

Tracking People's Minds during the Spread of 2019-nCoV

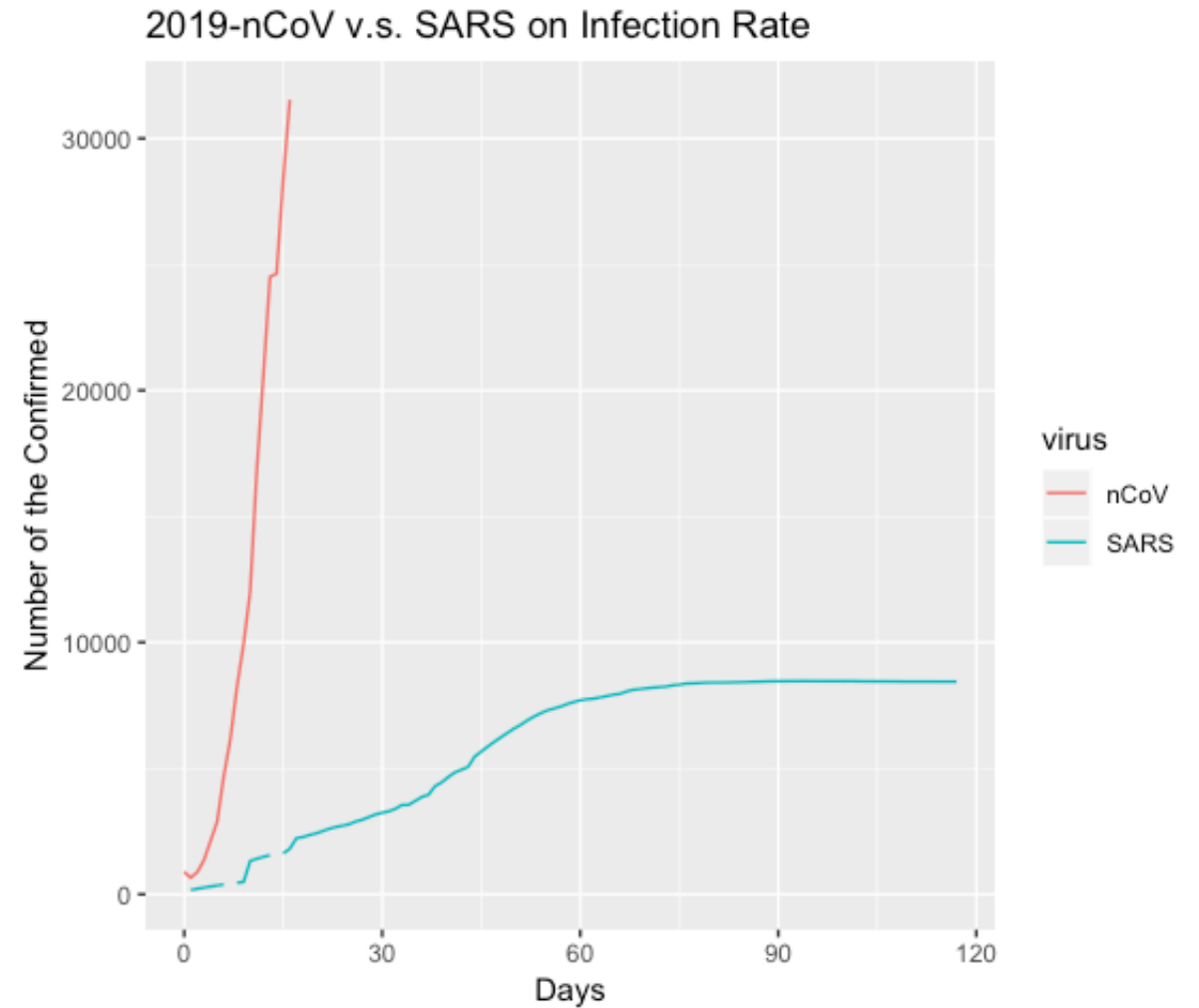
~Search Trends and Social Media Data Analysis~

