

News_Mood.py

The code below examines world mood according to the twitter profile of four five agencies: BBC News(World), CBS News, Fox News, and The New York Times. A sentiment analysis was completed on the 100 most recent tweets. Results are presented in a scatter plot and bar graph. Libraries used to complete this analysis include: Matplotlib, Pandas, Tweepy, VADER Sentiment Analysis, and Seaborn.

Set Up

```
In [2]: # Dependencies
%matplotlib inline
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import json
import tweepy
import time
import warnings
warnings.filterwarnings("ignore")
```

```
In [3]: # API keys
import config
```

```
In [4]: # Initialize Sentiment Analyzer
from textblob import TextBlob
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
analyzer = SentimentIntensityAnalyzer()
```

```
In [5]: # Twitter API Keys
consumer_key = config.consumer_key
consumer_secret = config.consumer_secret
access_token = config.access_token
access_token_secret = config.access_token_secret
```

```
In [6]: # Setup Tweepy API Authentication
auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)
api = tweepy.API(auth, parser=tweepy.parsers.JSONParser())

# Select News Sources (Twitter Accounts)
news_source = ["FoxNews", "CNN", "BBCWorld", "CBSNews", "nytimes"]
```

Compile Tweets

```
In [7]: # Create a generic dictionary for holding all tweet information
tweet_data = {
    "tweet_source": [],
    "tweet_text": [],
    "tweet_date": [],
    "tweet_vader_score": [],
    "tweet_neg_score": [],
    "tweet_pos_score": [],
    "tweet_neu_score": []
}
```

```
In [8]: # Grab 100 tweets from each site (total 500)
for x in range(5):

    # Loop through all news sources
    for source in news_source:

        # Grab the tweets
        tweets = api.user_timeline(source, page=x)

        # For each tweet store it into the dictionary
        for tweet in tweets:

            # All data is grabbed from the JSON returned by Twitter
            tweet_data["tweet_source"].append(tweet["user"]["name"])
            tweet_data["tweet_text"].append(tweet["text"])
            tweet_data["tweet_date"].append(tweet["created_at"])

            # Run sentiment analysis on each tweet using Vader
            tweet_data["tweet_vader_score"].append(analyzer.polarity_scores(tweet["text"])["compound"])
            tweet_data["tweet_pos_score"].append(analyzer.polarity_scores(tweet["text"])["pos"])
            tweet_data["tweet_neu_score"].append(analyzer.polarity_scores(tweet["text"])["neu"])
            tweet_data["tweet_neg_score"].append(analyzer.polarity_scores(tweet["text"])["neg"])
```

```
In [9]: # Store the final contents into a DataFrame
tweet_df = pd.DataFrame(tweet_data, columns=["tweet_source",
                                             "tweet_text",
                                             "tweet_date",
                                             "tweet_vader_score",
                                             "tweet_pos_score",
                                             "tweet_neu_score",
                                             "tweet_neg_score"])


# Export to CSV
file_name = str(time.strftime("%m-%d-%y")) + "-tweets.csv"
tweet_df.to_csv("analysis/" + file_name, encoding="utf-8")

tweet_df
```

Out[9]:

| | tweet_source | tweet_text | tweet_date | tweet_vader_score | tweet_pos_sc |
|---|--------------|---|--------------------------------|-------------------|--------------|
| 0 | Fox News | Bernie Sanders' 'Medicare for all' bill estima... | Mon Jul 30 13:32:42 +0000 2018 | 0.0000 | 0.000 |
| 1 | Fox News | MH370 report says someone deliberately 'manipu... | Mon Jul 30 13:19:03 +0000 2018 | 0.0000 | 0.000 |
| 2 | Fox News | New York brothers wake up to find intruder, be... | Mon Jul 30 13:17:40 +0000 2018 | -0.8271 | 0.000 |
| 3 | Fox News | India could strip 4 million of citizenship in ... | Mon Jul 30 13:04:45 +0000 2018 | 0.0000 | 0.000 |
| 4 | Fox News | Report: US Therapists See Increase in Patients... | Mon Jul 30 12:47:12 +0000 2018 | -0.2732 | 0.146 |
| 5 | Fox News | David Logan: "We are hopeful that we can have ... | Mon Jul 30 12:41:47 +0000 2018 | 0.7096 | 0.244 |
| 6 | Fox News | Meghan Markle's dad slammed by Chrissy Teigen ... | Mon Jul 30 12:29:43 +0000 2018 | 0.0000 | 0.000 |
| 7 | Fox News | Rudy Giuliani: "You could investigate an innoc... | Mon Jul 30 12:04:00 +0000 2018 | 0.3400 | 0.211 |
| 8 | Fox News | Rudy Giuliani: "Everything that's been release... | Mon Jul 30 12:01:00 +0000 2018 | 0.4005 | 0.130 |
| 9 | Fox News | Confederate monuments report raises idea of ch... | Mon Jul 30 11:54:19 +0000 2018 | 0.0000 | 0.000 |

| | tweet_source | tweet_text | tweet_date | tweet_vader_score | tweet_pos_sc |
|----|--------------|---|--------------------------------|-------------------|--------------|
| 10 | Fox News | Chick-fil-A's Canadian expansion sparks pro-LG... | Mon Jul 30 11:51:13 +0000 2018 | -0.2263 | 0.000 |
| 11 | Fox News | OPINION: @GiannoCaldwell: Chicago's gun violen... | Mon Jul 30 11:29:10 +0000 2018 | -0.5859 | 0.141 |
| 12 | Fox News | .@newtgingrich: "His policies are working." ht... | Mon Jul 30 11:19:42 +0000 2018 | 0.0000 | 0.000 |
| 13 | Fox News | .@newtgingrich: "Is it more radical to want to... | Mon Jul 30 11:16:21 +0000 2018 | 0.2795 | 0.135 |
| 14 | Fox News | Ginsburg hopes to serve 'at least' five more y... | Mon Jul 30 10:44:29 +0000 2018 | 0.7693 | 0.377 |
| 15 | Fox News | .@dbongino: "Tax reform was just monumental, w... | Mon Jul 30 10:39:15 +0000 2018 | 0.0000 | 0.000 |
| 16 | Fox News | .@kayleighmcenany: "There's this very disturbi... | Mon Jul 30 10:28:41 +0000 2018 | -0.2247 | 0.156 |
| 17 | Fox News | Martina White: "It's truly shocking that the m... | Mon Jul 30 10:21:47 +0000 2018 | 0.3400 | 0.206 |
| 18 | Fox News | Prankster puts Putin's portrait in Trump's pla... | Mon Jul 30 10:14:41 +0000 2018 | 0.0000 | 0.000 |
| 19 | Fox News | .@CarleyShimkus is live with the 'Fox News Upd... | Mon Jul 30 10:10:32 +0000 2018 | 0.0000 | 0.000 |

| | tweet_source | tweet_text | tweet_date | tweet_vader_score | tweet_pos_sc |
|-----|--------------|---|--------------------------------|-------------------|--------------|
| 20 | CNN | Twitter has hired researchers to examine the h... | Mon Jul 30 13:36:39 +0000 2018 | 0.0000 | 0.000 |
| 21 | CNN |  RECALL ALERT: \nA Chinese company is recalli... | Mon Jul 30 13:26:12 +0000 2018 | 0.4466 | 0.147 |
| 22 | CNN | RT @CNNPolitics: Giuliani: "I don't even know ... | Mon Jul 30 13:16:22 +0000 2018 | -0.8481 | 0.000 |
| 23 | CNN | President Donald Trump's lawyer, Rudy Giuliani... | Mon Jul 30 13:16:03 +0000 2018 | -0.6664 | 0.000 |
| 24 | CNN | In a previously undisclosed TSA program, feder... | Mon Jul 30 13:05:41 +0000 2018 | 0.1695 | 0.094 |
| 25 | CNN | The best travel photos of 2018 (so far) https:... | Mon Jul 30 12:55:11 +0000 2018 | 0.6369 | 0.318 |
| 26 | CNN | 'Rest in Power: The Trayvon Martin Story' lead... | Mon Jul 30 12:42:57 +0000 2018 | 0.0000 | 0.000 |
| 27 | CNN | Top Koch network official says, "The divisiven... | Mon Jul 30 12:32:44 +0000 2018 | 0.2023 | 0.107 |
| 28 | CNN | President Donald Trump's re-election campaign ... | Mon Jul 30 12:20:17 +0000 2018 | -0.4404 | 0.000 |
| 29 | CNN | Alcohol in breast milk may lead to lower cogni... | Mon Jul 30 12:08:02 +0000 2018 | -0.2960 | 0.000 |
| ... | ... | ... | ... | ... | ... |

