



pandemic pulse

Understanding Global Trends
During the Pandemic

COVID-19 changed the world in 2020 by affecting everyone's lives

82M

COVID 19 cases worldwide (as of 31 December 2020)¹

2.5%

Increase in unemployment rate from 4.8% globally ²

21%

Drop of S&P500 Index ³

2

¹ Johns Hopkins University, COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. <https://github.com/CSSEGISandData/COVID-19>

² International Monetary Fund (2020). Unemployment Rates, IMF - Updated: October 2020. https://www.imf.org/external/datamapper/LUR@WEO/OEMDC/ADVEC/WEO_WORLD

³ Mazur, M.; Dang, M.; Vega, M. (2020). COVID-19 and the march 2020 stock market crash. Evidence from S&P1500 https://ycharts.com/indicators/sp_500_monthly_return

⁴ Sundaram, A. (2020). Yelp data shows 60% of business closures due to the coronavirus pandemic are now permanent. <https://www.cnbc.com/2020/09/16/yelp-data-shows-60percent-of-business-closures-due-to-the-coronavirus-pandemic-are-now-permanent.html>

our motivation

3



What were the prominent pandemic news topics and how did they evolve in 2020?

- Aid companies identify new business opportunities by analysing news trends related to COVID-19
- Assist governments in identifying emerging public concerns (and indicators of social unrest) as everyone navigates the post-pandemic world.¹

4

¹ The International Monetary Fund has concluded that “pandemics can set off a vicious cycle of economic despair, inequality and social unrest”



**our method
of attack**

Our Data Science Methodology in a Big Data Framework



GDELT GKG data from Google Cloud

Global Knowledge Graph (GKG) data for 2020 was obtained from the Global Database of Events, Languages, and Tone (GDELT) hosted in Google BigQuery

Total Data Size: 182 GB
Total Entries: 45 Million



Identify COVID-19 Keywords

COVID-19 related **THEMES** were identified from the dataset and were used as a search query to the data

Total Data Size: 68 GB
Total Entries: 10 Million



Reconstruct articles using GKG columns

Using **THEMES**, **ALL NAMES** and **ORGANIZATIONS** in GKG, articles and events are reconstructed without the use the need for web scraping



Generate Topics using LDA

Topics per month were generated using a **Distributed Latent Dirichlet Allocation (LDA)** implementation in PySpark

