EVIDENCING THE USE OF SPATIAL ANALYSIS AND HOW THIS CAN BE APPLIED IN THE FOOTBALL INDUSTRY TO HIGHLIGHT THE MARKET POTENTIAL

Georgina Weaver
200840024
BA Geography
2017

Front Cover Photo = Author's Own

ACKNOWLEDGEMENTS

I would firstly like to thank my dissertation mentor, Luke Burns, for his continued support throughout the past year, as well as his patience in answering any questions I have had. Without him none of this work would have been anywhere near possible.

I would also like to thank my Dissertation Support Group for their help along the way, and for providing valuable assistance at all stages of our projects.

Special thanks to Bromley FC for supplying me with the data about their season ticket holders, and not to mention being my inspiration for my research work.

Finally, appreciation must go to my friends and family for their encouragement throughout the whole process, and their enthusiasm in my work over the years.

CONTENTS

ABSTRACT

Football is one of the most popular sports in the UK, with thousands of fans across the country going to watch their beloved team every week. Despite its popularity, it is becoming increasingly important for football clubs at all levels to find ways to continuously improve their attendances, which is what the research throughout this project aims to do. The use of geodemographics has been growing in recent years, however there is little evidence of its application to the sports industry. Literature may highlight what encourages people to attend football matches, but this information has not yet been used in methods to spatially analyse the different places where football supporters live. This study will create a geodemographic index to spatially analyse areas of Greater London, to establish which areas have a higher potential to be home to people who will attend matches of their local football club. Variables are included that are synonymous with characteristics of a typical football fan in England, as revealed by the surrounding literature. The creation of this index has firstly been applied to Bromley FC, with the general pattern being that a higher index score means that the population are more likely to attend a football match. The results are then spatially visualised using maps to portray an overall picture of where potential supporters of the football club are located across London. If targeting the areas that display high index scores proves to be a successful marketing strategy, then this index can be modified and utilised by other football clubs in a similar way, evidencing the way that spatial analysis and geodemographics can be applied to the football industry.

Keywords: GEODEMOGRAPHICS, INDEX, FOOTBALL, SPATIAL, MAPINFO, LONDON

Word Count: 10,000

ABBREVIATIONS

FC	Football Club
GIS	Geographic Information System
LSOA	Lower Layer Super Output Area
ONS	Office for National Statistics
SQL	Structured Query Language

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE - INTRODUCTION

1.1 –Background

Geodemographics is derived from the study of spatial information, and provides descriptions of people according to where they live (GKB, 2017), leading to further analysis. Its importance has been growing in numerous sectors throughout recent years, highlighting many benefits to the methods. By analysing people and their characteristics, target groups can be identified that best describe a certain behaviour, based on the findings of small groups of people (Webber and Longley, 2003). This is an important discipline within geography, due to the importance of understanding the changing demands and requirements of the population they are serving. Geodemographics has been chosen as a focus for this research project as it allows spatial analysis to be carried out to achieve the desired aims. Due to the increase in technology and available data, as well as the widespread use of GIS, geodemographics has become "well established as a data-driven analysis tool for marketers" (Leventhal, 2016, p1), and can be applied to several different sectors for specific purposes.

Football is the most popular spectator sport in England (Simmons, 1996), however despite this, football clubs have witnessed a decline in their attendances in recent years, with a decrease of around 430,000 fans throughout the Football League (Gupta, 2013). Statistics such as this indicate the need for football clubs to investigate this, and find ways of ensuring they continue to attract fans to their games, both old and new.

Football clubs work like businesses, meaning they want to be successful on and off the field. If a club has a high number of home supporters, then this can result in better performances for the team (Neave and Wolfson, 2003). For football clubs to achieve success and improve their revenue, they need to understand why fans come to games,

and ensure that their crowds remain as high as possible throughout the season (Gupta, 2013), resulting in higher overall average attendance at their games.

Although all football clubs want to improve their attendances, this is of primary interest to those who have a high potential for growth, especially in the lower leagues beyond the Premier League. Within the local community, a football club can become the main representation of that place, as well as creating a social network of fans (Bale, 2000). Clubs with low attendances rely on regularly attending fans, however many football clubs will have the potential to expand their current supporters base. The way they market themselves needs to be considered carefully on a spatial scale, due to the complex factors that determine who goes to watch football matches and why. If these lower league clubs cannot capture fans every week, then they risk going into administration, which would have a negative impact on the surrounding community, as well as the football industry.

In recognition of the importance of spatial analysis, the link between geodemographics and sport is an important topic that should be further examined. There is a wealth of literature and research on geodemographics, but little on its application to the football industry. Thus, this study will create a geodemographic index that can be applied to any football club in England and Wales to analyse which small-level geographical areas have a high potential of possible supporters of a football club. Bromley FC is being used as the by-product, as a lower-league team who average 1000 home fans per game, but have the potential to attract more on a regular basis.

The purpose of this research is to use the final index to target specific areas for marketing purposes for a football club. The emphasis throughout the project is on the methods and tools used to create the index, so that it can then be used as a framework for other football clubs to replicate for their own area, highlighting the index's reproducibility. As geodemographic methods have not been effectively applied to sport in the past, this research will provide an important contribution in this field.

1.2 – Aims and Objectives

The overarching aim of the project is to create a geodemographic index, using the characteristics of football supporters, to highlight the benefits these methods can have to the football industry and the football clubs.

To achieve this, a number of objectives will be addressed:

- Analyse the surrounding literature on geodemographics to understand their importance, and how it can provide links with football
- Use the literature to determine the characteristics that are synonymous with football attendance to be included within the index
- Collect lifestyle and socio-economic data to create an index that will identify areas of high potential for football support within Greater London
- Apply the index to one club, Bromley FC, to analyse how the index can be used to benefit local football clubs by targeting their marketing to certain areas
- Validate the index by obtaining data on season ticket holders of Bromley FC and identify the extent to which the index predicts correctly
- Analyse the marketing advantages this can have to clubs by increasing support

1.3 – Dissertation Outline

This dissertation will be spilt into 5 chapters. The preceding introduction has outlined the background context for the research project. The following chapter summarises the literature surrounding geodemographics, and the importance of football. Chapter 3 will define the study area, and outline the methods undertaken to create the index. Chapter 4 explains the patterns identified from the index, and highlights the areas more prone to football supporters, before validating the index with real-world data. The final chapter, summarises the findings, and explores further research.

CHAPTER TWO – LITERATURE REVIEW

This chapter will review the literature surrounding the themes that are being explored, and discuss how the creation of the index will provide a benefit to the football industry.

2.1 – What is Geodemographics?

The term geodemographics refers to the analysis of where people live (Sleight, 2004). By clustering areas based on demographic and lifestyle characteristics of the population, patterns can be identified to analyse how people living within a certain area can display similar characteristics (Leventhal, 2016). These methods can be used to predict the behaviour of people, or the information generated can aid the researcher achieve certain aims within their own work.

Geodemographic indexes have been seen to "bring together academic, public policy and private sector stakeholders, applying geographical thinking to tackle questions of social concern" (Harris, 2009, p277). Therefore, by creating indexes and analysing places in terms of their characteristics, policy makers and governments can truly understand local populations' needs, and thus improve quality of life for people. Knowing the location of customers is an important aspect of planning people's requirements, as well as evaluating any threats (Singleton, 2004). As this has become more prominent in recent years, the importance of geodemographics has grown in numerous sectors, as more see the benefits they hold.

From the 1970s, early geodemographics attempted to reduce the complexities involved in human settlements by converting them into simplified technologies, with the purpose of making the handling of data more efficient (Singleton and Spielman, 2014). As capitalism led to the evolution of cites and changed the urban processes and geographical phenomena that were occurring (Reibel, 2011), there was increasing

demand for techniques that classified areas, hence a move towards geodemographic methods in the 1980s. Moreover, the emerging social geography and global forces of capital accumulation caused the reshaping of areas (Burrows, 2013), thus leading to a focus on how specific areas within a region can differ from one another based on the characteristics of people who live there.

Improvements in the methods of collecting data and the inclusion of postcode information produced emerging geodemographic products resulting from 1981 census data (Batey and Brown, 1995). An example of geodemographic classifications created in recent decades include Acorn, where clusters of similar areas are compared to the rest of the nation, allowing companies to understand and target customers through a segmentation of the UK population (Acorn, 2013). Similarly, the MOASAIC classification (Experian, 2017), became the first market specific application to be developed, gaining traction as a private sector marketing tool, before growing and expanding to a wider audience (Singleton and Spielman, 2014)

The census of 2011 presented new opportunities for geodemographics, such as a higher level of geographical consistency for making comparisons between areas and their populations. (Furness and Sleight, 2013). With these advancements, different types of analysis could be utilised within the field of geodemographics. Distance analysis allows each area to be represented as a point in multi-dimensional space that is defined by specific variables, so that the norm can be located and used to distinguish between areas (Voas and Williamson, 2001). In comparison, principal components analysis is a tool where each component represents a weighted combination of variables, allowing an analysis of how they are associated with original variables, making it easy to interpret (Voas and Williamson, 2001). The range of methods highlights the various types of geodemographic tools that can be utilised.

The emergence of a huge volume of digital geo-referenced data throughout the 21st century has meant we can improve our understanding and knowledge of certain

geographic phenomena (Baçao et al, 2004). This data is becoming more accessible to wider parts of the population, and can thus be seen to have a range of purposes in both the private and public sector. The use of big data has meant that there is a change in the volume of data, as well as the velocity, complexity and variety (Leventhal, 2016). Traditional methods have been moving towards real-time geodemographics, integrating diverse and disparate spatial databases (Adnan et al, 2010), meaning areas and their populations can be analysed within time and space. This owes itself in part to an increase in technology, which has allowed a development of existing methods, as well as the introduction of new ones. With geodemographics advancing, it is becoming even more important to ensure it is included within more research.

2.2 – Importance of Neighbourhoods

As geodemographics has developed over the years, so has the importance of neighbourhoods. When geodemographic methods were first introduced, computer technology had developed enough to allow analysis of census data (Baker et al, 1979). The introduction of the postcode within a classification of areas was vital, since social space can be highly constructed (Burrows, 2008), and specific areas can be located by using postcodes. This, along with new forms of technology, are being utilised to construct ideas of more diverse neighbourhoods (Burrows, 2008), allowing policy makers and governments to tailor specific requirements to certain areas.

Using geodemographic methods allows researchers and the public to see the relationship between social geography and the digital representation of it (Burrows, 2008), linking to the neighbourhoods that they are a part of. Urban space can be viewed as a landscape for social and economic opportunity, as well as being a site for discrimination (Kearns and Parkinson, 2001), thus it is important to ensure it is always considered within policy making and decision making.

Despite the increasing focus on neighbourhoods, the degree of how important they are remains dependent on where you are (Kearns and Parkinson, 2001), emphasising the status of space within the work of geodemographics. Although geographic location and nearness is often seen to bring about the idea of neighbourhoods, people sharing space does not necessarily create a place (Kearns and Parkinson, 2001). This demonstrates the importance of characteristics of people within a certain location, that can then be used to truly determine how neighbourhoods can be clustered to allow spatial analysis.

2.3 – Uses of Geodemographics

Since the emergence of geodemographics within the social sciences, it has been utilised for a range of purposes due to the success of its applications in the 1980s (Adnan et al, 2010). The primary use of indexes that are created are to provide data about attributes in relation to a set scale for numerous areas (Donnay and Unwin, 2001). The overall theme of this research is that small areas that are close to each other often share similar characteristics, compared to those that are further away. Thus, areas can be classified to see more explicit differences between the populations. Throughout much of the literature, evidence shows that the main use of geodemographics has been in resource allocation, and thus is an important tool in planning and marketing strategies.

In the education sector, Butler and Hamnett examined the increasing way that social segregation is finding its way into the education system. These patterns are then compared to the educational provision and attainment of school pupils to see what causes the variation in the results produced (Butler and Hamnett, 2007). By using geodemographics to present the existence of class and ethnicity based residential segregation, a new geography of education emerges (Butler and Hammett, 2007). Similarly, using the Mosaic classification system (Experian, 2017), in 2007 Webber and Butler analysed school attainment levels to see patterns between neighbourhoods, concluding that the more affluent the area, the higher the average GCSE score. Policy

makers can examine what factors affect the performance of school pupils, and effectively manage these differences through responses that are targeted towards to right areas and pupils, to improve the attainment where most needed.

In the public sector, organisations have been looking to use consumer insight to understand the needs and behaviour of target populations, including within health. The growing pressure on health resources has meant that public health campaigns have been increasingly turning to the use of geodemographic tools (Peterson et al, 2009). General attributes of an area are summarised, and then linked to health variables to inform health service planners at a local level which areas resources should be allocated to (Abbas et al, 2009). This profiling provides the potential for benchmarking against similar populations. As health care is a product, the spatial distribution is of high importance, thus governments need to look at the relationship between socio-economic deprivation and poor health of an area (Abbas et al, 2009).

Crime can be analysed using geodemographics by exploring the characteristics of high crime neighbourhoods, and why crimes occur where they do (Chainey, 2005), aiding policy makers and police in targeting certain areas. The methods used support intelligence-led and evidence-based approaches to better inform (Chainey, 2005), and therefore can be linked to other information that is collected about crime. Similarly, road traffic accident risk provides growing evidence of the impact of geodemographics in public sector interventions (Anderson, 2010). Geodemographics gives researchers the potential to link socioeconomic variables to risk, allowing an exploration of what areas are most at risk (Anderson, 2010). Those without higher education were found to be at 60% greater risk of accidents (Anderson, 2010), providing an example of geodemographics offering a worthwhile solution to a complex problem by considering multiple factors, signposting where policy should be directed.

In the private sector, using geodemographic information for marketing purposes has been vital in recent years, where companies can obtain a better understanding of their

customer base to target campaigns to appropriate sectors of the population (Leventhal, 2016). The idea of targeting customers for business purposes provides more accurate sampling frames for data to be used by companies (Sleight, 2004), as well as emphasising plausible variations in social geography situations (Birkin, 1995). If consumption of a product is skewed towards a certain group, geotargeting can be used to improve sales by making the marketing specific (Inman et al, 2004). Within the UK, analysis of customers and geodemograhics are basically synonymous (Birkin, 1995), showing how valuable geodemographics can be for companies.

Retail marketing can also benefit from geodemographics. The population of each group of people within a specified distance can be calculated using GIS (Birkin, 1995). This will enable retailers to understand what types of people they will be attracting nearby and infer information about core customers (Harris, 2003), and thus target their products towards these demographics. Consumers now expect brands to deliver personalised experiences (Furness and Sleight, 2013), and therefore retailers must use more precise information to deliver these requirements for an ever-growing consumer base.

Other examples within the private sector have been identified within the literature, such as its uses in tourism, a diverse and complex industry (Kamarulzaman, 2010). The internet has provided an opportunity to reach new markets and allow more people to utilise these services. Geodemographic tools have been applied to better understand demand, highlighting that most customers of travel e-shoppers are females aged 25-44 with higher education, as well as young couples or those with considerations of children (Kamarulzaman, 2010). From this research and analysis, governments can use the information to collaborate with national companies to make technology advancements such as this accessible to more people.

2.4 – Benefits to Geodemographics

Analysis of the literature highlights many benefits to creating a geodemographic index. They can pinpoint exact locations and map them to show spatial patterns by including postcodes (Leventhal, 2016), allowing analysis of complete coverage of a specific area. Moreover, the techniques provide a rich description of distinguishing features of types of areas (Batey and Brown, 1995) that can be considered. As numerous variables are used within an index, they can unveil relationships between large numbers of characteristics rather than one specific feature (Batey and Brown, 1995). The urban fabric of cities can vary and be vastly different (Burrows, 2008), hence geodemographics allows a better understanding of the urban dynamics of places to be used by a variety of people and organisations. In a complex world, geodemographics has the potential to highlight problems of society (Baker et al, 1979), such as areas of high stress, and then policy recommendations can revolve around those areas.

The highly practical examples of geodemographic use since the 1980s have contributed to the success of the methods (Baker at al,1979), as they show what can be achieved. Once created, they are easily accessible for people, allowing the audience to paint a contextual picture about different areas (Leventhal, 2016). More recent uses, such as with social media data, can give an insight into the geography of individuals, rather than a whole group of people (Adnan, Lansley and Longley, 2012). This can improve the accuracy of analysis that is obtained from the creation of an index. Furthermore, the use of microdata and spatial clustering avoids difficulties arising from local district data (Reibel, 2011), meaning analysis can be carried out on a larger scale. Open data can become part of geodemographic tools to improve public sector efficiency and allow for better delivery of information (Furness and Sleight, 2013).

2.5 – Football and Sport

Geodemographics has been applied to a diverse range of applications, but has not yet been incorporated into the sporting sector, and thus there is a gap in the literature that needs to be explored. Football is a significant part of the culture & economy of the UK (Sport England, 2013), and with sport in the UK contributing to over 450,000 jobs within a £20 billion industry (Critchlow, 2015), there is further potential for the growth of the market. However, because people are spending less due to recent recessions (Wood and Allen, 2016), there is the possibility this may translate into less people spending their disposable income on leisure activities such as football. To ensure the sport economy remains successful, there needs to be efforts to ensure that football clubs continue to prosper within their local community.

The English Premier League is the largest football league by value in Europe, with 33% of the average turnover being derived from TV payments (Ofcom, 2005). An investigation found that although the number of live matches shown on TV since 1992 has increased from 60 to 138 in 2005, the average audience has stayed the same (Ofcom, 2005). If people are not watching football on the TV, even though a quarter of the population are fans of the Premier League (Ofcom, 2005), then there is the possibility that a certain demographic of people may choose to go and watch their local football club if they enjoy watching matches.

Although geodemographics has not been used to cluster areas and analyse the likelihood of people being football supporters, there has been some research that has involved the analysis of data. It is possible to look at gender, age, ethnicity and income in terms of supporters, to define what type of people support teams (Thompson, 2014). However, within this research, only one variable has been focused on at a time, whereas creating an index allows multiple factors to be considered. Furthermore, sport is an example of a sector that is highly personalised, however one that fluctuates and

is unpredictable (Tapp and Clowes, 2000). Due to this, there needs to be a focus on the consumers, which is possible through a quantitative approach such as building a model.

One of the main reasons for carrying out this research and creating the index is because it has not been done before. The value of geodemographics as a discipline can be applied to football in the same way it has been applied to other sectors, allowing football clubs to benefit from the index by attracting new fans. By applying specific marketing for different areas (Leventhal, 2016), those areas that are most prone to football fans can be focused on by a club so that their supporters base can be increased.

