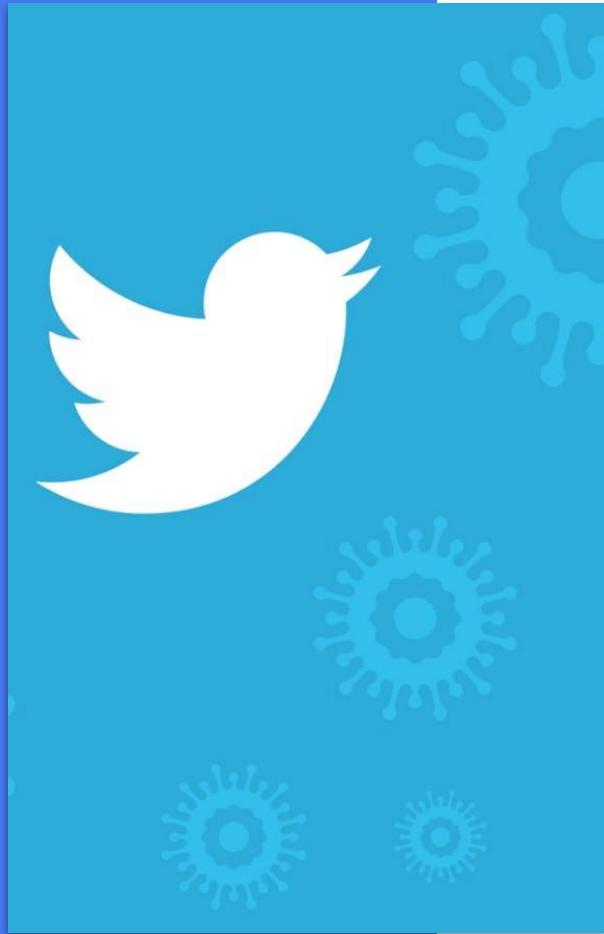


COVID-19 TWEETS

Simona Guida - Marco Pedrinazzi



**“We’re not just fighting an epidemic;
we’re fighting an infodemic.”**

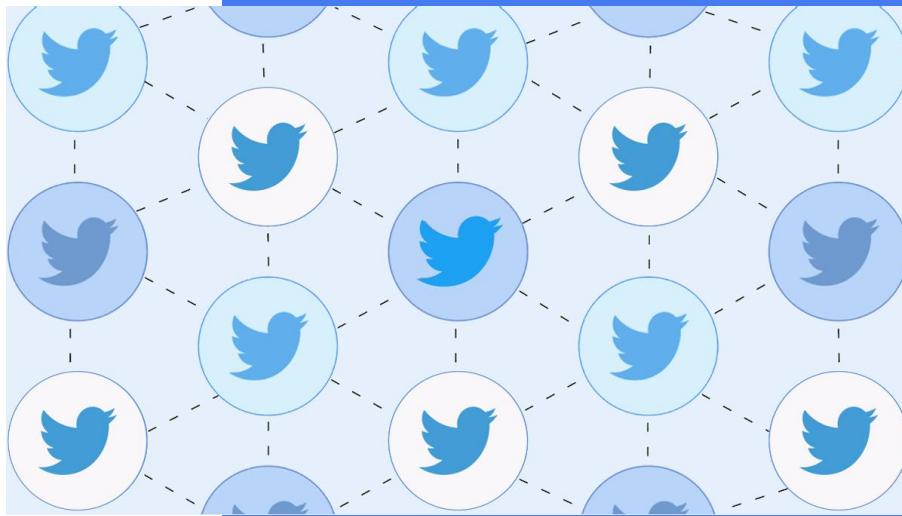
- **Tedros Adhanom Ghebreyesus**
(WHO General Director)

ABOUT THE PROJECT

Shahi G. K., Dirkson A. & Majchrzak T. A.
(2021).

**An exploratory study of COVID-19
misinformation on Twitter.**

Online social networks and media, 100104.



The project is licensed under a **Creative
Commons Attribution-ShareAlike 4.0
International License.**

01

ABOUT THE DATASETS

DATASET “GENERAL”

Background corpus of Covid-19 tweets

tweet_id

1217003561973760000

1217099725926797312

1217084298102214656

1216889242007138304

1217151361902620672

1216912067409526784

1217196291610087424

1217043503764713472

1217175838795239424

1217160810625150976

DATASET “FAKE”

Covid-19 misinformation tweets

Two tweet categories:

- **False**
- **Partially False**

tweet_id	tweet_class
1286643297402920960	FALSE
1262820117680918529	partially false
1263195773451239426	FALSE
1270765111020531718	FALSE
1294027549459218436	partially false
1276294100862545921	FALSE
1293194047037874178	partially false
1270487181652635652	FALSE
1271963436969349120	FALSE
1277787130362761217	partially false

DATASET HYDRATION

We've done the **hydratation** (to get complete details of a tweet from its ID) with [twarc](#).

The picture shows an **anonymized** example of the result of the hydration.

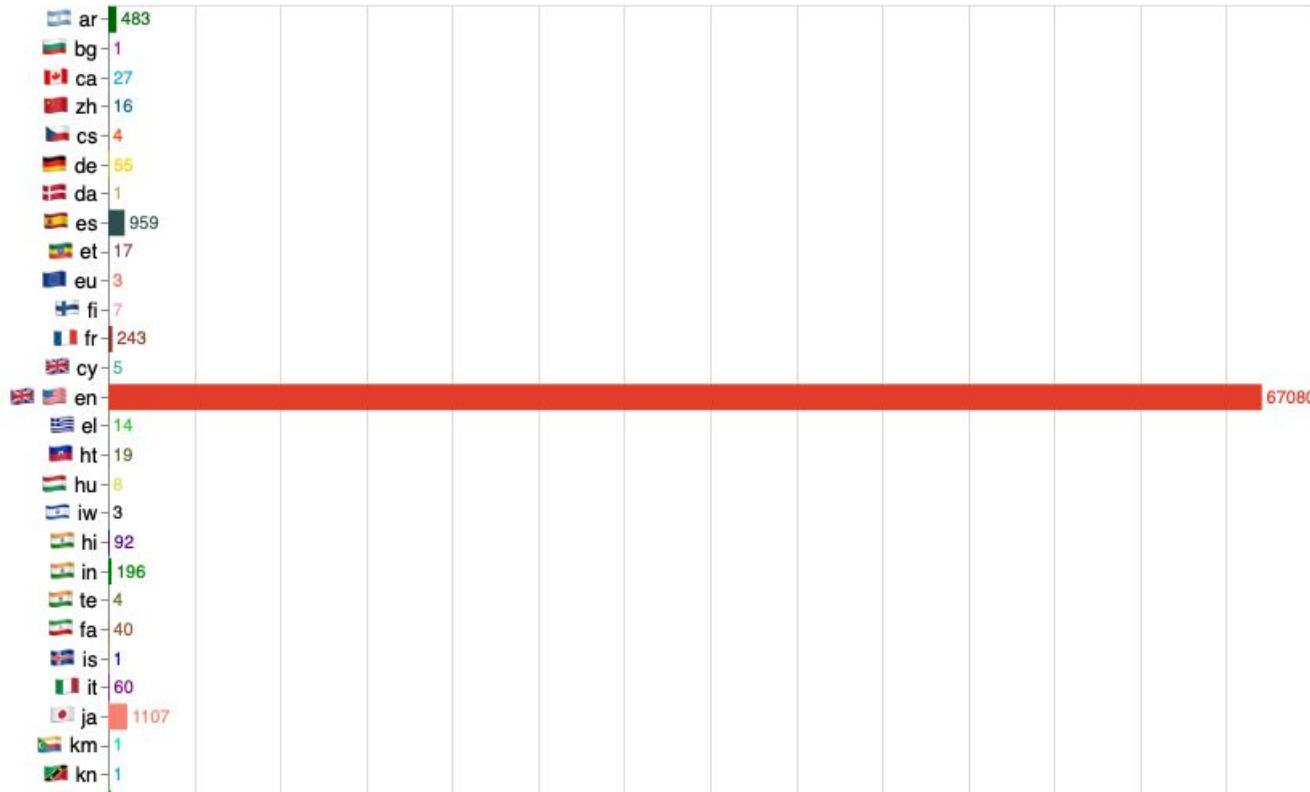
```
{  
  "created_at": "DATE",  
  "id": 1,  
  "id_str": "1",  
  "full_text": "text",  
  "truncated": false,  
  "display_text_range": [],  
  "entities": {},  
  "extended_entities": {},  
  "source": "source",  
  "in_reply_to_status_id": 2,  
  "in_reply_to_status_id_str": "2",  
  "in_reply_to_user_id": 3,  
  "in_reply_to_user_id_str": "3",  
  "in_reply_to_screen_name": "username",  
  "user": {},  
  "geo": null,  
  "coordinates": null,  
  "place": null,  
  "contributors": null,  
  "is_quote_status": false,  
  "retweet_count": 277,  
  "favorite_count": 411,  
  "favorited": false,  
  "retweeted": false,  
  "lang": "language"  
}
```

02

ABOUT THE CHARTS

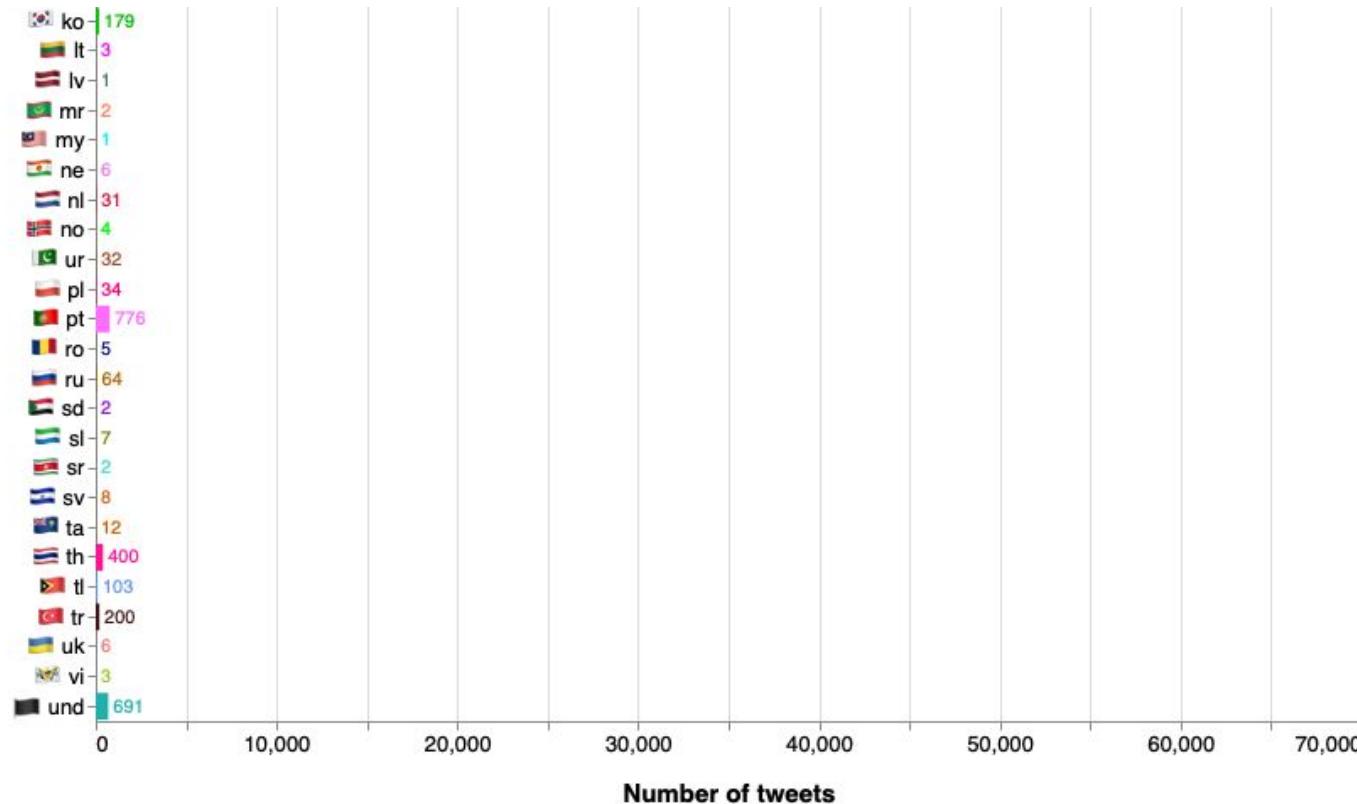
PRE PROCESSING DATASET “GENERAL”

Before the translation to **english**.



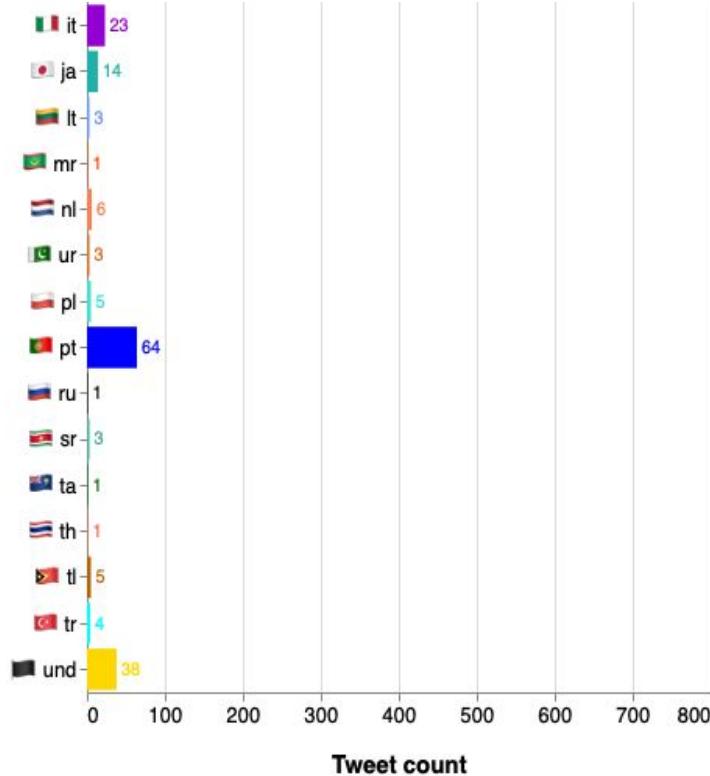
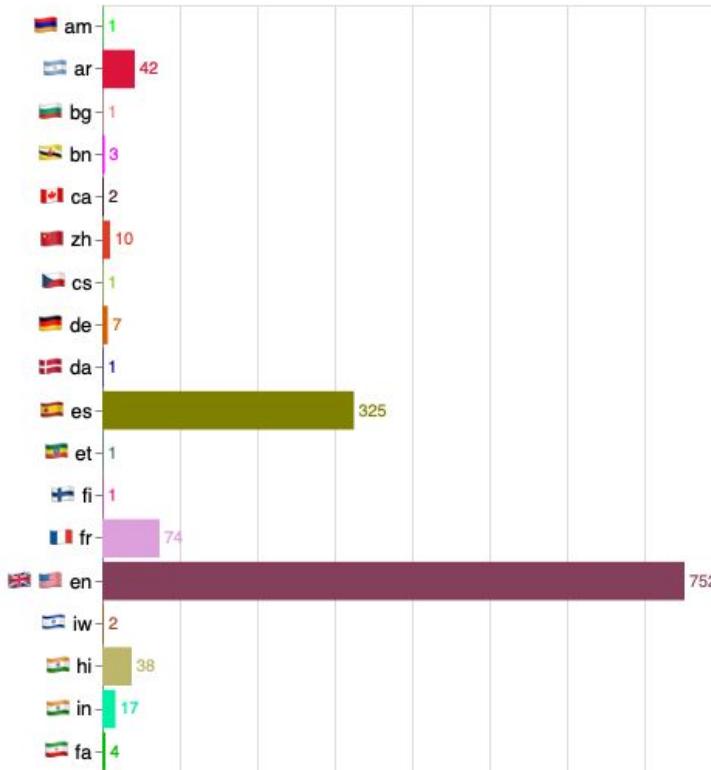
PRE PROCESSING DATASET “GENERAL”

Before the translation to **english**.



PRE PROCESSING DATASET “FAKE”

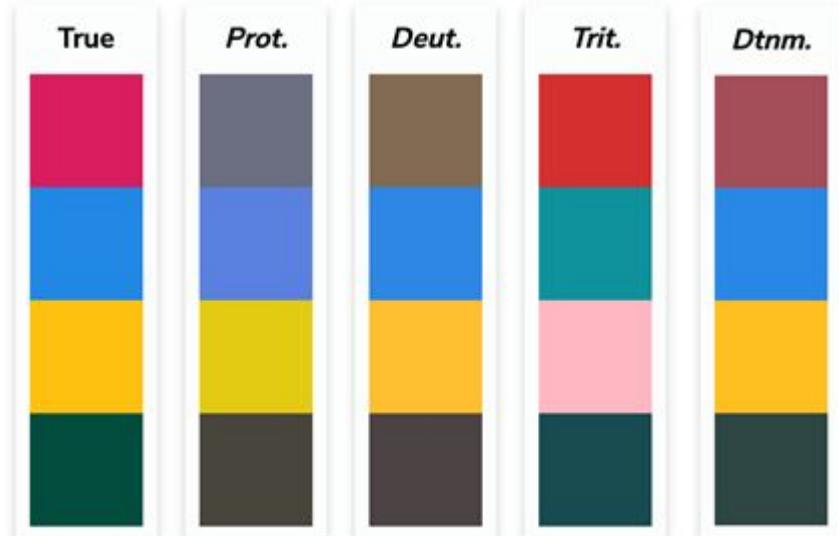
Before the translation to **english**.



COLORBLIND-FRIENDLY VISUALISATIONS

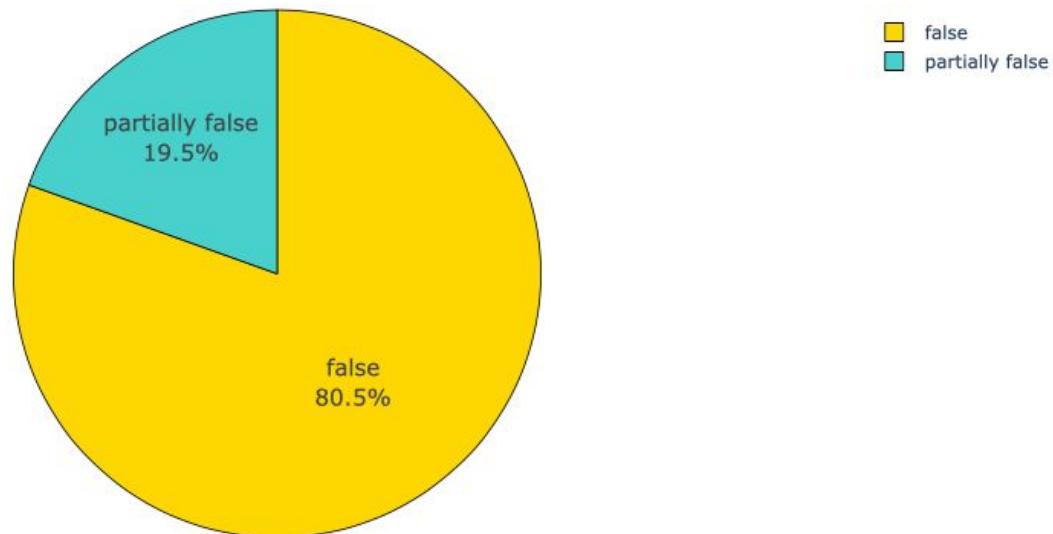
About **1 in 20 people are colorblind** in some way.

We've created **colorblind-friendly visualisations** by testing each color palette we've used in the [online color blindness simulation tool](#) created by David Nichols (UConn Math PhD Alumnus).



A DEEPER LOOK INTO THE DATASET “FAKE”

Tweets in Fake Covid-19 dataset classified by their category (false, partially false)



TWEETS VISUALISATION - DATASET “GENERAL”

Date	Tweet	Link
filter data...		
2020/01/14	@kpophappenings_ kim jaehyun from n.flying	http://twitter.com/anyuser/status/1216879817624211456
2020/01/14	LRT - For people who draw on an iPad and don't want to buy an expensive stand, buying doorstoppers works! If you buy 2 and put them under each side, then even if you use a lot of pressure when drawing (like me) it won't shift and its super cheap so I recommend it! You might just	http://twitter.com/anyuser/status/1216889242007138304
2020/01/14	Pretty equal booing to cheering ratio here at The Varsity as Trump makes an appearance	http://twitter.com/anyuser/status/1216889981446381568
2020/01/14	@esnewlife1 @PPersnickety Animals love unconditionally❤️ They do ❤️they can be obnoxious at times ❤️	http://twitter.com/anyuser/status/1216895118244044800
2020/01/14	Got savings goals? #BetterBanking can help you make it happen. Check out the Best Banks of 2020, plus enter to win \$500 while you're there! via @GOBankingRates	http://twitter.com/anyuser/status/1216896401688485888
2020/01/14	#GrandSolarMinimum Extreme Weather [Afghanistan]: Extreme cold and heavy snow hit Afghanistan, killing at least 17 people. Temperatures in parts of the country dropped to -12 °C, and heavy snowfall caused casualties.	http://twitter.com/anyuser/status/1216900045145927680
2020/01/14	Making something happen today! 💕 #MakeItHappenWithYou	http://twitter.com/anyuser/status/1216911371985514496
2020/01/14	The same thing that is happening in Portland OR. It's Very Sad.	http://twitter.com/anyuser/status/1216912067409526784

TWEETS VISUALISATION - DATASET “FAKE”

Type	Date	Tweet	Link
filter data...			
fake	2020/05/26	○ The Minister of Health of #burundi returned after hospitalization at #Kenya. He arrived at #Bujumbura around 17h. As at its departure, it has not been recorded as a passenger of a flight. His driver found him on the tarmac from the airport #Burundi 	http://twitter.com/anyuser/status/1265344818957484039
partiallyfalse	2020/05/27	the Go to Travel Campaign under consideration by the Japanese government is to stimulate domestic travel demand within Japan after the Covid-19 pandemic and only cover a portion of domestic travel expenses. (2 / 2)	http://twitter.com/anyuser/status/1265476469909676034
partiallyfalse	2020/05/27	#Breaking Watch: Patient films ‘horror’ in a Maharashtra hospital. Dead bodies, without even being wrapped up are found lying in rooms where patients are accommodated. Details by TIMES NOW’s Arunel. #RahulDumpsMaharashtra	http://twitter.com/anyuser/status/1265507998761463811
partiallyfalse	2020/05/27	See the background of the video shown by @ ramkadam! If the public representatives ask the administration of the administration before being accused, then people will not be scattered among the people. Also, in such efforts, the moral strength of the nurses and other employees, the moral of nurses and other staff were weak. But who will understand this?	http://twitter.com/anyuser/status/1265621485516963840

LINKS SHARED IN THE TWEETS - DATASET “FAKE”

Type	Link	First-Shared
filter data...		
fake	Bruxelles confirms: “The 5G technology is dangerous for the human health and for the environment”	11/06/2020
fake	An Effective Treatment for Coronavirus (COVID-19)	12/02/2020
fake	Trump's national security adviser says China 'covered up' coronavirus TheHill	12/03/2020
fake	A batch of medical supplies have arrived in Belgium.	12/03/2020

DOMAINS SHARED IN THE TWEETS

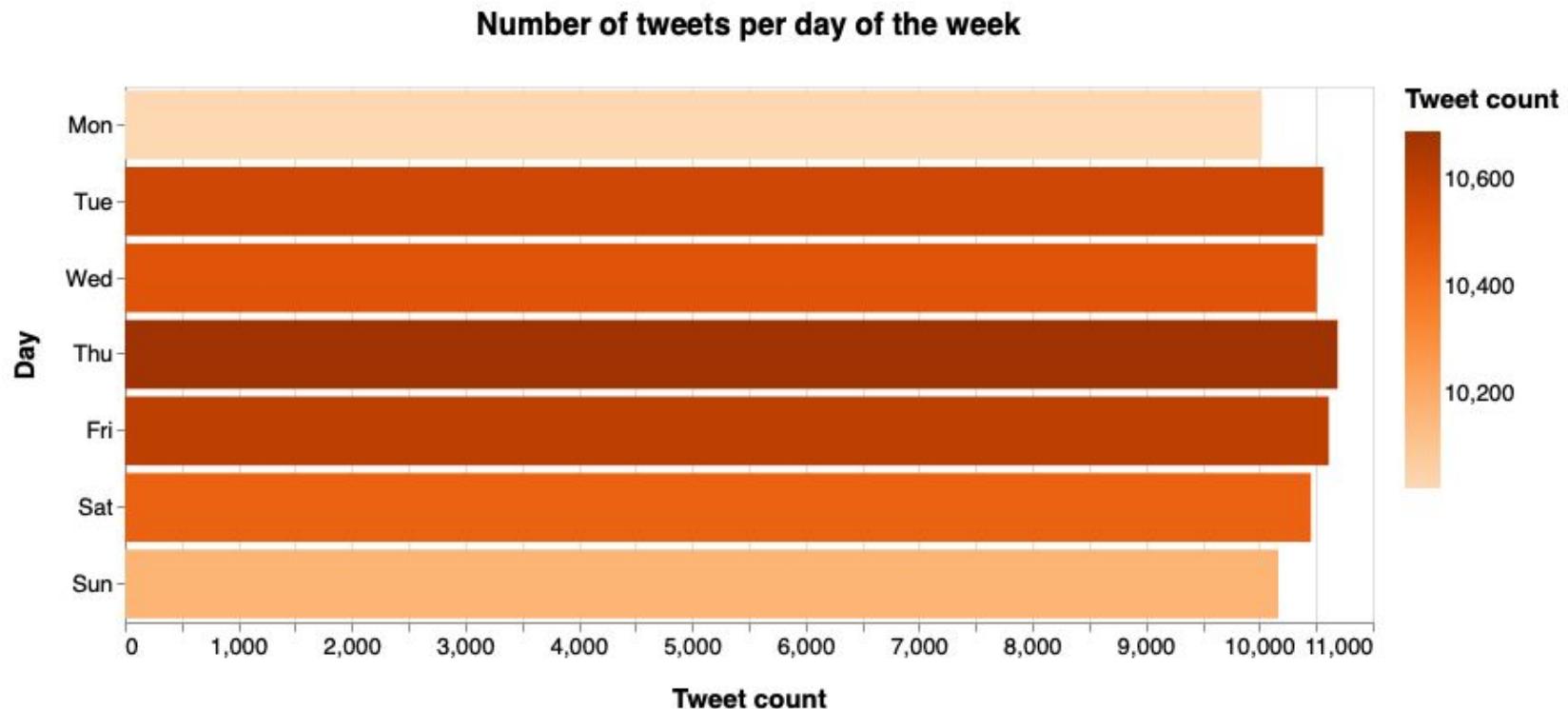
DATASET “GENERAL”

Domain	Count
filter data...	
bit.ly	1079
youtu.be	476
dlvr.it	322
www.instagram.com	303
www.theguardian.com	255
trib.al	253
ow.ly	234
www.nytimes.com	167