by Team DSI Furniture

Yuki Nishimura, Amaury Sudrie, Derek Chen, Wangzhi Li

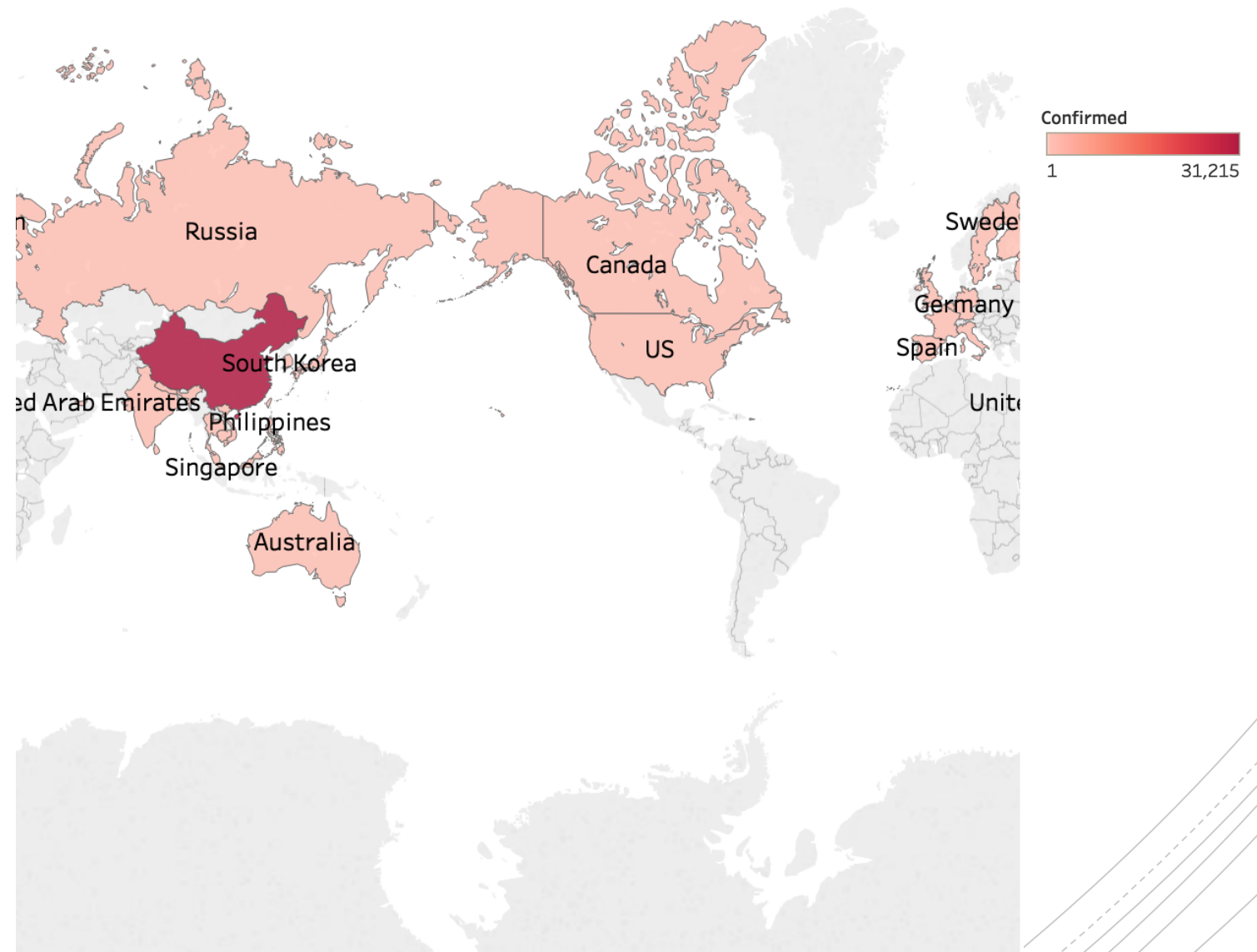
February 8, 2020

2019-nCoV has been
spreading at an
incredibly fast pace



Data from World Health Organization
<https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>
<https://www.who.int/csr/sars/country/en/>

