Evaluating public options of COVID-19 vaccine using natural language processing

**Problem description**

2020 is an unprecedented time and the whole world were almost shut down during the pandamic. Recently, two types of vaccines have been approved by FDA and they are like the light in the dark. Hopefully, vaccine can help end the pandemic and we can go back to normal life in 2021. However, there are different options about vaccines. People may welcome, oppose, or be skeptical about Vaccinations. It is important to find what people think.

**Approach**

I will utilize natural language processing approach in this project.

**Data sets**

My dataset is from https://ieee-dataport.org/open-access/coronavirus-covid-19-tweets-dataset. This website contains CSV files of the everyday tweets related to the COVID-19 starting from March 20, 2020. I will then clean the datasets by filtering only the COVID-19 vaccine related tweets.

**Applications and clients**

1. It would help government/companies to decide how many vaccines are needed
2. According to WHO, Community immunity requires at least 80% of the population to be vaccinated. As a result, knowing percentage of people who welcome, oppose, be skeptical is important.
3. It would help the government to decide whether they need to put more effort to persuade people to get vaccination.

**Deliverables**

Code, a final report