Title: A cross country analyses of sentiment and emotions, before and after COVID-19, using big data

Authors:
Talita Greyling
School of Economics, College of Business and Economics, University of Johannesburg, South Africa, email: talitag@uj.ac.za
Stephanié Rossouw
Faculty of Business, Economics and Law, Auckland University of Technology, Auckland, New Zealand, email: stephanie.rossouw@aut.ac.nz

Abstract: The main aim of this project is to analyse and compare sentiment and emotions of people in three countries, before and after the introduction of regulations to limit the spread of the COVID-19 virus. The uniqueness in the choice of the countries lies in the different approaches taken by these governments to enforce self-isolation. On the more lenient side, you have Australia who only enacted complete lockdown 68 days after their first COVID-19 patient was confirmed. In the middle of the spectrum, you have New-Zealand, who has the entire country on mandatory lockdown, but still allows people to leave their houses for the purpose of exercise. On the stricter side, you find South-Africa, who has their military and police enforcing lockdown with severe repercussions for those attempting to leave their house for anything less than essential errands. In the last couple of weeks, we have seen the far-reaching effects of COVID-19 on economies, financial markets and industries, such as tourism, hospitality and 'SME's. This, in turn, affects many people, especially those in the informal sector, part-time employees or those who are self-employed. However severe these economic implications, it does not reveal the cost this pandemic has had on human well-being namely: the trauma of people who face the death of loved ones, the physical effects of being infected, the uncertainty over the future, fear and anxiety related to uncertain circumstances, the mental consequences of social distancing, such as anxiety and loneliness. Therefore, it is of utmost importance for governments to measure not only the economic and health impacts of COVID-19, but also the mental effect, as reflected in happiness levels and the emotions of people. Furthermore, it is essential to know how different measures taken by governments, to curb the spread of the virus, impacts happiness levels and emotions of their citizens, to better understand the impacts of regulations in unprecedented situations.

In this project, we use organic (big) data, in the form of Tweets extracted from the social media platform Twitter. We analyse each tweet, firstly to determine the sentiment expressed and secondly to determine the underpinning emotion. The sentiment of a tweet is measured as either being positive, neutral or negative. This data is submitted to an algorithm, whereby happiness scores are derived and used to create happiness indices for the different countries. The happiness indices are measured on scales from 0 (very unhappy) to 10 (very happy), with 5 being neutral (neither happy nor unhappy). On the emotion side, we differentiate between eight emotions underpinning the Tweets; anger, anticipation, disgust, fear, joy, sadness and surprise. By using the results, we determine the dominant emotions of a nation and not only the mood or sentiment. As we are using organic data, all data are available at real-time (see https://www.ghan.today). We have been collecting these datasets since April 2019, and
currently have a year's data. As the process of extracting data is continuous, the datsets increase daily. On average 120 000 Tweets are extracted per day and at the time of writing the project proposal there were almost 42,5 million Tweets in total. We will compare the happiness levels and emotions in each country before and after COVID-19 was confirmed. Thereafter, we will compare the deviations across countries to establish whether different types of regulations employed in these different countries has had different impacts on the happiness levels and emotional adaptations of their people.

The preliminary analysis shows significant decreases in the levels of happiness in all three countries from the start of COVID-19, compared to levels pre-COVID-19. There are also marked differences in the emotions expressed by people before and after COVID-19 in all three countries, interestingly the changes of emotions over this period differs between the three countries. The emotions most prominent across all three countries, before COVID-19 were joy, trust and anticipation, but after the announcement of the first COVID-19- patient, emotions changed dramatically, to fear, anger and distrust, however the intensity of these adaptations differed across the three countries.

**Data:** We analyse the sentiment and emotions of Tweets extracted from Twitter. The project has been running since April 2019. On average 120 000 Tweets are extracted per day and at the time of writing the project proposal there were almost 42,5 million Tweets in total. As the process of extracting data is continuous, the datsets increase daily. The dataset is ideal for the current study as it gives as the opportunity to compare sentiment and emotions before and after COVID-19, in Australia, New-Zealand and South Africa, three countries that used different methods to curb the spread of the disease. Please see [https://www.gnh.today](https://www.gnh.today) for data portrayed in almost real-time.

**JEL classification codes:** I3, C43, C55, C88

**Keywords:** COVID-19: sentiment; emotions; big data; regulations