LSE_DA201 Assignment

COVID-19 Patterns & Trends

Prepared for UK Government

July 2022

Zikomo Smith

Agenda

1. Results & Recommendations

2. Key patterns and trends

3. The data analysis process

4. Process considerations moving forward
   a) Use of GitHub
   b) Improving data collection standards
   c) New types of analysis

Results and Recommendations

Results of analysis

Row Labels	Total Vaccinations	Sum of Eligible for Second Dose	Cases	Deaths
Anguilla	4,709,072	222,398	644	1
Bermuda	2,690,908	127,073	5548	95
British Virgin Islands	4,933,315	232,988	2725	37
Cayman Islands	3,363,624	158,852	1011	2
Channel Islands	3,139,385	148,261	12135	100
Falkland Islands (Malvinas)	3,587,869	169,438	69	0
Gibraltar	5,606,041	264,745	5727	97
Isle of Man	4,036,345	190,639	8343	54
Montserrat	5,157,560	243,568	41	1
Others	2,466,669	116,482	8317439	138237
Saint Helena, Ascension and Tristan da Cunha	2,242,421	105,889	4	0
Turks and Caicos Islands	2,915,136	137,686	2910	23

- **Hospitalisations reached peak across all regions in March of 2021** and have subsided since.
- **The Isle of Man** and **Bermuda** saw significant increases in deaths towards the end of 2021.
- **Deaths have plateaued** across the remainder of the regions.
- **Data quality remains an issue** – the total vaccinations number is not reflective of reality (Anguilla has a population of 15,090.
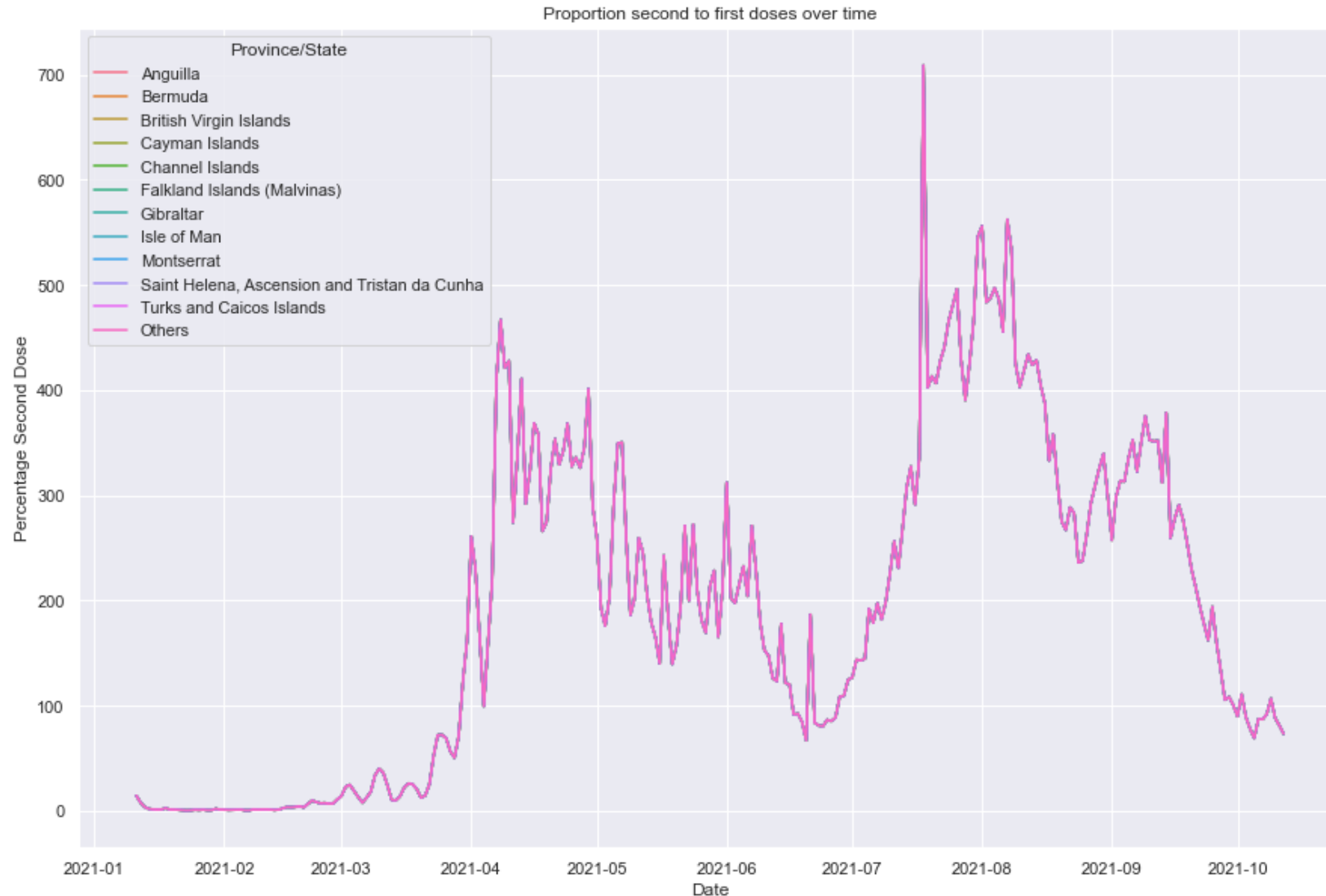- **COVID tweets tend to have negative sentiment**

