SGCI

OUR MISSION:

MyGeoHub

Collecting data from people in Atlanta, Georgia area that are tweeting about the current weather conditions. With this data we will create a Word Cloud that will display the most used words dealing with the weather.

#HackAttacks

Hackers:

DONOVAN DOSS

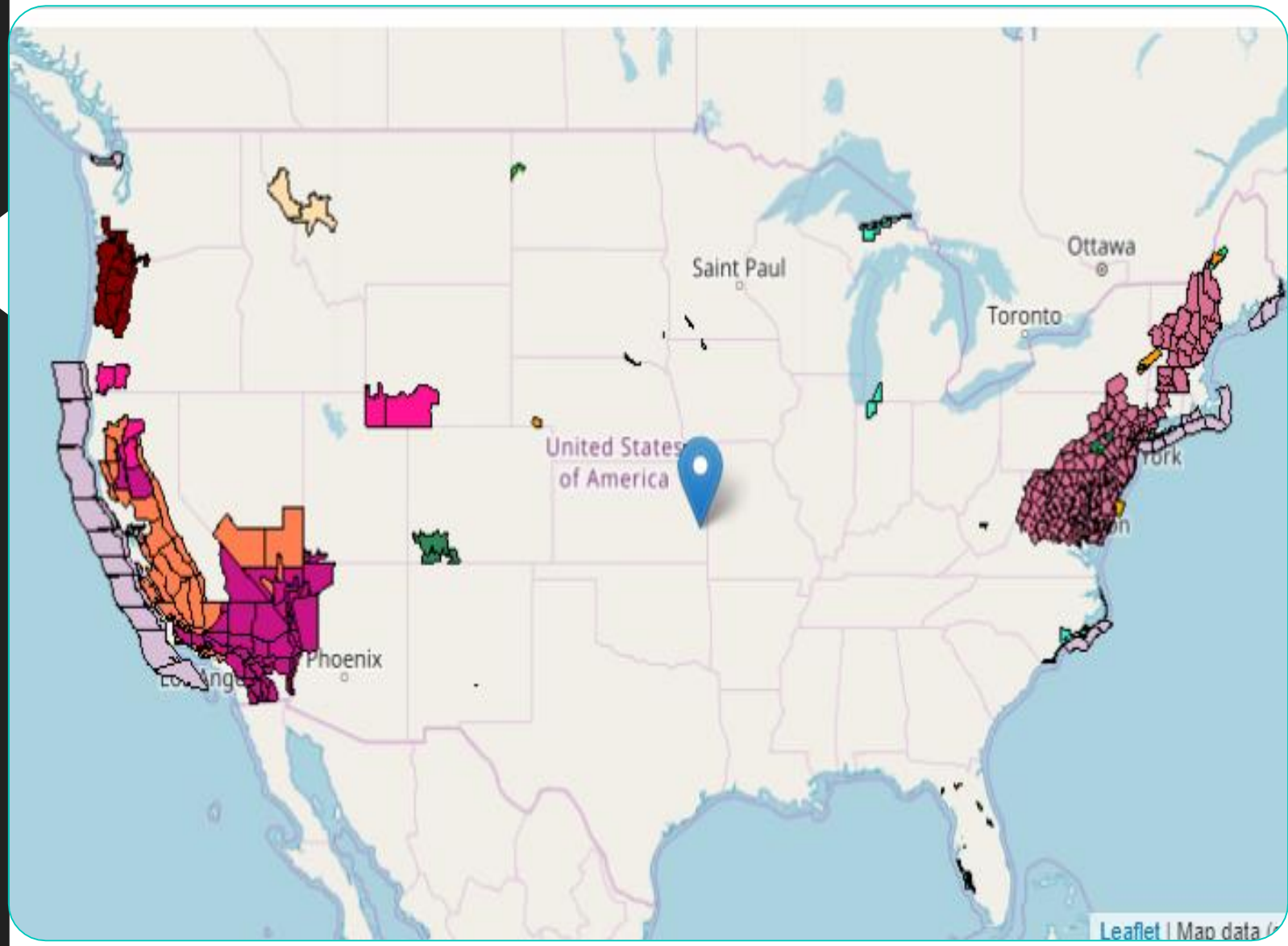
SOLOMON GEBRHANA

ALEXIS MARTIN

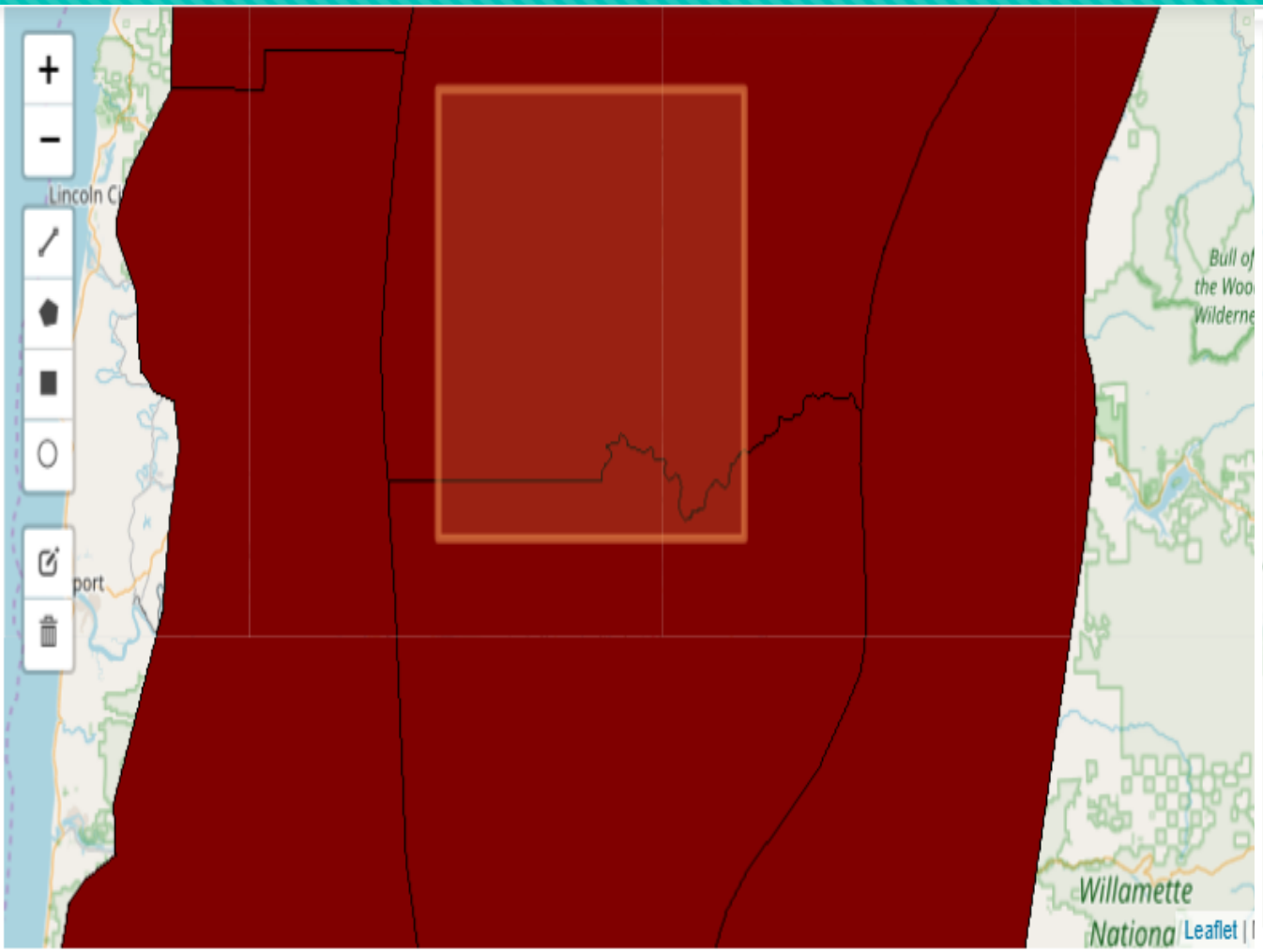
Steps To Achieve

Step 1:

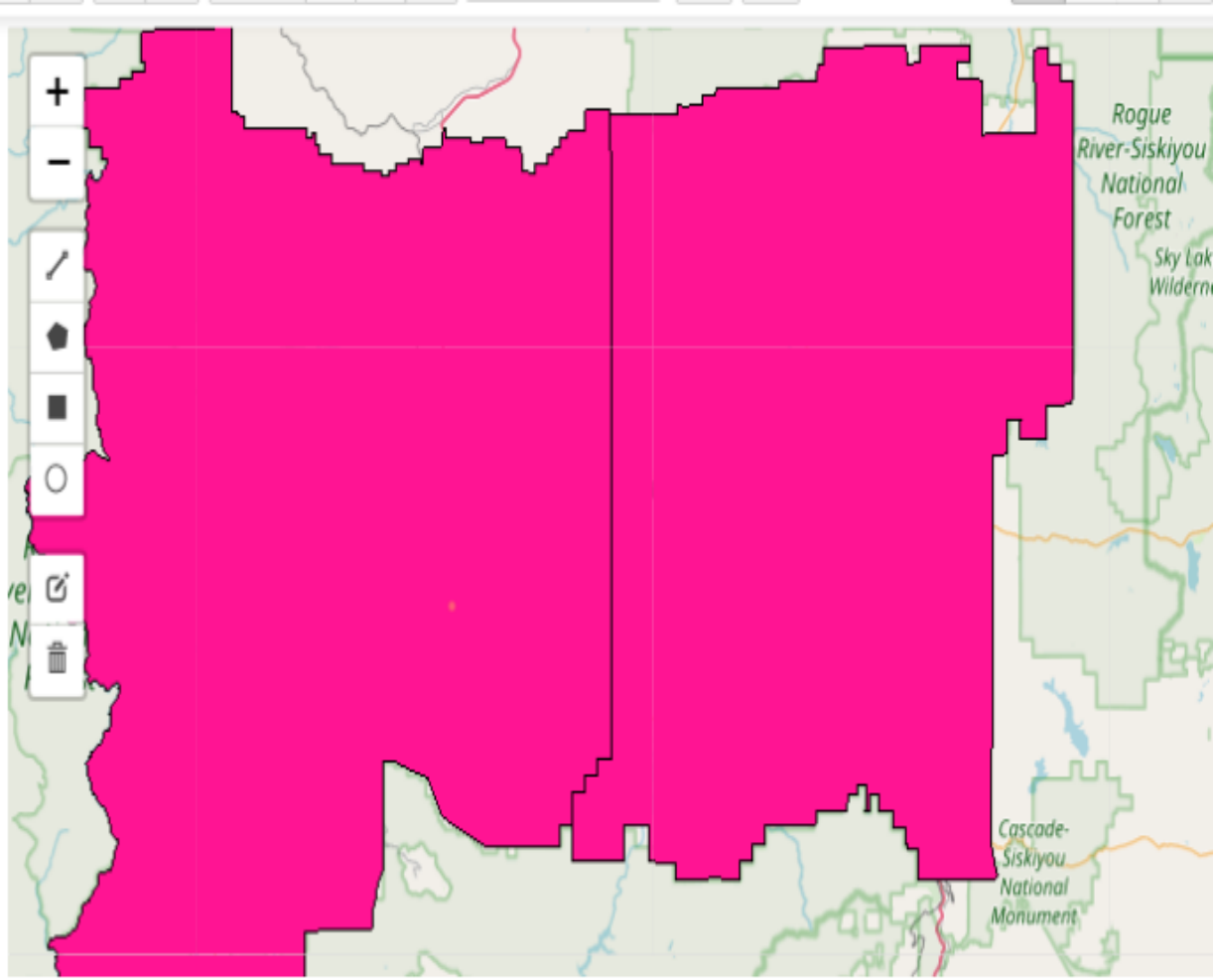
With the information that was given to us, we used XML to parse and extract the prod_type data, which gave us conditions such as; Flood Warning, Severe Thunderstorm, Heat Advisory, etc. With this information, we drew a rectangle on the affected areas to figure out the weather conditions.