| | tweet_source | tweet_text | tweet_date | tweet_vader_score | tweet_pos_sc |
|-----|--------------|---|--------------------------------|-------------------|--------------|
| 470 | CBS News | WATCH: California officials say containment ex... | Sun Jul 29 21:18:46 +0000 2018 | -0.3818 | 0.136 |
| 471 | CBS News | WATCH: The Carr Fire is the largest of 17 blaz... | Sun Jul 29 21:11:58 +0000 2018 | -0.5574 | 0.000 |
| 472 | CBS News | Fans were "Crazy in Love" with Michelle and Ba... | Sun Jul 29 20:39:46 +0000 2018 | 0.4215 | 0.171 |
| 473 | CBS News | Imagine your dog eating one month's rent. Now,... | Sun Jul 29 20:19:47 +0000 2018 | 0.0000 | 0.000 |
| 474 | CBS News | Eco-tourism expeditions to places like the Gal... | Sun Jul 29 20:00:42 +0000 2018 | 0.6124 | 0.250 |
| 475 | CBS News | Common wisdom about robots has long held that ... | Sun Jul 29 19:40:01 +0000 2018 | 0.3400 | 0.152 |
| 476 | CBS News | Nearly all strong supporters of President Trum... | Sun Jul 29 19:20:57 +0000 2018 | 0.8591 | 0.373 |
| 477 | CBS News | Fire officials in Greece raised the death toll... | Sun Jul 29 19:00:53 +0000 2018 | -0.8519 | 0.000 |
| 478 | CBS News | Authorities search for two suspects in New Orl... | Sun Jul 29 18:40:37 +0000 2018 | -0.8555 | 0.000 |
| 479 | CBS News | Sen. Ron Johnson says the Trump administration... | Sun Jul 29 18:20:31 +0000 2018 | -0.5859 | 0.000 |

| | tweet_source | tweet_text | tweet_date | tweet_vader_score | tweet_pos_sc |
|-----|--------------------|---|--------------------------------|-------------------|--------------|
| 480 | The New York Times | This stacked shortcake gives a classic summer ... | Sun Jul 29 23:12:58 +0000 2018 | 0.0000 | 0.000 |
| 481 | The New York Times | A renowned musician hated the music played at ... | Sun Jul 29 22:58:06 +0000 2018 | 0.4902 | 0.283 |
| 482 | The New York Times | Billy Joel recently celebrated his 100th show ... | Sun Jul 29 22:39:39 +0000 2018 | 0.5719 | 0.163 |
| 483 | The New York Times | The Enthusiast: In Praise of Elizabeth Hardwic... | Sun Jul 29 22:21:33 +0000 2018 | 0.7269 | 0.504 |
| 484 | The New York Times | Sulzberger recalled telling Trump that newspaper... | Sun Jul 29 22:16:42 +0000 2018 | 0.0000 | 0.000 |
| 485 | The New York Times | MoviePass, the subscription-based movie ticket... | Sun Jul 29 22:00:09 +0000 2018 | -0.4588 | 0.000 |
| 486 | The New York Times | Future of MoviePass Is Cast in Doubt After Ser... | Sun Jul 29 21:52:35 +0000 2018 | -0.3612 | 0.000 |
| 487 | The New York Times | With their eyes on blockchain jobs and revenue... | Sun Jul 29 21:45:09 +0000 2018 | 0.0000 | 0.000 |
| 488 | The New York Times | Imran Khan brings more star power and mystique... | Sun Jul 29 21:30:07 +0000 2018 | 0.5994 | 0.214 |
| 489 | The New York Times | Can you remove a color from an entire city? Th... | Sun Jul 29 21:15:04 +0000 2018 | 0.0000 | 0.000 |