Topics Found: 120 topics

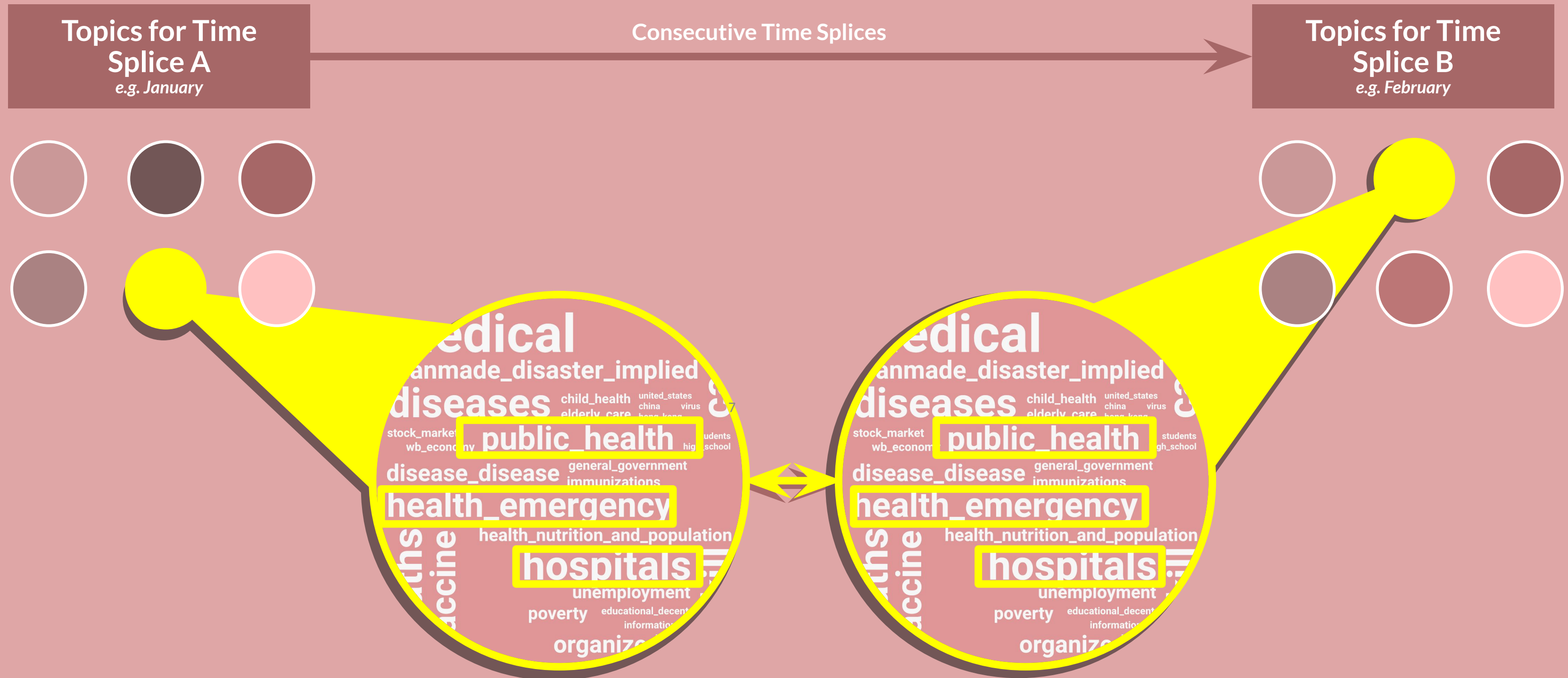


Trace topic evolution sequences

Trace topic evolution sequences using **Jaccard Similarity** for all topics in every consecutive time splices. Factored crowd sentiment using **TONE**

Total Topics Traced: 52
Topic Sequences

Tracking topic evolution requires similar topics to appear in consecutive time splices



