

Equipo: UMIZOOMI

Integrantes: Carlos Andres Barredez Rios Alejandro Hernández De la Torre Santiago Orozco Quintero Jorge Yepez Frutos Jorge Gerardo Inglesias Ortiz

```
In [14]: import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns
df = pd.read_csv("covid19_tweets.csv")

In [15]: df.head()
```

	user_name	user_location	user_description	user_created	user_followers	user_friends	user_favourites	user_verified	date	text	hashtags	source	is_retweet
0	Viola	astroworld	wednesday addams as a disney princess keepin i...	2017-05-26 05:46:42	624	950	18775	False	2020-07-25 12:27:21	If I smelled the scent of hand sanitizers toda...	NaN	Twitter for iPhone	False
1	Tom Basile	New York, NY	Husband, Father, Columnist & Commentator. Auth...	2009-04-16 20:06:23	2253	1677	24	True	2020-07-25 12:27:17	Hey @Yankees @YankeesPR and @MLB - wouldn't it...	NaN	Twitter for Android	False
2	Time4fisticuffs	Pewee Valley, KY	#Christian #Catholic #Conservative #Reagan #Re...	2009-02-28 18:57:41	9275	9525	7254	False	2020-07-25 12:27:14	@diane3443 @wdunlap @realDonaldTrump Trump nev...	['COVID19']	Twitter for Android	False
3	ethel mertz	Stuck in the Middle	#Browns #Indians #ClevelandProud #[] #Cavs ...	2019-03-07 01:45:06	197	987	1488	False	2020-07-25 12:27:10	@brookbanktv The one gift #COVID19 has give me...	['COVID19']	Twitter for iPhone	False
4	DIPR-J&K	Jammu and Kashmir	Official Twitter handle of Department of Inf...	2017-02-12 06:45:15	101009	168	101	False	2020-07-25 12:27:08	25 July : Media Bulletin on Novel #CoronaVirus...	['CoronaVirusUpdates', 'COVID19']	Twitter for Android	False

```
In [16]: df.describe()
```

	user_followers	user_friends	user_favourites
count	7.443600e+04	74436.000000	7.443600e+04
mean	1.059513e+05	2154.721170	1.529747e+04
std	8.222900e+05	9365.587474	4.668971e+04
min	0.000000e+00	0.000000	0.000000e+00
25%	1.660000e+02	153.000000	2.200000e+02
50%	9.600000e+02	552.000000	1.927000e+03
75%	5.148000e+03	1780.250000	1.014800e+04
max	1.389284e+07	497363.000000	2.047197e+06

```
In [39]: df["user_verified"].mode()[0]

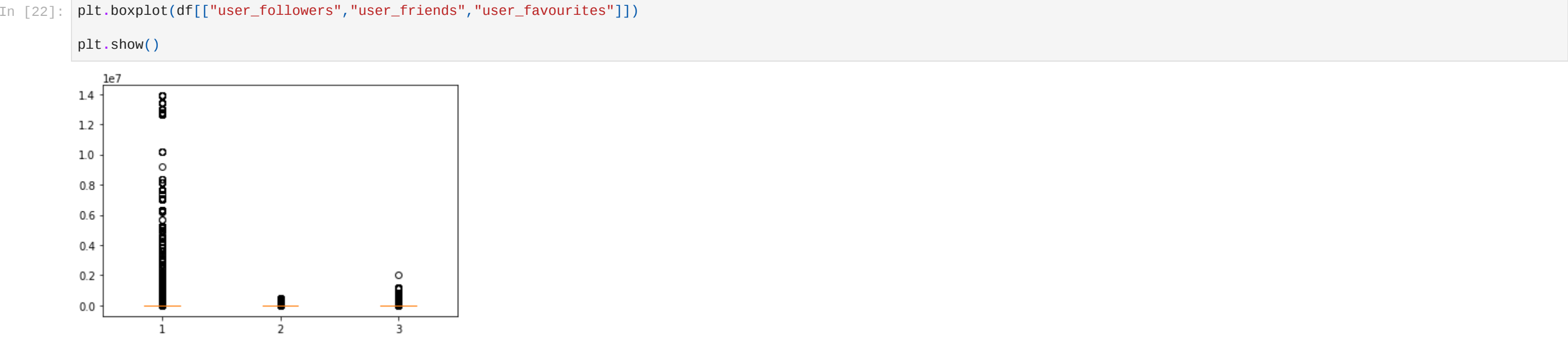
Out[39]: False
```

```
In [44]: df["user_followers"].mode()[0]

Out[44]: 0
```

```
In [43]: df["user_location"].mode()[0]

Out[43]: 'India'
```

