

# Rumour Spreading Analysis on Twitter *SydneySiege*



# Agenda

**01** Introduction

**02** The project steps

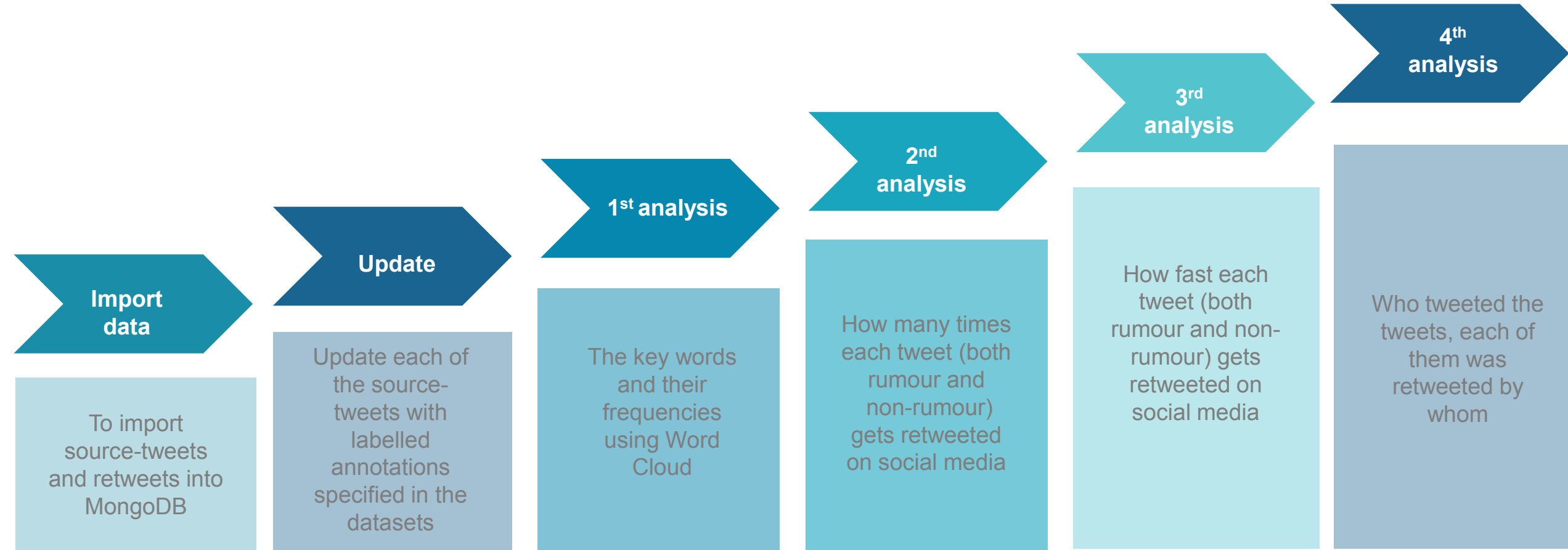
**03** results



# Rumour Spreading Analysis on Twitter

---

# The project steps



“

Where each tweet is tweeted by visualizing users' locations on the map using aforementioned tools or other visualization tools such as Google Maps API.

”

5<sup>th</sup> ANALYSIS  
EXTRA POINTS





# IMPORT DATA



```
import pymongo
import json
from pymongo import MongoClient
```

## #Connect with MongoDB Atlas

```
client = pymongo.MongoClient("mongodb://wisitsak:wp11223344@cluster0-shard-00-00.99fit.mongodb.net:27017,cluster0-shard-00-01.99fit.mongodb.net:27017,cluster0-shard-00-02.99fit.mongodb.net:27017/Bigdata?ssl=true&replicaSet=atlas-uuej29-shard-0&authSource=admin&retryWrites=true&w=majority")
db = client.sydneyseige
```

## # create Collection on

```
db.create_collection('source-tweets')
db.create_collection('retweets')
```

```
collection = db. source-tweets
collection2 = db. retweets
collection3 = db. annotation
```