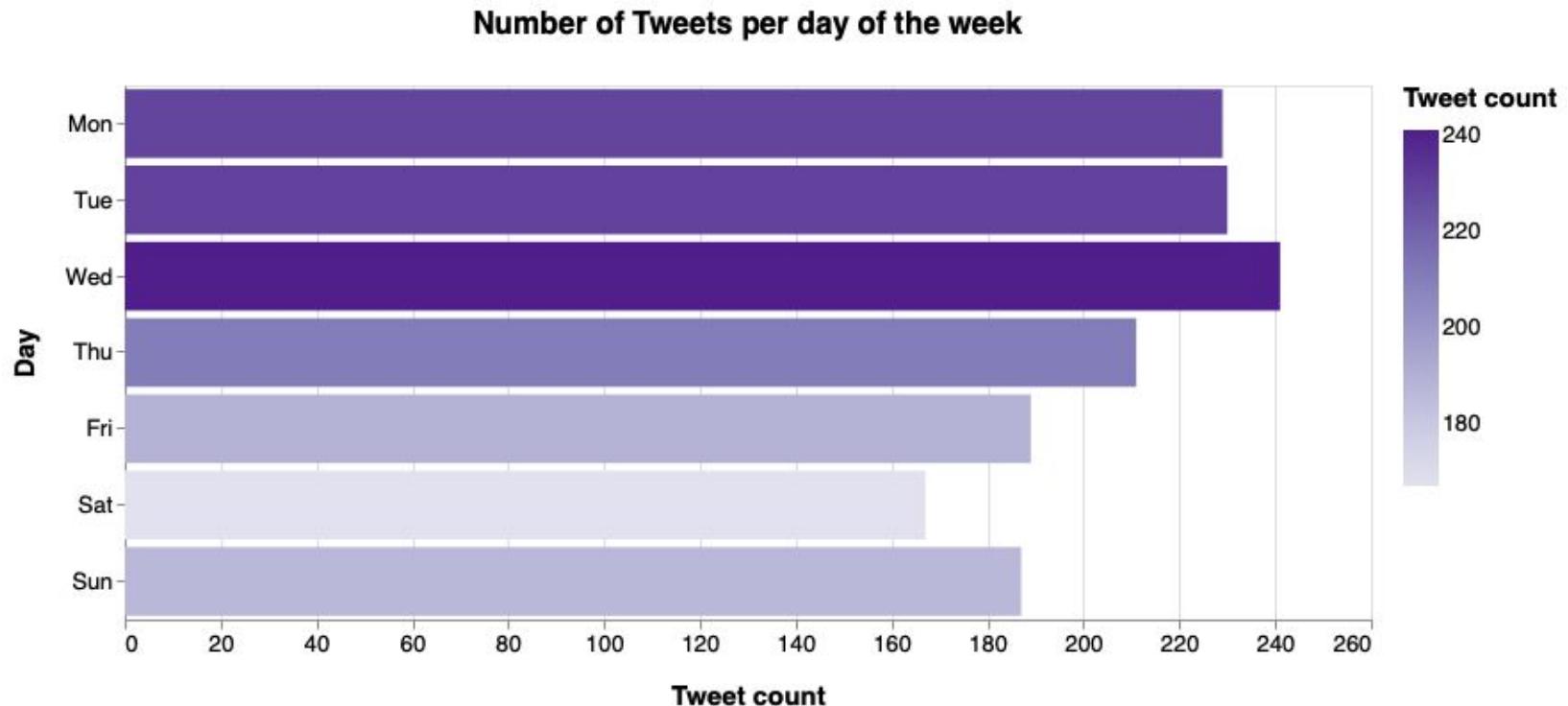
DATASET “FAKE”

Domain	Count
filter data...	
bit.ly	43
youtu.be	10
www.pscp.tv	9
www.who.int	5
www.youtube.com	5
buff.ly	4
www.cdc.gov	4
m.youtube.com	3

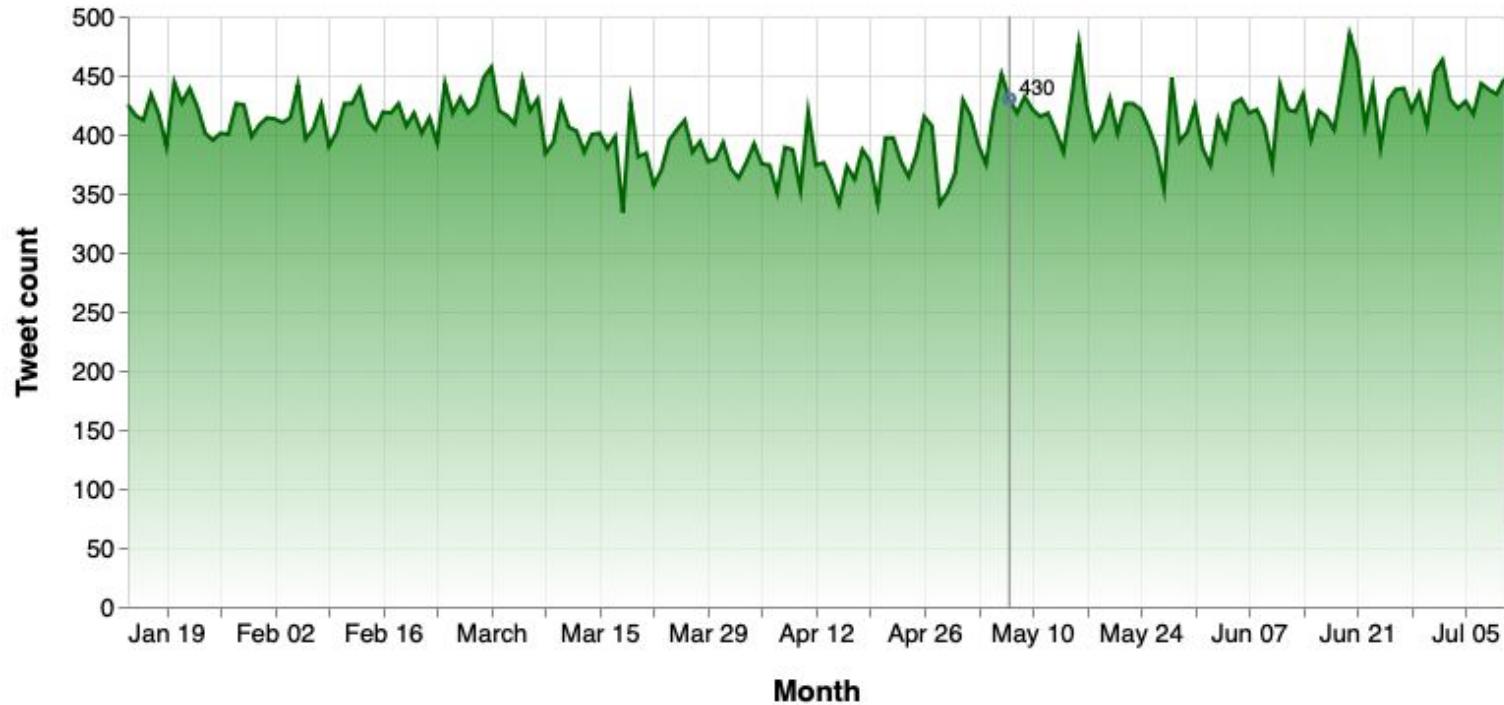
WHICH IS THE FAVORITE DAY OF THE WEEK TO TWEET ABOUT COVID-19? - DATASET “GENERAL”



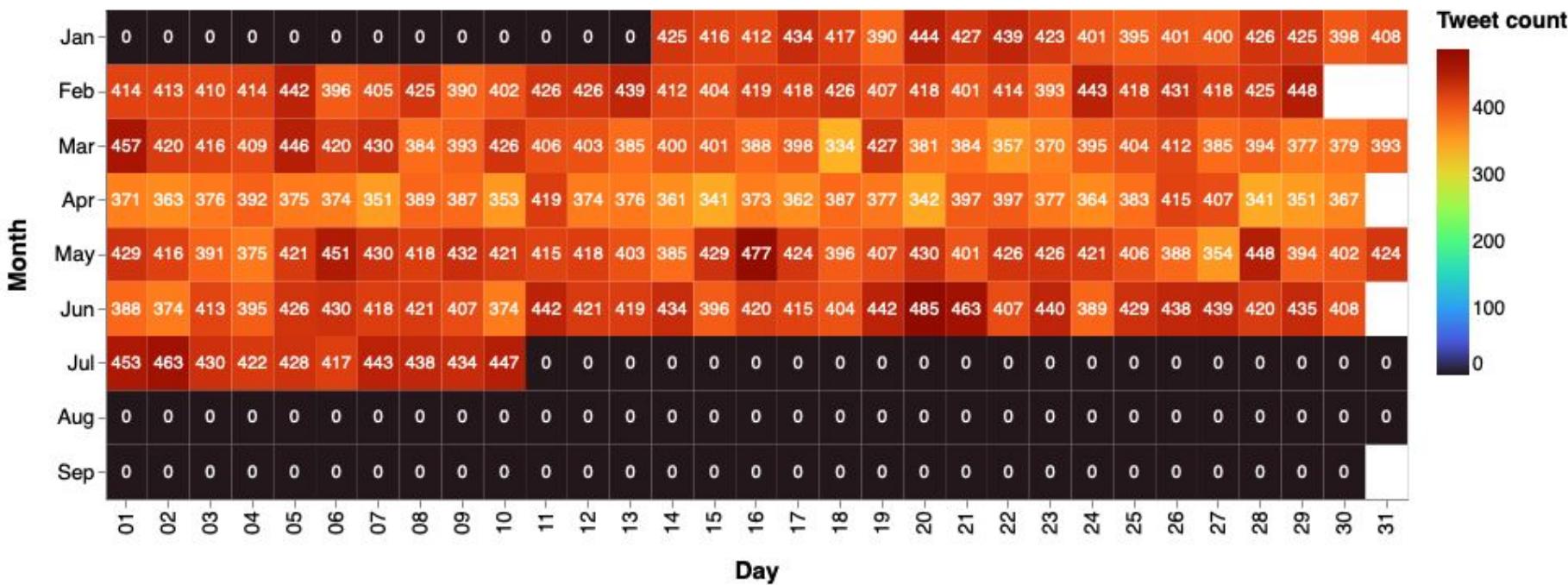
WHICH IS THE FAVORITE DAY OF THE WEEK TO TWEET ABOUT COVID-19? - DATASET “FAKE”



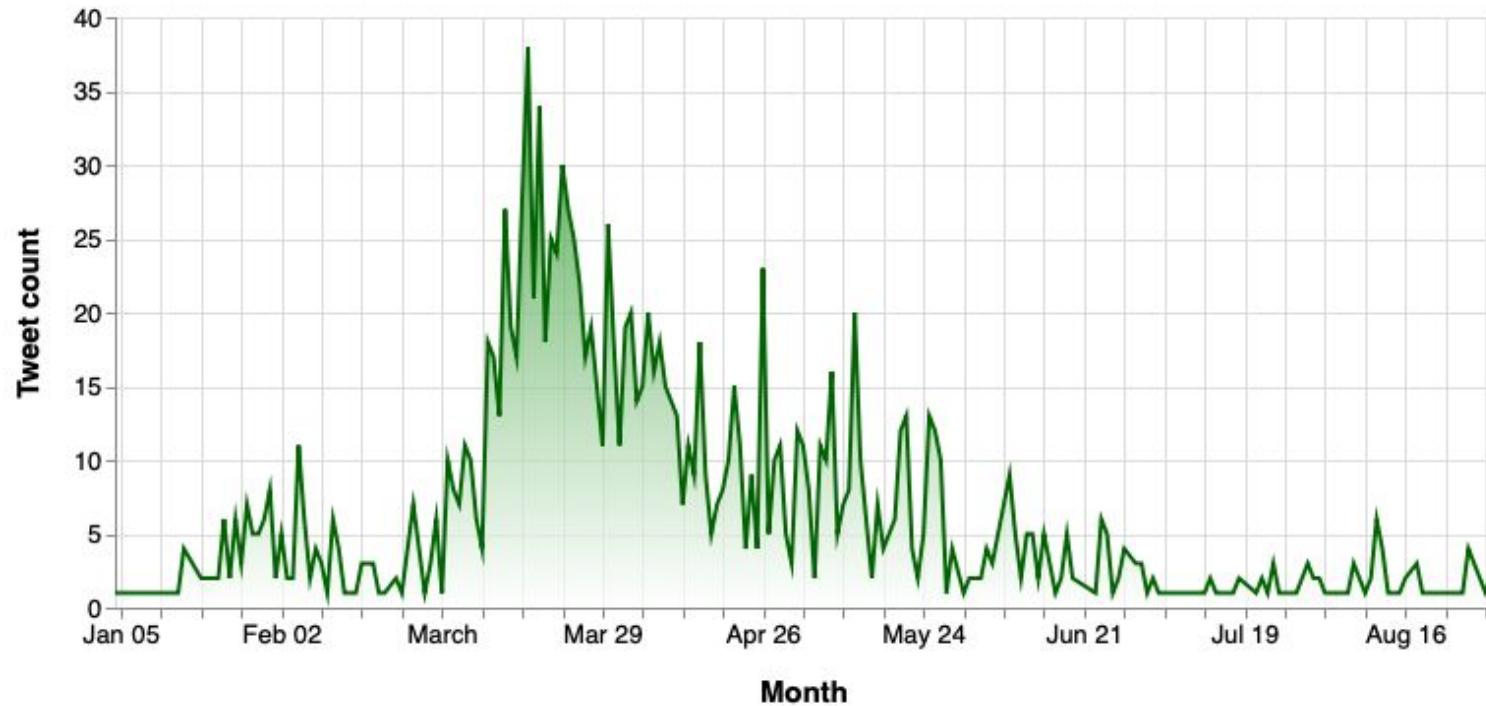
HOW MANY TWEETS HAVE BEEN CREATED DURING THE YEAR? - DATASET “GENERAL”



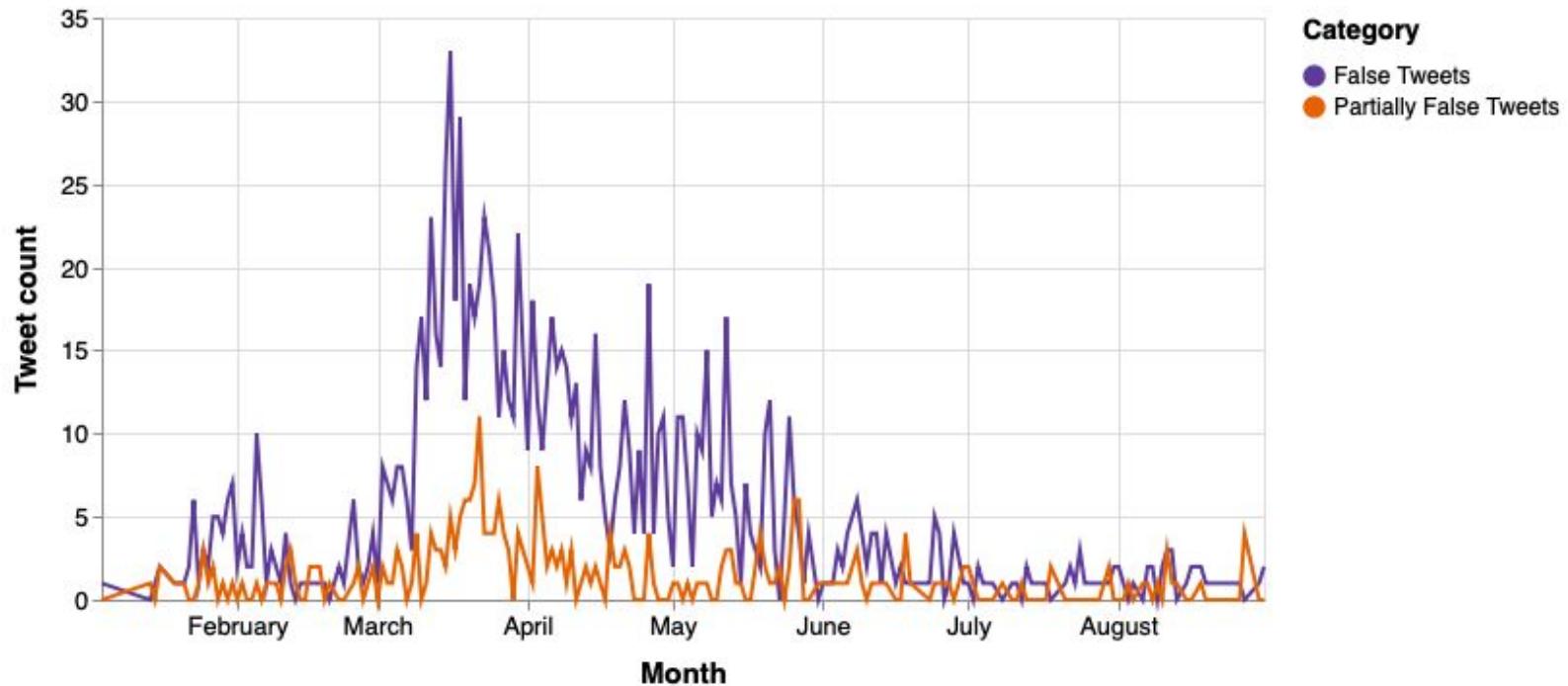
HOW MANY TWEETS HAVE BEEN CREATED DURING EACH DAY OF THE YEAR? - DATASET “GENERAL”



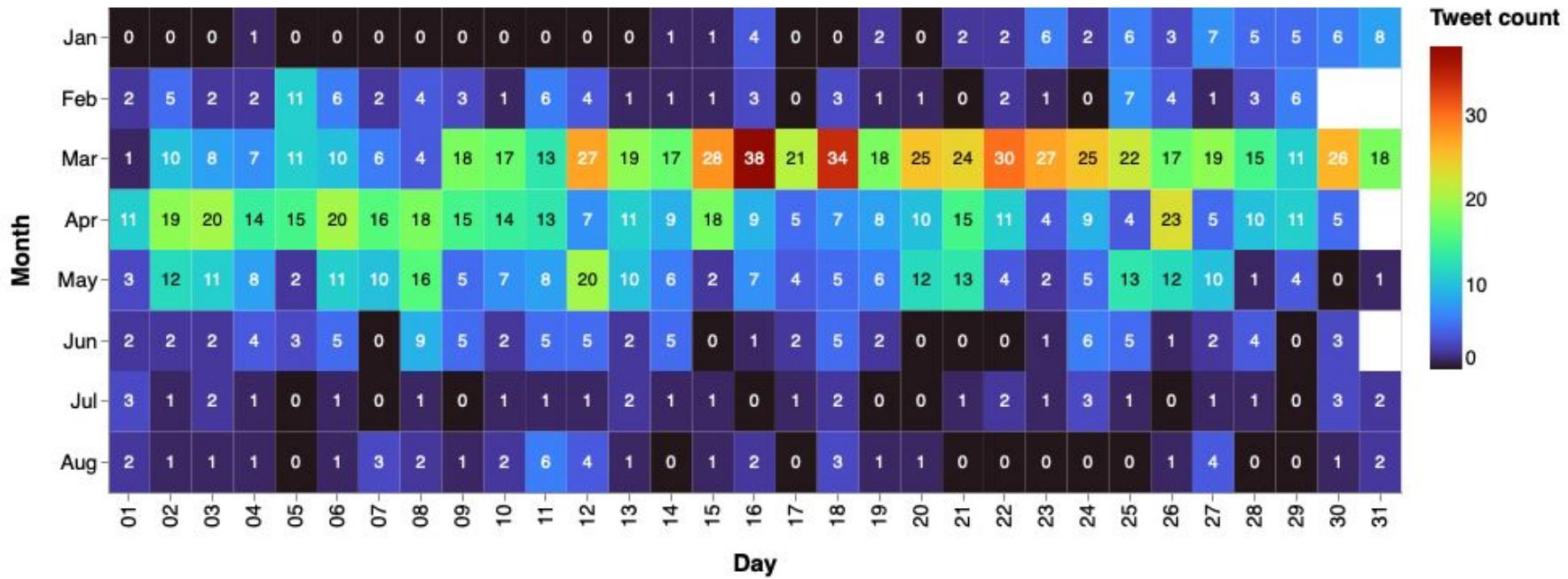
HOW MANY TWEETS HAVE BEEN CREATED DURING THE YEAR? - DATASET “FAKE”



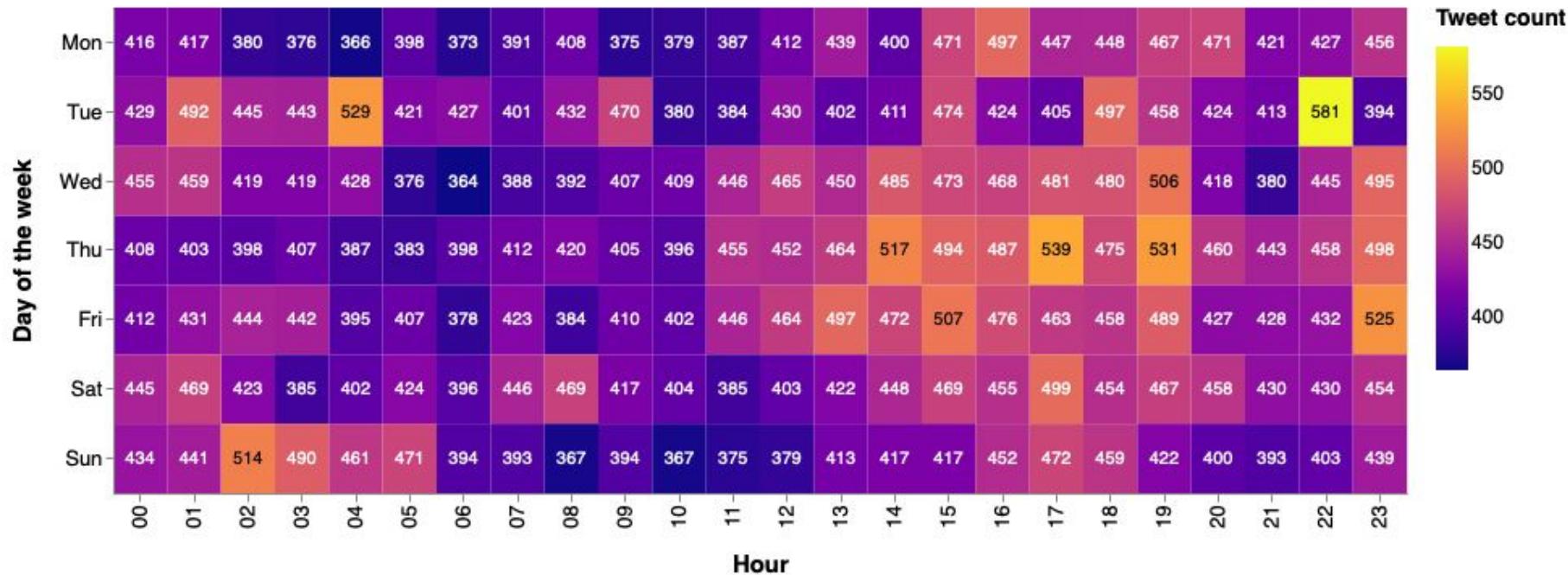
HOW MANY OF THOSE TWEETS WERE FALSE AND PARTIALLY FALSE? - DATASET “FAKE”



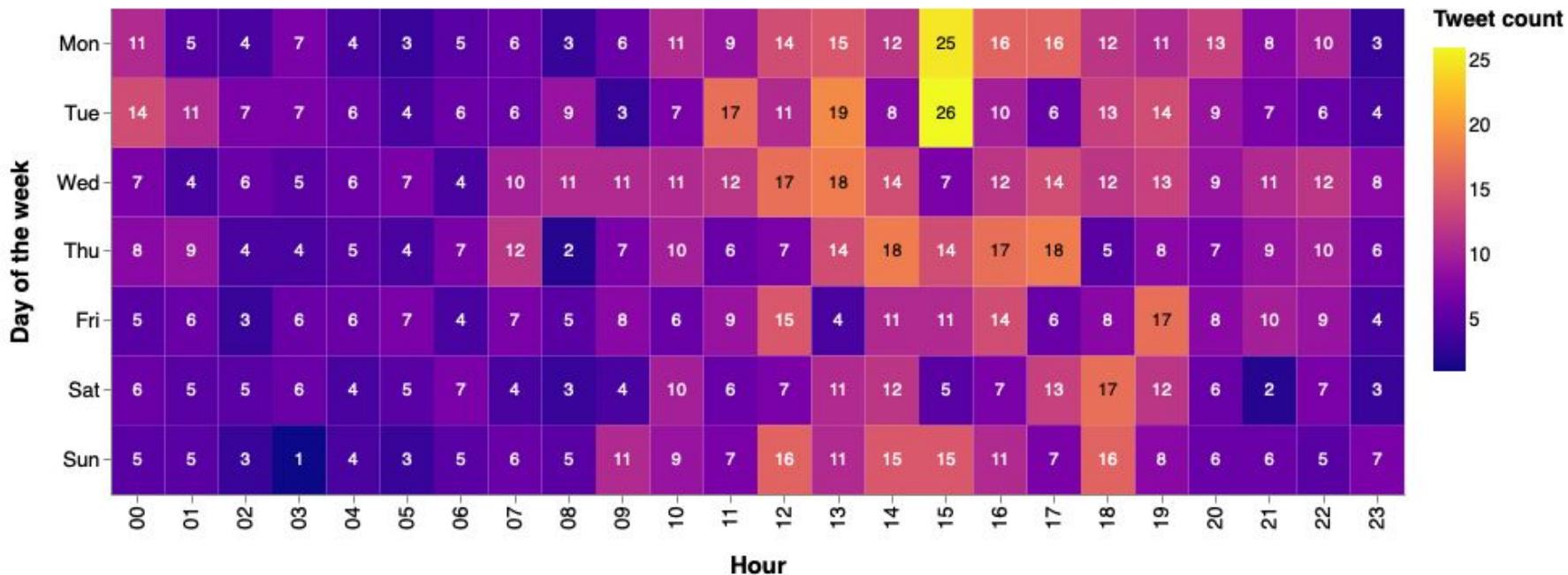
HOW MANY TWEETS HAVE BEEN CREATED DURING EACH DAY OF THE YEAR? - DATASET “FAKE”



WHAT TIME HAVE THE TWEETS BEEN POSTED DURING THE DAY? - DATASET “GENERAL”



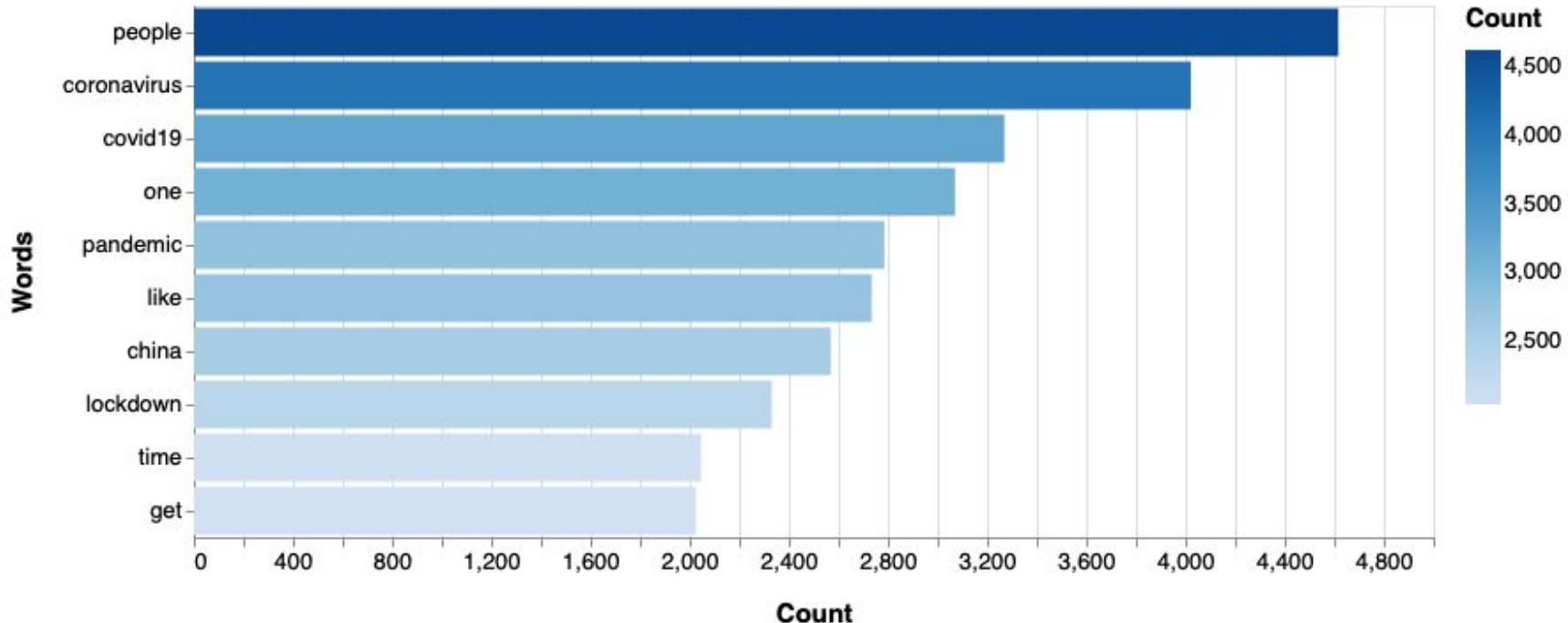
WHAT TIME HAVE THE TWEETS BEEN POSTED DURING THE DAY? - DATASET “FAKE”