**our
insights**

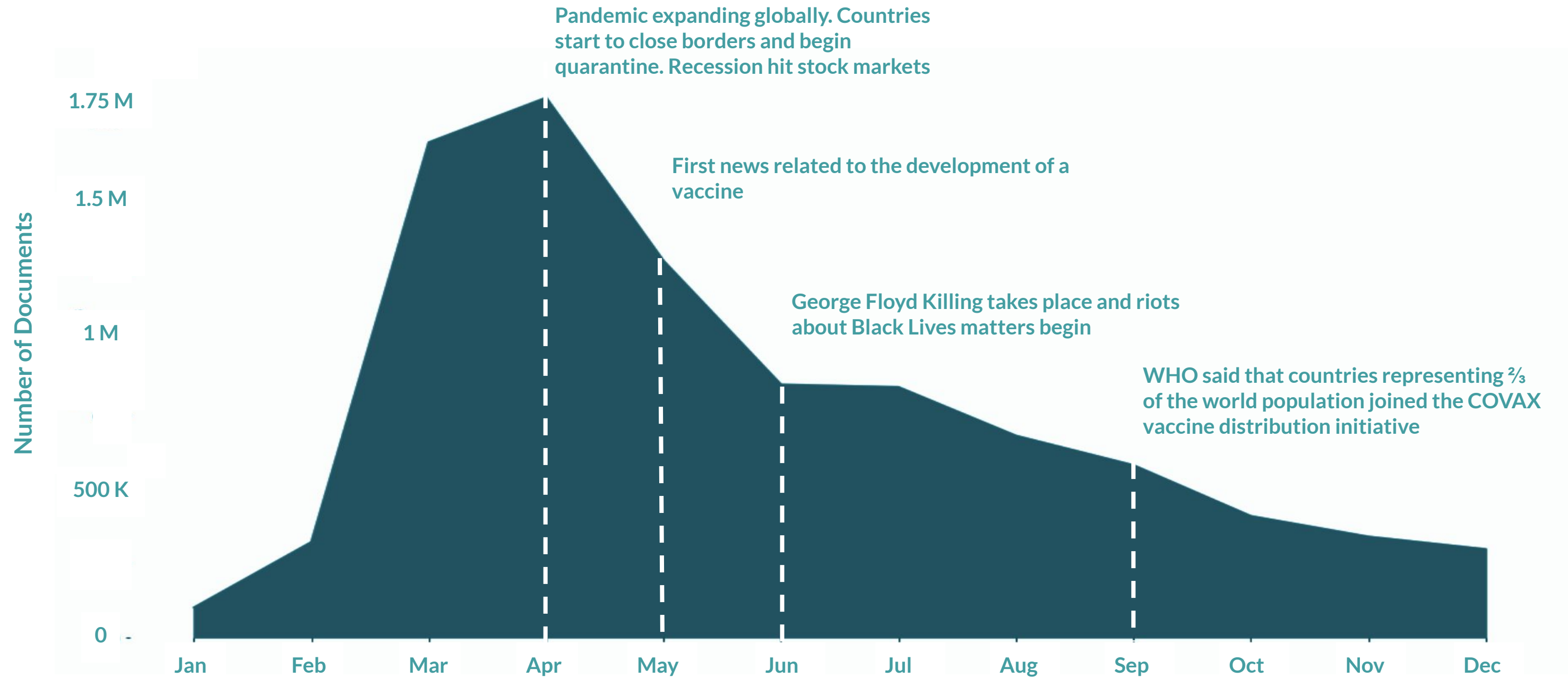
8



1

News related to
COVID-19 had peaked in
April 2020 and
continuously declined
since then

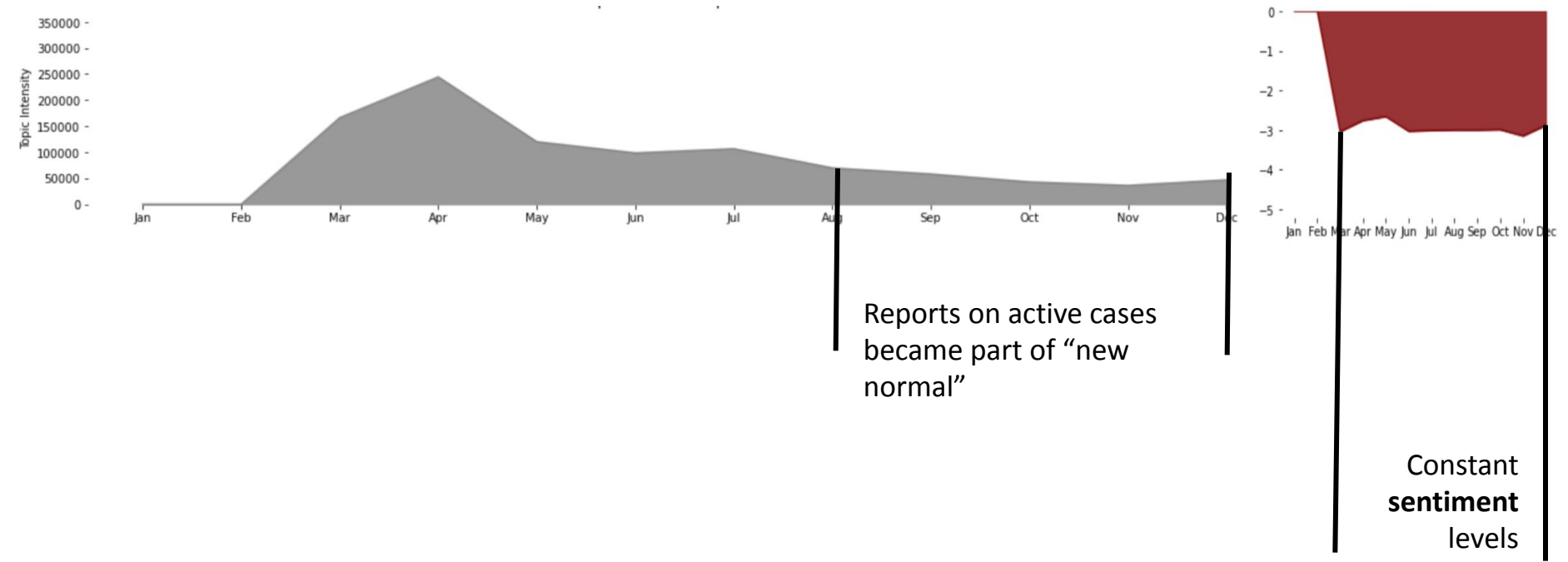
Evolution of news related to COVID-19 in 2020



