Sentimental Analysis On Pfizer Vaccine Tweets

In [1]:
```
1 import pandas as pd
2 data=pd.read_csv("D:/vaccination_tweets.csv")
3 data
```

Out[1]:

	id	user_name	user_location	user_description	user_created	user_followers	user_friends	user_favourites
0	1340539111971516416	Rachel Roh	La Crescenta-Montrose, CA	Aggregator of Asian American news; scanning di...	2009-04-08 17:52:46	405	1692	3247
1	1338158543359250433	Albert Fong	San Francisco, CA	Marketing dude, tech geek, heavy metal & '80s ...	2009-09-21 15:27:30	834	666	178
2	1337858199140118533	eli𝖑𝖙𝖊𝖚	Your Bed	heil, hydra ☺	2020-06-25 23:30:28	10	88	155
3	1337855739918835717	Charles Adler	Vancouver, BC - Canada	Hosting "CharlesAdlerTonight" Global News Radi...	2008-09-10 11:28:53	49165	3933	21853
		Citizen		Citizen News	2020-04-23			

In [2]:
```
1 data.columns
```

Out[2]:
```
Index(['id', 'user_name', 'user_location', 'user_description', 'user_created',
       'user_followers', 'user_friends', 'user_favourites', 'user_verified',
       'date', 'text', 'hashtags', 'source', 'retweets', 'favorites',
       'is_retweet'],
      dtype='object')
```

In [3]:
```
1 data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3882 entries, 0 to 3881
Data columns (total 16 columns):
 #   Column            Non-Null Count  Dtype 
---  ------            --------------  ----- 
 0   id                3882 non-null   int64 
 1   user_name         3882 non-null   object
 2   user_location     3070 non-null   object
 3   user_description  3640 non-null   object
 4   user_created      3882 non-null   object
 5   user_followers    3882 non-null   int64 
 6   user_friends      3882 non-null   int64 
 7   user_favourites   3882 non-null   int64 
 8   user_verified     3882 non-null   bool  
 9   date              3882 non-null   object
 10  text              3882 non-null   object
 11  hashtags          2884 non-null   object
 12  source            3881 non-null   object
 13  retweets          3882 non-null   int64 
 14  favorites         3882 non-null   int64 
 15  is_retweet        3882 non-null   bool  
dtypes: bool(2), int64(6), object(8)
memory usage: 432.3+ KB
```

In [4]:
```
1 data.isnull().sum()
```

Out[4]:
```
id                    0
user_name             0
user_location       812
user_description    242
user_created          0
user_followers        0
user_friends          0
user_favourites       0
user_verified         0
date                  0
text                  0
hashtags            998
source                1
retweets              0
favorites             0
is_retweet            0
dtype: int64
```

In [5]:
```
1 data['user_location'].fillna("Unknown", inplace=True)
2 data['user_description'].fillna("No Description", inplace=True)
3 data['hashtags'].fillna("No Hashtags", inplace=True)
4 data['source'].fillna("Twitter Web App", inplace=True)
```

In [6]:
```
1 data.isnull().sum()
```

Out[6]:
```
id                  0
user_name           0
user_location       0
user_description    0
user_created        0
user_followers      0
user_friends        0
user_favourites     0
user_verified       0
date                0
text                0
hashtags            0
source              0
retweets            0
favorites           0
is_retweet          0
dtype: int64
```

In [7]:
```
1 data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3882 entries, 0 to 3881
Data columns (total 16 columns):
 #   Column            Non-Null Count  Dtype 
---  ------            --------------  ----- 
 0   id                3882 non-null   int64 
 1   user_name         3882 non-null   object
 2   user_location     3882 non-null   object
 3   user_description  3882 non-null   object
 4   user_created      3882 non-null   object
 5   user_followers    3882 non-null   int64 
 6   user_friends      3882 non-null   int64 
 7   user_favourites   3882 non-null   int64 
 8   user_verified     3882 non-null   bool  
 9   date              3882 non-null   object
 10  text              3882 non-null   object
 11  hashtags          3882 non-null   object
 12  source            3882 non-null   object
 13  retweets          3882 non-null   int64 
 14  favorites         3882 non-null   int64 
 15  is_retweet        3882 non-null   bool  
dtypes: bool(2), int64(6), object(8)
memory usage: 432.3+ KB
```

In [8]:
```
1 data["user_location"].value_counts()
```

Out[8]:
```
Unknown                812
London, England         60
India                   55
United Arab Emirates    43
Canada                  41
                      ...
Malaga, Spain            1
Ice age                  1
Buffalo, NY              1
Quahog                   1
Santa Barbara, CA        1
Name: user_location, Length: 1271, dtype: int64
```

In [ ]:
```
1
```

Top 20 Locations

In [9]:
```
1  import seaborn as sns
2  import matplotlib.pyplot as plt
3  %matplotlib inline
4  plt.figure(figsize=(20,10))
5  sns.barplot(data["user_location"].value_counts().index[1:21], data["user_location"].value_counts().values[1
6  plt.title("Top 20 Locations", fontsize=15)
7  plt.xlabel("Locations", fontsize=15)
8  plt.xticks(rotation=60)
9  plt.ylabel("Total Number of Tweets", fontsize=15)
10 plt.show()
```