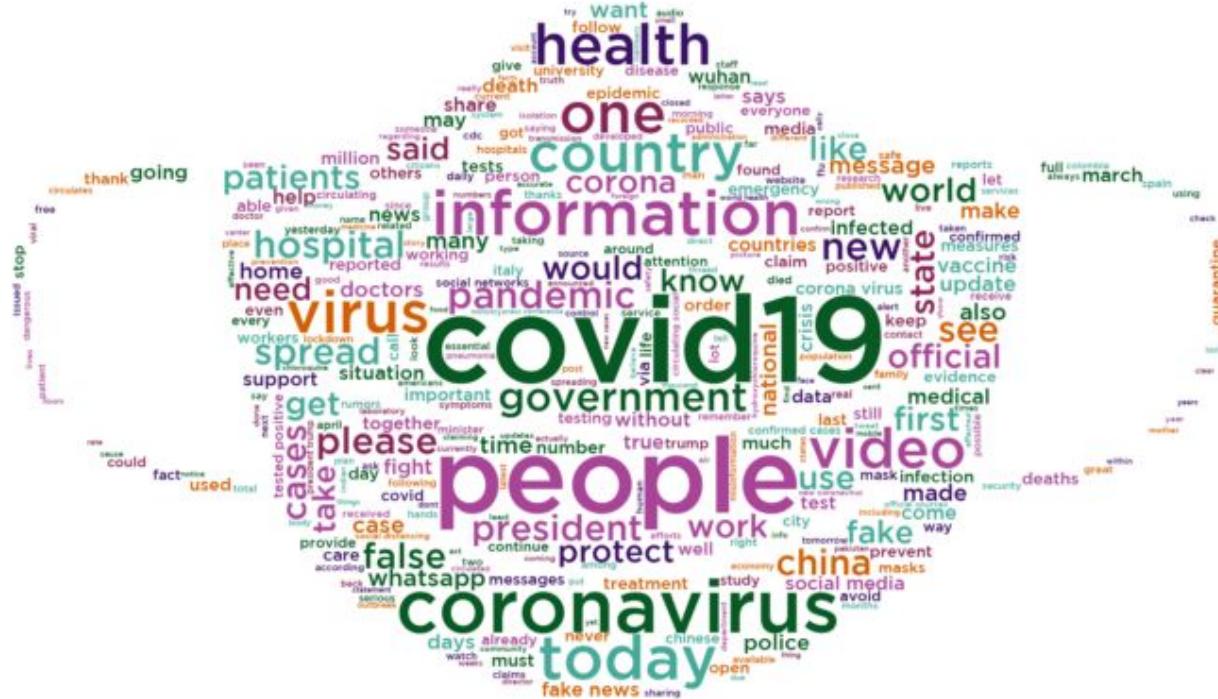
WHICH WORDS HAVE BEEN MORE USED IN THE TWEETS? (COVID-19 RELATED) - DATASET “GENERAL”



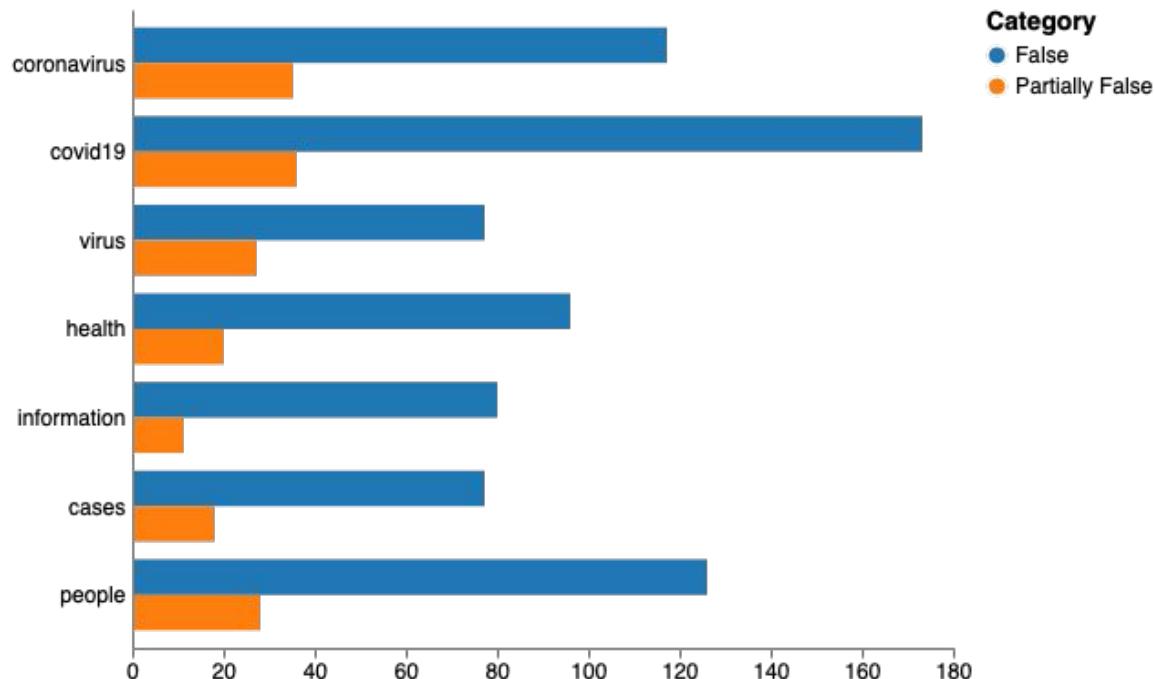
WHICH WORDS HAVE BEEN MORE USED IN THE TWEETS? (COVID-19 RELATED) - DATASET “GENERAL”



WHICH WORDS HAVE BEEN MORE USED IN THE TWEETS? (COVID-19 RELATED) - DATASET “FAKE”



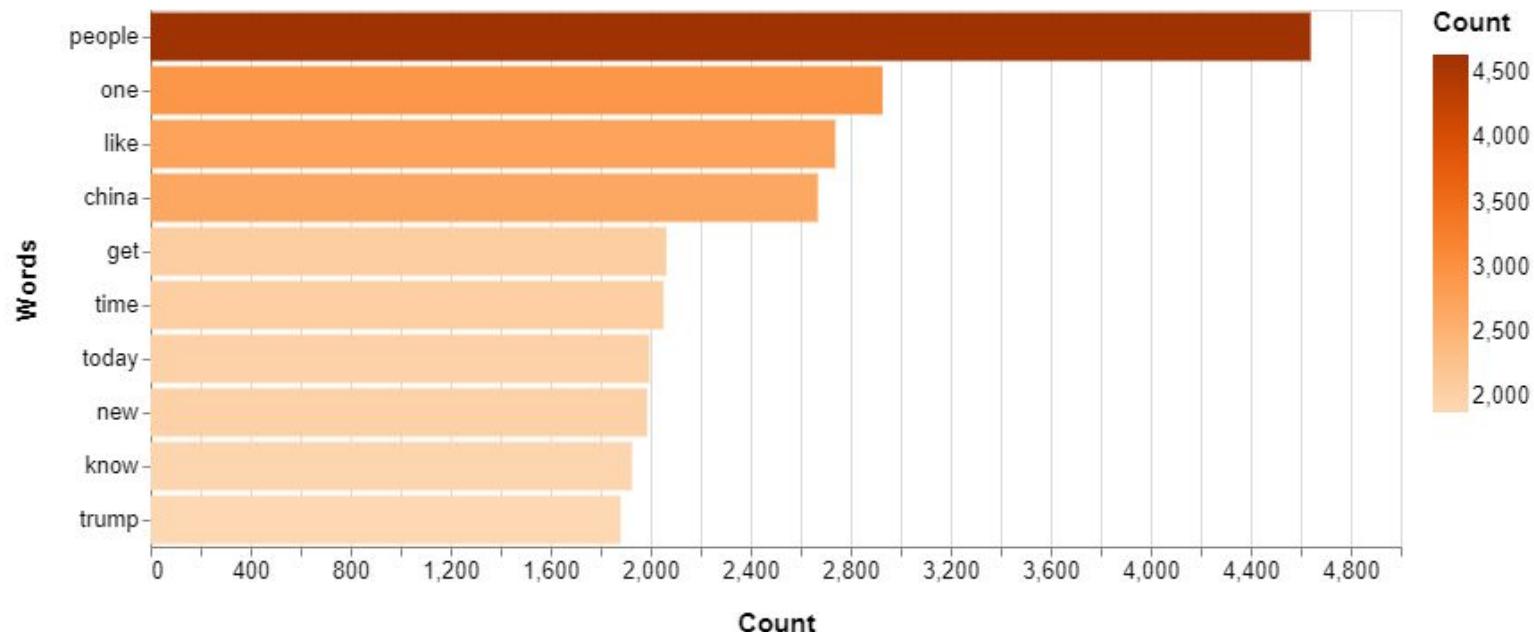
WHICH WORDS HAVE BEEN MORE USED IN THE TWEETS? (COVID-19 RELATED) - DATASET “FAKE”



WHICH WORDS HAVE BEEN MORE USED IN THE TWEETS? (NOT COVID-19 RELATED) - DATASET “GENERAL”



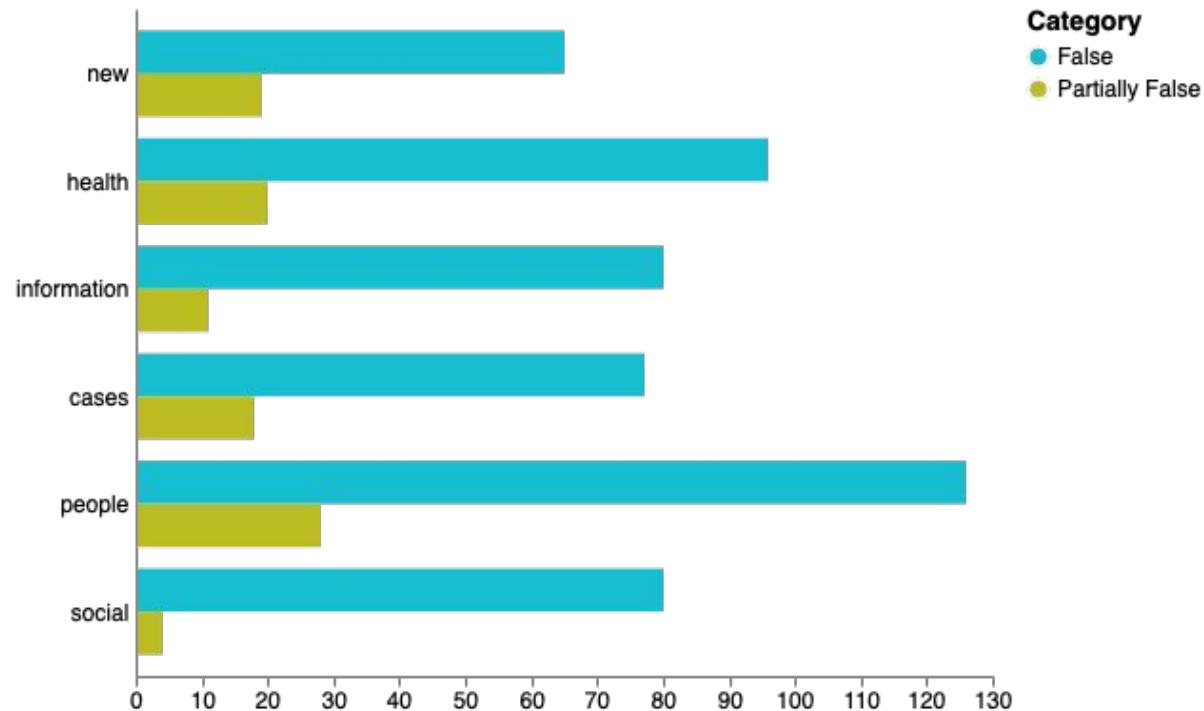
WHICH WORDS HAVE BEEN MORE USED IN THE TWEETS? (NOT COVID-19 RELATED) - DATASET “GENERAL”



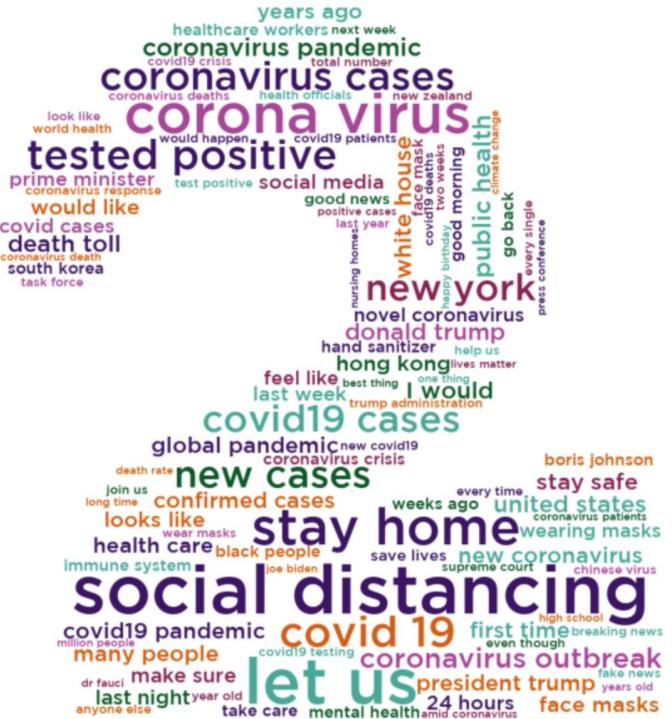
WHICH WORDS HAVE BEEN MORE USED IN THE TWEETS? (NOT COVID-19 RELATED) - DATASET “FAKE”



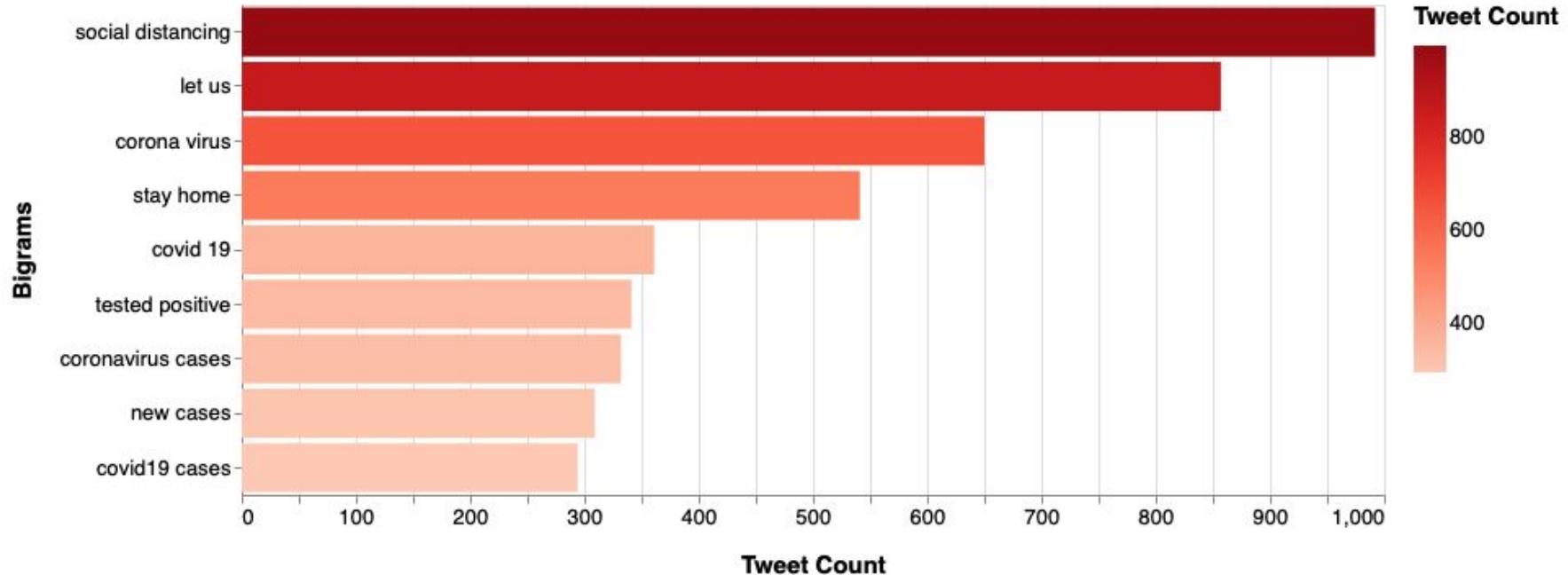
WHICH WORDS HAVE BEEN MORE USED IN THE TWEETS? (NOT COVID-19 RELATED) - DATASET “FAKE”



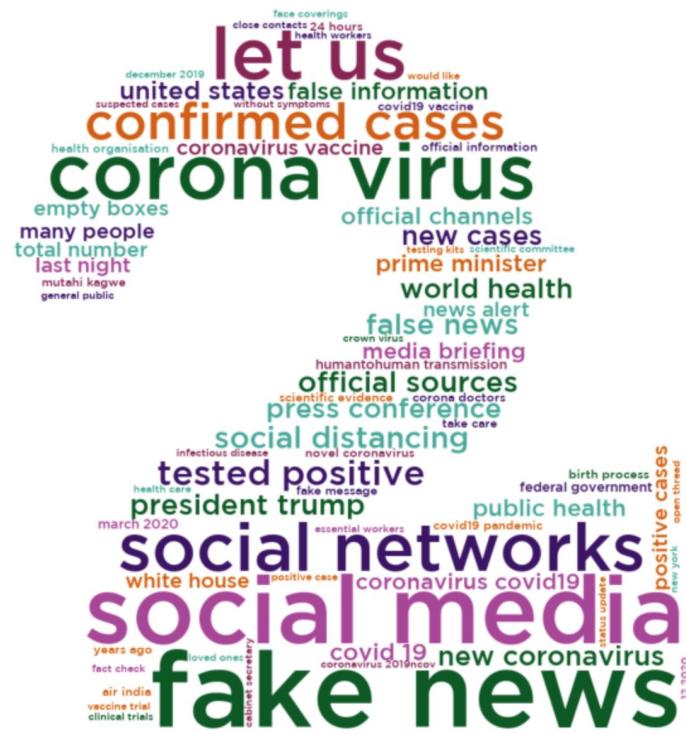
WHICH BIGRAMS HAVE BEEN MORE USED IN THE TWEETS? - DATASET “GENERAL”



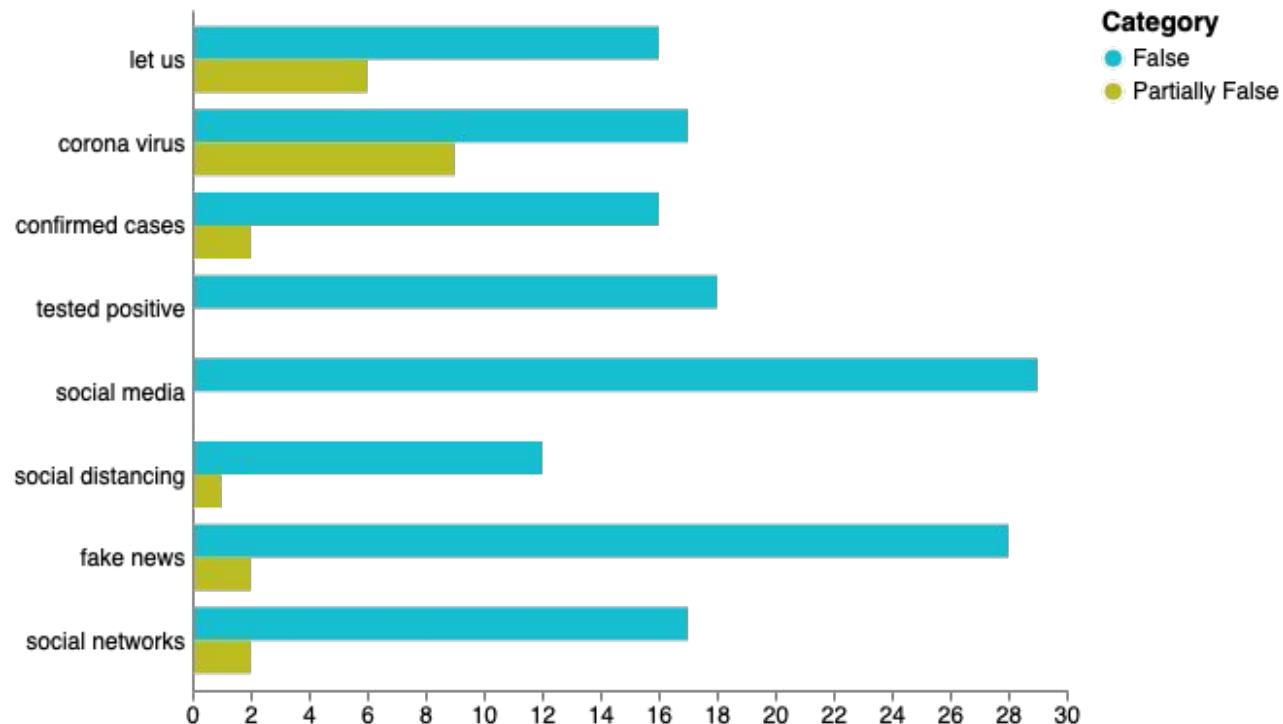
WHICH BIGRAMS HAVE BEEN MORE USED IN THE TWEETS? - DATASET “GENERAL”



WHICH BIGRAMS HAVE BEEN MORE USED IN THE TWEETS? - DATASET “FAKE”



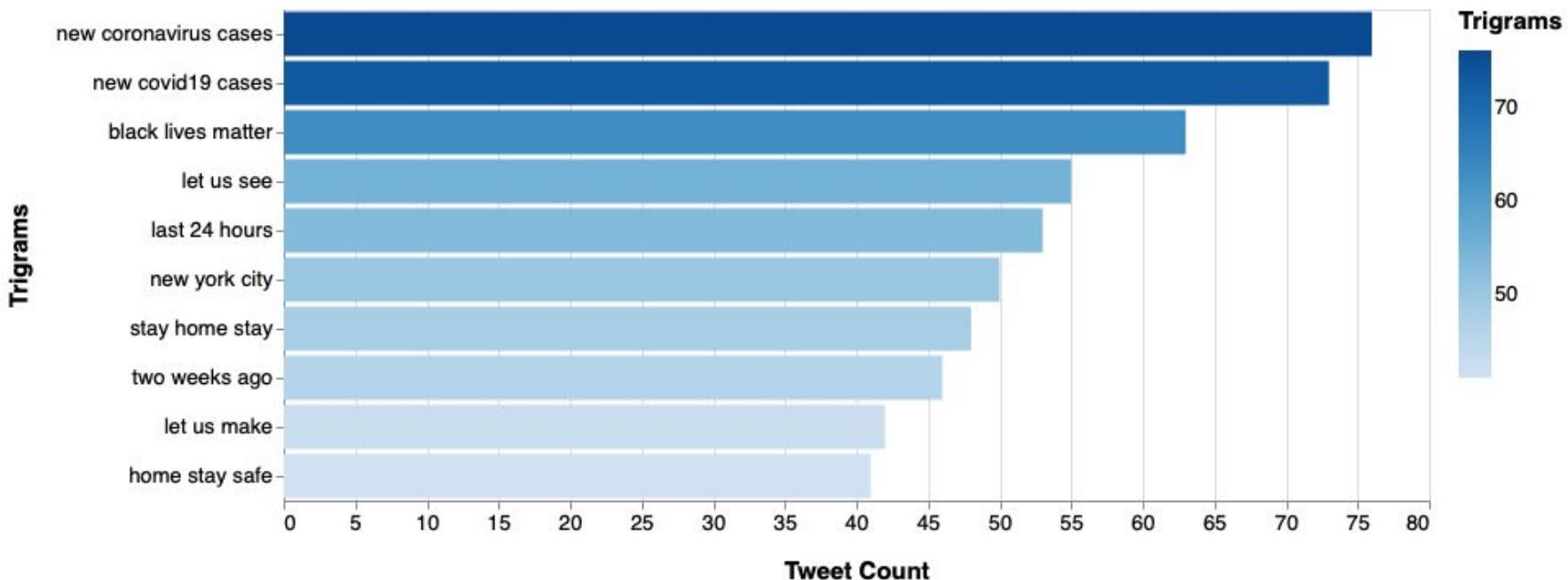
WHICH BIGRAMS HAVE BEEN MORE USED IN THE TWEETS? - DATASET “FAKE”



WHICH TRIGRAMS HAVE BEEN MORE USED IN THE TWEETS? - DATASET “GENERAL”



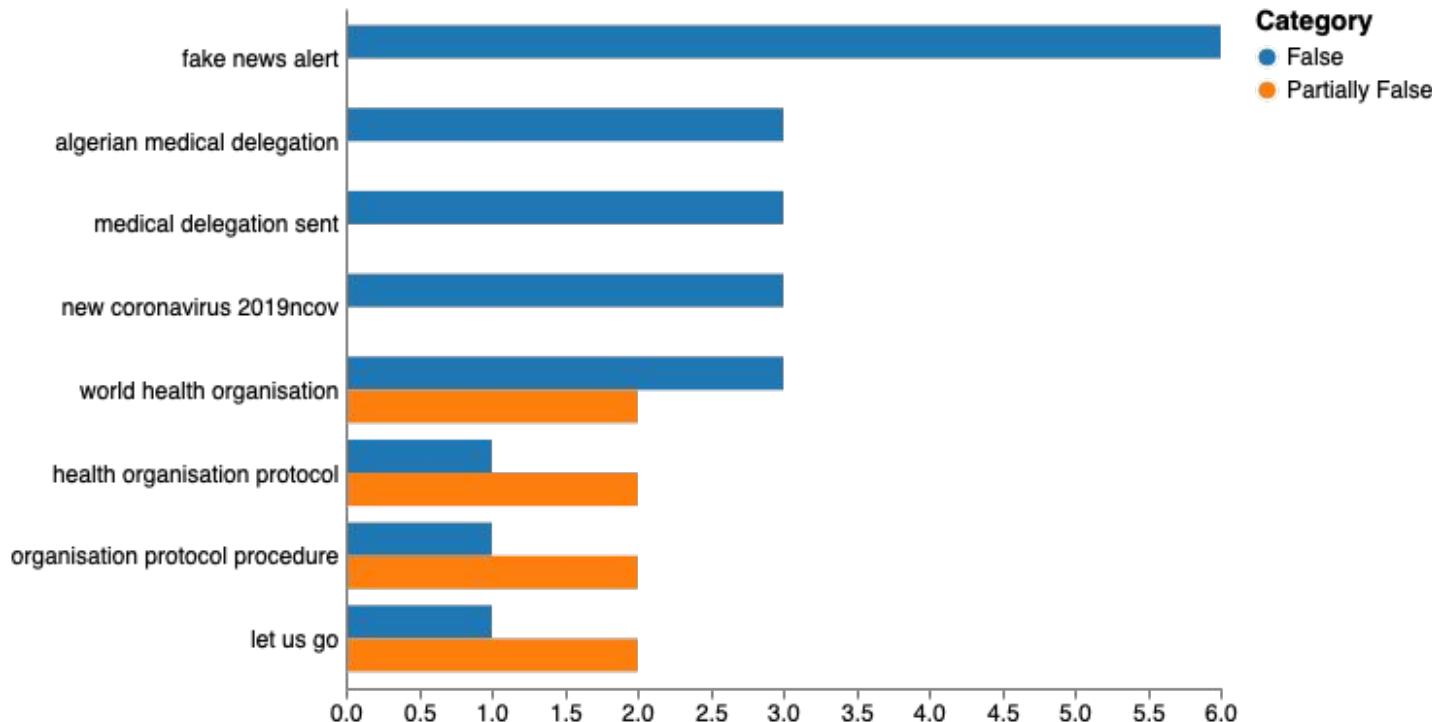
WHICH TRIGRAMS HAVE BEEN MORE USED IN THE TWEETS? - DATASET “GENERAL”



