Analyzing the national variations in the political sentiments using Twitter data – A case of Scottish Independence

Author: Monsuru Adepeju

In the three and half years since the Brexit referendum, the conversations around Scotland’s independence, dampened momentarily by a failed referendum in 2014, has been reinvigorated with the ‘certainty’ of Brexit, thanks to the December 2019 UK general election. Scottish independence is the political movement for Scotland to become a sovereign state, independent from the United Kingdom. Following the general election, there has been an incessant conversation, on Twitter, about the potentials of another Scottish referendum. These conversations (posts & tweets) can be identified based on the various tags on them (e.g. ‘#Indyref2’, ‘#scottishreferendum, etc) and downloaded for analysis. Using these data (having excluded ‘retweets’ and ‘replies’) between the January 1 and January 25 2020 (days before Brexit), I analyse and compare the expressed political sentiments in the conversations, across the four constituents nations (i.e. England, Wales, Northern Ireland and Scotland) of the UK.

Figure 1 shows that the majority of the tweets (73%) were sent out from the mainland of Scotland, while another 24% from England. Northern Ireland and Wales have small share of 1% and 2%, respectively (Inserted is the map of the UK, showing the relative position of the countries).

Figure 1. Percentage of tweets about Scottish Independence (on Twitter) in the United Kingdom between January 1st and January 31st 2020.

Importance of words

In Figure 2, the ‘Wordclouds’ is used to show the relative importance of the words used in the posts across each country. The bigger and bolder a word appears, the more often it is mentioned in the posts and the more important it is. Dominant words, such as “Indiref2”, “Scotland”, “Scottish”, “independence”, and all hashtags have been filtered out in order to enable clearer representation. There are similarities and differences between the four countries. Words such as “Brexit”, “Boris”, “Johnson”, “Sturgeon”, etc. can be seen to be of high importance across all countries. These are words that are directly connected with the discussion of a 2nd referendum in a near future. For example, the Prime Minister Boris Johnson has just officially rejected the request from the First Minister Nicola Sturgeon to grant the Scottish people another referendum (<https://www.telegraph.co.uk/politics/2020/01/14/boris-johnson-officially-rejects-second-independence-referendum/>).

In terms of the differences, words such as “indywales” and “union” can be identified in Wales, and words such as “snp” and “sturgeon” can be identified in Scotland. These are words of significance within the political context of each country. For example, “Indywales”, is a word that has been attributed to the rising nationalist sentiments in Wales for the past few years, while the word “snp” is an acronym for the Scottish National Party.