101001101001000010101  
0011110111011011011010  
101000011100101011001  
010100111010100010101  
0001011010110110110100  
010101110001010100010  
1000101110101100010011  
010011010010000101010  
0111101110110110110101  
010000111001010110010  
101001110101000101010  
0010110101101101101001





# IMPORT DATA



```
path = "*/source-tweets/*.json"
for filename in glob.glob(path, recursive=True):
    with open(filename) as f:
        file_data = json.load(f)
        collection.insert_one(file_data)

path2 = "*/retweets.json"
for filename in glob.glob(path2, recursive=True):
    with open(filename) as f:
        for line in f:
            file_data = json.loads(line)
            collection2.insert_one(file_data)
```

101001101001000010101  
0011110111011011011010  
101000011100101011001  
010100111010100010101  
0001011010110110110100  
010101110001010100010  
1000101110101100010011  
010011010010000101010  
0111101110110110110101  
010000111001010110010  
101001110101000101010  
0010110101101101101001



# Update python™

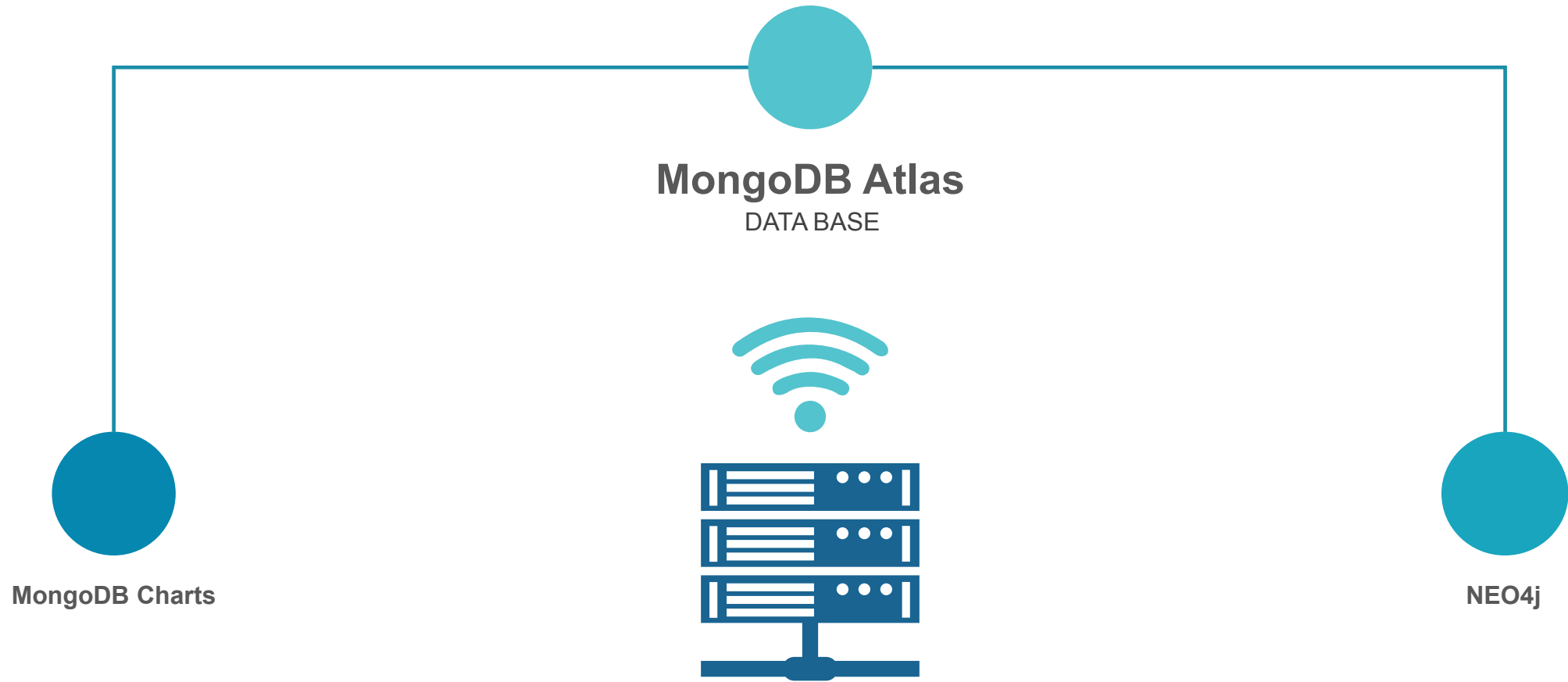
```
path3 = "*/annotation.json"
for filename in glob.glob(path3, recursive=True):
    with open(filename) as f:
        file_data = json.load(f)
        collection.update_many(
{"id":int(filename.strip("\\annotation.json")),{"$set":{"misinformation":int(file_
data["misinformation"])}}}

fake_word = []
for x in collection.find({"misinformation":1}):
    word = re.findall(r"[w]++", x['text'])
    for y in word:
        fake_word.append(y.lower())
print(fake_word)
```





