OSM-1: Social Media (Twitter) Analysis[1]

Attributes were selected based on policy discussions and government guidance observed in European countries where the infection waves were concentrated. To look more into the chosen attributes, we conducted a social media analysis. We extracted attributes-related tweets between 15th February 2020, and 19th May 2020, using the Twitter standard search application programming interface (API) consisting of a set of predefined expressions (see below), which are the most widely used news media terms relating to the novel coronavirus (COVID-19). 15,000 tweets with expressions related to the attributes were extracted for our social media analysis. Only English language tweets were extracted.

We analysed the extracted tweets using **sentiment analysis,** which is a standard procedure in text mining literature.[1,2,3] **Sentiment analysis**[2] is the process of computationally identifying and categorising opinions expressed in a piece of text, especially to determine whether the tweeter's attitude towards a particular topic, product, etc. is positive, negative, or neutral. The key objective of sentiment analysis is to gauge opinions, identify hidden sentiments and finally to classify their polarity into positive, negative or neutral.[4]

Although we selected attributes based on policy discussions and government guidance observed globally and previous DCE studies on preferences to control emerging infectious diseases, we used the twitter analysis to gain insights into how the selected attributes are discussed on social media. We used the sentiment analysis to categorise opinions in the text related to our attributes. We identified sentiments that people have when talking about the attributes of interest. For example, do people use the term "excess death" in their tweets? What sentiments do people have when talking about excess death? Should we use excess death or number of deaths as an attribute? Are people more concerned about excess death or the number of deaths? The sentiment scores helped us to gain insights into these questions. They allowed us to identify what attitudes or sentiments people have when communicating about the attributes we selected (lockdown restrictions, number of infections, excess death, hospital capacity, income loss and job loss) on social media.

The x-axis in all the histogram plots shows the sentiment score as a negative and positive integer or zero. The sentiment score is the sum of sentiment values assigned to parts of the sentence (or textual field) and can be less than -1 or more than 1, as shown in all the histogram plots. A positive score represents positive or good sentiments associated with a tweet. In contrast, a negative score represents negative or bad sentiments associated with that tweet. A score of zero indicates neutral sentiment. The more negative the score, the more negative the sentiments of the person tweeting and vice-versa.

[1] Date of Tweets Extraction: 19th May 2020 (Tweets from 15th February 2020-19th May 2020).
[2] Sentiment Scale: 0-Neutral >0-Positive <0-Negative

Type of lockdown: this refers to how strict the lockdown measures are. The sentiment score of each term related to the attribute 'lockdown type' (lockdown restrictions, lockdown rules, and lockdown policy) is presented in Figure 1 and Table 1. It can be seen that the significant portion of tweets of 'lockdown restrictions' and 'lockdown policy' fall in negative sentiment category. The 'lockdown restrictions' tweets displayed 53% of negative sentiment, 19% of positive sentiment, and 28% of neutral sentiment (Table 1). In the think-aloud (TA) interviews, the lockdown attribute was initially presented as "lockdown type". The group of attributes representing policy choices (lockdown severity, lockdown length, and postponed procedures) was described in a very similar way as "type of lockdown". Following the insights we get from the sentiment analysis and the TA interviews, we renamed the attribute to "lockdown restrictions".