2019-nCoV has been
spreading globally



Motivation

- 2019-nCoV has been spreading at an incredibly fast pace
- Were people mentally prepared to prevent infections and spreads?
 - Track people's minds during the spread

A large red speech bubble graphic with a white outline, pointing downwards. The word "Content" is written in white text inside the bubble.

Content

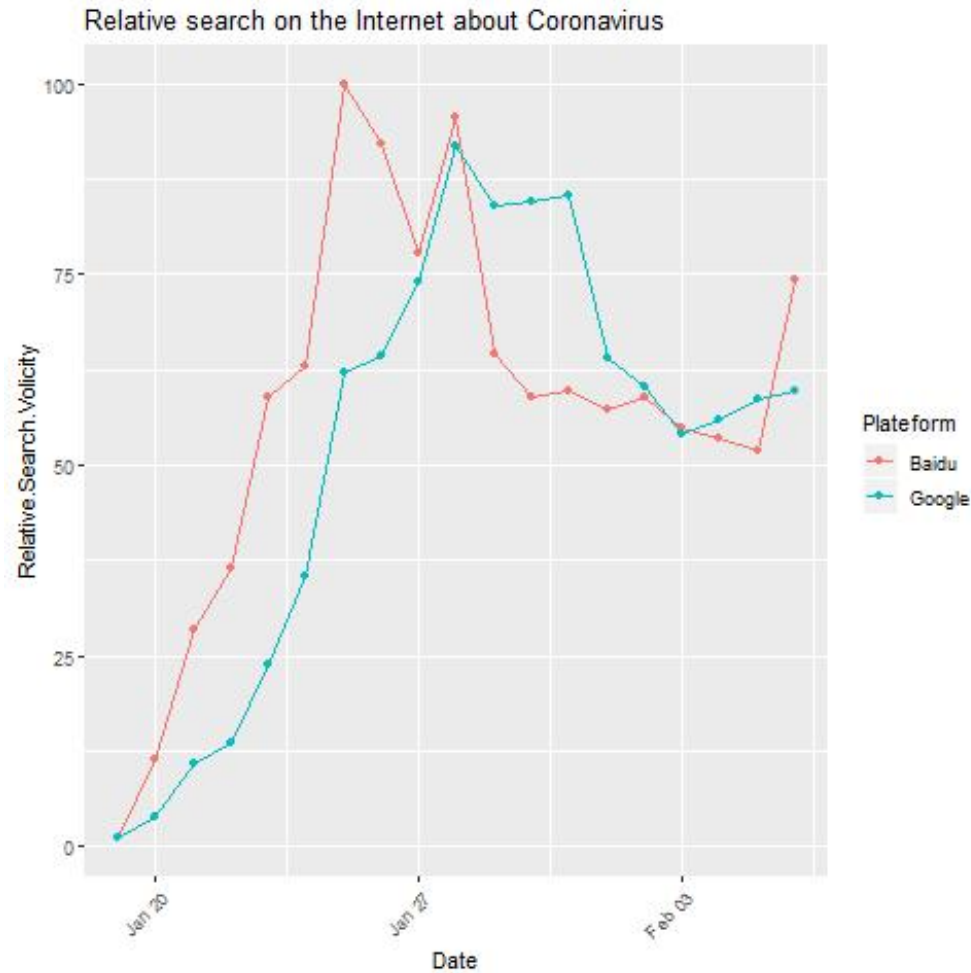
- **Tracking Minds**
 - Search Trends (Baidu and Google)
 - Tweet Sentiment (Weibo and Twitter)
- **References**

The background of the slide features a series of thin, curved lines in light gray and white, creating a sense of motion and depth. These lines are more prominent on the left side and fade towards the right.