In [10]:
```
df=pd.DataFrame(data['user_location'].value_counts()[1:50]).reset_index()
df
```

Out[10]:

	index	user_location
0	London, England	60
1	India	55
2	United Arab Emirates	43
3	Canada	41
4	United Kingdom	39
5	Malaysia	34
6	GP Partner, Watford	31
7	Email:talksavailable@gmail.com	30
8	London	29
9	Dubai, United Arab Emirates	28
10	Türkiye	27
11	New Delhi, India	25
12	Mumbai, India	25
13	England, United Kingdom	25
14	United States	23
15	Glasgow environs.	22
16	London, UK	21
17	Toronto, Ontario	20
18	Global	20
19	Atlanta, GA	19
20	California, USA	18
21	Ireland	18
22	Cornwall UK	18
23	Chicago, IL	18

	index	user_location
24	#KeepFightingMichael	17
25	France	17
26	Cork, Ireland	17
27	North West, England	16
28	UK	16
29	Chennai, India	15
30	New York, NY	15
31	Earth	14
32	Houston, TX	14
33	USA	13
34	Denmark	13
35	Los Angeles, CA	13
36	Cyprus	13
37	Miami, FL	13
38	Boston, MA	12
39	Sri Lanka	12
40	Qatar	12
41	Turkey-Israel	12
42	Michigan, USA	11
43	Tweets on #Asia & USA	11
44	Poole	11
45	Birmingham, England	10
46	New Delhi	10
47	La Belle Province	10
48	Singapore	10

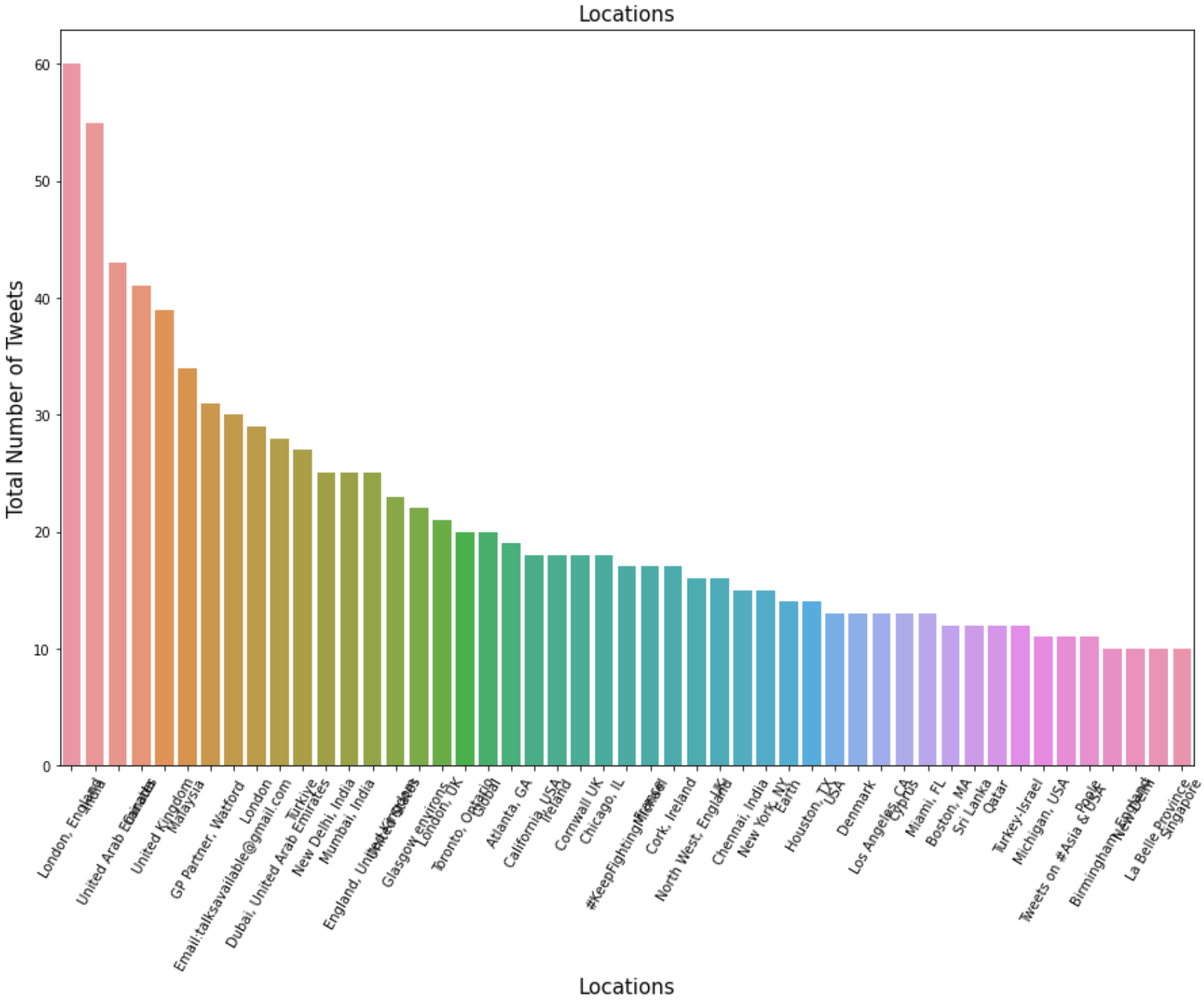
visualizing tweets/Locations

In [11]:

```
plt.figure(figsize=(15,10))
sns.barplot(df["index"], df["user_location"]);

plt.title("Locations", fontsize=15)

plt.xlabel("Locations", fontsize=15)
plt.xticks(rotation=60)
plt.ylabel("Total Number of Tweets", fontsize=15)
plt.show()
```