| | tweet_source | tweet_text | tweet_date | tweet_vader_score | tweet_pos_sc |
|-----|--------------------|---|--------------------------------|-------------------|--------------|
| 490 | The New York Times | Sulzberger recalled telling Trump that newspaper... | Sun Jul 29 21:10:26 +0000 2018 | 0.0000 | 0.000 |
| 491 | The New York Times | "We completely regret what happened," a spokes... | Sun Jul 29 21:00:07 +0000 2018 | -0.8221 | 0.000 |
| 492 | The New York Times | In Opinion\n\nContributing opinion writer Sand... | Sun Jul 29 20:45:08 +0000 2018 | -0.4678 | 0.000 |
| 493 | The New York Times | Here are some ideas for what to cook this week... | Sun Jul 29 20:30:07 +0000 2018 | 0.0000 | 0.000 |
| 494 | The New York Times | Living in a lighthouse was a childhood dream f... | Sun Jul 29 20:15:05 +0000 2018 | 0.2500 | 0.095 |
| 495 | The New York Times | Here are the week's top stories, and a look ah... | Sun Jul 29 20:00:13 +0000 2018 | 0.2023 | 0.167 |
| 496 | The New York Times | @Mikelsaac https://t.co/7vUC6VzNos | Sun Jul 29 19:54:09 +0000 2018 | 0.0000 | 0.000 |
| 497 | The New York Times | Want to feel old? Tom Cruise has played Ethan ... | Sun Jul 29 19:45:07 +0000 2018 | 0.4019 | 0.171 |
| 498 | The New York Times | RT @MarkLandler: Times publisher A.G. Sulzberg... | Sun Jul 29 19:36:43 +0000 2018 | 0.0000 | 0.000 |
| 499 | The New York Times | In Opinion \n\nEditorial board member Brent St... | Sun Jul 29 19:30:08 +0000 2018 | -0.1280 | 0.137 |

500 rows × 7 columns

```
In [10]: # Count the total number of tweets  
tweet_df.count()
```

```
Out[10]: tweet_source      500  
         tweet_text       500  
         tweet_date       500  
         tweet_vader_score 500  
         tweet_pos_score   500  
         tweet_neu_score   500  
         tweet_neg_score   500  
         dtype: int64
```

```
In [11]: # Obtain the source names for reference  
tweet_df["tweet_source"].unique()
```

```
Out[11]: array(['Fox News', 'CNN', 'BBC News (World)', 'CBS News',  
               'The New York Times'], dtype=object)
```

```
In [12]: # Convert dates (currently strings) into datetimes
tweet_df["tweet_date"] = pd.to_datetime(tweet_df["tweet_date"])

# Sort the dataframe by date
tweet_df.sort_values("tweet_date", inplace=True)
tweet_df.reset_index(drop=True, inplace=True)

# Preview the data to confirm data is sorted
tweet_df.head()
```

Out[12]:

| | tweet_source | tweet_text | tweet_date | tweet_vader_score | tweet_pos_score | tweet_ne |
|---|------------------|---|---------------------|-------------------|-----------------|----------|
| 0 | BBC News (World) | Trump 'open' to Moscow visit after Putin exten... | 2018-07-27 17:14:49 | 0.2732 | 0.279 | 0.721 |
| 1 | BBC News (World) | Killer whale spotted pushing dead calf for two... | 2018-07-27 18:17:35 | -0.8625 | 0.000 | 0.482 |
| 2 | BBC News (World) | Pakistan election: Rival parties reject result... | 2018-07-27 18:20:26 | 0.0000 | 0.175 | 0.649 |
| 3 | BBC News (World) | In pictures: Blood moon around the world https... | 2018-07-27 19:40:48 | 0.0000 | 0.000 | 1.000 |
| 4 | BBC News (World) | Les Moonves: CBS investigates report of 'misco... | 2018-07-27 20:21:01 | 0.0000 | 0.000 | 1.000 |

