Twitter Emoji Streams: A Time Series Visualization

Stefan Bund, bund@umich.edu

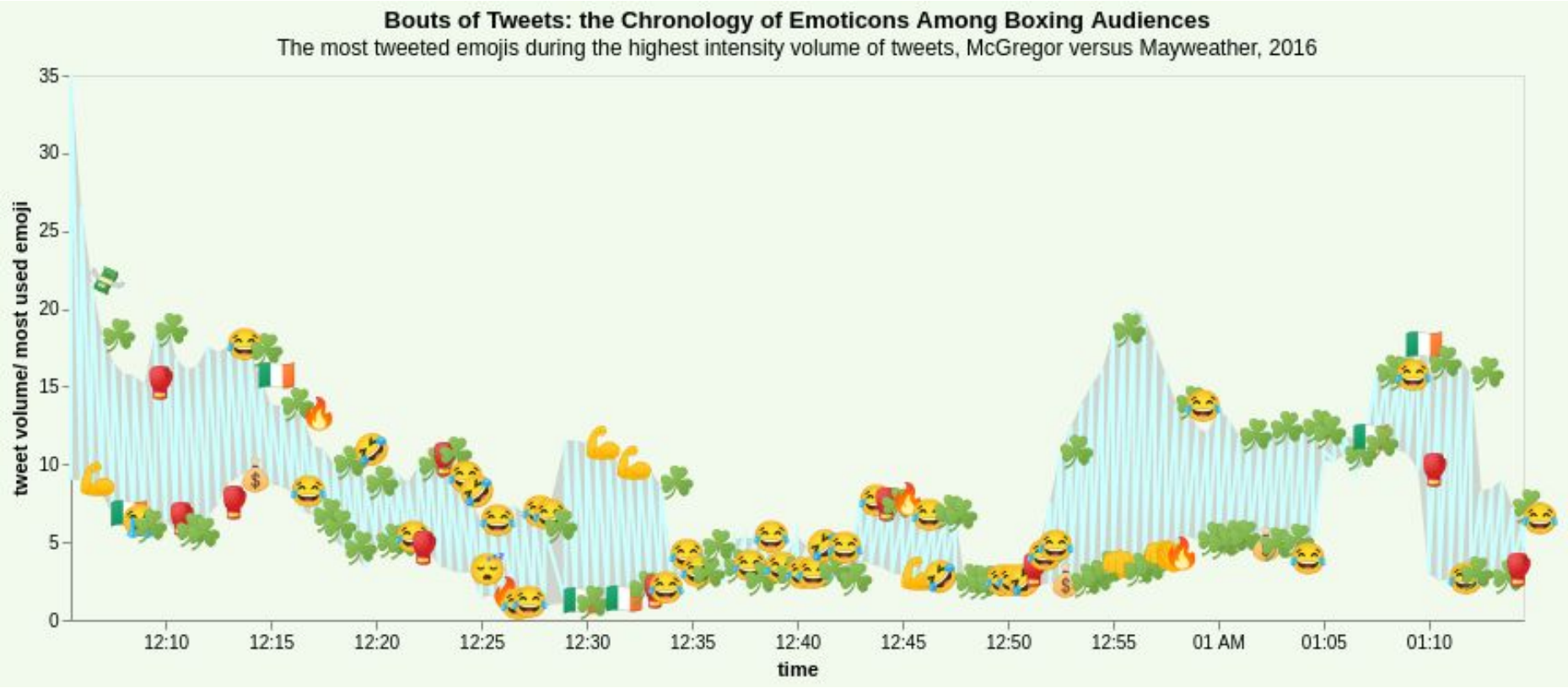
Abstract

Twitter is a constant stream of ASCII media, otherwise known as emojis. These emojis contain vital research content. The emphasis in this experimental notebook is to apply a grammar of graphics approach to exploring twitter streams, as it pertains to emoji content. A firehose, time series of tweets related to a televised boxing match was captured, then visualized.

Final notebook:

https://github.com/stefanbund/MADSPortfolio/blob/main/visualizing%20tweet%20time%20series%20via%20emojis.ipynb

The final graphic



Objectives

Retain the Time Series Encoding.

The original inks the details of the time series ticks, very precisely. This provides a pseudo-scientific 'factuality' to the graphic. Since the purpose of the study was to have fun, I wanted to abstract the precision of the time series line mark, retaining the preattentive, oppositional color encodings for each team.

Retain the cultural symbolism for the audience.

Since the spirit of the chart was to have fun, while illuminating the emoi types, I proposed displaying the most popular of each emoji, at periodic, important time events. Thus, the oppositional encoding of each line mark could be accompanied by the emojis which ranked at the top of each critical moment. This attempts to convey a sense of narrative, which clearly takes place within the 'battle of the emojis.' Again, I am trying to capture the fun spirit of the study.