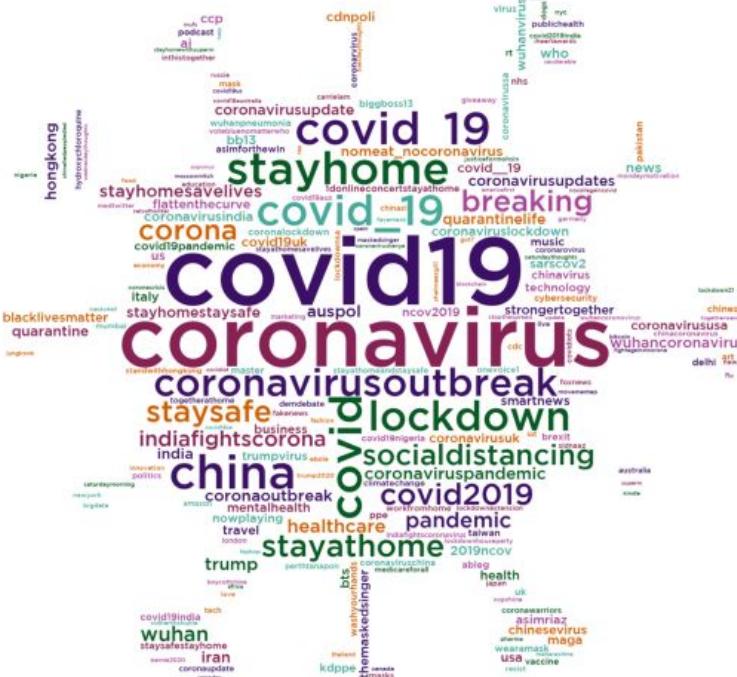
WHICH TRIGRAMS HAVE BEEN MORE USED IN THE TWEETS? - DATASET “FAKE”



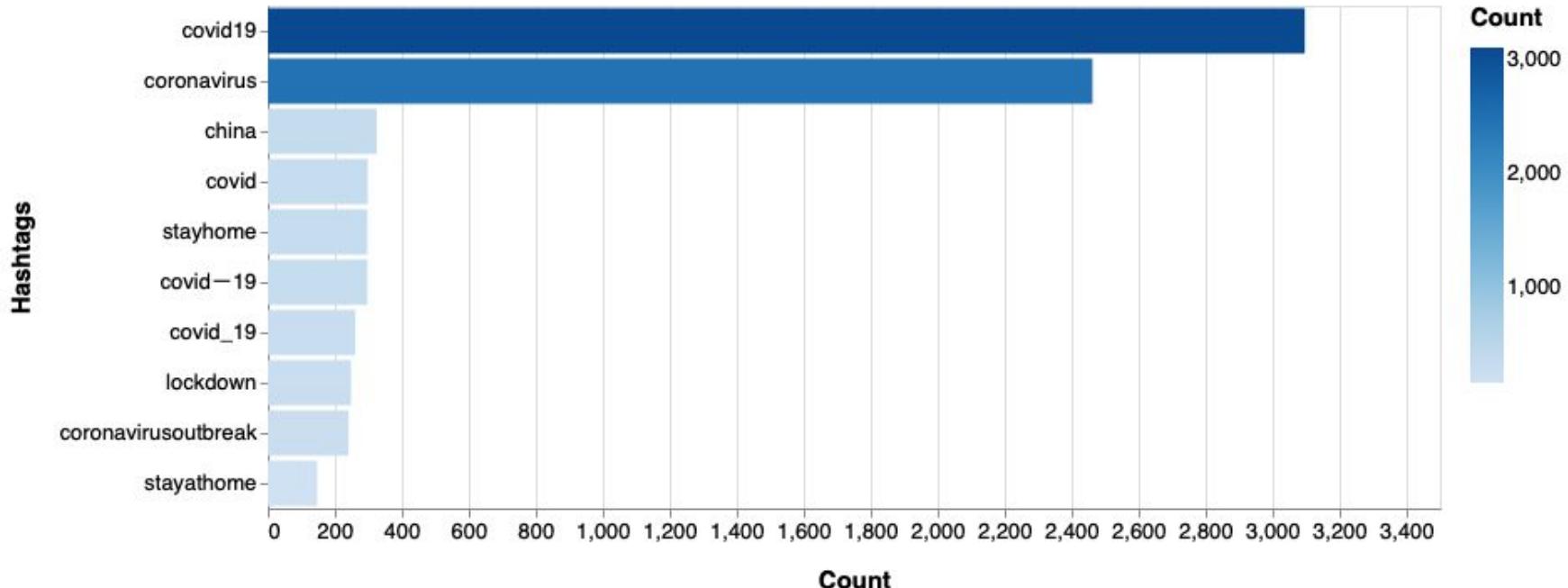
WHICH TRIGRAMS HAVE BEEN MORE USED IN THE TWEETS? - DATASET “FAKE”



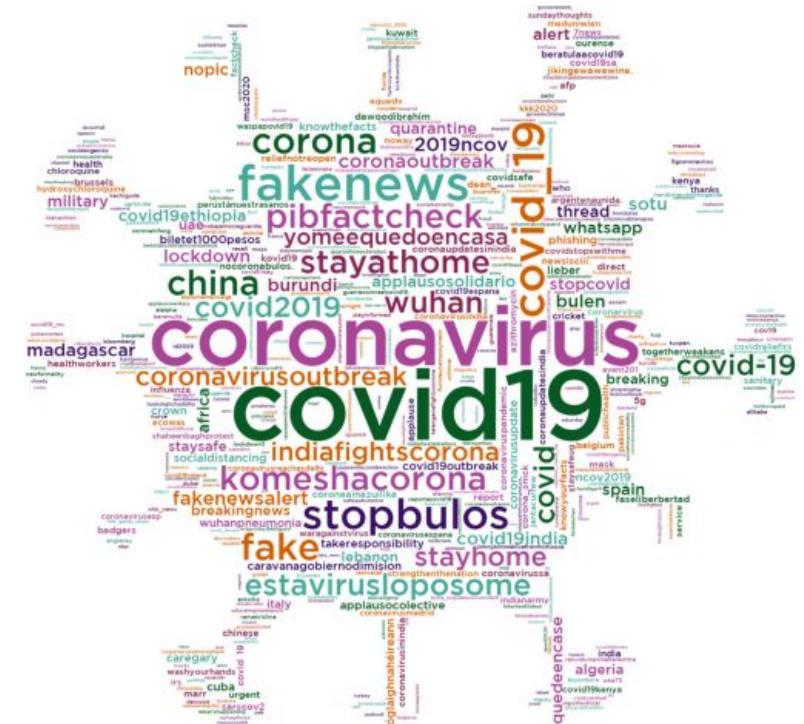
WHICH HASHTAGS HAVE BEEN MORE USED IN THE TWEETS? (COVID-19 RELATED) - DATASET “GENERAL”



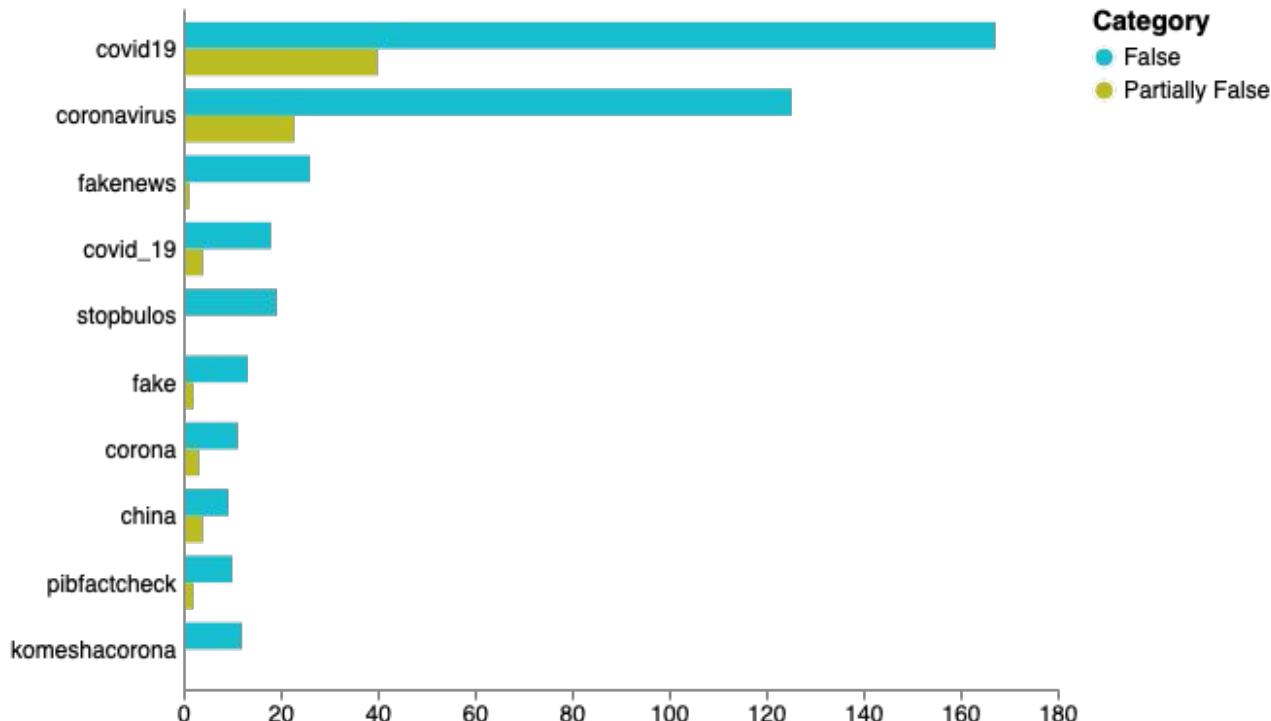
WHICH HASHTAGS HAVE BEEN MORE USED IN THE TWEETS? (COVID-19 RELATED) - DATASET “GENERAL”



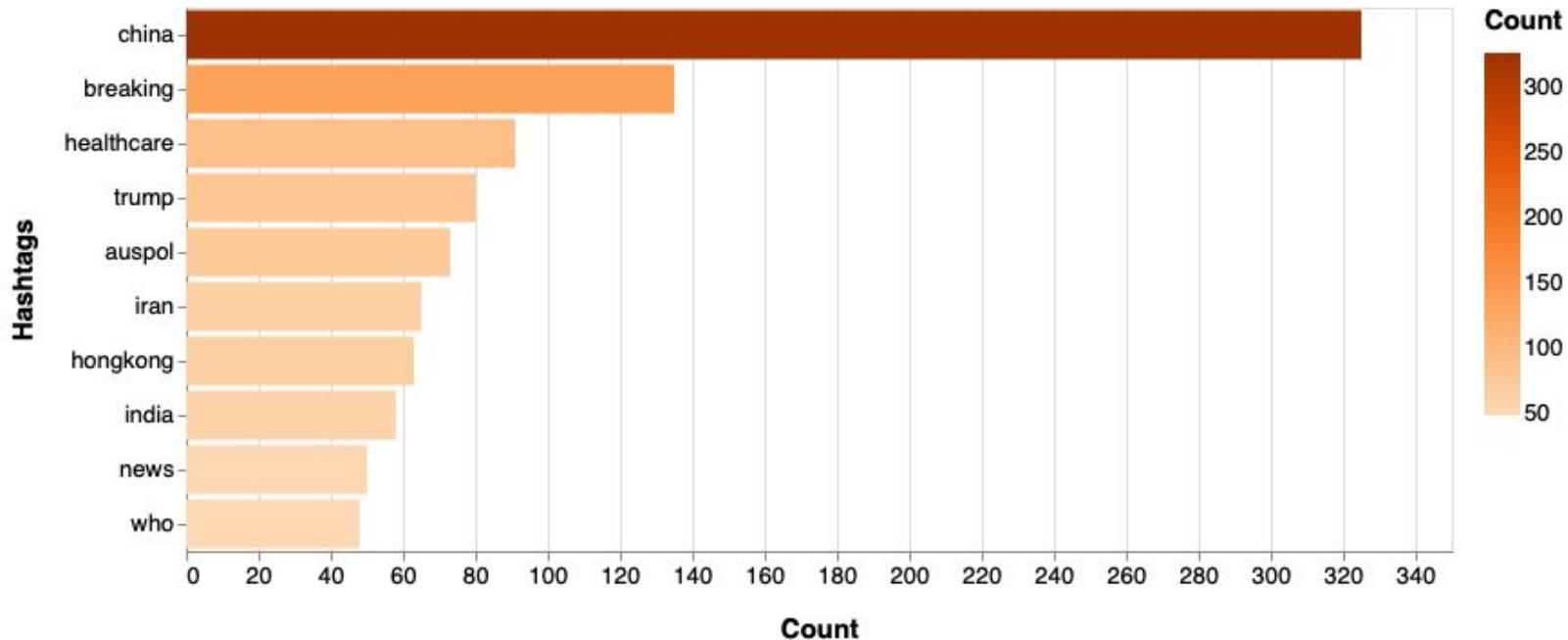
WHICH HASHTAGS HAVE BEEN MORE USED IN THE TWEETS? (COVID-19 RELATED) - DATASET “FAKE”



WHICH HASHTAGS HAVE BEEN MORE USED IN THE TWEETS? (COVID-19 RELATED) - DATASET “FAKE”



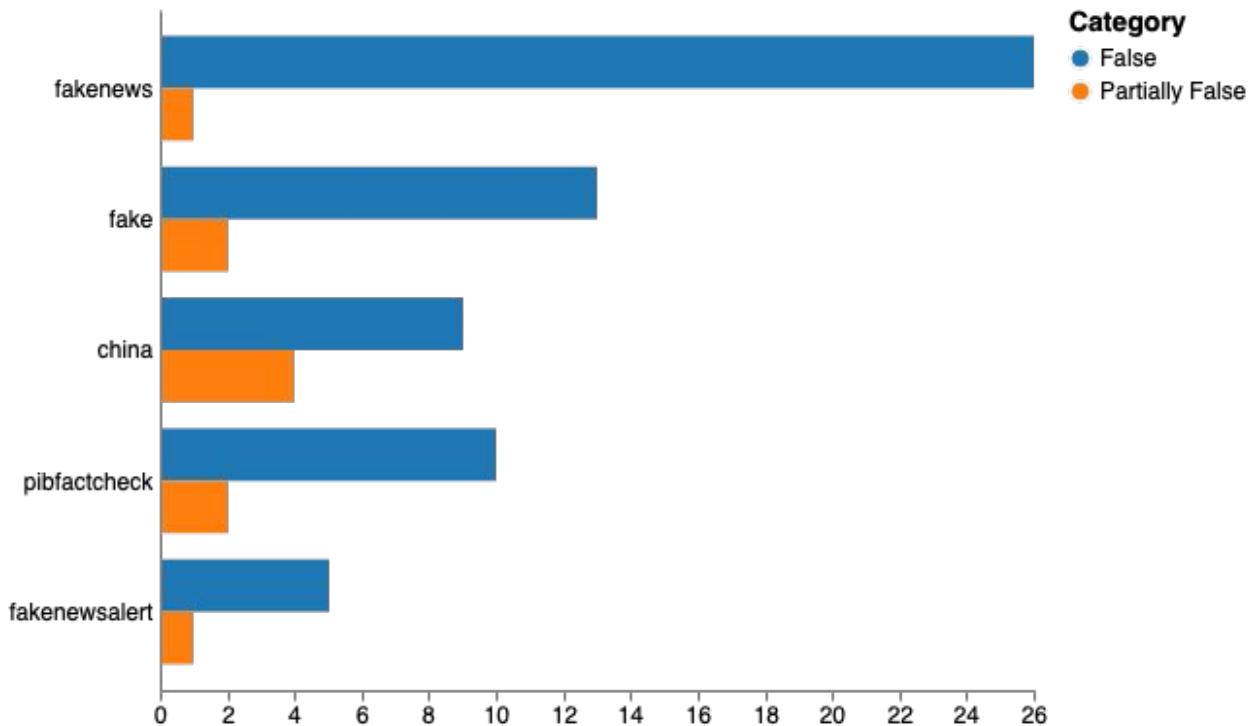
WHICH HASHTAGS HAVE BEEN MORE USED IN THE TWEETS? (NOT COVID-19 RELATED) DATASET “GENERAL”



WHICH HASHTAGS HAVE BEEN MORE USED IN THE TWEETS? (NOT COVID-19 RELATED) - DATASET “FAKE”



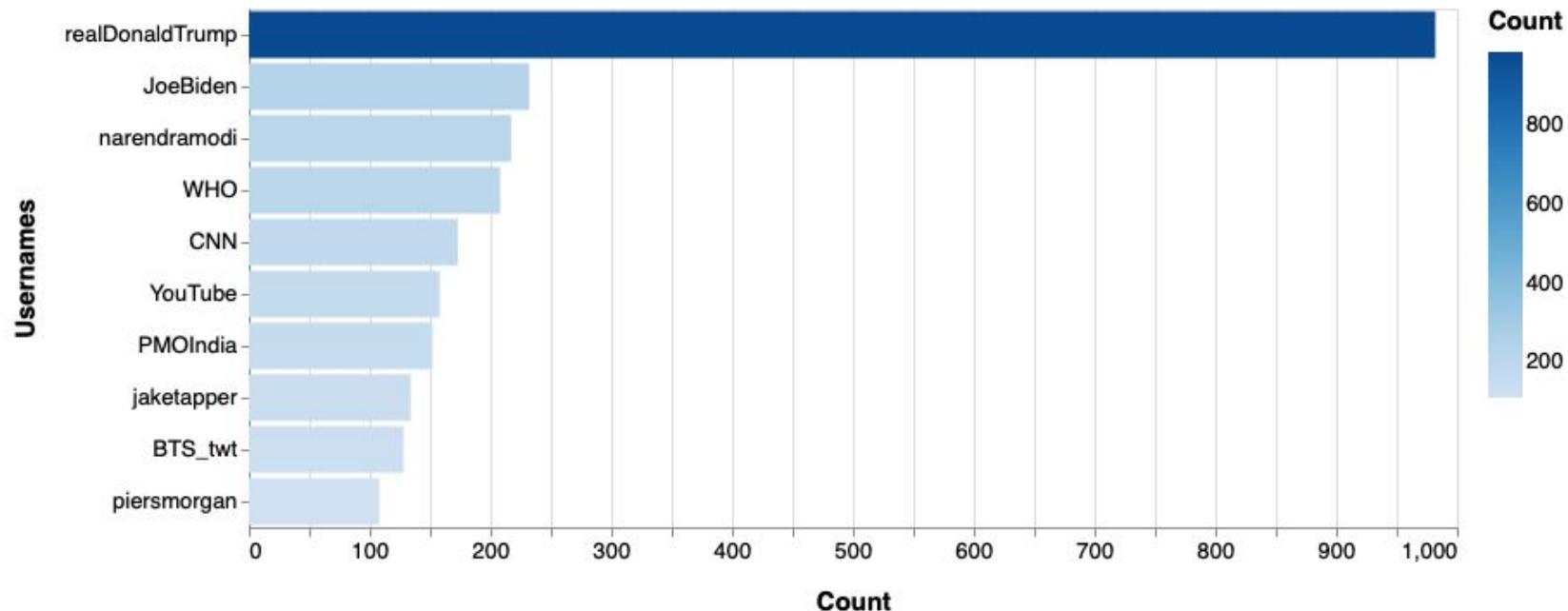
WHICH HASHTAGS HAVE BEEN MORE USED IN THE TWEETS? (NOT COVID-19 RELATED) - DATASET “FAKE”



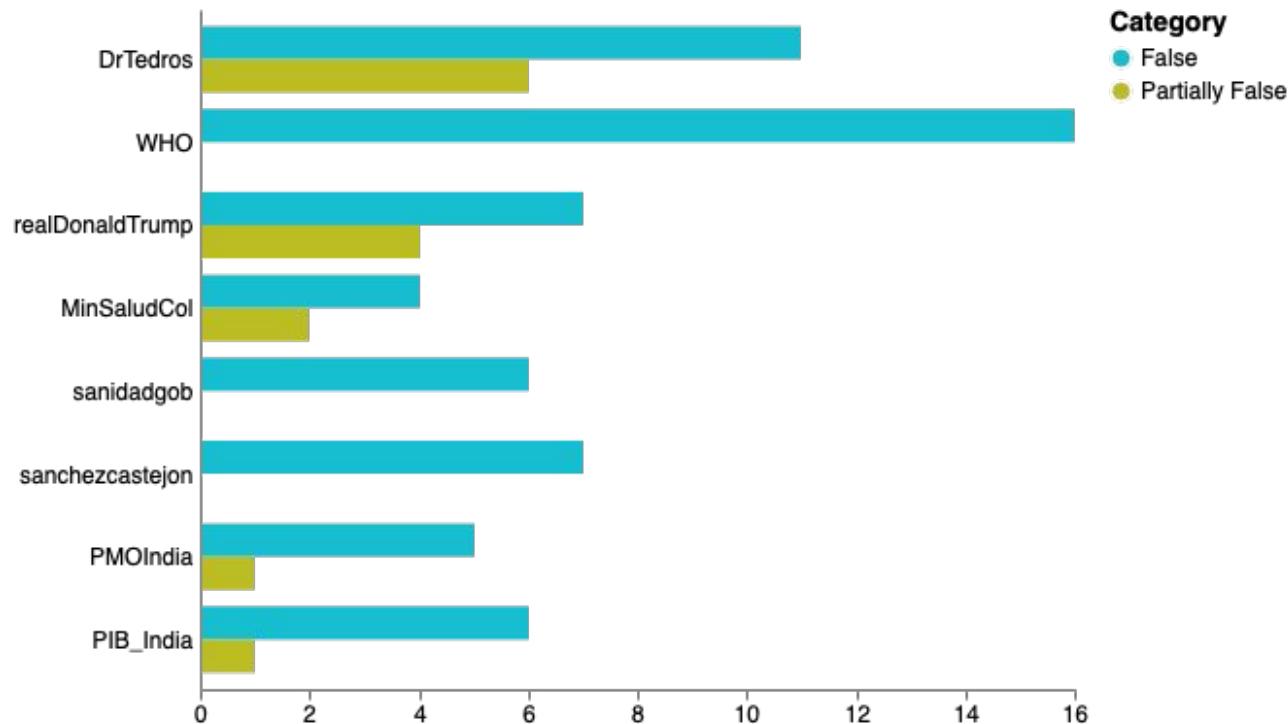
WHICH USERS HAVE BEEN MORE MENTIONED IN THE TWEETS? - DATASET “GENERAL”



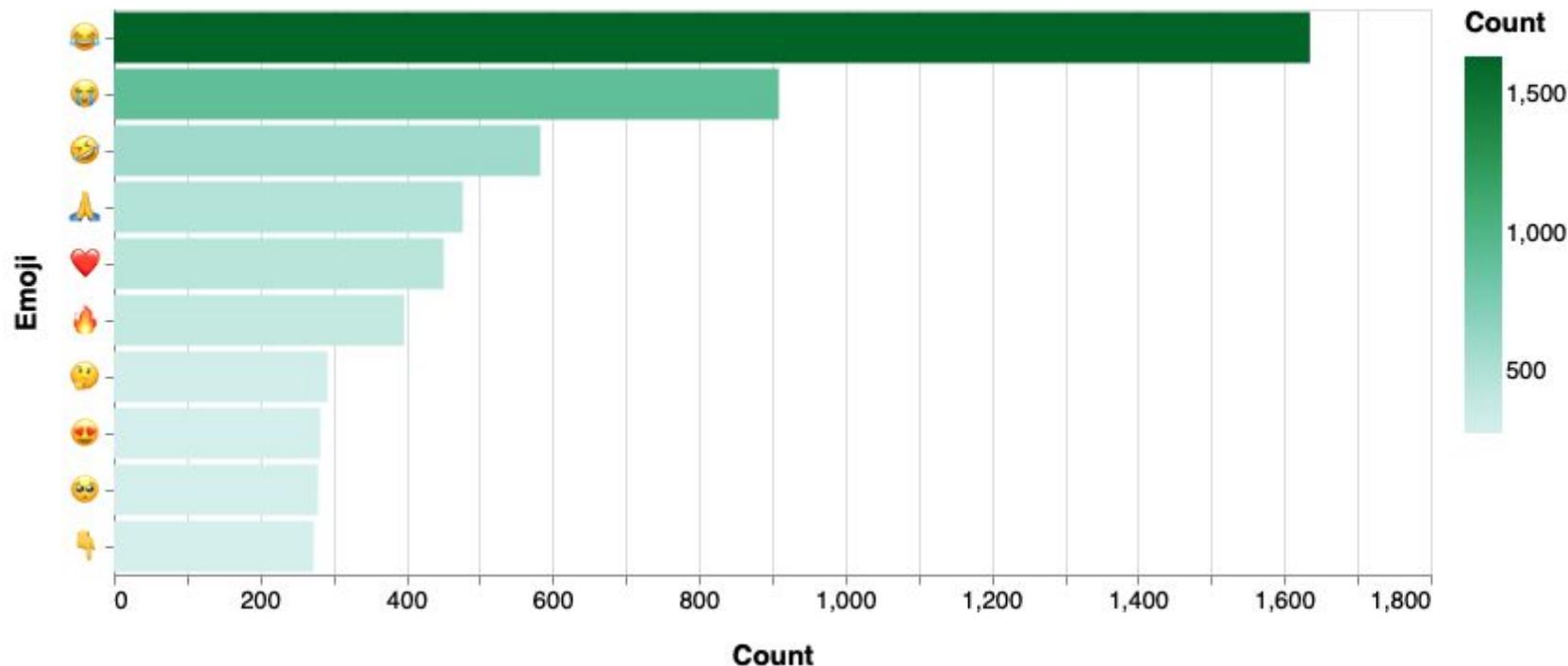
WHICH USERS HAVE BEEN MORE MENTIONED IN THE TWEETS? - DATASET “GENERAL”



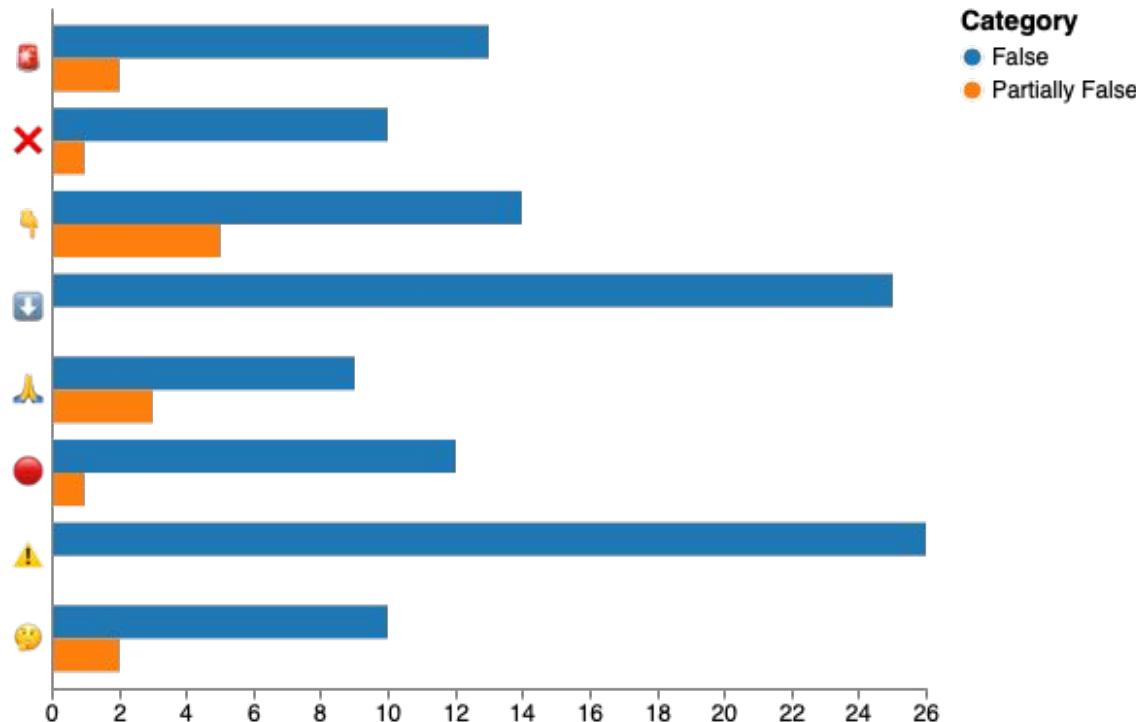
WHICH USERS HAVE BEEN MORE MENTIONED IN THE TWEETS? - DATASET “FAKE”



