05_AM_Geo_Visualization

July 13, 2021

[1]:
```
import pandas as pd
import numpy as np
from geopy import Nominatim
import math
import time
import plotly.express as px
import plotly.graph_objects as go
```

1 Vaccine sentiment visualization

[2]:
```
geo_tweets = pd.read_csv("../data/processed/geo_vaccine_tweets_with_sentiment.
↪csv", index_col=0)
```

[3]:
```
geo_tweets.head()
```

[3]:
```
                    id                 created_at  \
0  1362299993504145408  2021-02-18 07:16:43+00:00
1  1403714164401401856  2021-06-12 14:01:51+00:00
2  1377527649065594880  2021-04-01 07:45:59+00:00
3  1367118519754612736  2021-03-03 14:23:49+00:00
4  1366219869989969920  2021-03-01 02:52:54+00:00

                                                user  geo  \
0  {'id': 160763636, 'id_str': '160763636', 'name…  NaN
1  {'id': 140847625, 'id_str': '140847625', 'name…  NaN
2  {'id': 47333344, 'id_str': '47333344', 'name':…  NaN
3  {'id': 1252848393182822400, 'id_str': '1252848…  NaN
4  {'id': 842254045, 'id_str': '842254045', 'name…  NaN

                                           full_text  \
0  Hey dear friends my #COVAXIN is done and feel …
1  @BharatBiotech Myself, my lovely wife &amp; re…
2  #VaccineDone. #Covaxin at #KMC Jyoti, #Mangalu…
3  Congratulations @SuchitraElla and their scient…
4  Good morning dear PM Sir..wow! ð  ð ¹ð  Won…

                                            hashtags              user_id  \
```

```
0           ['covaxin', 'covid19', 'vaccinemaitri']            160763636
1                           ['covaxin', 'covid19']            140847625
2    ['vaccinedone', 'covaxin', 'kmc', 'mangaluru']             47333344
3             ['madeinindia', 'covaxin', 'covid19']  1252848393182822400
4  ['covaxin', 'covid19vaccine', 'love', 'respect…            842254045

   PfizerBiontech  SputnikV  Sinopharm  Sinovac  Moderna  AstraZeneca  \
0               0         0          0        0        0            0
1               0         0          0        0        0            0
2               0         0          0        0        0            0
3               0         0          0        0        0            0
4               0         0          0        0        0            0

   Covaxin  JandJ user_location              coordinates  \
0        1      0         India  [22.3511148, 78.6677428]
1        1      0         India  [22.3511148, 78.6677428]
2        1      0         India  [22.3511148, 78.6677428]
3        1      0         India  [22.3511148, 78.6677428]
4        1      0         India  [22.3511148, 78.6677428]

                                              corpus  \
0  hey dear friend covaxin done feel good thanks …
1  lovely wife respectful parent completed taking…
2  vaccinedone covaxin kmc jyoti mangaluru lovely…
3  congratulation scientist making great safe eff…
4  good morning dear pm sir wow wonderful see tak…

                                           sentiment  sentiment_compound
0  {'neg': 0.0, 'neu': 0.311, 'pos': 0.689, 'comp…              0.9876
1  {'neg': 0.129, 'neu': 0.291, 'pos': 0.58, 'com…              0.9771
2  {'neg': 0.0, 'neu': 0.248, 'pos': 0.752, 'comp…              0.9756
3  {'neg': 0.0, 'neu': 0.336, 'pos': 0.664, 'comp…              0.9709
4  {'neg': 0.0, 'neu': 0.508, 'pos': 0.492, 'comp…              0.9652
```

1.1 Average sentiment towards vaccine by country

Extract date from timestamp:

[4]:
```
geo_tweets["created_at"] = pd.to_datetime(geo_tweets["created_at"]).dt.
↪strftime('%Y-%m-%d')
```

