

# All the Feels

Emoji vs. Text Sentiment Analysis

Alexandra Plassaras  
QMSS Thesis

# Emo-what?

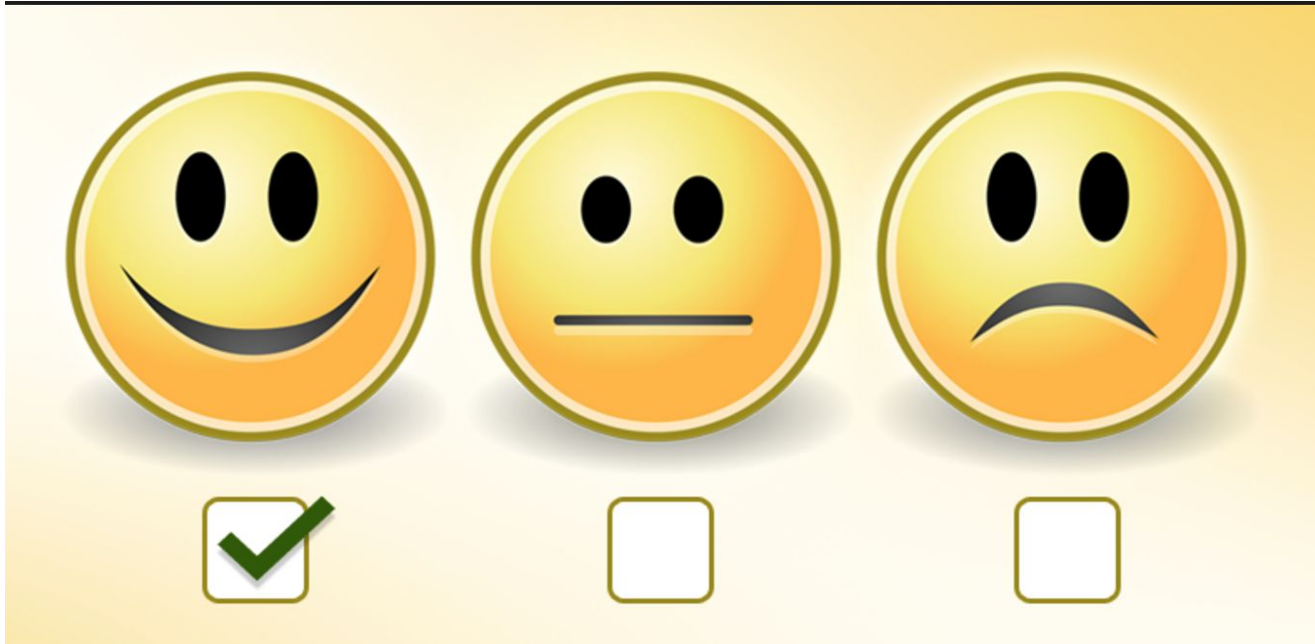
❤️	=	<3
💔	=	</3
😎	=	8)
😬	=	D:
😐	=	:
🐵	=	:o)

😊	:	=	=)	--)
😄	=	:D	:-D	
😏	=	;)	;-)	
😁	=	:>	:->	
😮	=	:O	:-O	
😡	=	>:(	>:-(	

😊	:	=	:)	(:	: -)
😞	=	: (	) :	: - (	
😓	=	: /	: - /	: \	: - \
😜	=	: p	: - p	: b :	: - b
😝	=	; p	; - p	; b	; - b
💋	=	: *	: - *		

..... Emoji NOT Emoticons










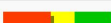


# Sentiment Analysis



# The Lexicons - Emoji

Source: Novak, K., Smailović, P.,  
Sluban, J., Mozetič, B. (2015)  
Sentiment of Emojis. PLoS ONE  
10(12): e0144296.  
doi:10.1371/journal.pone.0144296