Locations
Total Number of Tweets
0
10
20
30
40
50
60
London, England
United Arab Emirates
United Kingdom
Malaysia
GP Partner, Watford
London
Email:talksavailable@gmail.com
Türkiye
Dubai, United Arab Emirates
New Delhi, India
Mumbai, India
England, United Kingdom
Glasgow environs
London, UK
Toronto, Ontario
Global
Atlanta, GA
California, USA
Ireland
Cornwall UK
Chicago, IL
#KeepFightingMissouri
Cork, Ireland
North West, England
Chennai, India
New York, NY
Earth
Houston, TX
USA
Denmark
Los Angeles, CA
Cyprus
Miami, FL
Boston, MA
Sri Lanka
Qatar
Turkey-Israel
Michigan, USA
Tweets on #Asia & USA
Birmingham, England
La Belle Province
Singapore
Locations

In [12]:
```
import re
def cleanTweet(txt):
        txt = re.sub('\w*\d\w*', '', txt)
        txt = re.sub('\n', '', txt)

        # Remove mentions
        txt = re.sub('@[A-Za-z0-9_]+', '', txt)
        # Remove hashtags
        txt = re.sub('#', '', txt)
        # Remove punctuations:
        txt = re.sub('\[.*?\]', '', txt)
        # Remove urls
        txt = re.sub('https?:\/\/[A-Za-z0-9\.\/]+', '', txt)
        return txt

```

Cleaning tweets of text attribute

In [13]:
```
data['text'] = data['text'].apply(cleanTweet)
```

In [14]:
```
1 data['text']
```

Out[14]:
```
0       Same folks said daikon paste could treat a cyt...
1       While the world has been on the wrong side of ...
2       coronavirus SputnikV AstraZeneca PfizerBioNTec...
3       Facts are immutable, Senator, even when you're...
4       Explain to me again why we need a vaccine   wh...
                              ...
3877      dear dr. I would like to ask about when is t...
3878    Ok almost  since  dose of Pfizer covid vaccine...
3879    Katalin Karikó, a biochemist working for Biont...
3880      My husband and my aunt had their PfizerBioNT...
3881    E-learning on a Sunday in case I’m needed to w...
Name: text, Length: 3882, dtype: object
```

In [15]:
```
1 data.iloc[0,10]
```

Out[15]:
```
'Same folks said daikon paste could treat a cytokine storm PfizerBioNTech '
```

Using TextBlob determining subjectivity and polarity

In [16]:
```
1 from textblob import TextBlob
2 def getTweettSubjectivity(txt):
3      return TextBlob(txt).sentiment.subjectivity
4 def getTweetPolarity(txt):
5      return TextBlob(txt).sentiment.polarity
```

In [17]:
```
1 data['Subjectivity'] = data['text'].apply(getTweettSubjectivity)
2 data['Polarity'] = data['text'].apply(getTweetPolarity)
```

In [18]:
```
a=data[['Subjectivity','Polarity']]
a
```

Out[18]:

	Subjectivity	Polarity
0	0.125000	0.00
1	0.900000	-0.50
2	0.033333	0.00
3	0.550000	-0.05
4	0.000000	0.00
...	...	...
3877	0.300000	1.00
3878	0.741667	0.05
3879	0.000000	0.00
3880	0.000000	0.00
3881	0.000000	0.00

3882 rows × 2 columns

In [19]:
```
1 data
```

Out[19]:

	id	user_name	user_location	user_description	user_created	user_followers	user_friends	user_favourites	u
0	1340539111971516416	Rachel Roh	La Crescenta-Montrose, CA	Aggregator of Asian American news; scanning di...	2009-04-08 17:52:46	405	1692	3247	
1	1338158543359250433	Albert Fong	San Francisco, CA	Marketing dude, tech geek, heavy metal & '80s ...	2009-09-21 15:27:30	834	666	178	
2	1337858199140118533	eli𝖑𝖙𝖊𝖚👌	Your Bed	heil, hydra ✋☺	2020-06-25 23:30:28	10	88	155	
3	1337855739918835717	Charles Adler	Vancouver, BC - Canada	Hosting "CharlesAdlerTonight" Global News Radi...	2008-09-10 11:28:53	49165	3933	21853	
4	1337854064604966912	Citizen News Channel	Unknown	Citizen News Channel bringing you an alternati...	2020-04-23 17:58:42	152	580	1473	
...	...	...	...	...	...	...	...	...	
3877	1348284945358331906	Mounir Basalus (Βασιλιος)	Hengelo, Nederland	Cardiology, Cardiac devices, drug-eluting sten...	2010-05-02 19:12:33	680	77	2344	
3878	1348282854313230336	Meghana Chalasani, MD	San Antonio, TX	I like dogs, kidneys, and purses. In that order.	2010-01-23 06:09:37	443	721	25591	
3879	1348280046847406088	Hungary Today	Budapest	The latest tweets about #Hungary	2014-10-17 13:11:16	8520	1061	2370	
3880	1348279749144162305	Hilary	West Yorkshire	Wife, Mother, Nannie, Health Worker. Fave film...	2012-07-23 16:45:50	506	1123	31856	

	id	user_name	user_location	user_description	user_created	user_followers	user_friends	user_favourites	us
3881	1348277170213445632	Kimberley ♡	Leeds, England	Mum/Registered Mental Health Nurse/Specialist ...	2012-07-14 20:59:10	182	252	2428	