# ANALYSIS



# 1<sup>st</sup> ANALYSIS

## MongoDB Charts

The key words and their frequencies  
using Word Cloud

Chart Type

Text

Table

100

Number

T

Top Item

W

Word Cloud

[^ Show fewer charts](#)

Encode

Filter

Customize

Text

A text

ARRAY REDUCTIONS

hashtags

UNWIND ARRAY

Binning

SORT BY

VALUE

Limit Results

Size

# misinformation

AGGREGATE

COUNT BY VALUE

Color

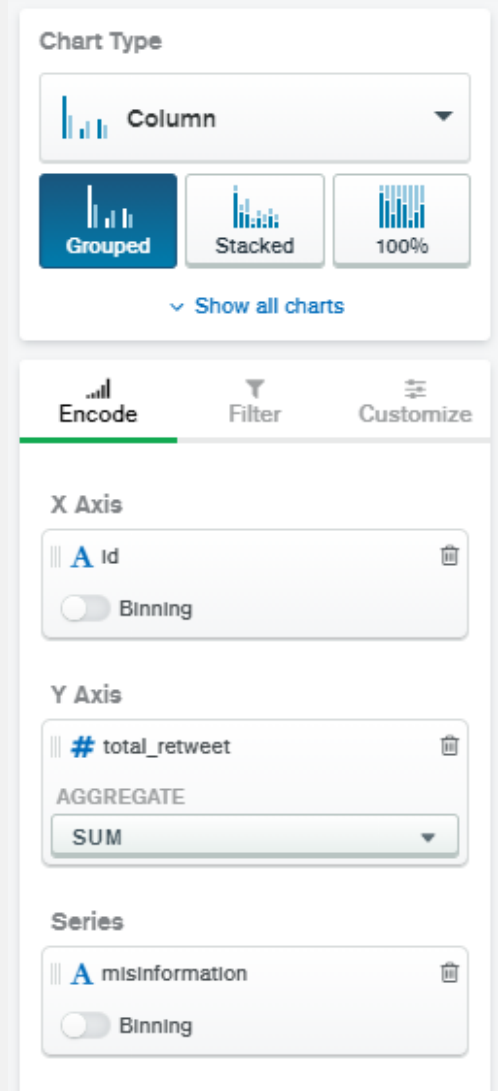
+ category

# 2<sup>nd</sup> ANALYSIS

## MongoDB Charts

How many times each tweet (both  
rumour and non-rumour) gets  
retweeted on social media

```
{
  "$lookup": {
    "from": "retweets",
    "localField": "id",
    "foreignField": "retweeted_status.id",
    "as": "retweets"
  },
  {
    "$project": {
      "id": 1,
      "text": 1,
      "city": 1,
      "misinformation": 1,
      "created_at": 1,
      "_id": 0,
      "total_retweet": {
        "$size": "$retweets"
      }
    }
  }
}
```





