The Citizens defensive line. However, Manchester City has this process very well automatized, being the team that suffered less goals in the League

Team psychology:

- **Team spirit and motivation-** This season, Everton invested massive sums in new players, targeting to a Champions League qualification spot (top 4), and also to a good performance in Europa League. Those goals are now out of reach, but The Toffees are still required to achieve the best possible place in the Premier League (realistically, a 7th place would be nice). In one hand, the club has to deal with the disappointment of missing the season's primary ambitions, and the demotivation of not competing for a real objective, but on the other hand, Everton's supporters are very demanding, and the team has to secure a place above mid-table, to continue its process of solidification as a contender for European qualification. Given this situation, we consider Everton's team spirit and motivation as medium. Pep Guardiola promotes a squad rotation policy, so that every player feels useful and integrated. There are no cases (at least publicly known) of indiscipline or internal issues between the players and the manager. The way the team is performing (and winning) provides strong evidence of a positive team culture. Besides that, Manchester City should be highly motivated to win the Premier League as soon as possible, in order to then focus its full attention on the latter stages of the Champions League
- **Confidence-** In 2017/2018, Manchester City has been perhaps the best team in

Europe, only losing to Shakhtar Donetsk (Champions League), Wigan (FA Cup) and Liverpool (Premier League), and tying against Everton, Crystal Palace and Burnley (all for Premier League). On these circumstances, we believe it is fair to say that The Citizens' confidence levels must be very high. Everton has its position in the Premier League perfectly stabilized: 9th place, relatively far from both relegation and European qualification zones. Also, the team was eliminated from Europa League (group stage), FA Cup (2-1 defeat vs Chelsea away) and League Cup (also a 2-1 defeat, in The Merseyside derby vs Liverpool, at Anfield Road). Although Everton has been performing below the initial expectations for this season, The Toffees have improved a lot since Koeman's departure (at least in terms of results), and so their confidence levels should be classified as medium

- **Managerial impact-** Everton sacked Ronald Koeman after the defeat at home against Arsenal (2-5), on the 22nd October, that left the Liverpool based side in relegation zone (18th place). Sam Allardyce succeeded Koeman as Everton's manager, and was able to put Everton in 9th place, with 7 wins, 4 draws and 6 losses. Because of this clear improvement in Everton's results, we consider that he had a high impact on the team's performance. Guardiola arrived at Manchester and after a mediocre first season, has been able to turn Manchester City into one of the most complete teams in Europe. The financial investment in the squad was huge, but nowadays City plays in the image of its coach, and is definitely one of the best sides that he built in the entire career. That being said, we have to consider Guardiola's impact on the club as very high

Team fatigue:

- **Toughness of previous match-** The last matches of both teams were against Stoke (away), because Manchester City did not play the match day 31, in which The Citizens were supposed to meet Brighton, at the Etihad Stadium. Because of that, we will not classify the toughness of Manchester City's last match. Everton, by its turn, had a tough match at The Britannia Stadium, although Stoke had a player sent off by the minute 30 (Charlie Adam). Everton scored first by Tosun (69th minute), but even with 10 players, Stoke was able to equalize the score. The winning goal was scored, once again, by Tosun (84th minute). Considering this, we decided to classify the toughness of Everton's last match as very high

Forecasting Football Outcomes to Invest in Betting Markets

- **Resting-** The last match day of Premier League ended at 17/03/2018 (14 days ago). Besides that, Manchester City's last match was delayed, and so this part of the report will not be considered to forecast what is likely to happen in this match
- **National team participation-** From the Manchester City's probable line-up for this match, the only player that was not called to the most recent double round of international matches was Fabian Delph, who is in doubt due to an injury. If Delph does not recover, he will be replaced by Zinchenko, who was called to play for Ukraine's national team. Therefore, the level of national team participation of the Manchester City's starting 11 can be classified either as many or all of its players. In Everton's case, the level of national team participation of its starting 11 should be classified either as half-team or many, as 6 of its players have been called to play for their respective national teams: Pickford for England, Coleman for Republic of Ireland, Gueye for Senegal, Bolasie for Congo, Tom Davies for England U21 and Tosun for Turkey

Poisson regression model:

	0	1	2	3	4	5	6	7+	Home
0	4,56%	4,89%	2,63%	0,94%	0,25%	0,05%	0,01%	0,00%	
1	9,18%	9,86%	5,29%	1,89%	0,51%	0,11%	0,02%	0,00%	
2	9,25%	9,93%	5,33%	1,91%	0,51%	0,11%	0,02%	0,00%	
3	6,21%	6,67%	3,58%	1,28%	0,34%	0,07%	0,01%	0,00%	
4	3,13%	3,36%	1,80%	0,65%	0,17%	0,04%	0,01%	0,00%	
5	1,26%	1,35%	0,73%	0,26%	0,07%	0,02%	0,00%	0,00%	
6	0,42%	0,45%	0,24%	0,09%	0,02%	0,01%	0,00%	0,00%	
7+	0,16%	0,17%	0,09%	0,03%	0,01%	0,00%	0,00%	0,00%	
Away									
	4,56%	9,86%	5,33%	1,28%	0,17%	0,02%	0,00%	0,00%	
Home win	19,61%	Odd	5,10						
Draw	21,22%	Odd	4,71						
Away win	59,12%	Odd	1,69						
Total	100%								
Under 1.5	18,63%	Odd	5,37						
Over 1.5	81,37%	Odd	1,23						
Under 2.5	40,37%	Odd	2,48						
Over 2.5	59,63%	Odd	1,68						
Under 3.5	62,74%	Odd	1,59						
Over 3.5	37,26%	Odd	2,68						
Under 4.5	80,01%	Odd	1,25						
Over 4.5	19,99%	Odd	5,00						

- **Analysis-** The Poisson regression model grants 59,12% chances to a home win, which can be converted into an odd of 1.69. The probability given by the Poisson regression to the match ending with more than 1.5 goals is 81,37%, which is equivalent to an odd of 1.23. The model also grants a chance of 59,63% to this match ends with over 2.5 goals, which corresponds to an odd of 1.68. The probability estimated for the match having less than 4.5 goals is 80,01%, which is equivalent to an odd of 1.25. It is important to understand that Everton's chances might be overestimated, as they performed considerably better in 2016/2017 than in 2017/2018, while City's chances, on the contrary, are probably underestimated (as its performance improved a lot in the current season).

Betting decision:

- Manchester City is the best team in the Premier League (and quite possibly, in the world) by far. The difference to the other teams is enormous, as The Citizens have the best (and more expensive) squad (in terms of quality and quantity), the best manager, the most balanced and refined style of play (best attack and defence, both at home and away). Manchester City is also an extremely confident team, that never renounces on its playing principles, regardless of the opponent's name or status. With the exception of the away match against Liverpool, every single team that played against Manchester City just assumed its superiority and gave away the control of the match, trying not to suffer a heavy loss. From our standpoint, every City's match on which the bookmakers are offering odds of around 1.5 for its win, delivers a great deal in terms of risk-reward trade-off.

Annex VIII- Arsenal vs Southampton (match day 33)

- 08/04/2018 at 2:15 pm, Emirates Stadium (London)
- Weather: 11°C; precipitation probability of 31%; humidity percentage of 84% and wind speed of 13 km/h
- **Analysis:** Acceptable weather conditions to play football in the UK, at this time of the year. It is likely that rains during the match, but UK-based players are used to that situation. Probably, it will not have a huge impact on the outcome of the event.

Potential bets:

- Arsenal- 1.50

Probable line-ups:

- **Arsenal**- 4-2-3-1: Cech; Bellerín, Mustafi, Chambers and Kolasinac; Elneny and Ramsey; Ozil, Wilshere and Welbeck; Aubameyang
- **Southampton**- 4-2-3-1: McCarthy; Soares, Stephens, Hoedt and Bertrand; Lemina and Romeu; Tadic, Davis and Redmond; Austin

Injuries and suspensions:

- **Arsenal**- Out: Cazorla, Mkhitarian and Ospina (all injured)
- **Southampton**- None
- **Analysis**- Besides the “chronically” injured Santi Cazorla, Arsenal has 1 important player missing (Mkhitarian), while Southampton has all of its players available for this match.