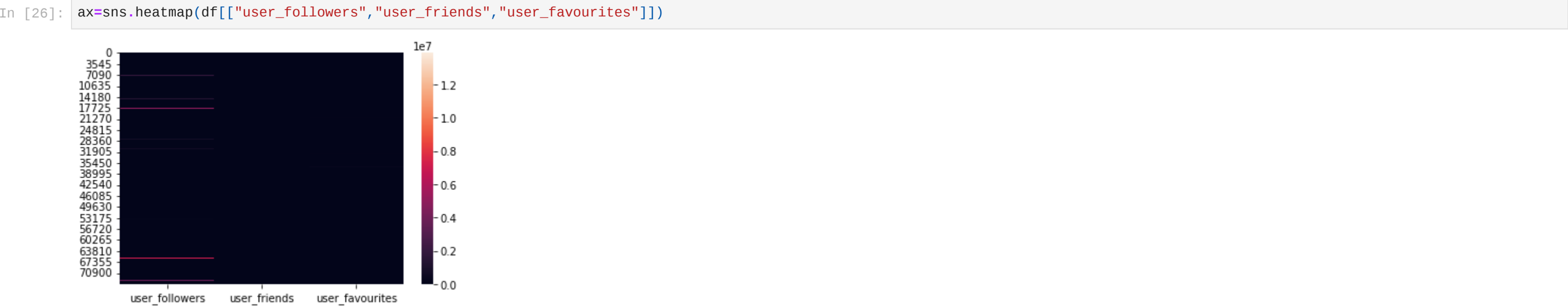


```
In [28]: df.corr(method="kendall")
```

	user_followers	user_friends	user_favourites	user_verified	is_retweet
user_followers	1.000000	0.410663	0.242534	0.415364	NaN
user_friends	0.410663	1.000000	0.439099	-0.003114	NaN
user_favourites	0.242534	0.439099	1.000000	-0.032220	NaN
user_verified	0.415364	-0.003114	-0.032220	1.000000	NaN
is_retweet	NaN	NaN	NaN	NaN	1.0

```
In [24]: df.corr(method="pearson")
```

	user_followers	user_friends	user_favourites	user_verified	is_retweet
user_followers	1.000000	-0.002722	-0.028724	0.322896	NaN
user_friends	-0.002722	1.000000	0.207825	0.013099	NaN
user_favourites	-0.028724	0.207825	1.000000	-0.060316	NaN
user_verified	0.322896	0.013099	-0.060316	1.000000	NaN
is_retweet	NaN	NaN	NaN	NaN	NaN



Mediante el metodo de correlacion podemos sacar diferentes interpreteaciones en la relacion de como estas personas se relacion con las personas que siguen o son seguidos. Si bien las redes sociales son tendenciosas, la relacion de cada uno de los usuarios es importantes y como estos estan clasificados en la red social. Por ejemplo, uno que se puede tomar en cuenta es el usuario vetificado. Estos usuarios suelen ser personas que son artistas, compositores, empresarios, politicos, economistas, etc. Si bien estas figuras publica transmiten cierta forma de veracidad en sus palabras, podemos inferir que la informacion de las personas no verificadas replican la misma informacion. Los usuarios verificados tienden a no seguir a sus propios seguidores, que agregando la mayoria de sus seguidores no son verificados, de la misma manera, no suelen interaccionar con los demas usuarios, mas bien, los seguidores mandan más favoritos. Podriamos visualizar. Otra conclusion que podemos obtener mediante la moda es que la mayoria de los usuarios que comentaron poseen 0 followers, son de India y que no son verificados!. Esto nos podría indicar que el uso de bots o de cuentas automaticas generaban twitts en relacion al covid 19 y su presencia fue mayora usuarios con caracteristicas que indicaran que son cuentas reales. Tambien el tema del pais de procedencia podría indicar que siendo de un pais el cual fue impactado por el covid podría indicar que el requerimiento de bots para cambiar la opinion o dar informacion falsa.