# 3<sup>rd</sup> ANALYSIS


## MongoDB Charts


How fast each tweet (both rumour and non-rumour) gets retweeted on social media

Chart Type

 Area

 Discrete

 100%

 Continuous

Encode

Filter

Customize

X Axis

A Id

☐ Binning

Y Axis

# total\_retweet

AGGREGATE

MIN

# total\_retweet

AGGREGATE

MAX

# total\_retweet

AGGREGATE

MEAN

+ aggregation

Series

+ category

# 4<sup>th</sup> ANALYSIS

NEO4j

Who tweeted the tweets, each of them was retweeted by whom by converting json objects

```
[19]: from py2neo import Graph
```

```
[33]: neo4j = Graph('bolt://localhost:7687', auth=('neo4j', '11223344'))
```



```
[24]: tweets = db.sourcetweets
      retweets = db.retweets
```


```
[34]: # Create Tweet Node
      for tweet in collection.find({}):
          neo4j.run('CREATE (:Tweet {id: $id, user_name: $user_name, text: $text, misinformation: $misinformation})',
                    id = tweet['id'], user_name = tweet['user']['name'], text = tweet['text'], misinformation = tweet['misinformation'])
```


```
[35]: # Create Retweet Node
      for retweet in collection2.find({}):
          neo4j.run('MATCH (t:Tweet {id: $tweet_id}) \
                    CREATE (r:Retweet{id: $id, user_name: $user_name, text: $text})-[:RETWEETKUB]->(t) RETURN r',
                    tweet_id = retweet['retweeted_status']['id'], id = retweet['id'], user_name = retweet['user']['name'], text = retweet['text'])
```