Discipline table:

- **Arsenal** (average per match at home)- Yellow cards: 1.3, red cards: 0.0, booking points: 12.5
- **Southampton** (average per match away)- Yellow cards: 1.9, red cards: 0.0, booking points: 18.7

Forecasting Football Outcomes to Invest in Betting Markets

- **Premier League** (team average per match overall)- Yellow cards: 1.5, red cards: 0.1, booking points: 16.5
- **Analysis-** At home, Arsenal is less aggressive than most Premier League's teams, while Southampton is above the average in terms of aggression, when playing away (although red cards were not shown to any of its players, so far). This means that The Saints may win more duels and recover the ball more swiftly, but on the other hand, they are probably going to be more exposed to disciplinary conditioning through yellow cards.

Head-to-head over the last 3 years:

- **10 matches-** 4 wins for Arsenal (40%), 2 draws (20%) and 4 wins for Southampton (40%)
- **Analysis-** There is not a significant bias favouring any of the teams, as each one was able to win the same number of matches against the other, at home and away. However, it is important to take into account that this season, Southampton is performing considerably worse than usually, as The Saints are fighting to avoid relegation, when over the last 4 seasons they finished always within the first half of the table.

Premier league tables overall, home and away:

- **Overall-** Arsenal: 6th place, 51 points, 58 goals scored and 41 conceded; Southampton: 18th place, 28 points, 29 goals scored and 47 conceded
- **Home-** Arsenal: 3rd place, 38 points, 42 goals scored and 17 conceded
- **Away-** Southampton: 13th place, 12 points, 13 goals scored and 26 conceded
- **Analysis-** Arsenal is one of the teams in the Premier League that depends more on the home advantage factor to score goals and win matches. Southampton, by its turn, has been the second worst team of the Premier League at home. Even so, The Saints conquered more points at home, also having a better goal difference than when playing away from the St Mary's Stadium. Overall, Arsenal conceded only 6 goals more than Southampton, but has a much stronger attack, composed by several world-class players, such as Aubameyang, Lacazette, or Ozil. As a

Forecasting Football Outcomes to Invest in Betting Markets

consequence of that, Arsenal has nearly the double of goals scored and points obtained.

Home/away performance tables:

- **Arsenal-** won 75% of its points at home and 25% when playing away, scoring also 72% of its goals at home and 28% when playing away, conceding 41% of its goals at home and 59% away
- **Southampton-** won 57% of its points at home and 43% when playing away, scored 55% of its goals at home and 45% when playing away, conceding 45% of its goals at home and 55% away
- **Premier League-** On average, 57% of the goals were scored at home and 43% were scored away
- **Analysis-** This table provides numerical evidence that Arsenal is very reliant on the home advantage, even in comparison with the Premier League's average. Southampton scored 10% more and conceded 10% less goals at home. Also, the team achieved 14% more points at St Mary's, in spite of being the second worst team at home, just above West Bromwich. Given so, we find these numbers consistent with the performance below the expectations that Southampton is demonstrating this season.

Points per game tables:

- Arsenal has a relative form of -0.15, while Southampton scores -0.02. This means that in the last 8 matches, Southampton conquered less points on average, than the average of points conquered over the course of the entire season, so far. Nonetheless, Arsenal has a relative form 0.13 points lower than The Saints. That being said, it is easy to understand that Southampton is facing a better moment than The Gunners, who lost 4 of their last 8 matches

Scoring tables:

- **Arsenal** scores, on average, 2.63 goals per match at home, while Southampton concedes 1.73 goals per match when playing away. If we average these numbers, we obtain Arsenal's "scoring strength", which corresponds to 2.18 goals

Forecasting Football Outcomes to Invest in Betting Markets

- **Southampton** scores, on average, 0.87 goals per match when playing away, while Arsenal concedes 1.06 goals per match at home. If we average these numbers, we obtain Southampton's "scoring strength", which corresponds to 0.97 goals

Team characteristics:

- Arsenal likes to play a possession-based football, using short passes to control the game in the opposition's half. However, Southampton is strong at stealing the ball from the opposition
- Both teams are weak in aerial duels, and also at defending counter-attacks
- Southampton is weak and Arsenal is even weaker at avoiding individual errors
- Arsenal is weak at defending attacks down the wings, and Southampton likes to play with width, specially down the left wing. In contrast, Arsenal prefers to conduct its attacks through the middle
- Southampton is weak at protecting the lead, while Arsenal is strong at coming back from losing positions
- Arsenal is strong at creating long shot opportunities, while Southampton is very weak at defending against long shots
- Arsenal is strong at finishing scoring chances, while Southampton is weak

Team psychology:

- **Team spirit and motivation-** In the Premier League, Arsenal is 5 points above the 7th placed Burnley and 5 points below the 5th ranked Chelsea. Arsène Wenger is a spent manager, that can no longer produce good results and inculcate a positive spirit among the players. The veteran Frenchman has a contract for another season, but will only remain as Arsenal's manager if he is able to win the Europa League, and to perform well in the last matches of the current season. The big question is if whether the players want Arsène to remain in charge or not, and because of this interrogation, we classify The Gunners' team spirit and motivation as medium. Arsenal is the 3rd best team in the League at home, and will face an opponent that is struggling to leave relegation zone. In fact, The Saints are desperate to get points, in order to secure a place in the next edition of the Premier

League, so they should be highly motivated to obtain at least 1 point from this match

- **Confidence-** Arsenal is performing below the expectations in the Premier League, ranking 6th, 13 points below the 4th place, that this season grants access to the 2018/2019 Champions league groups stage. Although The Gunners failed to defend the FA Cup title, losing away to Nottingham Forest in the 3rd round, they were able to reach the League Cup final (lost 3-0 to Manchester City), and still have good chances to qualify for the semi-finals of the Europa League, after defeating CSKA in the first leg (4-1), at The Emirates Stadium. The Saints are having their worst season since they got promoted back to the Premier League, in 2012/2013. The club has a relatively good squad, and was expected to finish comfortably above mid-table, as it has been the case over the last years. The fact is that Southampton is in relegation zone, and had to sack Mauricio Pellegrino, whose spell at St Mary's should be described as terrible, to say the least. When it comes to the national cups, The Saints were eliminated from the League Cup by Wolverhampton, in the 2nd round (2-0 at home), but managed to qualify for the semi-finals of the FA Cup, which is their last hope of saving the season. After evaluating the situation of both clubs, we believe that Arsenal's confidence should be classified as medium, and Southampton's as very low
- **Managerial impact-** Although Arsenal do not win the Premier League since 2003/2004, Arsène Wenger has been in charge for over 22 years. For better or worse, his impact on the club has to be considered very high. Mauricio Pellegrino was fired in the match day 31, after losing 3-0 away vs Newcastle, leaving Southampton in 17th place, just 1 point above the relegation zone. After that, Mark Hughes took charge and in his first match, he was able to qualify Southampton for the semi-finals of the FA Cup, defeating Wigan away (2-0). However, in the next match, Southampton categorically lost 3-0 away vs Newcastle, and was sent to relegation zone. Because Mark Hughes did only 2 matches at the helm of Southampton, we decided that it would be unfair to classify his managerial impact on the club

Team fatigue:

- **Toughness of previous match-** In the last match day, Arsenal beat Stoke City 3-

Forecasting Football Outcomes to Invest in Betting Markets

0, at home. However, Stoke's defence was a tough nut to crack, as The Gunners only scored the first goal by the minute 75, through a penalty kick converted by the newly-signed Pierre-Emerick Aubameyang. Southampton, by its turn, was destroyed by West Ham United, losing 3-0 away, with all the goals being scored within the first 45 minutes of the match. Having this in consideration, we consider Arsenal's toughness of previous match as medium, and Southampton's as the highest

