

(1) Create a JSON File of Your Credentials

```
In [ ]: ...
Q1: Create your own credential JSON file in the current directory.

    At first, you need to create a dictionary of your credentials.

    Second, you need to create a JSON file (i.e. twitter_credentials.json) of your cred
...

```

```
In [ ]: credential = {"API_KEY": "VqR9HJemANJ893BekWXTxXWT7",
    "API_SECRET": "gfAdxdfXj3eP9xH1XzwwK1cs7tQ6x2q68yzQRP4g0zuDu3t0iF",
    "ACCESS_TOKEN": "1244356645393707024-YgjM3PKRbwc65oFmmgdyEFNSkhoHqF",
    "ACCESS_TOKEN_SECRET": "LJLnwmw65iMuZnR2ULsU5RQq8Edfu4f1oHUilW3If9TTe"
}

credential

```

```
In [ ]: import json
outfile = open('twitter_credentials.json', 'w')
json.dump(credential, outfile)
outfile.close()

```

(2) Collect 1K Tweets of Your Keyword

```
In [ ]: ...
Q2: Choose your own keyword to collect 1K tweets.

    Please collect 1K tweets. Please use a code in Data_Collection_Twitter_API.ipynb.

    You will have "tweet_stream_easter_1000.json" in your working folder.
...

```

```
In [ ]: from twython import Twython
import sys

```

```
In [ ]: infile = open('twitter_credentials.json', 'r')
credentials = json.load(infile)
infile.close()

credentials

```

```
In [ ]: API_KEY = credentials['API_KEY']
API_SECRET = credentials['API_SECRET']

```

```
In [ ]: twitter = Twython(API_KEY, API_SECRET)

```

```
In [ ]: res = twitter.search(q = "Clemson", count = 1000)

for t in res['statuses']:
    if t['lang'] == 'en':
        print(t['text'])
```

```
In [ ]: outfile = open('tweet_stream_easter_1000.json', 'w')
json.dump(res, outfile)
outfile.close()
```

(3) Read the JSON file for Preliminary Analysis

```
In [ ]: ...
Q3: Read your JSON file (i.e. tweet_stream_easter_1000.json) and assign it to the "data"
...
```

```
In [ ]: infile = open('tweet_stream_easter_1000.json')
data = json.load(infile)
infile.close()
```

```
In [ ]: print(data)
```

```
In [ ]: from pprint import pprint
pprint(data)
```

(4) Preliminary Analysis

```
In [ ]: ...
Q4: What are the ten most popular hashtags?
...
```

```
In [ ]: import json
from collections import Counter
from pprint import pprint
```

```
In [ ]: infile = open('tweet_stream_easter_1000.json')
data = json.load(infile)
infile.close()
```

```
In [ ]: type(data)
```

```
In [ ]: hashtag_list = []
        for t in data:
            for h in t['entities']['hashtags']:
                if h != []:
                    hashtag_list.append(h['text'])
```

```
In [ ]: pprint(hashtag_list)
```

```
In [ ]: c = Counter(hashtag_list)
```

```
In [ ]: c.most_common(10)
```

```
In [ ]: ...
        Q5: Who is the most frequently tweeting person about the topic?
        ...
```

```
In [ ]: name_list = []
        for t in data:
            if t != '':
                name_list.append(t['user']['name'])
```

```
In [ ]: pprint(name_list)
```

```
In [ ]: c = Counter(name_list)
```

```
In [ ]: c.most_common(10)
```

(5) Create a WordCloud from Your 1K Tweets

```
In [ ]: ...
        Q6: (1) Read your JSON file (i.e. tweet_stream_easter_1000.json) and assign it to "contents"
            (2) Generate a WordCloud image from the "contents" variable.
        ...
```

```
In [ ]: from wordcloud import WordCloud
        import matplotlib.pyplot as plt
        %matplotlib inline
```

```
In [ ]: infile = open('tweet_stream_easter_1000.json')
        contents = infile.read()
        infile.close()
```

```
In [ ]: wordcloud = WordCloud(max_font_size = 80, collocations = False).generate(contents)
```

```
In [ ]: ...  
Q7: (1) Display the generated WordCloud image; and  
     (2) Save the image as wordcloud_new.pdf.  
     ...
```

```
In [ ]: %matplotlib inline  
plt.figure(figsize=(15,10))  
plt.imshow(wordcloud)  
plt.axis("off")  
plt.savefig('wordcloud.pdf')  
plt.show()
```