Leeds University
Business School

UNIVERSITY OF LEEDS

Dissertation/Project Coversheet

Student ID Number:	2	0	1	7	9	4	1	8	3
Student Name	Jayakumar Sukumar								
Module Code:	LUBS5579M								
Programme of Study:	Business Analytics and Decision Sciences Dissertation								
Supervisor:	Wessam Abouarghoub								
Title:	Factors affecting predictive analysis in forecasting election outcomes								
Declared Word Count:	8153								

Please Note:

Your declared word count must be accurate, and should not mislead. Making a fraudulent statement concerning the work submitted for assessment could be considered academic malpractice and investigated as such. If the amount of work submitted is higher than that specified by the word limit or that declared on your word count, this may be reflected in the mark awarded and noted through individual feedback given to you.

It is not acceptable to present matters of substance, which should be included in the main body of the text, in the appendices ("appendix abuse"). It is not acceptable to attempt to hide words in graphs and diagrams; only text which is strictly necessary should be included in graphs and diagrams.

By submitting an assignment you confirm you have read and understood the University of Leeds Declaration of Academic Integrity (http://www.leeds.ac.uk/secretariat/documents/academic_integrity.pdf).

ACKNOWLEDGEMENTS

I'd want to offer sincere gratitude to various people whose encouragement and direction made this dissertation possible.

First and foremost, I would like to express my gratitude to my supervisor, Mr. Wessam Abouarghoub, for his important advice, experience, and ongoing encouragement throughout this project. His observations and criticism were critical in influencing the course of my research, and his patience and dedication guided me through the difficulties of this subject.

Finally, I'd want to thank all the participants and professionals whose inputs enhanced my research and provided a larger perspective on the subject.

Thank you all.

ABSTRACT

This dissertation critically investigates the determinants of accuracy in models designed to predict electoral results in the United Kingdom, especially election outcome prediction models, as more specifically focuses on three hardcore problems: How Brexit has influenced the way people vote, How political discussions are facilitated by social media, and How the public health crisis like the COVID-19 check on the people's faith in the official bodies. These common group characteristics structure the electoral prediction process regarding the settlements accessed through the special survey and the application of multivariate statistics logistic regression and random forests. The outcomes demonstrate that the loyalty of people to the political parties and their political association has transformed to a new level on the part of Brexit and hence makes it difficult for the normal polls. Using social media as a communication tool is also knocking off the younger voters which presents new challenges and new opportunities for election forecasting. Additionally, the COVID-19 pandemic has brought home how critical public health management is about creating voter intention and voter confidence in the government.

This research with an increasing complexity of the political environment stresses that there is a need for the development of adequate approaches which consider these changing factors if there is to be any improvement on the accuracy of forecasting elections.

Table of Contents

CHAPTER 1

1.1 INTRODUCTION

In an age of extreme political volatility, it is increasingly harder to make accurate predictions regarding the election outcomes. The political forecasts do not always hold true in practice due to the emergence of these new types of problems. All this has not only transformed the UK political scene but also made evident the flaws of conventional election prognostication tools, with one of the most significant events in the recent past: the Brexit referendum – illustrating this point (Prosser, 2020). The political forecasting models which enjoyed some level of reliability in the past are now being contradicted due to developments in society and technology. Such unpredictability was forcefully illustrated in 2016 by the Brexit referendum, an event that emphasized the unworthiness of standard models of election predictions (Prosser, 2020). The outcome of the vote showed just how much the public's feelings about the vote can differ from the predictions of many experts, thereby revealing the challenges confronting models based on historical data in the face of changing political and social environments (Clarke et al. 2017). Polls about Brexit tended to fluctuate and be intangible, with several polls conducted before the referendum being drastically off the mark, including the result, pointing to an area of difficulty in understanding why voters make decisions in fierce political environments (Curtice, 2017).

The rapid change and perhaps even more pronounced is the effect of social media on voting patterns, particularly among young populations, adds usefulness as well as complexity to the efforts of those who seek to accurately predict election outcomes (Loader & Mercea, 2011). In addition, the social mediascape has enhanced the already existing political complexities. Especially with the younger voters, the use of platforms such as Facebook, Twitter and Instagram are important determinants of political attitudes and voting (Stier et al., 2018). More disturbingly, as time has passed Loader and Mercea have noted, the use of such tools has not only brought new tools that political actors can use in configuring their campaigns but instead makes it more difficult to accurately predict elections (Loader & Mercea, 2011). The

phenomenon of so called ‘echo chambers’, whereby only content reinforcing previously held beliefs is presented to the audience, also degrades public opinion and polling (Flaxman et al. 2016). Besides, information whether true or false coming out of these networks has also greatly affected how voters conduct themselves thereby introducing a higher level of complexity to predicting outcomes (Allcott & Gentzkow, 2017).

Such prediction scenario is further complicated by the COVID-19 pandemic which has brought deep-seated concerns over public health to the fore of political attention. Questions about how services provided by public authorities including the National Health Service (NHS), affect individuals’ intentions to vote have also been brought forth by the pandemic, which has also shaken the public trust in the government institutions (Fancourt et al.2020). New layers of complexities have arisen with the emergence of the COVID-19 pandemic, which brought health issues to the middle of political debates. The pandemic has changed the playbook on how the voters think, most importantly, on the views of citizens towards the effectiveness of government agencies like the NHS (Devine et al., 2021). Fancourt cited that the crisis has led to a great number of people reconsidering the extent to which they have faith in the government, which is an important determinant on whether one will vote. Rae found that during the health crisis, the government’s actions concerning the pandemic regard were essential to the general trust of the public towards the government and their voting intentions, elicits either confidence or disappointment towards the leaders, signaling increased or reduced support respectively (Rae et al. 2021). This pattern is common in many countries, which goes to show how, among other things crises such as pandemics can become issues during electoral campaigns altering the priorities of voters and their trust in public officials (Jennings et al., 2021).

Thus, this research has set out to examine the factors that influence the success of predictive models of elections, by looking at three areas of interest: how Brexit has affected voter’s party coherence; How social media has been used to shape political attitudes and voting behavior; and how epidemiological threats have influenced citizen’s perception and trust in authorities. By examining these features, this study looks to enhance understanding about the changing

factors that may be considered more carefully to increase the efficiency of the forecasting of elections in the present time.

1.2 RESEARCH OBJECTIVES

This research is aimed to analyse the factors concerning the prediction of the elections in relation to the recent events in the United Kingdom. It is quite clear that this study does not intend to be concerned with just how predictive opinions went wrong in predicting the outcome but, rather what went wrong in the voters' behavior. Moreover, it attempts to address the impact of social media in shaping the political views and voting behaviors of young people in the context of the UK. In addition, it shall deal with how public health crisis like the COVID-19 pandemic affect people's perception of government authority and whether or how this perception makes its way into various polling results. In such a manner, the literature review section will reasonably feature the difficulties incurred in getting effective approaches of campaign forecasting and suggest ways in which the modeling methodologies of predicting outcomes of elections in modern UK can be refined.

CHAPTER 2

2.1 LITERATURE REVIEW

2.1.1 BREXIT – IMPACT ON VOTER ALIGNMENT AND PARTY LOYALTY

The choice made by the United Kingdom to withdraw from the European Union which is popularly known as Brexit has significantly altered British politics and the trends in voting patterns as well as party loyalties with very damaging effects on election forecasts too. The referendum which came up in June in the year 2016 was able to demonstrate very clear divisions inside the British electorate by breaking the voters into two blocs, namely Leavers, who were in support of being out of the European Union and the Remainers who were advocates for staying. Together this division has had serious implications on party politics and electoral behavior consequently shaping perceptions and altering the political landscape.

In the article by Farrell and Newman, the authors investigate how Brexit has been influential in changing how the electorate votes. Prior to Brexit, parties were related to such diverse variables as social strata economic regions, etc.; now it is EU membership that has become the main marker of subdivision (Farrell & Newman, 2017). Particularly the leading parties have suffered from such a shift. The traditionally “laissez-faire”(free market) Conservative Party now is associated with supporters of Brexit. The Labour party now who has been trying to define its position on Brexit has been faced with internal disputes and difficulties to hold on to its roots while trying to win over new followers.

In support of this, Clarke, Goodwin, and Whiteley contend that Brexit intensified divisions in British politics, aligning voters with pro-EU and anti-EU perspectives, thereby altering party affiliations (Clarke et al. 2017). They highlight that the Conservative Party emerged as the representative for those in favor of leaving the EU, whereas Labour's uncertainty resulted in difficulties in maintaining its fundamental support.

The repercussions of Brexit have significantly affected the reliability of election polling. Dennis (Dennis.J. 2020) points out that conventional polling techniques have had difficulty adapting to

the evolving political landscape. Before the referendum, polls frequently undervalued the support for the Leave campaign, because of challenges in measuring populist feelings and the rising sway of parties such as UKIP. This was further illustrated in the 2019 UK General Election, in which numerous polls exaggerated Labour's backing while misjudging Brexit's effect on voter allocation, resulting in an unexpected Conservative majority.

Wroe and Williams similarly describe how because of Brexit the political environment has become more uncertain, where voters have become less committed to party lines (Wroe & Williams, 2018). Their study shows that Brexit has contributed to increased voter volatility such that individuals sought support regarding membership to the EU rather than political affiliations. Such volatility makes predictions of electoral outcomes quite hard judging from how wrong many opinion polls were during the 2016 referendum and the 2019 General Election.

Goodwin and Heath claim that on the contrary, the Brexit process has deepened the gap between the politicians and general population, complicating the prediction of voters' behavior. They note that Brexit has made people feel even more detached from politics than usual, which in turn, brings about additional difficulties with regards to polling accuracy (Goodwin & Heath, 2016).

To tackle these challenges, survey researchers have had to move up the ladder and employ other regimes. As pointed out by Farrell and Newman, the utilization of live data in the model, being sensitive to changes in demographics and using new statistical techniques to reflect the complexities in the changing political market (Farrell & Newman, 2017). Appreciating as well as contextualizing such variables as the impact of Brexit on voter realignment and electoral support for parties will contribute to better assessments of elections in post-Brexit landscape.

All in all, the Brexit has produced a tremendous change in British political life. It has transformed party systems, increased volatility of the electorate and created problems with traditional forecasting methodologies.

(Hypothesis-1): Brexit has permanently changed how people vote and their party loyalty in the UK.

2.1.2 SOCIAL MEDIA AND ITS INFLUENCE ON YOUNGSTERS

The emergence and expansion of social media platforms have changed the way young voters think about politics and seek to vote. Features provided by websites including Twitter, Facebook, Instagram and Tik Tok permit users to participate in political debates, develop as well as share material and take part in political activities. However, such increase in participation also has negative consequences that include the formation of echo chambers and extreme divisions in politics, and disinformation.