Figure 1. Histogram of tweets about lockdown restrictions and lockdown policy

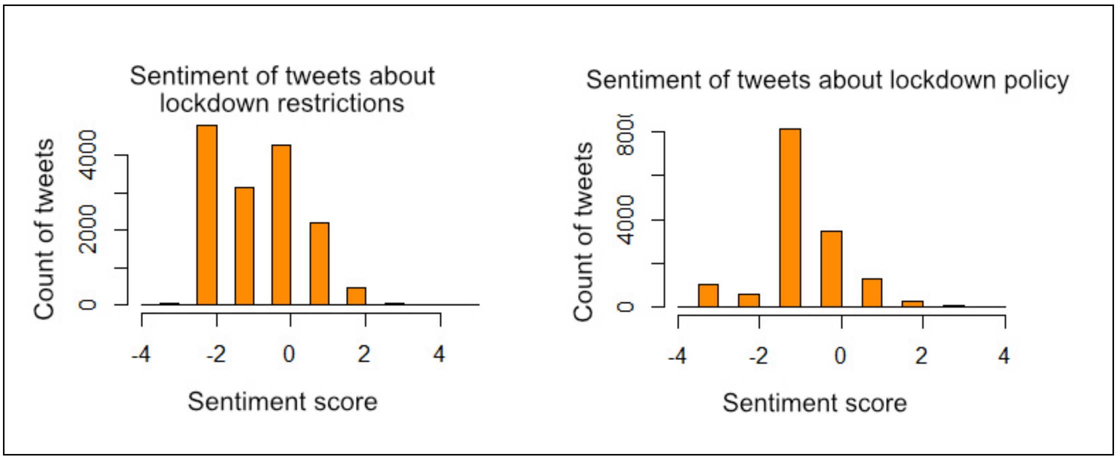


Table 1. Percentage of sentiment scores of tweets about lockdown restrictions and lockdown policy

Sentiment types	Lockdown restrictions	Lockdown policy
Negative	53%	65%
Neutral	28%	23%
Positive	19%	12%

Lockdown length: this refers to the number of months the lockdown will be in effect. For the sentiment analysis, we used related terms such as lockdown period and lockdown weeks. For the terms 'lockdown period' and 'lockdown weeks', the major portion of the tweets fall in the neutral sentiment category (Figure 2 and Table 2).

Figure 2. Histogram of tweets about lockdown period and lockdown weeks

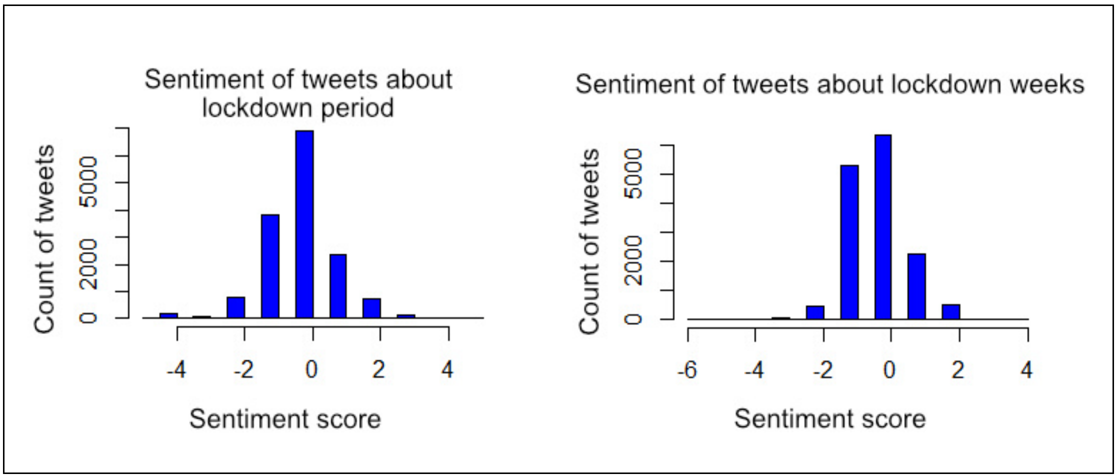
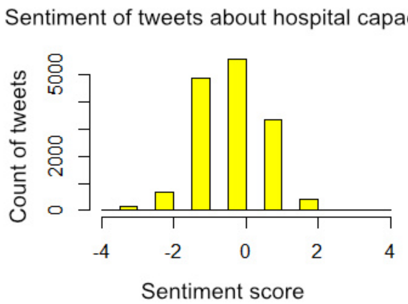


Table 2. Percentage of sentiment scores of tweets about lockdown period and lockdown weeks

Sentiment types	Lockdown period	Lockdown weeks
Negative	33%	39%
Neutral	46%	42%
Positive	21%	19%

Postponement of usual non-urgent medical care: this refers to whether hospitals will postpone non-pandemic related medical procedures. We used the expression 'hospital capacity' in our twitter search. The sentiment score for tweets of hospital capacity is fairly symmetric (Figure 3), with 38% of the tweets generating a negative sentiment, 37% a neutral sentiment and 25% positive sentiment.