3882 rows × 18 columns

In [20]:

```
def getTextAnalysis(x):
     if x<0:
        return "Negative"
     elif x == 0:
        return "Neutral"
     else:
        return "Positive"
```

In [21]:
```
data['Sentiment'] = data["Polarity"].apply(getTextAnalysis)
data.head(20)
```

Out[21]:

	id	user_name	user_location	user_description	user_created	user_followers	user_friends	user_f
0	1340539111971516416	Rachel Roh	La Crescenta-Montrose, CA	Aggregator of Asian American news; scanning di...	2009-04-08 17:52:46	405	1692	
1	1338158543359250433	Albert Fong	San Francisco, CA	Marketing dude, tech geek, heavy metal & '80s ...	2009-09-21 15:27:30	834	666	
2	1337858199140118533	eliLTEU	Your Bed	heil, hydra	2020-06-25 23:30:28	10	88	
3	1337855739918835717	Charles Adler	Vancouver, BC - Canada	Hosting "CharlesAdlerTonight" Global News Radi...	2008-09-10 11:28:53	49165	3933	
4	1337854064604966912	Citizen News Channel	Unknown	Citizen News Channel bringing you an alternati...	2020-04-23 17:58:42	152	580	
5	1337852648389832708	Dee	Birmingham, England	Gastroenterology trainee, Clinical Research Fe...	2020-01-26 21:43:12	105	108	
6	1337851215875608579	Gunther Fehlinger	Austria, Ukraine and Kosovo	End North Stream 2 now - the pipeline of corru...	2013-06-10 17:49:22	2731	5001	
7	1337850832256176136	Dr.Krutika Kuppalli	Unknown	ID, Global Health, VHF, Pandemic Prep, Emergin...	2019-03-25 04:14:29	21924	593	
8	1337850023531347969	Erin Despas	Unknown	Designing&selling on Teespring. Like 90s Disne...	2009-10-30 17:53:54	887	1515	
9	1337842295857623042	Ch.Amjad Ali	Islamabad	#ProudPakistani #LovePakArmy #PMIK @insafiansp...	2012-11-12 04:18:12	671	2368	
10	1337841934170255365	Tamer Yazar	Turkey-Israel	Im Market Analyst, also Editor... working (fre...	2009-09-17 16:45:16	1302	78	

	id	user_name	user_location	user_description	user_created	user_followers	user_friends	user_f
11	1337840331522453507	VoiceM	Unknown	campaigner & optimistic realist	2020-08-31 10:38:21	2	25	
12	1337815730486702087	WION	India	#WION: World Is One | Welcome to India's first...	2016-03-21 03:44:54	292510	91	
13	1337809196453081088	Dr.Krutika Kuppalli	Unknown	ID, Global Health, VHF, Pandemic Prep, Emergin...	2019-03-25 04:14:29	21924	593	
14	1337807109010780162	Opoyi	Unknown	High-quality trusted conversations around news...	2019-01-13 18:33:22	10332	49	
15	1337789467038523394	City A.M.	London, England	London's business newspaper - News, Opinion, a...	2009-06-09 13:53:06	66224	603	
16	1337789191598575616	STOPCOMMONPASS.ORG	Global	'Trust' is not carte-blanche for erosion of ou...	2020-10-25 20:33:33	406	176	
17	1337783770070409218	ILKHA	Türkiye	Official Twitter account of Ilke News Agency /	2015-05-22 08:31:12	4056	6	
18	1337781178665816064	Braderz73✊#GTTO EU GB	Bristol, UK	One of those lefty types, I believe in a faire...	2012-07-24 08:18:51	6430	6292	
19	1337776763384958976	Alex Vie	Los Angeles, CA	Marine vet. Yogi. Krav Maga. Runner. Climber. ...	2010-01-24 04:43:57	125	442	

In [22]:
```
positive_tweets = data[data['Sentiment'] == 'Positive']
positive_tweets
```

