Codestarters

Project 1 - Twitter Sentiment Analysis of the Ultra Music Festival 2019

Daniel Peña
Jose Rojas
Abraham Espinosa
Mark Breidsprecher

April 6th, 2019

Content

Motivation
&
Summary

Motivation & Summary

What is Ultra?

Ultra Music Festival is an annual outdoor electronic music festival that has taken place across three days during the month of March in Miami, Florida. The 2019 edition was held from March 29th through March 31st.

Considering this, we decided to gather information from Twitter to perform a sentiment analysis of the Ultra Music Festival 2019.

We were focused on respond the following questions:

1	What was the popularity & sentiment about the Ultra Music Festival 2019?
2	Identify if in the #Ultra we identified additional hashtags (i.e. #Artists, #Brands, etc.).
3	What was the popularity & sentiment about the additional hashtags identified?

Questions
&
Data Source

Questions & Data Source

Our main concerns to perform the sentiment analysis of the Ultra Music Festival 2019 were related to:

- How we can get the information needed from Twitter?
- What will be the timeframe that we should be analyzing?
- Will we get enough data to perform the analysis?
- What was the sentiment for the Ultra Music Festival?
- Is there any relationship between the artist that participated and the sentiment of the festival?
- What was the sentiment for the artist that participated in the festival?

Data Clean up
&
Exploration

Data Clean up & Exploration

A brief summary of the activities performed to gather the data and start the clean up & exploration process:

- ✓ We got a developer account from Twitter
- ✓ We requested the API keys to get the data
- ✓ We performed advanced searches in Twitter to identify the Ultra Music Festival mentions between March 23th and April 2nd
- ✓ We imported the data sets *(more details will be explained with the Data Analysis walkthrough)*
- ✓ We cleaned and prepared the data for analysis *(more details will be explained with the Data Analysis walkthrough*

Data Analysis

- Manipulate data analysis libraries
- Summarize data
- Build data figures

The results of our analysis are as follows:

Data Analysis

In [518]:
```
print(f"The Dataframe has {len(df_clean)} records")
df_clean.head(5)
```

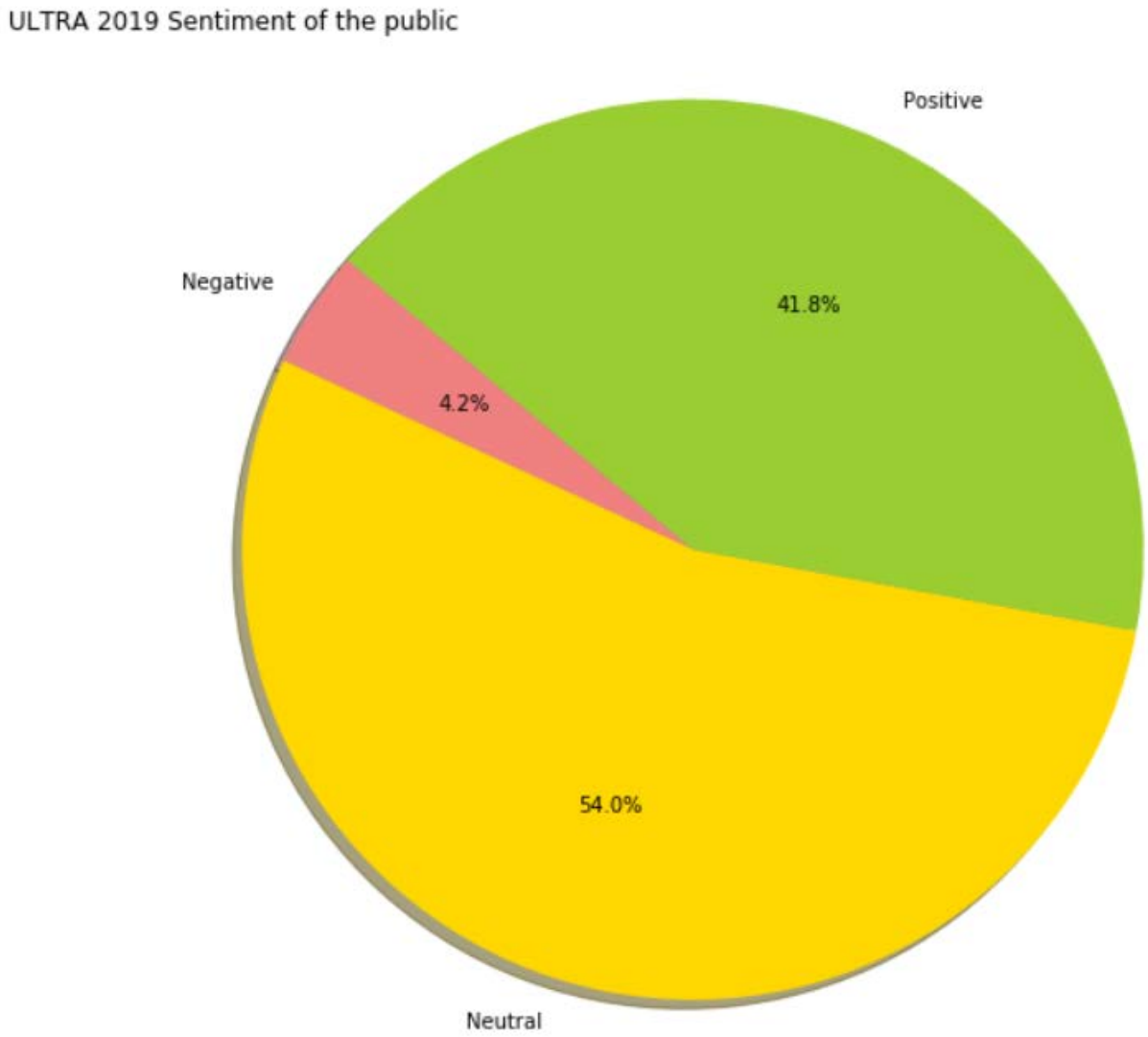
```
The Dataframe has 900 records
```

Out[518]:

	created_at	favorite_count	retweet_count	text	Sentiment	Polarity	Subjectivity
0	Sun Mar 31 07:05:03 +0000 2019	4826	995	Thank you for an incredible Day 2!\r\n\r\nSee ...	Positive	0.500000	0.483333
1	Wed Mar 27 16:23:20 +0000 2019	4135	1017	MIAMI, ARE YOU READY? \r\n\r\nOnly 2 more days ...	Positive	0.233333	0.666667
2	Sun Mar 31 15:46:24 +0000 2019	9947	2379	Shaq was turnt up at the #Ultra2019 in Miami! ...	Neutral	0.000000	0.000000

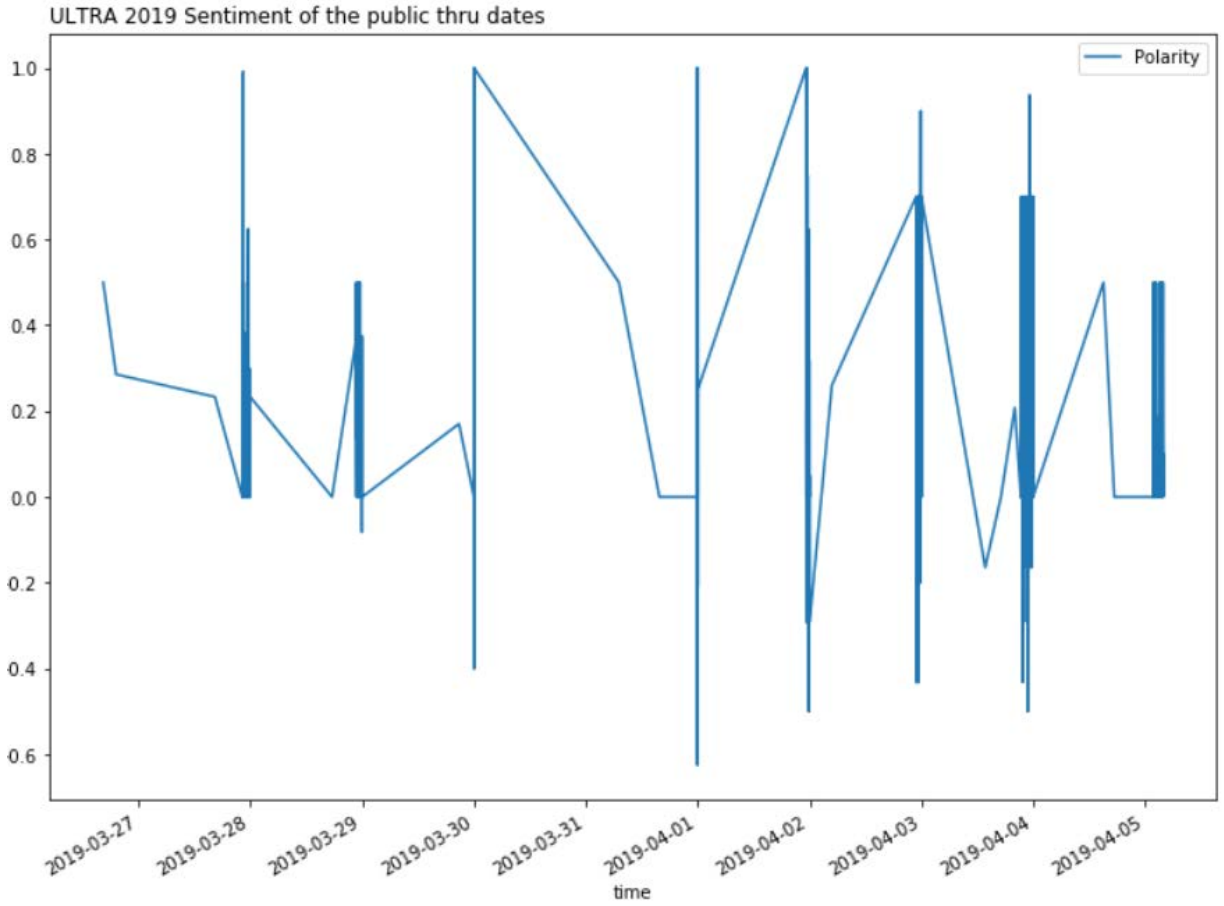
Data Analysis

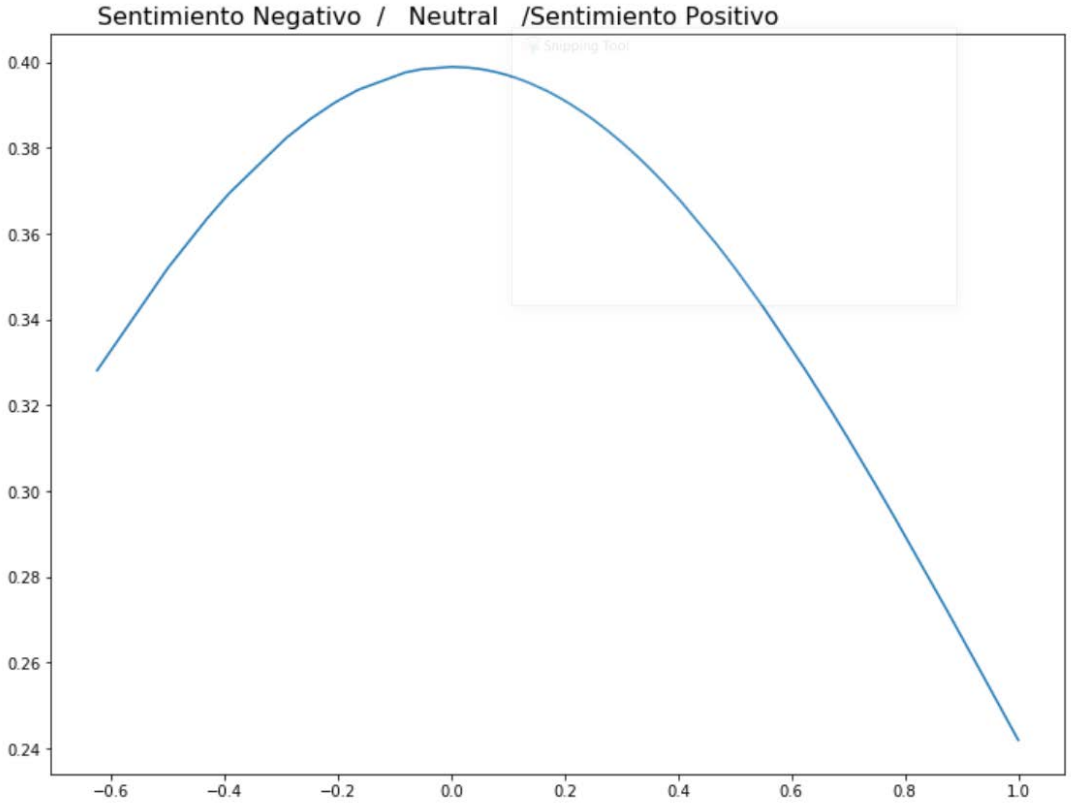
1. The chart below shows the Ultra 2019 sentiment of the public.