Tracking Minds: Search Trends (Baidu and Google)

- From the Search Trends of the keyword “coronavirus”, we may see **when people are getting serious about the virus**
- We collected data of the search trend in China from Baidu, and the search trend around the world from Google
- 2019-nCoV data was collected from Kaggle

Tracking Minds: Search Trends (Baidu and Google)



There is a time lag between Baidu and Google, which seems to represent the lag between the spread of the virus.

This means that people start to get serious after the virus arrives to their country, which is not a good thing.

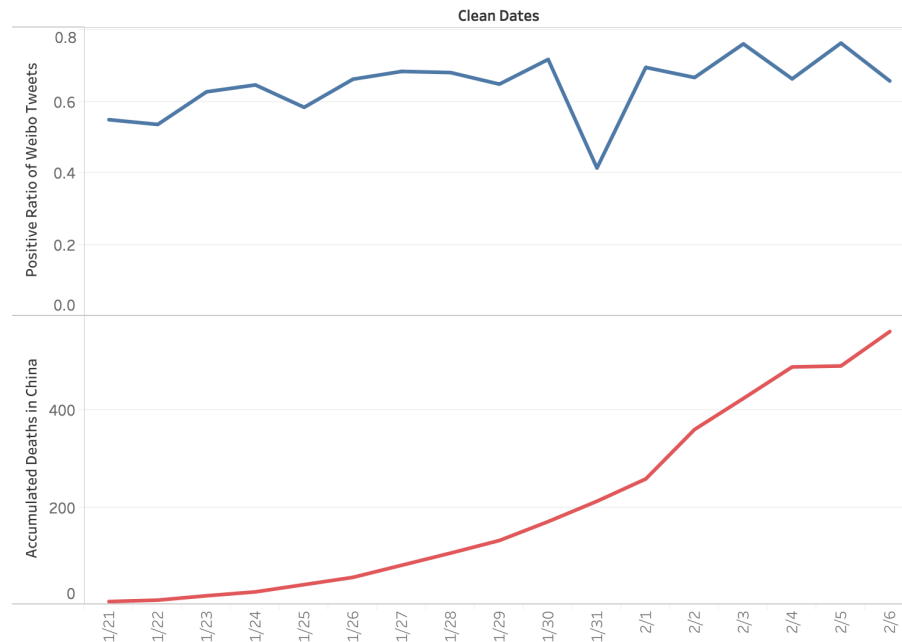
The background of the slide features a series of light gray, concentric curved lines that sweep across the frame, creating a sense of motion and depth. These lines are more prominent on the left side and fade towards the right.

Tracking Minds: Tweet Sentiment (Weibo and Twitter)

- From the Tweet Sentiment of tweets that include the keyword “coronavirus”, we may see **how people’s feelings toward the virus are changing over time**
- We collected tweet data (100 per day) in China from Weibo, and the tweet data (200 per day) in English-Speaking countries from Google
- Same 2019-nCoV data was used

Tracking Minds: Tweet Sentiment (Weibo and Twitter)