Key recommendations

1. Target campaigns at areas with highest number of people eligible for second doses **(Gibraltar, Montserrat, British Virgin Islands.)**

2. Target campaigns at areas with increasing death rates to get them to work to keep themselves as safe as possible **(Isle of Man & Bermuda)**

3. **Include positive Twitter sentiment as a new campaign KPI**

4. **Implement data quality assurance process to ensure clean data** and to prevent repeating past mistakes.
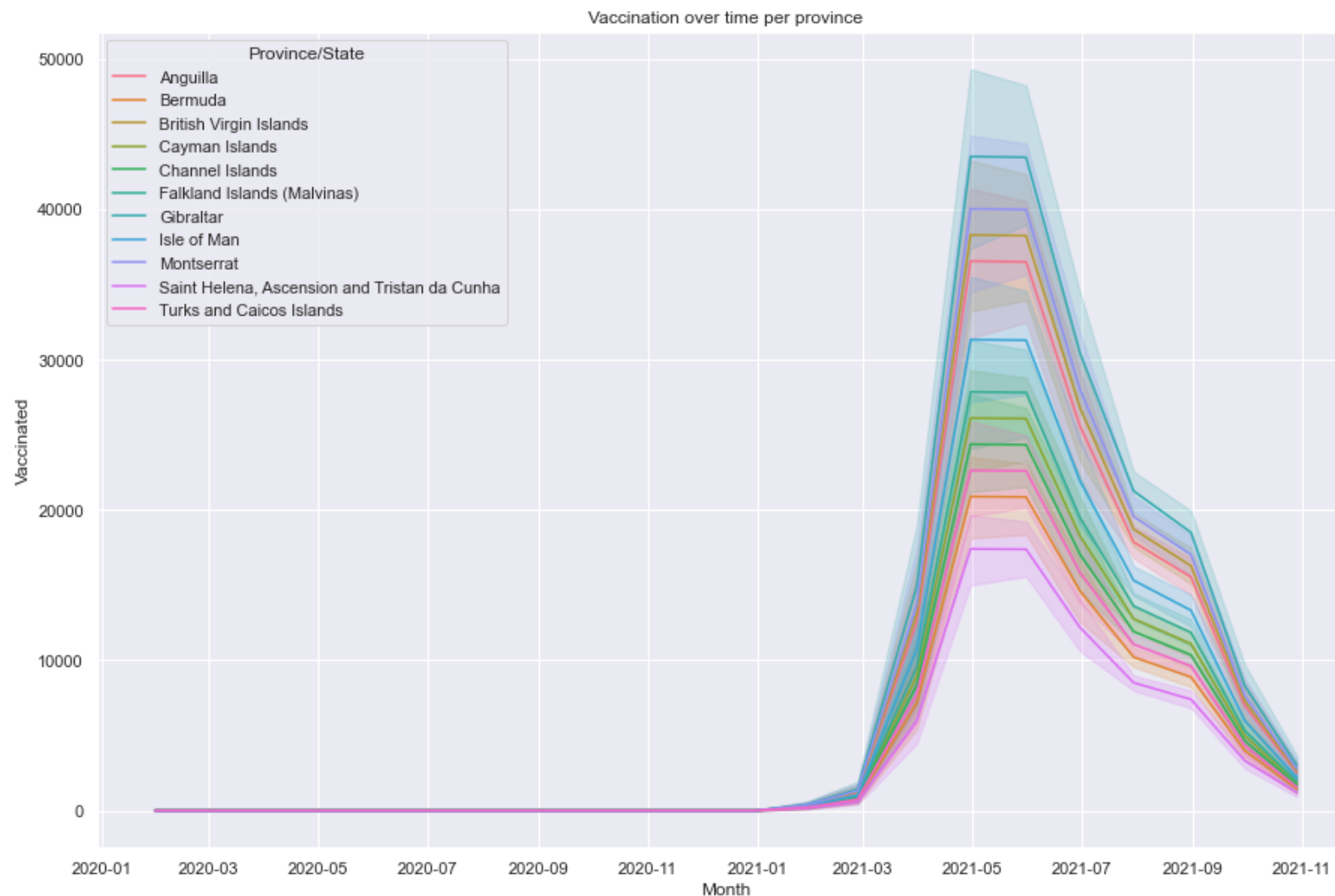
Key Trends

Second doses started to exceed first does in April 2021


Proportion second to first doses over time

Province/State
- Anguilla
- Bermuda
- British Virgin Islands
- Cayman Islands
- Channel Islands
- Falkland Islands (Malvinas)
- Gibraltar
- Isle of Man
- Montserrat
- Saint Helena, Ascension and Tristan da Cunha
- Turks and Caicos Islands
- Others

The UK Government is rolling out first and second vaccines in a similar manner across provinces

