

When do people talk about women's rights?

Timeline and sentiment in relationship with political or social events

Posted on December 11, 2021

Star

3,766

Fork

11,986

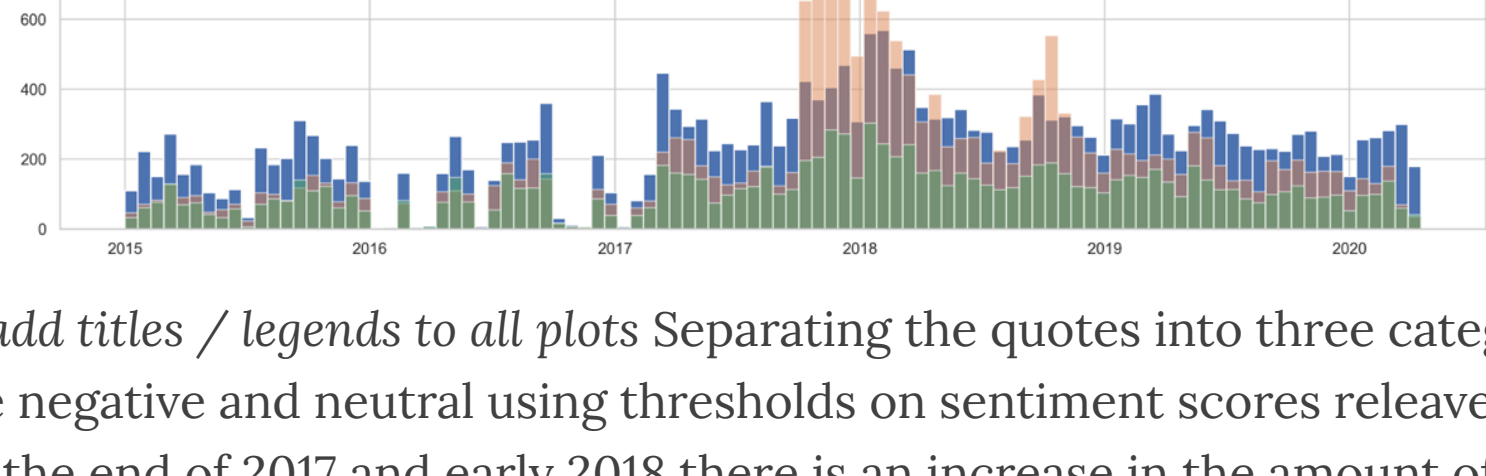
Follow @daattali

1,663

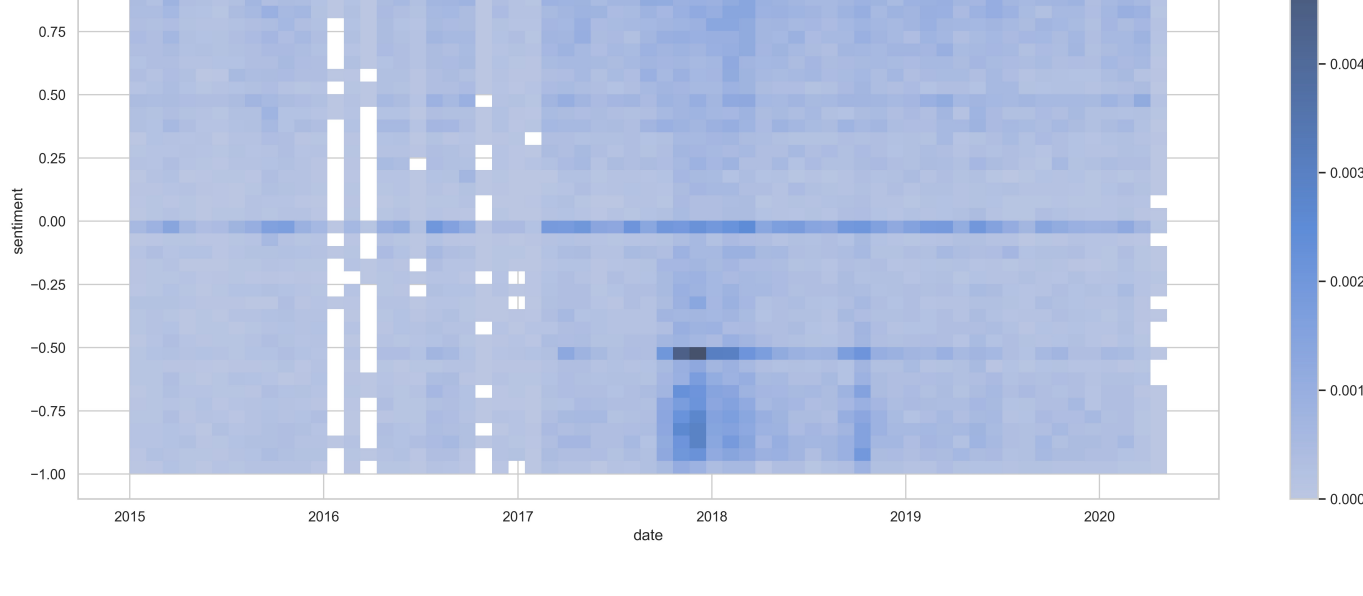
This section will investigate the timeline with sentiment and focus on specific days/months to highlight that quotes are related to political or social events. For now we can focus on the end of 2017 and show the subcategories found, i.e. that mostly it is sexual harassment quotes that are predominantly found around this time. We could also try to find other events.