# 5<sup>th</sup> ANALYSIS


## MongoDB Charts


Chart Type


 Geospatial


 Choropleth

 Scatter


 Heatmap


 Encode

 Filter

 Customize

Coordinates


 coordinates




Color

+ category

Size

 # favorite\_count



AGGREGATE

MEAN

Tooltip Detail

+ category





**RESULTS**



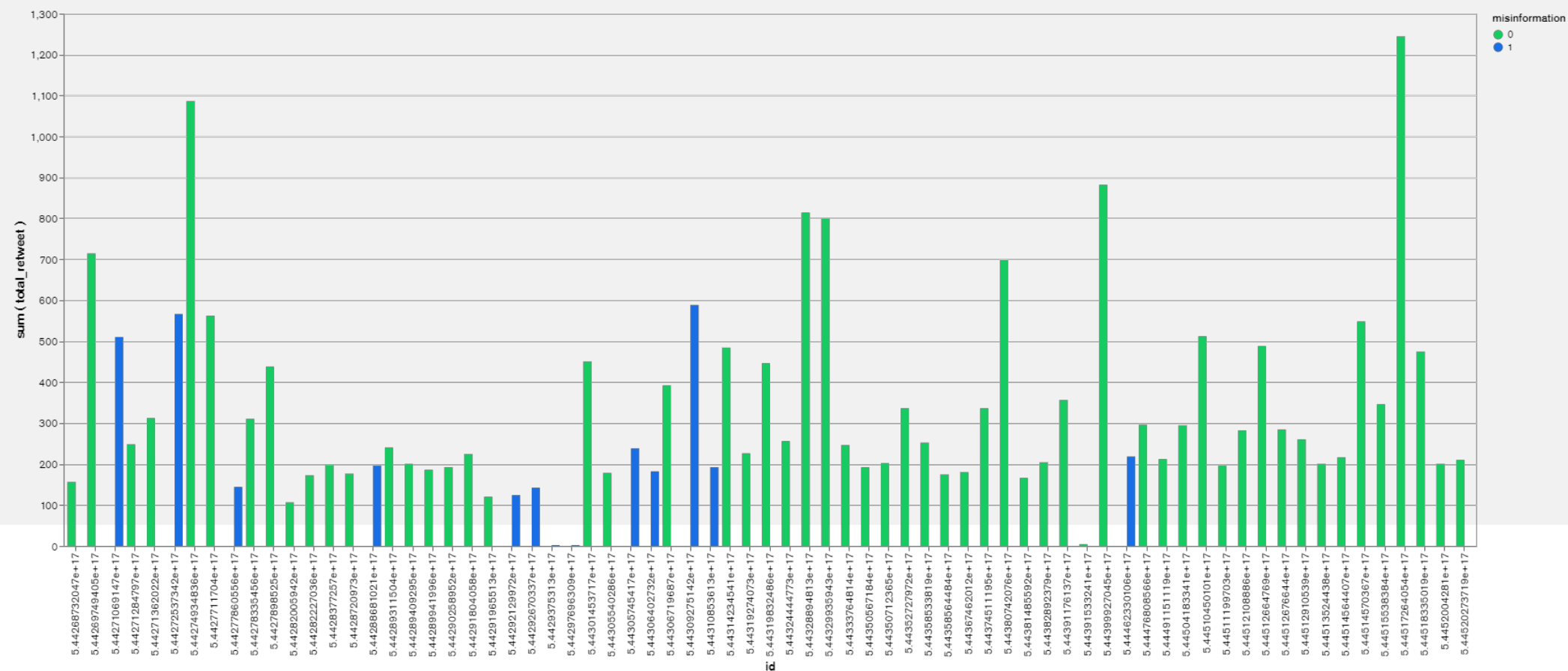
# 1<sup>st</sup> ANALYSIS

A word cloud visualization of tweets related to the Sydney Siege. The words are arranged in a circular pattern, with 'Sydney Siege' and 'Sydney' being the most prominent. Other visible words include 'BREAKING', 'SydneySiege', 'ISIS', 'MartinPlace', 'Lindt', 'Abbott', '9News', 'pmharper', 'MartinPl', 'Iran', 'sydneyseige', 'SYDNEYSIEGE', 'Lindt', 'ISISAttacks', 'SydneySiege', '7News', 'PrayForSydneyHostages', 'hostages', 'MartinPlaceSiege', 'Lindt', 'ABCNews24', 'DramaAlert', 'PrayForSydney', and 'SydneySiege'.