Data Analysis

2. The chart below shows the Ultra 2019 sentiment of the public between March 27th and April 5th.



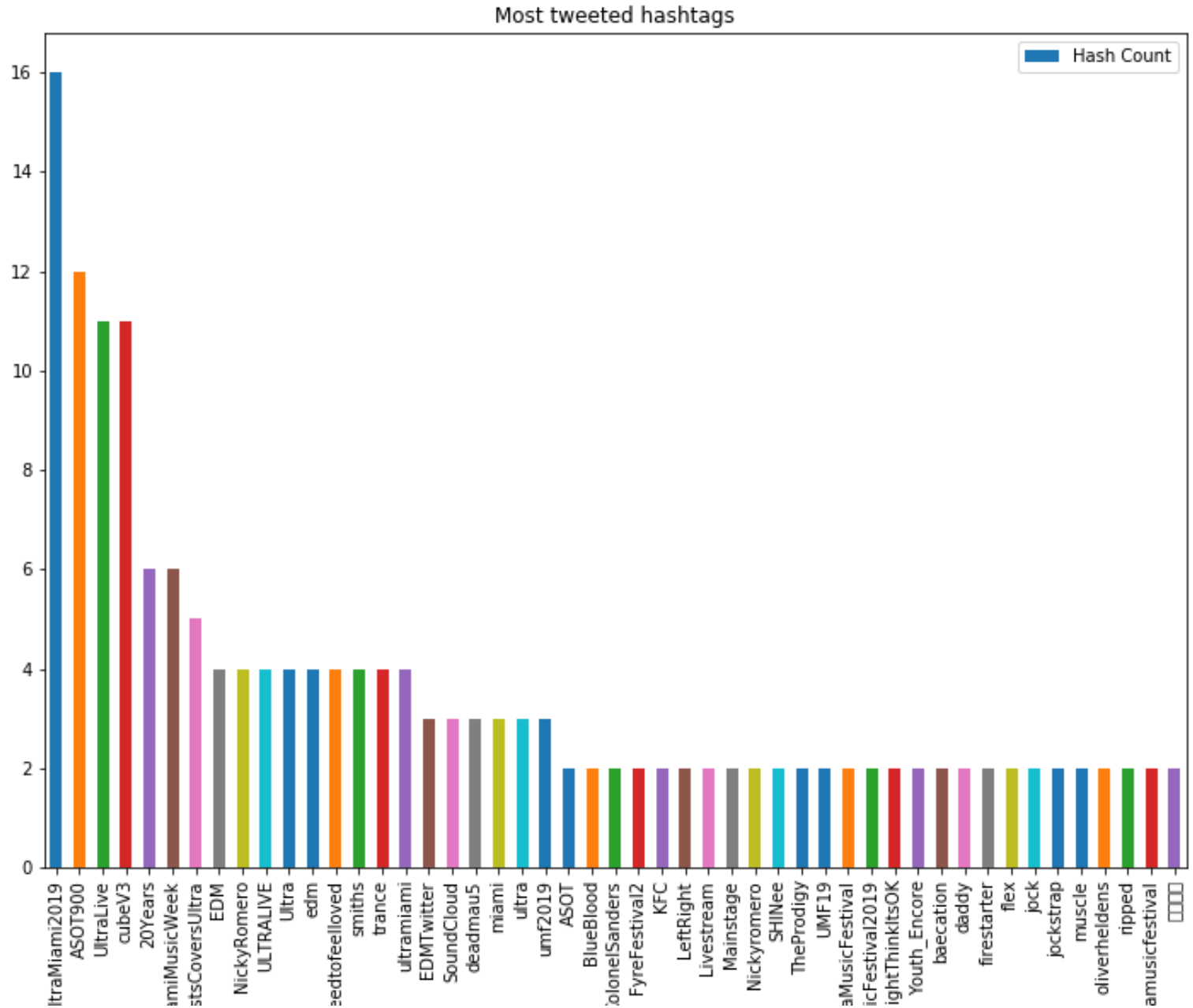


2.1 The chart above shows the Ultra 2019 sentiment of the public thru the period analyzed.

Data Analysis

3. This chart shows the most twitted hashtags:

    1. #UltraMiami2019
    2. #ASOT900
    3. #UltraLive
    4. #cubeV3
    5. #20Years