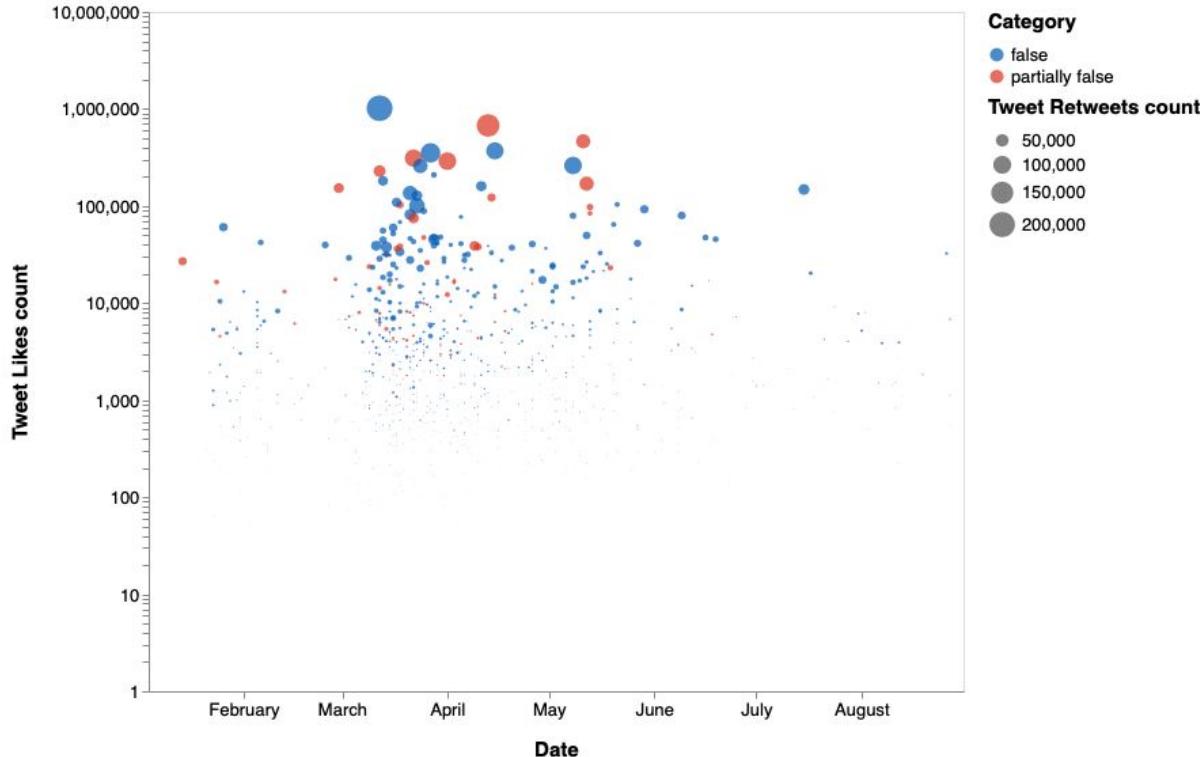
WHAT ARE THE MOST FREQUENT EMOJIS USED IN THE TWEETS? - DATASET “GENERAL”



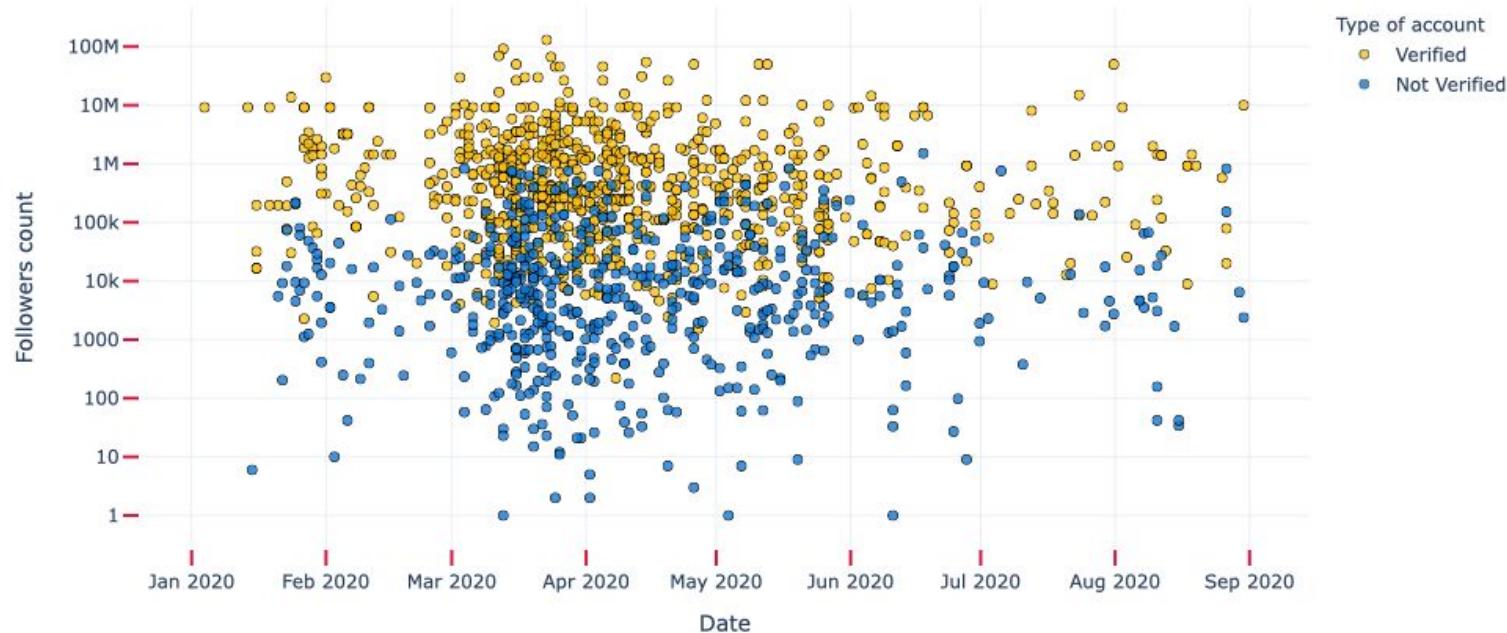
WHAT ARE THE MOST FREQUENT EMOJIS USED IN THE TWEETS? - DATASET “FAKE”



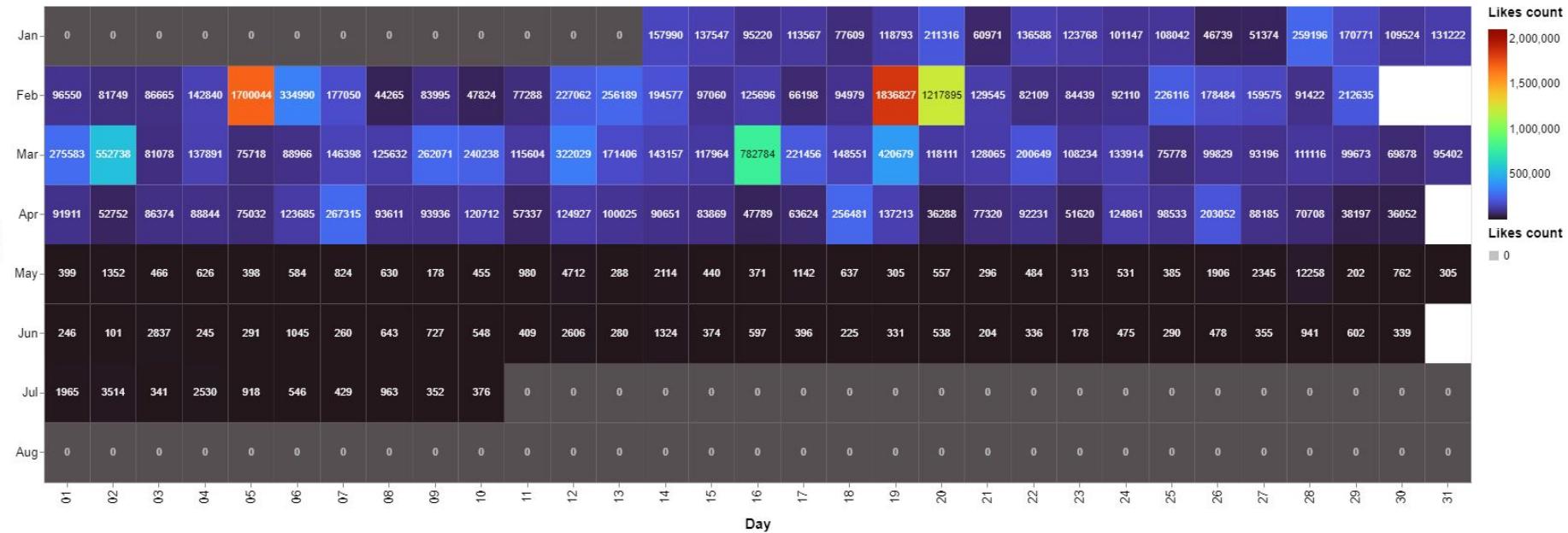
HOW POPULAR WERE THE TWEETS ABOUT COVID-19? - DATASET “FAKE”



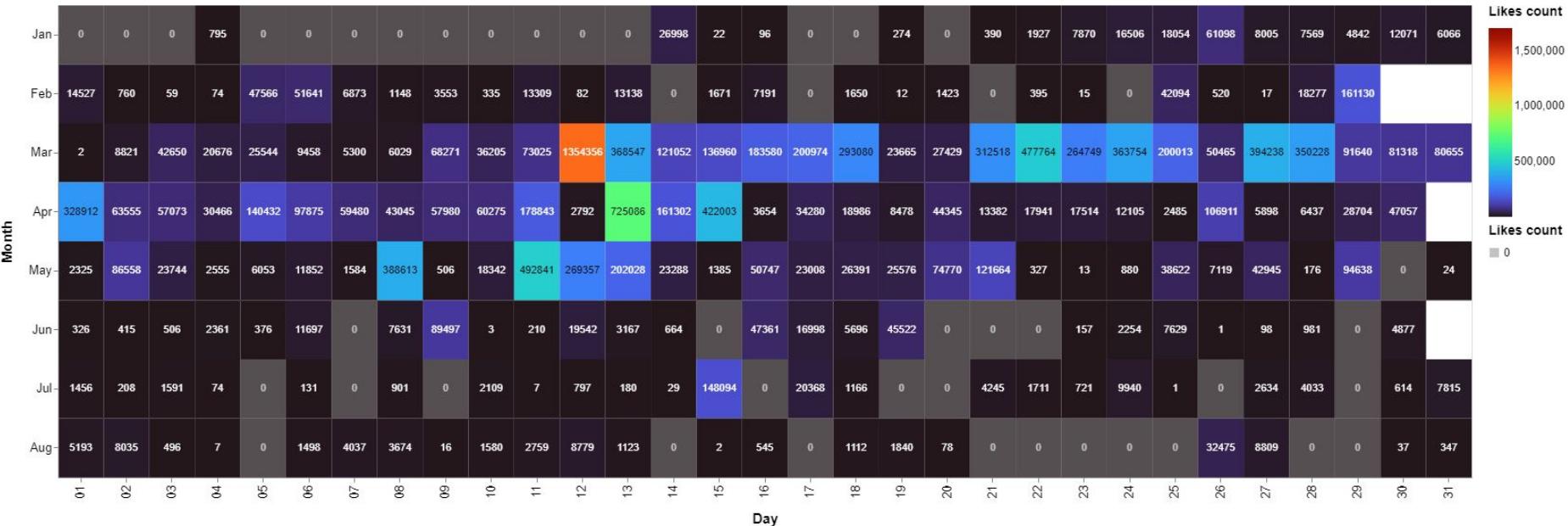
HOW FAMOUS WERE THE ACCOUNTS THAT TWEETED ABOUT COVID-19? - DATASET “FAKE”



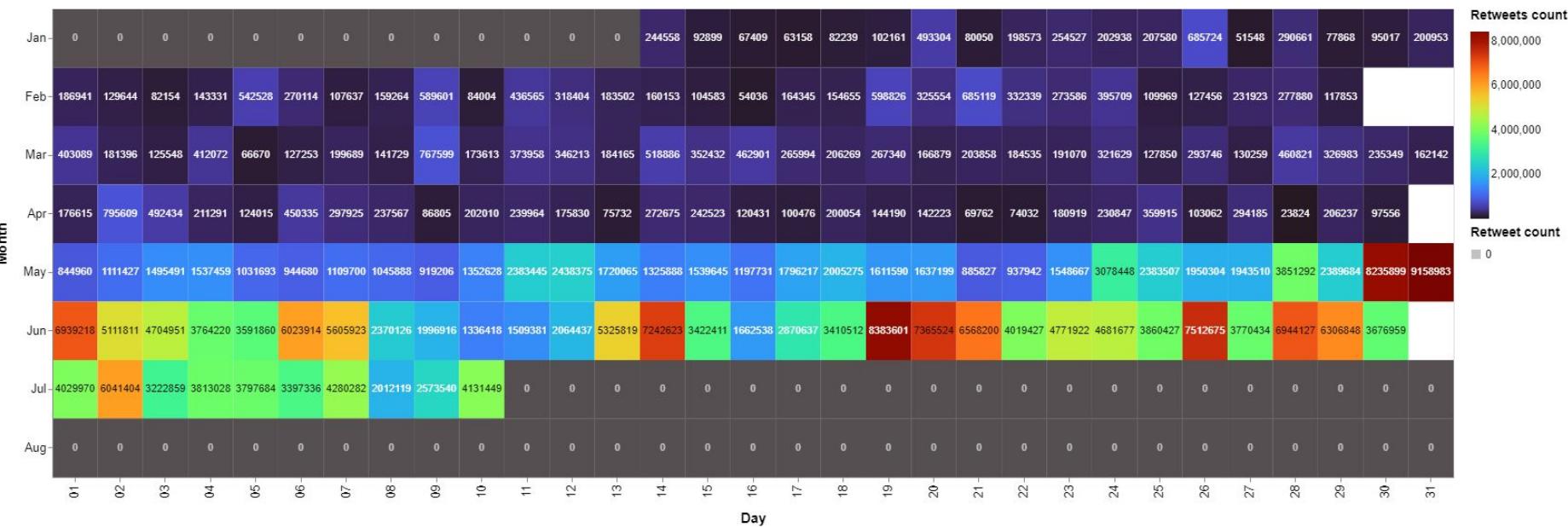
HOW MANY LIKES HAVE THE TWEETS RECEIVED DURING THE YEAR? - DATASET “GENERAL”



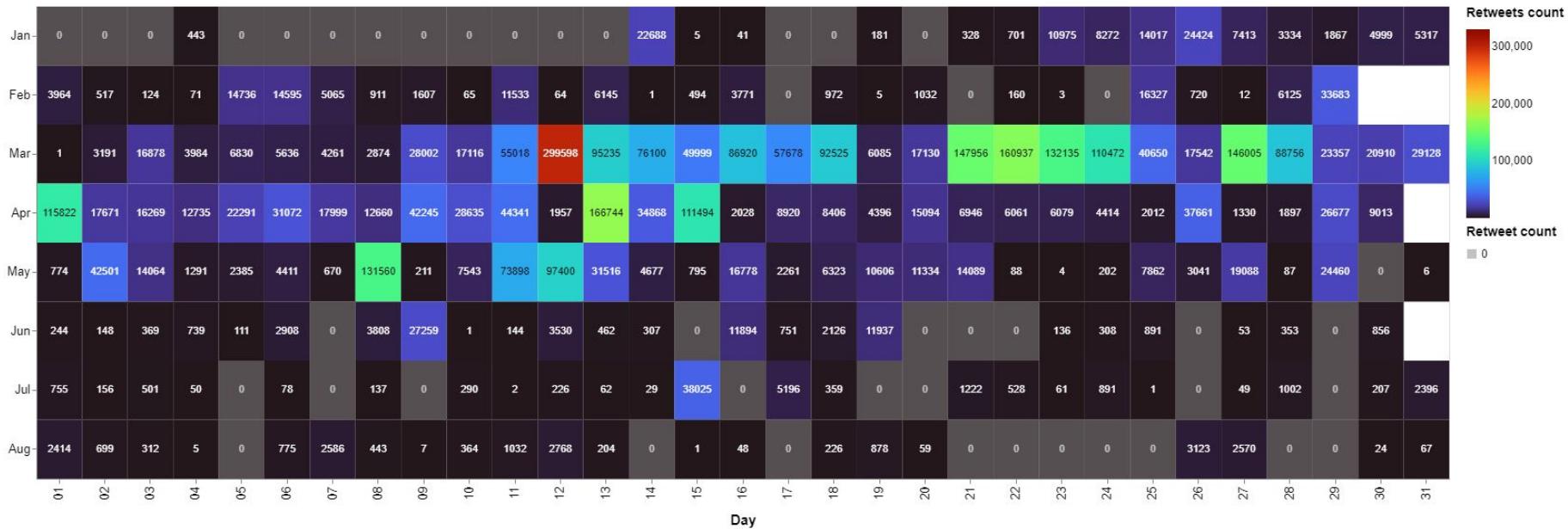
HOW MANY LIKES HAVE THE TWEETS RECEIVED DURING THE YEAR? - DATASET “FAKE”



HOW MANY RETWEETS HAVE BEEN CREATED DURING THE YEAR? - DATASET “GENERAL”



HOW MANY RETWEETS HAVE BEEN CREATED DURING THE YEAR? - DATASET “FAKE”



THE DIFFUSION OF COVID-19 TWEETS

According to the work of Stieglitz and Dang-Xuan^[1], **the diffusion of tweets can be analysed in terms of likes and retweets.**

We have visualised the number of likes and retweet gained by the tweets in the two datasets for each month.

There is a considerable variance in the count of retweet and favourite, so we decided to normalise the data.

We normalise the count of likes and retweet using the **Min-Max Normalization**.

THE DIFFUSION OF COVID-19 TWEETS

MIN/MAX NORMALIZATION

We normalised the count of retweet and likes for the overall month together using the **Min-Max Normalization**.

This is how this type of normalization works:

$$\frac{\text{value} - \text{min}}{\text{max} - \text{min}}$$

Considering a set of data:

- the **minimum value** of the set gets transformed into a **0**;
- the **maximum value** gets transformed into a **1**;
- **every other value** gets transformed into **a decimal between 0 and 1**.

THE DIFFUSION OF COVID-19 TWEETS DATASET “GENERAL”

[Check the notebook on GitHub](#)

THE DIFFUSION OF COVID-19 TWEETS

DATASET “FAKE”

[Check the notebook on GitHub](#)

SENTIMENT ANALYSIS

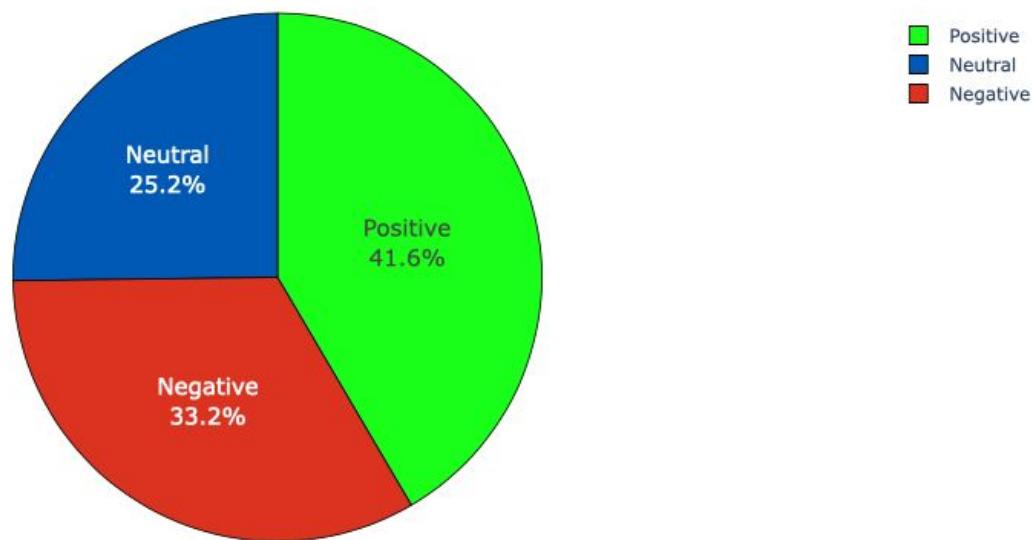
For the sentiment analysis, we used **VADER** (*Valence Aware Dictionary and sEntiment Reasoner*)^[2] is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in **social media**.

For the classification of the tweets as **positive**, **negative** and **neutral tweets**, we have used the **compound score** produced by VADER.

1. **Positive sentiment:** compound score ≥ 0.05
2. **Neutral sentiment:** (compound score > -0.05) and (compound score < 0.05)
3. **Negative sentiment:** compound score ≤ -0.05

CLASSIFICATION OF TWEETS BY SENTIMENT TYPE

DATASET “GENERAL”



EXAMPLE: A POSITIVE TWEET - DATASET “GENERAL”



Sonja @infiniterain · 10 lug 2020

...

The stadium where #BTS 🎤 was supposed to perform in Berlin this weekend, had there been no pandemic, will be lit up in purple for a couple of hours on both original concert dates! ARMYs brought this idea to the venue and they made it happen. That's so cool! 😊



BANGTAN GERMANY 🇩🇪 @BangtanGER_twt · 10 lug 2020

#NEWS | Das Olympiastadion in Berlin (@Oly_Berlin) wird am 11. und 12. Juli von 21:30 bis 24:00 Uhr lila erstrahlen - dieses Wochenende hätte eigentlich die MOTS Tour in Berlin stattgefunden.

Teilt eure Bilder von dem Moment mit dem Tag #OlympiastadionBerlin

©@RoseBangtans

Das Olympiastadion Berlin hätte Euch und BTS dieses Wochenende gern hier gehabt.

PS: Vom Teufelsberg hat man einen schönen Blick auf's Stadion. 😊

Viele Grüße vom Oly-Team,
Julia



EXAMPLE: A NEUTRAL TWEET - DATASET “GENERAL”



Mathi @iammathii · 7 lug 2020

...

Watch | How Dharavi curbed the spread of COVID-19:



Watch | How Dharavi curbed the spread of COVID-19

A video on how India's largest slum curbed the spread of the virus

🔗 [thehindu.com](https://www.thehindu.com)



EXAMPLE: A NEGATIVE TWEET - DATASET “GENERAL”



NancyNoCo  @NancyNoCo24 · 5 mar 2020

...

In Tennessee and other less affluent states, there are fewer hospitals.

Tennessee has coronavirus tests for ONLY 85 more people to test.

A state lawmaker is downplaying the severity of this virus by EQUATING IT WITH THE FLU!!!

GET SERIOUS!!!

#wtpTEAM

#wtp2020

@wtp_2020



Beth Joslin Roth @BethJoslinRoth · 5 mar 2020

In a state where hundreds of thousands don't have healthcare, there will be sick people, contagious people, who put off going to the doctor. Many communities no longer even have hospitals. Hard to see how lack of healthcare doesn't exacerbate spread of coronavirus
#ExpandMedicaid twitter.com/Tennessean/sta...

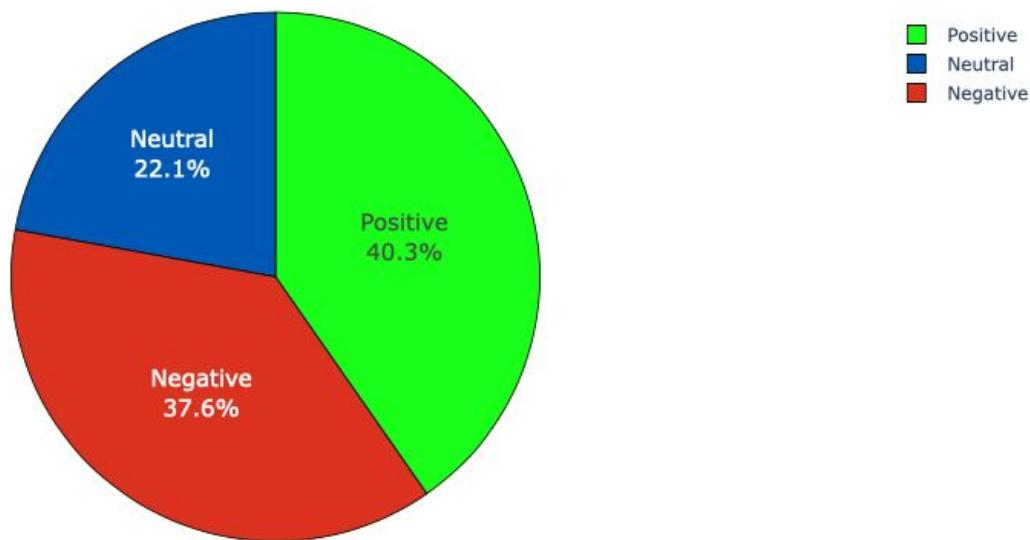
5

33

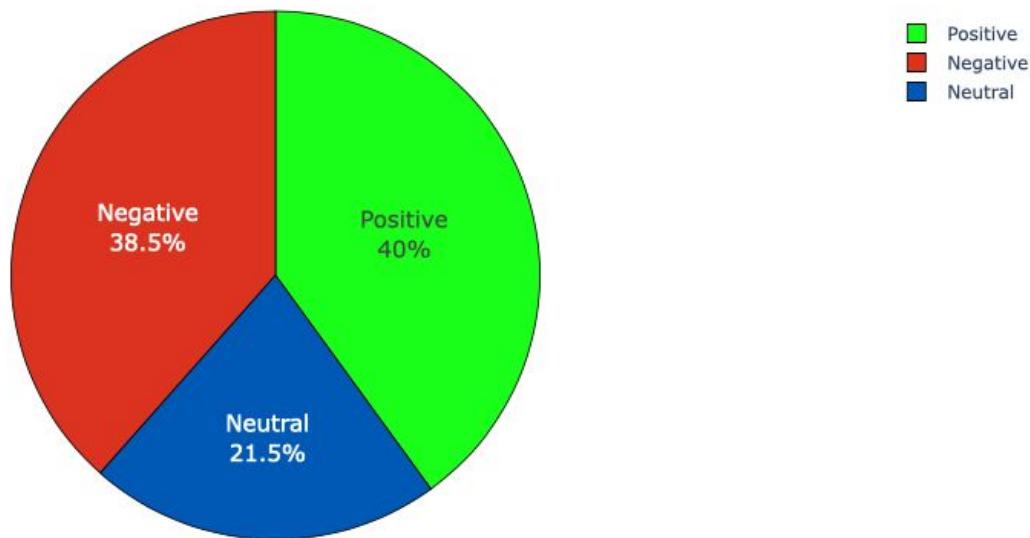
41

↑

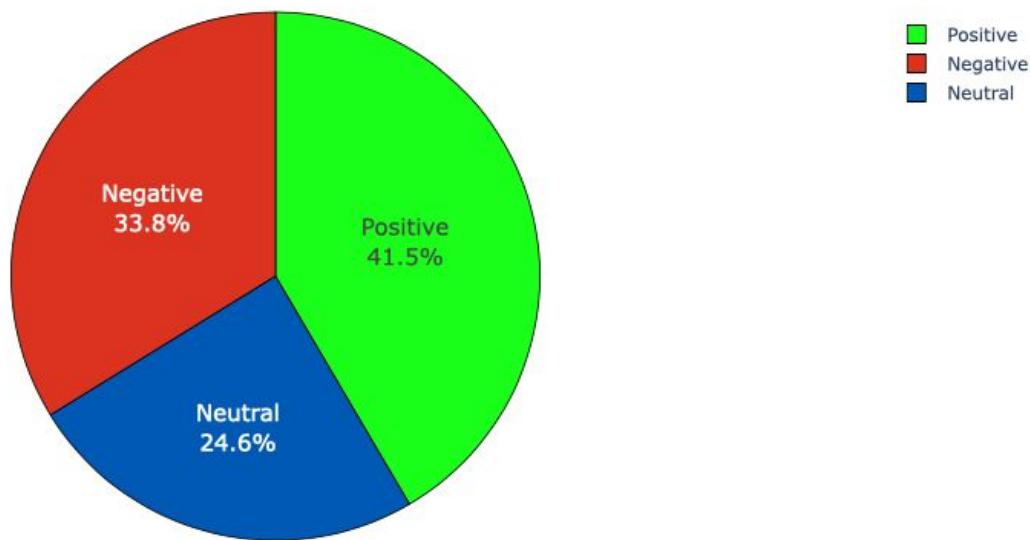
CLASSIFICATION OF TWEETS BY SENTIMENT TYPE (ALL TWEETS) - DATASET “FAKE”



CLASSIFICATION OF TWEETS BY SENTIMENT TYPE (FALSE TWEETS) - DATASET “FAKE”



CLASSIFICATION OF TWEETS BY SENTIMENT TYPE (PARTIALLY FALSE TWEETS) - DATASET “FAKE”



EXAMPLE: A POSITIVE TWEET - DATASET “FAKE”

This is a **false** tweet, read [here](#) why.