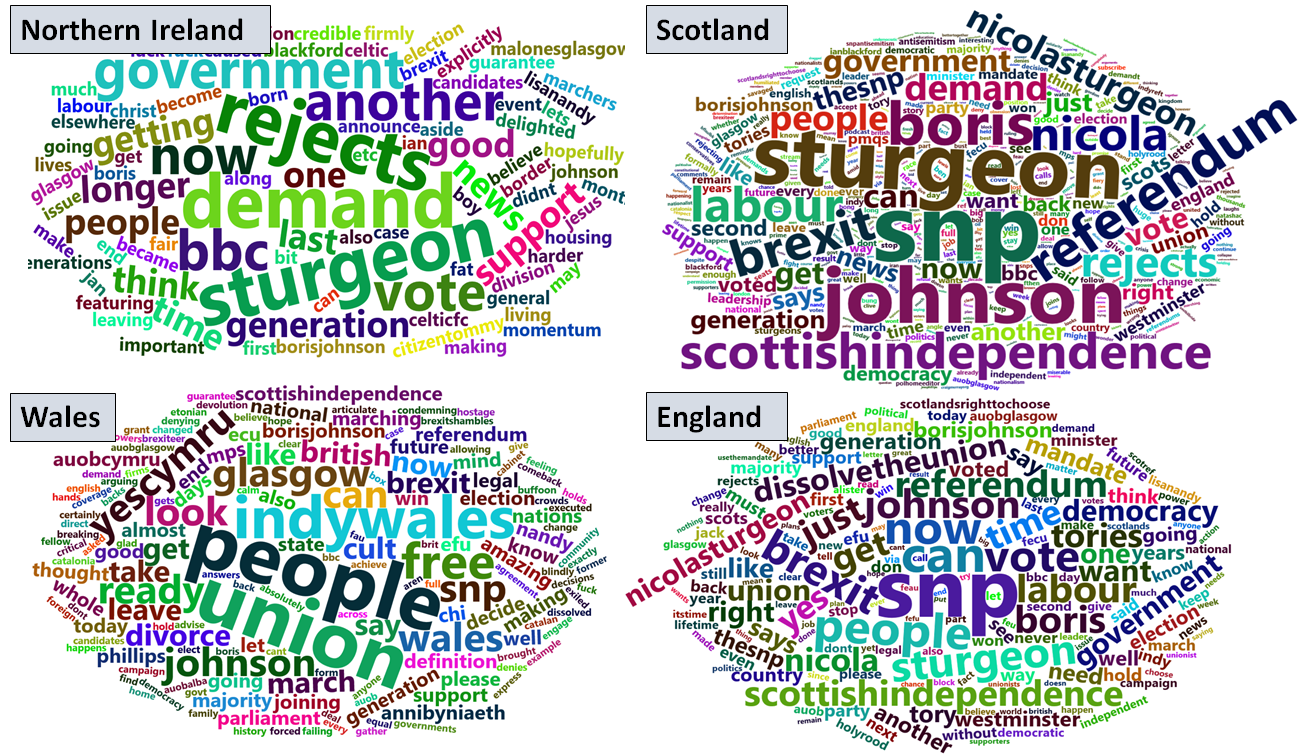


Figure 2. Importance of words

Mapping the sentiment of tweets

I used two sentiment lexicons in order to determine the type of emotion expressed in a tweet. First, is a polarity classifier and the second is the emotion classifier. The polarity classifier categorizes words in a binary fashion into positive and negative sentiments. The outcome of polarity classification is represented in Figure 3, which nicely compares the percentage of tweets with positive and negative sentiment for each country. In Wales and Northern Ireland, the majority of people expressed positive sentiments in their post, with 69% and 51%, respectively. This is in contrast to England and Scotland where the majority expressed negative sentiment, with 52% and 53%, respectively.

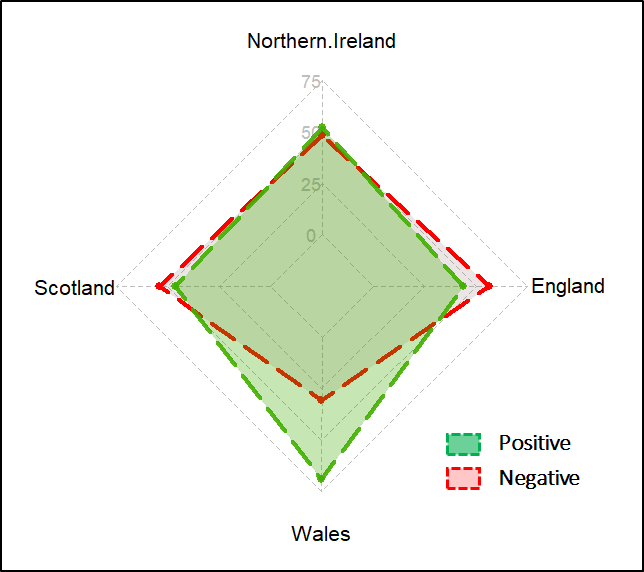


Fig. 3 Polarity sentiment (%)

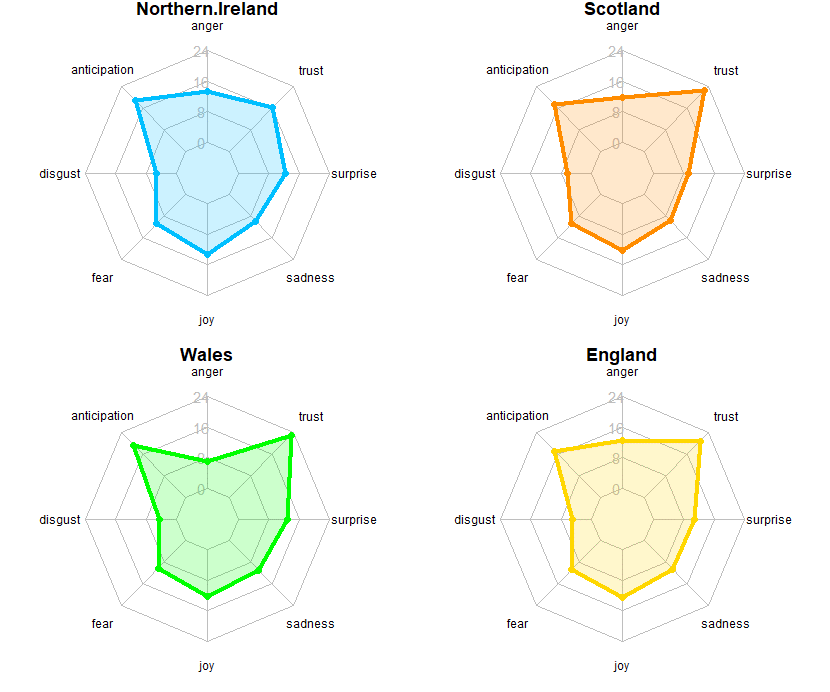


Figure 3. Emotion sentiment (%)

Second, the emotion classifier categorises words into several emotional statuses, allowing a broader insights into the emotion behind the posts. The categories of emotions include ‘anger’, ‘anticipation’, ‘disgust’, ‘fear’, ‘joy’, ‘sadness’, ‘surprise’, and ‘trust’. Overall, the patterns of the sentiments are comparable across all the four countries (see Figure 4). ‘Anticipation’ and ‘trust’ are two most expressed sentiments, with Wales showing slightly higher percentages in both categories. ‘Disgust’ is the least expressed sentiments in all the four countries. More analysis is needed in order to provide the justifications for these observed patterns, and their association with the polarity of sentiment in Figure 3.

CONCLUSION

Whilst the polarity of sentiment reveal that tweets originating from both England and Scotland have comparable percentages, with majority showing negative sentiments, the same is true between Northern Ireland and Wales, but with majority showing positive sentiment. The variances in the size of tweets with Wales and Northern Ireland having significantly small number of tweet on the subject of Scottish Independent. However, the content of the tweets remain similar in the most part.