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 a continuous conversations, mostly on Twitter, about the potentials of another Scottish referendum. The tweets relating to this subject can be identified using the associated tags, such as ‘#Indyref2’, and ‘#scottishreferendum. By downloading these data sets (excluding ‘retweets’ and ‘replies’) between the January 1 and January 25, 2020 (days before Brexit), I examine the variations in the sentiments expressed in these data sets, across the four constituents nations (i.e. England, Wales, Northern Ireland and Scotland) of the UK. Figure 1 shows that the majority (73%) of the tweets relating to Scottish independence were sent out from the mainland of Scotland, while another 24% were from England. Wales and Northern Ireland have small shares of 1% and 2%, respectively.

Figure 1. Percentage of tweets about Scottish Independence (on Twitter) in the United Kingdom between January 1st and January 31st 2020 (Inserted is the map of the UK, showing the relative position of the four countries).

Words used in tweets

In Figure 2, the ‘Wordclouds’ is used to highlight the most commonly used words in the conversation, across the country. The bigger and bolder a word appears, the more often it is mentioned in the posts and the more important it is. Expected words, such as ‘Indiref2’, ‘Scotland, ‘Scottish’, and ‘independence’, and hashtags have all been filtered out, in order to enable clearer visualization.

There are similarities and differences across the four countries. Words, such as ‘Brexit’, ‘Boris’, ‘Johnson’, and ‘Sturgeon’ are amongst the most commonly used words. These are words/names that are directly associated with the prospect of a 2nd referendum in a near future. A case in point, on the 14th of January 2020, the UK Prime Minister ‘Boris Johnson’ officially rejected the request from the Scottish 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/>).

Distinct words in each country with high importance are limited. The words tends to describe certain political sentiment in each country, in relation to the subject of Scottish independence. For example, the word ‘referendum’ in Scotland clearly emphasizes the discussion around the call for a second referendum. In Wales, words such as ‘Indywales’ and ‘yescymru’, can be attributed to similar calls for independence, due to rising nationalist sentiments across Wales over the past few years. Yet, word like ‘union’ can be seen to convey a sentiment that is directly opposite to the aforementioned. Although, it is difficult to derive the context of usage of these words, the technique provide a simplified rundown of conversations of a subject matter.

