Detecting Shocking Events on Twitter

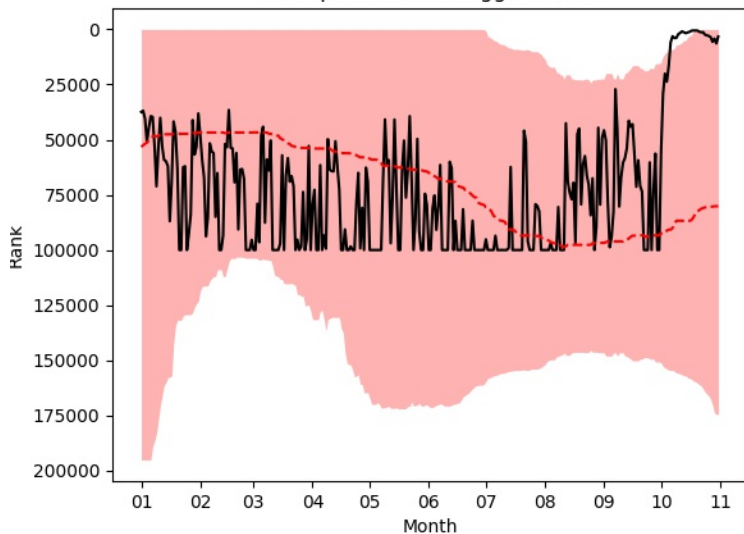## Or More Accurately Finding New Words on Twitter

Brendan Whitney

Tuesday, December 11th

What Happened

- Used Rank data from Twitter
  - Top 100000 words for each day
- Calculated median and IQR for previous 6 months
- Set cutoff at $25\% - 1.5 * IQR$
  - Ranks that went below this were "shocking"
  - Calculated the area between the actual ranks and the lower bound
    - Shock Event Score
- Recorded the largest shock event for each word
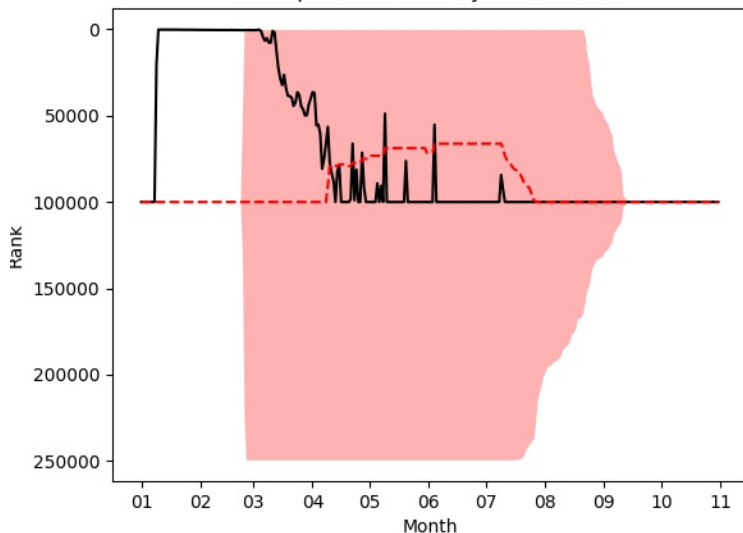- Minimized mean square error of the events to find the "story"