Pradeep Dwivedi® ✅ @PradeepDwivedi · 3 apr 2020

...

That's so true, @roshanabbas .. my friends sent me this from my lovely hometown #Chandigarh , literally 'spotted' near Lake Club 😊👉



3

1

9

↑

EXAMPLE: A NEUTRAL TWEET - DATASET “FAKE”

This is a **partially false** tweet, read [here](#) why.



Dón Grieshnak @DGrieshnak · 26 mar 2020

...

Spotted Malabar civet... A critically endangered mammal not seen until 1990 resurfaces for the first time in calicut town.. seems mother earth is rebooting! #COVID2019



400

10.399

25.915



EXAMPLE: A NEGATIVE TWEET - DATASET “FAKE”

This is a **false** tweet, read [here](#) why.

TIMES NOW  @TimesNow · 27 mag 2020 · ...
#Breaking | Watch: Patient films ‘horror’ in a Maharashtra hospital. Dead bodies, without even being wrapped up are found lying in rooms where patients are accommodated.

Details by TIMES NOW’s Arunee. | #RahulDumpsMaharashtra



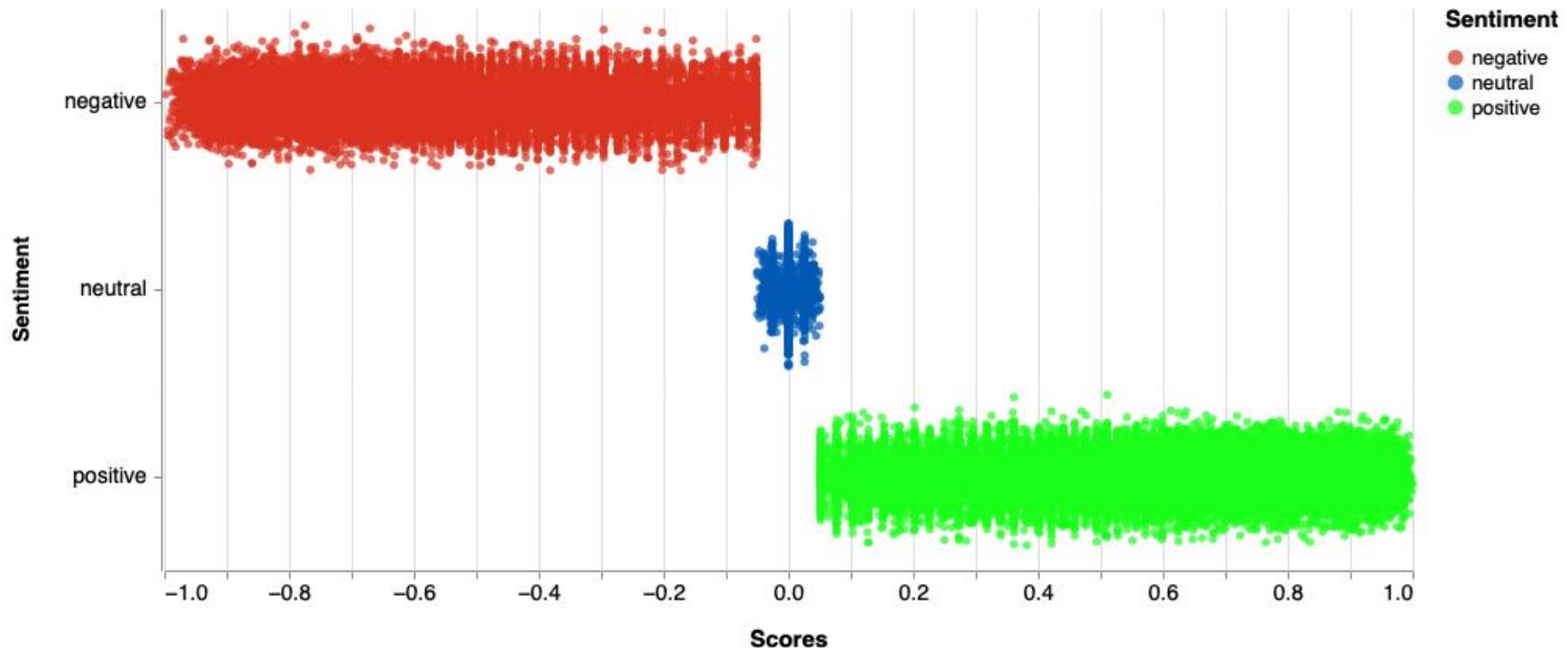
42

269

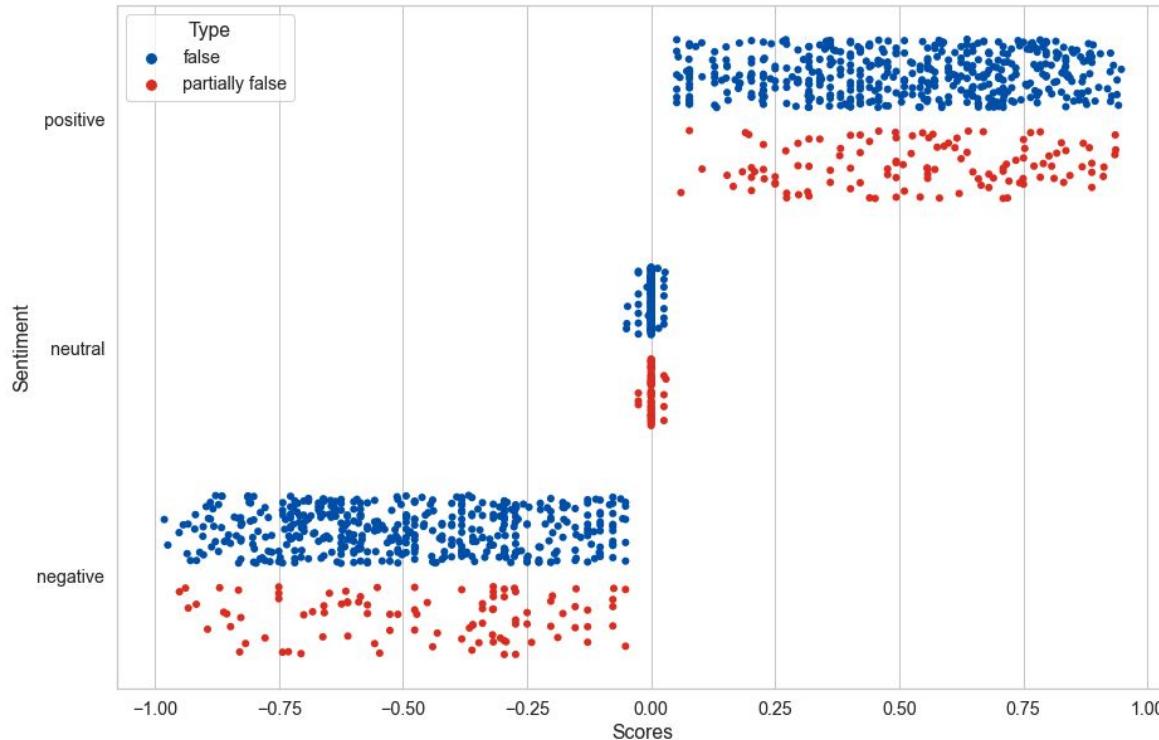
282

↑

SENTIMENT SCORES DISTRIBUTION CLASSIFIED BY SENTIMENT TYPE - DATASET “GENERAL”

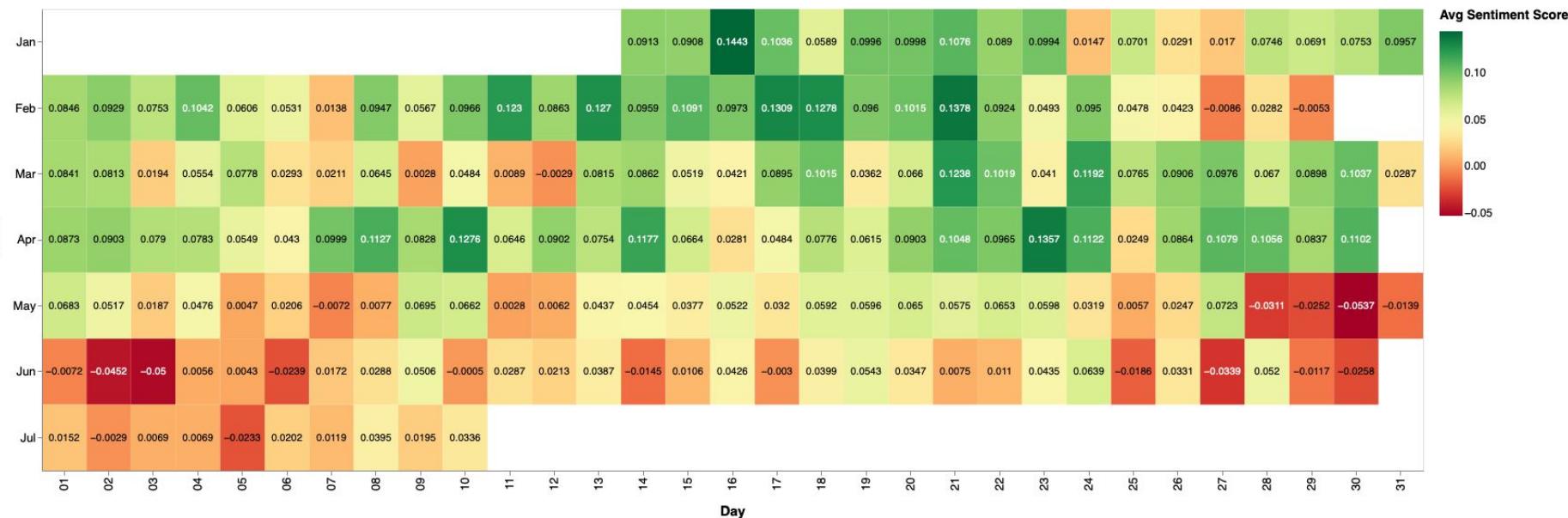


SENTIMENT SCORES DISTRIBUTION CLASSIFIED BY SENTIMENT TYPE AND TWEET CATEGORY - DATASET “FAKE”

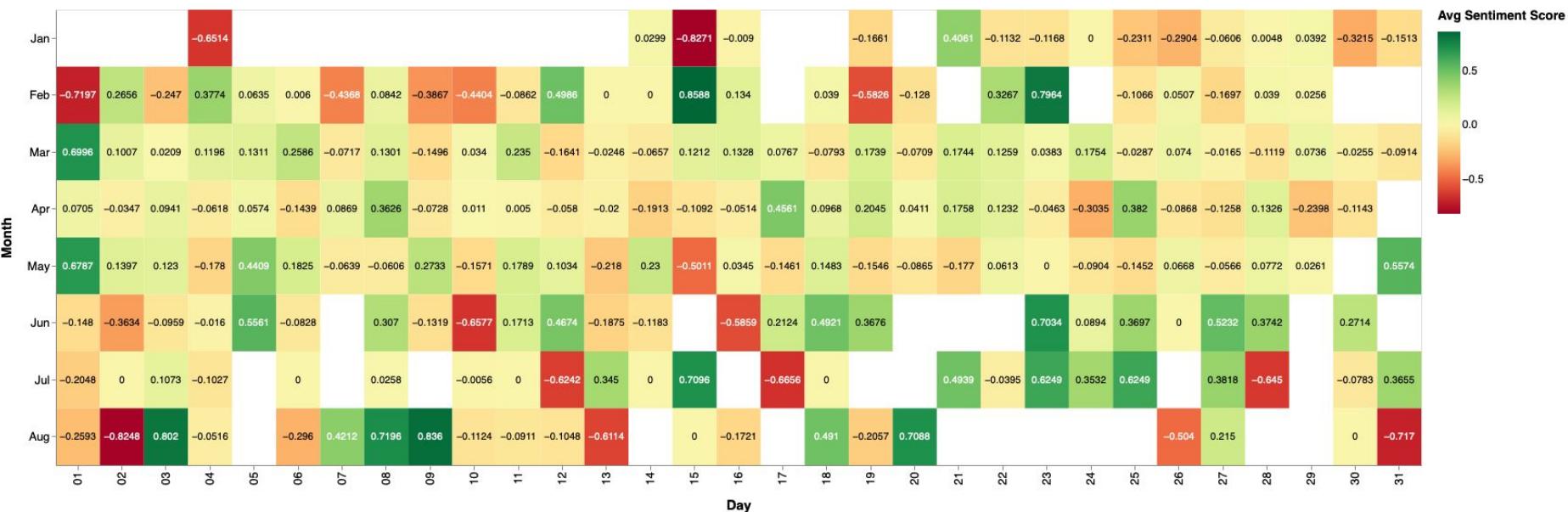


HEATMAP: DAILY AVERAGE SENTIMENT SCORE

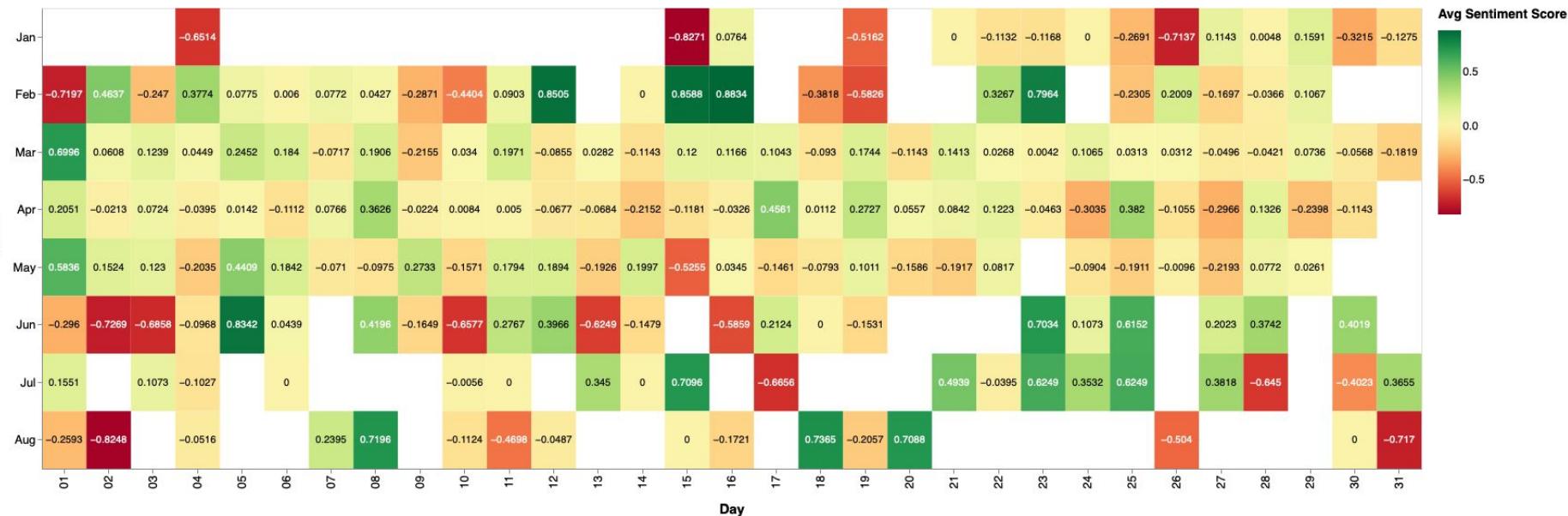
DATASET “GENERAL”



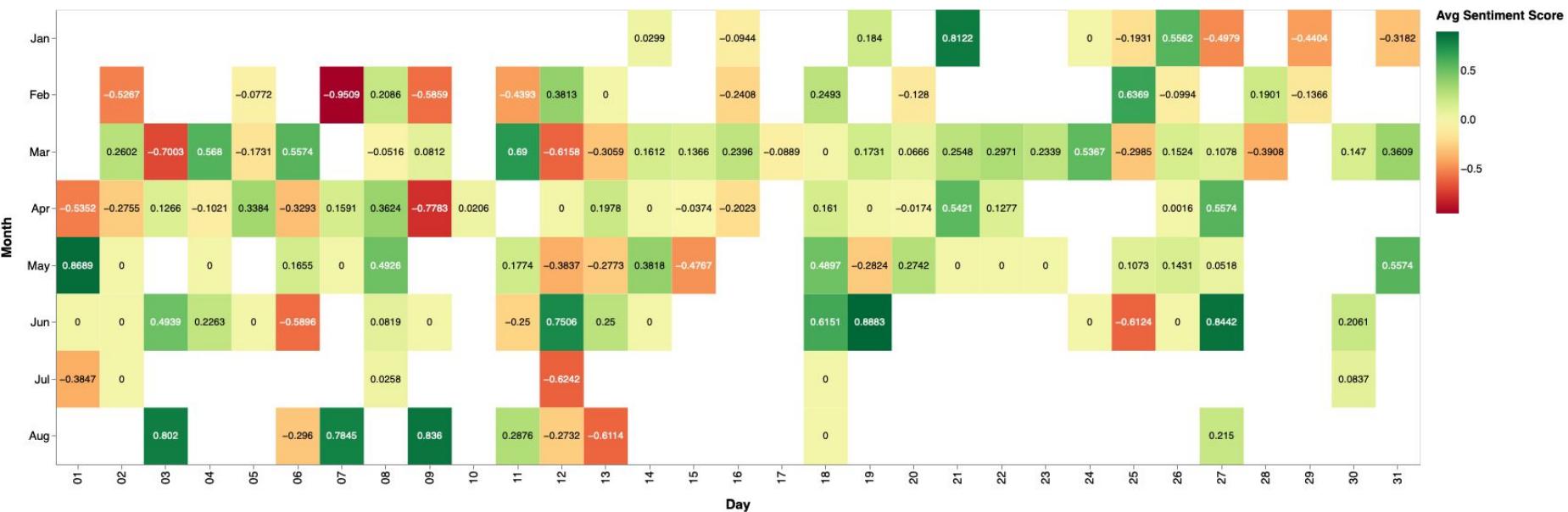
HEATMAP: DAILY AVERAGE SENTIMENT SCORE (ALL TWEETS) - DATASET “FAKE”



HEATMAP: DAILY AVERAGE SENTIMENT SCORE (FALSE TWEETS) - DATASET “FAKE”



HEATMAP: DAILY AVERAGE SENTIMENT SCORE (PARTIALLY FALSE TWEETS) - DATASET “FAKE”



WHAT IS THE SENTIMENT OF THE TWEETS WITH COVID-19 HASHTAGS?

We decided to do a deeper sentiment analysis in both datasets, so we picked **the most frequent Covid-19 related hashtags** in the two dataset and analyzed the sentiment of the tweets that have used them.

We used:

#coronavirus

#covid

#corona

#covid-19

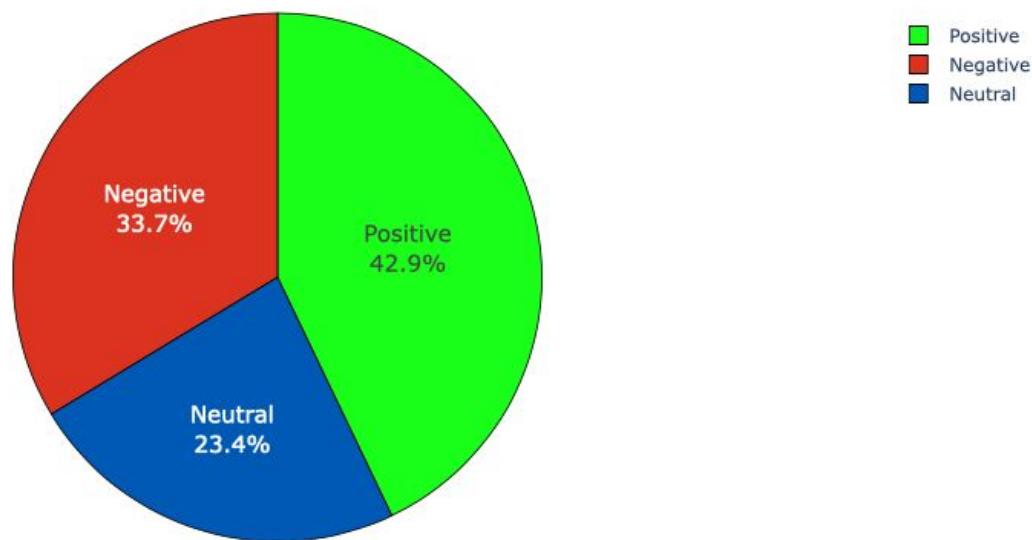
#covid19

#corona

#coronaoutbreak

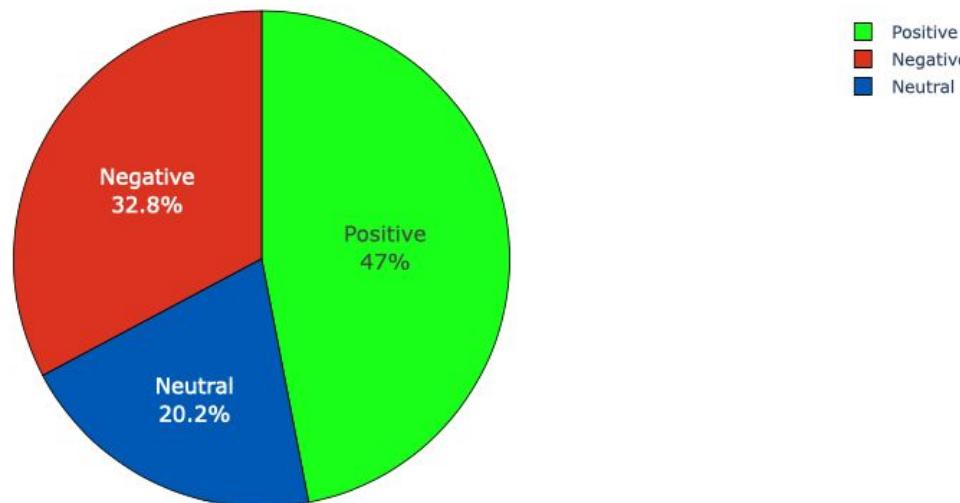
#covid_19

CLASSIFICATION OF TWEETS WITH COVID-19 RELATED HASHTAGS BY SENTIMENT TYPE - DATASET “GENERAL”



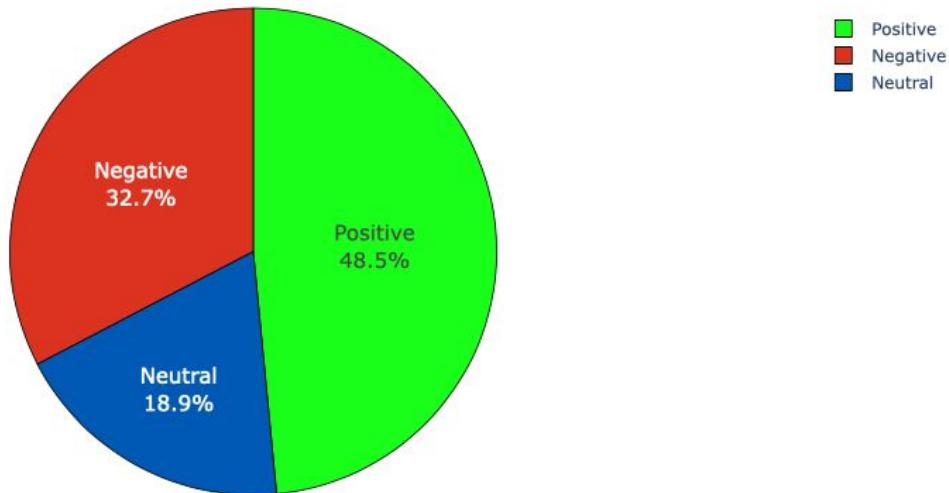
CLASSIFICATION OF TWEETS WITH COVID-19 RELATED HASHTAGS BY SENTIMENT TYPE (ALL TWEETS)

DATASET “FAKE”

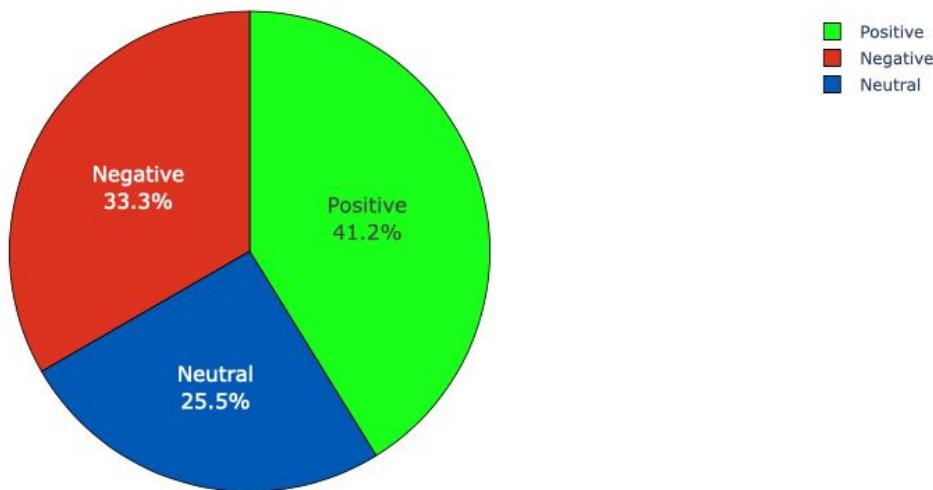


CLASSIFICATION OF TWEETS WITH COVID-19 RELATED HASHTAGS BY SENTIMENT TYPE (FALSE TWEETS)

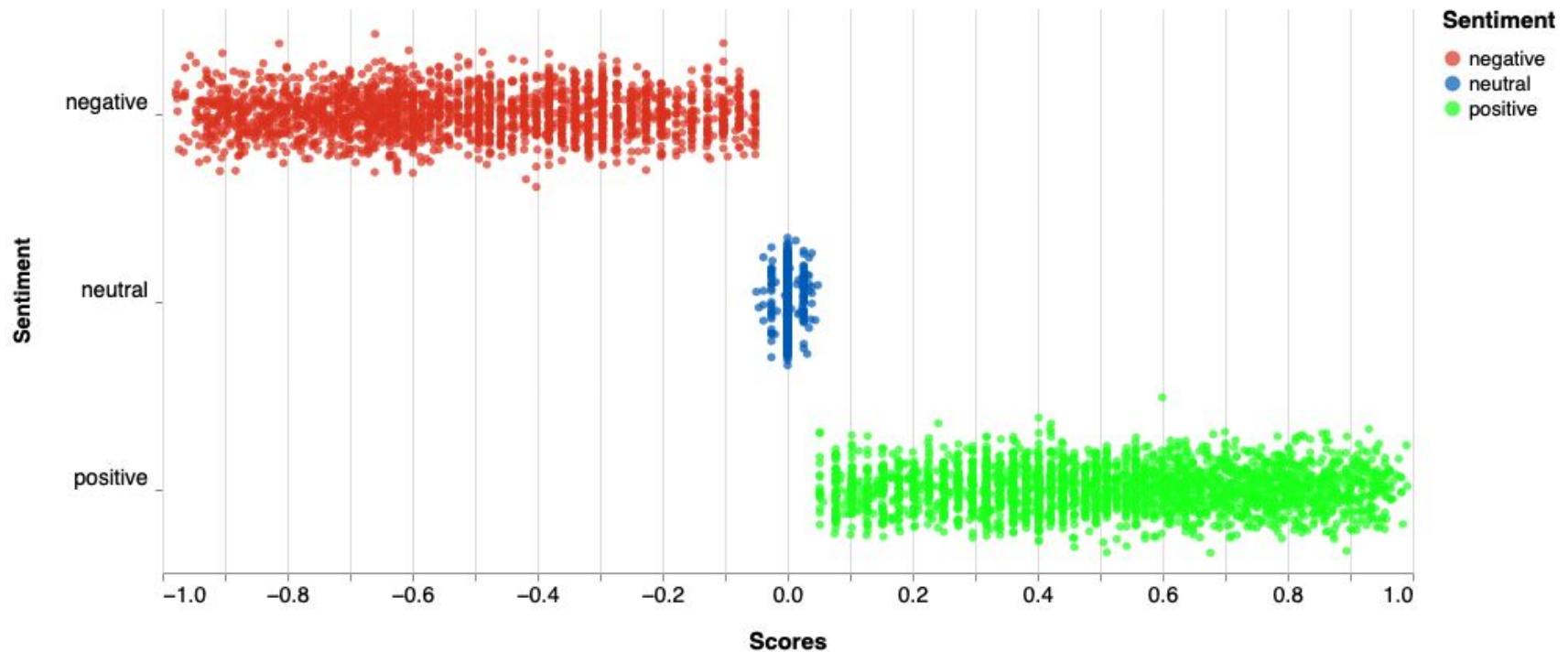
DATASET “FAKE”