- **Resting-** In the match *versus* Stoke, Petr Cech was unavailable, due to a groin injury, and Ospina was selected to replace him. Wenger also preferred to use Monreal as left full-back (played the entire match), instead of Kolasinac. Against West Ham, Southampton used a more offensive 4-4-2 formation, with Hojbjerg (box-to-box midfielder) instead of Oriol Romeu (defensive midfielder), and the Italian striker Gabbiadini instead of the attacking midfielder Davis. As Arsenal is, at least in theory, a stronger opponent than West Ham, it makes sense that Mark Hughes tries to reinforce the midfield area, risking less than against The Hammers. During the match, left-winger Redmond was replaced by Boufal (minute 67). In conclusion, from the probable line-ups for today, the number of players that rested was 2 for each side

- **National team participation-** The last round of international matches finished at 28/03/2018, and so this will not be considered

Poisson regression model:

	0	1	2	3	4	5	6	7+	Home
0	4,93%	10,97%	12,22%	9,07%	5,05%	2,25%	0,84%	0,36%	
1	3,86%	8,59%	9,57%	7,11%	3,96%	1,76%	0,65%	0,28%	
2	1,51%	3,37%	3,75%	2,78%	1,55%	0,69%	0,26%	0,11%	
3	0,39%	0,88%	0,98%	0,73%	0,40%	0,18%	0,07%	0,03%	
4	0,08%	0,17%	0,19%	0,14%	0,08%	0,04%	0,01%	0,01%	
5	0,01%	0,03%	0,03%	0,02%	0,01%	0,01%	0,00%	0,00%	
6	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	
7+	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	
Away									
	4,93%	8,59%	3,75%	0,73%	0,08%	0,01%	0,00%	0,00%	
Home win	70,22%		Odd	1,42					
Draw	18,09%		Odd	5,53					
Away win	11,67%		Odd	8,57					
Total	100%								

Forecasting Football Outcomes to Invest in Betting Markets

- **Analysis-** Our Poisson regression model grants more than 70% chances to an Arsenal's win, which can be converted into an odd of 1.42. However, Southampton chances might be overestimated, as the team is performing much worse this season than in the last one. If we made our decision just by looking at these numbers, betting on Arsenal at 1.50 would surely have positive expected value.

Betting decision:

- In 2017/2018, Southampton is performing highly below the initial expectations, currently facing the serious possibility of being relegated to the Championship. Because of that, the team must feel motivated (and pressured) to get points, but on the other hand, the players confidence levels should be very low, at this point. In contrast, Arsenal is in a much more neutral situation, as the club has its position in the Premier League perfectly stabilized (which is not a good thing, in this case). The Gunners have been a very consistent team at home, and have world-class players in the attack, so we can expect them to dominate the possession of the ball, with Southampton trying to hold their offensive line, and then exploiting Arsenal's defensive weaknesses through the wings, with crosses targeting the head of powerful striker Charlie Austin. Our Poisson regression model establishes the threshold of positive expected value at 1.42, which emphasizes even more our belief that backing an Arsenal's win at 1.50, given the context of both teams, has positive expected value.

Annex IX- Arsenal vs West Ham (match day 35)

- 22/04/2018 at 1:30 pm, Emirates Stadium (London)
- Weather: 21°C; precipitation probability of 40%; humidity percentage of 53% and wind speed of 23 km/h
- **Analysis:** Acceptable weather conditions to play football in the UK, at this time of the year. It is likely that rains during the match, but UK-based players are used to that situation. Probably, it will not have a huge impact in the outcome of the event.

Potential bets:

- Arsenal- 1.41
- Arsenal and over 1.5 goals- 1.58

Probable line-ups:

- **Arsenal**- 4-2-3-1: Cech; Bellerín, Mustafi, Chambers and Kolasinac; Elneny and Khaka; Nelson, Iwobi and Welbeck; Aubameyang
- **West Ham**- 3-4-2-1: Hart; Rice, Ogbonna and Cresswell; Zabaleta, Kouyaté, Noble and Masuaku; Mário and Lanzini; Arnautovic

Injuries and suspensions:

- **Arsenal**- Out: Cazorla and Mkhitarjan (all injured)
- **West Ham**- Out: Obiang, Reid, Byram and Antonio; doubtful: Collins (all injured)
- **Analysis**- Besides the “chronically” injured Santi Cazorla, Arsenal has 1 important player missing (Mkhitarjan). West Ham, by its turn, will miss the contributions of mid-importance players like the defender Winston Reid, the winger Michail Antonio or the midfielder Pedro Obiang.

Discipline table:

- **Arsenal** (average per match at home)- Yellow cards: 1.4, red cards: 0.1, booking

Forecasting Football Outcomes to Invest in Betting Markets

points: 15.3

- **West Ham** (average per match away)- Yellow cards: 2.0, red cards: 0.1, booking points: 22.8
- **Premier League** (team average per match overall)- Yellow cards: 1.6, red cards: 0.1, booking points: 16.7
- **Analysis-** At home, Arsenal is less aggressive than most Premier League's teams, while West Ham is highly above the average in terms of aggression, when playing away. This means that The Hammers may win more duels and recover the ball more swiftly, but on the other hand, they will probably be more exposed to disciplinary conditioning through yellow and red cards.

Head-to-head over the last 3 years:

- **8 matches-** 5 wins for Arsenal (62,5%), 2 draws (25%) and 1 win for West Ham (12,5%); 6 times over 1.5 (75%) and 2 (25%) times under
- **Analysis-** Considering the outcomes of the matches between these 2 teams over the last 3 years, it is possible to identify a bias favouring Arsenal. The last match ended with an Arsenal's home win (1-0), which eliminated West Ham from the League Cup. However, the last Premier League match between Arsenal and West Ham ended in a tie (0-0), at the London Stadium. When it comes to the number of goals, although the vast majority of the matches ended with over 1.5 goals, the last 2 matches (both in 2017/2018) ended 0-0 and 1-0, in Arsenal's favour.

Premier league tables overall, home and away:

- **Overall-** Arsenal: 6th place, 54 points, 62 goals scored and 45 conceded; West Ham: 14th place, 35 points, 41 goals scored and 59 conceded
- **Home-** Arsenal: 3rd place, 41 points, 45 goals scored and 19 conceded
- **Away-** West Ham: 15th place, 12 points, 21 goals scored and 38 conceded
- **Analysis-** Arsenal is one of the teams in the Premier League that depends more on the home advantage factor to score goals and win matches, as 41 out of 54 points, and 45 out of 62 goals were obtained at home. Besides that, Arsenal

Forecasting Football Outcomes to Invest in Betting Markets

conceded 16% more goals when playing away from the Emirates. On the other hand, The Hammers scored virtually the same number of goals at home and away, although they conceded 28% more goals when playing away. Also, almost 2/3 of West Ham's league points were achieved at the London Stadium. This attests West Ham's dependence on the home advantage factor, especially when it comes to its aptitude of defending solidly. In the Premier League table, West Ham is ranked 8 places below Arsenal, with less 19 points, less 21 goals scored and more 14 goals conceded. In fact, The Hammers have the worst defence away from home, and the third worst defence overall (just above Watford and Stoke). As Arsenal has a very strong offensive line, composed by several world-class players, such as Aubameyang, Lacazette, or Ozil, we anticipate that West Ham may experience severe difficulties at the Emirates Stadium.