The literature reviewed has focused on the way that geodemographics has emerged, and the way it can be utilised to create benefits. Research into football support has identified possible characteristics of a typical supporter, however has highlighted the fact that sport has seldom been used within this field. The next chapter will outline the way in which the index is created, and the data supplied for each variable selected.

CHAPTER THREE – METHODOLOGY

This chapter will explain and justify the methods and stages that have been undertaken to create the index and provide spatial analysis for the football industry.

3.1 – Study Area

Greater London is an urban conurbation with a population of 8.54 million in 2014, spread across 32 boroughs, as seen in figure 3.1, and has a projected growth of 2.35 million by 2041 (GLA Intelligence 2016). London is a dynamic city encompassing an area of 1,623km^2 (Pointer, 2005), and has been chosen as the study location for this research because of the wealth of information that can be gathered about the area. Moreover, across the region there are a variety of people that live within its boundaries, who display a diverse range of characteristics to be analysed.

Figure 3.1 – Map showing boundaries of London Boroughs
Source = Office for National Statistics, 2016b

To allow analysis, census boundaries were required rather than postal areas, since extensive census data will be used to create the index. Furthermore, the modifiable areal unit problem needs to be considered. Results are likely to be influenced by the scale of units chosen, therefore it is better to carry out analysis at a smaller scale (Leventhal, 2016), meaning it can be aggregated if necessary. Thus, Lower Layer Super Output Areas (LSOAs) have been chosen as the appropriate scale of units to be used, which is explained further in table 3.1, allowing local level analysis to be conducted.

Table 3.1 – Geography of LSOAs, Source: Adapted from (ONS, 2016c)

Minimum Population	Maximum Population	Minimum Households	Maximum Households	2011 Total in London
1000	3000	400	1200	4835

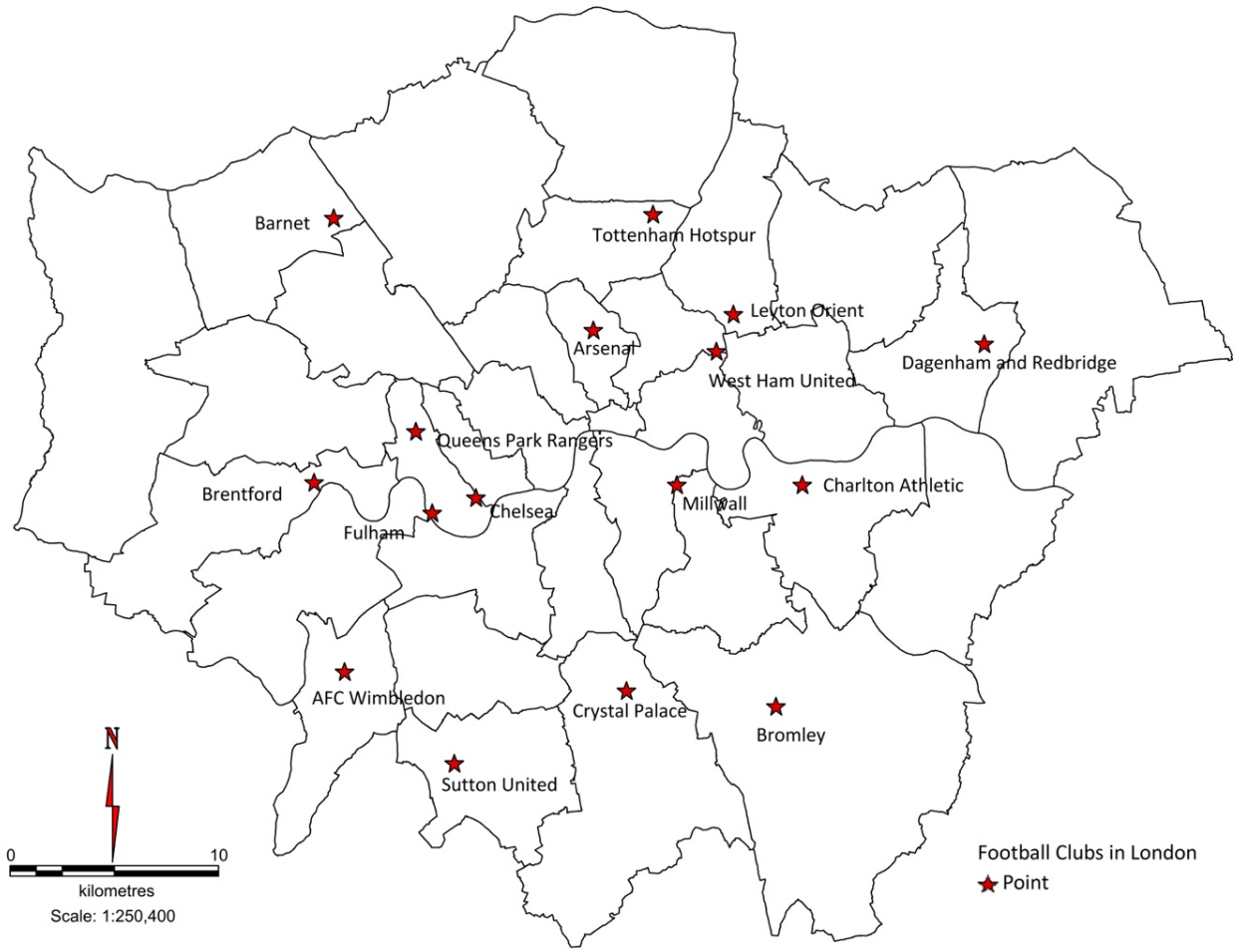
Greater London is home to numerous football clubs, both professional and semi-professional, ranging in popularity and ability. For this index, clubs that play in the top five tiers of English football are included. These leagues produced 16 clubs in London, as of the beginning of the 2016/17 season, displayed in table 3.2 and figure 3.2.

Table 3.2 – Football Clubs in Greater London

Football Club	Stadium Name	Postcode
AFC Wimbledon	Kingsmeadow	KT1 3PB
Arsenal	Emirates Stadium	N7 7AJ
Barnet	The Hive Stadium	HA8 6AG
Brentford	Griffin Park	TW8 0NT
Bromley	Hayes Lane	BR2 9EF
Charlton Athletic	The Valley	SE7 8BL
Chelsea	Stamford Bridge	SW6 1HS
Crystal Palace	Selhurst Park	SE25 6PU
Dagenham and Redbridge	Victoria Road	RM10 7XL
Fulham	Craven Cottage	SW6 6HH

Leyton Orient	Matchroom Stadium	E10 5NF
Millwall	The Den	SE16 3LN
Queens Park Rangers	Loftus Road	W12 7PJ
Sutton United	Gander Green Lane	SM1 2EY
Tottenham Hotspur	White Hart Lane	N17 0AP
West Ham United	Olympic Stadium	E9 5LN

Figure 3.2 – Map showing location of Football Clubs within Greater London



London is home to some of the most successful football clubs in the world, as well as the headquarters of the Football Association (Hidden London, 2005). However, this index will be tailored towards a lesser-known club. Bromley FC was formed in 1892 (Bromley FC, 2017), and in the 2015/16 football season they played in the highest tier of non-league football, the National League, for the first time in their history. As previously discussed, the aim of this index will be to aid Bromley FC in targeting areas that have the potential for having a high number of people who will attend their matches, thus increasing their attendances. Despite having regular attendances of around 1000 for home games, there is the opportunity to attract new customers within a crowded London market, matching their growth aspirations, making it an ideal football club for this research. Another benefit to using Bromley FC as the by-product is due to the availability of season ticket holder data that can be used in validation.

3.2 – Determining Variables for the Index

When a geodemographic index is created, an analysis of the literature is vital to confirm what variables are going to be used (Voas and Williamson, 2009), and what features best distinguish the different areas. Variables included in this index are socio-economic and demographic characteristics that are synonymous of a typical football fan. They have been chosen based on what is deemed to be the most important for the index to be able to identify those who will attend football games. Justification for each variable is shown in figure 3.3, however assumptions are inevitable in these decisions.

Table 3.3 – Justification for Index Variables

Variable	Justification
Age 18-44	Out of 2000 people interviewed, interest in football was most common amongst those aged 23-34 (Ofcom, 2005) Including ages either side of this as it is a census group that includes all these ages, with some flexibility either side
Male	Out of 2000 people interviewed, interest in football was most common amongst men (Ofcom, 2005) 77% of 5.2 million fans in 2011/2012 were male (White, 2013)
No Religion	Typical characteristics of football fans (YouGov, 2016)
White British	Typical characteristics of football fans (YouGov, 2016) 8% of Premier League fans from ethnic minorities (BBC News, 2010)
Income	Out of 2000 people interviewed, interest in football was most common amongst those in highest social group, (Ofcom, 2005), thus those who have a high level of disposable income Need disposable income to purchase season tickets, thus must be a certain level of affluence to attend football matches Economic variables are highly significant to determine football match attendance (Garcia and Rodriguez, 2002)
Access to Car	Need a mode of transport in order to attend the games Importance of car use, 43% travel to football matches by car (Campaign for Better Transport, 2013)
Distance to Club	Importance of distance when considering the changing dynamics of football attendances (Bale, 2000), what influences support Need to be close by football ground to have a strong connection with the identity of the club (Supporters Direct, 2010)
Distance to Other Clubs	Importance of distance when considering the changing dynamics of football attendances (Bale, 2000) If there is too much competition around, less likely to attend

Prior research highlighted other variables that would be appropriate for this purpose. Those who watch programmes such as Match of the Day are more likely to have an interest in football (YouGov, 2016), and thus this could have been an additional variable. However, it is not viably possible to collect this data, as the information is not easily accessible for London LSOAs. Other data that could have proved useful is the impact that televised Premier League matches have on local attendances (Buraimo, 2008), however this is not appropriate within the time scale. If this research was taken forward, then this could be explored in more depth for the population.

Another way of analysing whether people are likely to attend games is the success of the football club, such as the average number of points per game, or how attractive a match appears to spectators (Buraimo et al, 2009). However, the index being built is focusing on the characteristics of the population rather than the football club themselves, and thus this was not included. Moreover, a variable analysing the population within 5 miles of a club's stadium (Buraimo et al, 2009) is also relevant, as without a large catchment area, there would be limited opportunities to increase attendances. However, because the study area is London, this was deemed unnecessary, due to the high population density of the area.

3.3 – Data Collection

The first dataset that was collected was the boundary data for Greater London. The boundary data selector tool provided by the census allows the LSOAs within the area to be converted onto a file that can then be translated onto a mapping programme.

The census is the main source of data used within geodemographic work (Leventhal, 2016), as it is easily accessible and has all the required data for the boundaries used. Age, gender, religion, and ethnicity were all obtained from the 2011 census. Infuse allowed census data to be downloaded for individual LSOAs, which is then converted

onto an Excel spreadsheet. This stage was repeated for each variable that could be found on the census, alongside the code and name of each LSOA in Greater London.
For some of the data collected, modifications were required to ensure the data matched the chosen variables. When the age data was extracted from the census, there were several different categories that were obtained, and these had to be combined to create the 18-44 age group. Similarly, the car accessibility data from Infuse was divided into different number of cars. However, for the index, the focus is on whether a household has access to a car, and thus each separate category was combined to create the appropriate variable for the index.

Although it is the primary data source for geodemographics, the census does not include lifestyle data that is often required in an index. The London Data Store is an online portal that allows free access to over 600 open data sets about various aspects of Greater London. The gross disposable income variable data was obtained from this database, and is important for the index because household income offers a better indicator of geodemographic diversity at a finer scale (Harris, 2003). However, because this index is using a low-level geography of LSOAs, it is more difficult to find data at this level. The disposable income data found was for each London borough, and thus each value was translated to the appropriate LSOA.

The two distance variables differ from the other characteristics in terms of their method of data collection. Each football club involved in this research is allocated coordinates that allow them to be mapped onto the LSOA boundaries within MapInfo. For the first distance variable, only Bromley FC is mapped onto the LSOA boundaries. Following this, every individual LSOA is given specific coordinates in the centre of the LSOA boundary, shown in figure 3.3, that will be used to calculate distance. If a LSOA is close-by to the targeted club, in this case Bromley FC, then the population will be more likely to attend a match. The distance calculator tool is used to provide the distance from each LSOA to the football club, with the results being exported as an Excel file to then be added to the index as another variable, contributing to the overall score.

This process needs to be repeated for the distance to the other football clubs. If a LSOA is close by to another club, then the population will be less likely to attend a match, and thus will affect their index score. The same method is applied for this variable, with the rest of the football clubs being mapped instead of Bromley FC. The distance calculator is again used to calculate the relative distance to the nearest football club for each LSOA, whose spatial results are shown in figure 3.4, creating a list of values for each area that can be incorporated into the index as an additional factor.

If this index is replicated for other football clubs in London, the variables that would then require modification would be the distance ones, using the methods that were applied to clubs in Greater London. Then the relevant information for each variable would need to be obtained from online census data to match the football club being used. The different characteristics of the variables are displayed in table 3.4.

Table 3.4 – Variables Included in The Index

Variable	Source	Year	Scale of Units	Type
Age	Census	2011	LSOA	Individual
Male	Census	2011	LSOA	Individual
No Religion	Census	2011	LSOA	Individual
White British	Census	2011	LSOA	Individual
Income	London Data Store	2016	Borough	Household
Access to Car	Census	2011	LSOA	Household
Distance to Club	Club Websites Geoconvert	2016 2011	LSOA	Distance
Distance to Other Clubs	Club Websites Geoconvert	2016 2011	LSOA	Distance

Figure 3.3 – Map Showing Centre of Each LSOA in London
Source = Office for National Statistics, 2016b

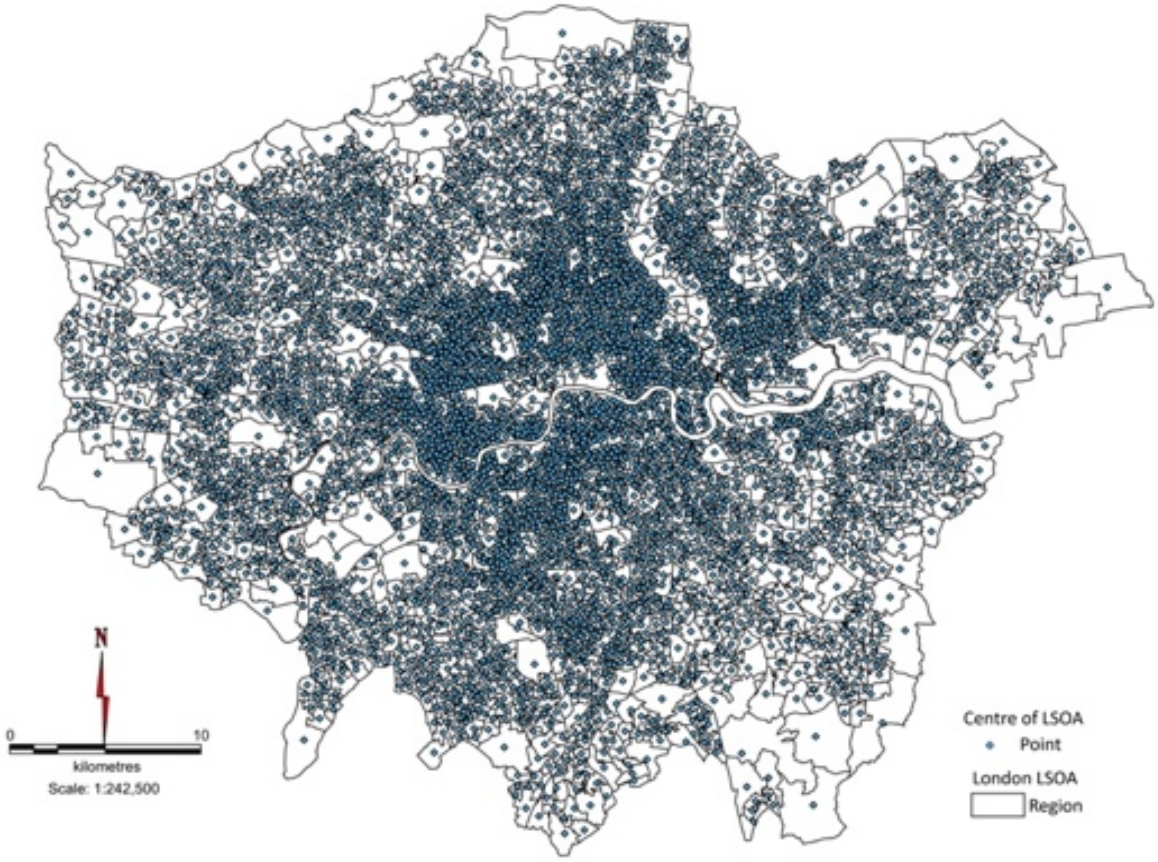
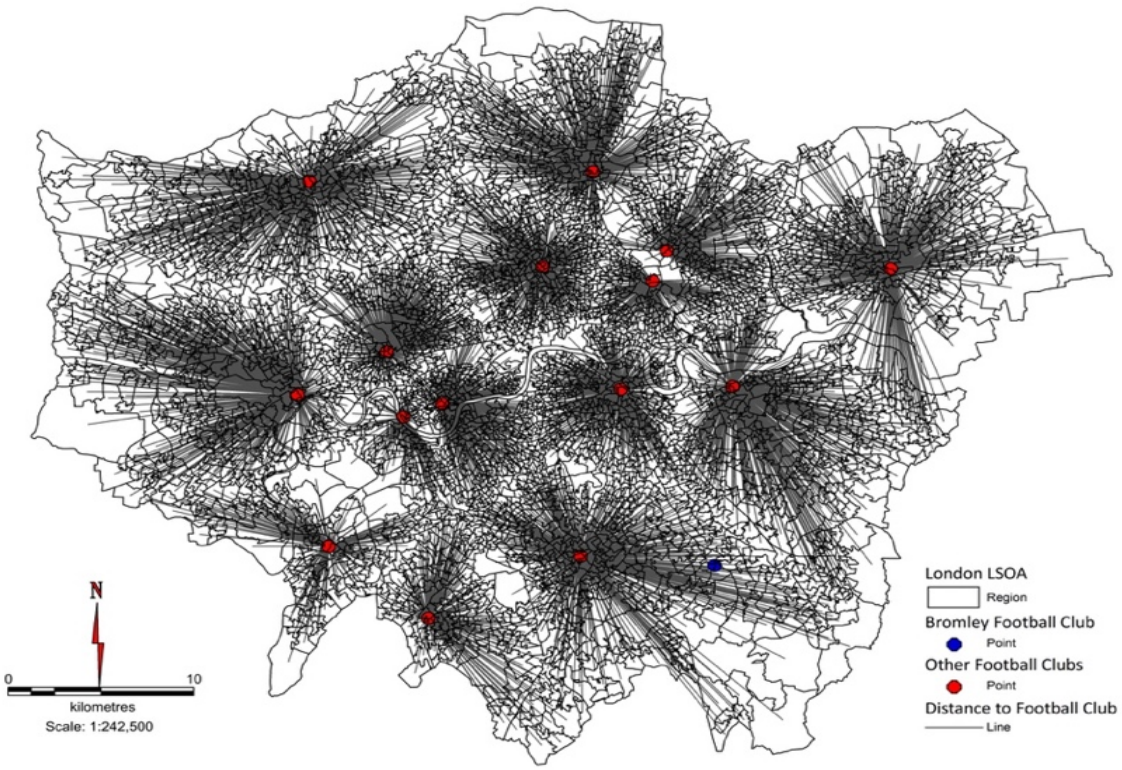


Figure 3.4 – Map Showing Distances from Each LSOA to Nearest Football Club
Source = Office for National Statistics, 2016b



3.4 – Creation of the Index

The method of creating a geodemographic index varies depending on the purpose, as they should be specific to the users, ensuring they are accessible and easy to understand (Singleton and Longley, 2009). Despite differences, the main goal of any index is to organise data into meaningful structures (Bacao et al, 2004). This index will be created using the multivariate classification technique, which involves using numerical methods to group areas together (Reibel, 2011). Characteristics of people in LSOAs in London will be combined to rank areas in their likelihood of having a high majority of potential football supporters of a specific football club, Bromley FC. Instead of mapping areas and then looking at the characteristics of the areas, creating an index through these stages makes it transferable across the UK. If the index proves successful, then it can be utilised by other football clubs in London, as well as across the country, through simple modifications. Bromley FC is being used a case study, allowing the index to be re-applied across the UK. The stages in creating this index can be utilised as a framework for other football clubs, using freely available data from the census.

