Code to get User Data

Used tweepy to collect data of all users who tweeted with the hashtag(#INDvsAUS)

User list derived from the tweets.json file collected

Import Required Modules-

In [1]:
```
import tweepy
import pandas as pd
import time
import json
from IPython.display import display
import matplotlib.pyplot as plt

import plotly.express as px

consumer_key = "Xrhspj6TP8dX7OcE1UI6sFCcx"
consumer_secret = "NWb6i1KIx4jn0HFPLsNvSU4PViRi0ZdeApKzPXcVZrQJ819F1i"
access_token  = "1339056562980024321-JKxnhaw8vszFxEV6SXnbKrmhzhV24D"
access_token_secret = "qNFHIHlIGer8XA1JUuwkbohzxGqfvlrccIzTB5t57c26H"
```

Authentication -

In [2]:
```
auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)
api = tweepy.API(auth,wait_on_rate_limit=True)
```

Collecting Users List from tweets.json file

In [3]:
```
df2 = pd.read_json("tweets.json",lines = True)
```

In [4]:
```
df3 = df2.drop_duplicates(subset = ['id'])
df3.index = range(len(df3))

df_temp = df3.username.unique()
users = df_temp.tolist()
```

Looping through User list to get User tweepy object

In [ ]:
```
user_info = []
for user in users:
    try:
        user_obj = api.get_user(user)
        user_info.append([user_obj.screen_name,user_obj.followers_count,
                          user_obj.verified, user_obj.friends_count,
                          user_obj.location,user_obj.statuses_count
                         ])
        print(user)


    except tweepy.TweepError as e:

        print (e)  # prints 34
```

In [ ]:
```

```

In [9]:
```
print(len(user_info))
```

```
260
```

In [ ]:
```
for user in users[260:]:
    try:
        user_obj = api.get_user(user)
        user_info.append([user_obj.screen_name,user_obj.followers_count,
                          user_obj.verified, user_obj.friends_count,
                          user_obj.location,user_obj.statuses_count
                         ])
        print(user)


    except tweepy.TweepError as e:

        print (e)  # prints 34
```

In [11]:
```
print(len(user_info))
```

```
6488
```

Dumping User dataframe into user.json file

In [19]:
```
df = pd.DataFrame(user_info,columns = ['Name','Followers','Verified','Friends','Location','Tweets'])
```

In [13]:
```
df.to_json('user.json')
```

In [17]:
```
display(df)
```

	Name	Followers	Verified	Friends	Location	Tweets
0	Sexyano_Donaldo	2607	False	970	Washington DC, Nepal	32221
1	interviewtimes2	134	False	28	Bhubaneshwar	1357
2	NarayanShastri	788	False	2518	ಬೆಂಗಳೂರು Bengaluru	53515
3	iamritikagarwal	1	False	144		559
4	Oneindia	64296	True	619	India	365502
...	...	...	...	...	...	...
6483	iamayush312	73	False	291		4586
6484	Kasturi_FanGirl	532	False	214		31472
6485	Rexxy_09	21	False	131		66
6486	SurajitTweet	3001	False	382	India	55482
6487	Msdhoni_183	2607	False	2305	India	13913

6488 rows × 6 columns

In [28]:
```
out = df.to_json(orient='records')[1:-1].replace('},{', '}\n {')
```

In [29]:
```
with open('user.json', 'w') as f:
    f.write(out)
```

In [30]:
```
df2 = pd.read_json("user.json",lines = True)
```

In [31]:
```
display(df2)
```

	Name	Followers	Verified	Friends	Location	Tweets
0	Sexyano_Donaldo	2607	False	970	Washington DC, Nepal	32221
1	interviewtimes2	134	False	28	Bhubaneshwar	1357
2	NarayanShastri	788	False	2518	ಬೆಂಗಳೂರು Bengaluru	53515
3	iamritikagarwal	1	False	144		559
4	Oneindia	64296	True	619	India	365502
...	...	...	...	...	...	...
6483	iamayush312	73	False	291		4586
6484	Kasturi_FanGirl	532	False	214		31472
6485	Rexxy_09	21	False	131		66
6486	SurajitTweet	3001	False	382	India	55482
6487	Msdhoni_183	2607	False	2305	India	13913

6488 rows × 6 columns