Vaccination over time per province

Cases no longer track hospitalisations, demonstrating the effectiveness of the vaccine.
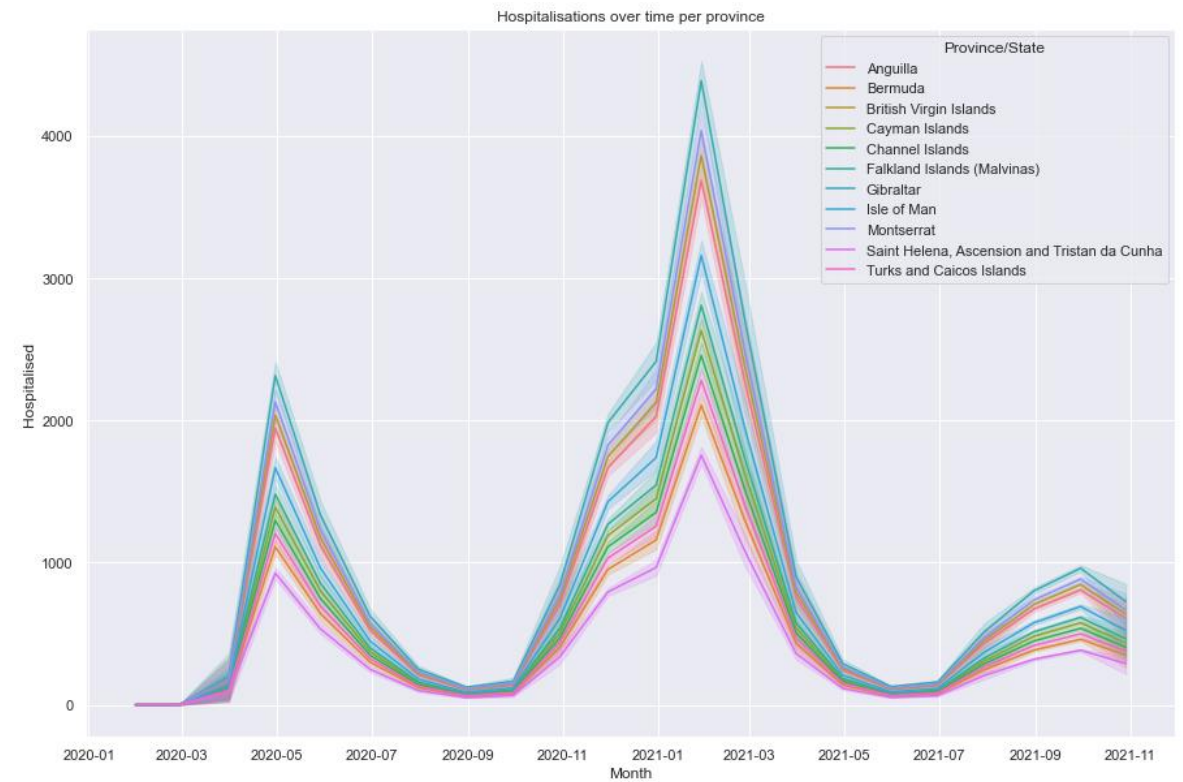