[5]:
```
geo_tweets.head()
```

[5]:
```
                   id  created_at  \
0  1362299993504145408  2021-02-18
1  1403714164401401856  2021-06-12
```

```
2  1377527649065594880  2021-04-01
3  1367118519754612736  2021-03-03
4  1366219869989969920  2021-03-01

                                                 user  geo  \
0  {'id': 160763636, 'id_str': '160763636', 'name…  NaN
1  {'id': 140847625, 'id_str': '140847625', 'name…  NaN
2  {'id': 47333344, 'id_str': '47333344', 'name':…  NaN
3  {'id': 1252848393182822400, 'id_str': '1252848…  NaN
4  {'id': 842254045, 'id_str': '842254045', 'name…  NaN

                                           full_text  \
0  Hey dear friends my #COVAXIN is done and feel …
1  @BharatBiotech Myself, my lovely wife &amp;amp; re…
2  #VaccineDone. #Covaxin at #KMC Jyoti, #Mangalu…
3  Congratulations @SuchitraElla and their scient…
4  Good morning dear PM Sir..wow! ð  ð ¹ð  Won…

                                            hashtags              user_id  \
0           ['covaxin', 'covid19', 'vaccinemaitri']            160763636
1                           ['covaxin', 'covid19']            140847625
2    ['vaccinedone', 'covaxin', 'kmc', 'mangaluru']             47333344
3             ['madeinindia', 'covaxin', 'covid19']  1252848393182822400
4  ['covaxin', 'covid19vaccine', 'love', 'respect…            842254045

   PfizerBiontech  SputnikV  Sinopharm  Sinovac  Moderna  AstraZeneca  \
0               0         0          0        0        0            0
1               0         0          0        0        0            0
2               0         0          0        0        0            0
3               0         0          0        0        0            0
4               0         0          0        0        0            0

   Covaxin  JandJ user_location              coordinates  \
0        1      0         India  [22.3511148, 78.6677428]
1        1      0         India  [22.3511148, 78.6677428]
2        1      0         India  [22.3511148, 78.6677428]
3        1      0         India  [22.3511148, 78.6677428]
4        1      0         India  [22.3511148, 78.6677428]

                                              corpus  \
0  hey dear friend covaxin done feel good thanks …
1  lovely wife respectful parent completed taking…
2  vaccinedone covaxin kmc jyoti mangaluru lovely…
3  congratulation scientist making great safe eff…
4  good morning dear pm sir wow wonderful see tak…

                                          sentiment  sentiment_compound
```

```
0  {'neg': 0.0, 'neu': 0.311, 'pos': 0.689, 'comp…              0.9876
1  {'neg': 0.129, 'neu': 0.291, 'pos': 0.58, 'com…              0.9771
2  {'neg': 0.0, 'neu': 0.248, 'pos': 0.752, 'comp…              0.9756
3  {'neg': 0.0, 'neu': 0.336, 'pos': 0.664, 'comp…              0.9709
4  {'neg': 0.0, 'neu': 0.508, 'pos': 0.492, 'comp…              0.9652
```

[6]:
```
#geo_tweets.groupby(["created_at","user_location"])["sentiment_compound"].
↪mean().reset_index()["user_location"].unique()
```

Plot the mean sentiment of a location on a given date:

[7]:
```
fig = px.choropleth(geo_tweets.
↪groupby(["created_at","user_location"])["sentiment_compound"].mean().
↪reset_index(),
                    locations = 'user_location',
                    color="sentiment_compound",
                    locationmode="country names",
                    animation_frame="created_at",
                    color_continuous_scale="sunsetdark",
                    scope="world",
                    range_color=(-1,1),
                    title='Mean sentiment over time by country',
                    height=600
                   )
fig.layout.updatemenus[0].buttons[0].args[1]["frame"]["duration"] = 80

fig.show()
fig.write_html("../reports/figures/VaccineGeo/
↪Mean_Sentiment_By_Country_Over_Time.html")
```

The day norway notified that 23 People died of PfizerBiontech (16. Jan 2021), the worldwide sentiment was negative (-0.27), especially in India, USA and Australia.