China Positive Ratio vs Deaths

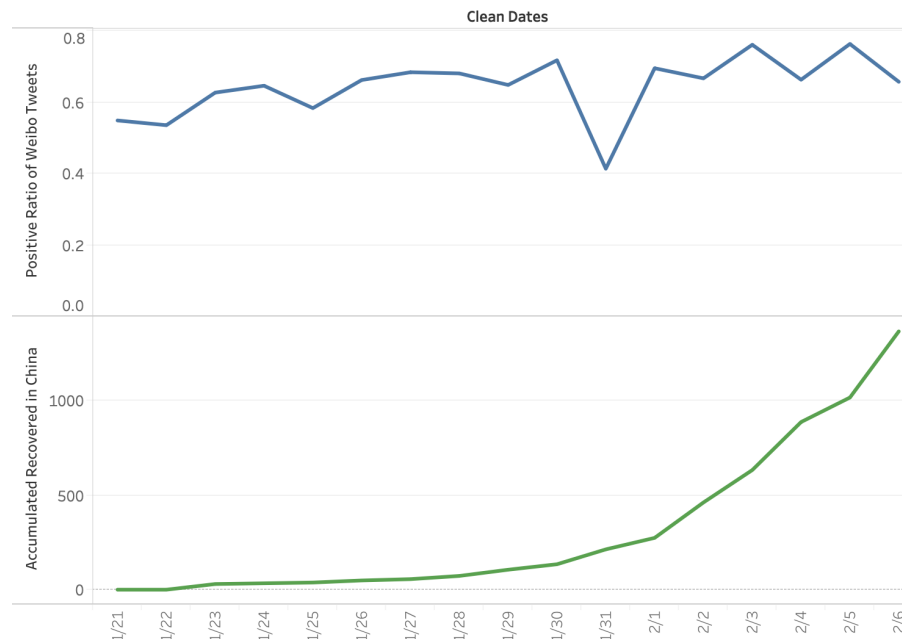


Weibo

Correlation:
0.51132316

There is a positive correlation against Weibo tweets for both Deaths and Recovered.

China Positive Ratio vs Recovered

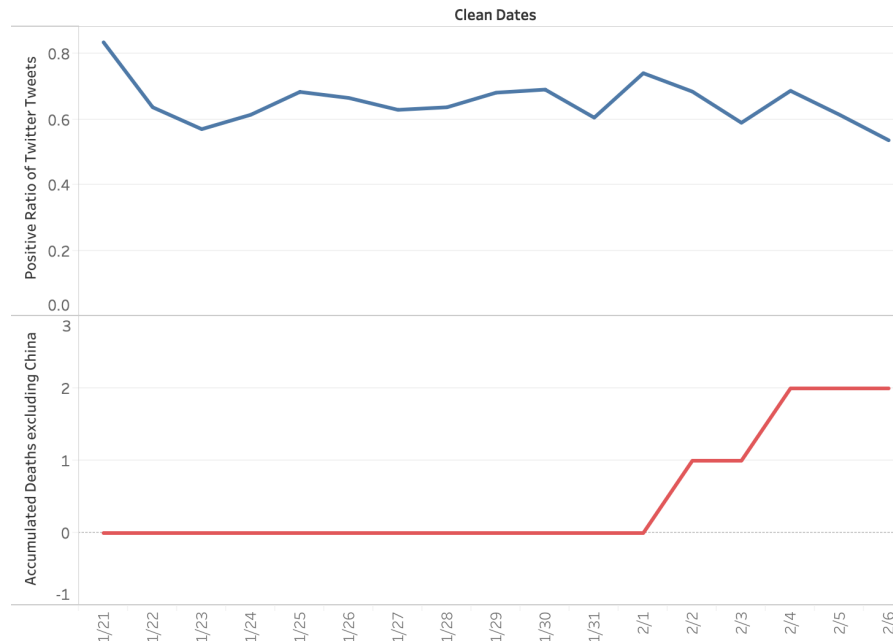


Correlation:
0.43046236

It may be that while the virus is spreading, more people are positively hoping for eradication.