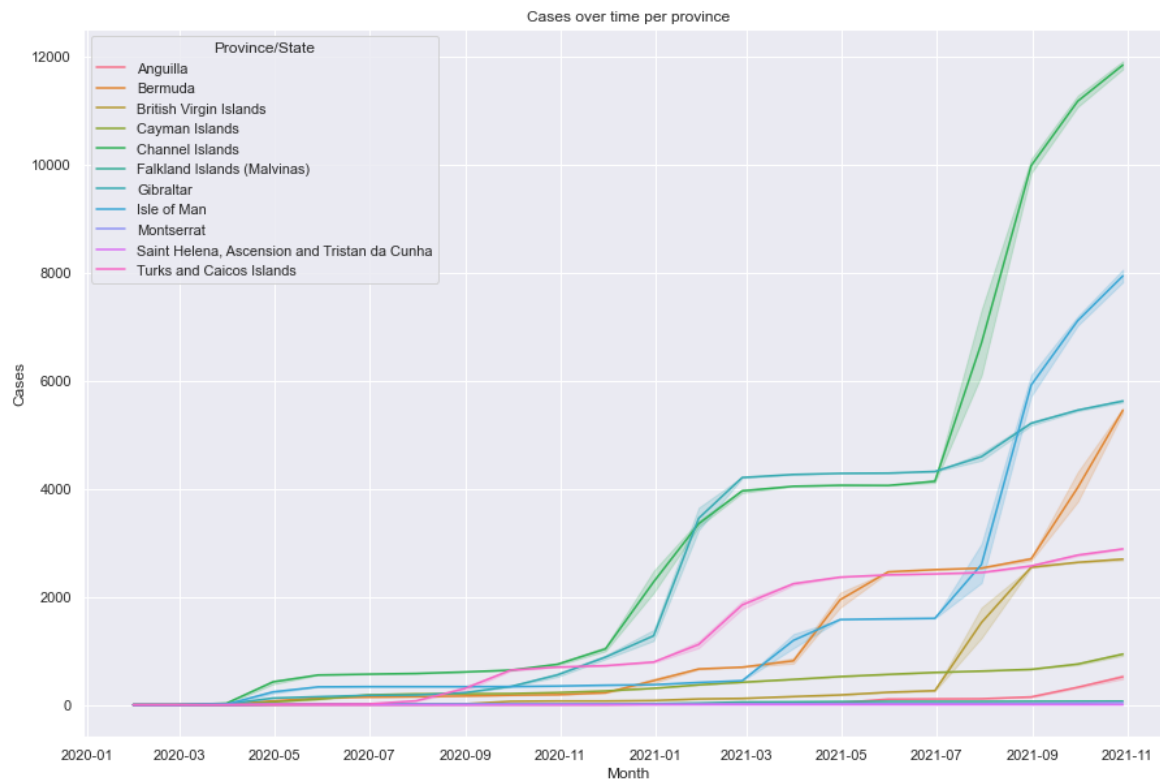
Each country is having a different COVID experience suggesting they require bespoke responses