[8]:
```
geo_tweets[geo_tweets["created_at"] == "2021-01-16"]["sentiment_compound"].
↪mean()
```

[8]:
```
-0.2708285714285715
```

Sample tweets from that day:

[9]:
```
geo_tweets[geo_tweets["created_at"] == "2021-01-16"]["full_text"].iloc[-2]
```

[9]:
```
'I\'m not on here to question the safety of the #PfizerBioNTech but as someone
who has a mother in a LTC home who\'s tired of hearing "that\'s just" or
"that\'s only" or "less than 1 in a 1000" when ppl speak abt the deaths of the
elderly. That\'s someone\'s mom, dad, grandma/grandpa. https://t.co/qyGIpiusFs'
```

[10]:
```
geo_tweets[(geo_tweets["created_at"] ==␣
↪"2021-01-16")&(geo_tweets["user_location"] == "India")]["full_text"].iloc[0]
```

[10]:
```
"I was totally optimistic and excited about the #VaccinationDrive  in India
until this came up. \n\nGlad that #India didn't approve of the #PfizerBioNTech
#vaccine for emergency use.\n\nNow I am totally uncertain about the #vaccine
https://t.co/44Sr3aI9S9"
```

[11]:
```
geo_tweets[(geo_tweets["created_at"] == "2021-01-16")]["full_text"].iloc[-1]
```

[11]:
```
'U MAY die if u become covid positive. But #PfizerBioNTech shot will DEFINITELY
kill youð\x9f\x98\xadð\x9f\x98\xad Stuck between devil and sea
https://t.co/BSLZSKn31O'
```

the tweets show clearly that people are scared despite knowing that myocarditis/thrombosis happens rarely. The fear, uncertainty and doubt (FUD) had a worldwide impact.

1.2 Sentiment towards vaccine by country

Build Dataframe for sentiment by country and vaccine:

[12]:
```
geo_df = pd.
 ↪DataFrame(columns=["user_location","Overall","PfizerBiontech","SputnikV","Sinopharm","Sinov
```

Get the overall sentiment of a country:

[13]:
```
geo_df.Overall = geo_tweets.sort_values("user_location").
 ↪groupby("user_location")["sentiment_compound"].mean().
 ↪reset_index()["sentiment_compound"]
geo_df.user_location = geo_tweets.sort_values("user_location").
 ↪groupby("user_location")["sentiment_compound"].mean().
 ↪reset_index()["user_location"]
```

Get the sentiment of a country (first loop) towards a vaccine(second loop):

[14]:
```
for i in range(len(geo_df)):
    for vaccine in geo_df.columns[2:]:
        geo_df[vaccine][i] = geo_tweets[(geo_tweets["user_location"] == geo_df.
 ↪user_location[i])&(geo_tweets[vaccine] == 1)].sentiment_compound.mean()
```

```
<ipython-input-14-a52dc8e8f16e>:3: SettingWithCopyWarning:


A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
```

[15]:
```
geo_df.head()
```

[15]:

	user_location	Overall	PfizerBiontech	SputnikV	Sinopharm	Sinovac	\
0	Argentina	0.147267	0.148783	0.137134	-0.0390286	-0.0683	
1	Australia	0.071498	0.0712732	0.217441	0.0619	0.121545	
2	Austria	0.098954	0.0740364	0.0472344	-0.0494	0.245262	
3	Belgium	0.139248	0.017287	0.11808	0.2238	0.437078	
4	Brazil	0.089853	0.5399	0.599733	0.3612	-0.333567	

	Moderna	AstraZeneca	Covaxin	JandJ
0	0.0399286	0	0.3321	NaN
1	0.0242799	0.0973958	0.105186	-0.335033
2	0.143522	0.0988625	0.6249	0
3	0.22886	0.0589133	-0.035	0.2202
4	-0.01742	0.0952875	0.00111818	NaN

Plot with dropdown menu:

[16]:
```
def sentiment_by_country(dataset):

    fig = go.Figure()

    for column in dataset.columns[1:].to_list():
        fig.add_trace(
            go.Bar(
                x = dataset.user_location,
                y = dataset[column],
                name = column
            )
        )

    # Add dropdown
    fig.update_layout(
        updatemenus=[
            dict(
                buttons=list([
                    dict(
                        args=[{'visible':␣
↪[True,True,True,True,True,True,True,True,True,True]},
                                {'title' : 'Mean sentiment by country (All)'}],
                        label="All",
                        method="update"
                    ),
                    dict(
                        args=[{'visible':␣
↪[True,False,False,False,False,False,False,False,False,False]},
                                {'title' : 'Mean sentiment by country (All␣
↪vaccines)'}],
                        label="Overall Sentiment",
```

```
                    method="update"
                ),
                dict(
                    args=[{'visible':␣
↪[False,True,False,False,False,False,False,False,False]},
                          {'title' : 'Mean sentiment by country for␣
↪PfizerBiontech'}],
                    label="PfizerBiontech",
                    method="update"
                ),
                dict(
                    args=[{'visible':␣
↪[False,False,True,False,False,False,False,False,False]},
                          {'title' : 'Mean sentiment by country for␣
↪SputnikV'}],
                    label="SputnikV",
                    method="update"
                ),
                dict(
                    args=[{'visible':␣
↪[False,False,False,True,False,False,False,False,False]},
                          {'title' : 'Mean sentiment by country for␣
↪Sinopharm'}],
                    label="Sinopharm",
                    method="update"
                ),
                dict(
                    args=[{'visible':␣
↪[False,False,False,False,True,False,False,False,False]},
                          {'title' : 'Mean sentiment by country for␣
↪Sinovac'}],
                    label="Sinovac",
                    method="update"
                ),
                dict(
                    args=[{'visible':␣
↪[False,False,False,False,False,True,False,False,False]},
                          {'title' : 'Mean sentiment by country for␣
↪Moderna'}],
                    label="Moderna",
                    method="update"
                ),
                dict(
                    args=[{'visible':␣
↪[False,False,False,False,False,False,True,False,False]},
```

```
                            {'title' : 'Mean sentiment by country for␣
↪AstraZeneca'}],
                        label="AstraZeneca",
                        method="update"
                    ),
                    dict(
                        args=[{'visible':␣
↪[False,False,False,False,False,False,False,True,False]},
                            {'title' : 'Sentiment by country for Covaxin'}],
                        label="Covaxin",
                        method="update"
                    ),
                    dict(
                        args=[{'visible':␣
↪[False,False,False,False,False,False,False,False,True]},
                            {'title' : 'Sentiment by country for JandJ'}],
                        label="JandJ",
                        method="update"
                    )
                ]),
                direction="down",
                pad={"r": 10, "t": 10},
                showactive=True,
                x=1.0,
                xanchor="right",
                y=1.2,
                yanchor="top"
            ),
        ],
        xaxis_title="Country",
        yaxis_title="Sentiment",
        legend_title="Vaccines"
    )

    fig.show()
    fig.write_html("../reports/figures/VaccineGeo/Mean_Sentiment_By_Vaccine.
↪html")
```

[17]:
```
sentiment_by_country(geo_df)
```

---

2 What region talks about what vaccine?

Aggregate the tweet volume by country and vaccine:

[18]:
```
tweet_amount_by_vaccine = geo_tweets.groupby(["user_location"])[geo_tweets.
↪columns[7:-5]].sum()
```

create the bool-List for the dropdown menu:

[19]:
```
button_settings=[
            dict(
                buttons=list([]),
                direction="down",
                pad={"r": 10, "t": 10},
                showactive=True,
                x=1.0,
                xanchor="right",
                y=1.2,
                yanchor="top"
            ),
        ]


for i in range(len(tweet_amount_by_vaccine.index)):
    visible_list = [True if country == tweet_amount_by_vaccine.index[i] else␣
↪False for country in tweet_amount_by_vaccine.index]
    button = dict(
        args = [dict(visible=visible_list),
        dict(title = f"Most talked about Vaccine in {tweet_amount_by_vaccine.
↪index[i]}")],
        label=tweet_amount_by_vaccine.index[i],
        method="update")

    button_settings[0]['buttons'].append(button)
```

[20]:
```
def amount_tweets_by_country(dataset, dropdown_settings):

    fig = go.Figure()

    for country in dataset.index.to_list():
        fig.add_trace(
            go.Pie(labels=dataset[dataset.index==country].columns,␣
↪values=dataset[dataset.index==country].values[0])
        )

    # Add dropdown
    fig.update_layout(
        updatemenus = dropdown_settings,
        xaxis_title="Time",
        yaxis_title="Sentiment",
        legend_title="Vaccines"
```

```
    )

    fig.show()
    fig.write_html("../reports/figures/VaccineGeo/
↪Most_Talked_About_Vaccine_By_Country.html")
```

[21]:
```
amount_tweets_by_country(tweet_amount_by_vaccine, button_settings)
```

Note: Keep in mind that the dataset was english only. Most other countries would probably tweet in their native language, thus the lack of “international” discourse.