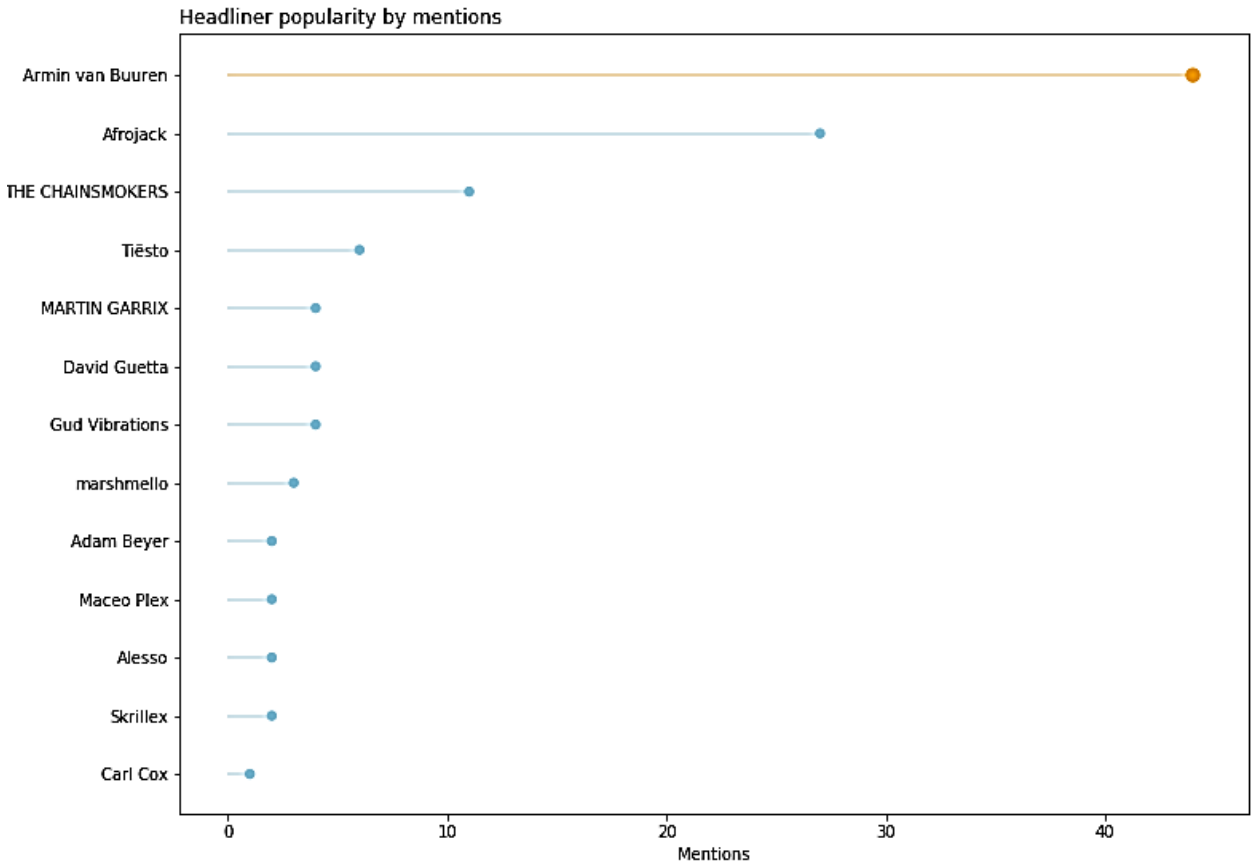


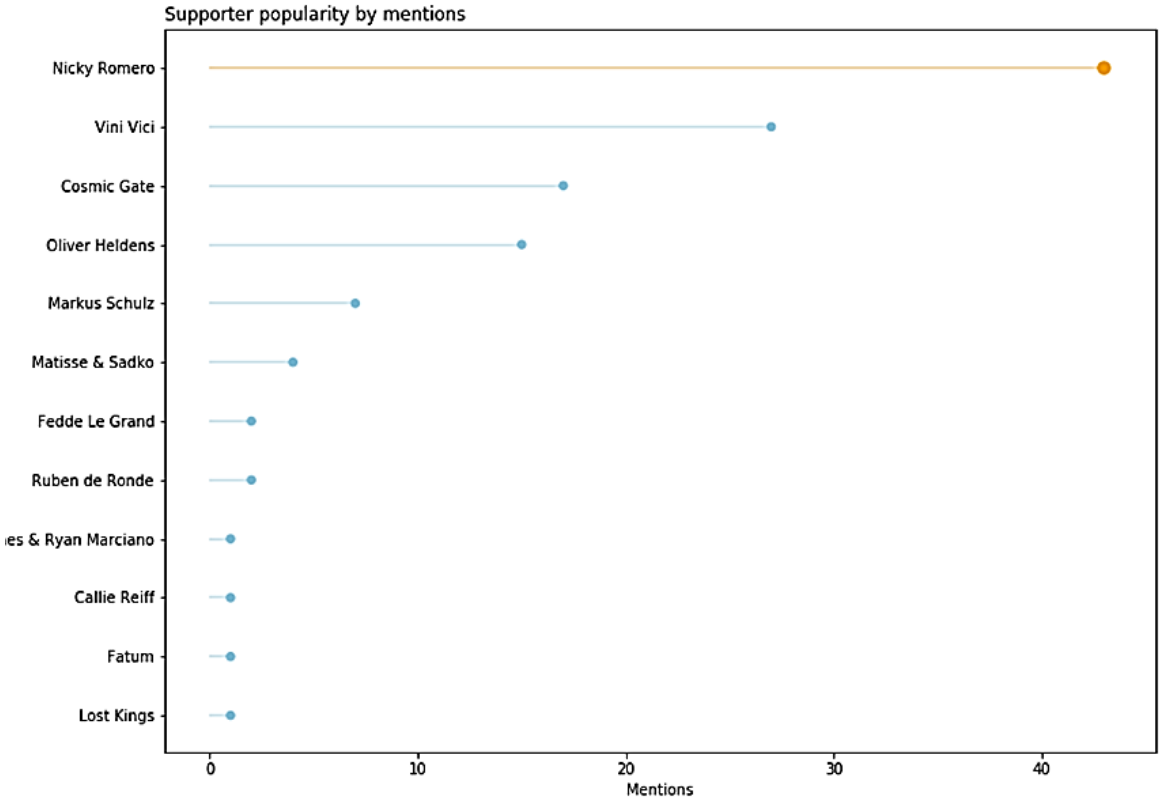
Data Analysis

4. This chart shows the total number of likes by tweet and retweets.

Data Analysis

5. The chart below shows the Headliner popularity by mention.

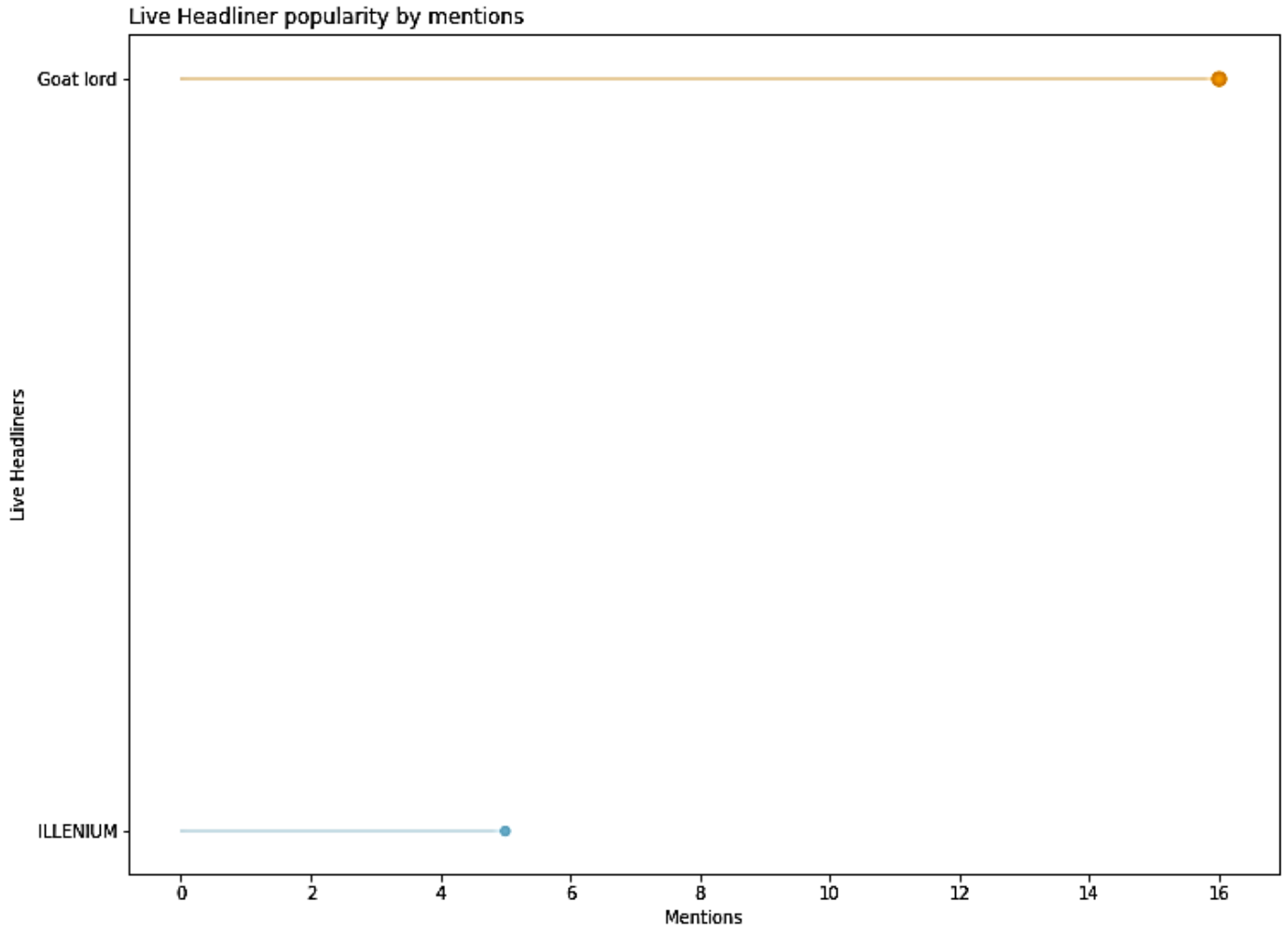




6. The chart above shows the Supporter popularity by mention.

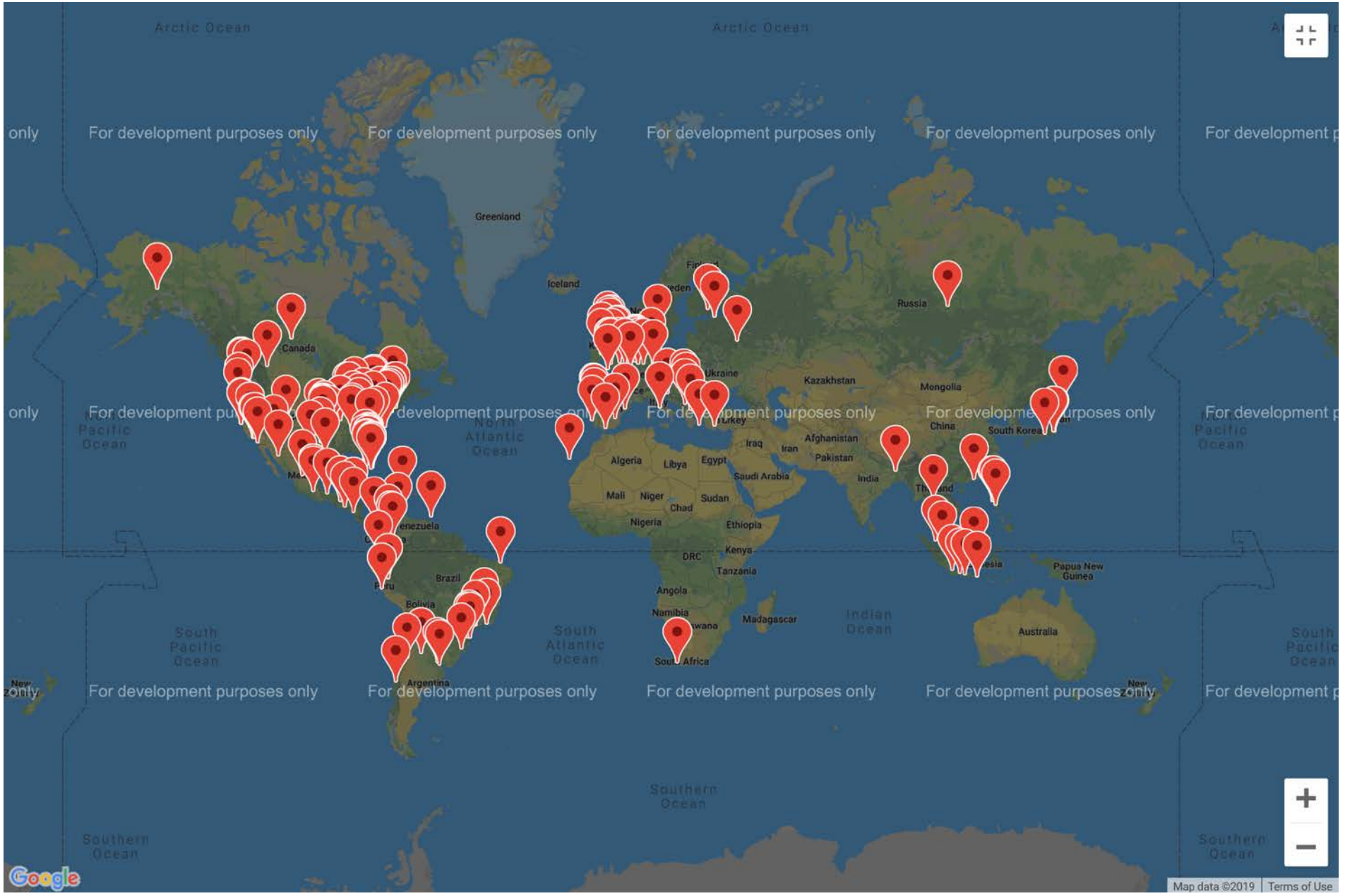
Data Analysis

7. The chart below shows the Ultra 2019 Live Headliner popularity by mention.