# Times each tweet

# RESULTS

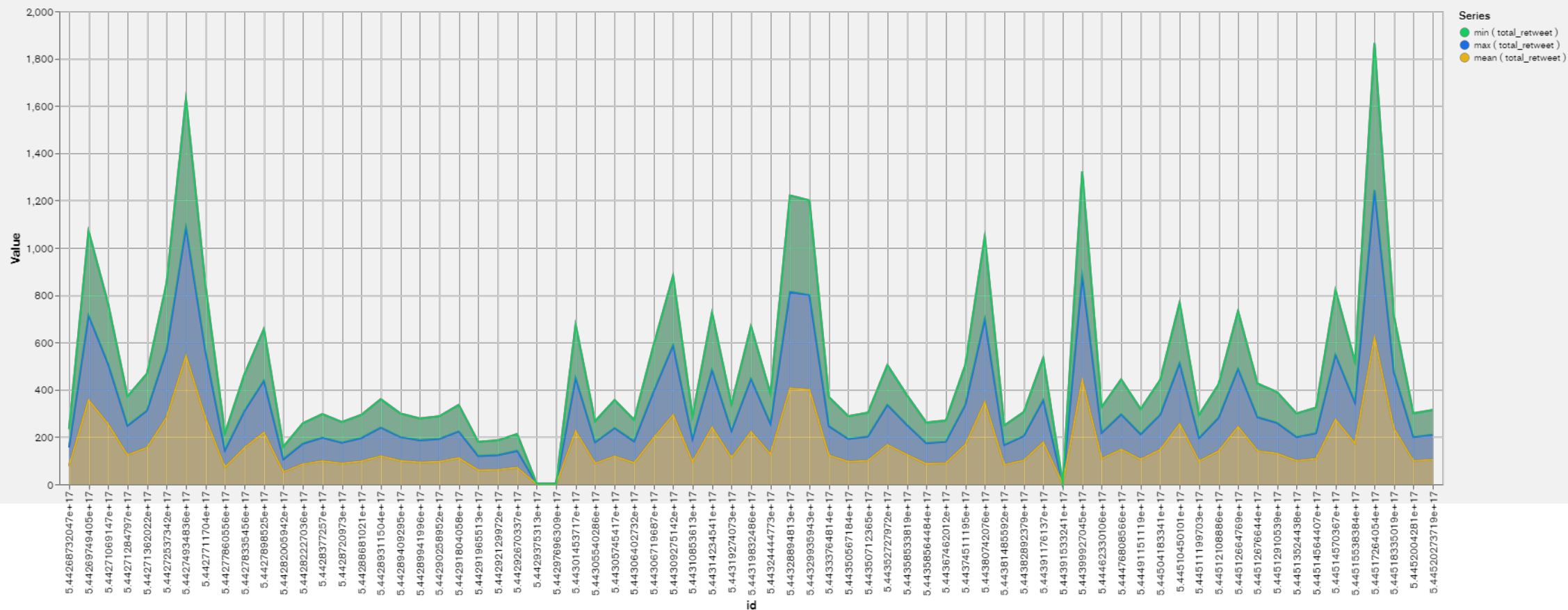
## 2<sup>nd</sup> ANALYSIS



# Times each tweet

## RESULTS

### 3<sup>rd</sup> ANALYSIS

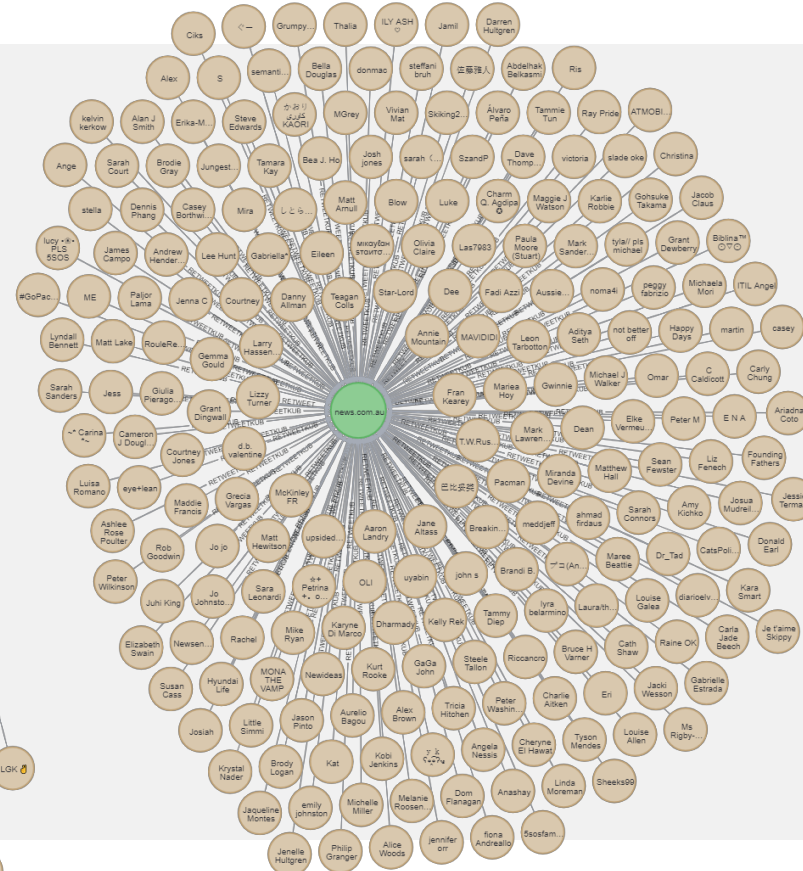


## 4<sup>th</sup> ANALYSIS

The graph visualization shows a central node labeled "Sky News Australia" (green circle) connected to a large number of peripheral nodes (tan circles). The nodes are arranged in a circular pattern around the center, with lines of varying thickness connecting them. The nodes are labeled with names, some of which are partially obscured or cut off. The graph is dense, with many overlapping lines and nodes.