Data Given

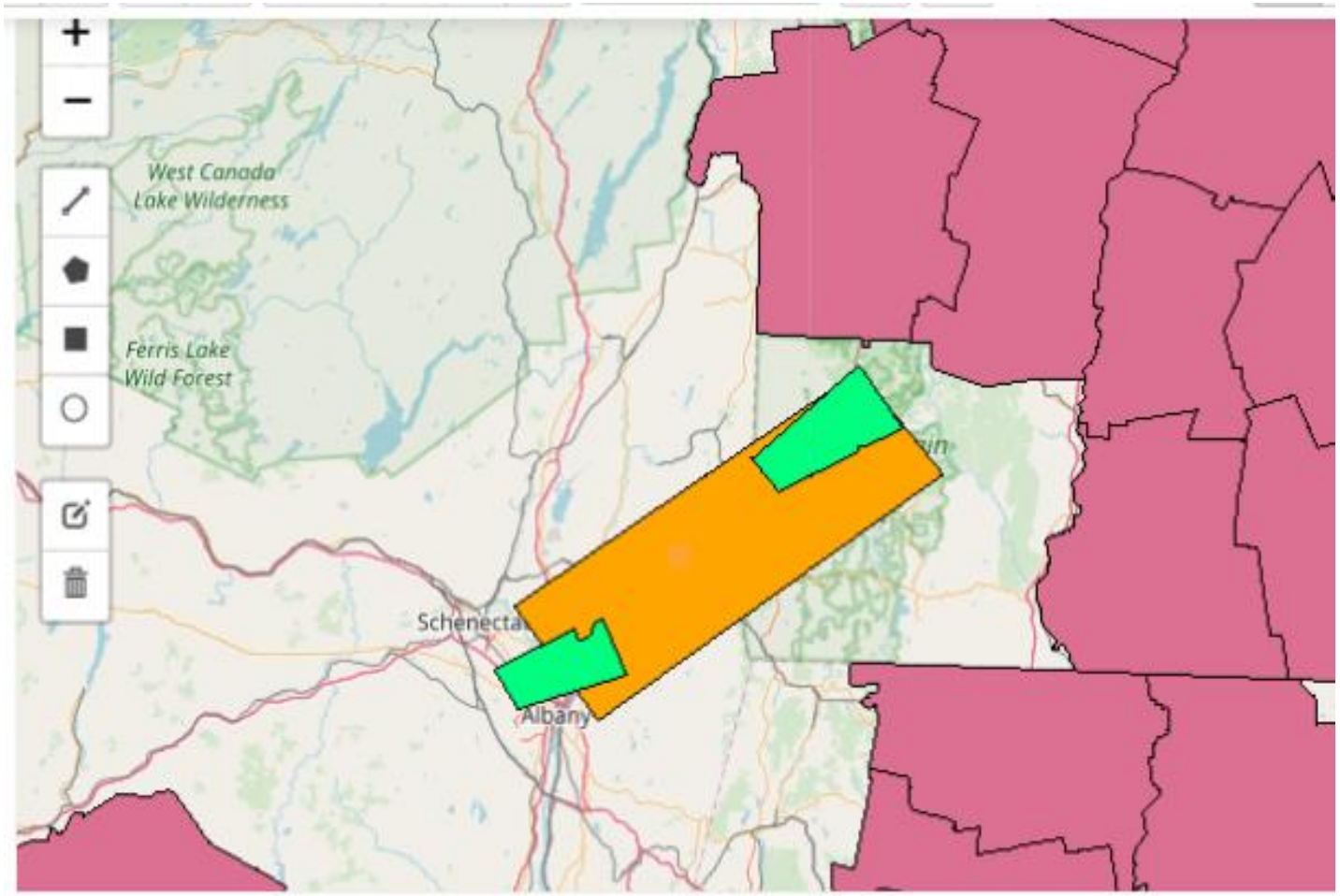


Excessive Heat Watch
Excessive Heat Watch



Excessive Heat Watch
Excessive Heat Watch
Red Flag Warning

Data Given



Beach Hazards Statement
Severe Thunderstorm Warning
Severe Thunderstorm Watch

Step 2

We imported and downloaded NLTK, Natural Language ToolKit. This allows us to tokenize a string. Tokenize means to separate every word used in that string. For example, the phrase, "Storm in Atlanta" turns into "Storm", "in", "Atlanta".

```
from ipyleaflet import *
import requests
import nltk

# import xml.etree.ElementTree as ET
# tree = ET.parse('country_data.xml')
# root = tree.getroot()

wfs_url = 'https://idpgis.ncep.noaa.gov/arcgis/services/NWS_Forecasts_Guidance_Warnings/watch_warn_adv/

m = Map(center=(39.82, -98.58), zoom=4, basemap=basemaps.OpenStreetMap.Mapnik)
wms = WMSLayer(url='http://idpgis.ncep.noaa.gov/arcgis/services/NWS_Forecasts_Guidance_Warnings/watch_w
               layers='0',
               format='image/png',
               transparent=True)
m.add_layer(wms)
mymark = Marker(location=[38,-95])
m.add_layer(mymark)
dc = DrawControl()
dc.rectangle = {
    "shapeOptions": {
        "fillColor": "#fca45d",
        "color": "#fca45d",
        "transparent": True
    }
}
```