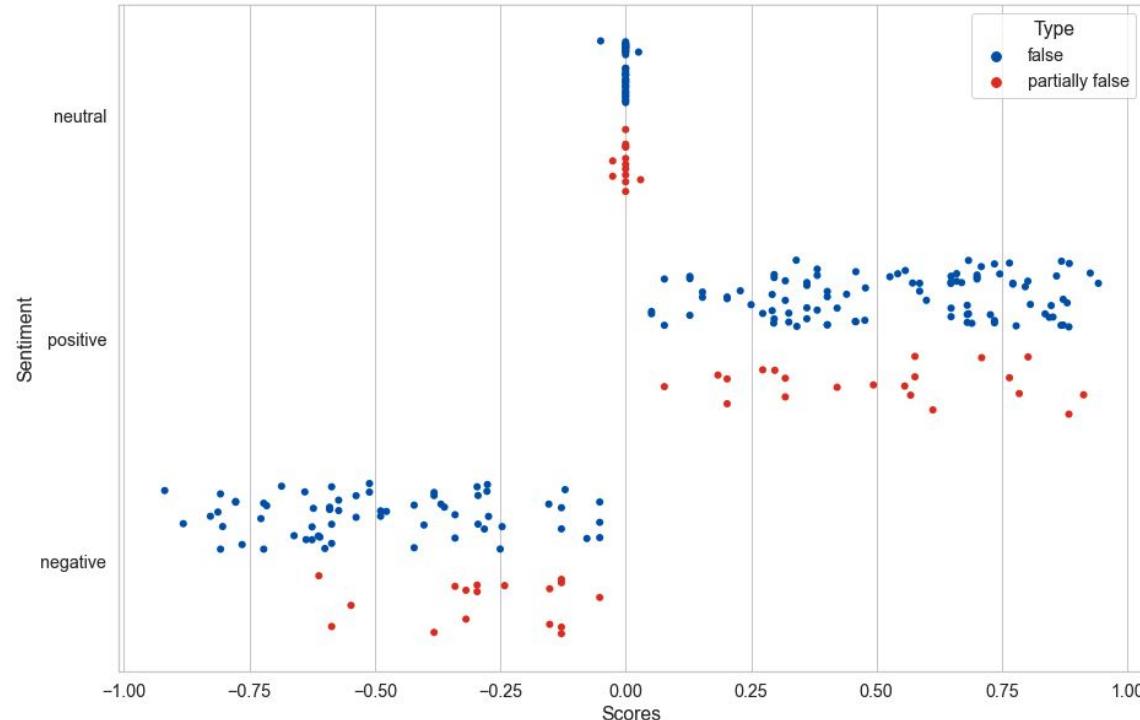
CLASSIFICATION OF TWEETS WITH COVID-19 RELATED HASHTAGS BY SENTIMENT TYPE (PARTIALLY FALSE TWEETS) DATASET “FAKE”



SENTIMENT SCORES DISTRIBUTION CLASSIFIED BY SENTIMENT TYPE (COVID-19 RELATED HASHTAGS) - DATASET “GENERAL”

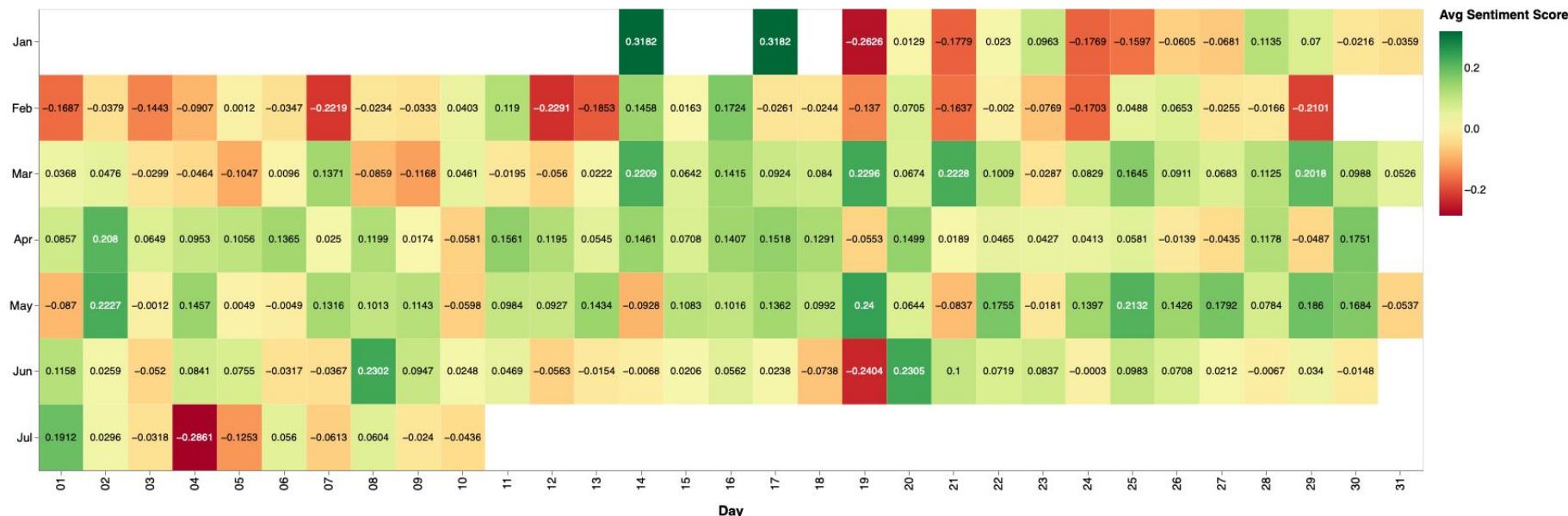


SENTIMENT SCORES DISTRIBUTION CLASSIFIED BY SENTIMENT TYPE AND TWEET CATEGORY (COVID-19 RELATED HASHTAGS) DATASET “FAKE”



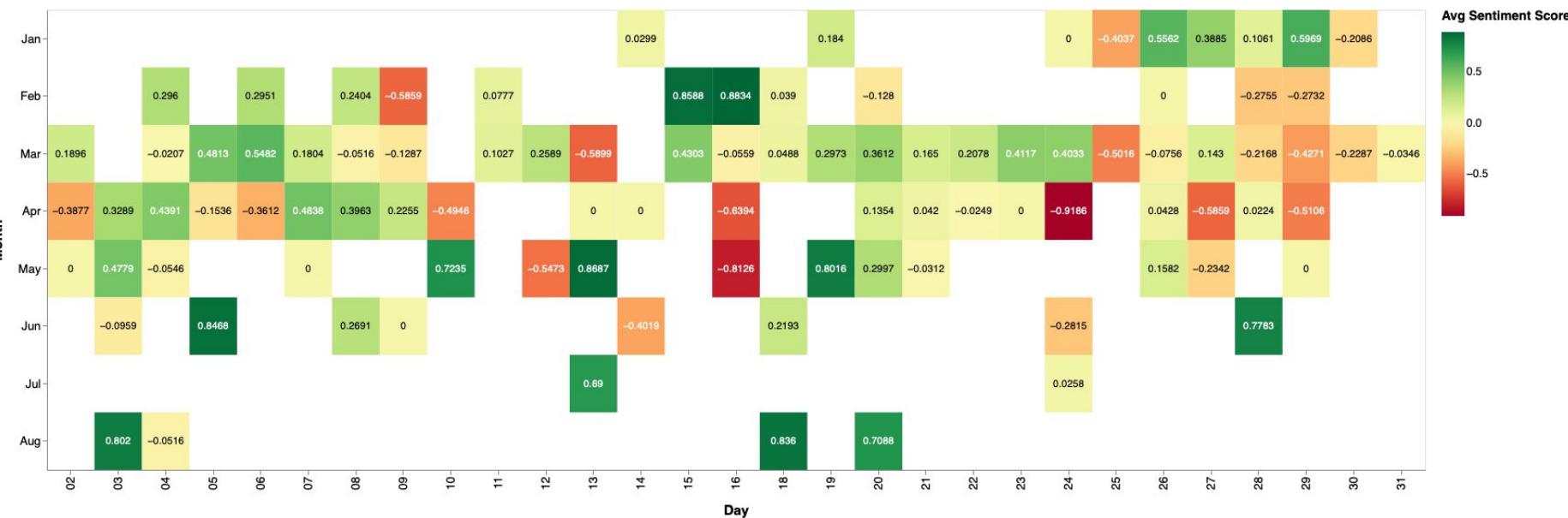
DAILY AVERAGE SENTIMENT SCORE (TWEETS WITH COVID-19 RELATED HASHTAGS)

Dataset “General”



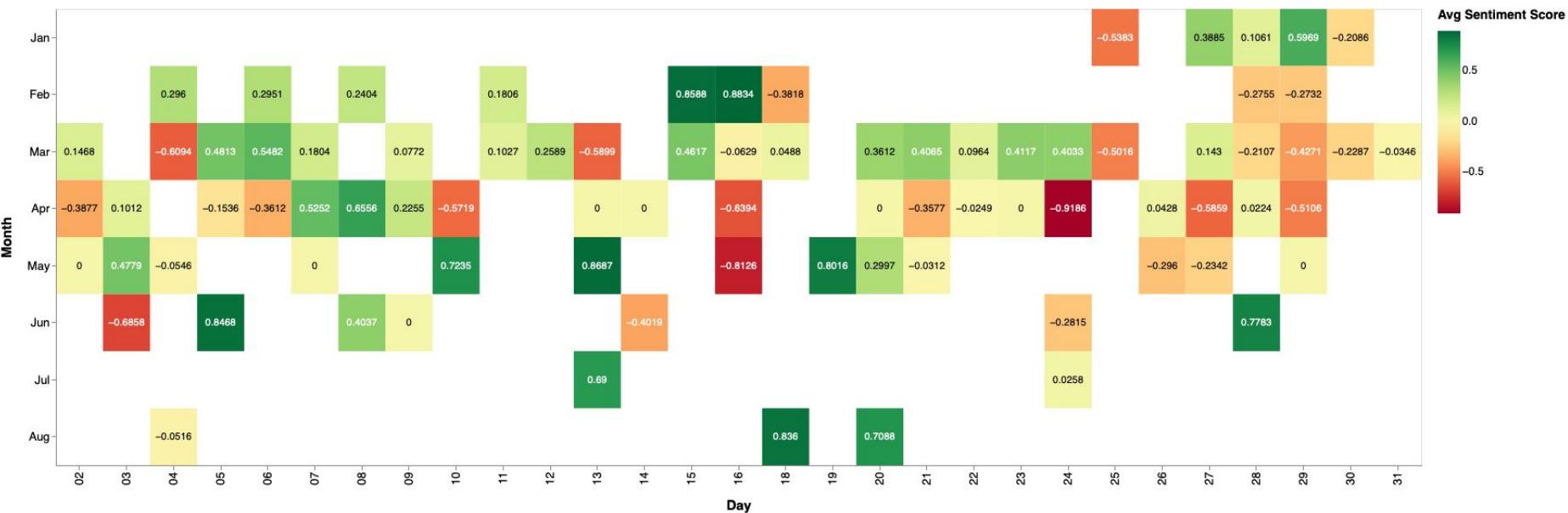
DAILY AVERAGE SENTIMENT SCORE (ALL TWEETS WITH COVID-19 RELATED HASHTAGS)

DATASET “FAKE”

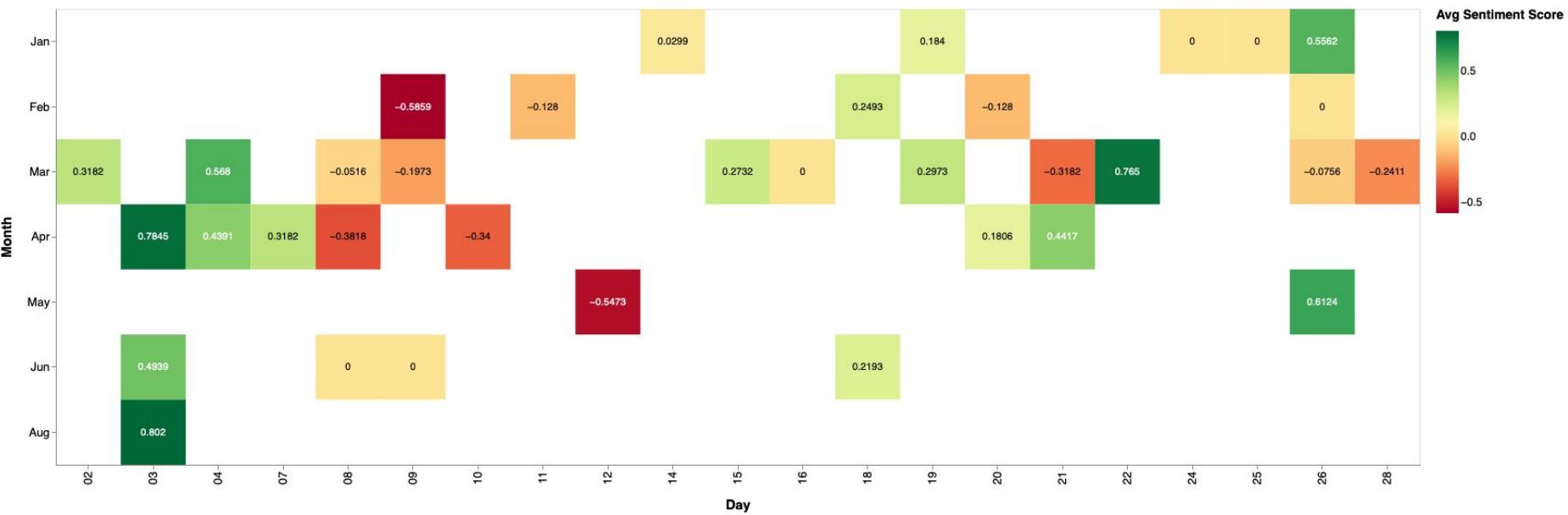


HEATMAP: DAILY AVERAGE SENTIMENT SCORE (FALSE TWEETS WITH COVID-19 RELATED HASHTAGS)

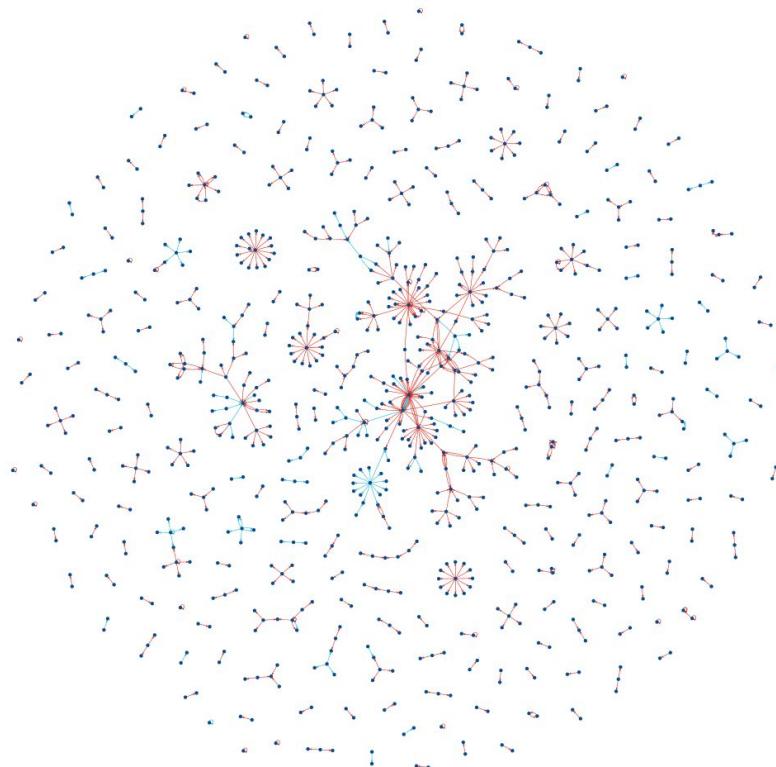
DATASET “FAKE”



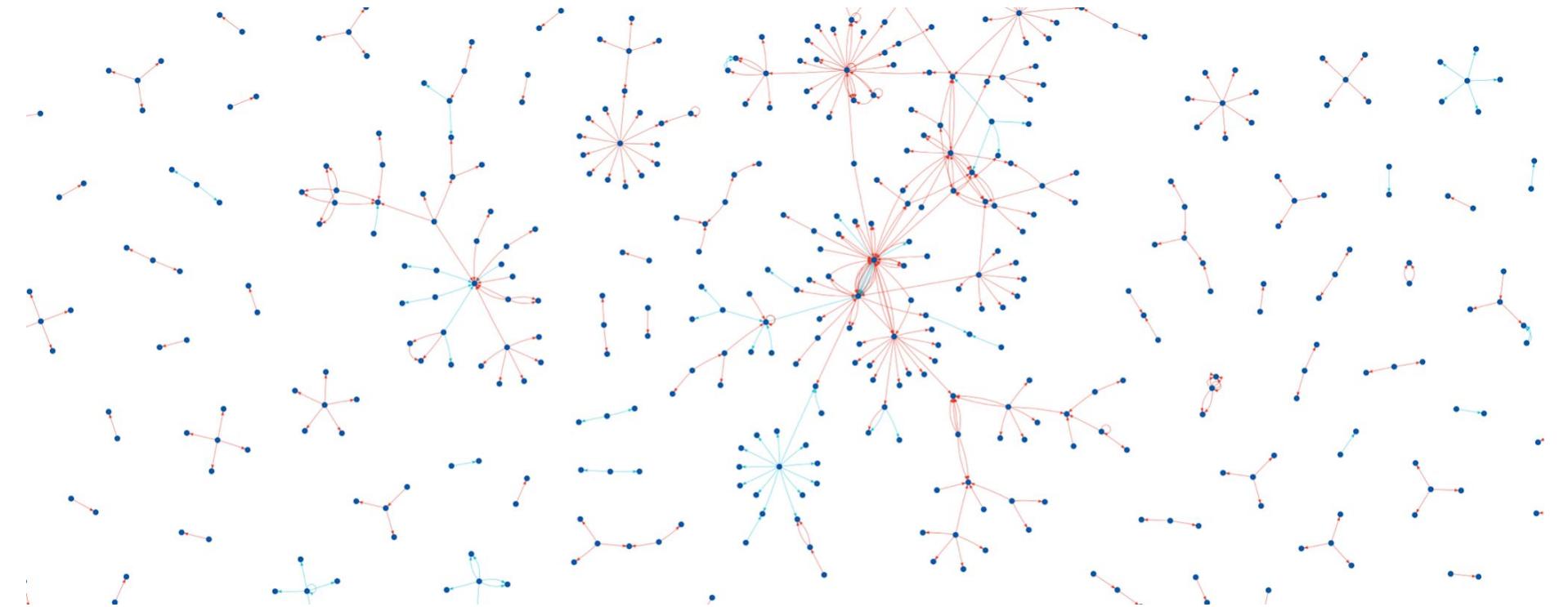
HEATMAP: DAILY AVERAGE SENTIMENT SCORE (PARTIALLY FALSE TWEETS WITH COVID-19 RELATED HASHTAGS) DATASET “FAKE”



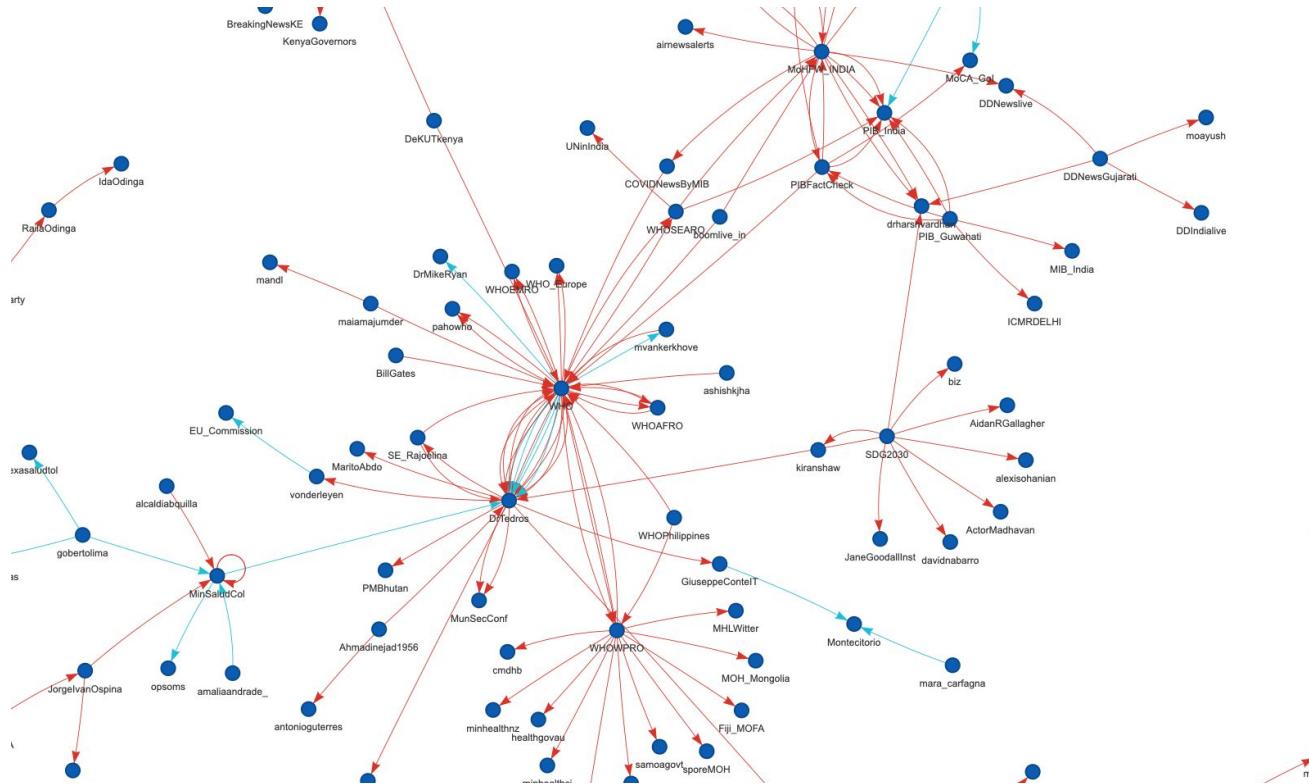
WHO MENTIONED WHO? - DATASET “FAKE”



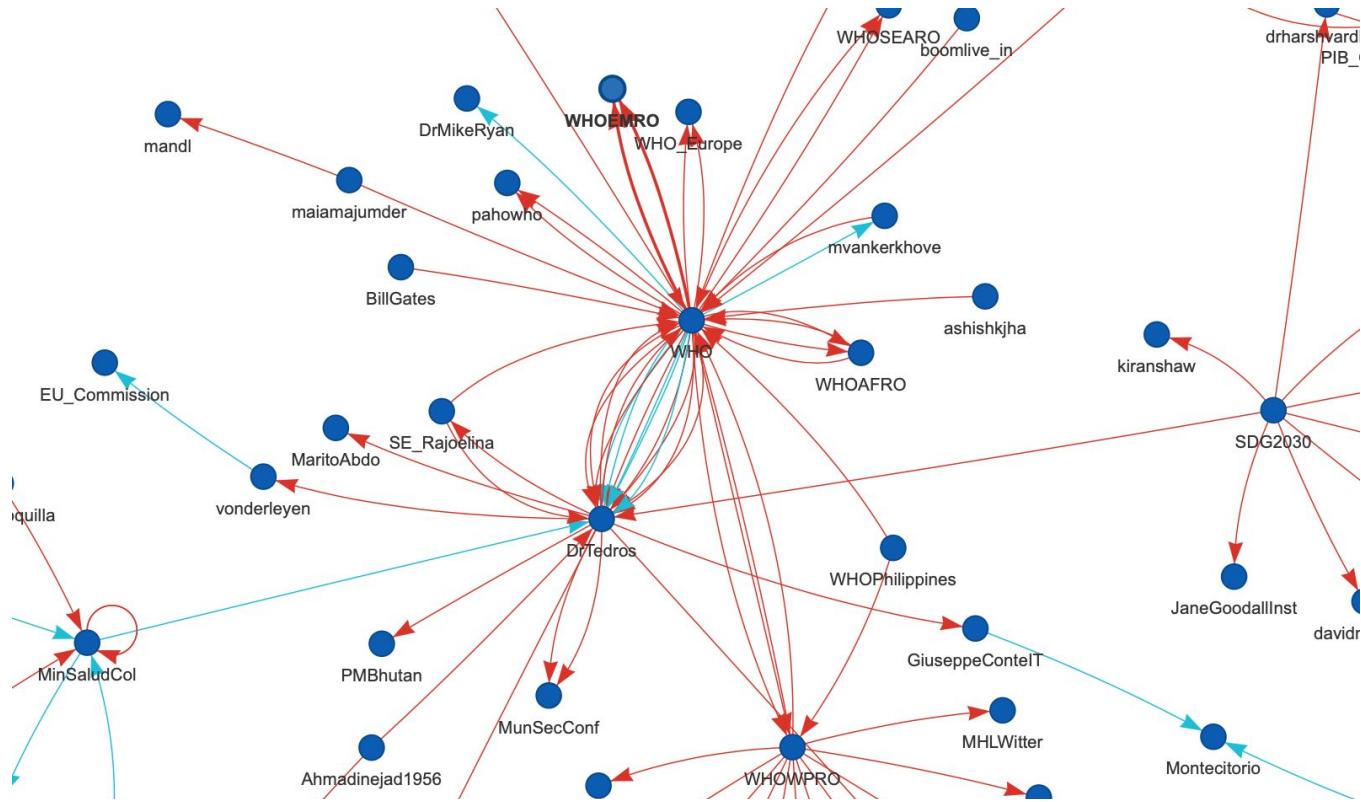
WHO MENTIONED WHO? - DATASET “FAKE”



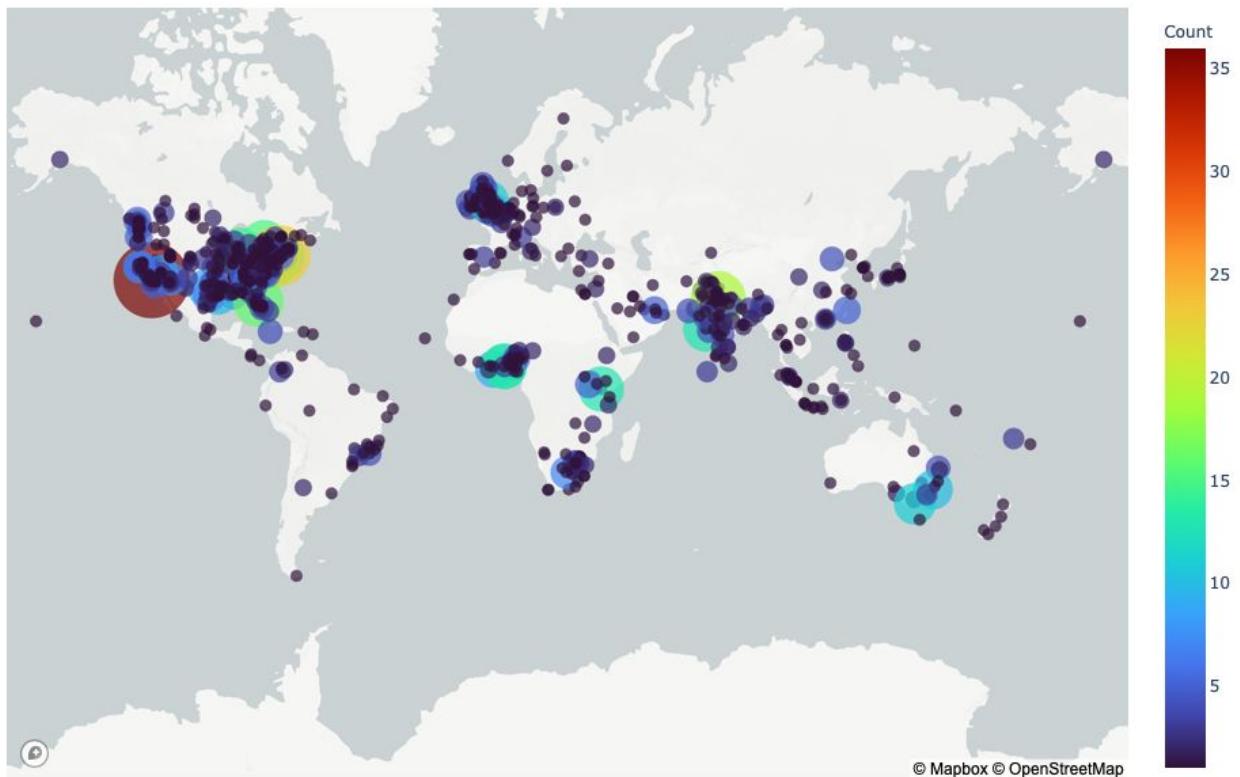
WHO MENTIONED WHO? - DATASET “FAKE”