Eligible for second dose

Deaths over time per province
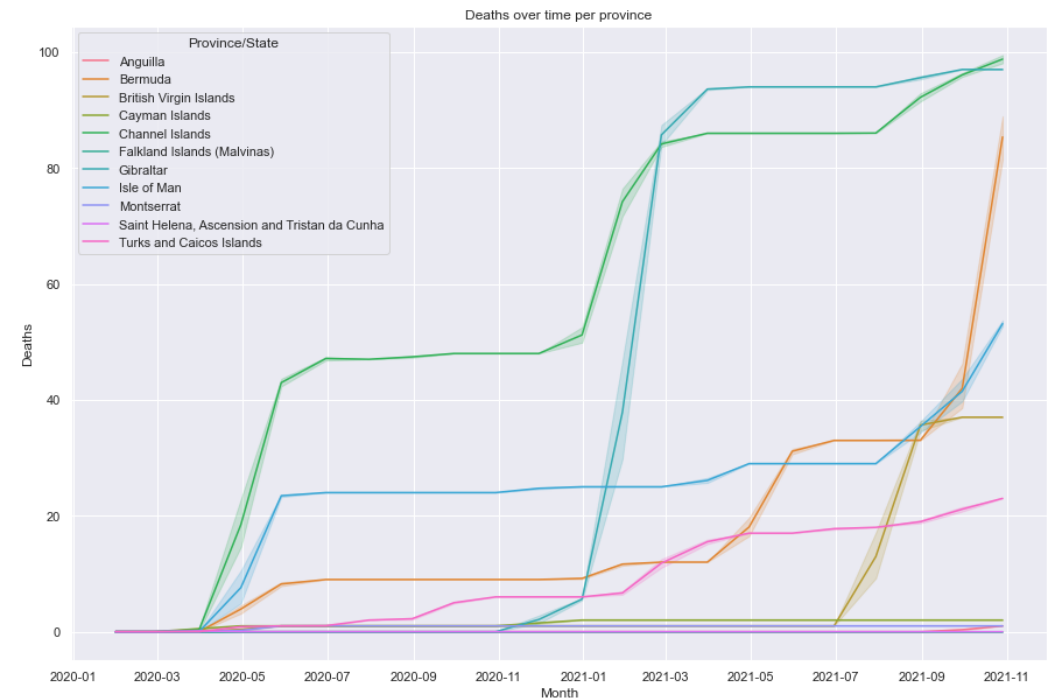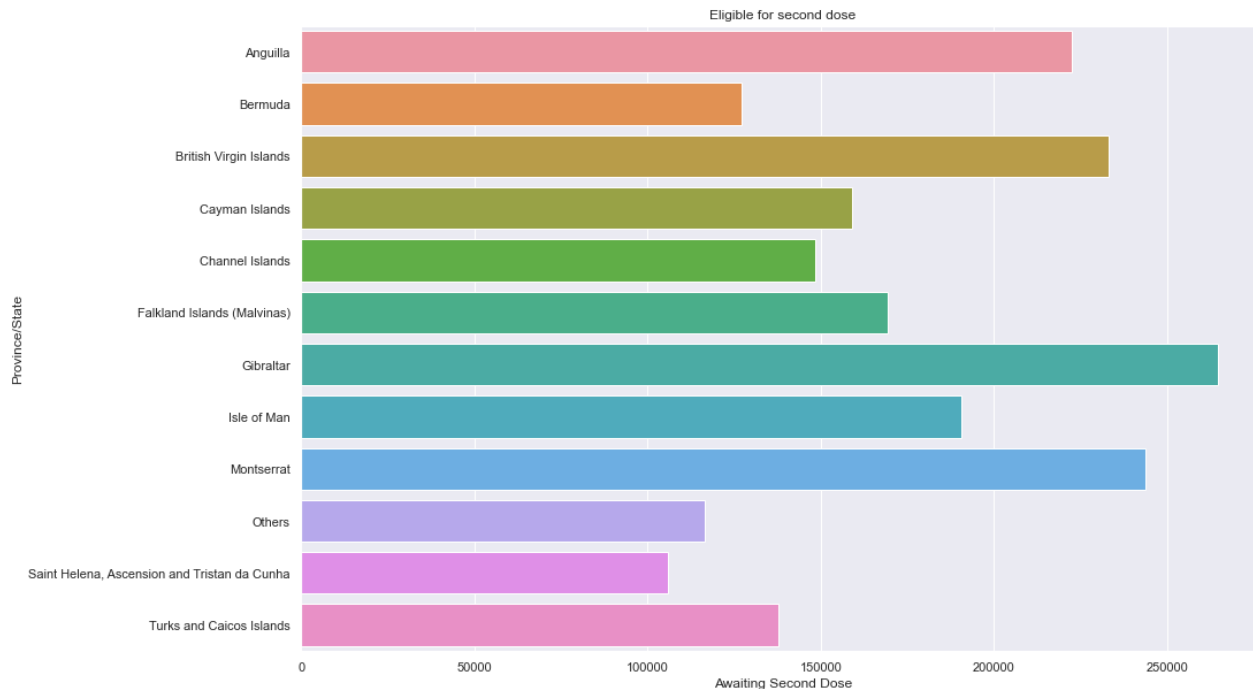