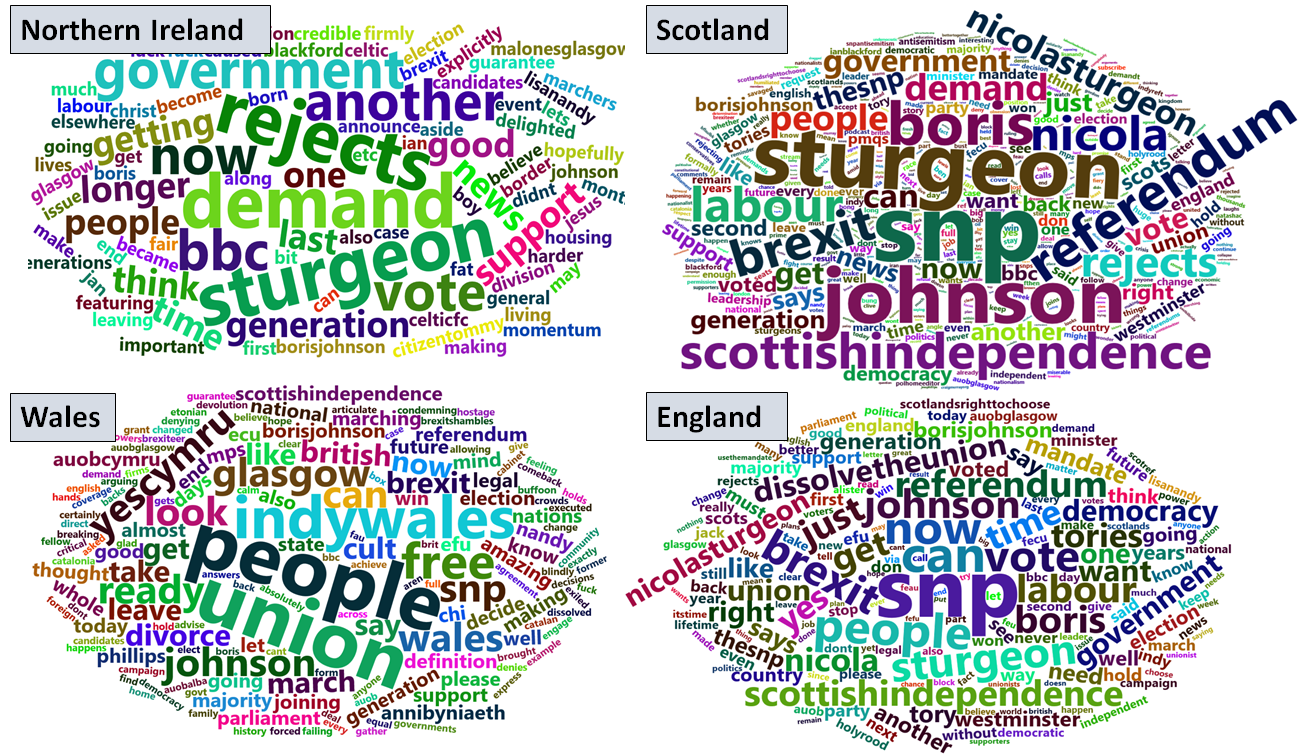


Figure 2. Words used

Sentiment extraction of tweets

An improved technique, *sentiment mapping,* is used extract the type of emotions expressed in a conversation, using a programmable sentiment lexicons. Two popular lexicons were used. First, a polarity classifier which categorizes words in a binary fashion into positive and negative sentiments. Figure 3 shows the percentage of tweets with positive and negative sentiment for each country. Wales and Northern Ireland have higher positive sentiments (69% and 51%, respectively) which contrast England and Scotland with higher negative sentiments (69% and 52%, respectively).

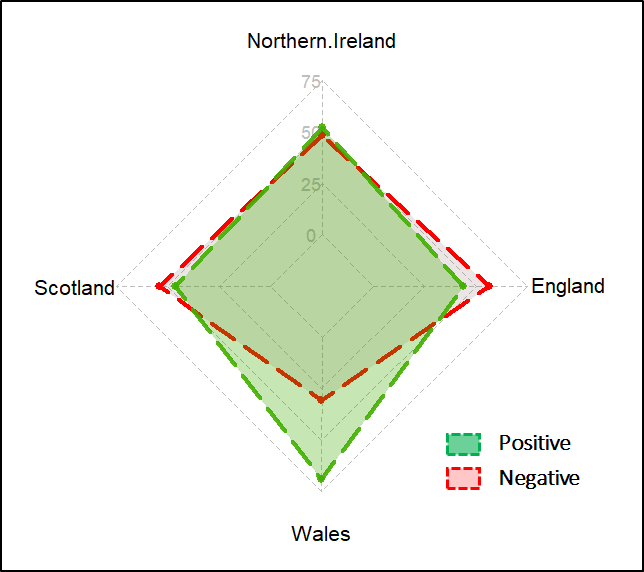


Figure 3 Polarity sentiment (%)

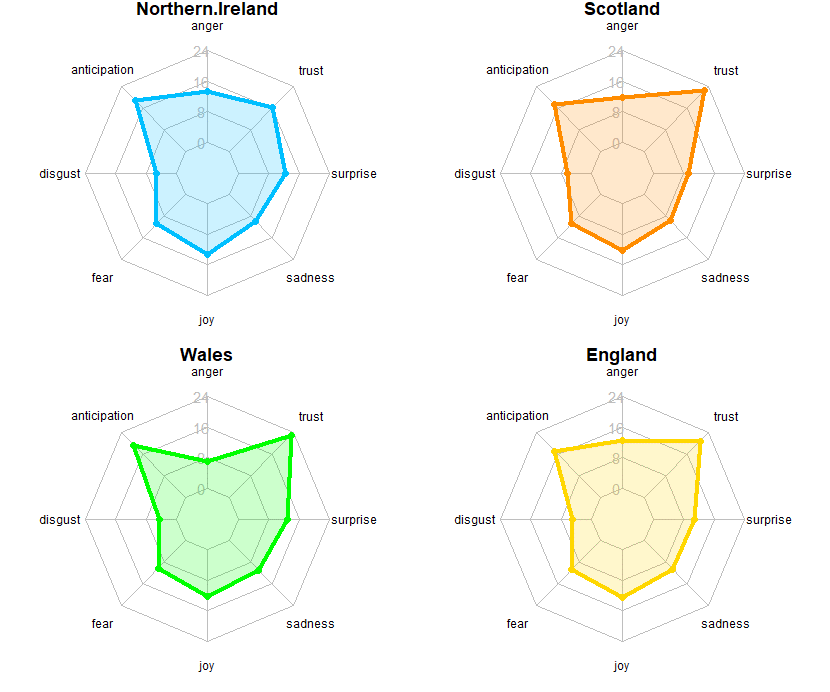


Figure 3. Emotion sentiment (%)

Second, an emotion classifier is used to categorize words into several emotional classes, allowing a broader insights into the underlying sentiments of the tweets. The emotional classes include ‘anger’, ‘anticipation’, ‘disgust’, ‘fear’, ‘joy’, ‘sadness’, ‘surprise’, and ‘trust’. Overall, the patterns of the sentiments are comparable across all the four countries (shown in Figure 4). ‘Anticipation’ and ‘trust’ are the two most expressed sentiments across all countries, 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.

Normalise each result by the count of each classes in each word.

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.

Notice:

The without any political influence. article does not have nor does the article sponsored by any political influence partu