Introduction

We're looking to see if there is a correlation of NYC Covid cases rate to keywords found in tweets from verified user in New York City.

Twint

In [1]:
```
import twint
import pandas as pd

import nest_asyncio
nest_asyncio.apply()
```

Covid Tweets

Create a function that has the appropriate criteria and parameters to obtain the desire tweets from New York City. Using keywords to feed through the function to obtain the tweets and saving it to a csv file named 'covidtweets.csv'. We're also looking for likes count as well for each tweets as a way to give more weight to the words since it is imply that the reader are affected by the tweets.

Six Keywords were used to obtain the tweets are:

- Covid
- Corona
- Coronavirus
- Mask
- Vaccine
- Quarantine

In [2]:
```
# c = twint.Config()
# def tweetsearch(keyword):
#     c.Search = keyword
#     c.Custom["tweet"] = ["date", "time", "username", "tweet", "likes_count"]
#     c.Geo = "40.730610, -73.935242, 25mi"
#     c.Since = '2020-01-01'
#     c.Until = '2021-02-14'
#     c.Verified = True
#     c.Min_likes = 50
#     c.Count = True
#     c.Limit = 10000
#     c.Store_csv = True
#     c.Output = 'covidtweets.csv'
```