Out[22]:

	id	user_name	user_location	user_description	user_created	user_followers	user_friends	user_favourites	us
5	1337852648389832708	Dee	Birmingham, England	Gastroenterology trainee, Clinical Research Fe...	2020-01-26 21:43:12	105	108	106	
7	1337850832256176136	Dr.Krutika Kuppalli	Unknown	ID, Global Health, VHF, Pandemic Prep, Emergin...	2019-03-25 04:14:29	21924	593	7815	
10	1337841934170255365	Tamer Yazar	Turkey-Israel	Im Market Analyst, also Editor... working (fre...	2009-09-17 16:45:16	1302	78	339	
11	1337840331522453507	VoiceM	Unknown	campaigner & optimistic realist	2020-08-31 10:38:21	2	25	20	
12	1337815730486702087	WION	India	#WION: World Is One | Welcome to India's first...	2016-03-21 03:44:54	292510	91	7531	
...	...	...	...	...	...	...	...	...	...
3872	1348291856166252546	Prashant Sankaye	London, England	Views own. Consultant MSK Radiologist with PGD...	2019-06-23 11:19:58	243	393	82	
3875	1348288244191031296	Chelsea Palmer 🤍	York, England	NHS Biomedical Scientist in Cytopathology/Andr...	2010-01-16 19:46:39	141	551	4161	
3876	1348286191418941451	Maryam	Canada	No Description	2011-11-14 03:59:34	187	373	19284	
3877	1348284945358331906	Mounir Basalus (Βασιλιος)	Hengelo, Nederland	Cardiology, Cardiac devices, drug-eluting sten...	2010-05-02 19:12:33	680	77	2344	
3878	1348282854313230336	Meghana Chalasani, MD	San Antonio, TX	I like dogs, kidneys, and purses. In that order.	2010-01-23 06:09:37	443	721	25591	

1658 rows × 19 columns

In [23]:
```
1 positive_tweets.shape
```

Out[23]: (1658, 19)

In [24]:
```
1 print(str(positive_tweets.shape[0]/(data.shape[0])*100) +"% of positive Tweets")
```

```
42.709943328181346% of positive Tweets
```

In [25]:
```
1 Neutral_tweets = data[data['Sentiment'] == 'Neutral']
2 print(str(Neutral_tweets.shape[0]/(data.shape[0])*100) +"% of Neutral Tweets")
```

```
47.37248840803709% of Neutral Tweets
```

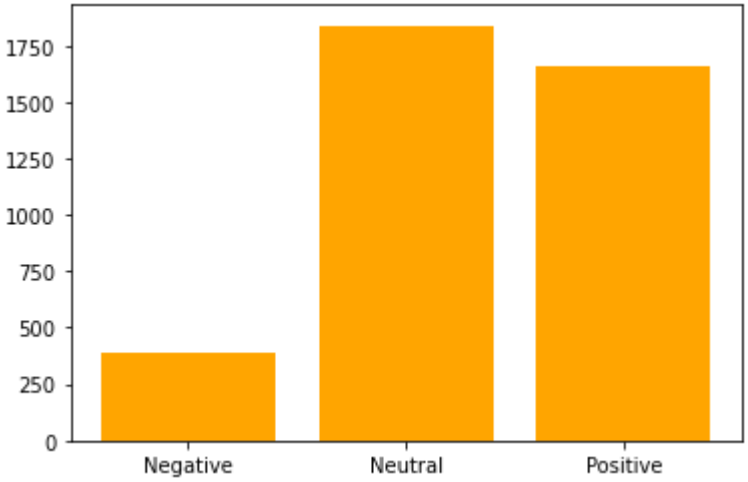
In [26]:
```
1 negative_tweets = data[data['Sentiment'] == 'Negative']
2 print(str(negative_tweets.shape[0]/(data.shape[0])*100) + " % of Negative Tweets")
```

```
9.917568263781556 % of Negative Tweets
```

In [27]:
```
1 data['Sentiment'].value_counts()
```

Out[27]:
```
Neutral     1839
Positive    1658
Negative     385
Name: Sentiment, dtype: int64
```

In [28]:
```
labels = data.groupby('Sentiment').count().index.values
values = data.groupby('Sentiment').size().values
plt.bar(labels,values,color='orange')
plt.show()
```



In [29]:
```
loc_df = data['user_location'].astype(str).str.split(',',expand=True)
loc_df=loc_df.rename(columns={0:'fst_loc',1:'snd_loc'})
```

In [30]:
```
1 loc_df.head(9)
```

Out[30]:

	fst_loc	snd_loc	2	3	4	5
0	La Crescenta-Montrose	CA	None	None	None	None
1	San Francisco	CA	None	None	None	None
2	Your Bed	None	None	None	None	None
3	Vancouver	BC - Canada	None	None	None	None
4	Unknown	None	None	None	None	None
5	Birmingham	England	None	None	None	None
6	Austria	Ukraine and Kosovo	None	None	None	None
7	Unknown	None	None	None	None	None
8	Unknown	None	None	None	None	None

In [31]:
```
1 #strip() use to remove before-after space
2 loc_df['snd_loc'] = loc_df['snd_loc'].str.strip()
3 loc_df['snd_loc'].head(10)
```

Out[31]:
```
0                  CA
1                  CA
2                None
3          BC - Canada
4                None
5             England
6    Ukraine and Kosovo
7                None
8                None
9                None
Name: snd_loc, dtype: object
```