Figure 3. Histogram of tweets about hospital capacity



Negative sentiment (38%), neutral (37%), positive (25)

Excess deaths: this refers to the difference between the number of people who die during the pandemic, and the historical average for the same place and time of the year. An analysis of the 'excess death' tweets displayed 81% of negative sentiment, 4% of positive sentiment 15% of neutral sentiment (Table 3). It can be seen that the major portion of tweets about excess death fall in negative sentiment category (Figure 4). As the 'excess death' attribute displayed 81% negative sentiment, we were careful in the framing of the excess death attribute levels. Initially, we tested the presentation of 'excess death' attribute in absolute numbers, but this inflated its importance relative to other attributes. Therefore, we changed the presentation of excess death as fractions of 10,000.

Figure 4. Histogram of tweets about excess death and number of deaths

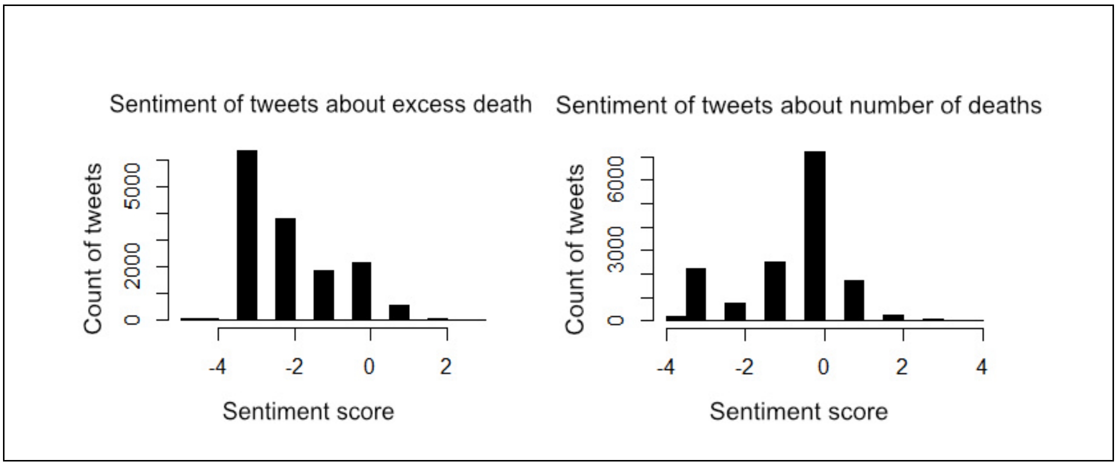


Table 3. Percentage of sentiment scores of tweets about excess death and number of deaths

Sentiment types	Excess death	Number of deaths
Negative	81%	38%
Neutral	15%	48%
Positive	4%	14%

The number of infections: the number of people who will be infected. We used the 'number of infections' and 'infection rate' in our tweet search. Tweets about the 'number of infections' displayed 53% of negative sentiment, 14% of positive sentiment, and 33% of neutral sentiment (Table 4). A major portion of tweets about 'excess death' fall in negative sentiment category (Figure 5). Initially, we tested the presentation of 'number of infections' attribute in absolute numbers, but this attribute, like the 'excess death' attribute, inflated its importance relative to other attributes. Therefore, we changed the presentation of the number of infections as fractions of 10,000.