Standardisation

After the data was obtained, the raw data was converted into percentages, allowing a standardisation of the data so that all the values for every LSOA run on a scale between 1 to 100, creating a percentage. To achieve this, each census value was divided by the total number of people within that LSOA, and then multiplied by 100. An exception to this was the access to car variable, which was instead divided by total number of households, due to the differences in the way the data is presented within the census. There are variables per person and variables per households, and thus they should be treated differently in the standardisation process. The disposable income variable did not need modification, as when downloaded it had already been converted to a value per household in each LSOA. Similarly, the distance calculations did not need

standardising, as this data does not need to be presented in percentages in the same way as demographic variables due to the fact there is no obvious denominator.

Multicollinearity

When variables are included in the index, if there are two that exhibit similar results, then it may not be necessary to include them both. The data analysis technique of correlation can be carried out in Excel to determine the similarity of the chosen variables. If high correlations are observed of a value above the threshold of 0.7, then one characteristic can be excluded, as both variables would be contributing to the same index, and thus one would not be needed. The decision to have a threshold of 0.7 is subjective, and can be altered depending on the opinion of the index creator. When this stage was executed, no high correlations were observed, as can be identified in table 3.5, and thus the decision was made to keep all variables. However, in the work of geodemographic index creation, it is an important aspect to consider.

Table 3.5 – Results of the Multicollinearity Process

	Age 18-44 Percent	***Male Percent***	***No Religion Percent***	***White British Percent***	***Disposable Income (£ per head 2014)***	***Access to Car Percent***	***Distance to Football Club (miles)***	***Distance to Other Clubs (miles)***
Age 18-44 Percent	1							
Male Percent	0.425182002	1						
No Religion Percent	0.401654999	-0.068086174	1					
White British Percent	-0.388526009	-0.31709827	0.482543916	1				
Disposable Income (£ per head 2014)	0.251605241	-0.002488239	0.226593073	-0.035760922	1			
Access to Car Percent	-0.694206128	-0.197230948	-0.296014153	0.514183567	-0.310524719	1		
Distance to Football Club (miles)	-0.097882805	0.032880702	-0.353992802	-0.190466015	0.032508198	0.209780763	1	
Distance to Other Clubs	-0.369728766	-0.075976339	-0.159900801	0.30847313	-0.199695006	0.453026411	0.164507516	1

Normalisation

The next stage was further standardisation of the data through employing normalisation to linearly re-scale each variable to the same scale of between 0 and 1. This was carried out for all variables, including the two distance calculations. Although the conversion to percentages adopted some degree of standardisation, the variable percentages will adopt different ranges, as some may go from 20 to 50% rather than ranging from 0 to 100%. The process of normalisation allows a countering of this by ensuring that the variation of the variables is captured equally, where the lowest value for each variable is 0, and the highest is 1, with every other value fitting in proportion. The equation inputted into Excel is as follows, and an example is displayed in table 3.6:

$$x_{norm} = (x_{raw} - min_1)/(max_1 - min_1)$$

x_{raw} *is the number being rescale*d
min_1*is the minimum number of the whole column*
max_1 *is the maximum number of the whole column*

Table 3.6 – Example of the Normalisation Process

LSOA	Normalisation Equation	Percentage Value	Normalised Value
Sutton 025E	=(A4668-MIN(A:A))/(MAX(A:A)-MIN(A:A))	20.71232877	0
City of London 001A	=(A3-MIN(A:A))/(MAX(A:A)-MIN(A:A))	40.13651877	0.279838462
Lewisham 00EF	=(A4739-MIN(A:A))/(MAX(A:A)-MIN(A:A))	90.12448133	1

Polarity

Once the data was normalised, the next stage was to determine the polarity of the data, where the direction of each data set is considered. Each variable included in the index needs to have values that mean that a high score is a positive aspect of the LSOA, hence a score of 1 for each variable will be desirable for an area to have a high potential for football supporters. Before this stage, every variable displayed this scale of 1 being the most preferable apart from the distance to football club variable. Due to this, the data was 'flipped' to run in the same direction, so that a value of 1 would be the area closest to Bromley FC, and thus a positive aspect of the index. This was achieved by subtracting the normalised value from 1, meaning the scale for each variable is reversed, and thus every variable in the index is going in the same direction. With the other distance variable, this stage was not necessary, because a high value means an area is further from other football clubs, thus are more likely to attend a football match.

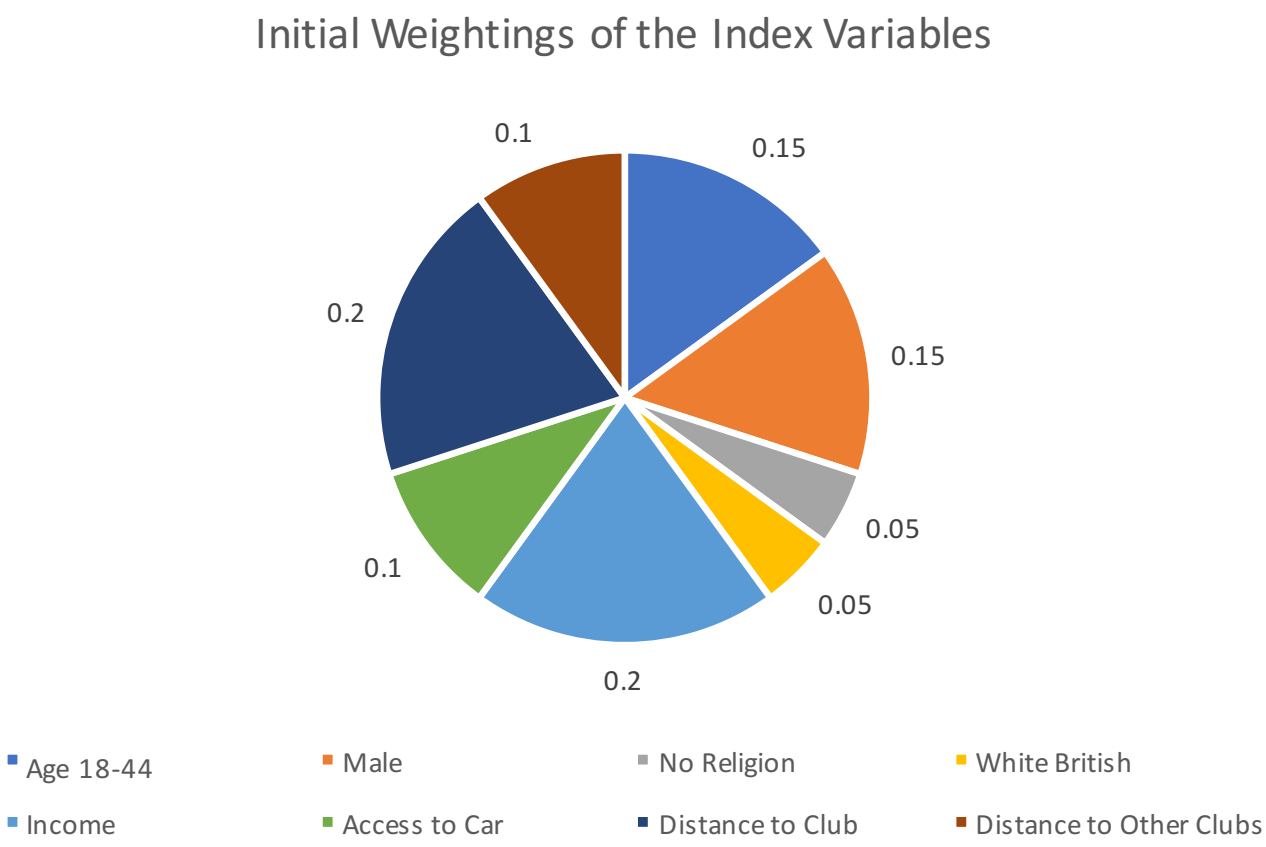
Although polarisation of the data is a necessary stage for this index, it is important to highlight that it is a decision based method that relies on the judgement of the creator of the index. Thus, it can often demonstrate the issue of subjectivity within the work of geodemographics that is important to consider when evaluating an index.

Weighting

Once the data is processed, decisions need to be made on what weighting to give each variable. When electing the weightings, a reflection on the literature highlights that income and distance are the most important factors. On the other hand, no religion and white British are the least important, and it can be argued that they are the most generalised. The results of the multicollinearity stage can be used to see if there are two variables that although are both important enough to be included in the index, could have smaller weightings. Age and access to car have the strongest correlation

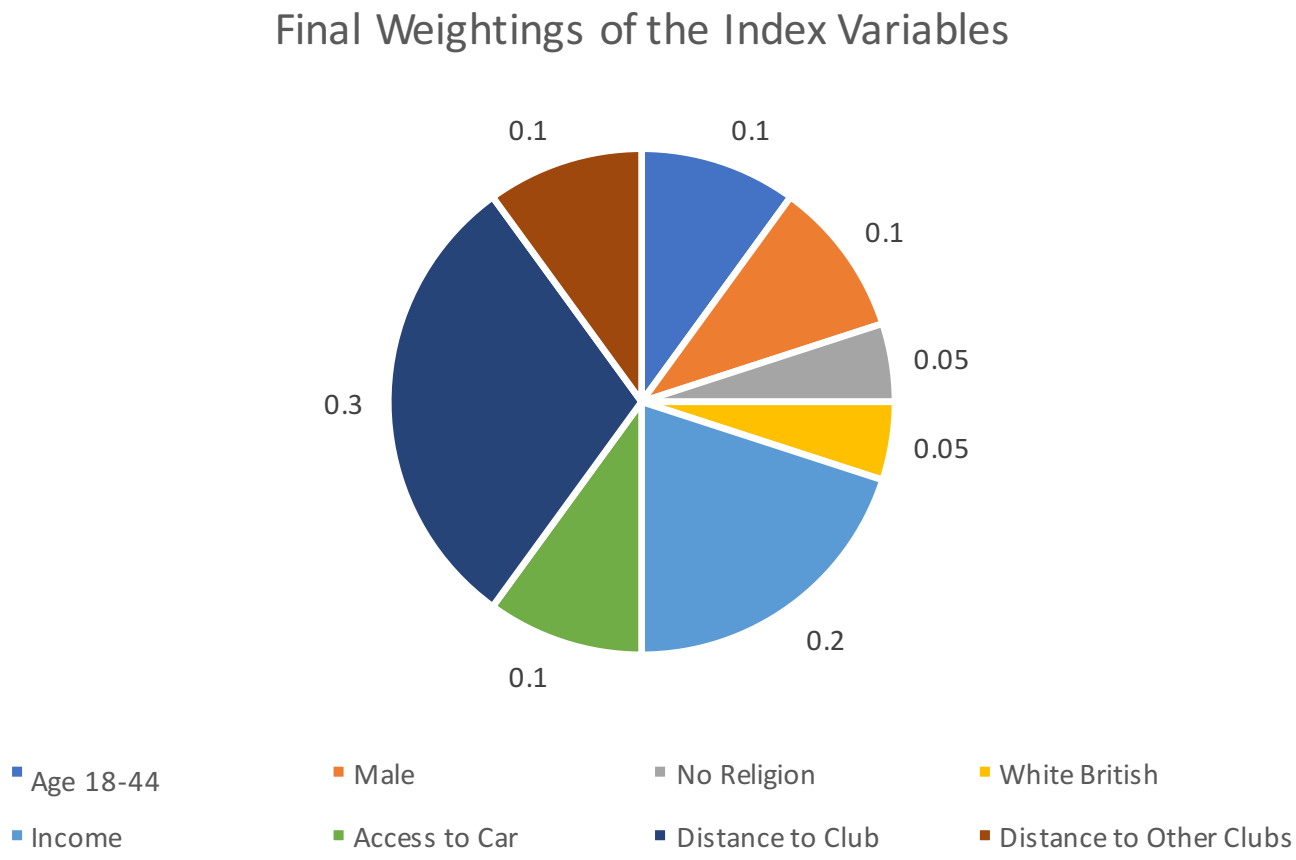
between any of the variables, and thus can have lower weightings than income and distance. The initial weightings are shown in figure 3.5.

Figure 3.5 – Chart Demonstrating Initial Weighting of the Index Variables



After the initial decision on the weightings of each variable, modifications were made for some of the variables, presented in figure 3.6, highlighting that the weighting stage is an iterative process of trial and error. This was done to allow a greater representation of the distance variable, as much of the prior research stated the importance of this (Bale, 2000). Different weightings would give different index results, and thus it is vital to have a firm justification behind the weightings that are chosen. If the index is replicated in another study area, alternative weightings may be more appropriate, for example distance being less represented of a less densely populated area.

Figure 3.6– Chart Demonstrating Final Weighting of the Index Variables



When the final weightings were confirmed, each normalised value in the index is multiplied by the weighting given to that variable. For each LSOA, the weightings of all the variables should combine to equal 100% of the score. However, the decision on weightings for the variables is another example of where subjectivity is involved. The importance of each variable is based on the beliefs of the person creating the index, and their interpretation of the literature, thus if the index is modified in future work then the weightings may vary.

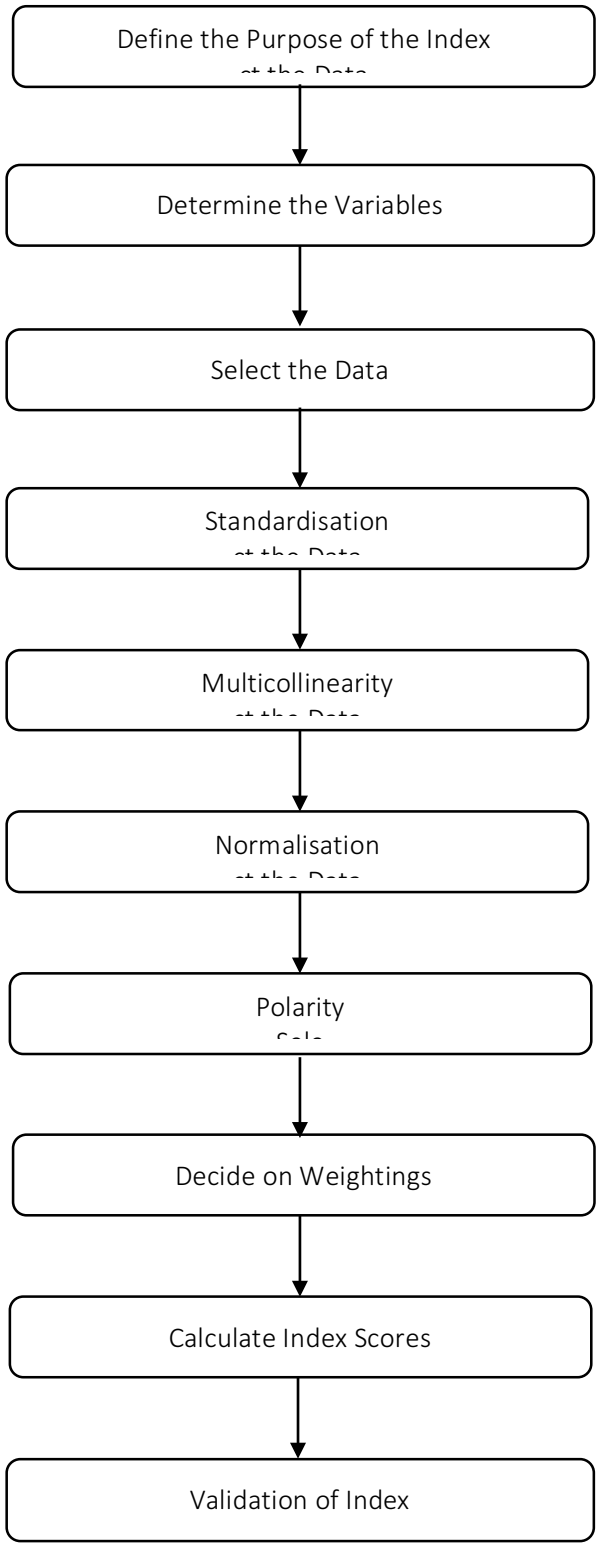
Creating the Index Score

The final stage of the index creation was to calculate the final score for each LSOA. To achieve this, each weighted score for all the variable values assigned to each area was added together. To make the results easier to explore and understand, each final index value has been multiplied by 100 before analysis.

When the final index value is formed, each LSOA is given a number between 1 and 0 based on their score, which represents the potential of that area having a high number of football supporters, and can be used to compare LSOAs with each other.

The overall process of creating the index can be seen in overview in figure 3.7, defining each stage that must be carried out to achieve the final scores.