2

Several topics have persisted throughout the year at varying intensities

Several topics have persisted throughout the year at varying intensities



14

COVID 19 ACTIVE CASES, RECOVERIES, AND FATALITIES

Global news on active cases, deaths, and recoveries for COVID-19 including news about healthcare workers, the healthcare industry, and the government handling of the pandemic

Topic Intensity (# of Documents)



Crowd Sentiment (Tone)



3

Numerous topics have
been emerging as the
year ends



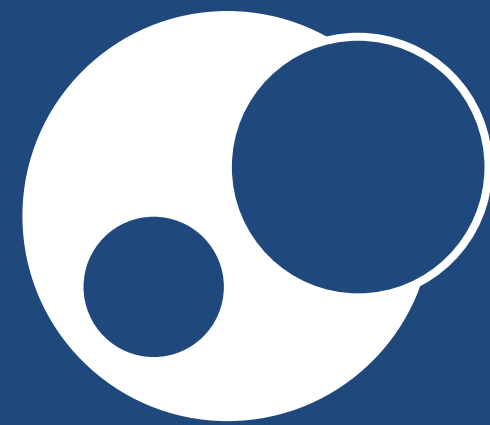
**our
concluding
thoughts**

Global news related to the pandemic has started to die down as economies slowly open and embrace the new normal; however, new trends emerge that may pose new challenges

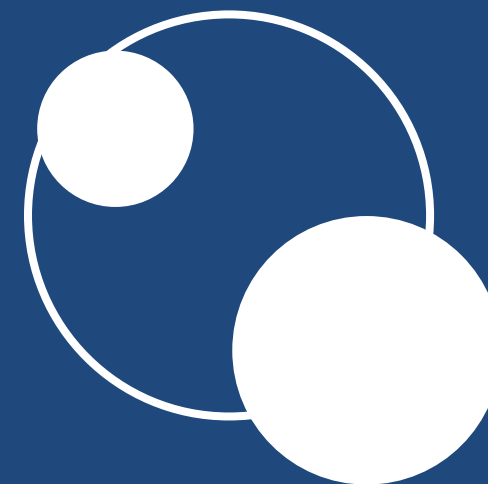
Business Conclusions



Public Health and COVID-19 related news are now deeply integrated into our daily lives



New discussions focus on how people will live in the post-pandemic world



Capture semantic relationship of words by looking at term co-occurrences



Conduct a more detailed analysis per country and determine how each one adapts to the new normal

Get to know more about us by connecting on LinkedIn!



**RAFAEL
MADRIGAL**

[/in/rgmadrigal/](#)



**TAL MIGUEL
NAVARRO**

[/in/tal-navarro/](#)



**DARYLL
TUMAMBING**

[/in/daryll-tumambing/](#)



**ALEJANDRO
WHITE**

[/in/alejandro-white/](#)



thanks!

You may visit us in our parallel session by going to bit.ly/pandemicpulse-aim