As remarked upon by Vaccari, social media has granted ordinary people the ability to participate in a political system that used to be limited to the elite by allowing people not just to consume but also contribute their own content on politics thus increasing the volume of content than what would be available in mass media (Vaccari, 2018). Younger voters are at an advantage with this wider availability of political content, but the interactive aspect of these sites also helps in quick propagation of ideas and rumors. This also impacts the way that voting patterns can be predicted since what is discussed online is only loosely based on facts and often on more sensationalist elements, since people are inherently emotional (Vaccari, 2018).

The concept of 'echo chambers' is a natural addition to this theory, and Barberá defines it as creating contexts within which users only consume information that corroborates their biases. The younger age group of voters seems to be affected by this effect more than the others as they tend to self-select the material they want to see in their social media accounts. This leads to the phenomenon where the political discourse seems to be on a single sided showing an increasing tendency of a decrease in framework exposure. The consequence of this polarisation is a variable in voter's intentions, which in turn leads to errors in the expectations of voting behavior as people may think everyone else supports those opinions that their circles do (Barberá, 2020).

Boulianne focuses on how certain types of technological platforms channel political behavior differently. Thus, instant commentary and dissemination of political events is mainly carried out via Twitter, while political advertisements in more visual platforms such as Instagram and Tik Tok, often take a lifestyle and identification approach (Boulianne, 2019). These related distinctions in platform use patterns among the younger voters introduce some complexities in trying to analyse

how political attitudes tend to be shaped and how such political attitudes have an impact on voting behavior (Boulianne, 2019).

On the one hand, social networking sites have provided the young constituency with avenues upon which they can engage in political debates. But this has its own negative perspectives, such as cyber bullying and hate speech. According to Awan and Zempi, such hijacking at the elite level has severe consequences for people's participation in social politics, particularly among minority populations. Younger people from impoverished origins might limit themselves out of fear that other users may personally attack them for sharing their political beliefs online. The holding of hands, though pleasurable, may result voices being missed when critic surveys are done, diminishing the ability of the pre-election losers and winners to be appraised (Awan & Zempi, 2016).

The (Bestvater et al, 2023) adds that this is made evident through the significance of social media in political activity of the Millennial generation. It can be by taking part in an internet-based challenges that give this generation a place in global politics. However, there are some counter arguments that the very same media, which is used for the political purposes, is employed for the spreading of false news, which in turn will influence politics and voting. Such opportunities for technological usage led to the emergence of echo chambers which support already existing views amongst users and narrow the range of information accessed by younger voters making their voting behavior hard to predict (Bestvater et al, 2023).

As per Alodat, Al-Qora’n and Abu Hamoud, younger users politically engage, rather emotionally and reactionally through the usage of Instagram and TikTok which are more of appealing platforms due to their more tagline (Alodat et al., 2023). Owing to the visually pleasing and emotionally appealing content on these sites, political fads easily fall into the category of the forwards hence affecting the youth's view of political issues, actors and processes. This type of participation in politics is easy to do and, although enhancing the citizen’s power, it leads to more dangerous reckless decisions when voting (Pavlovkin et al., 2023).

Also, Fletcher regarding the integration of other news sources mentions that social media algorithms within social media practitioners further promote the echo-tainment criticism. Targeting younger voters who rely on social media as a source for political information makes them a target to this because they tend to engage with only the beliefs that they already have. From such a point, people are not exposed to many political thoughts and arguments and hence are limited in forming and direction of their opinion about the specific issues most of the times, rather in an unconventional manner (Fletcher, 2023).

To sum up, social media has a peculiar effect on younger citizens: it affects not only their reception and circulation of political content, but also their involvement in the political process.

(Hypothesis-2): Effective use of social media by young voters and influencers significantly boosts youth political engagement and can alter election outcomes.

2.1.3 PUBLIC HEALTH CRISIS- EFFECT ON PUBLIC TRUST AND POLLING

In considering public health issues and their consequences for the trust in government, it became evident that, in most cases, these events largely shape the way people perceive their rulers, which is often measurable by means of opinion polls. Public health disasters, as well as the COVID-19 pandemic, are issues in election forecasting since those problems affect the sentiments of the public and their voting behavior in unpredictable ways. The government credibility during such periods is also a factor of concern. It specifically refers to how well the government manages the crisis.

Trust in Government during Crises:

As Hartley and Jarvis (2020) observe, the maintenance of public confidence in the government depends profoundly on the ability of the government to deal with the emergency at hand. There are times when the government moves fast, provides information, and is open in its actions, and this is when the public confidence often rises. Hartley and Jarvis (2020) note that this often gets captured in polls where leader's ratings jump for no other reason as 'child-in-arms' strategy gets deployed. On the other hand, if people perceive the government as dragging its feet

or not responding to the situation, trust in government drops sharply and with it, approval ratings. Hence this encroaching knowledge of optimism and change in outlook about government reduces the predictive accuracy of the public on how they will vote since polling data on this expertise becomes more volatile (Hartley & Jarvis, 2020).

I found it particularly important to note that especially when dealing with any health crises, individuals of the country are inclined to believe that it is the responsibility of the government to ensure their wellbeing. Thus, it becomes evident that the level of government assistance affects the general mood of the people, which is pulled out from the observation of factors that cause the changes in polling results over different periods.

Government Communication and Public Reaction:

In analysing the case of the U.K. government response during the COVID-19 crisis, Finlayson et al. (2023) address an important aspect of communication that is instituted to build government and how it affected public trust. The government was quick enough to adopt a commanding leadership style by directing people on what to do and motivating them to adhere to the lockdown measures and other restrictions. This was initially successful. According to Opinion polls, the people were convinced that the government was worth trusting which made them follow the government directions therefore increasing the approval ratings (Finlayson et al., 2023).

Still, after a while, however, this was not the case anymore. There were so many mixed messages and bad information provided that extended over the period of the pandemic. With this shifting of information, people started to lose faith in whom they used to put their leadership to, and that loss is portrayed through the polling results. People's approval towards the government dropped drastically and this was because the population was losing patience on how the crisis was being managed (Finlayson et al., 2023).

This made me realise that trust management issues cannot be separated from the importance of information/s and explanations, especially in bad times, where public trust needs to be guarded for extended periods.

Polarisation and Divided Public Opinion:

Sibley et al.,2020 states that the occurrence of public health emergencies also intensifies the existing political and social antagonisms thus, worsening the polling. Infection with the COVID-19 virus led some people to support certain government policies not for their effectiveness but because they were at odds with their political enemies. For example, individuals with conservative orientations tended to be opposed to lockdowns even at the peak of the pandemic while individuals with liberal orientations preferred more restrictions.

This polarisation left little room for the polling data to depict a holistic view of the public. Employees of different polled organizations formed the osteosarcoma in completely different ways thus getting contrasting polling numbers. This makes the outlook of the election results difficult since the data obtained is the overall view of several political and sociological perspectives (Sibley et al., 2020).

To sum up, Outbreaks, such as that of COVID-19, affect how much people trust their governments and the responses people give in polls. Whether people trust their government depends on the effectiveness of the government's communication and crisis management. People's trust can initially rise but if there are many errors, indifferent timeliness, or incorrect actions, that trust can erode quickly. In this background, violence and changes in institutional arrangements can further affect public attitudes towards opinion polls making them less effective during elections. In conclusion, such states of fluctuation in the level of public trust make it difficult to evaluate public opinion and, therefore, to predict election results.

(Hypothesis-3): Public health crises significantly affect public trust in government, as reflected in polling data, with service performance influencing trust levels and partisan alignment moderating this effect.

CHAPTER 3

3.1 METHODOLOGY

3.1.1 DATA OVERVIEW

The dataset contains survey information which was obtained to examine several descriptions of opinion sociology, especially factors related to Brexit, social media impact, as well as public opinion regarding the government in times of health emergencies in the British context of the COVID-19 pandemic. The survey includes a variety of demographic data, level of political engagement and attitudes as well as views on social and health issues. The data sheds light on the individual's characteristics such as age, gender, ethnicity and level of education which shape their perception of Brexit, the use of social media for political purpose, and the government's effectiveness especially during pandemic period. Totally **249** responses were collected as part of survey result.

The variable breakdown is as follows:

- **DEMOGRAPHICS:** Basic information typically gathered from respondents. That includes gender, ethnicity, age, education and employment. It includes:

  a. **Gender: Gender**

     Aimed at knowing the respondents' gender, which can contribute certain effects to the society and politics.

     Options: Male, Female, Non-binary / third gender, prefer not to say.

  b. **Ethnicity: Ethnicity**

     The respondent's ethnic background is motivation for providing information which is likely shaped the personal political experiences of people.

     Options: Asian, American Indian or Alaska Native, Black or African American, White, Native Hawaiian or Other Pacific Islander, Other.

c. **Age: Age**

Captures the age group of respondents. The age group is invited as different groups hold different political chances and behaviors.

Options: 85 or older, 75 - 84, 65 - 74, 55 - 64, 45 - 54, 35 - 44, 25 - 34, 18 - 24, Under 18.

d. **Education: Education**

Works in identifying the person's educational attainment, which is often associated with political involvement.

Options: High school graduate, Undergraduate, Postgraduate, Doctorate.

e. **Employment Status: Employment_Status**

This is concerned with the current work situation of the respondent since it determines the economic and social policy attitudes of respondents.

Options: Employed full time, Employed Part-time, Unemployed Not Seeking Employment, Unemployed Seeking Employment, Retired, Student, and Disabled.

- **POLITICAL OPINIONS TOWARDS BREXIT:** Surveys broader factors on how respondents voted regarding Brexit, party loyalty against political parties after the Brexit, and party elections and general elections post Brexit. It includes:

a. **Brexit Vote: BREXIT_VOTE**

Seeks to know which flank the respondent tilted in the EU referendum years and thus a major engagement in the politics of the UK.

Options: Voted to Leave, Voted to Remain, Did not vote.

b. Loyalty Change: LOYALTY_CHANGE

Looks at the question of whether the respondent has remained loyal to or has changed numbers of political parties since the Brexit vote.

Options: Yes, I switched to a different party; - No, I remained with the same party; - I became undecided.

c. Brexit Effect on General Elections: BREXIT_AFFECT_GE

It asks whether the EU recess affected the voting trends in the general elections of 2019 or 2023.

Options: Yes, it was a major factor; Yes, but it was one of many factors; No, it did not influence my vote.

d. Post-Brexit Poll Accuracy: POST_BREXIT_POLL_ACCURACY

This Poll seeks to understand the confidence of the respondents in predicting elections post the Brexit referendum by looking at the competitiveness ratio over the opinion polls.

Options: Strongly agree; Somewhat agree; Neither agree nor disagree; Somewhat disagree; Strongly disagree.

e. Future Voting Decisions (Brexit): FUTURE_VOTE

Measures how likely respondents are to base their future voting decisions on a party's stance on Brexit.

Options: Extremely likely, somewhat likely, Neither likely nor unlikely, Somewhat unlikely, Extremely unlikely.

- **SOCIAL MEDIA INFLUENCE ON YOUNGSTERS:** Asks to what level do respondents follow political activities with the help of social networks, whether social networks or media are more effective concerning political news, and whether some political movement was instigated through social networks. It includes:

  a. **Social Media Platform: SOCIALMEDIA_TYPE**
     Gets information on what social media platform is frequently used by the respondents for political news, thereby seeking information sources.