Figure 5. Histogram of tweets about infection rate and number of infections

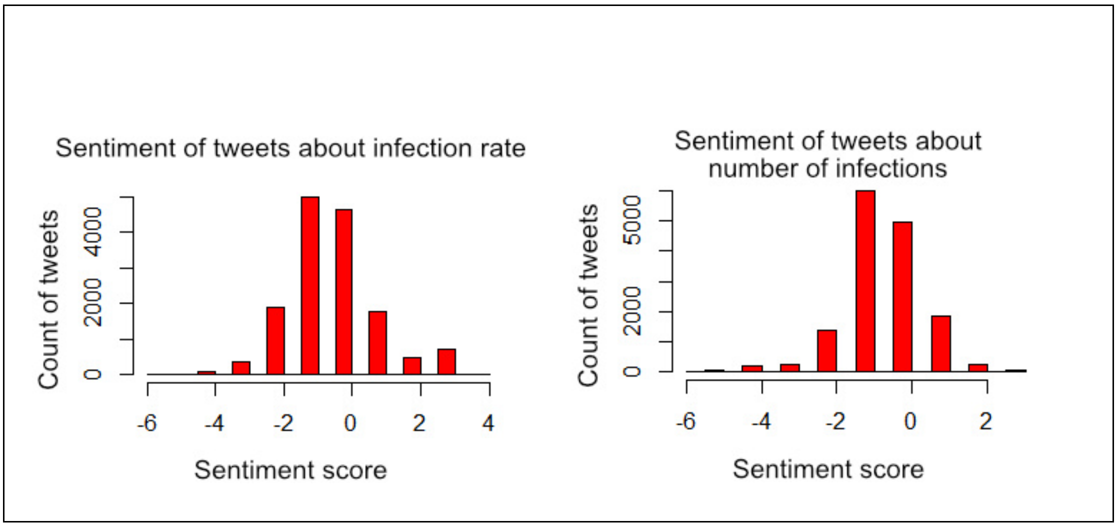


Table 4. Percentage of sentiment scores of tweets about the number of infection attribute

Sentiment types	Number of infections	Infection rate
Negative	53%	49%
Neutral	33%	31%
Positive	14%	20%

Ability to buy things: this refers to how much you will be able to afford one year from now compared to how much you would be able to afford normally. For this attribute, we used terms like 'inflation' and 'income loss' in our tweet search. A significant portion of (67%) of 'income loss' tweets generate negative sentiments. In comparison, a very small portion (12%) of tweets about 'income loss' suggested positive sentiments, while the remaining 21% are categorised as neutral tweets (Figure 6 and Table 5).

Figure 6. Histogram of tweets about inflation and income loss

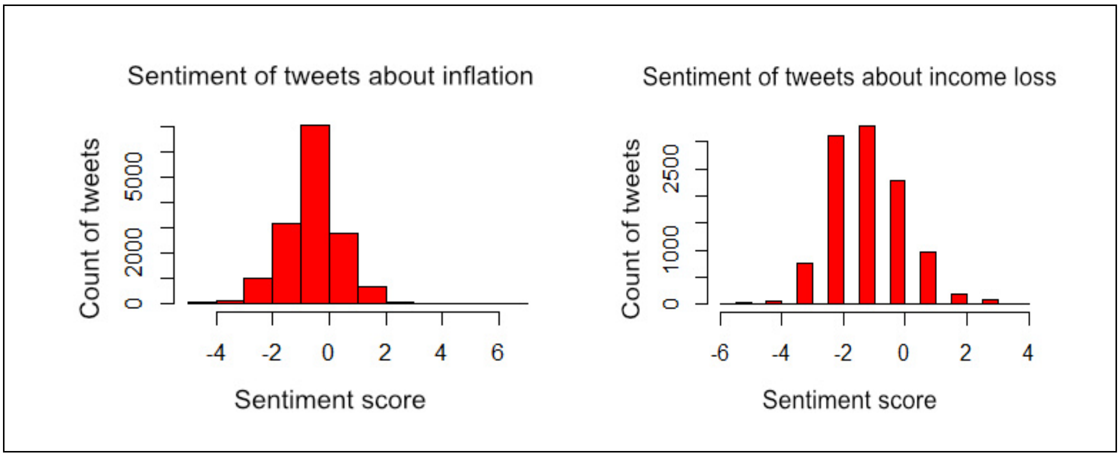


Table 5. Percentage of sentiment scores of tweets about inflation and income loss