Figure 3.7– Flowchart Highlighting the Stages of Creating the Index



3.5 – Creation of the Maps

The results of the index can be spatially analysed to demonstrate the areas where football supporters are most likely to live, and identify key patterns. This will allow users of the index to visually identify key spatial patterns created using geodemographics through a combination of GIS functions and the index (Harris et al, 2005). The GIS programme MapInfo allows the boundary data to be combined with the results of the index for each London LSOA, creating choropleth maps displaying the index scores.

This section has explored the study area that is being focused on within this research, and analysed the variables used to construct an index, before showing the stages completed to create the index. The following chapter will show the spatial patterns created by the index, and analyse the results that have been found.

CHAPTER FOUR – RESULTS & ANALYSIS

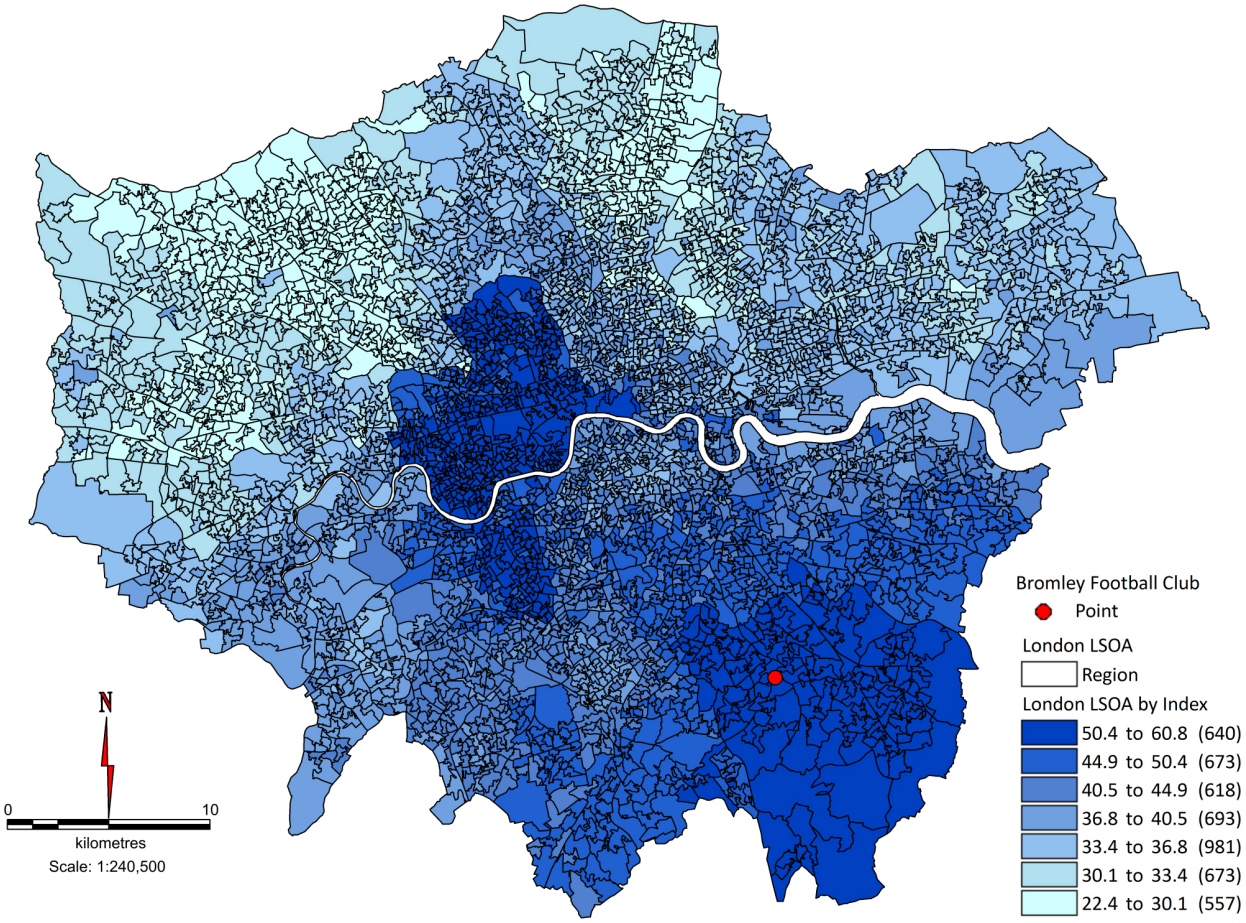
This chapter will present the results of the geodemogrpahic index, and analyse it in relation to the chosen by-product for this research project, Bromley FC.

4.1 – Analysis of the Index

The creation of this geodemographic index allows a comparison of LSOAs in Greater London, with the purpose of identifying specific areas that have the highest percentage of possible supporters of Bromley FC. To ascertain the areas that will be most suitable for the football club to target, mapping of the final index scores permits a spatial evaluation of the whole area. The higher the index score assigned to a LSOA, the more likely it is that people living within that area will attend a Bromley FC match, due to the socio-economic and demographic characteristics of the population.

The process of creating maps involves many decisions to be made that may influence the way the results are presented. One of these is what ranges to use for the choropleth mapping. To spatially visualise the index ranges, the ranges have been split by natural break when the results are mapped in MapInfo. This was the chosen way of separating the ranges as it is the most natural way of categorising the index scores, and thus is the most appropriate for this purpose, as it ensures there are similar number of LSOAs within each bracket of index scores, and thus shows a standard spatial distribution of the index. When translated into a choropleth map, it shows the way that various areas have different index scores, and thus allows further analysis to be carried out.

Figure 4.1 – Choropleth Map Showing the Spatial Pattern of the Final Index Scores
Source = Office for National Statistics, 2016b



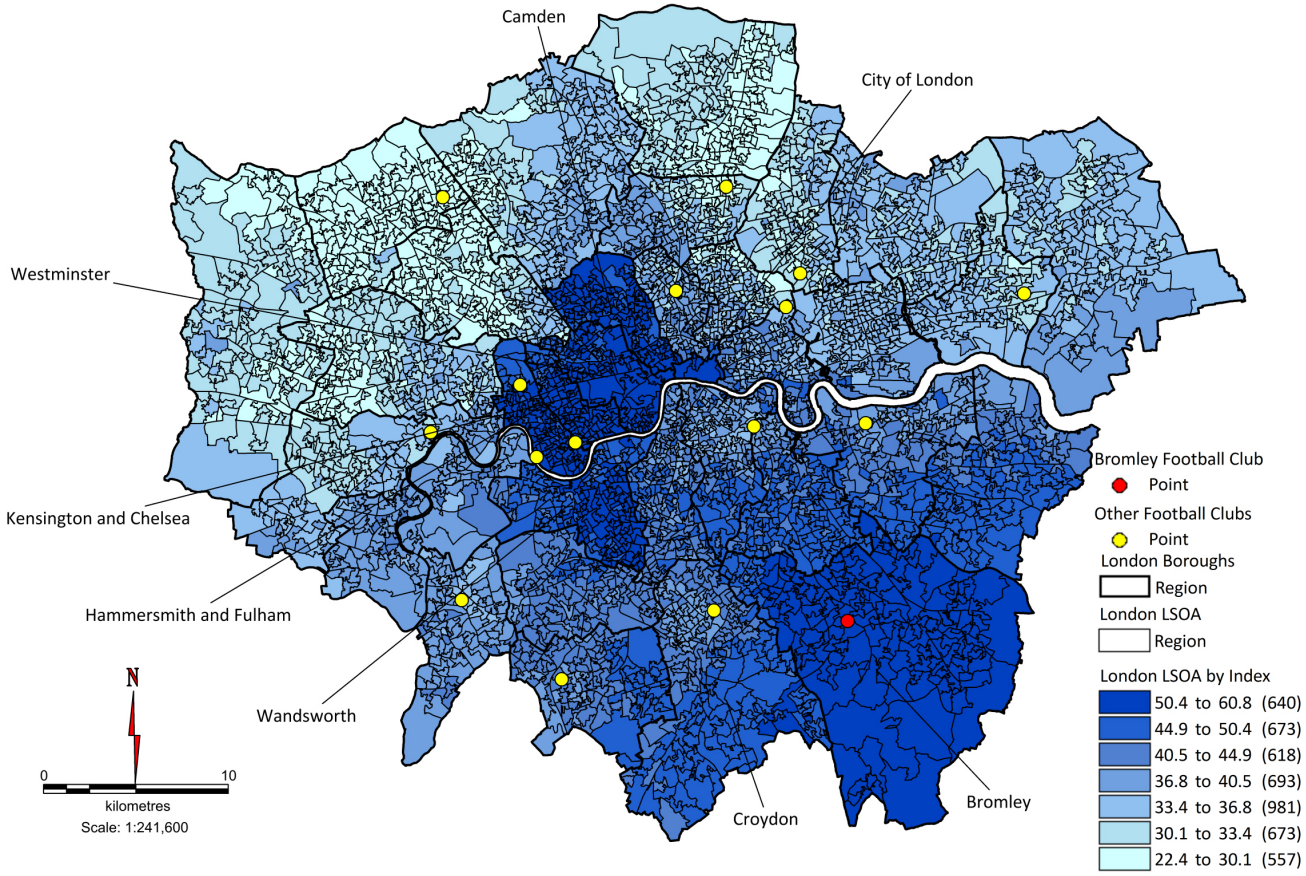
The darkest blue areas on the map in figure 4.1 indicate a LSOA that has a higher potential for supporters of Bromley FC, due to a greater index score than lighter areas. An obvious area of high index scores is in the borough of Bromley. The highest weighting within the index was given to distance to football club, meaning that areas surrounding the football club were going to score highly, and this is reflected in the map. The findings of this reflect the literature explicitly stating distance as a vital factor in determining football match attendance (Bale, 2000). Thus, one marketing strategy for Bromley FC should be to look at the areas near the football club to promote the attraction of local football within the area, and improve their attendances.

However, to allow a detailed targeting programme resulting from this index, specific areas need to be considered in more detail through Local-level analysis of LSOAs. The top 20 ranked areas are shown in table 4.1. Several Bromley LSOAs are included, and as previously mentioned this is expected. Other inclusions are those that are not easily identified before the creation of the index, and thus require more analysis.

Table 4.1 – Top 20 Ranked LSOAs Within The Geodemographic Index

GEO_CODE	GEO_LABEL	Index Score
E01004736	Westminster 018C	60.735
E01032739	City of London 001F	60.395
E01032740	City of London 001G	60.027
E01002838	Kensington and Chelsea 020A	59.966
E01004688	Westminster 019B	59.380
E01000851	Camden 026B	59.134
E01000668	Bromley 020B	59.097
E01004622	Wandsworth 024D	58.976
E01000679	Bromley 020E	58.747
E01000648	Bromley 016C	58.690
E01004763	Westminster 013B	58.588
E01001913	Hammersmith and Fulham 014C	58.481
E01000823	Bromley 008E	58.405
E01000680	Bromley 037A	58.242
E01004733	Westminster 020C	58.229
E01000647	Bromley 016B	58.158
E01000697	Bromley 007F	58.148
E01001937	Hammersmith and Fulham 025C	58.139
E01001926	Hammersmith and Fulham 024C	58.039
E01004742	Westminster 022A	57.984

Figure 4.2 – Map Showing the Final Index Scores, alongside the Football Clubs
Source = Office for National Statistics, 2016b



The best performing LSOA is in the borough of Westminster, which is also where four other top ranked LSOAs are located. Other areas that can be highlighted as having a high potential for people attending Bromley FC matches is in Kensington and Chelsea, Camden, Wandsworth, and Hammersmith and Fulham, shown in the map in figure 4.2. One reason for LSOAs in these areas showing their potential for football supporters is due to their disposable income. Those with the money to attend football matches regularly have been identified as being more likely to attend matches (Garcia and Rodriguez, 2002), as they have greater ability to do this compared to those who earn less money than them. This finding in the literature is mirrored in the index results by highlighting wealthy areas of London as potential places for Bromley FC to target due to the characteristics of the population that live in that area.

The boroughs with any LSOAs that have an index score of between 50.4 and 60.8 have been highlighted on this map to allow a detailed focus on the areas with the greatest potential for supporters of Bromley FC. Other than those previously identified, small regions of Croydon have a relatively high score. This can be explained through the proximity to Bromley FC, meaning people may be more likely to attend games. However, a restriction for areas in Croydon and nearby boroughs is the proximity of locations of other football clubs, such as Crystal Palace and Millwall. It is found that throughout the four leagues of English football, people are more likely to support a team who are in a higher league (Miller, 2016). Due to this, Bromley FC will lose fans to surrounding league clubs, even if people live close by. This emphasises issues in attracting supporters, and shows the range of factors that determine attendance.

With the identification of LSOAs, and hence boroughs, that are the most likely to have potential supporters of Bromley FC, the index highlights that the use of geodemographics can allow researchers and policy makers to see the importance of neighbourhoods, and how the characteristics of areas can be vital in what they require (Burrows, 2008). This index shows that this is not just the case in typical sectors, but also in the football industry, where different areas have varying interests. Within a dynamic city such as London, the urban fabric can be vastly different across the whole region (Burrows, 2008), and this index can aid in giving football clubs a better understanding of the range of people and places through LSOA geographies.

On the other hand, there are LSOAs that perform the worst within the index. As shown in table 4.2 these are Harrow, Enfield and Barnet. If these LSOAs are located on the map, then it can be argued that these are all boroughs that are furthest from Bromley FC. This therefore further emphasises the fact that distance is an important factor that determines match attendance. These areas should not be necessarily targeted by Bromley FC, as they have a low number of potential supporters due to the variables utilised in the index, and thus it is unlikely that the population would be willing to attend games. Besides the distance that would have to be travelled, reasons for these being

areas of low potential include a relatively low percentage of White British, and consequently a high number of ethnic minorities. It has been seen that football crowds are predominantly white, with a lack of social inclusion in society and matches (Conn, 2016), and thus it may be argued that the more diverse LSOAs in London will be less likely to have supporters of a club that utilises the framework of this index.

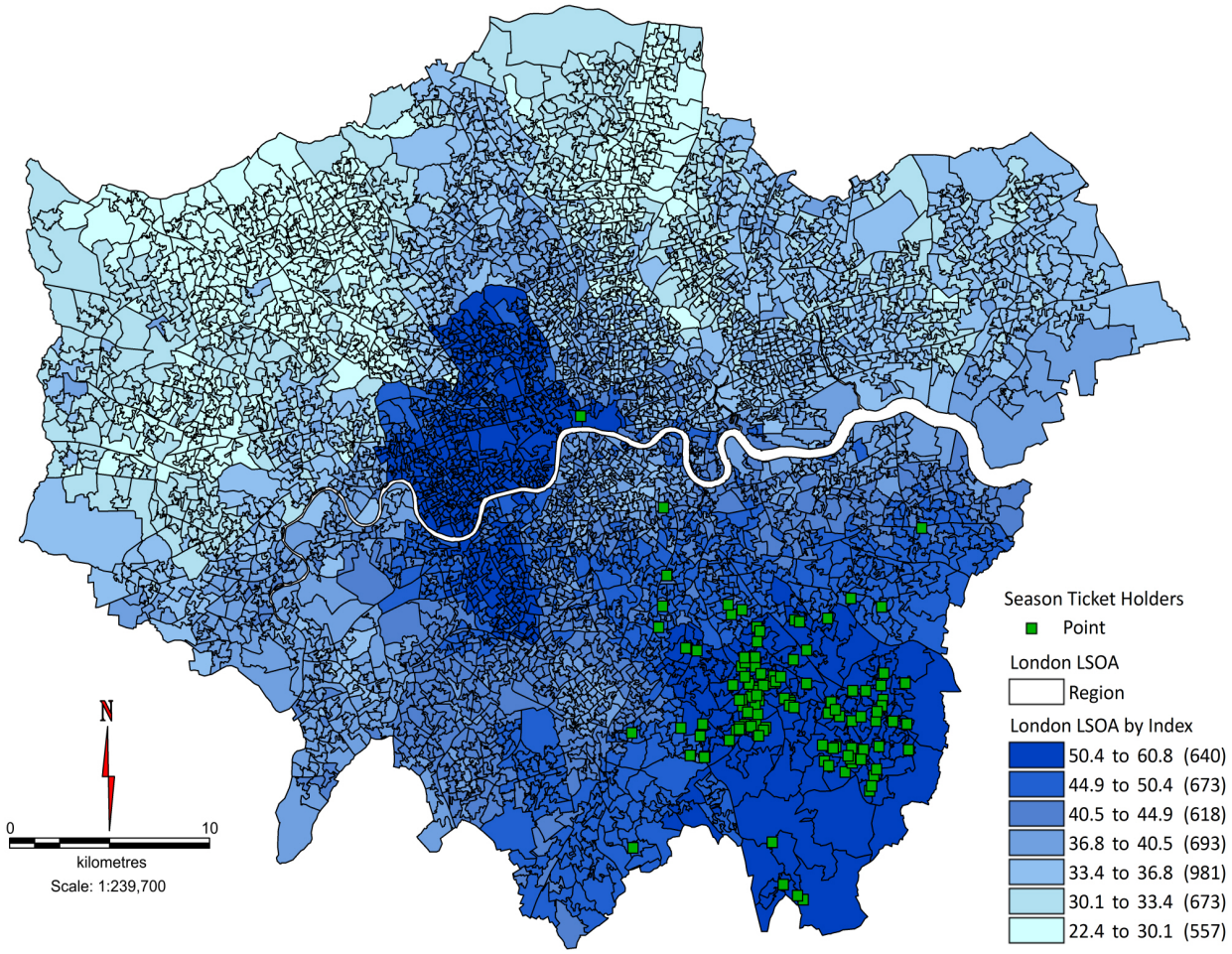
Table 4.2 – Bottom 20 ranked LSOAs within the geodemographic index

GEO_CODE	GEO_LABEL	Index Score
E01002120	Harrow 007B	22.403
E01002235	Harrow 011D	22.672
E01002114	Harrow 001B	23.011
E01001430	Enfield 030E	23.167
E01001429	Enfield 030D	23.560
E01002139	Harrow 006A	23.819
E01001554	Enfield 037C	24.170
E01002227	Harrow 003A	24.177
E01002233	Harrow 011B	24.182
E01000189	Barnet 014D	24.200
E01002113	Harrow 002A	24.249
E01002168	Harrow 017D	24.291
E01002202	Harrow 012E	24.330
E01001425	Enfield 027B	24.400
E01002145	Harrow 005E	24.437
E01002226	Harrow 001D	24.619
E01002203	Harrow 010E	24.620
E01002151	Harrow 006D	24.782
E01002141	Harrow 005B	24.850
E01002201	Harrow 012D	24.852

4.2 – Validation of the Index

In geodemographic methods, validation is an important stage, as an index can only be successful if it has been evaluated for the intended purpose (Singleton and Spielman, 2014). To see if the index is accurate, validation has been carried out using information from season ticket holders of Bromley FC. By mapping the location of current season ticket holders using their residential postcodes, their locations can be compared to areas of high index scores to see whether there is a correlation between the two. These results are shown in the map in figure 4.3, presenting where current supporters live.