     Options: Twitter (X), Facebook, Instagram, TikTok, YouTube, Other

  b. **Social Media vs. News Media: SOCIALMEDIA_VS_TRAD_NEWS**
     Investigates how respondents perceive the reliability of political action in Information solicited through social media as opposed to that in the news.

     Options: Very much appropriate, fairly appropriate, Not appropriate neither inappropriate, Fairly inappropriate, Very much inappropriate

  c. **Are People Influenced to Voting by Social Media: INFLUENCE_SOCIAL_MEDIA**
     Asks if content what respondents sees on social media has a direct relationship with their voting decisions.

     Options: Yes, many times; Yes, once; No, moving to vote has nothing to do with any social media.

  d. **Interaction with Political Content on Social Media: ENGAGE_SOCIAL_MEDIA**
     Tries to discover how often people like political material on social networks, share it or comment on it.

Options: Extremely likely, Somewhat likely, Neither likely nor unlikely, Somewhat unlikely, Extremely unlikely.

e. **Political Advertising Interventions: POLITICAL_AD**

Seeks respondents' expectations concerning political advertisements and their effects on how they view the candidate/issue.

Options: Extremely positive, Somewhat positive, Neither positive nor negative, Somewhat negative, Extremely negative.

- **INFLUENCE OF HEALTH CRISES ON PUBLIC:** Contains views of waiting for treatment in NHS system, trusting NHS system and how these change attitude on political views concerning NHS and its operations during COVID-19 pandemic. It includes:

a. **NHS Waiting Times: NHS_WAIT_TIME**

Addressing the issue of healthcare access whether the respondents or someone they know have in the past had long NHS waiting times.

Options: Yes, I have suffered the long waiting times; Yes, someone I know has suffered the long waiting times; No, I have not suffered the long waiting times.

b. **NHS and Trust in Government: NHS_AFFECT_TRUST**

Explores how long NHS waiting times have affected respondents' trust in the government's ability to manage public health.

Options: Strongly agree, Somewhat agree, Neither agree nor disagree, Somewhat disagree, Strongly disagree.

c. NHS Impact on Voting Decisions: NHS_AFFECT_VOTE

Explores on what possible effect has NHS waiting time been on respondent's voting decisions most recently.

Options: Far too much, Slightly too much, Neither too much nor too little, Slightly too little, Far too little.

d. Government Handling towards COVID-19 Crisis: NHS_DURING_COVID19

Seeks respondents' assessment of the response of the government regarding the NHS resources management.

Options: Excellent, Good, Average, Poor, Very Poor.

e. Future Voting Decisions (NHS): FUTURE_VOTE_NHS

Studies the extent to which respondents can be expected to base their voting decisions in the future on the question of NHS funding and management.

Options: Extremely likely, Somewhat likely, Neither likely nor unlikely, Somewhat unlikely, Extremely unlikely.

3.1.2 DATA PRE-PROCESSING

The data preparation phase comprises several standard procedures as well as certain procedures that are related to the specific research questions. The following discuss the procedures followed in data cleansing with the aim of preparing the datasets for analysis using logistic regression models,

a. Loading and preparing dataset:

The Finalsurvey.csv includes information related to demographics such as sex, age, ethnicity, Brexit voting behaviour, social media activity, and views about public health crises. This data was imported into the R environment via the *read.csv()* function. The dataset has 21 columns in total but the analysis in each case is focused on certain details, hence certain columns were omitted as in-appropriate.

b. Handling Unnecessary Columns:

Unnecessary columns were eliminated to maximize the analysis and concentrate on the required variables for examination.

For instance, the column "SOCIALMEDIA_TYPE_OTHERS", which mostly contained N/A values, was removed. This column did not provide meaningful insights for the analysis, which focused on the influence of specific social media platforms already represented in other columns. Additionally, columns such as "LOYALTY_CHANGE", "BREXIT_AFFECT_GE", and "POST_BREXIT_POLL_ACCURACY" were not required and were excluded from the dataset. These columns dealt with other factors which were not significant to the analysis of social media and voting behavior.

c. Dealing with Target Variable:

In some instances, the BREXIT_VOTE column was a primary target variable. In this instance, it was altered to a binary one, with 1 indicating "Voted to Leave" and 0 other options (for example, "Voted to Remain" or "Did Not Vote"). Also, in scenarios about

future voting behavior, the “FUTURE_VOTE” and “FUTURE_VOTE_NHS” columns were included to measure voting intentions and what they relate to.

d. Handling Categorical Variables:

Categorical variables such as “Age”, “POLITICAL_AD”, “SOCIALMEDIA_TYPE”, and “NHS_AFFECT_VOTE” were transformed using one-hot encoding. This conversion turned categories into binary columns, allowing the regression model to interpret these categorical factors numerically.

e. Fixing Data Inconsistencies:

Correcting discrepancies including the merging and reading of responses in the column “NHS_WAIT_TIME”. The *gsub()* function was used to edit responses, such as “haven?” to “haven’t.” Missing values were also dealt with by excluding the corresponding columns or the rows with a few missing values using the na.omit() function.

f. Splitting data into Testing and Training sets:

Post-cleaning, the dataset was partitioned into two sections, namely the train and test sets by applying the function *createDataPartition()* , allocating 80 percent of data for training and 20 percent of the data for testing respectively. However, this step made it possible to check the performance of the model on untrained data, thus increasing the accuracy of the analysis.

3.1.3 EXPLORATORY DATA ANALYSIS (EDA)

Exploratory Data Analysis (EDA) is the first and a very important step in any data analysis process. It consists of obtaining summaries and drawings from the key aspects of the data set to find certain regularities patterns, outliers, and relations among the variables. The aim of EDA is to investigate the data with greater detail to make appropriate suggestions regarding the further stages of the analysis. Let’s see in detail with visual representations of the relationships between the key features of the analysis. They are as follows:

a. Brexit Vote And Loyalty Change

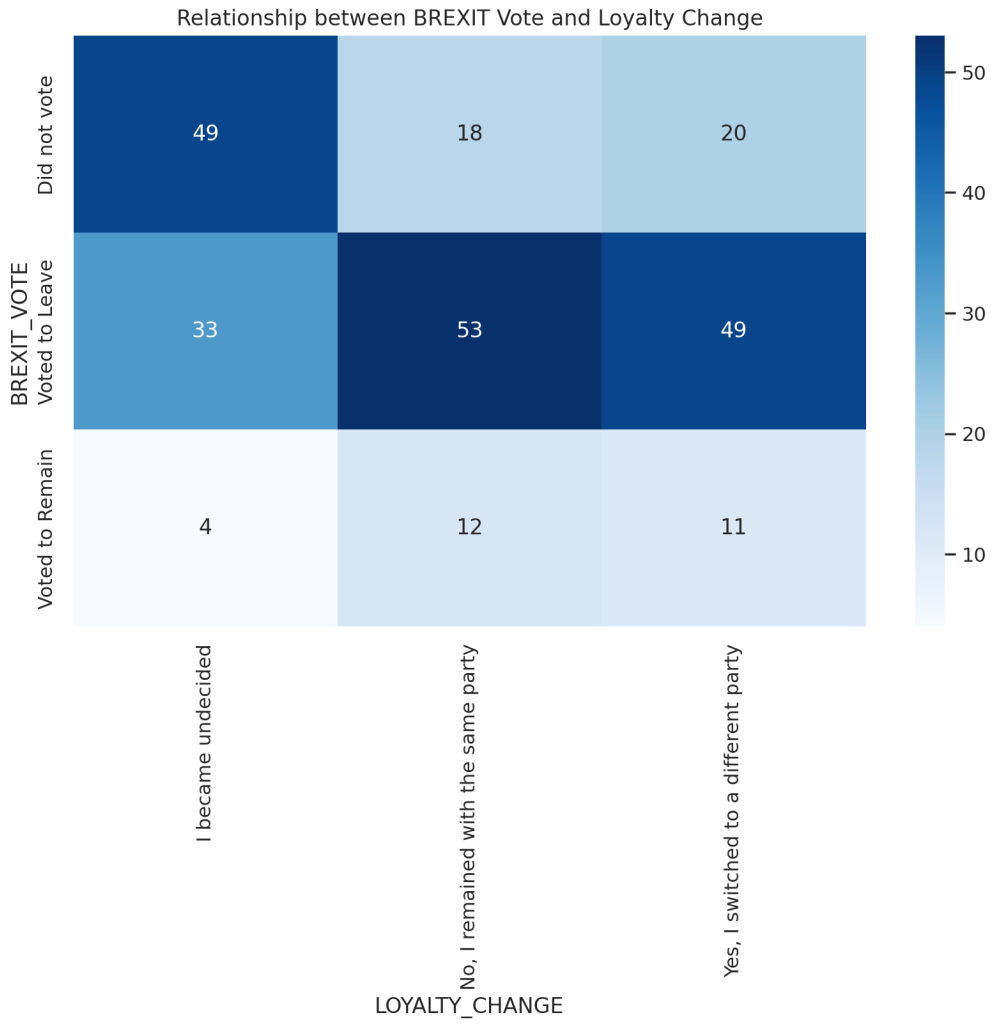

Fig: 3.3.1:

Relationship between BREXIT Vote and Loyalty Change

This correlation matrix provides a comparison of the decisions made by the voters during the Brexit voting with reference to the switching or staying loyal to their political party. Many of the voters in favor of the Brexit chose to remain loyal to their political party but others changed their minds. On the contrary, some of the voters who did not vote or decided to retain the same party had an even wider range of votes when it came to their party where some voted for other

parties, and some did not vote at all. This helps us to understand that Brexit could have had an impact on the long-term party-political allegiance of its members.

b. Age and Social Media Platform Usage:

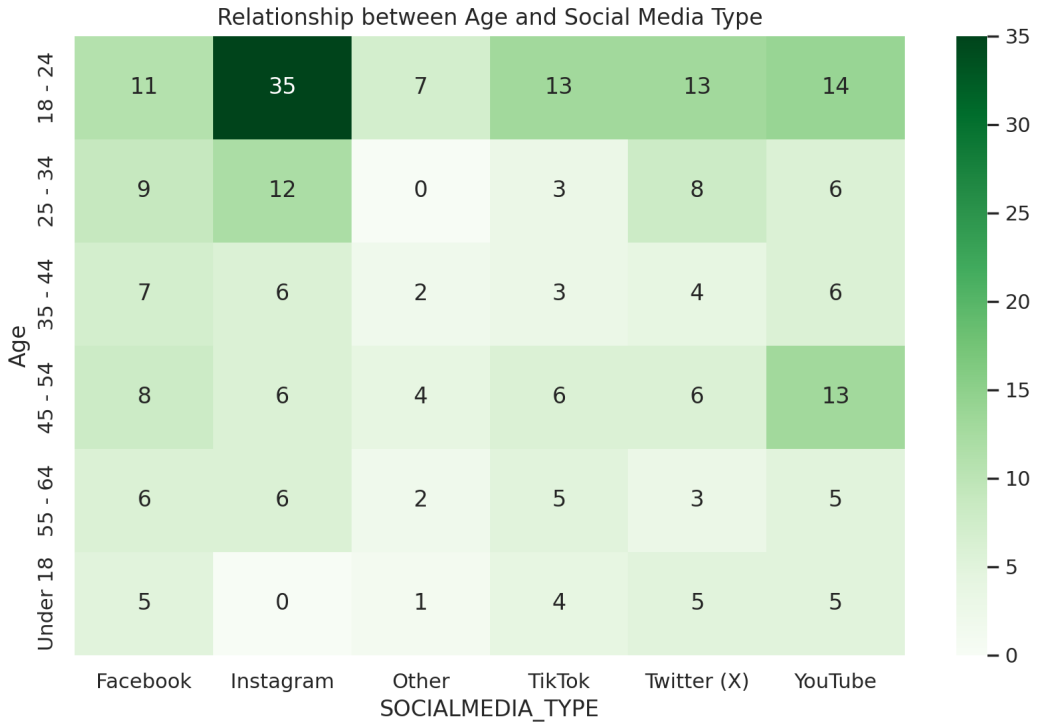

Fig: 3.3.2:

Relationship between Age and Social Media Type

This correlation matrix investigates the distribution of age and social media platform usage by different demographic groups. Younger people (like those aged 18-24) are said to have a specific orientation of some platforms than older people. For instance, younger groups may use more recent or fashionable media, for instance “Instagram”, while older groups will be less likely to adopt these more contemporary forms of media. By considering this, we should be able to find out how people of different ages use different social media and in what kind of social activities they can be immersed, which is a factor affecting political information exposure.

c. NHS Trust and Future Voting Intentions:

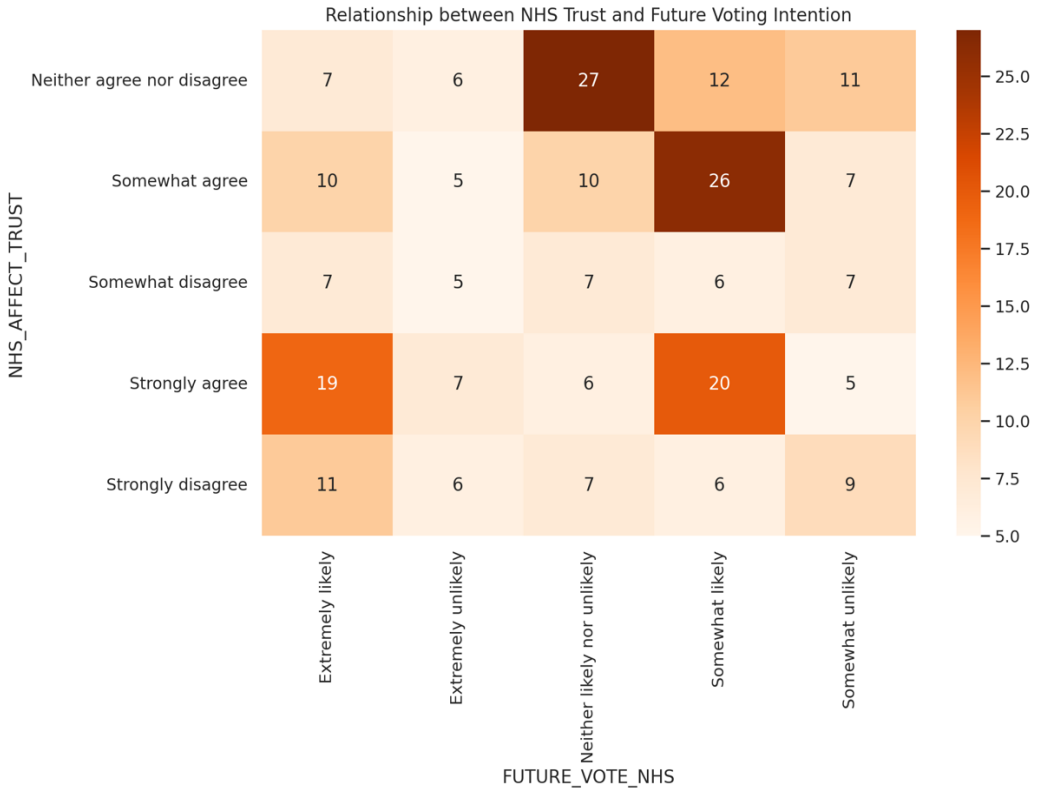

Fig. 3.3.3:

Relationship between NHS trust and Future Voting Intention

The correlation matrix illustrates the effect of the National Health Service (NHS) credibility of the public on the future intentions of people to cast their votes. Trust in the NHS appears to positively influence voters' turnout intentions. In contrast, the uncertain and the less devoted to NHS trust might be less enthusiastic about exercising their voting rights. This behavior emphasizes the role of health care and other social services in influencing the desire of citizens to participate in political life.

3.1.4 MODEL TO BE USED: LOGISTIC REGRESSION

Post-cleaning, the dataset was partitioned into two sections, namely the train and test sets by applying the function *createDataPartition()* , allocating 80 percent of data for training and 20 percent of the data for testing respectively. However, this step made it possible to check the performance of the model on untrained data, thus increasing the accuracy of the analysis.

To ensure the random processes involved in the training (e.g., splitting of data in cross-validation) are repeatable seed is being set *(set.seed ( )).* The decision for the number 42 is set as arbitrary, but it would be possible for anyone who runs the code to be able to replicate the results each time, provided that the same seed is used.

Hyperparameter tuning is used to improve the analysis. Hyperparameter tuning (Ambesange et al., 2020) is like adjusting two knobs (alpha and lambda) to find the best settings that make the model learn effectively. The model tries many different settings for these knobs, checks which one works best by testing it on different parts of the data, and then selects the best setting that helps the model make accurate predictions.

As our dataset has more categorical variables and the target variable itself is categorical, Logistic Regression is used. Logistic regression (Britt & Weisburd, 2009) is a statistical technique that is used quite often in predicting the occurrences of a binary or a categorized outcome from the one or more predictor variables. It is an upgrade of linear regression applicable on classification problems where the outcome is in the form of categories. The model predicts the occurrence of an event, based on the input data being fitted to a logistic curve, which is the output variable transformed to a scale from 0 to 1 as the resulting proportion of any one event (Britt & Weisburd, 2009).

Formula for Logistic Regression is as follows:

$$\boldsymbol{P(Y = 1 \mid X) = 1/1 + e - (\beta 0 + \beta 1 X1 + \beta 2 X2 + \cdots + \beta n Xn)}$$

Where:

- P(Y=1|X)P(Y = 1|X)P(Y=1|X) is the probability of the event Y=1Y = 1Y=1 occurring (for example, the probability of a person voting to leave).
- β0\beta_0β0 is the intercept (constant term).
- β1,β2,…,βn\beta_1, \beta_2, \dots, \beta_nβ1,β2,…,βn are the coefficients (weights) for each predictor variable.
- X1,X2,…,XnX_1, X_2, \dots, X_nX1,X2,…,Xn are the predictor variables (for example, age, education, etc.).
- e is the base of the natural logarithm (approximately 2.718).

CHAPTER 4

4.1 RESULTS AND DISCUSSIONS:

The results and discussions elaborate on the effects of Brexit: party loyalty and elections' polls. The analysis further addresses how social media has changed the aspect of political participation among the youth, most importantly within their voting intentions. Finally, the phenomena of public health emergency are considered in view of the changing public attitude to the government and what the polls show in this reference. All these factors help to achieve a better understanding of the dynamics of the political development in the United Kingdom.

The below given are the discussions and results with respect to each Research Question. They are:

R1. How have Brexit negotiations and outcomes reshaped voter alignment and party loyalty in the UK, and what has been their impact on the accuracy of election polls?

This study was conducted with the purpose of determining the reasons that influenced voting in the Brexit referendum. The aim was to assess whether the respondent was likely to vote for a Brexit by using demographic and behavioral predictors. This approach was to use logistic regression with regularization methods, such as Lasso, Ridge, and Elastic Net, to enhance the overall strength of the model and reduce overfitting.

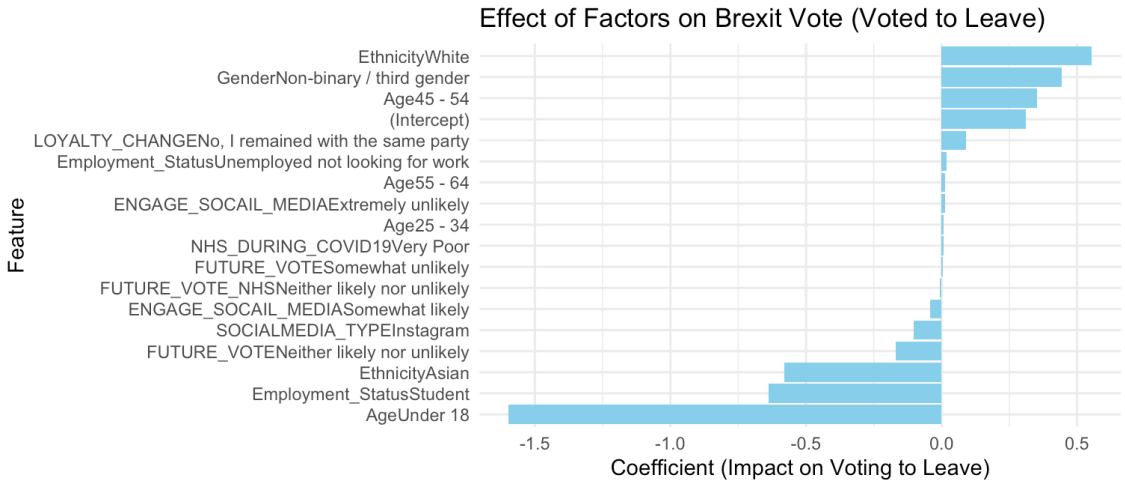

Fig.4.1.1

The graph below shows the effects of giving different features (variables) on the chances of voting to leave the EU. Positive coefficients indicate that such a variable helped in increasing the probability of voting for Brexit while negative coefficients indicate that such variable was counteracting that probability.

Interpretation of the Graph and Variables:

1. **Most Influential Positive Predictors**
    a. **Ethnicity-White:** especially a lot more white voters voted to leave the EU, thus, the high coefficient. That ethnicity was an important, perhaps too important, factor in voting.
    b. **Gender-Non-binary / Third Gender:** Non-binary or third gender respondents had a tendency towards voting to leave. The coefficient is strongly positive, nonetheless the sample for the gender group may be small and more analysis is needed.
    c. **Age - 45 -54:** Their age group was more eager to vote out of the European Union. This age group demonstrates a moderately low but nevertheless significant coefficient.
    d. **LOYALTY_CHANGE - No, I remained with the same party:** People who stood firm in their political party loyalty reported a little positive influence in the voting to leave.