In [32]:
```
# Rename States
state_fix = {'Ontario': 'Canada','United Arab Emirates': 'UAE','United Kingdom':'UK','Wales':'UK','London':
                  ,'FL': 'USA','England': 'UK','Watford': 'UK','GA': 'USA','IL': 'USA','New South Wales':'A
                  ,'Alberta': 'Canada','WA': 'USA','NC': 'USA','British Columbia': 'Canada','MA': 'USA','ON
            ,'OH':'USA','MO':'USA','AZ':'USA','NJ':'USA','CA':'USA','DC':'USA','AB':'USA','PA':'USA','SC':'
            ,'VA':'USA','TN':'USA','New York':'USA','Dubai':'UAE','CO':'USA' ,'BC - Canada':'Canada','Delhi
             'Germany':'Europe','France':'Europe','Nederland':'Europe','Belgium':'Europe','Italy':'Europe',
            'Switzerland':'Europe','Deutschland':'Europe','BC':'Canada','Polska':'Europe','LA':'USA','Calif
loc_df = loc_df.replace({"snd_loc": state_fix})
loc_df['snd_loc'].value_counts()[:20]
```

Out[32]:
```
USA               460
UK                401
India             114
Canada            102
Europe             77
UAE                43
Australia          19
Pakistan            8
Malaysia            7
Turkey              6
UT                  6
RI                  5
B No 4              5
Philippines         4
MD                  4
Chile               4
South Africa        4
Nigeria             4
MN                  4
earth sky moon      4
Name: snd_loc, dtype: int64
```

In [33]:
```
1 loc_df['snd_loc']
```

Out[33]:
```
0          USA
1          USA
2         None
3       Canada
4         None
          ...
3877    Europe
3878       USA
3879      None
3880      None
3881        UK
Name: snd_loc, Length: 3882, dtype: object
```

In [34]:
```
1 data = data.drop('user_location', axis = 1)
```

In [35]:
```
1 data
```

Out[35]:

	id	user_name	user_description	user_created	user_followers	user_friends	user_favourites	user_verified
0	1340539111971516416	Rachel Roh	Aggregator of Asian American news; scanning di...	2009-04-08 17:52:46	405	1692	3247	False
1	1338158543359250433	Albert Fong	Marketing dude, tech geek, heavy metal & '80s ...	2009-09-21 15:27:30	834	666	178	False
2	1337858199140118533	eli𝖑𝖙𝖊𝖚👌	heil, hydra 👋☺	2020-06-25 23:30:28	10	88	155	False
3	1337855739918835717	Charles Adler	Hosting "CharlesAdlerTonight" Global News Radi...	2008-09-10 11:28:53	49165	3933	21853	True
4	1337854064604966912	Citizen News Channel	Citizen News Channel bringing you an alternati...	2020-04-23 17:58:42	152	580	1473	False
...	...	...	...	...	...	...	...	...
3877	1348284945358331906	Mounir Basalus (Βασιλιος)	Cardiology, Cardiac devices, drug-eluting sten...	2010-05-02 19:12:33	680	77	2344	False
3878	1348282854313230336	Meghana Chalasani, MD	I like dogs, kidneys, and purses. In that order.	2010-01-23 06:09:37	443	721	25591	False
3879	1348280046847406088	Hungary Today	The latest tweets about #Hungary	2014-10-17 13:11:16	8520	1061	2370	False
3880	1348279749144162305	Hilary	Wife, Mother, Nannie, Health Worker. Fave film...	2012-07-23 16:45:50	506	1123	31856	False

	id	user_name	user_description	user_created	user_followers	user_friends	user_favourites	user_verified
3881	1348277170213445632	Kimberley ♡	Mum/Registered Mental Health Nurse/Specialist ...	2012-07-14 20:59:10	182	252	2428	False

3882 rows × 18 columns

In [36]:
```
data['users_location']=loc_df['snd_loc']
data
```

Out[36]:

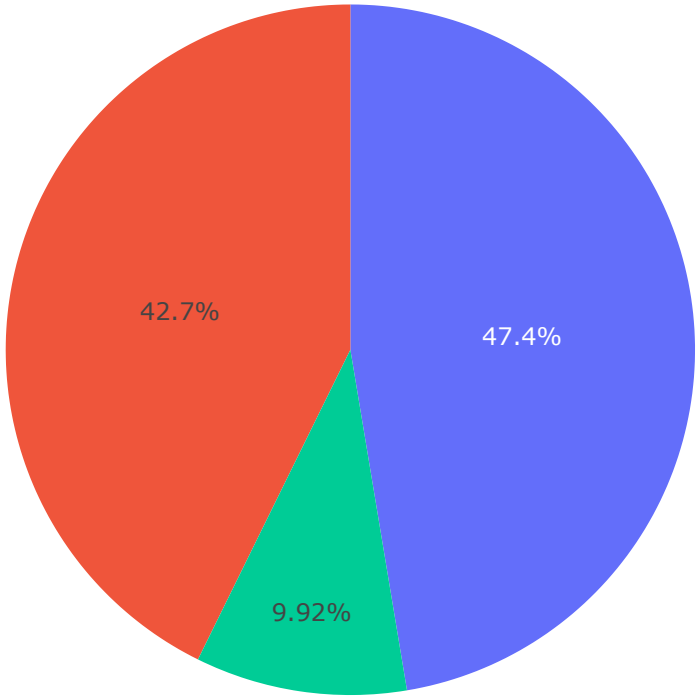
	id	user_name	user_description	user_created	user_followers	user_friends	user_favourites	user_verified	
0	1340539111971516416	Rachel Roh	Aggregator of Asian American news; scanning di...	2009-04-08 17:52:46	405	1692	3247	False	06
1	1338158543359250433	Albert Fong	Marketing dude, tech geek, heavy metal & '80s ...	2009-09-21 15:27:30	834	666	178	False	16
2	1337858199140118533	eliLTEU	heil, hydra ☺	2020-06-25 23:30:28	10	88	155	False	20
3	1337855739918835717	Charles Adler	Hosting "CharlesAdlerTonight" Global News Radi...	2008-09-10 11:28:53	49165	3933	21853	True	20
4	1337854064604966912	Citizen News Channel	Citizen News Channel bringing you an alternati...	2020-04-23 17:58:42	152	580	1473	False	20
...	...	...	...	...	...	...	...	...	
3877	1348284945358331906	Mounir Basalus (Βασιλιος)	Cardiology, Cardiac devices, drug-eluting sten...	2010-05-02 19:12:33	680	77	2344	False	15
3878	1348282854313230336	Meghana Chalasani, MD	I like dogs, kidneys, and purses. In that order.	2010-01-23 06:09:37	443	721	25591	False	14
3879	1348280046847406088	Hungary Today	The latest tweets about #Hungary	2014-10-17 13:11:16	8520	1061	2370	False	14
3880	1348279749144162305	Hilary	Wife, Mother, Nannie, Health Worker. Fave film...	2012-07-23 16:45:50	506	1123	31856	False	14