Rank Example
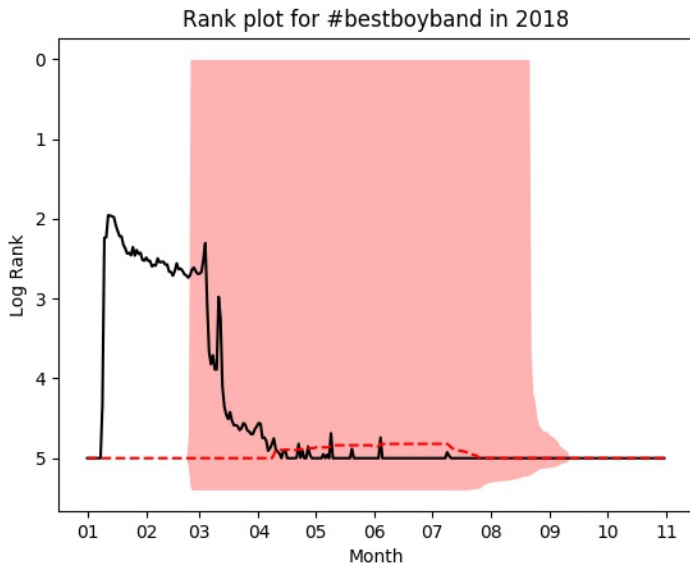

Rank plot for khashoggi in 2018

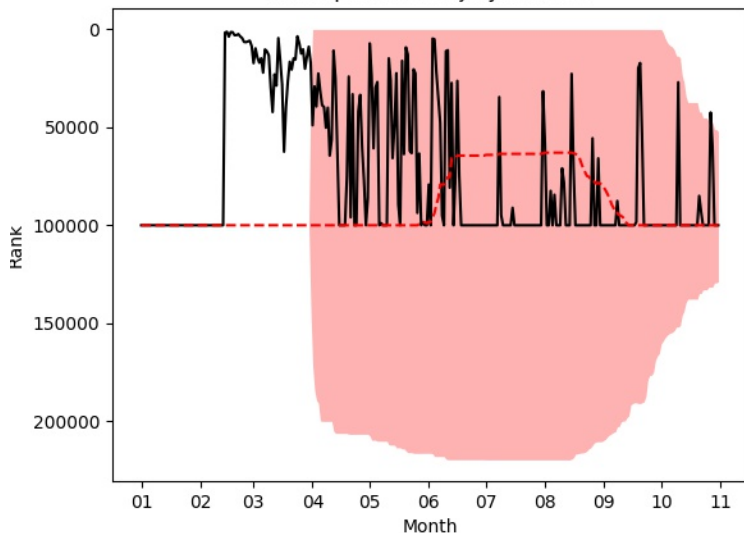# Most Shocking Word

Rank plot for #bestboyband in 2018

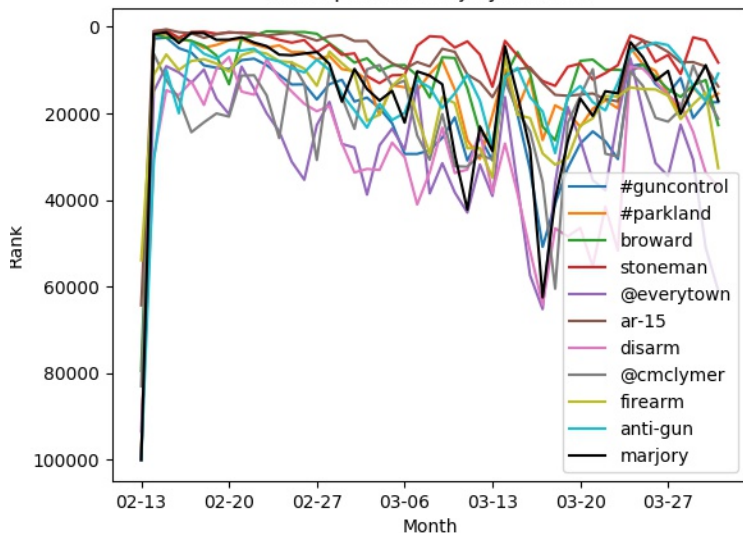# Most Shocking Word

Rank plot for #bestboyband in 2018

Story Detection

Rank plot for marjory in 2018

Story Detection

Rank plot for marjory in 2018

Initial Results

Rank	Word	Hashtag	User
1	'tear'	#bestboyband	@davidhogg111
2	rate]	#marchforourlives	@emma4change
3	marjory	#parkland	@sjofficial
4	come99	#ninepercent	@jewepri_staff
5	stc79	#nct2018	@shindeshilpas
6	4/23	#greatawakening	@cameron_kasky
7	killmonger	#poshmark	@247jimin
8	descargables	#mdrive	@neiljedcastro
9	fessy	#fakelove	@jaclyncorin
10	stoneman	#nbaplayoffs	@projetocabrabr

Known Words Results

Rank	Word	Hashtag	User
1	010x2163x8847	#bbnaija	@chuuzus
2	gander	#cncowners	@mariobautistamb
3	berinteraksi	#wrestlemania	@huniverses
4	boo'd	#toulouse	@vivo_india
5	oylar	#backtoyou	@dismiss
6	sdcc	#mvsales	@baneandlewis
7	[08	#shameless	@debryanshow
8	kanye's	#hwpl	@digvijaya_28
9	currentcet	#mbs	@soonsvt
10	esquentar	#stevenuniverse	@amas