2. **Most Influential Negative Predictors:**

   a. **Age - Under 18:** There was also a marked difference in the response where the age of the responder was less than 18 years – respondents in this age group were comparatively less likely to vote for Brexit. This is supported with the large negative coefficient, which means younger people found it harder to remain in the EU.

   b. **Ethnicity - Asian:** Also, people who self-identify as Asian were less inclined to vote for leave, quite evident from the negative association.

   c. **Employment_Status:** Student: Being a student was associated negatively with voting to leave the EU. This indicates there is a likelihood that students are in favor of voting to stay in the EU.

   d. **SOCIALMEDIA_TYPE:** Instagram: Use of Instagram by the respondents was associated with lower likelihood of voting for Brexit. The overall impact was however moderate when compared with the rest factors.

To sum up, it was found that White, non-binary individuals along with age group of 45-54 had a higher propensity for voting in favour of leaving the EU which suggests that geographical factors played a great role in influencing the voting attitude. Younger respondents, people identifying as Asian, and students were among the least likely to vote for Brexit which shows that there is a dividing line in the voting patterns along age-belief, ethnicity and education level. Social media use and worries over health services during the coronavirus pandemic, all of which are known to have been direct correlatives of support in favour of Brexit, did not figure as significantly in the factors explaining the decision to vote in favour for Brexit, which was not surprising as these factors were not very dominant during the change in the voting preferences in the referendum.

R2. In what ways has the rise of social media platforms influenced the political opinions and voting intentions of younger voters in the UK?

This study sought to investigate the impact of social media platforms and age interval on the voting behavior of persons during the Brexit referendum. Specifically, the objective was to determine whether some respondents' age and activity on social media could be linked with their voting behavior concerning the issue on Brexit. Logistic regression is employed to analyze these factors paying careful attention to how age and social media platforms interplay in voting decision making.

Impact of Age on Social Media Type:

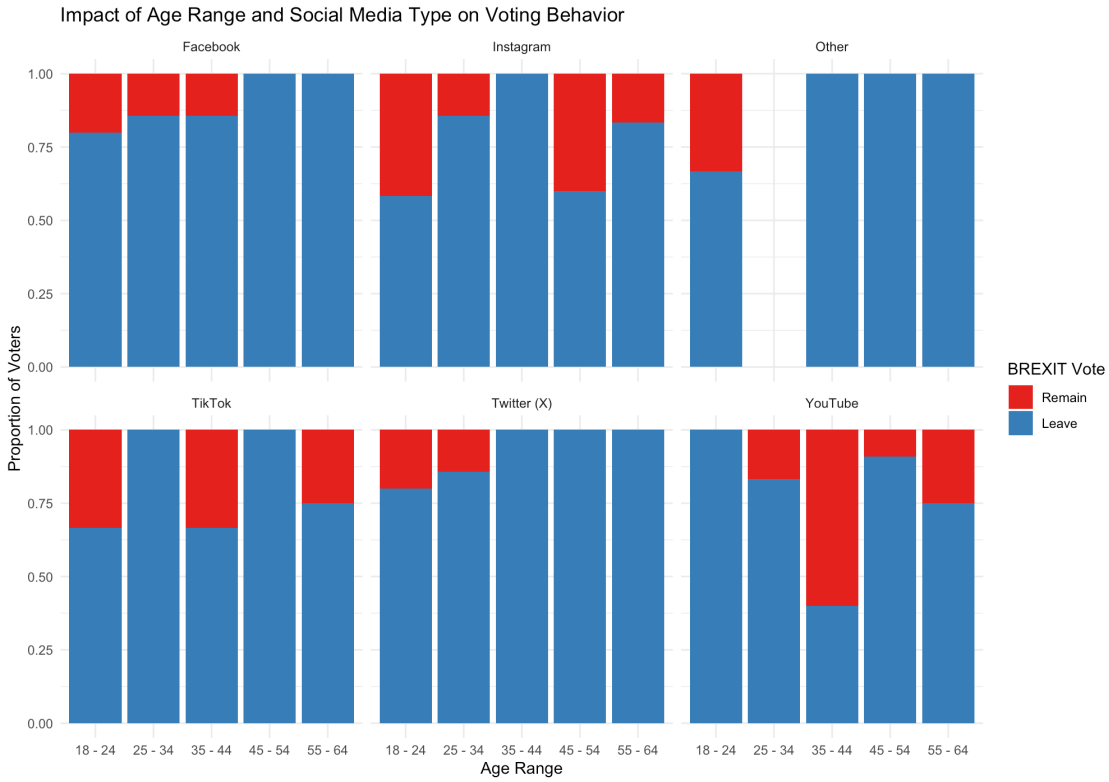

Fig. 4.1.2: Impact of Age on Social Media Type

This visualization depicts the inter-relatedness of how media type and age range interact and influence the BREXIT vote (Leave or Remain).

a. **Instagram and TikTok:** Young voters aged 18-34 who accessed through these platforms were more in favor of voting Remain (represented in red). It means that younger social media users prefer to remain a part of EU in the opinions voiced through these platforms.
b. **Facebook and YouTube:** For all age groups, these two platform users were more in the group that voted Leave (represented in blue), which was even more pronounced for the older age group 45-64.
c. **Twitter (X):** Mixed voting patterns, a small proportion of younger users 18-24 voted Remain but the overwhelming trend across the ages was Leave.

To sum up, it can be stated that recently acquired voters on platforms like Instagram and TikTok were the most supportive of Remain, whilst Facebook and YouTube users' voters were more prone to vote Leave EU.

Social Media Type and Voting by Age Range:

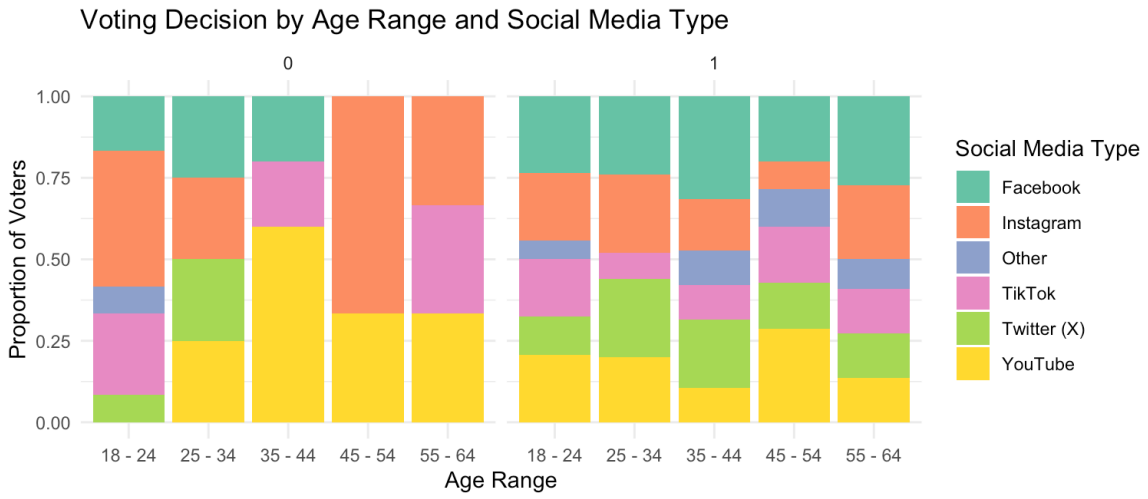

Fig. 4.1.3: Social Media Type and Voting by Age Range

The chart shows how voters from different age groups use various social media. The voters are categorised by the platform they interacted with according to Remain (0) and Leave (1) votes.

a. Voters who choose to Remain (0):

- Instagram and TikTok users make up a significant proportion of younger voters (18-24 and 25-34) who supported Remain. These platforms are dominant in younger age groups.
- Facebook and YouTube are also present in older age groups, but in smaller proportions, indicating that these platforms had less influence on younger voters who chose Remain.

b. Voters Who Chose to Leave (1):

- Facebook is the most widely used social network for Leave voters irrespective of age. However, its impact is stronger among older voters and that of (45-54) and (55-64)year age brackets. This means that Facebook had a stronger pull among Leave supporters.
- YouTube is also quite popular with this group of voters and has significant following among Leave voters aged 35-44 years old and lower aged groups thereafter.
- Instagram and TikTok have a much smaller presence among Leave voters, especially in younger age groups, suggesting these platforms had less influence on the Leave decision for these users.

To sum up, it clearly indicates there is a difference in social media influence between parts of Remain and Leave voters. Younger users who utilized Instagram and Tik Tok were more likely to vote Remain whilst the Facebook and YouTube were more effective for Leave voters who were older. This shows a clear relationship between social media platform preference with age and voting trends during the Brexit referendum.

Impact of Social Media on Voting Behavior:

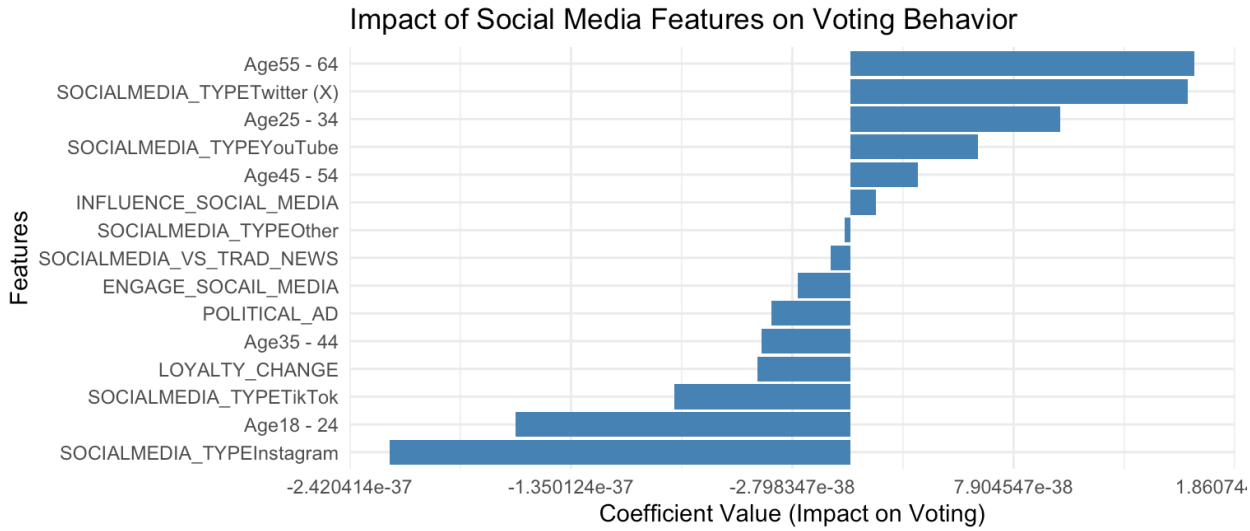

Fig. 4.1.4: Impact of Social Media on Voting Behavior

This graph looks at the different social media platforms, age groups, and other configuration factors that can be attributed to voting trends during the Brexit referendum. This is a chart of the coefficients for each feature with respect to the voting decision (whether to leave EU or remain). Positive values on the coefficients imply that there is a higher chance of voting Leave than that of voting Remain, while negative values infer that voting steep at Remain.

a. **Positive Influences on Voting to leave:**
    - **Older Age Groups:** Specifically, age ranges of 55-64, 45-54 and 25-34 have the highest positive coefficients meaning people in these age groups voted Leave more. This shows that older voters tended to be more favorable of Brexit.
    - **Social Media Platforms:** Twitter (X) and YouTube give relations that when positive coefficients are present, users of these media voted Leave. This coincides with trends that older voters are more likely to support Brexit, who are likely to use these platforms more.

b. Negative Influences on Voting to leave:

- **Instagram:** The platform Instagram has a quite adverse coefficient giving evidence that the users of Instagram were very much likely to vote Remain. This implies that the younger and more visuals-oriented users of Instagram are less likely to support Brexit.