Enable an Abstract Time Series Encoding, for a more aesthetic outcome

Instead of encoding the time series with highly precise marks, I am making a smoothed line. This captures the quantitative value of the time series, but illustrates the series as a shape or trend.

Data Prep

The firehose of tweets had to be organized into a dataframe where the raw tweet with the emojis became available

datetime	created_at	emojis	id	link	retweeted	screen_name	text	☘	🇮🇪	🌴
2017-08-27 00:05:34	2017-08-27 00:05:34	1	901656910939770881	https://twitter.com/statuses/901656910939770881	False	aaLiysr	Ringe çıkmadan ateş etmeye başladı 😃 #McGregor ...	0	0	
2017-08-27 00:05:35	2017-08-27 00:05:35	5	901656917281574912	https://twitter.com/statuses/901656917281574912	False	zulmafrancozaf	😲😲😲😲😲 @lalylourbet2 https://t.co/ERUGHhQINE	0	0	
2017-08-27 00:05:35	2017-08-27 00:05:35	2	901656917105369088	https://twitter.com/statuses/901656917105369088	False	Adriana11D	🇮🇪🇮🇪🇮🇪💪💪 #MayweathervMcgregor	0	3	
2017-08-27 00:05:35	2017-08-27 00:05:35	2	901656917747142657	https://twitter.com/statuses/901656917747142657	False	Nathan_Caro_	Cest partit #MayweatherMcGregor 💪🏾	0	0	
2017-08-27 00:05:35	2017-08-27 00:05:35	2	901656916828594177	https://twitter.com/statuses/901656916828594177	False	sahouraxox	Low key feeling bad for ppl who payed to watch...	0	0	

Divining suitable aggregates

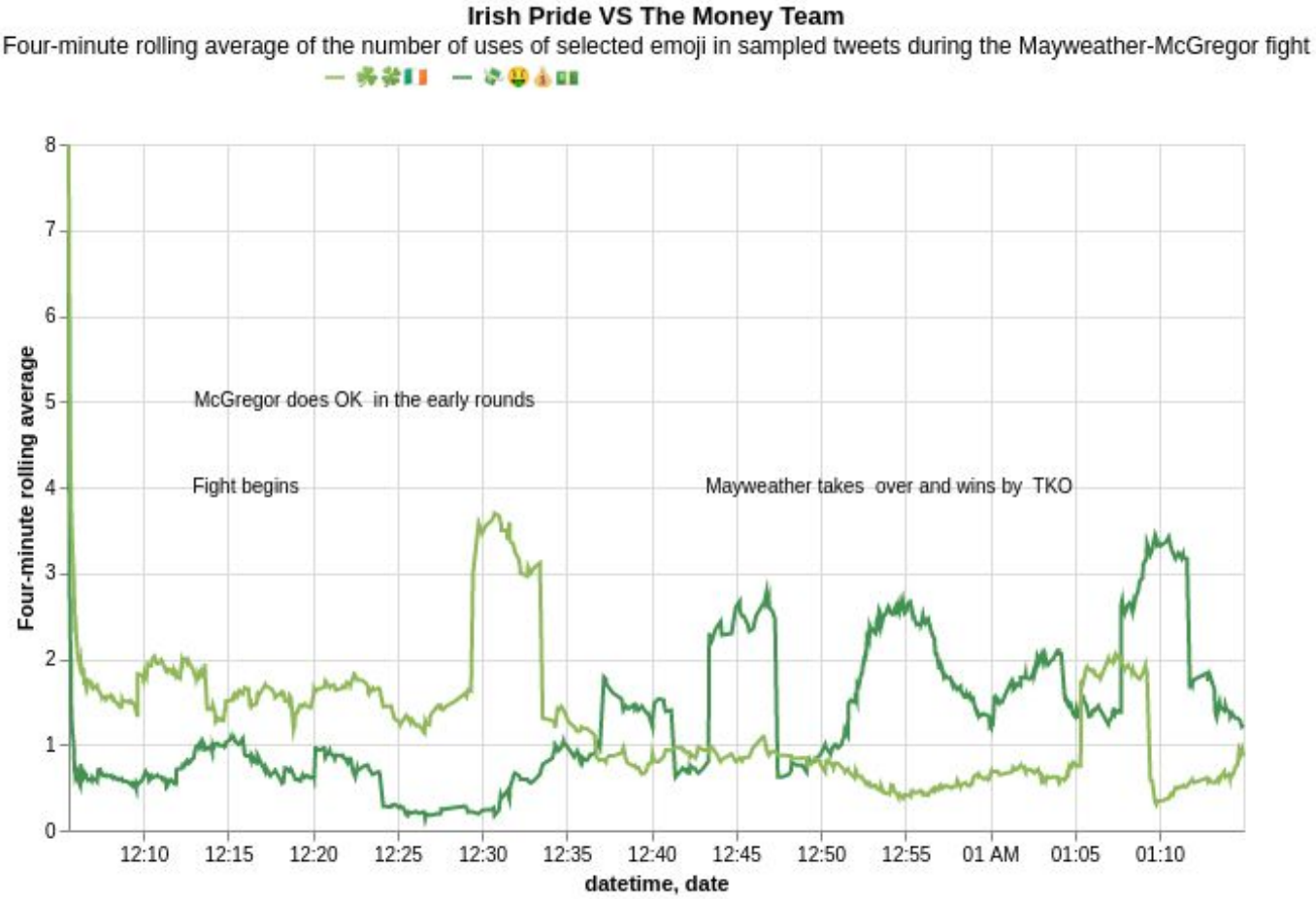
From the morass of twitter text, a simple set of usable statistics had to be gleaned, per emoji