## Emoji Sentiment Ranking v1.0

Char	Image [twemoji]	Unicode codepoint	Occurrences [5...max]	Position [0...1]	Neg [0...1]	Neut [0...1]	Pos [0...1]	Sentiment score [-1...+1]	Sentiment bar (c.i. 95%)	Unicode name	Unicode block
😭		0x1f602	14622	0.805	0.247	0.285	0.468	0.221		FACE WITH TEARS OF JOY	Emoticons
♥		0x2764	8050	0.747	0.044	0.166	0.790	0.746		HEAVY BLACK HEART	Dingbats
♥		0x2665	7144	0.754	0.035	0.272	0.693	0.657		BLACK HEART SUIT	Miscellaneous Symbols
😍		0x1f60d	6359	0.765	0.052	0.219	0.729	0.678		SMILING FACE WITH HEART-SHAPED EYES	Emoticons
😭		0x1f62d	5526	0.803	0.436	0.220	0.343	-0.093		LOUDLY CRYING FACE	Emoticons
😘		0x1f618	3648	0.854	0.053	0.193	0.754	0.701		FACE THROWING A KISS	Emoticons

# The Lexicons - Text

Source: Hu, M., and Liu, B. "Mining and Summarizing Customer Reviews." Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004), Aug 22-25, 2004, Seattle, Washington, USA

## Positive

36	a+
37	abound
38	abounds
39	abundance
40	abundant
41	accessible
42	accessibile
43	acclaim
44	acclaimed
45	acclamation
46	accolade
47	accolades
48	accommodative
49	accomodative
50	accomplish
51	accomplished
52	accomplishment
53	accomplishments
54	accurate
55	accurately
56	achievable
57	achievement
58	achievements
59	achievable
60	acumen

## Negative

36	2-faced
37	2-faces
38	abnormal
39	abolish
40	abominable
41	abominably
42	abominate
43	abomination
44	abort
45	aborted
46	aborts
47	abrade
48	abrasive
49	abrupt
50	abruptly
51	abscond
52	absence
53	absent-minded
54	absentee
55	absurd
56	absurdity
57	absurdly
58	absurdness
59	abuse
60	abused

# The Data Set

Context: Presidential Primaries

Date Range: February 7th, 2016 to  
April 2nd, 2016

Total number of Tweets: 1,816,475

Tweets with emoji: 86,258 (~ 4.75%)

Tweets with emojis and reference to a  
political party: **41,834 (~2.30%)**

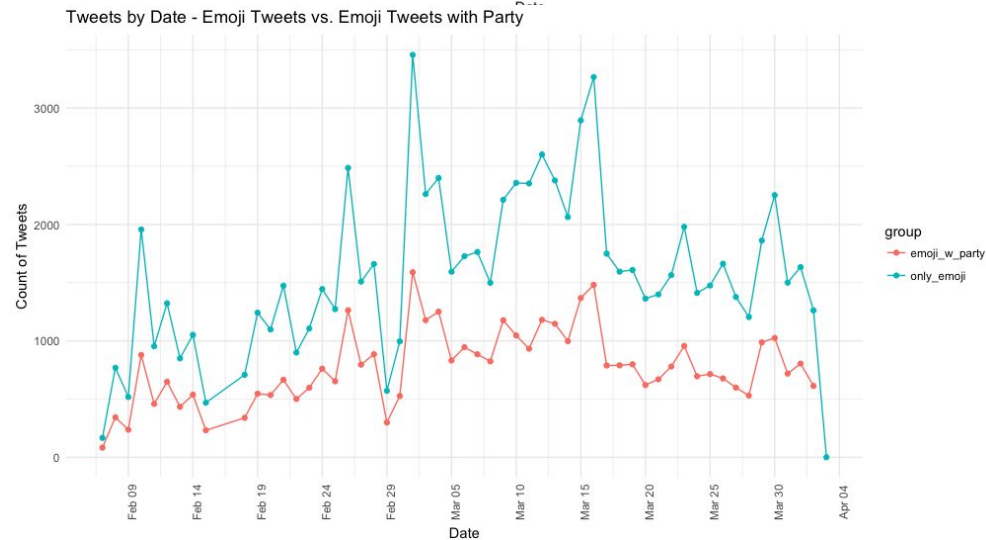
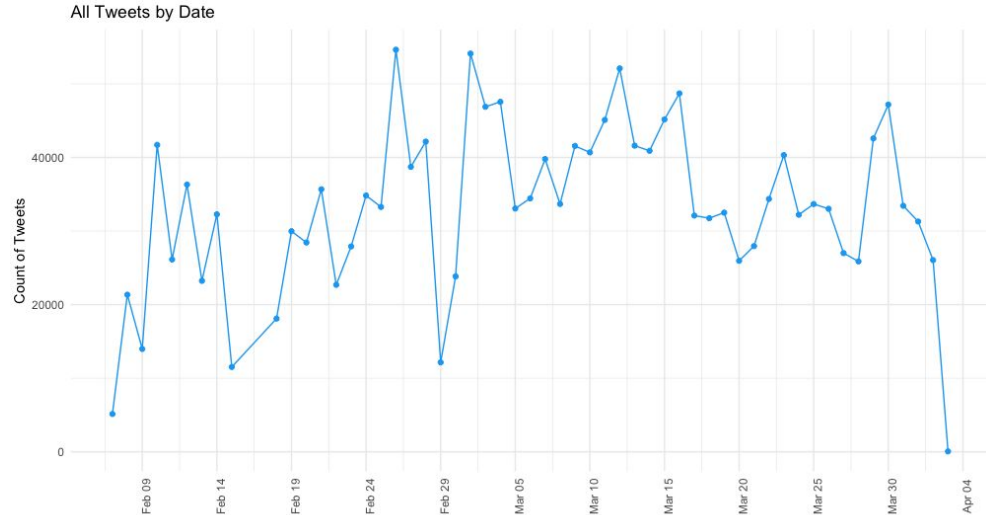
Candidate's Full Name	Key words associated to locate Tweets
Hillary Clinton	Clinton, clinton, Hillary, hillary, Hillaryclinton, hillaryclinton, Hillary Clinton, hillary clinton
Bernie Sanders	Berniesanders, berniesanders, Bernie Sanders, bernie sanders, Bernie, Bernie, Sensanders, sensanders
Ted Cruz	Cruz, cruz, Ted, ted, Tedcruz, tedcruz, Ted Cruz, ted cruz
Donald Trump	Donaldtrump, donaldtrump, Donald Trump, donald trump, Trump, trump, Donald, Donald, Trumpf, trumpf
Marco Rubio	Marcorubio, marcorubio, Marco Rubio, marco rubio