Step 3

We imported tweepy which allowed us to filter the information. We used myStream.filter, which points out specific tweets with keywords that they will like to know about.

```
import nltk
nltk.download('punkt')
sentence = "Heat Advisory Storm  "
tokens=nltk.word_tokenize(sentence)
tokens
```

```
[nltk_data] Downloading package punkt to
[nltk_data]     /home/mygeohub/mygeohub02/nltk_data...
[nltk_data]   Package punkt is already up-to-date!

['Heat', 'Advisory', 'Storm']
```

```
import tweepy
from tweepy import*
consumer_key = 'NqYmN1yKycA0iXNggwtee5sze'
consumer_secret = 'NKG9FqJ1SYGeuC6IIitQTmjpPjFZ2bJPn1v9aTjOph7xnD8B6v'
access_token_key = '756563398284935170-Ks0kvwIEVc14v8lkyWN4tgHh8DHRLuY'
access_token_secret = 'Dg6E8vSbHqpjxKzzP4fF1EOBRUiEUkQHdxojnO99SuCAF'

auth = OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token_key, access_token_secret)
api = API(auth)


#override tweepy.StreamListener to add logic to on_status
class MyStreamListener(tweepy.StreamListener):
```

Step 4

We used hashtag dictionaries. When running hashtag_dict() this gives you the list and count of all hashtags currently used.

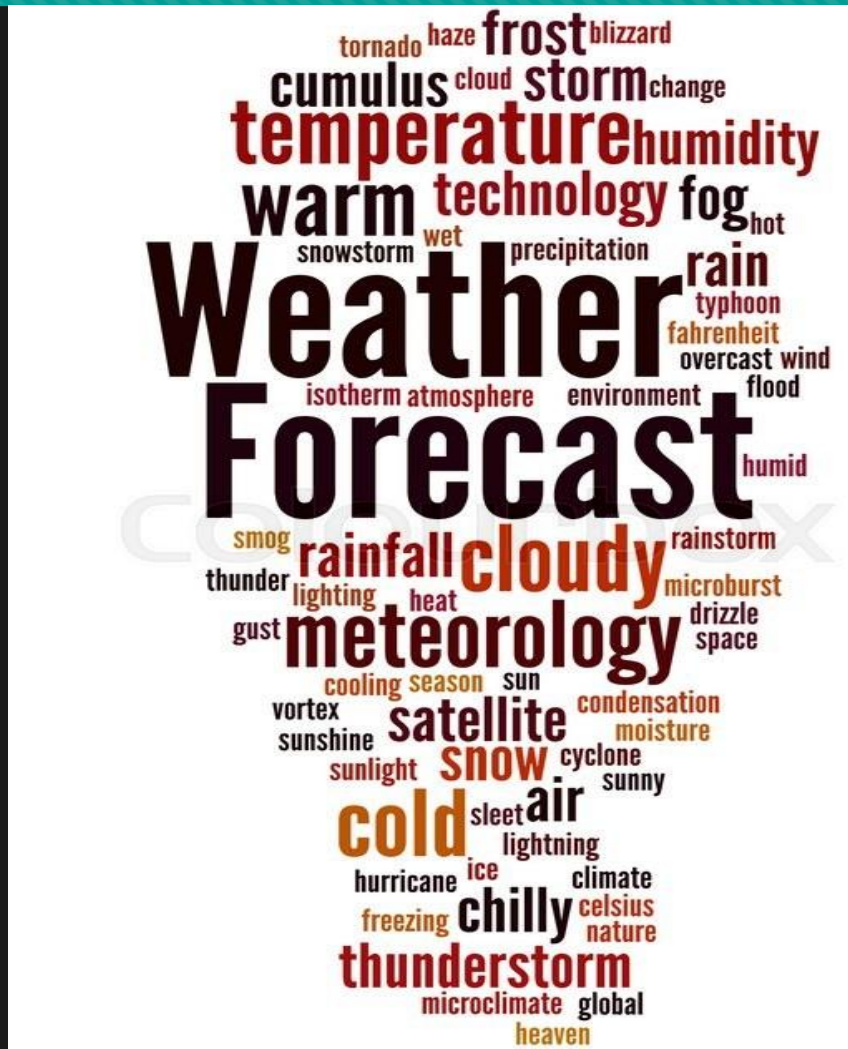
Running “for key in hashtag”, this sends the data to a new hashtag file.

When running word cloud, this takes the words from the hashtag file and stores the data in the cloud. The bigger words means they are used the most.

Future Works

If we had more time we would have:

- **Added another team member to incorporate unique ideas.**
- **Track the number of retweets and favorites to compare the data.**
- **Displayed a word cloud outcome involving weather.**



Special Thanks

Rajesh Kalyanam – Mentor and Creator of MyGeoHub!

SGCI Mentors!

HackAttacks Mission Continues…

“As SGCI Members, we will not give up on this production”.