Tracking Minds: Tweet Sentiment (Weibo and Twitter)

Other Countries Positive Ratio vs Deaths

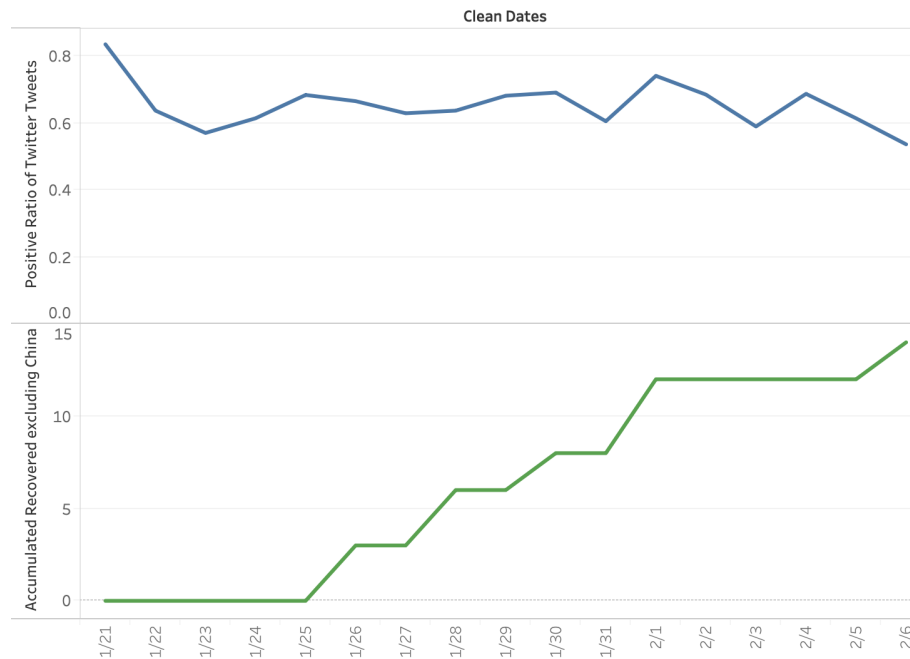


Twitter

Correlation:
-0.14499143

Was not able to find a strong correlation between Sentiment vs Deaths or Recovered.

Other Countries Positive Ratio vs Recovered



Correlation:
-0.10349502

But it seems to have a negative correlation which may mean that more people are becoming more worried.

The background of the slide features a series of thin, curved lines in light gray and white, creating a sense of motion and depth. These lines are more prominent on the left side and fade towards the right.

Tracking Minds: Tweet Sentiment (Weibo and Twitter)

From the Tweet sentiment results, we may be able to say that in countries where the virus is already spreading widely, people tend to think optimistically such as eradication and curing.

On the other hand, in countries where the virus hasn't spread too much yet, people may be thinking more negatively, such as fear of infection.



Difficulties during the DevFest

- 2019-nCoV data had incorrect values
- Limited # in Tweets that can be collected
 - + Noisy tweets (Retweets)
 - + Other languages
- Prediction of the spread in coronavirus is too difficult
- Amaury spilled water and broke his MacBook
- Amaury got ill

References

- 1. Barnaghi, Peiman, Parsa Ghaffari, and John G. Breslin. "Opinion mining and sentiment polarity on twitter and correlation between events and sentiment." 2016 IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService). IEEE, 2016. Limited # in Tweets that can be collected
- 2. <https://www.nytimes.com/2020/02/09/world/asia/coronavirus-china.html> Amaury lost his voice
- 3. <https://zhuanlan.zhihu.com/p/70153413>
- 4. <https://zhuanlan.zhihu.com/p/30061051>
- 5. <https://www.rollingstone.com/culture/culture-news/coronavirus-wuhan-china-virus-quarantine-death-toll-941417/>
- 6. <https://news.163.com/special/epidemic/>
- 7. <https://towardsdatascience.com/twitter-sentiment-analysis-classification-using-nltk-python-fa912578614c>
- 8. Ji, Xiang, Soon Ae Chun, and James Geller. "Monitoring public health concerns using twitter sentiment classifications." 2013 IEEE International Conference on Healthcare Informatics. IEEE, 2013.
- 9. Pollacci, Laura, et al. "Sentiment spreading: an epidemic model for lexicon-based sentiment analysis on twitter." Conference of the Italian Association for Artificial Intelligence. Springer, Cham, 2017.
- 10. <https://github.com/jonbakerfish/TweetScraper>