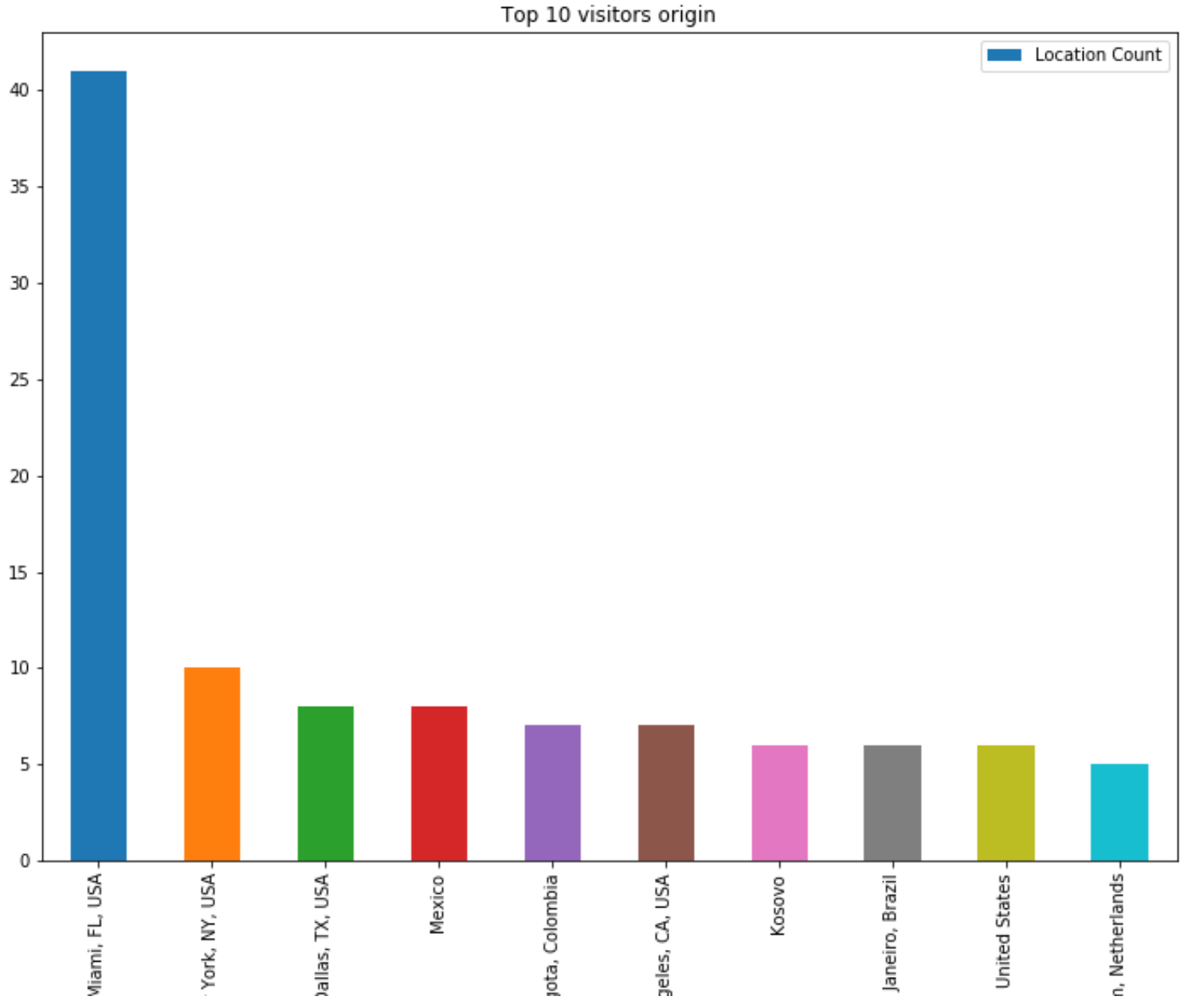
Data Analysis

8. The map shows the location of attendees to the Ultra 2019 .



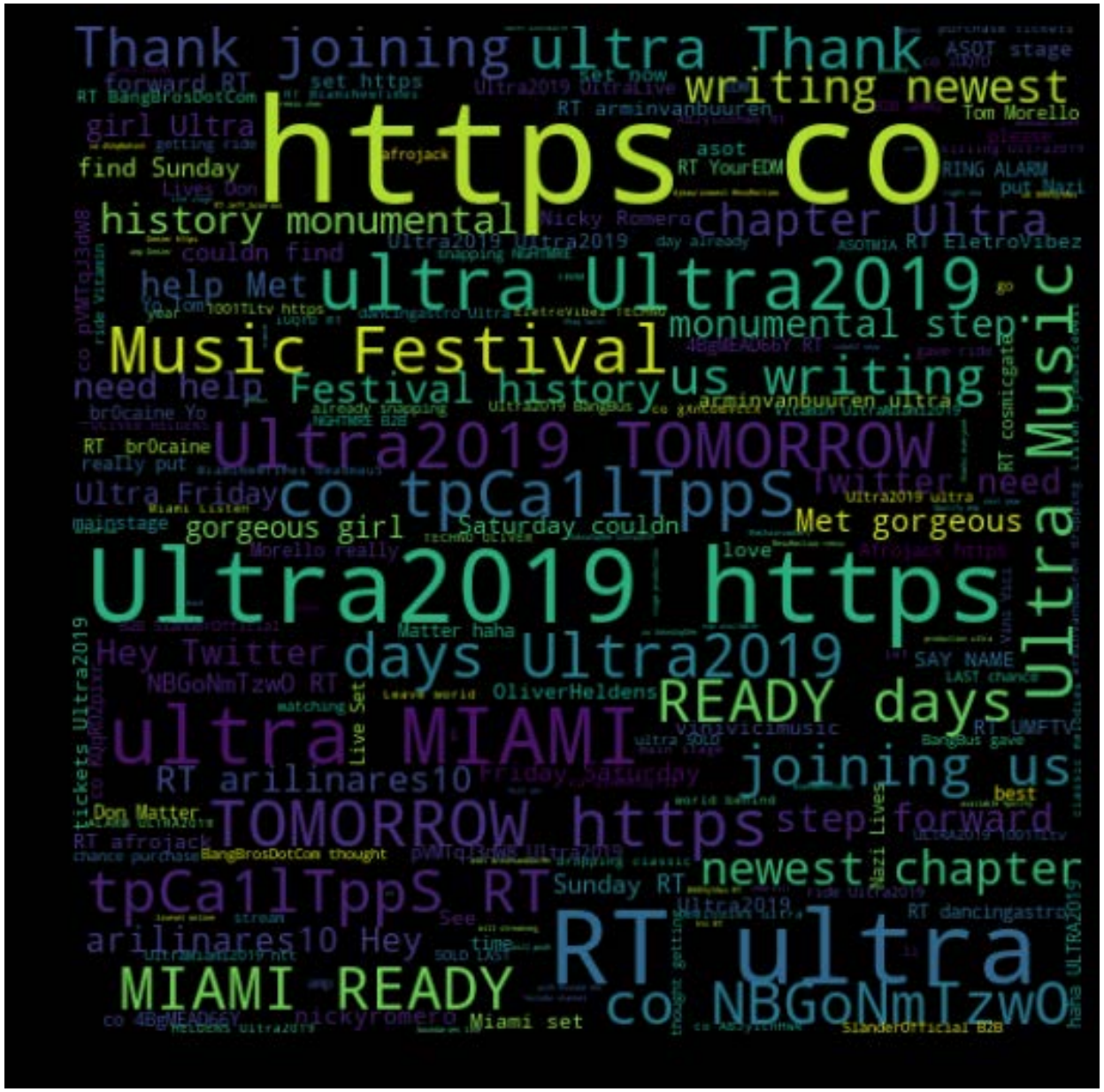
Data Analysis

9. The chart shows the top 10 places of origin.



Data Analysis

10. Finally, we present a cloud collage with the words included in the tweets.



Data Analysis Conclusions

The main conclusions regarding the analysis performed are as follows:

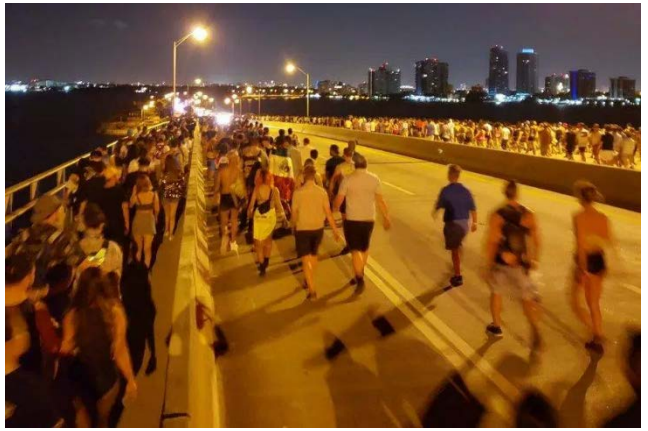
- ✓ We were able to conclude that the most popular artist during the festival was **Armin Van Buuren**. This correlate with the most popular #'s.

- ✓ It was surprising to find a particular event that cause big sensation. **"Shaquile O'Neal dancing with the public"**.

- ✓ From the three different lineups of the event (Headliners, Live Headliners and Supporter DJ's) we were able to identify the most successful artist during the event

Data Analysis Conclusions

The main conclusions regarding the analysis performed are as follows:



- ✓ We were also able to find out a particular event that causes the mood of the attendees to drop. **A problem with the transportation** after the event causes big delays, therefore, attendees heavily twitted their angry on social media.

- ✓ Finally a top ten origin location from the attendees came out from the data. This tell us from where people are coming from to the event.

About the code:

- ✓ We found that open source sentiment analysis libraries, **require big tune ups** and usually perform just OK with regular data. To be able to perform a good sentiment analysis, it is needed to develop and train a dedicated algorithm using meaningful data for the particular analysis.

Questions?

Thank You!

Codestarters – Project 1

Daniel Peña
Jose Rojas
Abraham Espinosa
Mark Breidsprecher

April 6th, 2019