

Twitter Text Analytics:

1. Report and analyse the top words from TF-IDF matrix (Home exercise)

We use keywords, “climate change”, “gas”, “oil” and “renewable energy”, to track the topics. The top words for important tweets are ポルトガルgp決勝予想 (which means Portugal gp final forecast, related to F1 competition, group, hot, hope, etc. For the first word, F1 competition may use lots of oil or gas, so the prediction for the competition appears in such topics. “Hot” and “heater” are frequently used when people talk about climate changes. The word “green” may be related to the “green house effect”, which are also related to climate and renewable energy. Other words are common words that people used in daily life.

2. Analyse and report whether the following method can be used as a reliable and accurate indicator of the important or trending topic. (Home exercise)

Based on the results, it seems that the method does not provide a reliable indicator of the trending topics on Twitter. There were some nonsensical words (e.g. high, hold, mark), so overall the words don't yield much actionable insight. To make this useful for further analysis, we may need to refine the code so as to ensure that we take into account the various nuances in tweets that could skew the results. For example, tweet rates could be misleading because the majority of tweets come from a small group of Twitter users; the collection of tweets or words is not necessarily indicative of importance or popularity.

Twitter Sentiment Analytics:

1. Analyze the difference between the sentiment of the public towards the vaccination as a whole versus the COVID-19 vaccination.

The sentiment of the public towards (general) vaccination is relatively neutral when compared with that towards COVID-19 vaccination. From our sentiment analysis on Vaccination Tweets, we use “vaccination” and “vaccine”. We can see that ‘inject’ and ‘free’, “hospital” seem to be some of the more common words when people think about getting a vaccine. On the other hand, for COVID-19 Vaccination Tweets, we use names of three main covid vaccinations among American, “covid” and “vaccination” to track the topics. We see words like ‘fight’, ‘lockdown’ and ‘yönetmeliği’(means regulation in English), which seems to carry more negative and anxious sentiments and be disliked by the public.