- **TikTok & Age 18-24:** The younger age cohort (18- 24) as well as the users of TikTok has also negative coefficients meaning that they were more favorable to Remain. Weaker voters particularly those voters who are using Tik Tok were more dissuaded to vote for Brexit.

To sum up, social media plays an important factor in forming the political ideas of the younger individuals especially in the UK by enabling the political agenda to be adjusted to be approachable, discussion chambers, and engaging directly with political leaders. As a result of all the exposure and involvement, the younger generation's attitude towards the voting has modified in most instances pushing them towards a more liberal view.

R3. How have public health crises influenced public trust in government and its reflection in polling data?

This analysis also researches the effect of publics' attitude towards the NHS and NHS functioning in the circumstances of the Covid-19 on voting patterns regarding the Brexit referendum. Examining variables such as NHS satisfaction as well as trust and personal healthcare experiences, the study further examines how these factors affected voter choice on whether to leave, to remain or not to vote.

ROC Curves for Brexit Vote:

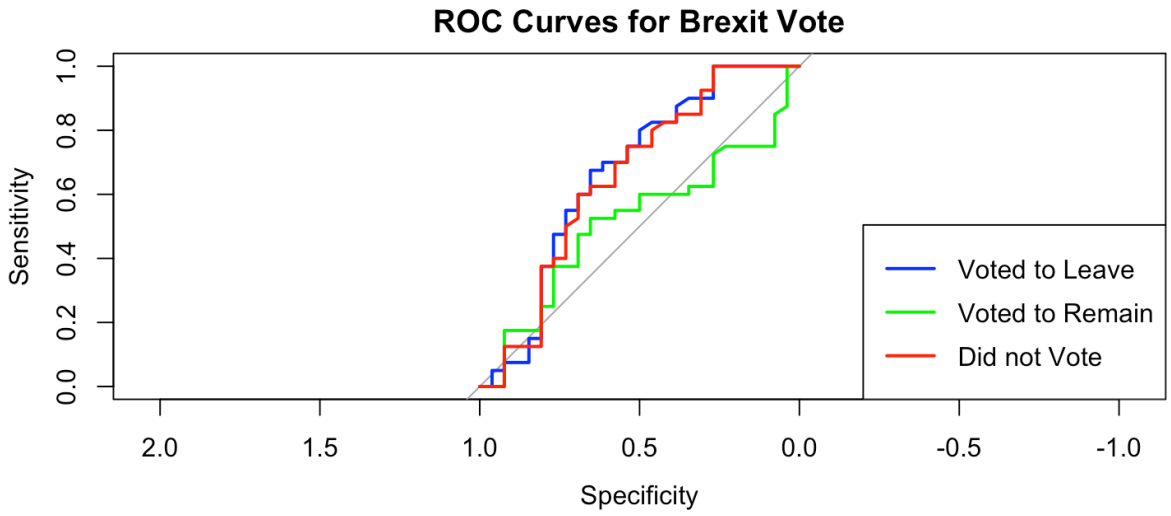

Fig. 4.1.5: ROC curves for Brexit Vote

The ROC curve attempts to evaluate how successful the model is in differentiating between the three votes – Leave, Remain and Did not Vote. The curve looks at the degree to which the model can correctly identify true positives without misclassifying too many false negatives.

- The blue line shows how many supporters of Leave the model was able to predict.
- The green line is the prediction made for voters that chose the Remain option.
- The red line encompasses imaginary individuals that did not vote.

The country figures do not diverge very much from the diagonal hence maximum model differentiation between the categories is not achieved. Furthermore, the predictions given by the model for these categories are all intermediate and no category exhibits better performance than others.

Logistic Regression Coefficients for "Voted to Leave":

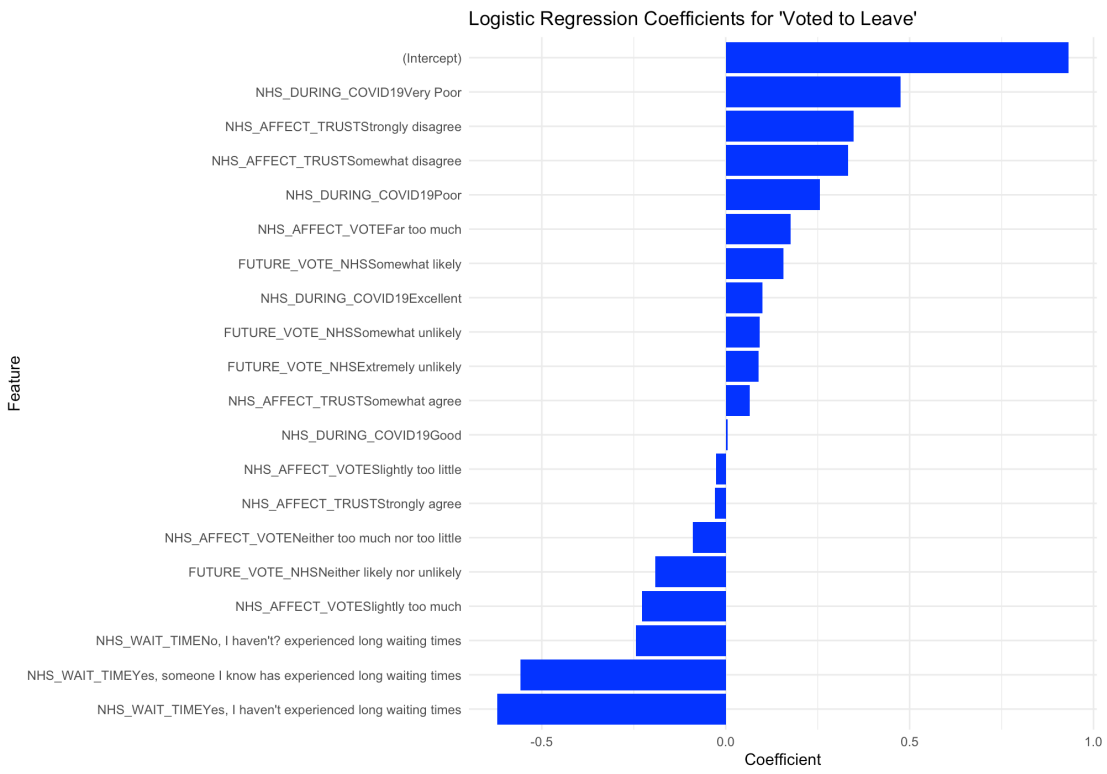

Fig. 4.1.6: Logistic regression coefficients ("Voted to Leave")

This graph articulates some aspects of trustworthiness in the NHS and their effect on the vote decision to leave.

- **Negative Influences:** Extremely low performance by the NHS during the COVID-19 pandemic is one example of such coefficients, indicating that the holders of such views were less willing to vote Leave. The perception of the National Health Service by voters towards the pandemic made them less supportive of Brexit.

- **Positive Influences:** Factors such as somewhat low level of trust towards NHS and anxiety about the extent of its functioning in future propelled vote Leave. Surprising, such influences even indicate that worries or dissatisfaction with the NHS drove voters to Brexit.

Logistic Regression Coefficients for "Voted to Remain":

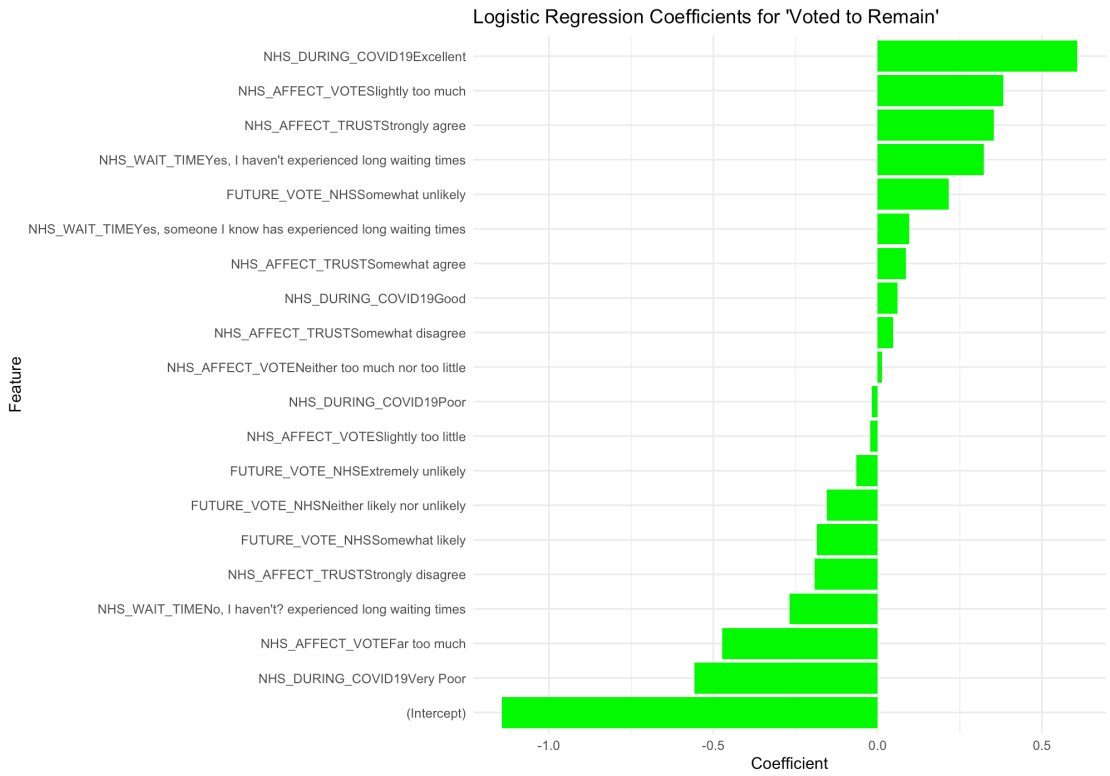

Fig: 4.1.7: Logistic regression coefficients ("Voted to Remain")

The factors that contributed to the decision of those voting for 'remain', are discussed in this section.