Home/away performance tables:

- **Arsenal-** won 76% of its points at home and 24% when playing away, scoring also 73% of its goals at home and 27% when playing away, conceding 42% of its goals at home and 58% away
- **West Ham-** won 66% of its points at home and 34% when playing away, scored 49% of its goals at home and 51% when playing away, conceding 36% of its goals at home and 64% away
- **Premier League-** On average, 56% of the goals were scored at home and 44% were scored away
- **Analysis-** This table provides numerical evidence that Arsenal is very reliant on the home advantage, even in comparison with Premier League's average. As mentioned above, West Ham scored 49% of its goals at home and 51% when playing away. Also, the team conceded 28% more goals and achieved 32% less points away from the London Stadium. Summing up, both teams are considerably dependent on the home advantage factor, but Arsenal is even more dependent than The Hammers, which is interesting once that, at least in theory, better teams should be less dependent of the home advantage to win.

Points per game tables:

Forecasting Football Outcomes to Invest in Betting Markets

- Arsenal has a relative form of -0.14, while West Ham scores -0.06. This means that in the last 8 matches, West Ham conquered less points on average, than the average of points conquered over the course of the entire season, so far. Nonetheless, Arsenal has a relative form 0.09 points lower than The Hammers. That being said, it is easy to understand that The Gunners are facing an even worse moment than West Ham United

Scoring tables:

- **Arsenal** scores, on average, 2.65 goals per match at home, while West Ham concedes 2.24 goals per match when playing away. If we average these numbers, we obtain Arsenal's "scoring strength", which corresponds to 2.45 goals. Besides that, Arsenal's matches at home ended with over 1.5 goals 94% of the times
- **West Ham** scores, on average, 1.24 goals per match when playing away, while Arsenal concedes 1.12 goals per match at home. If we average these numbers, we obtain West Ham's "scoring strength", which corresponds to 1.18 goals. Besides that, West Ham's away matches ended with over 1.5 goals 94% of the times, as well
- **Analysis-** If we sum Arsenal and West Ham scoring strengths, we obtain a total of 3.63 goals expected in the forthcoming match. Besides that, if we average both teams' percentage of matches that ended with over 1.5 goals, we get a 94% probability that this match in particular ends with over 1.5 goals, which can be converted into an odd of 1.06

Team characteristics:

- Arsenal likes to play a possession-based football, using short passes to control the game in the opposition's half. West Ham, by its turn, uses the typical British style of play, often attempting long balls and crosses to the head of an offensive reference, such as Andy Carroll. As Arsenal defends high on the pitch, using the offside trap strategy, there might be some space to exploit behind The Gunners defensive line
- Both teams are weak at avoiding offside, and also at defending counter-attacks
- Arsenal is weak at defending attacks down the wings, and West Ham likes to play

with width. In contrast, Arsenal prefers to conduct its attacks through the middle

- Arsenal is weak at protecting the lead, while both teams are strong at coming back from losing positions
- Arsenal is strong at attacking and defending set pieces, while West Ham is weak at defending set pieces

Team psychology:

- **Team spirit and motivation-** In the Premier League, Arsenal is 2 points above the 7th placed Burnley and 6 points below the 5th ranked Chelsea. Arsène Wenger is a spent manager, that can no longer produce good results and inculcate a positive spirit among the players. For that reason, the veteran Frenchman will not finish his contract (expires in the summer of 2019), leaving Arsenal's helm by the end of this season. Arsenal is the 3rd best team in the League at home, and will face an opponent that is struggling to avoid relegation. In fact, The Hammers need to get points in order to secure a place in the next edition of the Premier League, hence they should be highly motivated to obtain at least 1 point from this match. Furthermore, this is a London derby, which by itself is a factor of motivation, due to the historical rivalry between 2 teams that are based in the same city. Consequently, we believe that both teams' spirit and motivation should be considered high
- **Confidence-** Arsenal is performing below the expectations in the Premier League, ranking 6th, 13 points below the 4th place, that this season grants access to the 2018/2019 Champions league groups stage. Although The Gunners failed to defend the FA Cup title, losing away to Nottingham Forest in the 3rd round, they were able to reach the League Cup final (lost 3-0 to Manchester City), and still have the chance to qualify for the final of Europa League, after drawing with Atlético de Madrid in the first leg of the semi-finals (1-1), at The Emirates Stadium. West Ham, by its turn, is fighting to avoid relegation, placing 14th in the league with 35 points (7 points above the relegation zone). The Hammers' performances in this season have been very irregular, as they have been capable of the best and the worst. In the domestic cups, West Ham was eliminated by Wigan in the 4th round of the FA Cup, and by Arsenal in the quarter-finals of the

Forecasting Football Outcomes to Invest in Betting Markets

League Cup. After evaluating the situation of both clubs, we believe that Arsenal's confidence should be classified as medium, and West Ham's as low

- **Managerial impact-** Although Arsenal do not win the Premier League since 2003/2004, Arsène Wenger has been in charge for over 22 years. For better or worse, his impact on the club has to be considered very high. This summer, West Ham invested a considerable sum in well-known players, such as Arnautovic or Chicharito, targeting for a place in the top half of the table, maybe even a European qualification spot. The expectations were high, but West Ham began the season terribly, and Slaven Bilic was fired after being defeated by Liverpool (1-4 at home, in the match day 11), leaving The Hammers in relegation zone. In his replacement, the board appointed David Moyes, who was able to take the club out of the relegation zone, although a place among the teams that will dispute the 2018/2019 edition of the Premier League is not secured yet. In spite of Moyes' ability to improve the club's situation in the league table, the reality is that West Ham has a fairly good squad, and should be performing much regularly. For that reason, we decided to classify Moyes' managerial impact on the club as medium

Team fatigue:

- **Toughness of previous match-** In the last match day Arsenal scored first, but eventually loss 2-1 to Newcastle, at St. James Park. West Ham, by its turn, disappointingly tied at home with Stoke City, in a match where the giant veteran Peter Crouch scored first for The Potters (by the minute 79), but the home team was able to equalize 10 minutes later, by Andy Carroll, who entered the pitch just 4 minutes earlier. Having this in attention, we consider that the toughness of previous matches was high, in both cases
- **Resting-** In the match *versus* Newcastle, Arsène Wenger used a 4-3-3 formation, with an offensive line composed by Lacazette (as striker), Aubameyang and Iwobi (as wingers, instead of Nelson and Welbeck). Besides that, the Frenchman reinforced the midfield with Willock playing alongside Elneny and Xhaka. Wenger also preferred to use Monreal as the left full-back (played the entire match), instead of Kolasinac, with Chambers playing as right full-back and Rob Holding as central defender, instead of Bellerín. During the match, Willock was

Forecasting Football Outcomes to Invest in Betting Markets

replaced by Welbeck (68th minute), Chambers was replaced by Maitland-Niles (78th minute) and Iwobi was replaced by Nketiah (86th minute). Against Stoke City, West Ham began with the same 11 as the probable line-up for this match, but with Fernandes instead of Lanzini. During the match, Fernandes was replaced by Chicharito (minute 76), Masuaku was replaced by Lanzini (minute 81) and João Mário was replaced by Andy Carroll (minute 86)

- **National team participation-** The last round of international matches finished at 28/03/2018, and so this will not be considered

Poisson regression model:

	0	1	2	3	4	5	6	7+	Home
0	2,22%	6,24%	8,77%	8,22%	5,78%	3,25%	1,52%	0,92%	
1	2,21%	6,22%	8,74%	8,19%	5,76%	3,24%	1,52%	0,92%	
2	1,10%	3,10%	4,35%	4,08%	2,87%	1,61%	0,76%	0,46%	
3	0,37%	1,03%	1,45%	1,36%	0,95%	0,54%	0,25%	0,15%	
4	0,09%	0,26%	0,36%	0,34%	0,24%	0,13%	0,06%	0,04%	
5	0,02%	0,05%	0,07%	0,07%	0,05%	0,03%	0,01%	0,01%	
6	0,00%	0,01%	0,01%	0,01%	0,01%	0,00%	0,00%	0,00%	
7+	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	
Away									
	2,22%	6,22%	4,35%	1,36%	0,24%	0,03%	0,00%	0,00%	
Home win	74,99%	Odd		1,33					
Draw	14,42%	Odd		6,93					
Away win	10,61%	Odd		9,43					
Total	100%								
Under 1.5	10,67%	Odd		9,37					
Over 1.5	89,33%	Odd		1,12					
Total	100%								

- **Analysis-** Our Poisson regression model grants almost 75% chances to an Arsenal's win, which can be converted into an odd of 1.33. However, West Ham chances might be overestimated, as the team is performing worse this season than in the last one. The model also attributes a probability of 89,33% to this match ends with more than 1.5 goals, which corresponds to an odd of 1.12. If we sum $1.33 + 0.12$ (the threshold of positive expected value for an Arsenal's win, plus the threshold for the match ending with more than 1.5 goals), we obtain an odd of 1.46, which means that according to this model, betting on Arsenal at 1.41, and on Arsenal and over 1.5 goals at 1.58 would have positive expected value.