Figure 4.3 – Map Showing the Current Season Ticket Holders of Bromley FC
Source = Office for National Statistics, 2016b



Once the season ticket holders were mapped, Structured Query Language was used in MapInfo to calculate the number of season ticket holders in each LSOA. This tool allows a table to be extracted, as shown in table 4.3, highlighting the areas where the current season ticket holders live at present, which are then compared to the index results.

Table 4.3 – Number of season ticket holders within each London LSOA

LSOA	Season Tickets	LSOA	Season Tickets
Bromley 036D	5	Bromley 002B	1
Bromley 036C	4	Bromley 030F	1
Bromley 036B	4	Lewisham 027E	1
Croydon 022C	4	Bromley 019B	1
Bromley 021D	3	Bromley 035C	1
Bromley 028E	3	Bromley 030E	1
Bromley 035B	3	Bromley 023D	1
Bromley 007D	3	Bromley 031E	1
Bromley 028D	3	Bromley 022D	1
Bromley 025E	3	Bromley 006C	1
Bromley 040A	3	Bromley 027D	1
Bromley 018C	3	Bromley 018E	1
Bromley 033D	2	Bromley 037D	1
Bromley 018B	2	Bromley 035D	1
Bexley 027E	2	Bromley 030D	1
Bromley 014B	2	Bromley 031A	1
Bromley 011E	2	Bromley 025D	1
Bromley 032D	2	Bromley 039D	1
Bromley 029C	2	Bromley 001F	1
Bromley 029A	2	Lewisham 038B	1
Bromley 042B	2	Bromley 016B	1
Bromley 032A	2	Croydon 026E	1

Lewisham 006A	2	Lewisham 019C	1
Bromley 027A	2	Bromley 040C	1
Lewisham 035B	2	Bromley 026A	1
Bexley 025D	2	Bromley 021B	1
Bromley 025A	2	Bromley 007A	1
Bromley 014E	2	Bromley 023A	1
Lewisham 030C	2	Bromley 020C	1
Bromley 031D	2	Bromley 008B	1
Croydon 025A	2	Bromley 042C	1
Bromley 023B	2	Bromley 035A	1
Bexley 015D	1	Bromley 018D	1
Bromley 025C	1	Bromley 034D	1
Croydon 039A	1	Bromley 027B	1
Bromley 030G	1	Bromley 018A	1
Bromley 020E	1	Bromley 020D	1
Bromley 022A	1	City of London 001G	1
Bromley 008A	1	Bromley 027C	1
Lewisham 026C	1		

As previously argued, the validation of the index highlights that the most likely places of supporters of Bromley FC to live is within the borough of Bromley, or in nearby surrounding areas. The LSOAs that are home to at least one season ticket holder, as well as being in the top 20 ranked LSOAs from the index, are Bromley 020E, City of London 001G, and Bromley 016B. As there is a current supporter of Bromley FC within the City of London, it highlights how the index can show areas that have the potential to be home to potential supporters. Despite these results, other LSOAs that were ranked highly in the index, such as the boroughs of Westminster and Kensington and Chelsea, do not have any current season ticket holders of Bromley FC.

Although the index does show a certain level of similarities with the current location of season ticket holders, there are reasons that there may be issues with the validation. During the 2016/17 season, Bromley FC had 133 season ticket holders, with only 128 being within Greater London. Thus, there is not enough data to explicitly validate the success of the index. If another club was being investigated, that had more season ticket holders, then a better pattern may have been identified. However, one of the reasons for using Bromley FC as the by-product of the index was because of their potential for growth, and thus there was going to be a relatively small number of season ticket holders to be used in the validation. Furthermore, the validation process does not consider those who do not own a season ticket, but still attend many of the games. If other supporters of Bromley FC were mapped based on their residential location, then there may be a better comparison between the final index results and the distribution of supporters of Bromley FC.

However, these issues do not render the index unusable, and it can still be utilised by football clubs to target certain areas for marketing purposes. The areas that are ranked highly can be still be targeted, as even though they are not current locations of season ticket holders, it does not mean they cannot be in the future. Moreover, as supporters of Bromley FC mainly come from the Bromley area, there is still potential for more growth in these LSOAs. Sport is a sector that is highly personalised and unpredictable (Tapp and Clowes, 2000). Even if the validation cannot be directly compared to the index results, it can still aid football clubs in predicting where their ideal customers would be located to a certain extent, allowing them to grow and improve attendances.

As Bromley is one of the most important regions for supporters, it may be necessary to modify the weightings of the variables within the index, so that there is more emphasis on distance. However, this is another aspect of subjectivity within the creation of the index, and can be decided on depending on the football club in question.

This section has analysed the geodemographic index that has been created, and discussed the major spatial patterns identified. The results have then been validated against data from Bromley FC about their current season ticket holders to see if the index produces similar results. The following chapter will summarise the findings of this study, and consider the future of this research and the index that has been created.

CHAPTER FIVE – CONCLUSION

This chapter will summarise the findings of this research project, and identify further research from this dissertation, as well as state any possible limitations.

5.1 – Overall Findings

The research throughout this project has aimed to highlight the socio-economic and lifestyle characteristics of football supporters, using this to create a geodemographic index that is able to locate LSOAs with a high number of potential supporters of a particular club. The literature that has been reviewed has analysed the use of geodemographics to understand their importance, and how neighbourhoods can be diverse (Burrows, 2008), displaying a range of characteristics of people. A major finding is that those who live nearby each other are more likely to have similarities than those who live further away. Advantages of these methods include complete coverage of a specific area (Leventhal, 2016), and providing a rich description of distinguishing features of different geographical units (Batey and Brown, 1995).

The links between the sport industry and geodemographic methods have been under-researched since the increase in spatial analysis in the past few decades. Despite the range of uses of geodemographics, little was found on the way they had been applied to football. Despite the lack of focus in the past, the football industry is an important sector of the economy by being a £20 billion industry (Critchlow, 2015), and to ensure the continuing success of football clubs, they need to fully understand their supporters and where they live. This research has aimed to aid football clubs in this task.

An investigation into the demand for football and what contributes to attendances at matches emphasised the determinants of being a football supporter, including social, economic and lifestyle variables that can be used within the index. The creation of this

index has provided evidence of the use of spatial analysis, and how these methods can be applied in the football industry to highlight the market potential, and the benefits this can have to football clubs. By ranking areas based on variables, the value of geodemographics can be applied to football as it has been before to other sectors.

In this project, the by-product of the index has been Bromley FC. As a relatively small club, they have aspirations for growth and the ability for more supporters and thus were an ideal club to target. Once the index was created, the results showed that the higher the final index score, the more likely a LSOA is to have a higher potential of supporters of Bromley FC. This general pattern was in line with what was expected in the initial stages of the research, with high scores in the chosen variables meaning a higher likelihood of football attendance. LSOAs immediately surrounding Bromley FC were seen to have the highest index scores, thus highlighting the fact that the football club should tailor their marketing strategy to areas near the club. Other areas of high potential for supporters included Camden, Wandsworth, Kensington and Chelsea, and Hammersmith and Fulham. This analysis highlights the range of factors influencing football attendances, and shows areas that can be targeted by local football clubs. The general benefits of geodemographics have been seen with the creation of this index, including the fact it is easily accessible, hence emphasising its reproducibility. This means that it can be used in different contexts by various football clubs (Leventhal, 2016), whilst having the same overall aim of evidencing the use of spatial analysis within the football industry by using variables to create a geodemographic index.

Validation stages, through the mapping of current season ticket holders, did not completely match the results of the index. Despite this, it highlighted that those LSOAs that were ranked highly can still be targeted by football clubs, with modifications depending on which football club is utilising the index. One of the major benefits of the creation of this index is its replicability, broadening the use of geodemographics within sports, and more specifically within the football industry across all leagues.

The football industry is significant and has numerous opportunities for growth, including using the geodemographics of places to benefit football clubs financially by gaining further income. This wealth of opportunities can be used to help them understand their supporters, which in turn can make targeting the relevant groups easier (Birkin et al, 1996). Research has found that season ticket holders are now making up a greater proportion of the supporter bases of football clubs, accounting for 10 million of the 15 million admissions to Football League matches each season (BBC Sport, 2015). If this trend continues, then there is the opportunity to promote further purchasing of season tickets in the targeted areas, utilising the results of this index. Using geodemographics can only benefit a football club, especially those that have such a great opportunity for expansion, such as Bromley FC. The targeting of areas through this index will allow them to gain further income by enhancing ticket sales.

5.2 – Implications and Future Work

As previously mentioned, the methods for creating this index have not been applied in sport previously, thus making this research innovative and imperative for the football industry. This index will have a high value for football clubs in London as it can be modified and utilised by any club that feels they want to increase their match-day attendances, and attract more supporters. If LSOAs that have a high potential of supporters of Bromley FC are targeted by their marketing campaigns in a successful manner, then the index has the ability to be applied to other football clubs in London and across the country. By highlighting areas that are more likely to have potential football supporters of a club, marketing and promotion can be more efficient and successful, thus improving profit of football clubs, and helping them to have a positive impact within the local community they are in.

It is argued by many that commercial geodemographics has reached stagnation in terms of the tools and methods that are being used (Singleton and Longley, 2009),

Thus, there needs to be further integration of new technology to improve the methods that are being used, and ensure they remain successful. However, despite this potential issue, geodemographics will continue to remain important in solving spatial problems (Singleton and Longley, 2009) as society continues to change. In terms of the index, this means that it must strive to use up to date information and data, ensuring it will maintain its aim of aiding football clubs with the issue of their attendances.

Geodemographics and spatial analysis are dynamic fields within geography, and thus the data and patterns that are involved are constantly evolving alongside technological advancements. Future research that can emerge out of this index may utilise the growing trend of social media. Data from platforms such as Twitter have allowed more of an understanding of people's behaviour in virtual as well as observable space (Adnan, Lansley and Longley, 2012). Tweets can have location tags to see what people are talking about and where, allowing an observation of the differences between residential geographies of various groups within society (Adnan, Lansley and Longley, 2012). This advanced technology of geodemographics could be incorporated into the index to further highlight where potential football supporters would be located within a certain region. If there is social media data that includes information about football attendances, the source of this may be areas with a greater likelihood of having people who would support a football club and attend matches regularly.

5.3 – Limitations to the Research

Although there are obvious benefits to this index, there are issues and limitations that need to be considered. The variables included in the index are highly generalising, such as the fact that not every football fan is going to be male, or fall within the chosen age category. This means that there will be areas that have a high potential for supporters of a football club, but will have a low index score, or vice versa. Similarly, there needs to be a consideration of ecological fallacy when analysing the index. Not everyone

within an LSOA will possess the same characteristics (Leventhal, 2016), and thus assumptions are being made about individuals based on the analysis of a group (Birkin, 1995). Moreover, if a different geography had been chosen, and different variables had been included in the index, then it would have generated different results. As this index is going to be transferable across the country, those utilising it will need to make their own decisions, and consider these implications within their analysis.

There are also issues with indexes and geodemographics in general. When census data is used, there can be accuracy issues (Sleight, 2004). This is because it is not possible for the data to be completely up to date, causing degradation over time, thus making the results not as accurate as they could be. Furthermore, the theoretical frameworks that underpin many geodemographic systems remain weak (Singleton and Longley, 2009), making it difficult to truly determine the effects of the index that has been created. The validation stage is useful and vital when creating an index, however because the strength of geodemographics is to explore and describe rather than to explain, then the analysis may fall short (Peterson et al, 2009). Therefore, the index may not fully help football clubs to understand their fans, and those people who are more likely to attend their games.

The analysis of this index shows areas that are going to be targeted because they have a higher likelihood of being home to people who display characteristics synonymous with football attendance. However, even if these areas have the potential for football supporters, it is not guaranteed that those people who live there will want to attend matches. An increase in televised matches has been seen to be one of the major obstacles in attracting fans to live games (Soberg and Mehus, 2014), and thus even if areas are targeted, it may not result in any successful improvement in attendances.

To conclude, despite some limitations, the index can support football clubs in their aim of increasing their support. The final index scores, and the spatial analysis of these results, highlights the use of geodemographic methods within the football industry, and how they can benefit football clubs. As the Premier League becomes more popular, lesser supported clubs need to find ways to improve people's identification to their football club (Soberg and Mehus, 2014). The index created can aid football clubs in this by targeting specific areas, and encouraging people to attend games, continuing the success of the football industry on a local level, focusing on the population.

BIBLIOGRAPHY

Abbas J, Ojo A and Orange S. (2009). 'Geodemographics – a tool for health intelligence?'. Public Health. 123. pp.35-39.

Acorn. (2013). 'Acorn Classification'. [Online] [Accessed on 16th January 2017] Available from: http://acorn.caci.co.uk

Adnan, M et al. (2010). 'Towards real-time geodemographics: clustering algorithm performance for large multidimensional spatial databases'. Transactions in GIS. 14(3). pp.283-297.

Adnan M, Lansley G and Longley P. (2012). 'Twitter geodemographic analysis of ethnicity and identity in Greater London'. University College London, Department of Geography. [Online] [Accessed on 12th March 2016] Available from: http://www.geocomputation.org/2013/papers/51.pdf

Anderson, T. (2010). 'Using geodemographics to measure and explain social and environment differences in road traffic accident risk'. Environment and Planning A. 42. pp.2186-2200.

Baçao F Lobo V and Painho M. (2004). 'Clustering Census Data: Comparing the Performance of Self-Organising Maps and K-means Algorithms'. KDNet Symposium, Bonn, Germany. [Online] [Accessed on 22nd October 2016] Available from: http://www.novaims.unl.pt/labnt/geosom/Public/o8-2_11_lobo04_som_kmeans_kdnet.pdf

Baker K, Bermingham K and McDonald C. (1979). 'The utility to market research of the classification of residential neighbourhoods'. Proceedings of the Market Research Society Conference, Brighton. pp. 253-271.

Bale, S. (2000). 'The changing face of football: stadiums and communities.' Soccer and Society. 1(1). pp.91-101.

Batey, P and Brown, P. (1995).'From human ecology to customer targeting: the evolution of geodemographics'. In: Longley, P and Clarke, G, eds, *GIS for Business and Service Planning*. London: Taylor & Francis. pp.115-144.

BBC News. (2010). 'More diverse crowds at Premier League football' – BBC News, 25th August 2010. [Online] [Accessed on 28th January 2017] Available from: http://www.bbc.co.uk/news/uk-11079597

BBC Sport. (2015). 'Price of Football: Ticket Increases Outstrip Cost of Living' – BBC Sport, 15th October 2015. [Online] [Accessed on 24th January 2017] Available from: http://www.bbc.co.uk/sport/football/29614980

Birkin, M. (1995). 'Customer targeting, geodemographics and lifestyle approaches'. In: Longley, P and Clarke, G, eds, *GIS for Business and Service Planning*. London: Taylor & Francis. pp.104-149.

Birkin M, Clarke G, Clarke M, and Wilson A. (1996). '*Intelligent GIS'*. Cambridge: GeoInformation International.

Bromley FC. (2017). 'Bromley FC Website'. [Online] [Accessed on 26th January 2017] Available from: http://www.bromleyfc.tv/site/

Buraimo, B. (2008). 'Stadium Attendance and Television Audience Demand in English League Football'. Managerial and Decision Economics. 29(6). pp.512-523.

Buraimo B, Forrest D, and Simmons R. (2009). 'Insights for clubs from modelling match attendance in football'. Journal of the Operational Research Society'. 60(2). pp.147-155.

Burrows, R. (2008). 'Geodemographics and the construction of differentiated neighbourhoods'. In: Flint, J and Robinson, D, eds, *Community Cohesion in Crisis? New Dimensions of Diversity and Difference*. [Online] [Accessed on 10th October 2016] Available from: http://policypress.universitypressscholarship.com/view/10.1332/policypress/9781847420244.001.0001/upso-9781847420244-chapter-12

Burrows, R. (2013). 'The New Gilded Ghettos: The Geodemographics of the Super-Rich'. Discover Society. 3. [Online] [Accessed on 12th January 2017] Available from: http://discoversociety.org/2013/12/03/the-new-gilded-ghettos-the-geodemographics-of-the-super-rich/

Butler, T and Hamnett, C. (2007). 'The Geography of Education: Introduction'. Urban Studies. 44(7). pp.1161-1174.

Campaign for Better Transport. (2013). 'Door to Turnstile, Improving Travel Choices for Football Fans'. [Online] [Accessed on 28th January 2017] Available from: http://www.bettertransport.org.uk/sites/default/files/research-files/Door_to_Turnstile_CfBT_FINAL_web.pdf

Chainey, S. (2005). '3rd National Crime Mapping Conference' – Jill Dando Institute of Crime Science, 12th-13th April 2005. [Online] [Accessed on 15th October 2016] Available from: https://www.ucl.ac.uk/jdi/events/mapping-conf/conf-2005/conf2005-downloads/spencer-chainey-crime-mapping-guide.pdf

Conn, D. (2006). 'The clubs lost the trust of the ethnic minority people' – The Guardian, 18th October 2006. [Online] [Accessed on 16th February 2017] Available from: https://www.theguardian.com/sport/blog/2006/oct/18/theclubslostthetrustofet

Critchlow, A. (2015). 'Britain's sport industry hitting top form' – The Telegraph, 3rd May 2015. [Online] [Accessed on 23rd October 2016] Available from: http://www.telegraph.co.uk/finance/11580773/Britains-sport-industry-hitting-top-form.html

Donnay JP and Unwin, D. (2001). 'Modelling geographical distributions in urban areas. In: Donnay JP, Barnsley M and Longley P, eds, *Remote Sensing and Urban Analysis*. London: Taylor & Francis. pp.205-224.

Experian. (2017). 'MOSAIC Geodemographic Classification System'. [Online] [Accessed on 17th January 2017] Available from: http://www.experian.co.uk/marketing-services/products/mosaic-uk.html

Furness, P and Sleight, P. (2013). 'Tracking a Decade of Changing Britain' – Conference Report, 18th December 2013. [Online] [Accessed on 23rd October 2016] Available from: https://www.mrs.org.uk/pdf/Chairmen's%20Conference%20Report%20061113.pdf

Garcia, J and Rodriguez, P. (2002). 'The Determinants of Football Match Attendance Revisited: Empirical Evidence from Spanish Football League'. Journal of Sports Economics. 3. pp.18-38.