# Hypotheses

1. Difference between sentiment analysis of text compared to emojis
2. Tweets referencing Republican candidates will contain more negative sentiment for emojis than tweets referencing Democratic candidates
3. Tweets that mention Democratic candidates are more likely to have higher negative text sentiment than tweets mentioning Republican candidates

# Exploratory Data Analysis

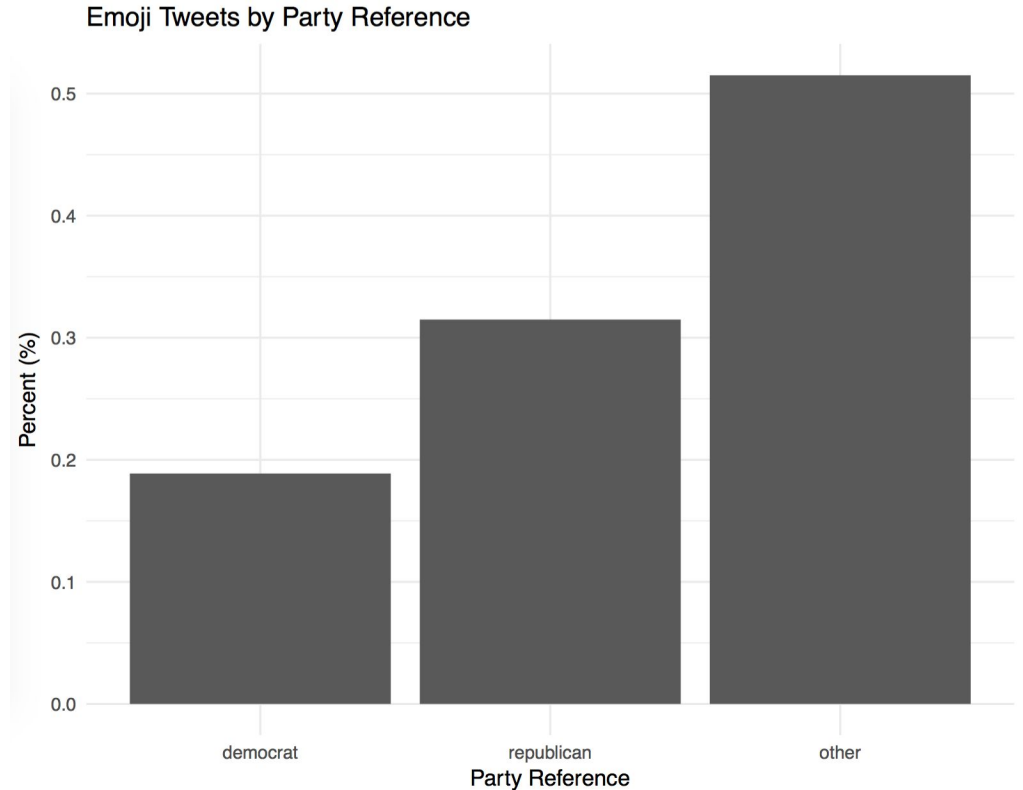
Subsets of the data appear to have similar distribution over time as original sample of ~1.8 million tweets.





