DS105 - Sentipedia

Q: How can we use data science to analyse the sentiment of different groups regarding Brexit?

Introduction

Since 2016, Brexit has arguably emerged as the most polarizing subject in UK politics, triggering the resignation of then-Prime Minister David Cameron and exerting a profound influence on the policies of subsequent Prime Ministers.

The Brexit campaign was historic in its use of technology, with debates raging on all the notable social media platforms, including Twitter, Facebook, and YouTube. Due to the widespread usage of these platforms and the ease of communication they provide, people discussed their opinions here in vast numbers. Beyond social media, other uses of technology include targeted ads using data analytics and profiling as well as campaign specific apps being created.

We remembered that the outcome of the Brexit referendum came as a shock to many pollsters – many of which were confident that overall people would vote to stay within the European Union. We thought that this surprise may be due to the methods of calculating public sentiment used by pollsters.

Despite massive changes in the way people communicate and express their opinions, pollsters relied heavily on traditional methods of finding out about public opinion. We thought that perhaps if we used data from social media sites, specifically Reddit, we may get a better idea about public sentiment.

We chose to analyse Reddit threads because it offers us an insight into the minds of a demographic which may be underrepresented in traditional polls. The majority of Reddit users are aged between 18 and 29, a group that we thought may not be extremely open to answering polling questions. Additionally, since Reddit users responded in real time to key events, speeches, and advertisement campaigns, we can also aim to observe the impact they have on sentiment.

Wider implications of our research

If we are able to use data science techniques to gain a better understanding of public sentiment with regards to Brexit, this could have much wider implications. As Economists studying at the London School of Economics, we learn about the importance of expectations on macroeconomics. For example, inflation expectations impact the price level of the economy, which in turn effect the ideal rate of interest which the Bank of England should set. The Bank of England, alongside many other financial institutions, spend a vast amount of resources forecasting expectations. If data science can help them make more accurate forecasts, this could have real benefits in terms of GDP growth and reducing unemployment.

Furthermore, if we can analyse the responsiveness of sentiment to different political manoeuvres, politicians can make more informed decisions on where to allocate their campaign budgets. During the 2016 Brexit campaign, Vote Leave and other Leave campaigns spent a combined £4.45 million. If future campaigns are able to see what outlets provide the most impact per pound, political campaigners could save money by avoiding ineffective expenses and focus on what actually influences people.

https://thrivemyway.com/reddit-statistics/#:~:text=The%20average%20age%20of%20Reddit,of%20getting%20their%20college%20degrees.

https://www.statista.com/statistics/1031993/eu-referendum-campaign-groups-advertising-spend-in-the-uk/