	EMOJI	PERCENT	rank	PERCENT_TEXT
0	😂	23.1	1	23.1 %
1	🥊	5.7	2	5.7 %
2	👊	3.5	3	3.5 %
3	👋	3.0	4	3.0 %
4	💪	2.5	5	2.5 %
5	🇮🇪	2.4	6	2.4 %
6	🤣	2.3	7	2.3 %
7	🔥	2.3	8	2.3 %
8	😭	2.0	9	2.0 %
9	💰	1.8	10	1.8 %

Exploration 2: leader boards for emojis, per moment

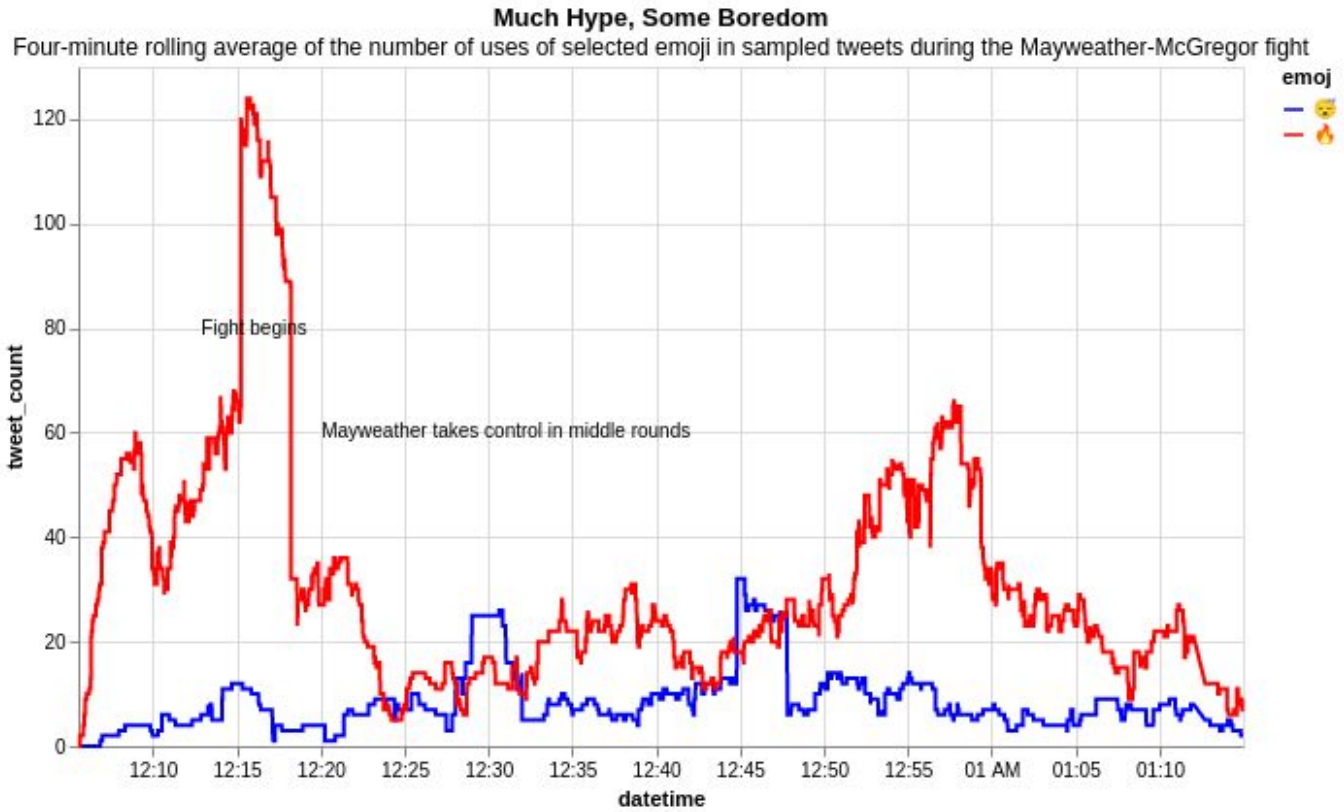
Moments of prime viewership would need to be visualized, with the leading tweet, thus if a large traffic moment takes place, a leading emoji can express it