Sentiment types	Inflation	Income loss
Negative	30%	67%
Neutral	47%	21%
Positive	23%	12%

Job losses: this refers to the proportion of people who will lose their jobs as a result of the lockdown. We used terms like unemployment and job loss in our tweet search. A major portion of the 'job loss' attribute (68%) generate negative sentiments (Figure 7 and Table 6). To make this attribute easier to understand, we used job loss instead of unemployment. Further, the higher sentiment attached to the job loss attribute would make the attribute easier to be traded-off when combined with other attributes.

Figure 7. Histogram of tweets about job loss and unemployment

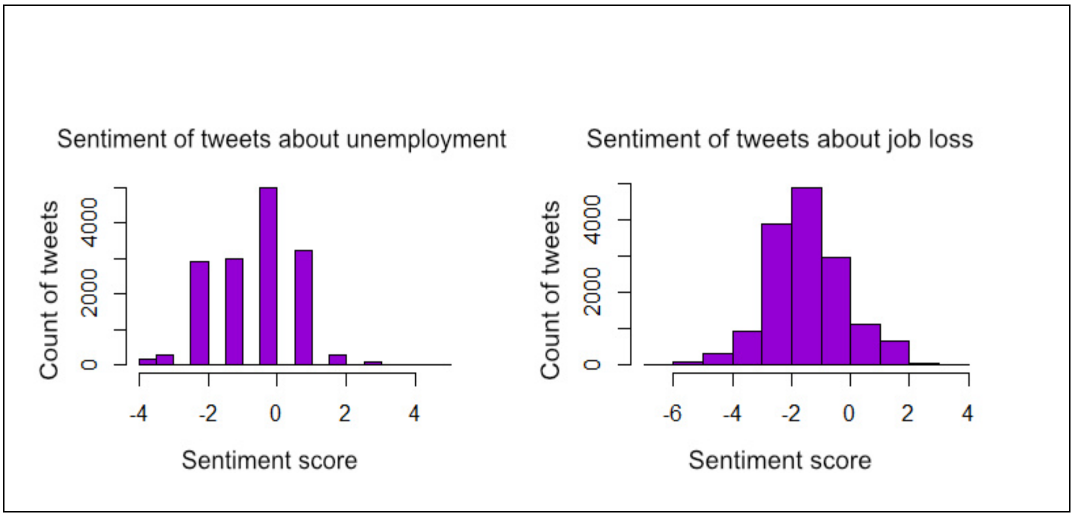


Table 6. Percentage of sentiment scores for the job loss attribute

Sentiment types	Unemployment	Job loss
Negative	43%	68%
Neutral	33%	20%
Positive	24%	12%

Additional References

1. Saif, H., He, Y., & Alani, H. (2012, November). Semantic sentiment analysis of twitter. In *International semantic web conference* (pp. 508-524). Springer, Berlin, Heidelberg.
2. Kumar, A., & Sebastian, T. M. (2012). Sentiment analysis on twitter. *International Journal of Computer Science Issues (IJCSI)*, *9*(4), 372.
3. Nirmala, C. R., Roopa, G. M., & Kumar, K. N. (2015, October). Twitter data analysis for unemployment crisis. In *2015 International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)* (pp. 420-423). IEEE.
4. Rajput, N. K., Grover, B. A., & Rathi, V. K. (2020). Word frequency and sentiment analysis of twitter messages during Coronavirus pandemic. *arXiv preprint arXiv:2004.03925*.