	id	user_name	user_description	user_created	user_followers	user_friends	user_favourites	user_verified	
3881	1348277170213445632	Kimberley ♡	Mum/Registered Mental Health Nurse/Specialist ...	2012-07-14 20:59:10	182	252	2428	False	14

3882 rows × 19 columns

In [37]:
```
Neutral = len(data[data['Sentiment']=='Neutral'])
Negative = len(data[data['Sentiment']=='Negative'])
Positive = len(data[data['Sentiment']=='Positive'])
labels = ['Negative','Positive','Neutral']
values = [Negative,Positive,Neutral]
#====
print(values)
import plotly.graph_objects as go
fig = go.Figure(data=[go.Pie(labels=labels, values=values)])
fig.show()
```

```
[385, 1658, 1839]
```



In [38]:

```
data_USA=data[data['users_location']=='USA']
data_USA
Neutral = len(data_USA[data_USA['Sentiment']=='Neutral'])
Negative = len(data_USA[data_USA['Sentiment']=='Negative'])
Positive = len(data_USA[data_USA['Sentiment']=='Positive'])
labels = ['Negative','Positive','Neutral']
values = [Negative,Positive,Neutral]

print(values)
import plotly.graph_objects as go
fig = go.Figure(data=[go.Pie(labels=labels, values=values)])
fig.show()
```

```
[37, 235, 188]
```

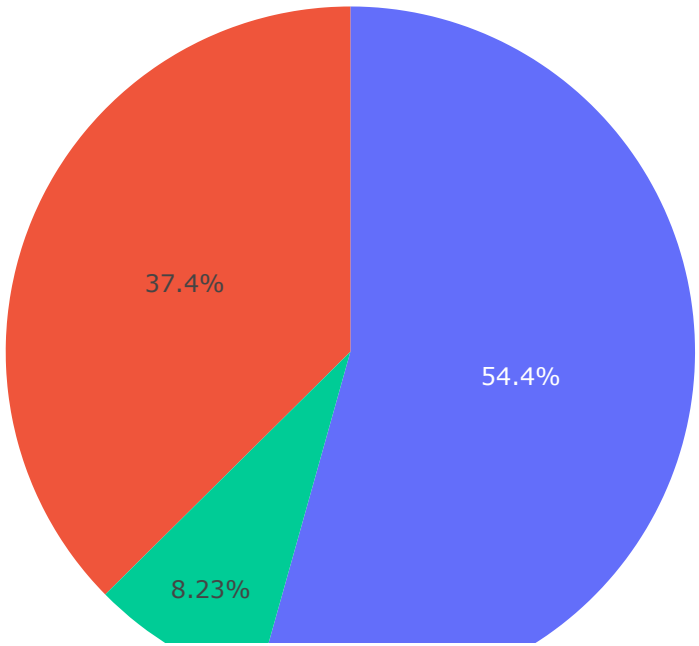
In [39]:
```
data_UK=data[data['users_location']=='UK']
data_UK
Neutral = len(data_UK[data_UK['Sentiment']=='Neutral'])
Negative = len(data_UK[data_UK['Sentiment']=='Negative'])
Positive = len(data_UK[data_UK['Sentiment']=='Positive'])
labels = ['Negative','Positive','Neutral']
values = [Negative,Positive,Neutral]

print(values)
import plotly.graph_objects as go
colors = ['darkred','green', 'darkblue' ]

fig = go.Figure(data=[go.Pie(labels=labels,values=values)])
fig.show()