# Exploratory Data Analysis

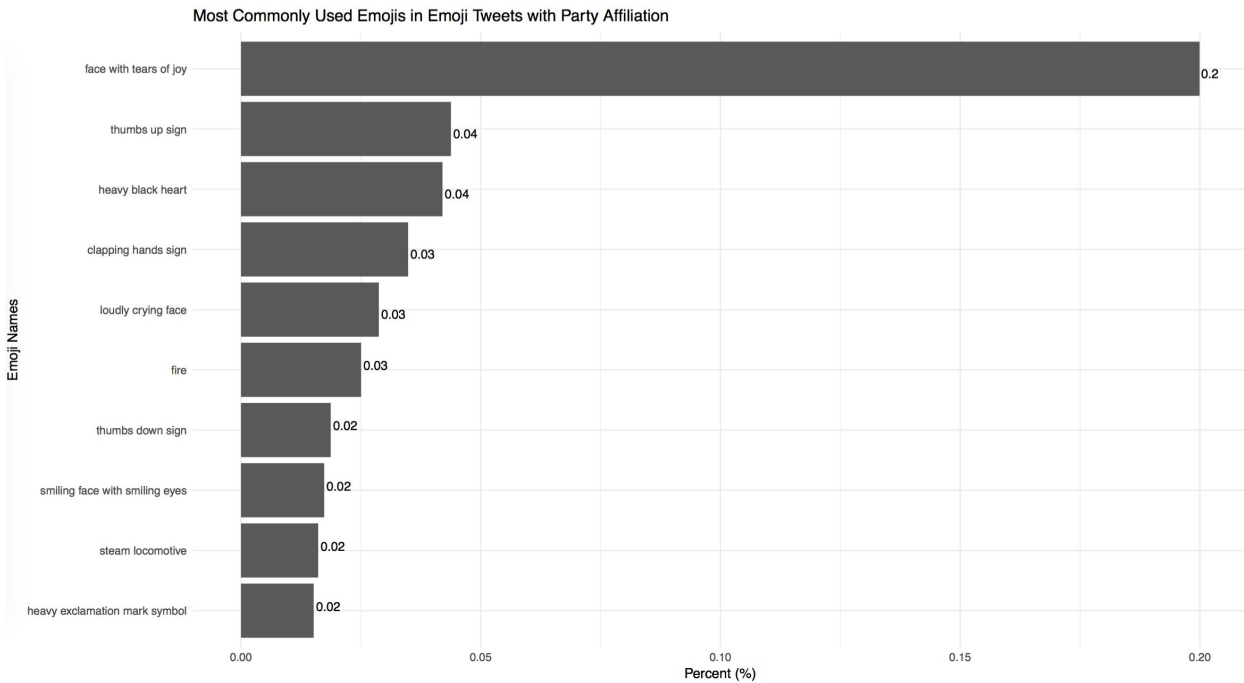
Of the emoji tweets that contain a political party reference, about 30% referred to the republican party and about 18% referred to the democratic party.



# Exploratory Data Analysis

Rank	Emoji	Percent (%)
1	😭	20%
2	👍	4%
3	❤️👉	4%
4	🌀	3%
5	🌍	3%
6	🔥	3%
7	👎	2%
8	😬	2%
9	👉	2%
10	!	2%

Total of 89,759 emojis used with 658 unique values



## Initial text EDA

Next Steps...

Calculate sentiment of text and compare  
different text lexicons