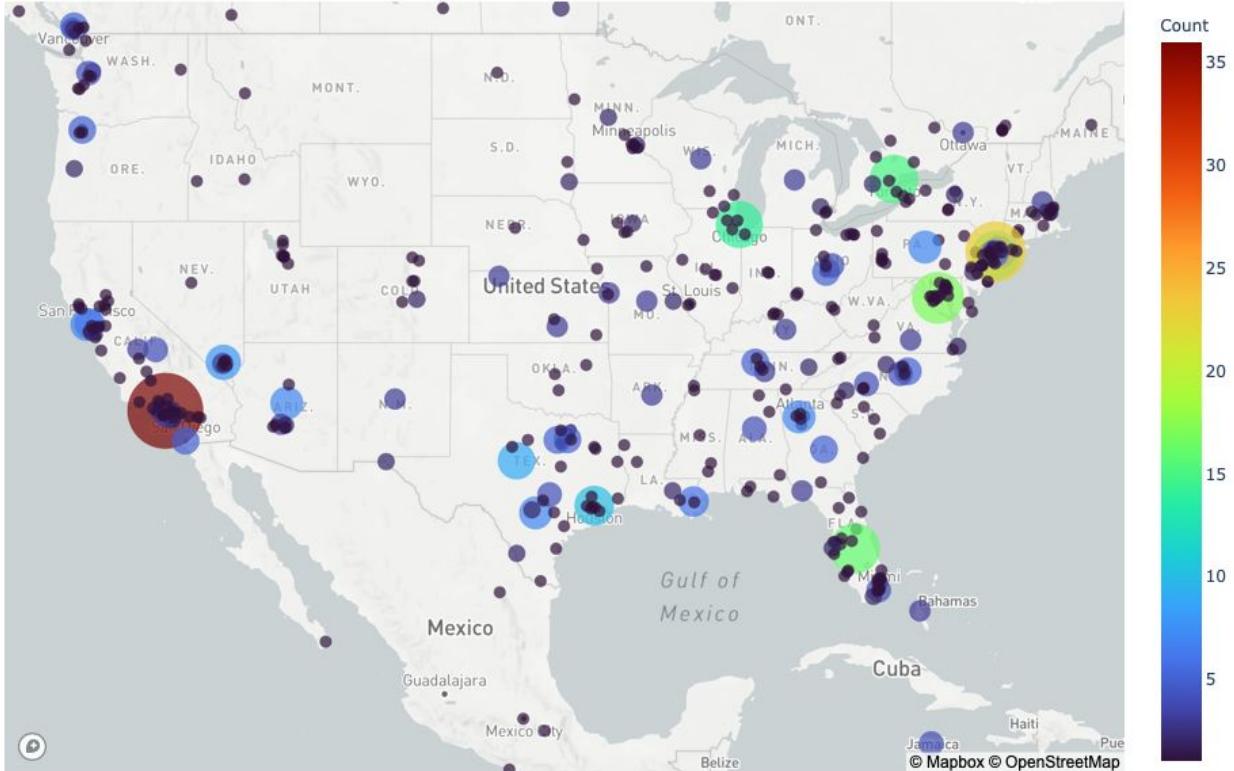
WHO MENTIONED WHO? - DATASET “FAKE”



WHERE DID PEOPLE TWEET ABOUT COVID-19? DATASET “GENERAL”

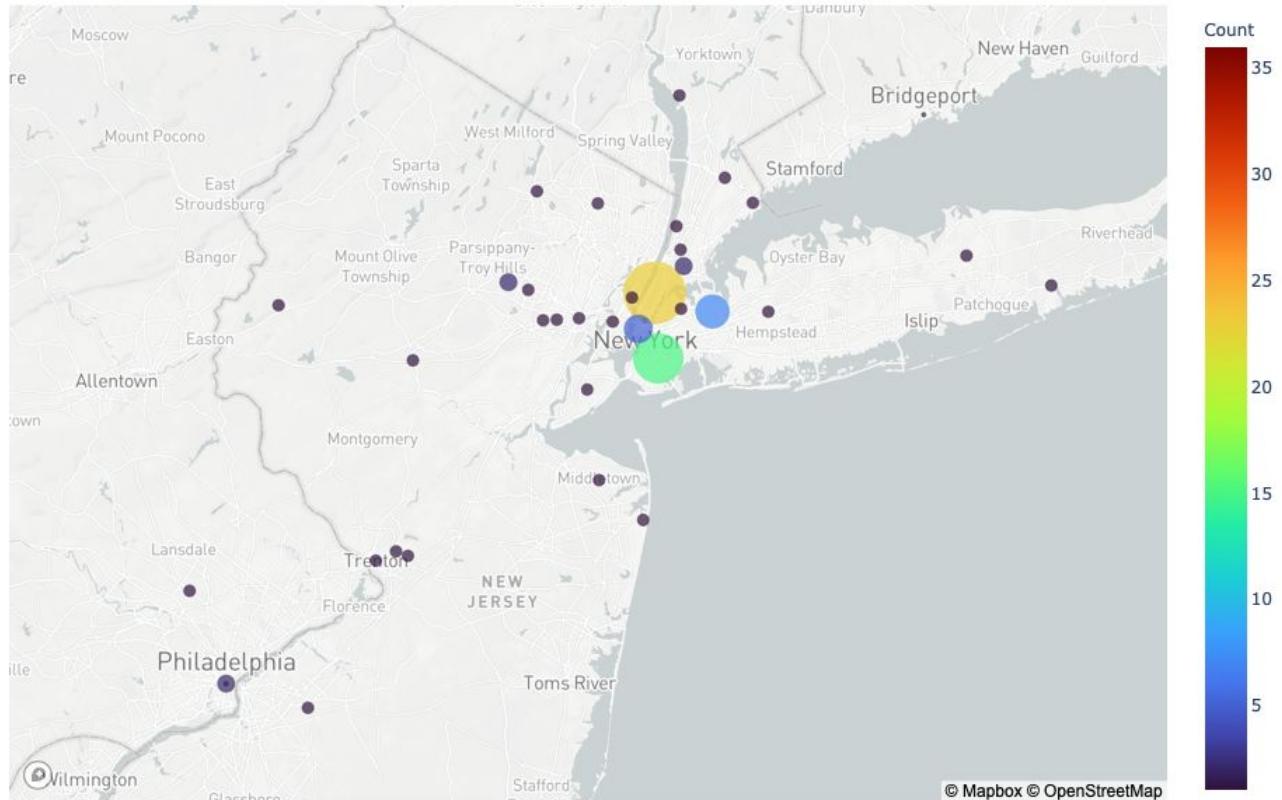


WHERE DID PEOPLE TWEET ABOUT COVID-19? DATASET “GENERAL”



WHERE DID PEOPLE TWEET ABOUT COVID-19?

DATASET “GENERAL”



A FREQUENCY COMPARISON BETWEEN THE TWO DATASETS

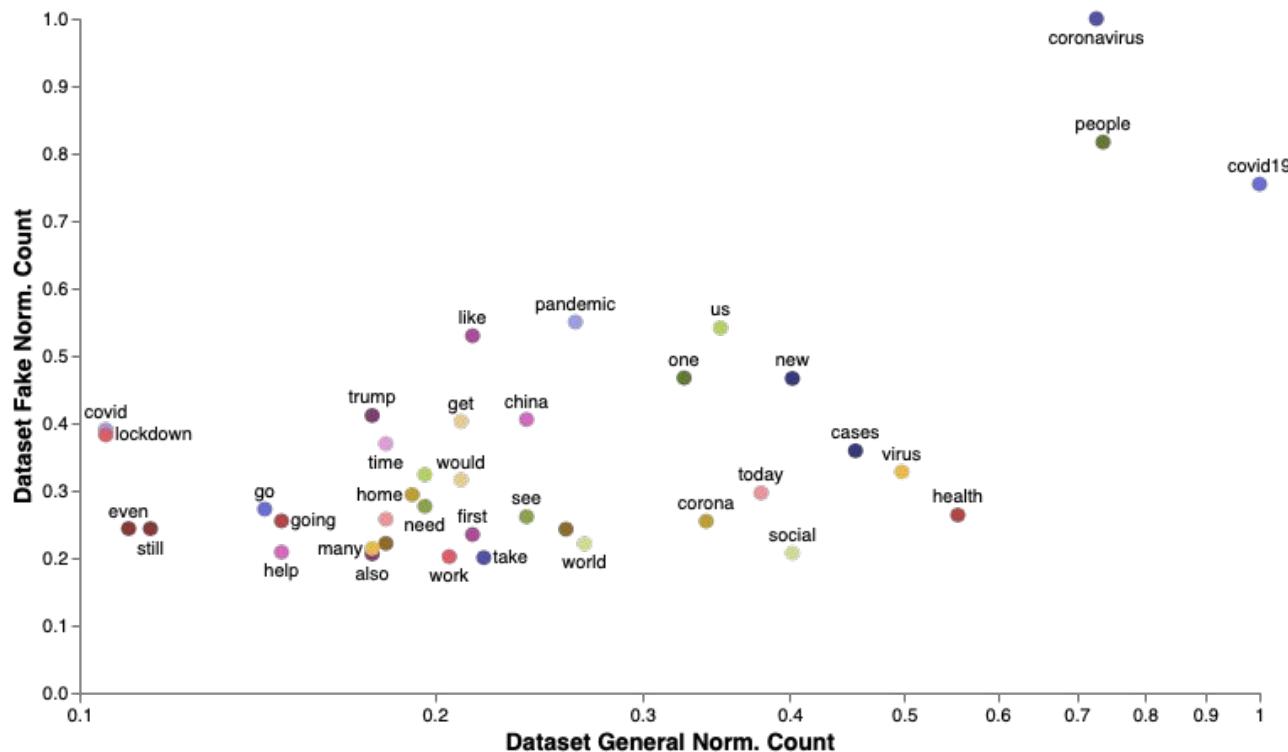
In order to compare the frequencies of emojis, hashtags, user mentions, words, bigrams, trigrams in the two datasets, we had to use the **min/max normalization** to normalize the data of each chart.

$$\frac{\text{value} - \text{min}}{\text{max} - \text{min}}$$

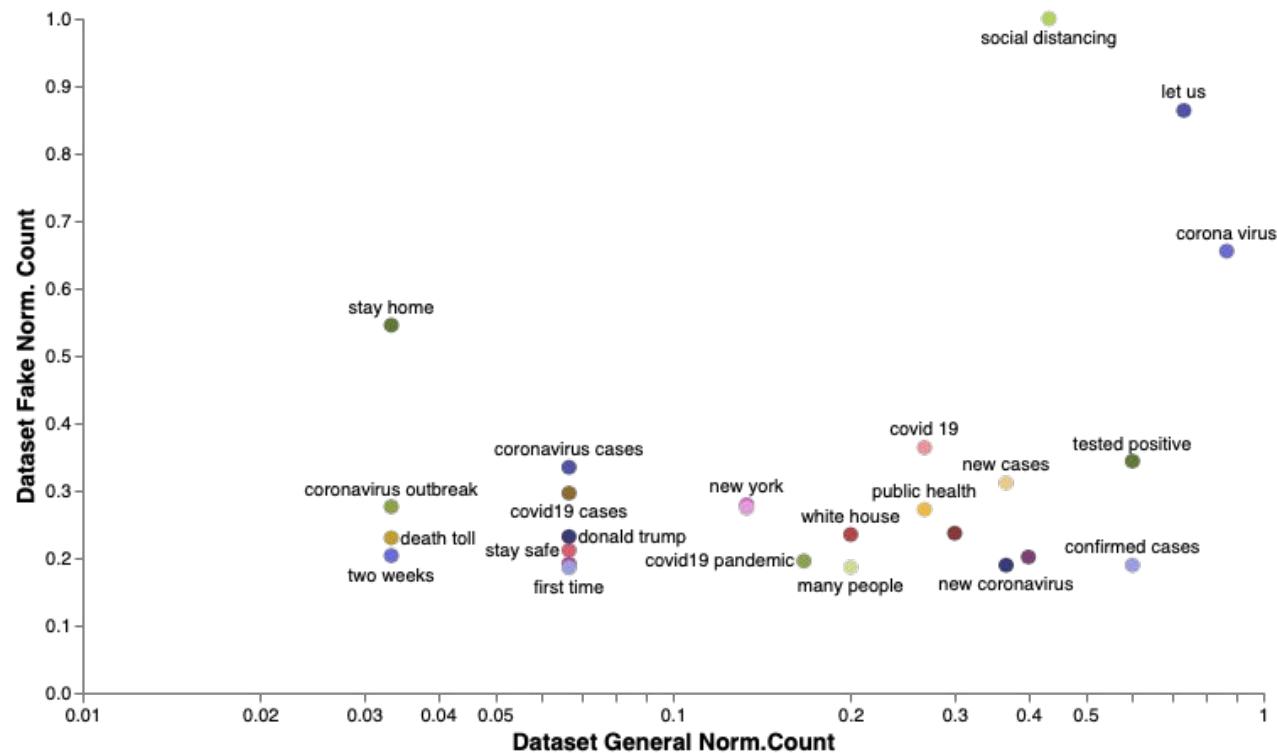
Considering a set of data:

- the **minimum value** of the set gets transformed into a **0**;
- the **maximum value** gets transformed into a **1**;
- **every other value** gets transformed into **a decimal between 0 and 1**.

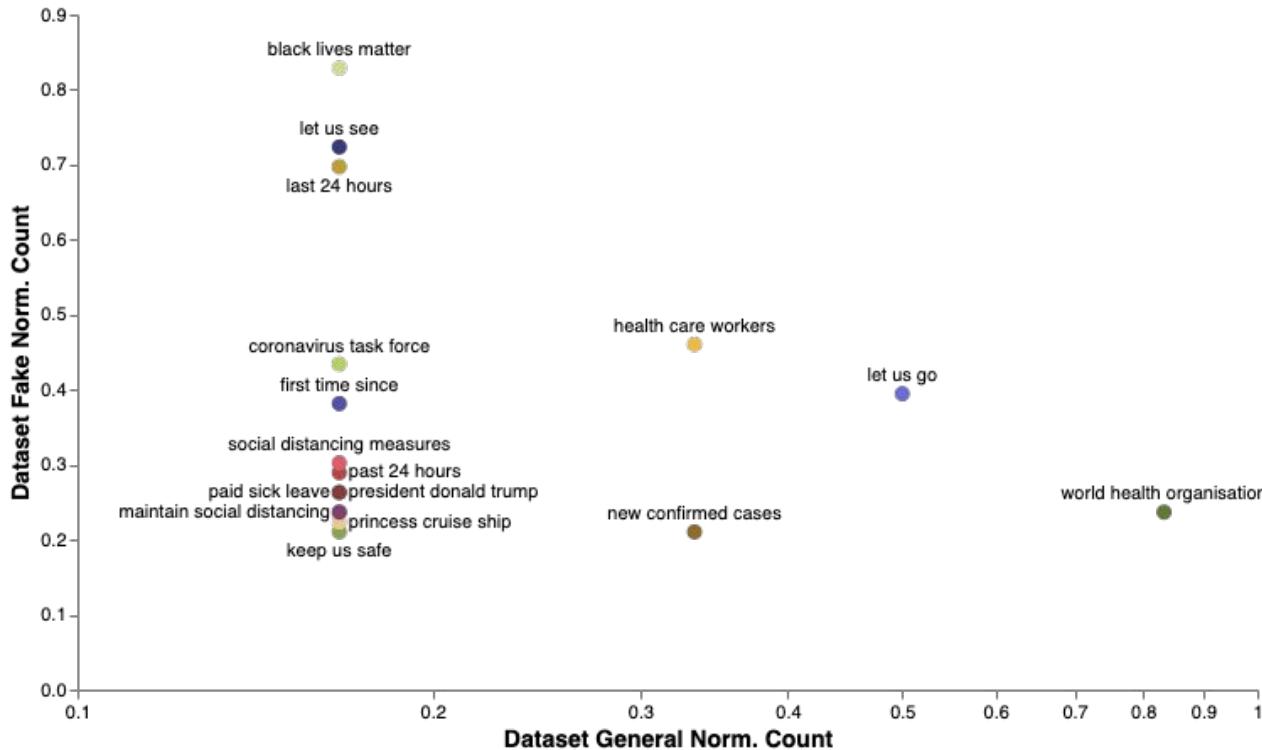
A FREQUENCY COMPARISON BETWEEN THE WORDS USED IN THE TWO DATASETS



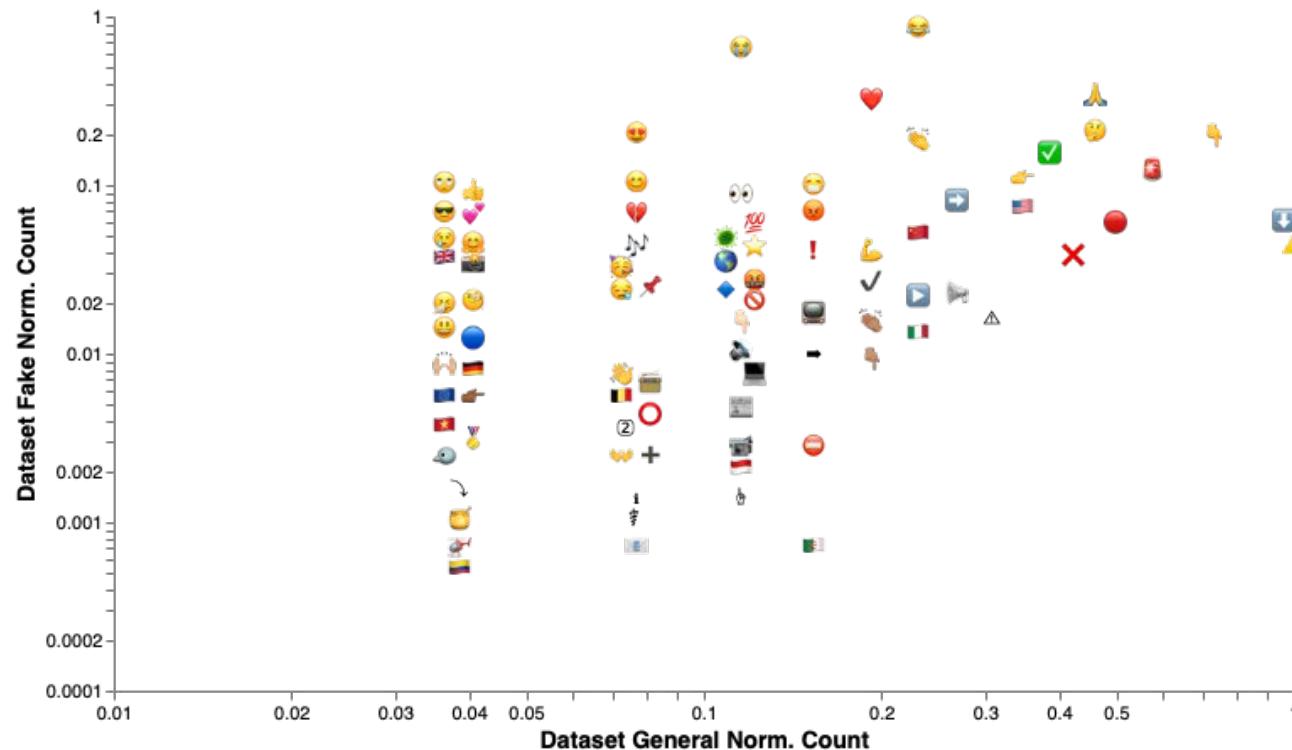
A FREQUENCY COMPARISON BETWEEN THE BIGRAMS USED IN THE TWO DATASETS



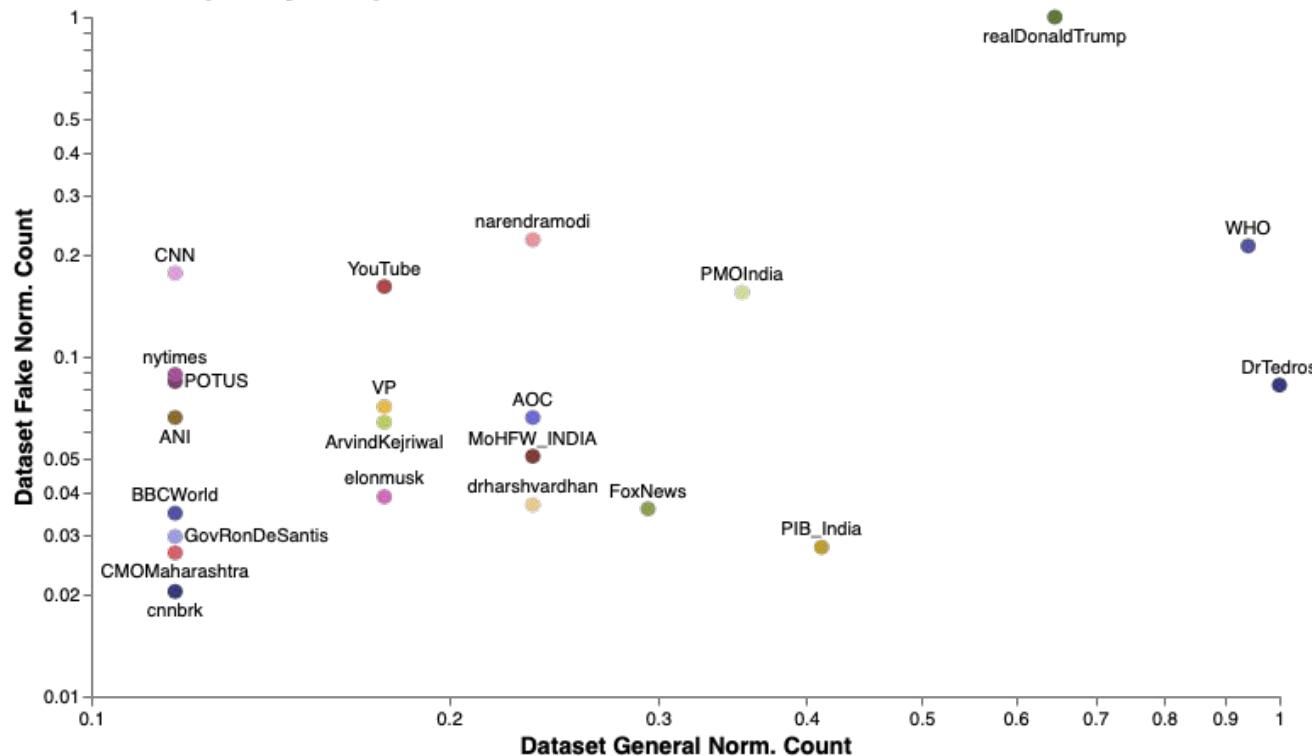
A FREQUENCY COMPARISON BETWEEN THE TRIGRAMS USED IN THE TWO DATASETS



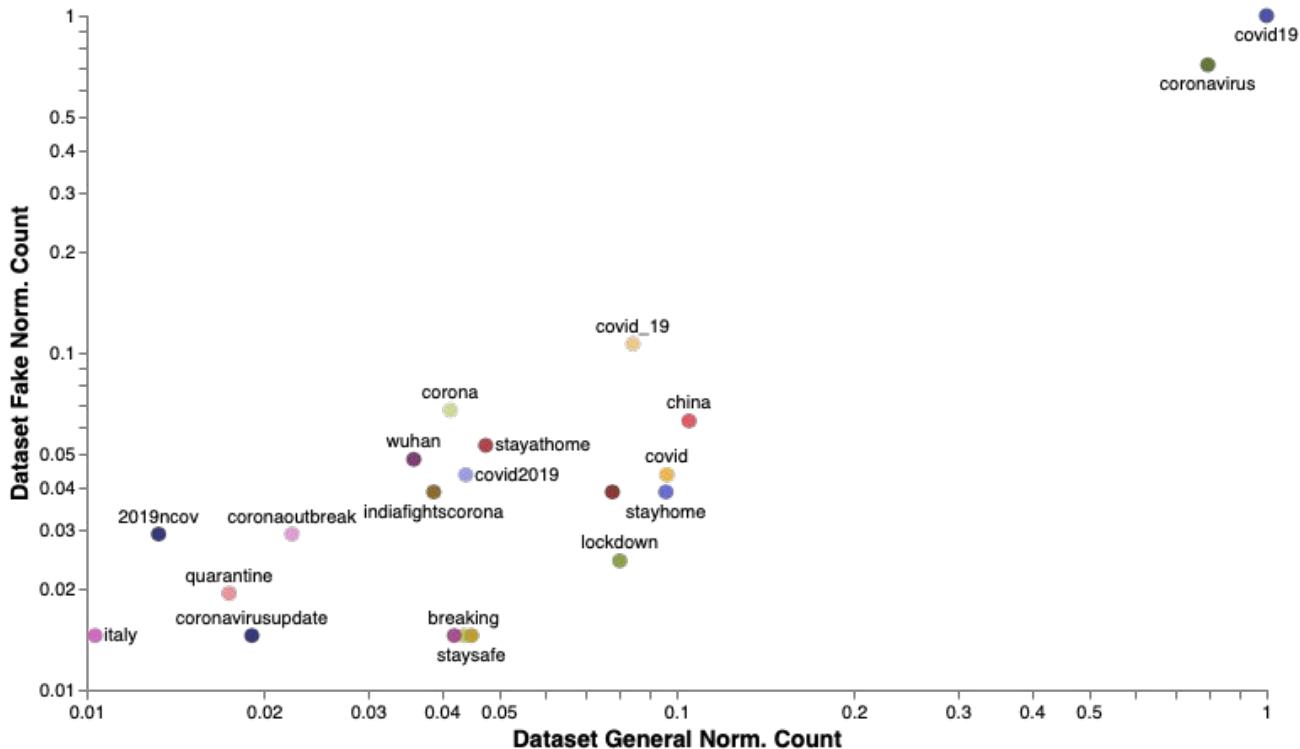
A FREQUENCY COMPARISON BETWEEN THE EMOJIS USED IN THE TWO DATASETS



A FREQUENCY COMPARISON BETWEEN THE USERNAMES MENTIONED IN THE TWO DATASETS



A FREQUENCY COMPARISON BETWEEN THE HASHTAGS USED IN THE TWO DATASETS



RESOURCES

1. S. Stieglitz, L. Dang-Xuan, **Emotions and information diffusion in social media- sentiment of microblogs and sharing behavior**, J. Manag. Inf. Syst. 29 (4) (2013) 217–248.
2. Hutto, C.J. & Gilbert, E.E. (2014). **VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text**. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.
3. Shahi G. K., Dirkson A. & Majchrzak T. A. (2021), **An exploratory study of COVID-19 misinformation on Twitter**, Online social networks and media, 100104.
4. [Paul Tol's Notes - Colour schemes and templates](#)
5. [Coloring for Colorblindness by David Nichols](#)

GRAZIE

github.com/marcopedrinazzi/twitter-covid19-vis