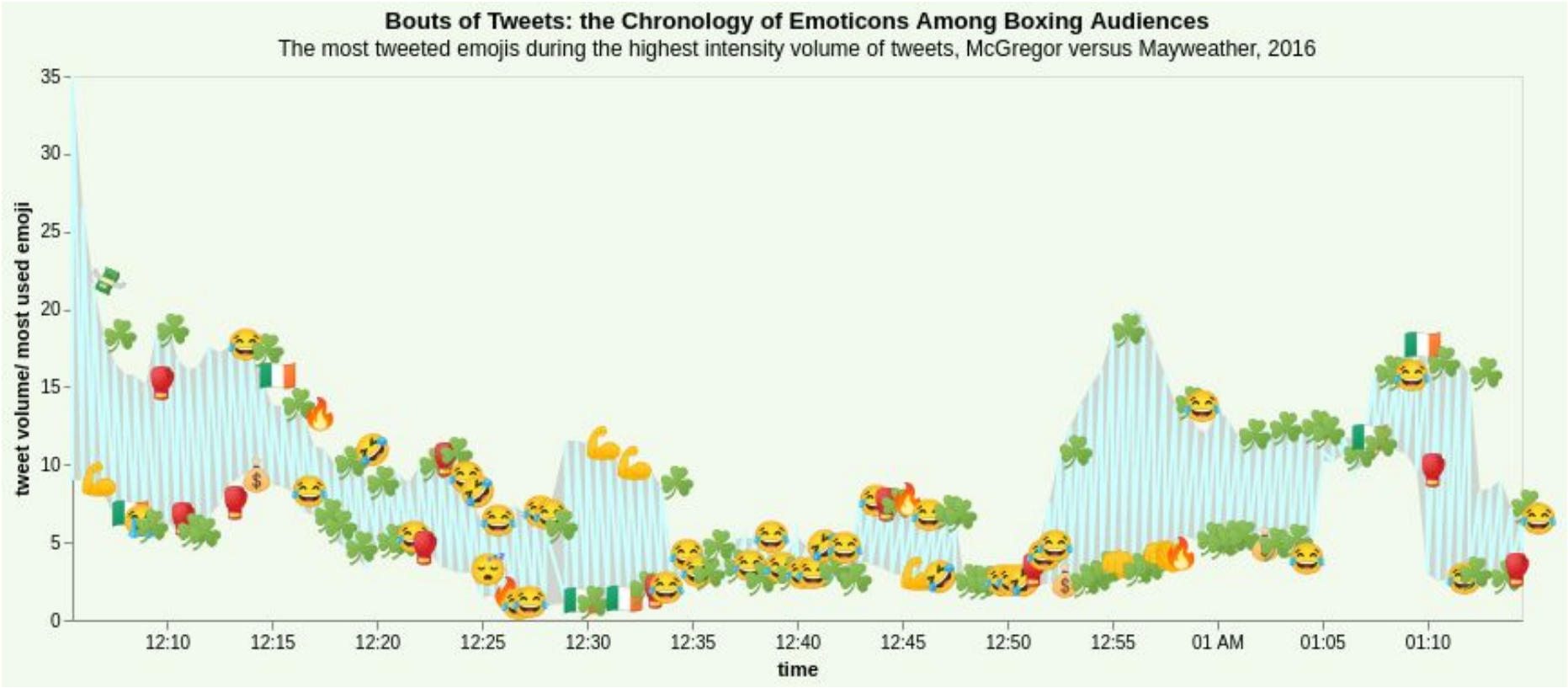
I rebuilt imagery from its original publication on 538, where certain clustered emoji appear, thereby capturing the leading emotion of the fight



Major emoji were found to be capable of signifying the major emotion.

Data trends could be signified using these main, leading emoji




Bouts of Tweets: the Chronology of Emoticons Among Boxing Audiences
The most tweeted emojis during the highest intensity volume of tweets, McGregor versus Mayweather, 2016
tweet volume/ most used emoji
35
30
25
20
15
10
5
0
12:10
12:15
12:20
12:25
12:30
12:35
12:40
12:45
12:50
12:55
01 AM
01:05
01:10
time