```

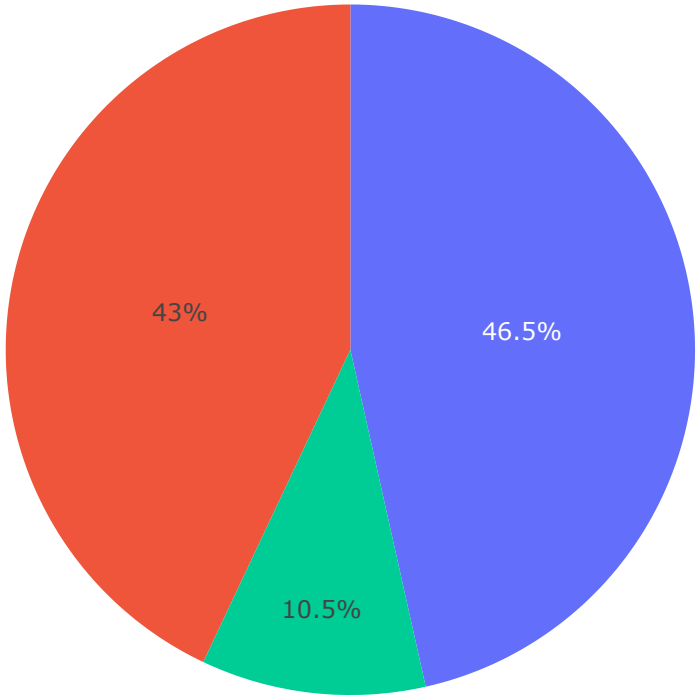
```
[33, 218, 150]
```



In [40]:
```
data_India=data[data['users_location']=='India']
data_India
Neutral = len(data_India[data_India['Sentiment']=='Neutral'])
Negative = len(data_India[data_India['Sentiment']=='Negative'])
Positive = len(data_India[data_India['Sentiment']=='Positive'])
labels = ['Negative','Positive','Neutral']
values = [Negative,Positive,Neutral]

print(values)
import plotly.graph_objects as go
fig = go.Figure(data=[go.Pie(labels=labels,values=values)])
fig.show()
```

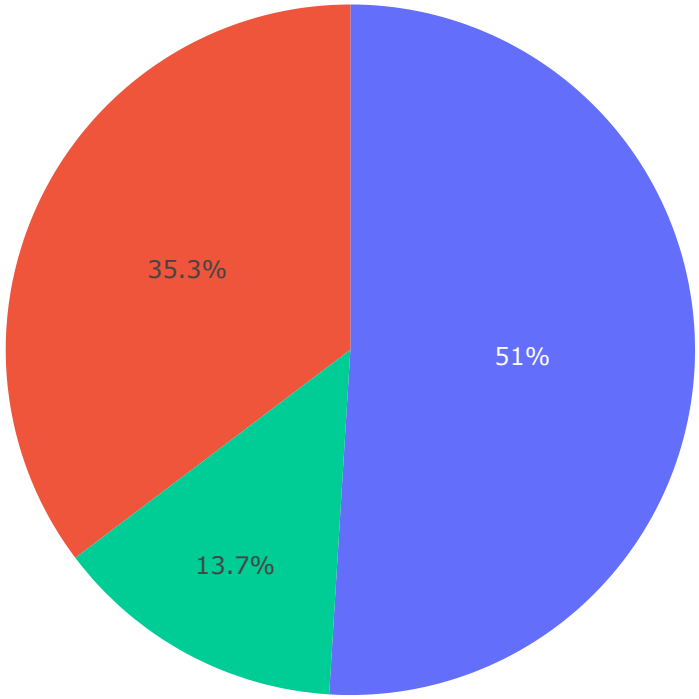
```
[12, 49, 53]
```



In [41]:
```
data_Canada=data[data['users_location']=='Canada']
data_Canada
Neutral = len(data_Canada[data_Canada['Sentiment']=='Neutral'])
Negative = len(data_Canada[data_Canada['Sentiment']=='Negative'])
Positive = len(data_Canada[data_Canada['Sentiment']=='Positive'])
labels = ['Negative','Positive','Neutral']
values = [Negative,Positive,Neutral]

print(values)
import plotly.graph_objects as go
fig = go.Figure(data=[go.Pie(labels=labels,values=values)])
fig.show()
```

```
[14, 52, 36]
```



In [42]:

```
data_Europe=data[data['users_location']=='Europe']
data_Europe
Neutral = len(data_Europe[data_Europe['Sentiment']=='Neutral'])
Negative = len(data_Europe[data_Europe['Sentiment']=='Negative'])
Positive = len(data_Europe[data_Europe['Sentiment']=='Positive'])
labels = ['Negative','Positive','Neutral']
values = [Negative,Positive,Neutral]

print(values)
import plotly.graph_objects as go
fig = go.Figure(data=[go.Pie(labels=labels,values=values)])
fig.show()
```

```
[7, 32, 38]
```

In [43]:
```
data_UAE=data[data['users_location']=='UAE']
data_UAE
Neutral = len(data_UAE[data_UAE['Sentiment']=='Neutral'])
Negative = len(data_UAE[data_UAE['Sentiment']=='Negative'])
Positive = len(data_UAE[data_UAE['Sentiment']=='Positive'])
labels = ['Negative','Positive','Neutral']
values = [Negative,Positive,Neutral]

print(values)
import plotly.graph_objects as go
fig = go.Figure(data=[go.Pie(labels=labels,values=values)])
fig.show()
```

```
[0, 26, 17]
```