GKB. (2017). 'Geodemographics Knowledge Base' – MRS, Evidence Matters. [Online] [Accessed on19th February 2017] Available from: https://www.geodemographics.org.uk/geodemographics

GLA Intelligence. (2016). 'GLA 2015 round trend-based projections: Results'. – Greater London Authority, July 2016. [Online] [Accessed on 26th January 2017] Available from: https://files.datapress.com/london/dataset/2015-round-population-projections/2016-07-19T13:06:46/2015%20trend%20projection%20results.pdf

Gupta, R. (2013). 'Clubs like Brighton & Millwall take steps to halt declining attendances' – BBC Sport, 24th January 2013. [Online] [Accessed on 19th February 2017] Available from: http://www.bbc.co.uk/sport/football/21142999

Harris, R. (2003). 'Population mapping by geodemographics and digital imagery', In: Mesev, V, ed*, Remotely Sensed Cities*. London: Taylor & Francis. pp.223-241.

Harris, R. (2009). 'Geodemographics'. In: Gregory, D et al, 5th ed*, The Dictionary of Human Geography.* Singapore. Wiley: Blackwell. pp.276-277.

Harris R, Sleight P and Webber R. (2005). '*Geodemographics, GIS and Neighbourhood Targeting'*. Chichester: Wiley

Hidden London. (2005). 'A history of London football'. [Online] [Accessed on 26th January 2017] Available from: http://hidden-london.com/miscellany/football/history/

Inman J, Shankar V and Ferraro R. (2004). 'The Roles of Channel-Category Associations and Geodemographics in Channel Patronage'. Journal of Marketing. 68(2). pp.51-71

Kamarulzaman, Y. (2010). 'Geodemographics of Travel E-shoppers: An Empirical Analysis of UK Consumers'. European Journal of Social Sciences'. 16(2). pp.195-205

Kearns, A and Parkinson, M. (2001). 'The significance of neighbourhoods'. Urban Studies. 12. pp.2103-2110.

Leventhal, B. (2016). '*Geodemographics for Marketers'*. Croydon: KoganPage.

London Data Store. (2017). 'London Data Store' – London Assembly, DataPress. [Online] [Accessed on 21st February 2017] Available from: https://data.london.gov.uk

Miller, M. (2016). 'Stat shows Manchester United and Liverpool fans are glory supporters' – The Metro, 22nd February 2016. [Online [Accessed on 22nd February 2017] Available from: http://metro.co.uk/2016/02/22/stat-shows-manchester-united-and-liverpool-fans-are-glory-supporters-5709814/

Neave, N and Wolfson, S. (2003). 'Testosterone, territoriality and the 'home advantage' Physiology and Behaviour. 78(2). pp.269-275.

Ofcom. (2005). 'Premier League Football: Research Into Viewing Trends, Stadium Attendances, Fans' Preferences and Behaviour, and the Commercial Market' – Commission of the European Communities. [Online] [Accessed on 13th January 2017] Available from: http://ec.europa.eu/competition/antitrust/cases/dec_docs/38173/38173_104_7.pdf

Office for National Statistics. (2016a). '2011 Census Aggregate Data'. UK Data Service, Edition June 2016. [Online] [Accessed on 21st February 2017] Available from: http://dx.doi.org/10.5257/census/aggregate-2011-1

Office for National Statistics. (2016b). 'ONS Census Digitised Boundary Data (England and Wales) – Computer File. UK Data Service Census Support. [Online] [Accessed on 21st February 2017] Available from: https://borders.ukdataservice.ac.uk

Office for National Statistics. (2016c). '2011 ONS LSOA Geography'. [Online] [Accessed on 21st February 2017] Available from: http://webarchive.nationalarchives.gov.uk/20160105160709/http://neighbourhood.statistics.gov.uk/dissemination/info.do?page=nessgeography/superoutputareasexplained/output-areas-explained.htm

Peterson, J et al. (2009). 'Geodemographics as a tool for targeting neighbourhoods in public health campaigns'. J Geogr Syst. 13. pp.173-192.

Pointer, G. (2005). 'The UK's Major Urban Areas' – Chapter 3, Focus on People and Migration. [Online] [Accessed on 26th January 2017] Available from: https://www.ons.gov.uk

Reibel, M. (2011). 'Classification Approaches in Neighbourhood Research: Introduction and Review'. Urban Geography. 32(3). pp.305-316.

Simmons, R. (1996). 'The demand for English league football: a club-level analysis'. Applied Economics. 28(2). pp.139-155.

Singleton, A. (2004). 'A State of the Art Review of Geodemographics and their Applicability to the Higher Education Market'. Working Paper Series. [Online] [Accessed on 13th January 2017] Available from: http://discovery.ucl.ac.uk/212/1/paper74.pdf

Singleton, A and Longley, P. (2009). 'Geodemographics, visualisation, and social networks in applied geography'. Applied Geography. 29. pp.289-298.

Singleton, A and Spielman, S. (2014). 'The Past, Present and Future of Geodemographic Research in the United States and United Kingdom'. The Professional Geographer. 66(4). pp.558-567.

Sleight, P. (2004). '*Targeting Customers: How to Use Geodemographic and Lifestyle Data in Your Business*'. 3Rd Edition. Henley-on-Thames: World Advertising Research Centre.

Soberg, H and Mehus, I. (2014). 'The Challenge of Attracting Football Fans to Stadia?' International Journal of Sport Finance. 9(1). pp.3-19

Sport England. (2013). 'Economic value of sport in England' – Sport England, July 2013. [Online] [Accessed on 23rd October 2016] Available from: https://www.sportengland.org/media/3174/economic-value-of-sport-summary.pdf

Supporters Direct. (2010). 'The Social and Community Value of Football' – Summary Report June 2010. [Online] [Accessed on 28th January 2017] Available from: http://www.supporters-direct.org/wp-content/uploads/2012/08/svoff-summary-report.pdf

Tapp, A and Clowes, J. (2002). 'From "carefree casuals" to "professional wanders". Segmentation possibilities for football supporters'. European Journal of Marketing. 36(11/12). pp.1248-1269.

Thompson, D. (2014). 'Which sports have the whitest/richest/oldest/fans?' – The Atlantic, 10th February 2014. [Online] [Accessed on 12th March 2016] Available from: http://www.theatlantic.com/business/archive/2014/02/which-sports-have-the-whitest-richest-oldest-fans/283626/

Voas, D and Williamson, P. (2001). 'The diversity of diversity: a critique of geodemographic classification'. Area. 33(1). pp.63-76.

Webber, R and Longley, P. (2003). 'Geodemographic analysis of similarity and proximity: their roles in the understanding of the geography of need'. In: Longley, P and Batty M, eds, *Advanced Spatial Analysis*. California: ESRI. pp233-266.

Webber, R and Butler, T. (2007). 'Classifying Pupils by Where They Live: How Well Does This Predict Variations in Their GCSE Results?' Urban Studies. 44(7). pp.1229-1253.

White, J. (2013). 'Football grounds aren't full of old white men after all' – The Telegraph, 10th January 2013. [Online] [Accessed on 28th January 2017] Available from: http://www.telegraph.co.uk/sport/football/9793689/Football-grounds-arent-full-of-old-white-men-after-all-....html

Wood, Z and Allen, K. (2016). 'Where have all Britain's shoppers gone?' – The Guardian, 23rd April 2016. [Online] [Accessed on 16th January 2017] Available from: https://www.theguardian.com/business/2016/apr/23/uk-high-street-spending-shoppers-slowdown

YouGov (2017). 'Football Fan – YouGov Profiles'. [Online] [Accessed on 21st February 2017] Available from: https://yougov.co.uk/profileslite#/Football/media

APPENDIX A: EXTRACT OF THE INDEX

LSOA Name and Code	Age	Male	No Religion	White British	Disposable Income	Access to Car	Distance to Football Club	Distance to Other Clubs	Total Score	Final Index Score
Islington 022H	0.071679521	0.026709409	0.0413977	0.019570552	0.038228068	0.011978127	0.184234624	0.022327551	0.416125551	41.613
Islington 011H	0.062003078	0.033048771	0.039789982	0.020443642	0.038228068	0.011642995	0.159387184	0.001173065	0.365716785	36.572
Islington 011I	0.084836841	0.03624567	0.039775756	0.025509736	0.038228068	0.022742257	0.159147042	0.000912395	0.407397766	40.740
Islington 006F	0.093962751	0.047570303	0.043153491	0.030881123	0.038228068	0.03117717	0.158444905	0.003830868	0.447248679	44.725
Islington 006G	0.053061404	0.029593285	0.040251019	0.031565961	0.038228068	0.043002063	0.157611381	0.003602284	0.396915465	39.692
Kingston upon Thames 004F	0.060288523	0.027961135	0.026433865	0.02191307	0.027402804	0.058047516	0.142918934	0.012024613	0.37699046	37.699
Kingston upon Thames 004G	0.056409864	0.03104108	0.028956987	0.029274365	0.027402804	0.056619939	0.144893802	0.010703412	0.385302253	38.530
Richmond upon Thames 004G	0.047020854	0.033228702	0.025599516	0.021537965	0.040359879	0.073169815	0.143358985	0.014077289	0.398353002	39.835
Richmond upon Thames 004H	0.017089866	0.025311422	0.025643488	0.034611437	0.040359879	0.071570558	0.142484317	0.014255808	0.371326775	37.133
Barnet 030E	0.037748475	0.037247855	0.010695719	0.009822703	0.040280387	0.075549145	0.095991233	0.018368248	0.325703767	32.570
Barnet 030F	0.080919643	0.021973805	0.020479087	0.011218707	0.040280387	0.05353412	0.097153419	0.020507845	0.346067012	34.607
Newham 013F	0.083363823	0.036679141	0.025830393	0.010904305	0.006409886	0.021889986	0.188672678	0.008096227	0.381846438	38.185

APPENDIX B: DSG FORMS

GEOG3600 DISSERTATION

DSG REPORT FORM

Name of Student: Georgina Weaver

Meeting Number: 1 **Date:** 27/09/16

Members present: Georgie, Amber, Jordan, and mentor

Apologies for absence: Els

Meeting With Mentor

Issues to be raised (complete prior to each meeting)**:**

Confirming what the dissertation topic is, and what progress has been made

Solutions discussed (to be completed during/after the meeting)**:**

What the main elements of a dissertation are
What needs to be done for the first progress report due in week 5
Begin on literature reading and organizing aims/objectives
Sorting out when the deadlines are

Questions for mentor (record after peer-group meeting only; bring to next meeting with mentor)**:**

How do I go about deciding on what variables would be the most appropriate

GEOG3600 DISSERTATION

DSG REPORT FORM

Name of Student: Georgina Weaver

Meeting Number: 2 **Date:** 19/10/16

Members present: Georgie, Jordan, Amber, Els

Apologies for absence:

Peer Group Meeting

Issues to be raised (complete prior to each meeting)**:**

How people have progressed with their progress report
How people have decided to collect their information/data
Confirmation of what project other members are doing for their dissertation

Solutions discussed (to be completed during/after the meeting)**:**

How to go about creating an index – all of which the group members are doing
Discuss literature found about geodemographics in general
How many questionnaires need to be carried out during the research

Questions for mentor (record after peer-group meeting only; bring to next meeting with mentor)**:**

Can all the DSG forms be the same for every group member?

GEOG3600 DISSERTATION

DSG REPORT FORM

Name of Student: Georgina Weaver

Meeting Number: 3 **Date:** 17/11/17

Members present: Georgie, Amber, Els, Jordan

Apologies for absence:

Peer Group Meeting

Issues to be raised (complete prior to each meeting)**:**

What stage are people at with their dissertations
What variables have been chosen for each index
How much data has been collected so far and how/where from?

Solutions discussed (to be completed during/after the meeting)**:**

What needs to be collected for next week, confirmation of data
How data is going to be used for the index, and why it is the most appropriate
What is going to be completed during the rest of the semester
Complete getting the data to be used, can then start sorting this out
Support each other for mapping and creation of the index

Questions for mentor (record after peer-group meeting only; bring to next meeting with mentor)

Find the best way to carry out an index
Ask about literature and methodology sections

GEOG3600 DISSERTATION

DSG REPORT FORM

Name of Student: Georgina Weaver

Meeting Number: 4 **Date:** 2/12/16

Members present: Georgie, Amber, Els and Jordan

Apologies for absence:

Meeting with Mentor

Issues to be raised (complete prior to each meeting)**:**

Workshop on how to create an index
What are the most suitable variables to use?
How to find the distance from each football club to a LSOA

Solutions discussed (to be completed during/after the meeting)**:**

Use the UK data service for boundary, census data from Infuse
Convert data from each variable into a percentage, standardizing the data
Work out correlation through excel calculations, can remove a variable
Normalization, stretch out every scale between 0 and 1
Polarity, ensure all data for variables is going in the same direction
Think about weighting for the variables what's the most important
Distance, add points to MapInfo, find the centre of LSOA, use distance calculator
Add all weighted scores for each area to give final index score, then times by 100

Questions for mentor (record after peer-group meeting only; bring to next meeting with mentor)

Confirm whether chosen variables are appropriate for this index
Is it possible to use disposable household income for each borough

GEOG3600 DISSERTATION

DSG REPORT FORM

Name of Student: Georgina Weaver

Meeting Number: 5 **Date:** 23/1/17

Members present: Georgie, Amber, Els, Jordan, and mentor

Apologies for absence:

Meeting With Mentor

Issues to be raised (complete prior to each meeting)**:**

How do I reference Infuse etc.
Do we include all references on the interim report?
How far has everyone progressed with their index/dissertation
How many references need to be in the methodology?

Solutions discussed (to be completed during/after the meeting)**:**

Reference Infuse by simply referencing Census Data 2011 etc.
In interim report, can just include references that have been used in that
Only need a handful of references within the methodology
Can include methodology section within the interim report
Analysis section will be short, just maps and explaining indexes
Create flow diagram to show different sections of the index creation
Link findings from this index to the relevant literature fields
At the end of each chapter, summarize, and explain what is in next section

Questions for mentor (record after peer-group meeting only; bring to next meeting with mentor)

Why am I struggling to map my index on MapInfo?
Check whether my index looks plausible

GEOG3600 DISSERTATION

DSG REPORT FORM

Name of Student: Georgina Weaver

Meeting Number: 6 **Date:** 27/1/17

Members present: Georgie, Amber, Els and Jordan

Apologies for absence:

Peer Group Meeting

Issues to be raised (complete prior to each meeting)**:**

Discuss our feedback from the interim report
What people feel is the most urgent to do in the coming weeks

Solutions discussed (to be completed during/after the meeting)**:**

Methodologies are going well, emphasizing the stages that we went through
Now need to focus on the results and analysis section
How to show the results obtained from the creation of the index
What needs to change in our aims and objectives

Questions for mentor (record after peer-group meeting only; bring to next meeting with mentor)

How many maps is appropriate to include in the dissertation?

GEOG3600 DISSERTATION

DSG REPORT FORM

Name of Student: Georgina Weaver

Meeting Number: 7 **Date:** 9/2/17

Members present: Georgie, Amber, Els and Jordan

Apologies for absence:

Peer Group Meeting

Issues to be raised (complete prior to each meeting)**:**

How is everyone's index creation stage going
What is being included within the results and analysis section for each person?
Confirming what each person's final title will be

Solutions discussed (to be completed during/after the meeting)**:**

Planning from now and until submission, 1 month to go
How we are going to deal with the pressure of the deadline
What is still left to do, need to focus on the results and analysis section

Questions for mentor (record after peer-group meeting only; bring to next meeting with mentor)

Whether or not there is information on the final stages of the overall process

GEOG3600 DISSERTATION

DSG REPORT FORM

Name of Student: Georgina Weaver

Meeting Number: 8 **Date:** 24/2/17

Members present: Georgie, Amber, Els and Jordan

Apologies for absence:

Peer Group Meeting

Issues to be raised (complete prior to each meeting)**:**

Questions about the final draft of the dissertation
What still needs to be done before the deadline

Solutions discussed (to be completed during/after the meeting)**:**

Mainly just about the final touches and modifications of each section
Ensuring that all references are included, and all tables/figures are labelled
How we go about binding and printing the final version
How people have found the overall process of the dissertation

Questions for mentor (record after peer-group meeting only; bring to next meeting with mentor)

No questions because close to the deadline

APPENDIX C: DSG REFLECTIVE LOG

The Dissertation Support Group (DSG) has been an important aspect of my overall dissertation work. Throughout the process, we met at regular intervals, either with or without our mentor, in order to discuss what progress was being made at each stage of the research, and to ask any questions, or talk about any issues we were encountering.

I have found the DSG meetings extremely useful whilst carrying out my dissertation. The support network that was created meant that whenever there was an issue, or concern over what stage everyone was at, there was guidance and help available. Having a Facebook group set up meant that we had the ability to discuss and share ideas at any point throughout the year. Moreover, discussions before meeting with our mentor, Luke Burns, meant that we could sort any issue more efficiently, and allowed us to fully understand what was being expected of us.

Another major benefit of our DSG was that we were all undertaking similar projects in terms of the methods that were being carried out. This was most important in the methodology stages, for example with the data collection, as we were all creating our own version of a geodemographic index. As all our projects involved secondary data sources, we could find out how people had found their data, and what the best methods of collection were, to ensure our variables were relevant. As a result, it can be argued that the DSG has improved my understanding of data collection, and allowed us to enhance our own projects.

In conclusion, I have highly benefited from having a Dissertation Support Group throughout the dissertation process. I believe it is a worthwhile aspect for all Geography students, so that they can have discussions with people they might not have encountered before, allowing group conversations about relevant topics. I am very appreciative of having such a nice group of people to support me along the way.

APPENDIX D: FIRST INTERIM REPORT

GEOG3600 – Dissertation First Interim Report

Title

Creating a geodemographic classification index to highlight areas of Greater London with high potential for supporters of football clubs

Literature Review

The term geodemographics refers to the analysis of where people live (Sleight, 2004) through the use of a data-based classification of their residential location (Harris, 2009). By clustering areas based on the demographic and lifestyle characteristics, patterns can be identified and predictions of behaviour can be made for particular areas. Within the social sciences, and in particularly Geography, the overall theme is that small geographical areas are different to one another, and thus can be classified to see more explicit differences between populations. Evidence shows that geodemograhics' main use has been in resource allocation, and it is thus an important tool in planning and marketing strategies.