Covid tweets focus on political issues and topics likely to incur negative sentiment
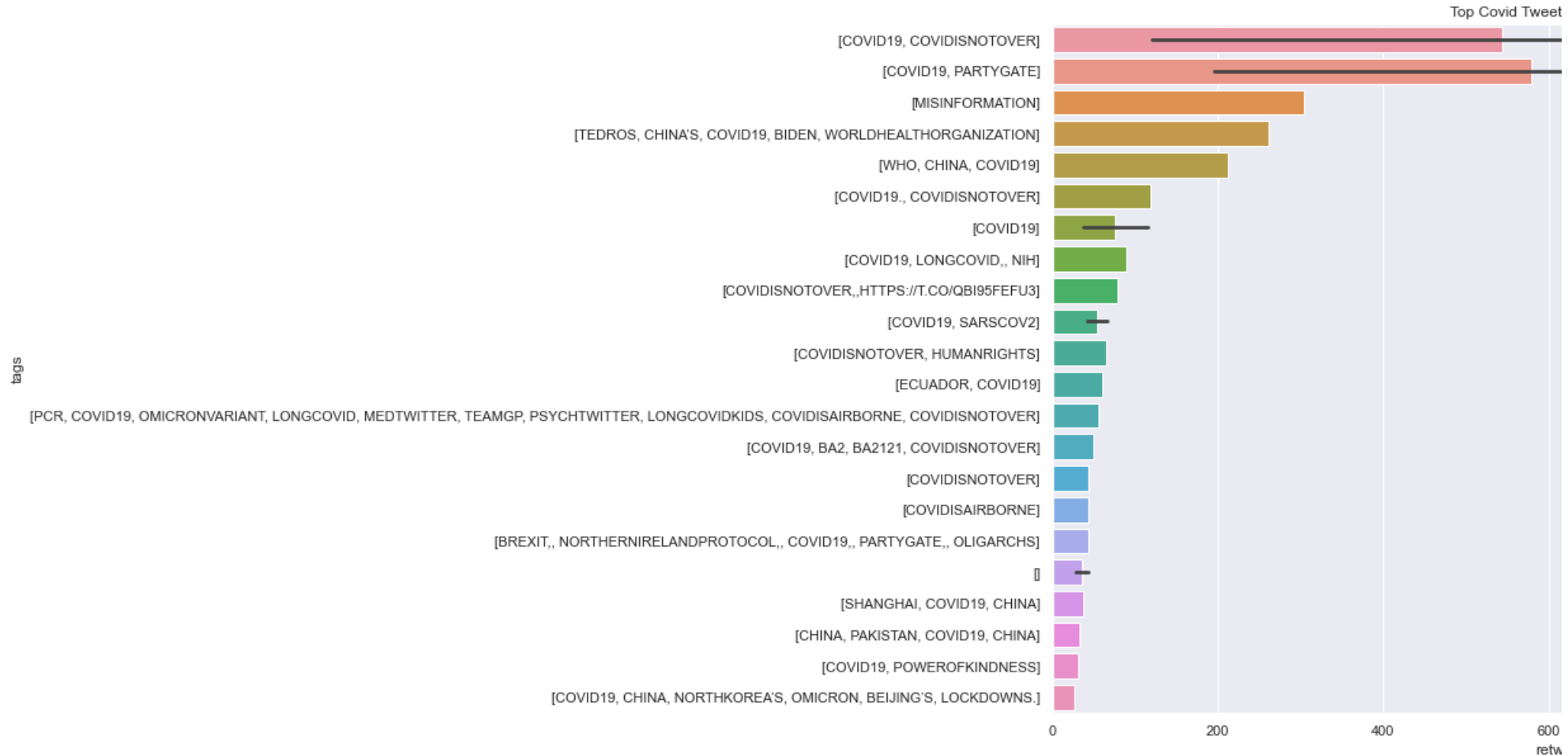
The Data Analysis Process