Betting decision:

- In 2017/2018, West Ham is performing below the initial expectations, currently facing the possibility of being relegated to the Championship. Because of that, the team must feel motivated (and pressured) to get points, especially due to the rivalry inherent to a derby, and the importance that the fans attribute to win against a club from the same city (this last part is even more true for Arsenal, once that this is the last derby in Wenger's era). In contrast, Arsenal is in a much more neutral situation, as the club has its position in the Premier League perfectly stabilized (which is not a good thing, by the way). The historical head-to-head over the last 3 years is clearly favourable to Arsenal, and both teams' dependence on the home advantage factor also favours Arsenal, in this specific case (furthermore because the League's second-best attack at home will face-off the worst defence when playing away). Besides that, The Gunners have been a very consistent team at home, and have world-class attacking players, thus we can expect them to dominate the possession of the ball, with West Ham trying to hold their offensive line, and then exploiting Arsenal's defensive weaknesses through crosses delivered from the wings, and also through poisonous counter attacks. The scoring table analysis tells us that there is a 94% probability that this match ends with over 1.5 goals (equivalent to an odd of 1.06), and our Poisson regression model establishes the threshold of positive expected value at 1.33 for an Arsenal's win, and at 1.12 to the match ends with more than 1.5 goals. Considering all these factors, we believe that the possibility of returning 58% on this investment seems highly attractive.

Annex X- Tottenham vs Watford (match day 36)

- 30/04/2018 at 8:00 pm, Wembley Stadium (London)
- Weather: 5°C; precipitation probability of 4%; humidity percentage of 74% and wind speed of 14 km/h
- **Analysis:** Great weather to play football in the UK, at this time of the year. It is unlikely that rains, which allows for a fast circulation of the ball, and consequently for an open match.

Potential bets:

- Under 3.5 goals in the match- 1.58

Probable line-ups:

- **Tottenham**- 4-2-3-1: Lloris; Trippier, Sánchez, Verthongen and Davies; Dier and Dembelé; Eriksen, Alli and Son; Kane
- **Watford**- 4-2-3-1: Karnezis; Janmaat, Mariappa, Kabasele and Britos; Doucoure and Capoue; Femenía, Hughes and Richarlison; Deeney

Injuries and suspensions:

- **Tottenham**- Out: Winks (injured)
- **Watford**- Out: Hoban, Chalobah, Cleverley, Deulofeu and Okaka; doubtful: Prodl, Carrillo and Pereyra (all injured)
- **Analysis**- Watford has 4 long-time injured players (Hoban, Chalobah, Cleverley and Deulofeu), plus 3 rotation players in doubt (Prodl, Carrillo and Pereyra), while Tottenham will only miss its young talent Harry Winks (who is injured since the 20th of March).

Discipline table:

- **Tottenham** (average per match at home)- Yellow cards: 1.0, red cards: 0.0, booking points: 10.0
- **Watford** (average per match away)- Yellow cards: 1.3, red cards: 0.1, booking

Forecasting Football Outcomes to Invest in Betting Markets

points: 15.9

- **Premier League** (team average per match overall)- Yellow cards: 1.6, red cards: 0.1, booking points: 16.7
- **Analysis-** Both teams are below the Premier League average in terms of aggressiveness, therefore, it seems unlikely that cards or injuries might condition the outcome of this match.

Head-to-head over the last 3 years:

- **5 matches-** 4 wins for Tottenham (80%) and 1 draw (20%); 3 times under 3.5 (60%) and 2 times over (40%)
- **Analysis-** Given these numbers, it is fair to conclude that there is a clear bias favouring Tottenham, however, the last Premier League's match between these teams ended up in a draw, at Vicarage Road.

Premier league tables overall, home and away:

- **Overall-** Tottenham: 4th place, 68 points, 66 goals scored and 31 conceded, Watford: 12th place, 34 points, 42 goals scored and 60 conceded
- **Home-** Tottenham: 5th place, 34 points, 32 goals scored and 12 conceded
- **Away-** Watford: 11th place, 14 points, 17 goals scored and 30 conceded
- **Analysis-** These numbers demonstrate that Tottenham is a much stronger team, investing a lot more to be placed among the Champions League places (top 4), while Watford is just a team that fights for achieving its stability in the Premier League. Overall, Tottenham has the double of Watford's points, more than that in terms of the goals scored, and conceded virtually half as much goals as The Hornets did, which is a very considerable difference.

Home/away performance tables:

- **Tottenham-** won 50% of its points at home and 50% when playing away, scoring also 48% of its goals at home and 52% when playing away, conceding 39% of its goals at home and 61% away

Forecasting Football Outcomes to Invest in Betting Markets

- **Watford-** won 63% of its points at home and 37% when playing away, scoring also 60% of its goals at home and 40% when playing away, conceding 50% of its goals at home and 50% away
- **Premier League-** On average, 57% of the goals were scored at home and 43% were scored away
- **Analysis-** By looking at these numbers, it is possible to conclude that Tottenham does not rely on the home advantage factor to get points, and also to score goals, which is a sign of consistency. However, Tottenham depends on the home advantage to avoid conceding goals. In the summer of 2017, Tottenham's board made an effort to keep all the most important players from the previous season, but sold the right full-back Kyle Walker to Manchester City, for a fee that was, at the time, the highest ever paid for a defender. Besides that, the left full-back Danny Rose was injured most of the season, and the newly-signed young prodigy Davinson Sánchez needed some time to adjust to his new team, and to the English style of football, as well. This conjugation of factors provides an explanation for the lack of defensive reliability shown in some of The Spurs' matches. Watford, by its turn, is significantly dependent on the home advantage to obtain points, and is slightly above the Premier League's average when it comes to the number of goals scored at home. The Hornets conceded exactly the same number of goals at home and away, which in this case suggests defensive permeability, as Watford has the 3rd worst defence of all the Premier League' teams.

Points per game tables:

- Tottenham has a relative form of +0.38, while Watford scores -0.09. This means that in the last 8 matches, the latter conquered less points on average, than the average of points conquered over the course of the entire season, so far. On the other hand, Tottenham has a positive relative form, that is +0.47 points higher than the one relative to The Hornets. That being said, it is easy to understand that The Spurs are facing a much better moment than Watford

Scoring tables:

- **Tottenham** scores, on average, 2.00 goals per match at home, while Watford concedes 1.76 goals per match when playing away. If we average these numbers,

Forecasting Football Outcomes to Invest in Betting Markets

we obtain Tottenham's "scoring strength", which corresponds to 1.88 goals. Besides that, Tottenham's matches at home ended with under 3.5 goals 69% of the times

- **Watford** scores, on average, 1.00 goals per match when playing away, while Tottenham concedes 0.75 goals per match at home. If we average these numbers, we obtain Watford's "scoring strength", which corresponds to 0.88 goals. Besides that, Watford's away matches ended with under 3.5 goals 71% of the times
- **Analysis-** If we sum Tottenham and Watford scoring strengths, we obtain a total of 2.76 goals in the forthcoming match. Besides that, if we average both teams' percentage of matches that ended with under 3.5 goals, we get an 70% probability that this match in particular ends with under 3.5 goals, which can be converted into an odd of 1.43.