Specific examples of where geodemographics has been used include within education. Butler and Hamnett examined the changing geographies of areas, with an increase in social segregation finding its way into the education system (Butler & Hamnett, 2007). These patterns are then compared to the educational provision and attainment of school pupils to see what causes the variation in results (Butler and Hamnett, 2007). Similarly, Webber and Butler analysed school attainment levels to see patterns between neighbourhoods, concluding that the more affluent the area, the higher the average GCSE score (Webber and Butler, 2007). As a result, policy makers are able to examine what factors affect the performance of pupils, and effectively manage these differences.

In the public sector, organisations have been looking to use consumer insight to understand the needs and behaviour of target populations, including within health. Geodemographic tools are being used to summarise general attributes of an area, and then link in health variables, to inform health service planners at a local level which areas resources should be focused on (Abbas et al, 2009). This profiling and clustering provides the potential for benchmarking against similar populations. Another example is in the mapping of crime, analysing the characteristics of high crime neighbourhoods, and why crimes occur where they do (Chainey, 2005).

Using geodemographic information for marketing purposes has been vital in recent years, where a classification can be used by companies to obtain a better understanding of their customer base (Leventhal, 2016). A code can be obtained to show clusters of locations that "can receive different marketing treatments" (Leventhal, 2016, p46), emphasising the importance of targeting specific areas. For business purposes, this provides more accurate sampling frames for data to be used by companies (Sleight, 2004), as well as emphasising plausible social geographies (Birkin, 1995).

Retail marketing and analysis can also benefit from geodemographics. The population of each group of people within a specified distance can be calculated using GIS (Birkin, 1995). This will enable retailers to understand what types of people they will be attracting nearby and infer information about core customers (Harris, 2003), thus target their products towards these demographics. As well as traditional datasets being utilised, the use of social media has allowed more of an understanding of people's behaviour in virtual space (Adnan, Lansley and Longley, 2012).

The surrounding literature highlights many benefits to creating an index. Classifications can analyse exact locations and map them to show spatial patterns, with postcodes giving complete coverage of regions (Leventhal, 2016). Moreover, clustering techniques provide a rich description of distinguishing features of types of areas (Batey and Brown 1995) that can be used for analysis. More recent uses, such as with social media data, can give an insight into the geography of individuals, rather than a whole group of people (Adnan, Lansley and Longley, 2012). This has the ability to improve the accuracy of the analysis that is obtained. The clustered nature of the analysis is important as the urban fabric of cities can be vastly different (Burrows, 2008), and thus geodemographics allows a better understanding of the dynamics of locations.

As with any form of analysis, utilising geodemographics within research does bring about potential issues. Assumptions are often used, as when clusters are created, not everyone within the area will possess the same characteristics (Leventhal, 2016). Due to this problem, when data analysis is carried out from the index, there needs to be attempts to avoid ecological fallacy. Furthermore, a major issue for classification systems is often that when census data is used, there can be accuracy issues (Sleight, 2004), as it isn't possible for it to constantly be up to date.

Geodemographics has not yet been incorporated into football, and thus there is a gap in the literature that needs to be explored. Football is a significant sector in the culture & economy of the UK (Sport England, 2013), and with sport in the UK contributing to over 450,000 jobs in a £20 billion industry (Critchlow, 2015), there is further potential for the growth of the market. Football clubs will be able to benefit from a new index by improving their marketing to target certain areas and increase their supporters base.

Aims and Objectives

The overarching aim of the project is to create a geodemographic classification index to cluster areas of Greater London based on the characteristics of people who live there, highlighting those areas that have a high potential for football supporters.

In order to achieve this, a number of objectives will be addressed:

- Analyse the literature on geodemographics to understand its importance
- Collect lifestyle and socio-economic data from census to cluster areas based on what characteristics are common in football fans
- Propose a framework to identify potential areas where football supports would most likely be located
- Map areas within Greater London using the rankings created in the index
- Produce thematic maps to show any patterns that may appear

- Analyse how the classification can be used to benefit local football clubs
- Validate the classification by obtaining Twitter data from football fans in Greater London to see the key areas in which they are living
- Compare this to the classification to evaluate whether it can be used to highlight areas with a high potential for football fans

Methodology

The first stage of this research project will involve continuing with the research into the existing literature surrounding geodemographics. By developing a literature review with a detailed overview of the use of geodemographics and the creation of indexes, the new index will prove to be successful and worthwhile within this field.

Literature allows a consideration of what variables are appropriate to use and it is important to think about what features best distinguish between areas (Voas & Williamson, 2009). The variables that are currently going to be used to classify football fans have been taken from general research about census input into classifications, as well as data from YouGov (YouGov, 2016). From the census, each area will be ranked in terms of percent of males, percent of those aged between 25 and 39, and the percent of those of no religion. Other socio-economic variables that will be included are the percent with £125 and £499 of disposal income per month, and what political party is most common in that particular area. As well as these traditional variables, how many watch Match of the Day, as well as how many own a Sky Sports account, will be used, highlighting the people who have an interest in football and will attend a game.

Following on from the initial research, the next stage will be to collect the lifestyle and socio-economic data from the census and other sources. This data will then be used to by create a spatial index by ranking areas in London based on the geodemographics of the neighbourhoods though the chosen variables. For each Lower Layer Super Output Area (LSOA) in Greater London, the data for the variables will be converted into a table, where rows represent the output area and the columns represent a variable. A score for each LSOA will then be calculated, from which each area can be ranked. As a result, the difference between areas of London will be identified, as those that are higher ranked will have a greater likelihood of people attending football games.

Once the classification has been created, MapInfo will be used to map the results from the index to see the patterns more clearly. This is the most appropriate software to use as attribute data and geographical data can be mapped together. This will allow users of the index to visually identify key spatial patterns created by the use of geodemographics through a combination of GIS functions and the classification (Harris et al, 2005).

The final stage will focus on the validation of the index. This will be done by obtaining the data from one football club in London, and looking at the people who buy their season tickets. By mapping where these people live, it will be examined whether there is a correlation between their residential locations and their likelihood of being a supporter. Furthermore, further validation can be carried out by analysing Twitter data from people in London who are using social media to highlight their interest in football.

Timetable of Dissertation

	Sep	Oct	Nov	Dec	Jan	Feb	Mar
Read Literature							
Detailed Plan							
Collect Data							
Literature Review							
Writing							
Creation of Maps							
Classification							
Analysis of Data							
Intro & Conclude							
Finalisation							

Reference List

Abbas J, Ojo A and Orange S. (2009). 'Geodemographics – a tool for health intelligence?'. Public Health. 123. pp.35-39.

Adnan, M et al. (2010). 'Towards real-time geodemographics: clustering algorithm performance for large multidimensional spatial databases'. Transactions in GIS. 14(3). pp.283-297.

Baçao F Lobo V and Painho M. (2004). 'Clustering Census Data: Comparing the Performance of Self-Organising Maps and K-means Algorithms'. KDNet Symposium, Bonn, Germany. [Online] [Accessed on 22nd October 2016] Available from: http://www.novaims.unl.pt/labnt/geosom/Public/o8-2_11_lobo04_som_kmeans_kdnet.pdf

Batey, P and Brown, P. (1995).'From human ecology to customer targeting: the evolution of geodemographics'. In: Longley, P and Clarke, G, eds, *GIS for Business and Service Planning*. London: Taylor & Francis. pp.115-144.

Birkin, M. (1995). 'Customer targeting, geodemographics and lifestyle approaches'. In: Longley, P and Clarke, G, eds, *GIS for Business and Service Planning*. London: Taylor & Francis. pp.104-149.

Burrows, R. (2008). 'Geodemographics and the construction of differentiated neighbourhoods'. In: Flint, J and Robinson, D, eds, *Community Cohesion in Crisis? New Dimensions of Diversity and Difference*. [Online] [Accessed on 10th October 2016] Available from: http://policypress.universitypressscholarship.com/view/10.1332/policypress/9781847420244.001.0001/upso-9781847420244-chapter-12

Butler, T and Hamnett, C. (2007). 'The Geography of Education: Introduction'. Urban Studies. 44(7). pp.1161-1174.

Chainey, S. (2005). '3rd National Crime Mapping Conference' – Jill Dando Institute of Crime Science, 12th-13th April 2005. [Online] [Accessed on 15th October 2016] Available from: https://www.ucl.ac.uk/jdi/events/mapping-conf/conf-2005/conf2005-downloads/spencer-chainey-crime-mapping-guide.pdf

Critchlow, A. (2015). 'Britain's sport industry hitting top form' – The Telegraph, 3rd May 2015. [Online] [Accessed on 23rd October 2016] Available from: http://www.telegraph.co.uk/finance/11580773/Britains-sport-industry-hitting-top-form.html

Harris, R. (2003). 'Population mapping by geodemographics and digital imagery', In: Mesev, V, ed*, Remotely Sensed Cities*. London: Taylor & Francis. pp.223-241.
Harris, R. (2009). 'Geodemographics'. In: Gregory, D et al, 5th ed*, The Dictionary of Human Geography.* Singapore. Wiley: Blackwell. pp.276-277.

Leventhal, B. (2016). '*Geodemographics for Marketers'*. Croydon: KoganPage.

Singleton, A and Longley, P. (2009). 'Geodemographics, visualisation, and social networks in applied geography'. Applied Geography. 29. pp.289-298.

Sleight, P. (2004). '*Targeting Customers: How to Use Geodemographic and Lifestyle Data in Your Business*'. 3Rd Edition. Henley-on-Thames: World Advertising Research Centre.

Sport England. (2013). 'Economic value of sport in England' – Sport England, July 2013. [Online] [Accessed on 23rd October 2016] Available from: https://www.sportengland.org/media/3174/economic-value-of-sport-summary.pdf

Voas, D and Williamson, P. (2001). 'The diversity of diversity: a critique of geodemographic classification'. Area. 33(1). pp.63-76.

Webber, R and Butler, T. (2007). 'Classifying Pupils by Where They Live: How Well Does This Predict Variations in Their GCSE Results?' Urban Studies. 44(7). pp.1229-1253.

Further references to be used in dissertation

Adnan M, Lansley G and Longley P. (2012). 'Twitter geodemographic analysis of ethnicity and identity in Greater London'. University College London, Department of Geography. [Online] [Accessed on 12th March 2016] Available from: http://www.geocomputation.org/2013/papers/51.pdf

Baker K, Bermingham K and McDonald C. (1979). 'The utility to market research of the classification of residential neighbourhoods'. Proceedings of the Market Research Society Conference, Brighton. pp. 253-271.

Birkin M, Clarke M and Clarke G. (2002). 'Retail Geography and Intelligent Network Planning'. Chichester: Wiley.

Donnay JP and Unwin, D. (2001). 'Modelling geographical distributions in urban areas. In: Donnay JP, Barnsley M and Longley P, eds, *Remote Sensing and Urban Analysis*. London: Taylor & Francis. pp.205-224.

Harris, R. (1999). 'Geodemographics and geolifestyles: a comparative review'. Journal of Targeting, Measurement and Analysis for Marketing'. 8. pp.164-78.

Harris R, Sleight P and Webber R. (2005). '*Geodemographics, GIS and Neighbourhood Targeting*'. Chichester: Wiley

Kearns, A and Parkinson, M. (2001). 'The significance of neighbourhoods'. Urban Studies. 12. pp.2103-2110.

Leventhal, B. (1995). 'Evaluation of Geodemographic Classifications'. Journal of Targeting, Measurement and Analysis for Marketing. pp.1-11.

Martin D, Cockings S and Harfoot A. (2013). 'Development of a geographical framework for census workplace data'. Journal of the Royal Statistical Society'. 176(2). pp.585-602

Moy, C and Meier, E. (2004). 'Social grading and the Census'. International Journal of Market Research. 46(2). pp.2-29.

Openshaw, S. (1989). 'Making geodemographics more sophisticated'. Journal of the Market Research Society. 31. pp.111-131.

Peterson, J et al. (2009). 'Geodemographics as a tool for targeting neighbourhoods in public health campaigns'. J Geogr Syst. 13. pp.173-192.

Reibel, M. (2011). 'Classification Approaches in Neighbourhood Research: Introduction and Review'. Urban Geography. 32(3). pp.305-316.

Sheffield Hallam University. (2016). 'Sport Industry Research Centre' – Sheffield Hallam University'. [Online] [Accessed on 12th March 2016] Available from: https://www.shu.ac.uk/research/specialisms/sport-industry-research-centre

Singleton, A and Spielman, S. (2014) 'The Past, Present and Future of Geodemographic Research in the United States and United Kingdom'. The Professional Geographer. 66(4). pp.558-567.

Thompson, D. (2014). 'Which sports have the whitest/richest/oldest/fans?' – The Atlantic, 10th February 2014. [Online] [Accessed on 12th March 2016] Available from: http://www.theatlantic.com/business/archive/2014/02/which-sports-have-the-whitest-richest-oldest-fans/283626/

APPENDIX E: SECOND INTERIM REPORT

GEOG3600 – Dissertation Second Interim Report

Title

Creating a geodemographic index to highlight areas of Greater London with a high potential for supporters of their local football club

Aims and Objectives

The overarching aim of the project is to create a geodemographic index using the characteristics of people, to identify areas of Greater London that have a high number of potential football supporters.

To achieve this, a number of objectives will be addressed:

- Analyse the surrounding literature on geodemographics and football to understand their importance
- Use the literature to determine the values that are synonymous with football attendances that can be used within the index
- Collect lifestyle and socio-economic data to create an index that will identify areas of high potential for football support within Greater London
- Apply the index to one club, Bromley FC, to analyse how the index can be used to benefit local football clubs by targeting their marketing to certain areas
- Validate the index by obtaining data on season ticket holders of Bromley FC and identifying what areas they tend to live in
- Analyse the advantages this can have to football clubs by increasing the support

Methodology

Determining Variables for the Index

Variable	Justification
Age 18-44	Out of 2000 people interviewed, interest in football was most common amongst those aged 23-34 (Ofcom, 2005)
Male	Out of 2000 people interviewed, interest in football was most common amongst men (Ofcom, 2005) 77% of 5.2 million fans in 2011/2012 were male (White, 2013)
No Religion	Typical characteristics of football fans (YouGov, 2016)
White British	Typical characteristics of football fans (YouGov, 2016) Only 8% of Premier League fans from ethnic minorities (BBC News, 2010)

Income	Out of 2000 people interviewed, interest in football was most common amongst those in highest social group, (Ofcom, 2005), thus having high level of disposable income Need disposable income to purchase season tickets Economic variables are highly significant to determine football match attendance (Garcia and Rodriguez, 2002)
Access to Car	Need a mode of transport in order to attend the games Importance of car use, 43% travel by car (Campaign for Better Transport, 2013)
Distance to Club	Importance of distance when considering the changing dynamics of football attendances (Bale, 2000) Need to be close by football ground to have a connection with the identity of the club (Supporters Direct, 2010)
Distance to Other Clubs	Importance of distance when considering the changing dynamics of football attendances (Bale, 2000) If there is too much competition around, from other clubs, then will be less likely to attend

Prior research highlighted other variables that would be appropriate for this purpose. Those who watch programmes such as Match of the Day are more likely to have an interest in football (YouGov, 2016), and thus this could have been an additional variable to include. However, it is not viably possible to collect this data for this research project, as the information is not easily accessible. Other data that could have proved useful is the impact that televised premier league matches has on local attendances (Buraimo, 2008), however as mentioned previously, it is not appropriate within the time scale.

Creation of the Index

The method of creating a geodemographic index varies depending on the purpose, as they should be specific to the end users, ensuring they are accessible and easy to understand (Singleton and Longley, 2009). Despite differences, the main goal of any index is to organise data into meaningful structures (Bacao et al, 2004). This index will be created using the multivariate classification technique, which involves numerical methods to group areas together (Reibel, 2011). Characteristics of people within LSOAs in London will be combined to rank areas in their likelihood of having a high majority of potential football supporters. Instead of mapping areas and then looking at the characteristics of the areas, creating an index through these stages makes it transferable across the UK. If the index proves successful, then it can be utilised by other football clubs in London, as well as across the country, through simple modifications.

Standardisation

After obtained, the data was obtained, the raw data was converted into percentages, allowing a standardisation of the data so that all values for every LSOA runs on a scale between 1 to 100. To achieve this, each census value was divided by the total number of people within that LSOA, and then multiplied by 10. An exception to this was the access to car variable, which was instead divided by total number of households, due to the differences in the way the data is presented within the census. The disposable income variable did not need modification, as when downloaded it had already been converted to a value per person in

each LSOA. Similarly, the distance calculations did not need standardising, as this data does not need to be presented in percentages in the same way as demographic variables due to the fact there is no obvious denominator.

Multicollinearity

When variables are included in the index, if there are two that exhibit similar results, then it may not be necessary to include them both. The data analysis technique of correlation can be carried out in Excel to determine the similarity of the chosen variables. If high correlations are observed of a value above 0.7, then one characteristic can be excluded. When this stage was executed, no high correlations were observed, and thus the decision was made to keep all variables that were initially chosen, however in the work of geodemographic index creation, it is an important aspect to consider.

Normalisation

The next stage of the process was the standardisation of the data through employing normalisation in order to linearly re-scale each variable to the same scale of between 0 and 1. Although the conversion to percentages adopted some degree of standardisation, the variable percentages will adopt different ranges, as some may go from 20 to 50% rather than all ranging from 0 to 100%. The process of normalisation allows us counter this by ensuring that the variation of the variables is captured equally, where the lowest value for each variable is 0, and the highest being 1, with every other value fitting in proportion. The equation inputted into Excel is as follows:

$$x_{norm} = (x_{raw} - min_1)/(max_1 - min_1)$$

x_{raw} *is the number being rescale*d
min_1*is the minimum number of the whole column*
max_1 *is the maximum number of the whole column*

Polarity

Once the data was completely standardised, the next stage in creating the index was to determine the polarity of the data, where the direction of each data set is considered. Each variable included in the index needs to have values that mean that a high score is a positive aspect of the LSOA, hence a score of one for each variable will be desirable for an area to have a high potential for football supporters. Before this stage, every variable displayed this scale of 1 being the most preferable apart from the distance to football club variable. Due to this, the data was 'flipped' to run in the same direction, so that a value of 1 would be the area closest to Bromley FC, and thus a positive aspect of the index. This was achieved by subtracting the normalised value from 1, meaning the scale for each variable is reversed, and thus every variable in the index is going in the same direction. With the other distance variable, this stage was not necessary, because a high value means an area is further from other football clubs, thus are more likely to attend a match of the football club being targeted, which in this case is Bromley FC.