OSM-2: Opportunistic Think-Aloud Interviews

Virtual think-aloud (TA) interviews were conducted using MS Teams with colleagues from the University of Aberdeen (n=10) and members of our Stakeholder Advisory Group (SAG, n=3). Participants were asked to share their computer screen with the interviewers and to verbalise their thought processes whilst responding to the survey. As a warm-up, they were asked to think aloud whilst responding to the question: *"How many windows are there in your house?"* Respondents were told to consider the interviewer as a silent observer of their thought process. Interviewers did, however, encourage respondents to verbalise their thoughts if they were silent for a short period. Respondents were told there were no right or wrong answers. The interviews lasted between 45 and 90 minutes.

A number of changes were made to the DCE survey.

1. **Presentation of the excess death, number of infections, and job loss attributes**

In the TA interview used for internal testing, the attributes for excess death, number of infections, and job losses were presented differently. The number of jobs lost, and the number of people infected were presented as fractions of 100. In contrast, the excess death attribute was presented as absolute numbers of additional people dying over the expected figure during a normal year. This led to the excess death attribute dominating the choices of a significant number of participants, with some participants stating that they ignored all other attributes and considered the lower number of excess deaths presented in the choice task.

While this might be an expression of a valid preference, the feedback we received included evidence that the presentation of the excess death attribute in absolute numbers inflated its importance relative to other attributes. One participant stated that, while they recognised that job losses were presented as fractions, in their mind they ignored the denominator of the job loss attribute and directly compared its numerator to the absolute figures presented for the excess death attribute.

We changed the presentation of the excess death and number of infections attributes to be uniform across the choice task. The number of infections and excess deaths are now presented as fractions of 10,000.

2. **Presentation and placement of lockdown type attribute**

In the TA interview for internal testing, the colour-coded visual for the lockdown type attribute was prominently presented at the top of each choice option. Some participants interpreted the visual as a summary of the choice option rather than as an independent attribute.

We changed the position of the visual for the lockdown type attribute to make it appear next to the visual for the lockdown duration attribute.

3. **Visual presentation of the number of infections attribute**

The TA for internal testing displayed a *static* visual for the number of infections attribute that did not change according to the attribute level presented. Several participants stated that a changing visual would help them make better sense of the attribute level. We changed the visual to change with an increasing number of infections.

4. **Presentation of the shopping trolley attribute**

Initially, the text under the visual for the 'shopping trolley' attribute read "X% of the trolley." Some participants interpreted this to mean the economic impact on society rather than the economic impact on themselves. We changed the text to read "You can buy X% of the trolley."

5. Explanation of the shopping trolley attribute

Some participants were concerned that the initial explanation of the shopping trolley focused on consumption rather than the general cost of living. One participant expressed concerns that this might not accurately reflect the experiences of impoverished respondents. We expanded the explanation of the shopping trolley attribute to include housing costs and utility bills.

6. MFQ20: Likert scale anchors

The initial presentation of the MFQ20 presented the anchors for different points on a 6-point Likert scale ("not at all relevant" to "extremely relevant" and "strongly disagree to "strongly agree") at the top of the page. For the selection matrix, points on the scale were labelled with numbers running from 0-5 to mimic the presentation of the paper-based MFQ 20.

We observed that the top of the page was not visible for participants while they were answering the questions, leading them to spend a lot of time scrolling up and down on the page. We amended the selection matrix to display the anchors next to the numbered points on the Likert scale.

7. Ease-of-use updates

To make the survey more engaging, we made various improvements to the interface and presentation formats. This included a progress bar at the top of the screen, mouse-hover explanations for different selection options, and input prompts.

OSM-3: Social Media Ad for Think Aloud

Want to take part in a survey development about interventions to control a future pandemic?

Participate in our study!

- **We are trying to understand public preferences for interventions to control a future pandemic.**
- **We are asking for volunteers who are willing to support the design of a questionnaire using a process called "Think Aloud".**
- **A small gratuity (£20) will be offered for your participation.**

Where? Over Video Call
How long? 30-40 minutes
Who? 18 years or over living in the UK
If you are interested in supporting this research project in this way, please email heru@abdn.ac.uk to arrange a suitable time for this to take place.