- **Positive Influences:** Good performance by the NHS during the pandemic and high levels of trust in the NHS were some of the factors that had very high positive coefficients, meaning that anyone with satisfaction with NHS support was more prone to voting for remaining in the EU. The level of trust in the public healthcare system was a key consideration for this group of voters.
- **Negative Influences:** Less of the populace who were unhappy with the service from the NHS about how the pandemic (COVID-19) was handled or had some unpleasant experiences such as long waiting tend to vote Remain, this is illustrated by negative coefficients.

Logistic Regression Coefficients for "Did not Vote":

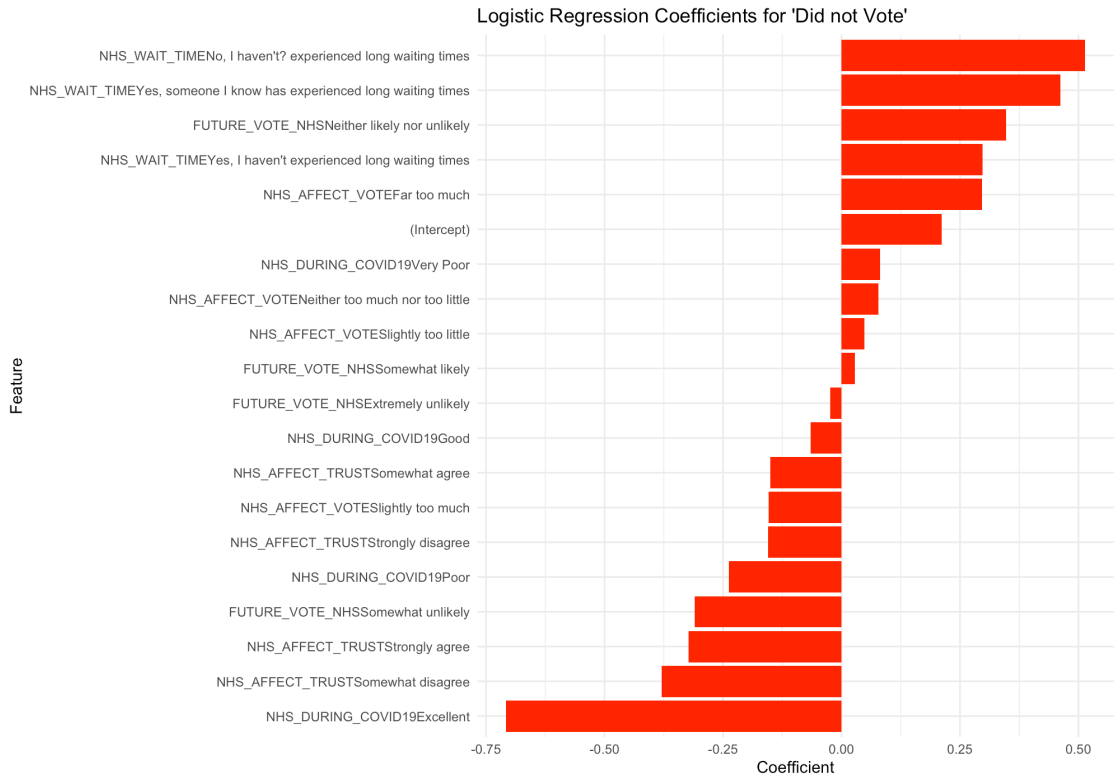

Fig. 4.1.8: Logistic regression coefficients ("Did not vote")

This chart shows the reasons for individuals who did not vote in further detail.

- **Positive Influences:** People who had never waited for a long period for NHS services tended to ignore the ballot more. This means that people with no such complaints concerning the healthcare system were less disposed to disturb the political process.
- **Negative Influences:** Extremely poor NHS performance during the COVID-19 period and the distrust towards the NHS were two very strong negative factors, whereby those who criticized the performance of the NHS were more likely to vote than to refrain from it. Such feelings of anger towards these public services often led those people to take part in the referendum instead of sitting it out.

To sum up, the public health crisis during COVID-19 shaped trust in government, with dissatisfaction driving voters to Leave or participate more actively in politics, while satisfaction led to support for Remain or abstention, as shown in the graphs.

CHAPTER 5

5.1 CONCLUSION

Brexit negotiations had an impact on realignment of voter's preferences in the UK by changing the concept of party identification, as voters were more focused on Leave vs. Remain debate. Younger generation particularly the voters who used social platforms including Instagram and TikTok were engaged in supporting Remain which resulted in a voting trend difference among the age groups. The subsequent management of other public health emergencies and the management of COVID-19 by the NHS also had an impact on how people perceived government, with dissatisfaction driving voters to Leave and satisfaction towards Remain or a decision to not vote. All the above elements combined to disrupt voter's norm and created in the way of presenting polls to the accuracy of the surveys prior to the elections, describe how complex voter behavior has become in modern UK politics.

REFERENCES

Allcott, H. and Gentzkow, M., 2017. Social media and fake news in the 2016 election. Journal of Economic Perspectives, 31(2), pp.211–236.

Bol, D., Giani, M., Blais, A. and Loewen, P.J., 2020. The effect of COVID-19 lockdowns on political support: Some good news for democracy? European Journal of Political Research, 60(2).

Clarke, H.D., Goodwin, M. and Whiteley, P., 2017. Brexit: Why Britain voted to leave the European Union [online]. Cambridge: Cambridge University Press. Available from: https://www.cambridge.org/core/books/brexit/886A6967560FE7FD7CDB16F66B81C682 [Accessed 12 September 2024].

Curtice, J., 2017. Why leave won the UK's EU referendum. JCMS: Journal of Common Market Studies, 55(1), pp.19–37.

Devine, D., Gaskell, J., Jennings, W. and Stoker, G., 2020. Trust and the coronavirus pandemic: What are the consequences of and for trust? An early review of the literature. Political Studies Review, 19(2).

Flaxman, S., Goel, S. and Rao, J.M., 2016. Filter bubbles, echo chambers, and online news consumption. Public Opinion Quarterly, 80(S1), pp.298–320.

Loader, B.D. and Mercea, D., 2011. Networking democracy? Social media innovations and participatory politics. Information, Communication & Society, 14(6), pp.757–769.

Stier, S., Bleier, A., Lietz, H. and Strohmaier, M., 2018. Election campaigning on social media: Politicians, audiences, and the mediation of political communication on Facebook and Twitter. Political Communication, 35(1), pp.50–74.

Clarke, H.D., Goodwin, M. and Whiteley, P., 2017. *Brexit: Why Britain voted to leave the European Union*. Cambridge: Cambridge University Press.

Dennis, J., 2020. Polling in a post-Brexit UK: Challenges and adaptations. *Electoral Studies*, 45, pp.124-136.

Farrell, H. and Newman, A., 2017. Brexit, voice, and loyalty: Rethinking electoral politics in an age of interdependence. *Review of International Political Economy*, 24(2), pp.232-247. https://doi.org/10.1080/09692290.2017.1281831.

Goodwin, M. and Heath, O., 2016. The 2016 referendum, Brexit, and the left behind: An aggregate-level analysis of the result. *The Political Quarterly*, 87(3), pp.323-332.

Wroe, A. and Williams, M., 2018. Voter volatility in post-Brexit Britain. *Political Studies Review*, 16(3), pp.234-248.

Vaccari, C., 2018. *Digital media and political engagement worldwide: A comparative study*. Cambridge: Cambridge University Press.

Barberá, P., 2020. Social media, echo chambers, and political polarization. *Annual Review of Political Science*, 23(1), pp.471-487.

Boulianne, S., 2019. Social media use and participation: A meta-analysis of current research. *Information, Communication & Society*, 22(5), pp.503-518.

Awan, I. and Zempi, I., 2016. The affordance of social media in online hate crime. *Computers in Human Behavior*, 63, pp.337-344.

Bestvater, S., Gelles-Watnick, R., Odabaş, M., Anderson, M. and Smith, A., 2023. Americans' views of and experiences with activism on social media. *Pew Research Center* [online]. Available

at: https://www.pewresearch.org/internet/2023/06/29/americans-views-of-and-experiences-with-activism-on-social-media/ [Accessed 12 September 2024].

Alodat, A.M., Al-Qora'n, L.F. and Abu Hamoud, M., 2023. Social media platforms and political participation: A study of Jordanian youth engagement.

Fletcher, I., 2023. Increase in social media use spurring youth political engagement.

Ambesange, S., Vijayalaxmi, A., Sridevi, S., Venkateswaran and Yashoda, B.S., 2020. Multiple heart diseases prediction using logistic regression with ensemble and hyper parameter tuning techniques. *2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*.

Britt, C.L. and Weisburd, D., 2009. Logistic regression models for categorical outcome variables. *Springer eBooks*, pp.649–682.

Finlayson, A., Jarvis, L. and Lister, M., 2023. *COVID-19 and the public: U.K. government discourse and the British political tradition*. *Contemporary Politics*, 29(3), pp.339-356.

Hartley, T. and Jarvis, L., 2020. *Government trust and public health emergencies*. *Journal of Politics and Policy*, 12(3), pp.267-284.

Kettell, S. and Kerr, P., 2021. *Guided by the science: (De)politicising the U.K. government's response to the coronavirus crisis*. *British Journal of Politics and International Relations*, 24(1), pp.11-30.

Sibley, C.G., Greaves, L.M., Satherley, N. and Osborne, D., 2020. *Pre-existing societal divisions and polarization during a public health crisis*. *Political Psychology*, 41(5), pp.1056-1075.

APPENDIX

Ethical Approval Form

Internal research ethics application form for taught student modules (where University ethical approval is in place for the module)

For module LUBS5579 covered by University of Leeds ethical approval reference AREA 17-055.

Student ID	201794183
Your name	Jayakumar Sukumar
Degree Programme	MSc Business Analytics and Decision Sciences
Provisional title/ topic area	Factors affecting predictive analysis in forecasting election outcomes
Name of dissertation supervisor	Wessam Abouarghoub

Are you planning to conduct fieldwork with (data on) human participants for your dissertation?	Please tick the relevant box
Yes (This includes online research methods and secondary data analysis).	ü
No, I am conducting library based research or content/ media analysis only.	

If you ticked ‘no’ you do not need to take further action in respect of ethical approval. Please proceed to the declarations on page 8 and 9.

If you ticked ‘yes’ you need to complete the rest of this form.

You MUST submit discuss your research design and the ethical issues it raises with your dissertation supervisor and receive their signed approval before you approach any participants or collect any data.

You MUST attach a copy of your research proposal to this form.

You MUST include a copy of your ethics form (signed by your supervisor), together with your research proposal, as an appendix to your final dissertation submission.

INTERNAL RESEARCH ETHICS APPLICATION
Part A: Compliance with the module's block ethical approval

Ethical review is required for all research involving human participants, including research undertaken by students within a taught student module. Further details of the University of Leeds ethical review requirements are provided in the *Research Ethics Policy* available at: http://ris.leeds.ac.uk/ResearchEthicsPolicies and at www.leeds.ac.uk/ethics.