Sentiment Scatter Plot

```

In [13]: # Build scatter plot for tracking tweet polarity by tweet history
# Note how a few data munging tricks were used to obtain (-100 -> 0 tick marks)
plt.scatter(np.arange(-len(tweet_df[tweet_df["tweet_source"] == "BBC News (World)"]), 0, 1),
            tweet_df[tweet_df["tweet_source"] == "BBC News (World)"]["tweet_vader_score"],
            edgecolor="black", linewidths=1, marker="o", color="skyblue", s=75,
            alpha=0.8, label="BBC")

plt.scatter(np.arange(-len(tweet_df[tweet_df["tweet_source"] == "CBS News"]), 0, 1),
            tweet_df[tweet_df["tweet_source"] == "CBS News"]["tweet_vader_score"],
            edgecolor="black", linewidths=1, marker="o", color="green", s=75,
            alpha=0.8, label="CBS")

plt.scatter(np.arange(-len(tweet_df[tweet_df["tweet_source"] == "CNN"]), 0, 1),
            tweet_df[tweet_df["tweet_source"] == "CNN"]["tweet_vader_score"],
            edgecolor="black", linewidths=1, marker="o", color="red", s=75,
            alpha=0.8, label="CNN")

plt.scatter(np.arange(-len(tweet_df[tweet_df["tweet_source"] == "Fox News"]), 0, 1),
            tweet_df[tweet_df["tweet_source"] == "Fox News"]["tweet_vader_score"],
            edgecolor="black", linewidths=1, marker="o", color="b", s=75,
            alpha=0.8, label="Fox")

plt.scatter(np.arange(-len(tweet_df[tweet_df["tweet_source"] == "The New York Times"]), 0, 1),
            tweet_df[tweet_df["tweet_source"] == "The New York Times"]["tweet_vader_score"],
            edgecolor="black", linewidths=1, marker="o", color="gold", s=75,
            alpha=0.8, label="New York Times")

# Incorporate the other graph properties
plt.title("Sentiment Analysis of Media Tweets (%s)" % time.strftime("%x"))
plt.ylabel("Tweet Polarity")
plt.xlabel("Tweets Ago")
plt.xlim([-105, 5])
plt.xticks([-100, -80, -60, -40, -20, 0], [100, 80, 60, 40, 20, 0])
plt.ylim([-1.05, 1.05])
plt.grid(True)

# Create a Legend
lgnd = plt.legend(fontsize="small", mode="Expanded",
                 numpoints=1, scatterpoints=1,
                 loc="upper left", bbox_to_anchor=(1,1), title="Media Sources",
                 labelspace=0.5)

# Save the figure (and account for the legend being outside the plot when saving)

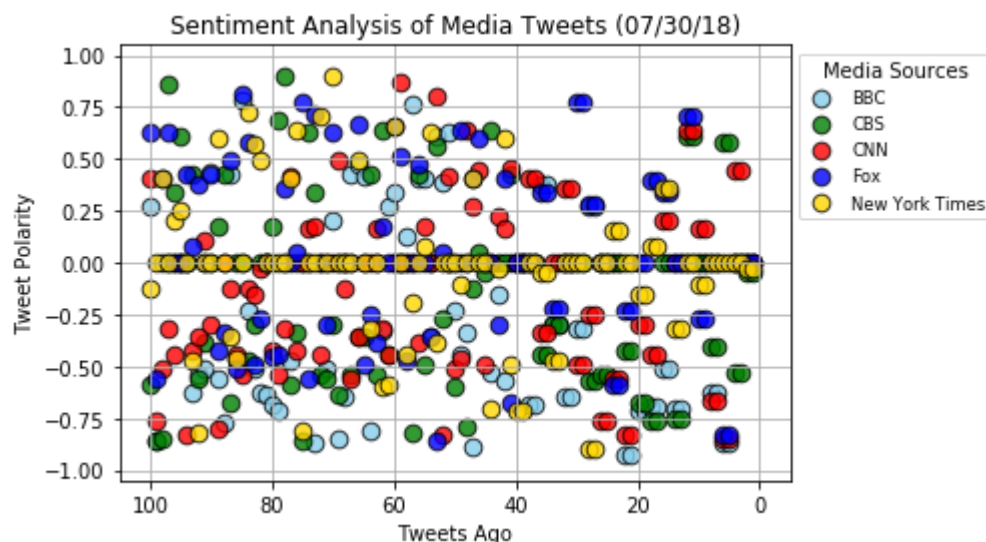
```

```

file_name = str(time.strftime("%m-%d-%y")) + "-Fig1.png"
plt.savefig("analysis/" + file_name, bbox_extra_artists=(lgnd, ), bbox_inches=
'tight')

# Show plot
plt.show()