Visible names include:

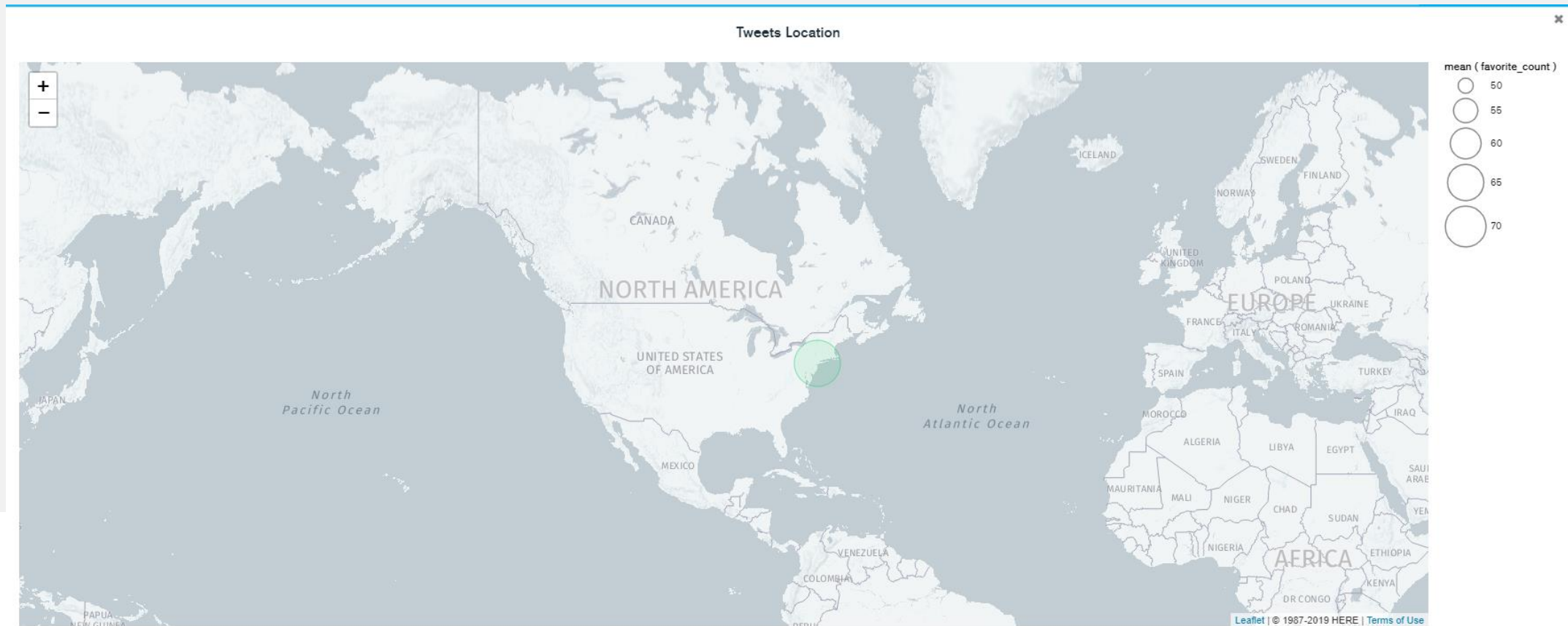
- Pete Mitt
- Usha M Rodriguez
- Catherine
- Michelle Roberts
- A.S. Dulac
- MyFosh
- Betty T
- Chris Jones
- Sarah Wiley
- Daisy Farmer
- James Ian Rogers
- Earle Court History
- Matthew Hewson
- Daniel Stevens
- Nathan Allen
- Champa
- Kimmy
- Kyle the worst
- Jake Green
- Elizabeth J Horns
- Claudius Holler
- Natalie Dick
- Julian Carter
- Ben Shute
- Anna Arizono
- Vince K
- Robert
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf
- Nina Yanne James
- Hedding
- leon
- tylviol
- Kendall Forbes
- Mareike
- rochshot
- Victor Hugo
- Jamal
- Jericho Pueblo
- Adam Hanna
- Steph
- Trudy Monty
- Aaron Cornely
- Jenna Good
- Carol South
- Kees de Graaf</



# Tweeted locations

# RESULTS

## 5<sup>th</sup> ANALYSIS







THANK YOU

