Keyword ideas:

Covid

Covid-19

Corona

Corona-virus

Coronavirus

Vaccination

Vaccinated

Vaccination

Vaccine

Pfizer

Johnson & johnson

Janssen

Sinovac

Sinon-vac

Moderna

Astrazeneca

Biontech

Sinopharm

Covilo

Coronavac

Covaxin

Vaxicode

Keywords used:

"vaccine hesitancy",

"covaxin",

"vaccination",

"vaccine",

"vaccinated"

Conditions that have to be satisfied to be be added to our set of tweets:

1. No retweet
2. Within 3 days counting back from the day of collection
3. Language has to be “en”
4. Stop when getting 1000 tweets

In parameters, use tweet.fields => language, created\_at for timezone,

For emotional notation,

0 -> negative

1 -> neutral

2 -> positive

Topics draft:

1. Variant
2. Children vaccination
3. Pro-vax
4. Stats / info / news
5. Vaccine policies
6. Encourage others to get vaccinated
7. Side effects

Final topics:

1. News/Facts/Stats/Science/questions (pure objective posts) ----------- news
2. Anything subjective :

2.1 Government-related opinion (including criticisms about politicians, and policies,

vaccine passport, mandate) ---------- gov

2.2 **Personal** life:

2.2.1 pro-vax personal opinion, any comment that is positive about the vaccine (include encouragement) ---------- pro

2.2.2 anti-vax person opinion (ineffectiveness, safeness, ) ---------- anti

(including opinion (eg. others should (not) get vaccinated), announcement (eg. i got vac/ i will never get vac.)

*For methods:*

We found out it was difficult to annotate without knowing the original tweet that the “replied\_to” tweet was answering to, making it difficult to annotate correctly.

^ This way of doing it is excluded from our method because it puts too much emphasis on one person’s idea.

*Some remarks that may be added to the report:*

* *Sometimes, it is incredibly difficult to class a tweet into a single category. Some of them are really long and talk about multiple subjects. Is it impossible to put them in a single one, hence, we have devised a decision hierarchy. This way, we can ALWAYS put a tweet into a single category.*
* Sometimes, the tweets don’t mention a clear context, making it difficult to understand what position they stand for.
* Some tweets are sarcastic. It can be impossible to determine the sarcasm without the tweet that they are replying to. You would think that removing tweets that are replied\_to could solve this issue, but this method would bring another issue: we may be putting too much emphasis on public figures who tweet often.
* Of course, we can be somewhat biased towards what the definition of a fact is since as student annotators, we are most likely to be pro vaccine. A so-called “fact” by us may be considered a “lie” by an unvaccinated person.
* Some people send a tweet: let’s say TWEET, but send another one that is almost the same, say (TWEET + @someone). It is the same message, but technically, it’s a different tweet.
* Some tweets contain emojis. We decided to not include them in the tf-idf calculations since emojis aren’t that common in tweets and there are too many of them for one of them to be significant enough

To do tf-idf:

1. Replace all punctuations with space
2. Then, put every word into lowercase
3. Remove words that are in stop\_words and those that still contain non alphabetic characters

| **Date** | **task** | **notes** |
| --- | --- | --- |
| Wednesday 24th (or lastest Friday 26th) | Collect 1000 tweets |  |
| Saturday 27th 7pm  zoom | Open coding / topics discussion |  |
| Tuesday 30th | Finish the annotations |  |
| Wednesday 1st 13h45 on campus after munch and mingle (mcconner 103 12h30) | Tf-idf, copy code from last assignments,  Data visualization with R/matplotlib,  Decide report separation | Not reached |
| Sunday 5th 20:00 | Having done all annotations, go for tf-idf |  |
| Thursday 9th 20:00 | Do some more data analysis and start report writing division of tasks. | It’s incredibly important to have beautiful data visualizations. It will be the main selling point of our report. |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
| Sunday 12th | Everything done |  |

Collections of tweets that were funny

“@Dragonwolflord @CyrilRamaphosa People need Jesus they don't need the vaccine. If they receive him we will forget about this pandemic.”

“I don't care about anyone stands in politics if they start out saying 'I am pro vaccine'. If you are pro vaccine you are against God. You are a follower of Satan and have willingly taken the mark of the beast. I don't care if you are a republican whose other policies align.”

“@USAFacts 200 deaths a day in the US. In Uttar Pradesh India zero deaths each day because of ivermectin and zinc. 5% vaccinated in that state-The pandemic has ended because of ivermectin! IVM, HCQ, quercetin and zinc stop C -19 from entering cells!”

Ah yes, zinc prevents covid from entering the body.

“@laralogan If the vaccine works why have more people died in 2021 than 2020. Before vaccine. It’s not real vaccine, it more like a flu shot.”

Of course, since a flu shot is not a vaccine.