Team characteristics:

- Tottenham likes to play a possession-based football, using short passes to control the game in the opposition's half. However, Tottenham is also very strong in counter attacks, often using through balls to create scoring chances. In contrast, Watford is very weak at defending counter attacks. As both teams use an offside trap defensive strategy, there might be some space to exploit behind the defensive lines' backs, although The Spurs are weak at avoiding offside
- Tottenham is strong at attacking and defending set pieces, while The Hornets are just weak at defending them
- Tottenham is strong at protecting the lead, while Watford is very weak
- Watford preferentially conducts its attacks down the left flank, where its most talented player, Richarlison, is deployed. On the contrary, Tottenham usually attacks through the middle

Team psychology:

- **Team spirit and motivation-** Watford suffered upsetting early eliminations from domestic cups, but at this stage has already mathematically avoided relegation (the main goal for this season). Now, the only focus is to achieve the best possible

classification in the Premier League, preferentially placing within the first half of the table. The team seems united enough around Gracia, although some important players expressed, at the time, their dissatisfaction about Silva's dismissal (e.g. Richarlison), and so we consider that Watford's team spirit and motivation should be classified as medium, at this point. This has been an unusual season for Tottenham, who is playing all of its home games at the national Wembley Stadium, while the new White Hart Lane is under construction. Although The Spurs still have the support of its fans, being temporarily separated from their home ground probably reduces the home advantage benefits to the team's spirit and motivation. However, Tottenham is clearly favourite and still has chances of reaching the 2nd place. Therefore, its players should be highly motivated to win this match

- **Confidence-** The Hornets have been performing much better in 2017/2018 than in the last season, where they ranked 17th, the last place above relegation zone. In the national cups, Watford was surprisingly eliminated from the League Cup by Bristol (2-3 at home), in the 3rd round, and defeated by Southampton (1-0 away) in the 4th round of FA Cup. Additionally, in the last 6 matches for the Premier League, Watford lost 4 times and tied twice, hence we think that the team's level of confidence should be classified as low. When it comes to the internal cups, Tottenham suffered a surprising early elimination from the League Cup, at the hands of city rivals West Ham (2-3 at home), but reached the semi-finals of the FA Cup (eliminated by Manchester United). In the Champions League, Tottenham passed the group stage and then had an electrifying clash against Italian titans Juventus, in which The Spurs performed incredibly well, but eventually ended up being eliminated (2-3 on aggregate), perhaps due to the team's lack of experience. At the moment, Tottenham is 4th in the Premier League, 5 points above rivals Chelsea, and 3 points below Liverpool (with a game in hand). As one can observe, the difference of scenarios between these teams is huge, and so, Tottenham should approach this match very confident in obtaining a good score
- **Managerial impact-** Mauricio Pochettino began his coaching career at Barcelona, being appointed as the manager of RCD Espanyol (the last team where

he played). After two and a half seasons, his work caught Southampton's attention, and he moved on to the Premier League. After just one season at St Mary's (2013/2014), the attractive style of play that Pochettino was able to implement granted him a place as the head coach of Tottenham Hotspurs, a club with higher ambitions, but struggling to reach Champions League qualification spots (6th place in 2013/2014). In his first season at White Hart Lane, Pochettino repeated the 6th place, but The Spurs' board decided to give him another shot. In the next 2 seasons (2015/2016 and 2016/2017), the Argentinian put The Spurs back on track, placing 3rd and 2nd respectively, and fighting for the Premier League title until the last match days. Furthermore, this season Tottenham qualified for the knockout stages of the Champions League for the first time since 2010/2011 (lost dramatically to last year's finalist Juventus). With Pochettino, Tottenham achieved consistent results, playing stylish football while investing relatively less than its direct rivals, consequently, his managerial impact must be considered very high. Marco Silva was fired on the 20th January, after suffering a categorical away defeat against Leicester (2-0), being replaced by Javi Gracia, who started by changing the usual formation from 3-4-2-1 to 4-2-3-1. Besides that, some players that were important under Silva's management, like Hughes, Carrillo or Gray, lost relevance to others, such as Roberto Pereyra or Troy Deeney. The Spanish coach began his spell by drawing away, against Stoke (0-0). After that, he achieved an incredible win at home against Chelsea (4-1), but then lost away to direct competitor West Ham (2-0). In the following 3 matches, Watford was able to beat Everton and West Brom (both 1-0) at home, but then lost 3-0 away vs Arsenal, and 5-0 vs Liverpool (both times when playing away). Since Javi Gracia took the helm, Watford has been stable in the mid-table, therefore we think that his managerial impact at Vicarage Road has been normal

Team fatigue:

- **Toughness of previous match-** In the previous match, Watford and Crystal Palace tied 0-0 at Vicarage Road. The Hornets were superior to Palace, but were unable to materialize that superiority in goals, although Crystal Palace had some chances of its own. On the same day (21st of April), Tottenham and Manchester United disputed a place in the FA Cup final, at Wembley. The Spurs scored first,

Forecasting Football Outcomes to Invest in Betting Markets

by Dele Alli (11th minute), dominated possession (64%) and had more attempts on goal, but were less effective, and allowed United to turn things in its favour, with goals scored by Alexis Sánchez (24th minute) and Ander Herrera (62nd minute). Having all this in consideration, we decided to classify Tottenham's toughness of last match as the highest, and Watford's as medium

- **Resting-** Against Crystal Palace, the only players from Watford's probable line-up that rested entirely were the full-backs, Janmaat and Britos. Besides that, Pereyra was replaced by Richarlison (55th minute), Okaka was replaced by Femenía (70th minute) and Hughes was replaced by Gray, 9 minutes later. Tottenham began the semi-final of the FA Cup exactly with the same probable line-up for the match against Watford, but replaced Ben Davies for Lucas Moura as a response to the goal scored by Herrera (68th minute), Wanyama for Dembelé (10 minutes later) and Lamela for Son (by the minute 86, in a final attempt to obtain the equalizing goal)
- **National team participation-** The last round of international matches finished at 28/03/2018, and so this will not be considered

Poisson regression model:

	0	1	2	3	4	5	6	7+	Home
0	2,78%	8,38%	12,64%	12,72%	9,60%	5,79%	2,91%	1,96%	
1	1,57%	4,74%	7,16%	7,20%	5,43%	3,28%	1,65%	1,11%	
2	0,44%	1,34%	2,02%	2,04%	1,54%	0,93%	0,47%	0,31%	
3	0,08%	0,25%	0,38%	0,38%	0,29%	0,18%	0,09%	0,06%	
4	0,01%	0,04%	0,05%	0,05%	0,04%	0,02%	0,01%	0,01%	
5	0,00%	0,00%	0,01%	0,01%	0,00%	0,00%	0,00%	0,00%	
6	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	
7+	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	
Away									
	2,78%	4,74%	2,02%	0,38%	0,04%	0,00%	0,00%	0,00%	
Home win	85,78%		Odd	1,17					
Draw	9,96%		Odd	10,04					
Away win	4,23%		Odd	23,64					
Total	100%								
Under 3.5	51,85%		Odd	1,93					
Over 3.5	48,15%		Odd	2,08					
Total	100%								

- **Analysis-** Our Poisson regression model grants more than 85% chances to a home

win, less than 10% to a draw, and only 4,23% to an away win. Although the model strongly indicates that Tottenham has very high probabilities of winning, when it comes to the number of goals in the match, the probabilities for under and over 3.5 goals are more or less even. In fact, the model estimates a probability of 51,85% for the former (equivalent to an odd of 1.93), and of 48,15% for the latter (corresponds to an odd of 2.08). If we based our decision only on the data provided by the Poisson regression, betting on under 3.5 goals in this match at 1.58 could not be seen as a reasonable choice.