Analysis techniques and assumptions

1. The analysis used line plot graphs to assess whether the data was cumulative or not.
    a) Based on the analysis of Gibraltar, we found that death data was cumulative.

2. The analysis aimed to compare spikes in the non-cumulative date, with the plateaus and rises in the cumulative data to generate insights.

3. The analysis uncovered rows without data and backfilled them to ensure the completeness of the entire dataset.
    a) This was so the new numbers did not disturb the cumulative nature of the dataset.
    b) This was based on the assumption that data would be similar on days that were clustered together

4. The analysis aimed to provide multiple views of the dataset to allow us to find the most compelling insights.
    a) The analysis found that the 'Others' Province/State dataset was skewing the overall data so that it was removed from the final analysis.

5. The analysis aimed to assess the context of COVID tweets.
    a) This determined whether tweets focused on positive or negative stories by assessing clusters of hashtags.

Process Considerations Moving Forward

Future-proofing the analysis

1. Continue the use of GitHub
   a) Version control system reduces errors
   b) We can make data public to crowdsource solutions in the future

2. Improve data collection standards
   a) Vaccine data was not realistic
   b) The government needs to avoid reputational damage as has occurred in the past due to its COVID data collection practices

3. Incorporate machine learning techniques in future analysis
   a) The Twitter API allows us to access huge amounts of qualitative data.
   b) Machine learning will give us deeper insight into trends and sentiment beyond the top trending tweets.

THANK YOU!

- Zikomo Smith
- Zikomo.smith@mail.mcgill.ca