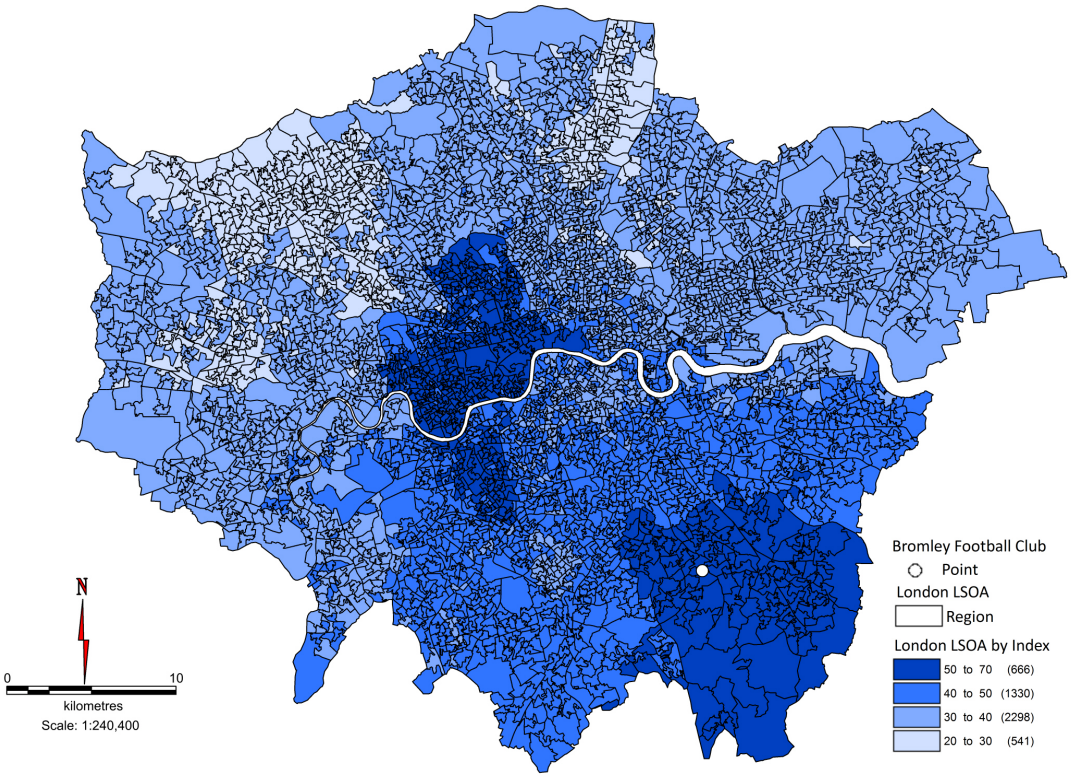
Weighting

Once all the data is processed, the decision needs to be made on what weighting to give each variable within the index. When electing the weightings, a reflection on the literature studied highlights that income and distance are the most important variables included in this index. On the other hand, no religion and white British are the least important, as it can be argued that they are the most generalised. When deciding on the weightings, the results of the multicollinearity stage can be used to see if there are two variables that although are both important enough to be included in the index, could have smaller weightings. Age and access to car have the strongest correlation between any of the variables, and thus can have lower weightings than income and distance.

When the final weightings were confirmed, each normalised value in the index is multiplied by the weighting given to that variable. For each LSOA, the weightings of all the variables should combine to equal 100% of the score. However, the decision on weightings for the variables is another example of where subjectivity is involved in the process of creating an index. The importance of each variable is based on the beliefs of the person creating the index, and their interpretation of the literature.

Findings and Analysis

The creation of this geodemographic index allows a comparison of LSOAs in Greater London, with the purpose of identifying specific areas that have the highest percentage of possible supporters of Bromley FC. To ascertain the areas that will be most suitable for the football club to target, mapping of the final index scores permits a spatial evaluation of the whole study area. The higher the index score assigned to a LSOA, the more likely it is that people living within that area will attend a Bromley FC match, due to the socio-economic and demographic characteristics of the population.



The dark blue areas indicate a LSOA that has a higher potential for supporters of Bromley FC, as a result of a higher index score than the lighter areas. An obvious area of high index scores is in the borough of Bromley. The highest weighting within the index was given to distance to football club, meaning that areas surrounding the location of the football club were going to score highly, and this is reflected in the choropleth map. The findings of this reflect the literature explicitly stating distance as a vital factor in determining football match attendance (Bale, 2000).

Table: Top 20 ranked LSOAs within the geodemographic index

GEO_CODE	GEO_LABEL	Index Score
E01004736	Westminster 018C	60.735
E01032739	City of London 001F	60.395
E01032740	City of London 001G	60.027
E01002838	Kensington and Chelsea 020A	59.966
E01004688	Westminster 019B	59.380
E01000851	Camden 026B	59.134
E01000668	Bromley 020B	59.097
E01004622	Wandsworth 024D	58.976
E01000679	Bromley 020E	58.747
E01000648	Bromley 016C	58.690
E01004763	Westminster 013B	58.588
E01001913	Hammersmith and Fulham 014C	58.481
E01000823	Bromley 008E	58.405
E01000680	Bromley 037A	58.242
E01004733	Westminster 020C	58.229
E01000647	Bromley 016B	58.158
E01000697	Bromley 007F	58.148
E01001937	Hammersmith and Fulham 025C	58.139
E01001926	Hammersmith and Fulham 024C	58.039
E01004742	Westminster 022A	57.984

The best performing LSOA is in the borough of Westminster, which is also where four other top ranked LSOAs are located. Other areas that can be highlighted as having a high potential for people attending Bromley FC matches is in the Kensington and Chelsea, Camden, Wandsworth, and Hammersmith and Fulham. One reason for LSOAs in these areas showing their potential for football supporters is due to their disposable income. Those with the money to attend football matches regularly have been identified as being more likely to attend Bromley FC matches (Garcia and Rodriguez, 2002), as they have greater ability to do this compared to those who earn less money than them.

Dissertation Structure

Title Page
Acknowledgements
Abstract
List of Figure and Tables

Chapter 1: Introduction
- Background
- Aims and Objectives
- Dissertation Outline

Chapter 2: Literature Review
- What is Geodemographics?
- Evolution of Geodemographics
- Importance of Neighbourhoods
- Uses of Geodemographics
- Benefits to Geodemographics
- Negatives of Geodemographics
- Football and Sport

Chapter 3: Methodology
- Study Area
- The Use of GIS
- Determining Variables for the Index
- Data Collection
- Creation of the Index
- Creation of the Maps

Chapter 4: Results
- Maps and Patterns of the Index
- Discussion of Index
- Validation of the Index

Chapter 5: Conclusion
- Overall Findings
- Implications and Future Work
- Limitations to Research

Bibliography

Appendices
- Extract of in the Index
- DSG Forms
- DSG Reflective Log
- Interim Reports
- Risk Assessment Form

Outstanding Work

- Introduction, including the outline of the dissertation
- Finishing results and analysis section – expanding on what has already been discussed and relating it to the wider literature
- Conclusion, including overall findings, implications, and limitations
- Contents, title page, list of figures, appendices,

Reference List for Dissertation

Abbas J, Ojo A and Orange S. (2009). 'Geodemographics – a tool for health intelligence?'. Public Health. 123. pp.35-39.

Acorn. (2013). 'Acorn Classification'. [Online] [Accessed on 16th January 2017] Available from: http://acorn.caci.co.uk

Adnan, M et al. (2010). 'Towards real-time geodemographics: clustering algorithm performance for large multidimensional spatial databases'. Transactions in GIS. 14(3). pp.283-297.

Adnan M, Lansley G and Longley P. (2012). 'Twitter geodemographic analysis of ethnicity and identity in Greater London'. University College London, Department of Geography. [Online] [Accessed on 12th March 2016] Available from: http://www.geocomputation.org/2013/papers/51.pdf

Anderson, T. (2010). 'Using geodemographics to measure and explain social and environment differences in road traffic accident risk'. Environment and Planning A. 42. pp.2186-2200.

Baçao F Lobo V and Painho M. (2004). 'Clustering Census Data: Comparing the Performance of Self-Organising Maps and K-means Algorithms'. KDNet Symposium, Bonn, Germany. [Online] [Accessed on 22nd October 2016] Available from: http://www.novaims.unl.pt/labnt/geosom/Public/o8-2_11_lobo04_som_kmeans_kdnet.pdf

Baker K, Bermingham K and McDonald C. (1979). 'The utility to market research of the classification of residential neighbourhoods'. Proceedings of the Market Research Society Conference, Brighton. pp. 253-271.

Bale, S. (2000). 'The changing face of football: stadiums and communities.' Soccer and Society. 1(1). pp.91-101.

Batey, P and Brown, P. (1995).'From human ecology to customer targeting: the evolution of geodemographics'. In: Longley, P and Clarke, G, eds, *GIS for Business and Service Planning*. London: Taylor & Francis. pp.115-144.

BBC News. (2010). 'More diverse crowds at Premier League football' – BBC News, 25th August 2010. [Online] [Accessed on 28th January 2017] Available from: http://www.bbc.co.uk/news/uk-11079597

BBC Sport. (2014). 'Price of Football: Ticket Increases Outstrip Cost of Living' – BBC Sport, 15th October 2015. [Online] [Accessed on 24th January 2017] Available from: http://www.bbc.co.uk/sport/football/29614980

Birkin, M. (1995). 'Customer targeting, geodemographics and lifestyle approaches'. In: Longley, P and Clarke, G, eds, *GIS for Business and Service Planning*. London: Taylor & Francis. pp.104-149.

Bromley FC. (2017). 'Bromley FC Website'. [Online] [Accessed on 26th January 2017] Available from: http://www.bromleyfc.tv/site/

Buraimo, B. (2008). 'Stadium Attendance and Television Audience Demand in English League Football'. Managerial and Decision Economics. 29(6). pp.512-523.
Buraimo B, Forrest D, and Simmons R. (2009). 'Insights for clubs from modelling match attendance in football'. Journal of the Operational Research Society'. 60(2). pp.147-155.

Burrows, R. (2008). 'Geodemographics and the construction of differentiated neighbourhoods'. In: Flint, J and Robinson, D, eds, *Community Cohesion in Crisis? New Dimensions of Diversity and Difference*. [Online] [Accessed on 10th October 2016] Available from: http://policypress.universitypressscholarship.com/view/10.1332/policypress/9781847420244.001.0001/upso-9781847420244-chapter-12

Burrows, R. (2013). 'The New Gilded Ghettos: The Geodemographics of the Super-Rich'. Discover Society. 3. [Online] [Accessed on 12th January 2017] Available from: http://discoversociety.org/2013/12/03/the-new-gilded-ghettos-the-geodemographics-of-the-super-rich/

Butler, T and Hamnett, C. (2007). 'The Geography of Education: Introduction'. Urban Studies. 44(7). pp.1161-1174.

Campaign for Better Transport. (2013). 'Door to Turnstile, Improving Travel Choices for Football Fans'. [Online] [Accessed on 28th January 2017] Available from: http://www.bettertransport.org.uk/sites/default/files/research-files/Door_to_Turnstile_CfBT_FINAL_web.pdf

Chainey, S. (2005). '3rd National Crime Mapping Conference' – Jill Dando Institute of Crime Science, 12th-13th April 2005. [Online] [Accessed on 15th October 2016] Available from: https://www.ucl.ac.uk/jdi/events/mapping-conf/conf-2005/conf2005-downloads/spencer-chainey-crime-mapping-guide.pdf

Critchlow, A. (2015). 'Britain's sport industry hitting top form' – The Telegraph, 3rd May 2015. [Online] [Accessed on 23rd October 2016] Available from: http://www.telegraph.co.uk/finance/11580773/Britains-sport-industry-hitting-top-form.html

Donnay JP and Unwin, D. (2001). 'Modelling geographical distributions in urban areas. In: Donnay JP, Barnsley M and Longley P, eds, *Remote Sensing and Urban Analysis*. London: Taylor & Francis. pp.205-224.

Experian. (2017). 'MOSAIC Geodemographic Classification System'. [Online] [Accessed on 17th January 2017] Available from: http://www.experian.co.uk/marketing-services/products/mosaic-uk.html

Furness, P and Sleight, P. (2013). 'Tracking a Decade of Changing Britain' – Conference Report, 18th December 2013. [Online] [Accessed on 23rd October 2016] Available from: https://www.mrs.org.uk/pdf/Chairmen's%20Conference%20Report%20061113.pdf

Garcia, J and Rodriguez, P. (2002). 'The Determinants of Football Match Attendance Revisited: Empirical Evidence from the Spanish Football League'. Journal of Sports Economics. 3(1). pp.18-38.

GLA Intelligence. (2016). 'GLA 2015 round trend-based projections: Results'. – Greater London Authority, July 2016. [Online] [Accessed on 26th January 2017] Available from: https://files.datapress.com/london/dataset/2015-round-population-projections/2016-07-19T13:06:46/2015%20trend%20projection%20results.pdf

Harris, R. (2003). 'Population mapping by geodemographics and digital imagery', In: Mesev, V, ed*, Remotely Sensed Cities*. London: Taylor & Francis. pp.223-241.

Harris, R. (2009). 'Geodemographics'. In: Gregory, D et al, 5th ed*, The Dictionary of Human Geography.* Singapore. Wiley: Blackwell. pp.276-277.

Harris R, Sleight P and Webber R. (2005). '*Geodemographics, GIS and Neighbourhood Targeting*'. Chichester: Wiley

Hidden London. (2005). 'A history of London football'. [Online] [Accessed on 26th January 2017] Available from: http://hidden-london.com/miscellany/football/history/

Inman J, Shankar V and Ferraro R. (2004). 'The Roles of Channel-Category Associations and Geodemographics in Channel Patronage'. Journal of Marketing. 68(2). pp.51-71

Kamarulzaman, Y. (2010). 'Geodemographics of Travel E-shoppers: An Empirical Analysis of UK Consumers'. European Journal of Social Sciences'. 16(2). pp.195-205

Kearns, A and Parkinson, M. (2001). 'The significance of neighbourhoods'. Urban Studies. 12. pp.2103-2110.

Leventhal, B. (1995). 'Evaluation of Geodemographic Classifications'. Journal of Targeting, Measurement and Analysis for Marketing. pp.1-11.

Leventhal, B. (2016). '*Geodemographics for Marketers'*. Croydon: KoganPage.

Neave, N and Wolfson, S. (2003). 'Testosterone, territoriality and the 'home advantage' Physiology and Behaviour. 78(2). pp.269-275.

Ofcom. (2005). 'Premier League Football: Research Into Viewing Trends, Stadium Attendances, Fans' Preferences and Behaviour, and the Commercial Market' – Commission of

the European Communities. [Online] [Accessed on 13th January 2017] Available from: http://ec.europa.eu/competition/antitrust/cases/dec_docs/38173/38173_104_7.pdf

Peterson, J et al. (2009). 'Geodemographics as a tool for targeting neighbourhoods in public health campaigns'. J Geogr Syst. 13. pp.173-192.

Pointer, G. (2005). 'The UK's Major Urban Areas' – Chapter 3, Focus on People and Migration. [Online] [Accessed on 26th January 2017] Available from: https://www.ons.gov.uk

Reibel, M. (2011). 'Classification Approaches in Neighbourhood Research: Introduction and Review'. Urban Geography. 32(3). pp.305-316.

Simmons, R. (1996). 'The demand for English league football: a club-level analysis'. Applied Economics. 28(2). pp.139-155.

Singleton, A. (2004). 'A State of the Art Review of Geodemographics and their Applicability to the Higher Education Market'. Working Paper Series. [Online] [Accessed on 13th January 2017] Available from: http://discovery.ucl.ac.uk/212/1/paper74.pdf

Singleton, A and Longley, P. (2009). 'Geodemographics, visualisation, and social networks in applied geography'. Applied Geography. 29. pp.289-298.

Singleton, A and Spielman, S. (2014). 'The Past, Present and Future of Geodemographic Research in the United States and United Kingdom'. The Professional Geographer. 66(4). pp.558-567.

Sleight, P. (2004). '*Targeting Customers: How to Use Geodemographic and Lifestyle Data in Your Business*'. 3Rd Edition. Henley-on-Thames: World Advertising Research Centre.

Soberg, H and Mehus, I. (2014). 'The Challenge of Attracting Football Fans to Stadia?' International Journal of Sport Finance. 9(1). pp.3-19

Sport England. (2013). 'Economic value of sport in England' – Sport England, July 2013. [Online] [Accessed on 23rd October 2016] Available from: https://www.sportengland.org/media/3174/economic-value-of-sport-summary.pdf

Supporters Direct. (2010). 'The Social and Community Value of Football' – Summary Report June 2010. [Online] [Accessed on 28th January 2017] Available from: http://www.supporters-direct.org/wp-content/uploads/2012/08/svoff-summary-report.pdf

Tapp, A and Clowes, J. (2002). 'From "carefree casuals" to "professional wanders". Segmentation possibilities for football supporters'. European Journal of Marketing. 36(11/12). pp.1248-1269.

Thompson, D. (2014). 'Which sports have the whitest/richest/oldest/fans?' – The Atlantic, 10th February 2014. [Online] [Accessed on 12th March 2016] Available from: http://www.theatlantic.com/business/archive/2014/02/which-sports-have-the-whitest-richest-oldest-fans/283626/

Voas, D and Williamson, P. (2001). 'The diversity of diversity: a critique of geodemographic classification'. Area. 33(1). pp.63-76.

Webber, R and Butler, T. (2007). 'Classifying Pupils by Where They Live: How Well Does This Predict Variations in Their GCSE Results?' Urban Studies. 44(7). pp.1229-1253.

White, J. (2013). 'Football grounds aren't full of old white men after all' – The Telegraph, 10th January 2013. [Online] [Accessed on 28th January 2017] Available from: http://www.telegraph.co.uk/sport/football/9793689/Football-grounds-arent-full-of-old-white-men-after-all-....html

Wood, Z and Allen, K. (2016). 'Where have all Britain's shoppers gone?' – The Guardian, 23rd April 2016. [Online] [Accessed on 16th January 2017] Available from: https://www.theguardian.com/business/2016/apr/23/uk-high-street-spending-shoppers-slowdown

APPENDIX F: RISK ASSESSMENT APPROVAL

Sarah Burdall
Dissertion Risk Assessment -BA
To: Georgina Weaver [gy14gw]

Inbox - Exchange 11:04

SB

Dear Georgina

Thank you for completing the risk assessment for your dissertation. Due to some technical difficulties please print and insert this email into your dissertation (instead of a copy of your risk assessment) as proof of an approved risk assessment.

Best Wishes
Sarah

Sarah Burdall
Health and Safety Manager (Faculty of Environment) Environmental Compliance Manager Biological Safety

Health and Safety Services
5-9 Willow Terrace Road
University of Leeds, LS2 9JT

0113 3438042 | 0750 6707339 | s.e.burdall@leeds.ac.uk