Analyzing How Triggering and Activism Events Impact Hashtag Activism

#team: jwhang3, tnguye72, kkwan, lwilson7

Hypothesis

We believe that on average, triggering events for social movements will cause a greater increase in related hashtag usage than activism events. We will measure the tweet count increase by calculating percent changes between the number of tweets from the day before each event and the day of the event, and we will use t-tests to determine statistical significance.

Data

We primarily used the Twitter API to collect our data. To create our Events dataset, we chose 3 movements (MeToo, Black Lives Matter, and Gun Violence) to focus our study on. For each movement, we selected 10 "triggering" events (shootings, scandals, etc) and 10 "activism" events (marches, protests). In addition, we used the counts endpoint of the Twitter Full Archive Search API to collect the number of tweets per day that contain a specific hashtag. Using this data, we created an events csv with the following columns of interest: id, desc, movement, utc, type. We also created a tweet counts csv with the following columns of interest: timestamp, movement, count.

Using our data, we conducted nine t-tests. These t-tests included tests between triggering and activism events, triggering events and control days, and activism events and control days for #MeToo, #BlackLivesMatter or #BLM, and #NeverAgain or #EnoughIsEnough. We created our set of control days by randomly selecting 10 days that were not within 1 day of any event and calculated the percent change in the number of tweets the same way that we did for the triggering and activism events.

Findings

Claim #1: There is not enough evidence to suggest that on average, triggering events cause a greater increase in hashtag usage for the related movement than activism events.
Support for Claim #1: We conducted t-tests for all three movements to see if there was a statistically significant difference in the number of tweets between triggering and activism events. We then compared these results to a significance level of 0.05, per standard practice. We found that for all the movements, the p-values for each t-test were above 0.05 significance level.

Triggering vs. Activism		
Movement	**t-stat**	**p-value**
MeToo	1.1690	0.2595
Black Lives Matter	-0.5742	0.5734
Gun Violence	-1.5907	0.1291

Claim #2: There is not enough evidence to suggest that on average, the events for #MeToo and #BlackLivesMatter caused a greater increase in hashtag usage than non-event days.
Support for Claim #2: For all three movements, we conducted t-tests between the triggering events and the control as well as between activism events and the control. We found that none of our p-values for MeToo and Black Lives Matter were less than a significance level of 0.05.

Movement	Statistic	Triggering vs. Control	Activism vs. Control
MeToo	t-stat	1.6856	1.1727
	p-value	0.1091	0.2581
Black Lives Matter	t-stat	0.9388	0.9319
	p-value	0.3603	0.3644

Claim #3: Triggering events are associated with significantly higher increases in hashtag usage (as compared to the control group) in the Gun Violence movement.
Support for Claim #3: We ran a t-test to determine if there is a significant difference in the percent changes of hashtag usage associated with triggering vs. control, activism vs. control, and triggering vs. activism events. We found that the p-value for the triggering vs. control for the Gun Violence movement was the only one that fell under the 0.05 significance level.

However, because we conducted many separate t-tests, we also used Bonferroni correction to test significance and avoid false positives. Thus, we divided the significance value of 0.05 by the number of tests that we ran, which was 9. This gave us a new significance level of 0.0056, which suggests that the t-test for gun violence between triggering and control events is not statistically significant once multiple comparisons have been corrected for.

Movement	Statistic	Triggering vs Control	Activism vs Control	Triggering vs Activism
Gun Violence	t-stat	2.6676	1.7667	-1.5907
	p-value	**0.0152**	0.0952	0.1291