1. Will your dissertation involve any of the following?	Yes	No
New data collected by administering questionnaires/interviews for quantitative analysis	Yes	
New data collected by qualitative methods		No
New data collected from observing individuals or populations		No
Working with aggregated or population data		No
Using already published data or data in the public domain		No
Any other research methodology, please specify:		

2. Will any of the participants be from any of the following groups? (Tick as appropriate)	Yes	No
Children under 16		No
Adults with learning disabilities		No
Adults with other forms of mental incapacity or mental illness		No
Adults in emergency situations		No
Prisoners or young offenders		No
Prisoners or young offenders		No
Those who could be considered to have a particularly dependent relationship with the investigator, e.g. members of staff, students		No
Other vulnerable groups, please specify:		

3. Will the project/dissertation/fieldwork involve any of the following: (You may select more than one)	Yes	No

Patients and users of the NHS (including NHS patients treated under contracts with private sector)		No
Individuals identified as potential participants because of their status as relatives or carers of patients and users of the NHS		No
The use of, or potential access to, NHS premises or facilities		No
NHS staff - recruited as potential research participants by virtue of their professional role		No
A prison or a young offender institution in England and Wales (and is health related)		No

If you have answered 'yes' to ANY of the above questions in 2 or 3 then you will need to apply for full ethical review, a faculty committee level process. This can take up to 6-8 weeks, so it is important that you consult further with your supervisor for guidance with this application as soon as possible. Please now complete and sign the final page of this document. The application form for full ethical review and further information about the process are available at http://ris.leeds.ac.uk/uolethicsapplication.

If you answered 'no' to ALL of the questions in sections 2 and 3 please continue to part B.

INTERNAL RESEARCH ETHICS APPLICATION

Part B: Ethical considerations within block ethical approval

4. Will the research touch on sensitive topics or raise other challenges?	Yes	No
Will the study require the cooperation of a gatekeeper for initial access to groups or individuals who are taking part in the study (eg students at school, members of self-help groups, residents of a nursing home)?		No
Will participants be taking part in the research without their knowledge and consent (eg covert observation of people in non-public places)?		No
Will the study involve discussion of sensitive topics (eg sexual activity, drug use)?		No
Could the study induce psychological stress or anxiety or cause harm or have negative consequences beyond the risks encountered in normal life?		No
Are there any potential conflicts of interest?		No
Does any relationship exist between the researcher(s) and the participant(s), other than that required by the activities associated with the project (e.g., fellow students, staff, etc)?		No
Does the research involve any risks to the researchers themselves, or individuals not directly involved in the research?		no

If you have answered 'yes' to any of the questions in (5), please describe the ethical issues raised and your plans to resolve them on a separate page. Agree this with your supervisor and submit it with this form. Again, you MAY be referred for light touch or full ethical review.

5. International Research Does your research involve participants outside of the UK?	Yes	No
Are any of your research participants located outside of the UK, e.g., will you be gathering data through Skype interviews with participants located overseas?		No
Will any of the fieldwork or research require you to travel outside of the UK to collect data?		No

If you have answered 'yes' to either part of question (5), please describe the ethical issues raised with: gaining consent and gathering data from participants located overseas, securely storing and transferring data from the field back to the UK, any cultural issues that may be relevant. Please outline your plans to resolve this on a separate page and ensure that you have completed a risk assessment form. Agree this with your supervisor and submit it with this form.

You MAY be referred for light touch or full ethical review if you are unable to demonstrate that you have resolved the ethical issues relating to international research.

6. Personal safety Where will any fieldwork/ interviews/ focus groups take place?	Yes	No
At the university or other public place (please specify below).		No
At my home address		No
At the research subject's home address		No
Some other location (please specify below).		

If you conduct fieldwork anywhere except at the university or other public place you need to review security issues with your supervisor and have them confirmed by the Module Leader who may refer you for light touch or full ethical review. Write a brief statement indicating any security/personal safety issues arising for you and/or for your participants, explaining how these will be managed. Agree this with your supervisor and submit it with this form.

Please note that conducting fieldwork at the research subject's home address will require strong justification and is generally not encouraged.

A risk assessment is required before any data is gathered for any dissertation project, please view the Health and Safety advice on the module's Minerva pages.

7. Anonymity	Yes	No
Is there any potential for data to be traced back to individuals or organisations, for instance because it has been unanonymised or anonymised in such a way that there remains risk (eg highlighting people's positions within an organisation, which may reveal them).		No

If you have answered 'yes' to question 7, please discuss this further with your supervisor. You need to provide a strong justification for this decision on a separate sheet. This application will need to be reviewed by the dissertation Module Leader and may require a full ethical review.

8. Data management issues

Will the research involve any of the following activities at any stage (including identification of potential research participants)?		Yes	No
Examination of personal records by those who would not normally have access			No
Sharing data with other organisations			No
Use of personal addresses, postcodes, faxes, e-mails or telephone numbers			No
Publication of direct quotations from respondents			No
Publication of data that might allow identification of individuals to be identified			No
Use of audio/visual recording devices			No
Storage of personal data on any of the following:			No
	FLASH memory or other portable storage devices		
	Home or other personal computers		
	Private company computers		
	Laptop computers		

If you have answered 'yes' to any of the questions under 8, you must ensure that you follow the University of Leeds Information Protection Policy: http://www.leeds.ac.uk/informationsecurity and the Research Data Management Policy: http://library.leeds.ac.uk/research-data-policies#activate-tab1_university_research_data_policy.

You are obliged to provide a copy of your anonymised data to your supervisor for their records and to destroy other copies of your data when your degree has been confirmed.

Dissertation Research Ethical Approval: Declaration

For students	Please tick as appropriate
Option 1: I will NOT conduct fieldwork with (data on) human participants for my dissertation.	
Option 2: I will conduct fieldwork with (data on) human participants for my dissertation.	ü

For options 1 and 2 - I confirm that:

The research ethics form is accurate to the best of my knowledge.

I have consulted the University of Leeds Research Ethics Policy available at http://ris.leeds.ac.uk/ResearchEthicsPolicies.

I understand that ethical approval will only apply to the project I have outlined in this application and that I will need to re-apply, should my plans change substantially.

For option 2 only:

I am aware of the University of Leeds protocols for ethical research, in particular in respect to protocols on informed consent, verbal consent, reimbursement for participants and low risk observation. If any are applicable to me, signing this form confirms that I will carry out my work in accordance with them. http://ris.leeds.ac.uk/PlanningResearch

Student's signature: Jayakumar Sukumar

Date: 23-08-2024

For supervisors	Yes	No
No further action required		
I confirm that the dissertation is in line with the module's block ethical approval (Part A & question 8).	x	
I have discussed the ethical issues arising from the research with the student and agree that these have been accurately and fully addressed.	x	
I have reviewed the student's research proposal.	x	
I have reviewed the student's Risk Assessment Form.		
Further actions required		
Refer to dissertation Module Leader for further review / discussion.		
The dissertation falls outside the module's block ethical approval and the student was advised to apply for full ethical review.		

Supervisor’s signature: …W Abouarghoub.....…….……..

Date: ……24/08/2024…….....……..

QUALTRICS SURVEY:

Dear Participant,

My name is Jayakumar Sukumar, and I am currently pursuing an MSc in Business Analytics and Decision Sciences. As part of my dissertation, I am conducting research on the topic "Factors Affecting Predictive Analysis in Forecasting Election Outcomes."

This study aims to explore how various factors, such as Brexit, social media, and NHS waiting times, influence voter behaviour and election predictions in the UK. Your participation in this survey is crucial, as your insights will significantly contribute to the understanding of these influences.

The survey will take approximately 5–10 minutes to complete. Your responses will provide valuable data that will help in analysing current trends and improving the accuracy of election forecasts.

Data Privacy Notice:

Please be assured that your privacy is of the utmost importance. All responses will be kept strictly anonymous, and no personally identifiable information will be collected. The data gathered will be used solely for academic research purposes and will be securely stored. You may choose to skip any question that you do not feel comfortable answering.

Thank you for your time and participation. Your contribution is greatly appreciated and will have a meaningful impact on this research.

- Yes, I agree

Gender

- Male
- Female
- Non-binary / third gender
- Prefer not to say

Ethnicity

- White
- Black or African American

- American Indian or Alaska Native
- Asian
- Native Hawaiian or Pacific Islander
- Other

Age

- Under 18
- 18 - 24
- 25 - 34
- 35 - 44
- 45 - 54
- 55 - 64
- 65 - 74
- 75 - 84
- 85 or older

Education

- High school graduate
- Undergraduate
- Postgraduate
- Doctorate

Employment Status

- Employed full time
- Employed part time
- Unemployed looking for work
- Unemployed not looking for work
- Retired
- Student
- Disabled

How did you vote in the 2016 Brexit referendum?

- Voted to Leave
- Voted to Remain
- Did not vote

Have you changed your political party affiliation since the Brexit referendum?

- Yes, I switched to a different party
- No, I remained with the same party
- I became undecided

Did your Brexit mentality affect how you voted in the general elections of 2019 or 2023?

- Yes, it was a major factor
- Yes, but it was one of many factors
- No, it did not influence my vote

How accurate do you believe election polls have been in predicting outcomes since the Brexit referendum?

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

How likely are you to base your future voting decisions on a party's stance on Brexit?

- Extremely likely
- Somewhat likely
- Neither likely nor unlikely
- Somewhat unlikely
- Extremely unlikely

Which social media platform do you use most frequently for political news?

- Twitter (X)
- Facebook
- Instagram
- TikTok
- YouTube
- Other

How do you assess the trustworthiness of political information on social media compared to traditional news outlets?

- Extremely appropriate
- Somewhat appropriate
- Neither appropriate nor inappropriate
- Somewhat inappropriate
- Extremely inappropriate

Can you recall a specific instance where the content you encountered on social media directly influenced your voting intention?

- Yes, multiple times
- Yes, once
- No, social media has not influenced my voting intention

How often do you engage with political content on social media (liking, sharing, commenting)?

- Extremely likely
- Somewhat likely
- Neither likely nor unlikely
- Somewhat unlikely
- Extremely unlikely

How do you feel about political advertising on social media, and how does it impact your perception of the candidates or issues?

- Extremely positive
- Somewhat positive
- Neither positive nor negative
- Somewhat negative
- Extremely negative

In the past two years, have you or someone you know experienced long waiting times for NHS services?

- Yes, I have experienced long waiting times
- Yes, someone I know has experienced long waiting times
- No, I haven’t experienced long waiting times

How have long waiting times for NHS services affected your trust in the government's ability to manage public health?

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

To what extent did NHS waiting times influence your voting decision in the most recent election?

- Far too much
- Slightly too much
- Neither too much nor too little
- Slightly too little
- Far too little

How would you rate the government's handling of NHS resources during the COVID-19 pandemic?

- Excellent
- Good
- Average
- Poor
- Very Poor

How likely are you to base your future voting decisions on a party's stance on NHS funding and management?

- Extremely likely
- Somewhat likely
- Neither likely nor unlikely
- Somewhat unlikely
- Extremely unlikely

We thank you for your time spent taking this survey.

Your response has been recorded.