Betting decision:

- After carefully analysing all the factors that may influence the outcome of this match, one thing that becomes clear is The Spurs' favouritism to win. As a matter of fact, unlike Watford, Tottenham has no relevant players missing or at risk for this match, and is going through a positive moment of recent form. In addition, the historical head-to-head over the last 3 years suggests that there is a bias in favour of the home team. The Poisson regression grants Tottenham 85,78% chances of winning this match, which is equivalent to an odd of 1.17, and a probability of 51,85% that this match ends with under 3.5 goals, which corresponds to an odd of 1.93. However, the scoring tables tell us that there is a 70% probability that the match ends with less than 3.5 goals, which can be converted into an odd of 1.43. Moreover, the sum of both teams' scoring strengths is equal to 2.76 goals. We believe that Tottenham will win this match, but will not try very hard to score a lot of goals, in an attempt to manage the players' effort, as many of them will be called for the FIFA World Cup 2018. Furthermore, we expect Watford to approach this match with caution, aware of Tottenham's goal-scoring ability. In this case, the subjective analysis will outweigh the Poisson quantitative analysis in our decision-making process. In combination with the data retrieved from the scoring tables, it makes us feel confident in the success of betting on under 3.5 goals, at 1.58.

Annex XI- Southampton vs Manchester City (match day 38)

- 13/05/2018 at 3:00 pm, St Mary's Stadium (Southampton)
- Weather: 16°C; precipitation probability of 2%; humidity percentage of 61% and wind speed of 7 km/h
- **Analysis:** Weather as perfect as it gets to play football in the UK, allowing for a fast circulation of the ball, and consequently for an open match.

Potential bets:

- Manchester City- 1.46
- Manchester City and over 1.5 goals- 1.55

Probable line-ups:

- **Southampton-** 3-4-2-1: McCarthy; Stephens, Yoshida and Hoedt; Soares, Romeu, Hojbjerg and Bertrand; Redmond and Tadic; Austin
- **Manchester City-** 4-3-3: Ederson; Walker, Stones, Otamendi and Mendy; De Bruyne, Fernandinho and Gundogan; Sterling, Jesus and Sané

Injuries and suspensions:

- **Southampton-** Out: Davis (injured) and Boufal (internal disciplinary issues); doubtful: Lemina and Bednarek (both injured)
- **Manchester City-** Out: Aguero (injured)
- **Analysis-** Southampton will miss 2 important players: the central midfielder Steven Davis and the attacking midfielder Soufiane Boufal, who was banished from The Saints' first team by Mark Hughes, after refusing to enter the pitch against Chelsea (2-3 loss, at home). Additionally, Mario Lemina and Bednarek are both having physical problems, and consequently in doubt for this match. On the other hand, Manchester City will only miss its striker Sergio Aguero. Although this is an important casualty, The Citizens have Gabriel Jesus, who is more than capable of performing at a similar level.

Discipline table:

Forecasting Football Outcomes to Invest in Betting Markets

- **Southampton** (average per match at home)- Yellow cards: 1.4, red cards: 0.0, booking points: 14.4
- **Manchester City** (average per match away)- Yellow cards: 1.7, red cards: 0.1, booking points: 18.6
- **Premier League** (team average per match overall)- Yellow cards: 1.6, red cards: 0.1, booking points: 16.7
- **Analysis-** At home, Southampton is less aggressive than most Premier League's teams, while Manchester City is way above the average in terms of aggression, when playing away. This means that The Citizens may win more duels and recover the ball more swiftly, but on the other hand, they are probably going to be more exposed to disciplinary conditioning through yellow or red cards.

Head-to-head over the last 3 years:

- **7 matches-** 1 win for Southampton (14%), 1 draw (14%) and 5 wins for Manchester City (72%); 7 times over 1.5 (100%)
- **Analysis-** Given these numbers, it is fair to conclude that there is a clear bias favouring Manchester City. On top of that, every single match between these teams over the last 3 years ended with over 1.5 goals.

Premier league tables overall, home and away:

- **Overall-** Southampton: 17th place, 36 points, 28 goals scored and 56 conceded, Manchester City: 1st place, 97 points, 105 goals scored and 27 conceded
- **Home-** Southampton: 19th place, 19 points, 20 goals scored and 25 conceded
- **Away-** Manchester City: 1st place, 47 points, 44 goals scored and 13 conceded
- **Analysis-** These numbers demonstrate that Manchester City is the best team in the Premier League, and of course, a much stronger team than Southampton. Overall, The Citizens have more 61 points, 77 goals scored and less 29 conceded than Southampton, which is a huge difference.

Home/away performance tables:

Forecasting Football Outcomes to Invest in Betting Markets

- **Southampton**- won 53% of its points at home and 47% when playing away, scoring also 54% of its goals at home and 56% when playing away, conceding 47% of its goals at home and 53% away
- **Manchester City**- won 52% of its points at home and 48% when playing away, scoring also 58% of its goals at home and 42% when playing away, conceding 52% of its goals at home and 48% away
- **Premier League**- On average, 57% of the goals were scored at home and 43% were scored away
- **Analysis**- This data allows one to understand that neither of the teams depend much on the home advantage factor to gain points, and also to score and prevent goals.

Points per game tables:

- The Saints have a relative form of -0.03, while City scores -0.17. At first glance, it may seem weird that Southampton has a better level of relative form than Manchester City. This is because The Citizens have been very consistent over the course of the season, only losing against Liverpool and Manchester United, and drawing against Everton, Crystal Palace, Burnley and Huddersfield. The defeat vs United and the draw vs Huddersfield happened over the last 8 matches. Southampton performed very poorly under Pellegrino's guidance. When the Argentinian was dismissed, Mark Hughes took charge and the results improved significantly, as Southampton was able to leave relegation zone, after winning 0-1 against Swansea (away). In the last 8 matches, The Saints won 2, lost 4 and tied 2

Scoring tables:

- **Southampton** scores, on average, 1.11 goals per match at home, while Manchester City concedes 0.72 goals per match when playing away. If we average these numbers, we obtain Southampton's "scoring strength", which corresponds to 0.92 goals. Besides that, Southampton's matches at home ended with more than 1.5 goals 72% of the times
- **Manchester City** scores, on average, 2.44 goals per match when playing away,

Forecasting Football Outcomes to Invest in Betting Markets

while Southampton concedes 1.39 goals per match at home. If we average these numbers, we obtain City's "scoring strength", which corresponds to 1.92 goals. Besides that, The Citizens' away matches ended with more than 1.5 goals 83% of the times

- **Analysis-** If we sum Southampton and Manchester City scoring strengths, we obtain a total of 2.84 goals in the forthcoming match. Moreover, if we average both teams' percentage of matches that ended with over 1.5 goals, we get a 78% probability that this match in particular ends with over 1.5 goals, which can be converted into an odd of 1.28

Team characteristics:

- Both teams are weak in aerial duels
- Manchester City is very strong at creating scoring chances, especially by using through balls, and also by using the individual skill of its players. Furthermore, City is very strong at finishing scoring chances, while Southampton is weak
- Man City is very strong at creating long shot opportunities, while The Saints are very weak at defending against long shots
- Manchester City is very strong at attacking and defending set pieces, and strong at shooting from direct free kicks
- Manchester City style is related with keeping possession of the ball, using short passes to control the game in the opposition's half. Moreover, The Citizens defend high on the pitch, using a very well automatized offside trap strategy, which allows them to be the team that suffered the less number of goals in the League, this season

Team psychology:

- **Team spirit and motivation-** Pep Guardiola promotes a squad rotation policy, so that every player feels useful and integrated. There are no cases (at least publicly known) of indiscipline or internal issues between the players and the manager. The way the team is performing (and winning) provides strong evidence of a positive team culture. The Citizens already assured the main objective of winning

the Premier League, but even so, they should be highly motivated to win it in style, targeting for an unprecedented record of 100 points. Although Southampton's season was a huge disappointment to its supporters, the club was able to duck the menace of relegation, which due to the team's underperformance, became the main objective for 2017/2018. In spite of the Soufiane Boufal's incident, the group showed union, tenacity and mental strength when pursuing the permanence in the Premier League, especially when was able to recover from the traumatic losses in the matches against Arsenal and Chelsea, obtaining a positive final sequence that culminated in the amazing win against Swansea. However, in this final match, there is nothing else to fight for, and consequently, there might be some propensity to relaxation. Given so, we consider that, for this match in particular, Southampton's levels of motivation should be low

- **Confidence-** In 2017/2018, The Citizens have been perhaps the best team in Europe, winning the Premier League comfortably in the match day 34, after Man United surprising defeat vs West Bromwich, at Old Trafford. As a matter of fact, Manchester City only lost to Shakhtar Donetsk (Champions League), Wigan (FA Cup), Liverpool (twice for the quarter-finals of the Champions League and once for the Premier League) and Manchester United (Premier League), and tied against Everton, Crystal Palace, Burnley and Huddersfield (all for Premier League). On these circumstances, we believe it is fair to say that Man City's confidence levels must be very high. On the contrary, The Saints are having their worst season since they got promoted back to the Premier League, in 2012/2013. The club has a relatively good squad, and was expected to finish comfortably above mid-table, as it has been the case over the last years. The fact is that Southampton barely avoided relegation, and had to sack Mauricio Pellegrino, whose spell at St Mary's should be described as terrible, to say the least. When it comes to the national cups, The Saints were eliminated from the League Cup by Wolverhampton, in the 2nd round (2-0 at home), but managed to qualify for the semi-finals of the FA Cup, in which they were eliminated by Chelsea (2-0 at Wembley). For those reasons, we believe that the team's confidence levels should be classified as low
- **Managerial impact-** Mauricio Pellegrino was fired in the match day 31, after

losing 3-0 away vs Newcastle, leaving Southampton in 17th place, just 1 point above the relegation zone. After that, Mark Hughes took charge and in his first match, was able to qualify Southampton for the semi-finals of the FA Cup, defeating Wigan away (2-0). However, in the next match, Southampton categorically lost 3-0 away vs Newcastle, and was sent to relegation zone. After that, a sequence of 3 losses in a row anticipated the worst scenario for The Saints, although they were a tough nut to crack against strong opponents like Arsenal and Chelsea. In the last 4 matches for the Premier League, Southampton was able to bounce back, achieving 2 wins and 2 draws, including a dramatic victory in the last match against Swansea, that guaranteed Southampton the relegation avoidance. Because of that, and also because the team improved significantly the quality of its football, we decided to classify Hughes' managerial impact in Southampton as high. Guardiola arrived at Manchester and after a mediocre first season, has been able to turn Manchester City into one of the most complete teams in Europe. The financial investment in the squad was huge, but nowadays City plays in the image of its coach, and is definitely one of the best sides that he built in the entire career. That being said, we have to consider Guardiola's impact in the club as very high

Team fatigue:

- **Toughness of previous match-** The Citizens' last match took place on the 9th of May, at the Etihad Stadium, where they achieved a comfortable win against Brighton (3-1). As usual, Man City dominated its opponent, and the final victory was never at risk, therefore, we decided to rate the toughness of City's last match as low. One day earlier, at Wales, Southampton disputed the first place above the relegation zone with Swansea. In a well-balanced match, Southampton was able to strike the only goal by the minute 72. Gabbiadini, who entered the match just 4 minutes earlier, was the goal scorer. Due to the importance of what was there at stake, and also because of the evenness of the match itself, we considered the toughness of Southampton's last match as the highest
- **Resting-** Against Brighton, Manchester City started with a completely different defensive line, in comparison with the probable line-up for this match: Danilo, Kompany, Laporte and Zinchenko began the match, instead of Walker, Stones,

Forecasting Football Outcomes to Invest in Betting Markets

Otamendi and Mendy, although the latter replaced Zinchenko by the minute 76. Besides that, Guardiola decided to spare the key midfielder Kevin De Bruyne, that was replaced by the veteran Yaya Touré (who was substituted by Nmecha, on the 86th minute). 3 minutes earlier, the Catalan manager already have changed Jesus for Diaz. When it comes to Southampton, the only difference between the starting 11 of the game against Swansea and the probable line-up for this match is that Yoshida is expected to substitute Bednarek (who is in doubt, as mentioned above). During the match at the Liberty Stadium, Redmond was replaced for Shane Long by the minute 64, and 4 minutes later, the defender Bednarek was replaced by the striker Manolo Gabbiadini. These changes were made with the aim of providing the firepower needed to score, and were proven successful, as Gabbiadini netted the winning goal that assured to The Saints a place in the next edition of the Premier League. The final substitution was made by the minute 83, when Mark Hughes decided to replace Tadic for McQueen

- **National team participation-** The last round of international matches finished at 28/03/2018, and so this will not be considered

Poisson regression model:

	0	1	2	3	4	5	6	7+	Home
0	5,63%	4,47%	1,77%	0,47%	0,09%	0,01%	0,00%	0,00%	
1	11,74%	9,31%	3,69%	0,97%	0,19%	0,03%	0,00%	0,00%	
2	12,23%	9,69%	3,84%	1,01%	0,20%	0,03%	0,00%	0,00%	
3	8,50%	6,73%	2,67%	0,70%	0,14%	0,02%	0,00%	0,00%	
4	4,43%	3,51%	1,39%	0,37%	0,07%	0,01%	0,00%	0,00%	
5	1,84%	1,46%	0,58%	0,15%	0,03%	0,00%	0,00%	0,00%	
6	0,64%	0,51%	0,20%	0,05%	0,01%	0,00%	0,00%	0,00%	
7+	0,25%	0,20%	0,08%	0,02%	0,00%	0,00%	0,00%	0,00%	
Away									
	5,63%	9,31%	3,84%	0,70%	0,07%	0,00%	0,00%	0,00%	
Home win	13,10%		Odd	7,63					
Draw	19,55%		Odd	5,12					
Away win	67,28%		Odd	1,49					
Total	100%								
Under 1.5	21,84%		Odd	4,58					
Over 1.5	78,16%		Odd	1,28					
Total	100%								

Forecasting Football Outcomes to Invest in Betting Markets

- **Analysis-** The Poisson regression model grants 67,28% chances to an away win, which can be converted into an odd of 1.49. The probability given by the Poisson regression to the match ends with more than 1.5 goals is 78,16%, which is equivalent to an odd of 1.28. It is important to understand that Southampton's chances might be overestimated, as they performed much better in 2016/2017 than in this season, while City's chances, on the contrary, are probably underestimated (as its performance improved a lot in the current season).

Betting decision:

- Manchester City is the best team in the Premier League (and quite possibly, in the world) by far. The difference to the other teams is enormous, as The Citizens have the best (and more expensive) squad (in terms of quality and quantity), the best manager, the most balanced and refined style of play (best attack and defence, both at home and away). Manchester City is an extremely confident team, that never renounces on its playing principles, regardless of the opponent's name or status. With the exception of the away matches against Liverpool, every single team that played against Manchester City just assumed its superiority and gave away the control of the match, trying not to suffer a heavy loss. Both teams already accomplished their main goals: Man City was crowned Premier League champion and Southampton avoided relegation, after an uncommon struggled season. However, City still wants to be the first team to surpass the barrier of the 100 points in the league. Also, the Poisson regression analysis allocates a probability of 67,28% to an away win (odd of 1.49), and of 78,16% (odd of 1.28) to the match ends with over 1.5 goals. The scoring table analysis grants the same probability when it comes to the total number of goals in this match, plus a scoring potential of 2.84 goals, resulting from the sum of both teams' scoring strengths. Given this context, our decision will be to bet on Manchester City and over 1.5 goals at 1.55, in an attempt to maximize the return on investment, as we are very confident that more than 1.5 goals will occur.