```



Overall Sentiment Bar Graph

```

In [14]: # Average all polarities by news source
tweet_df_polarity = tweet_df.groupby(["tweet_source"]).mean()["tweet_vader_score"]

# View the polarities
pd.DataFrame(tweet_df_polarity)

```

Out[14]:

| | tweet_vader_score |
|--------------------|-------------------|
| tweet_source | |
| BBC News (World) | -0.209864 |
| CBS News | -0.136684 |
| CNN | -0.129184 |
| Fox News | 0.052225 |
| The New York Times | -0.024634 |

```

In [15]: # Store all polarities in a tuple
tweets_polarity = (tweet_df_polarity["BBC News (World)"],
                    tweet_df_polarity["CBS News"],
                    tweet_df_polarity["CNN"],
                    tweet_df_polarity["Fox News"],
                    tweet_df_polarity["The New York Times"])

# Generate bars for each news source
fig, ax = plt.subplots()
ind = np.arange(len(tweets_polarity))
width = 1
rect1 = ax.bar(ind[0], tweets_polarity[0], width, color="skyblue")
rect2 = ax.bar(ind[1], tweets_polarity[1], width, color="green")
rect3 = ax.bar(ind[2], tweets_polarity[2], width, color="red")
rect4 = ax.bar(ind[3], tweets_polarity[3], width, color='blue')
rect5 = ax.bar(ind[4], tweets_polarity[4], width, color='gold')

# Generate labels for each news source
def autolabelpos(rects):
    # attach some text labels
    for rect in rects:
        height = rect.get_height()
        ax.text(rect.get_x() + rect.get_width()/2., 1*height,
                '+%.2f' % float(height),
                ha='center', va='bottom')

def autolabelneg(rects):
    # attach some text labels
    for rect in rects:
        height = rect.get_height()
        ax.text(rect.get_x() + rect.get_width()/2., -1*height-0.015,
                '-%.2f' % float(height),
                ha='center', va='bottom')

autolabelpos(rect1)
autolabelneg(rect2)
autolabelneg(rect3)
autolabelpos(rect4)
autolabelneg(rect5)

# Orient widths, add labels, tick marks, etc.
ax.set_ylabel("Tweet Polarity")
ax.set_title("Overall Media Sentiment based on Twitter (%s) " % (time.strftime("%x")))
ax.set_xticks(ind + 0.5)
ax.set_xticklabels(("BBC", "CBS", "CNN", "Fox", "NYT"))
ax.set_autoscaley_on(True)
ax.grid(False)

# Save Figure
file_name = str(time.strftime("%m-%d-%y")) + "-Fig2.png"
plt.savefig("analysis/" + file_name, bbox_extra_artists=(lgnd, ), bbox_inches='tight')

# Show Figure
fig.show()

```