The selected quotes for our study have been attributed dates between 2015 and 2020 approximately. Some periods were more active than others, hence one can wonder how did the sentiment of these quotes evolved over these time frames. The next analysis are done on quotes that have been selected (.. TODO : add precise details about which quotes we used here after deciding).

How has Sentiment Evolved over time ?



To do : add titles / legends to all plots Separating the quotes into three categories, positive negative and neutral thresholds on sentiment scores reveals that around the end of 2017 and early 2018 there is an increase in the amount of negative quotes. Neutral quotes are also present but generally less than each of the other categories, this could be something that is sensitive the threshold used to select neutral / positive / negative quotes. (todo: check this)



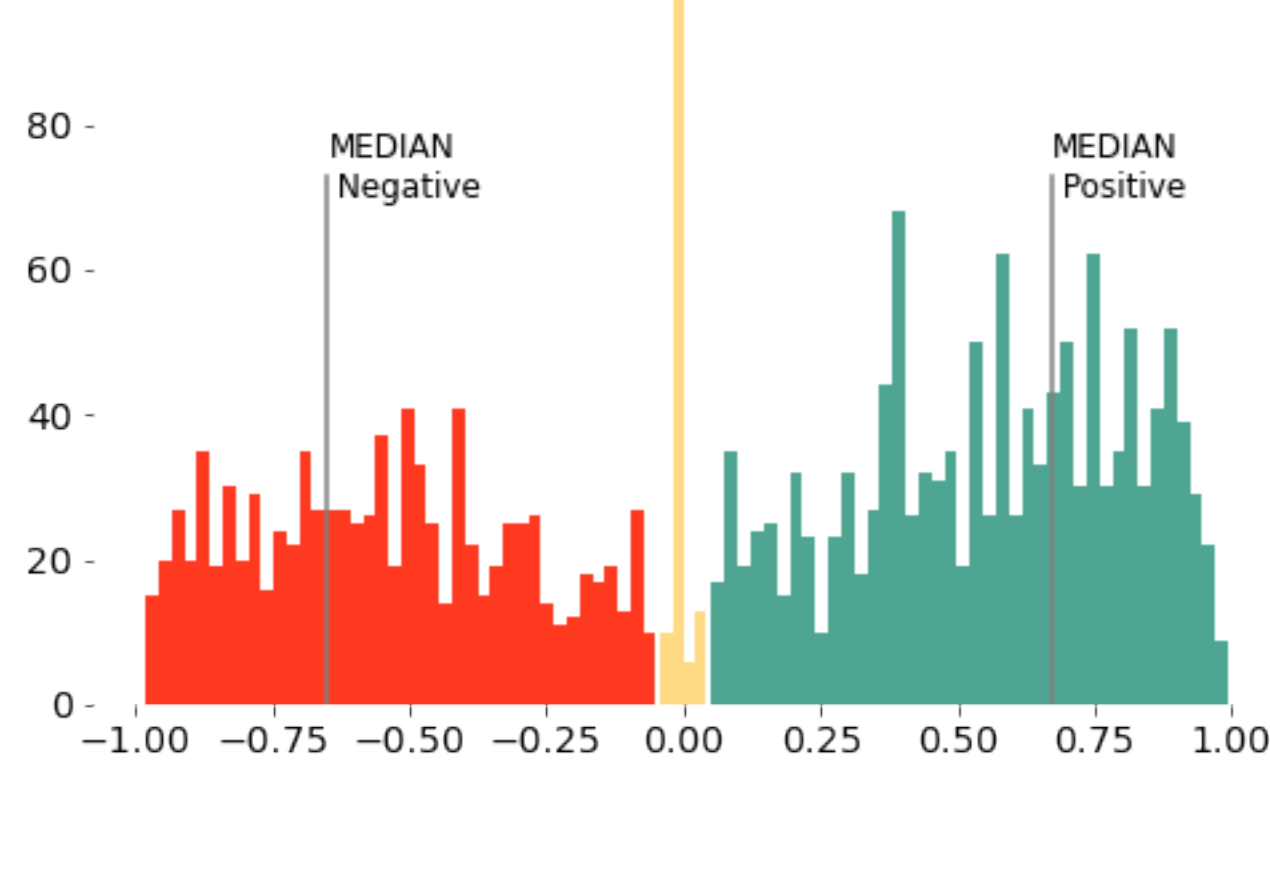
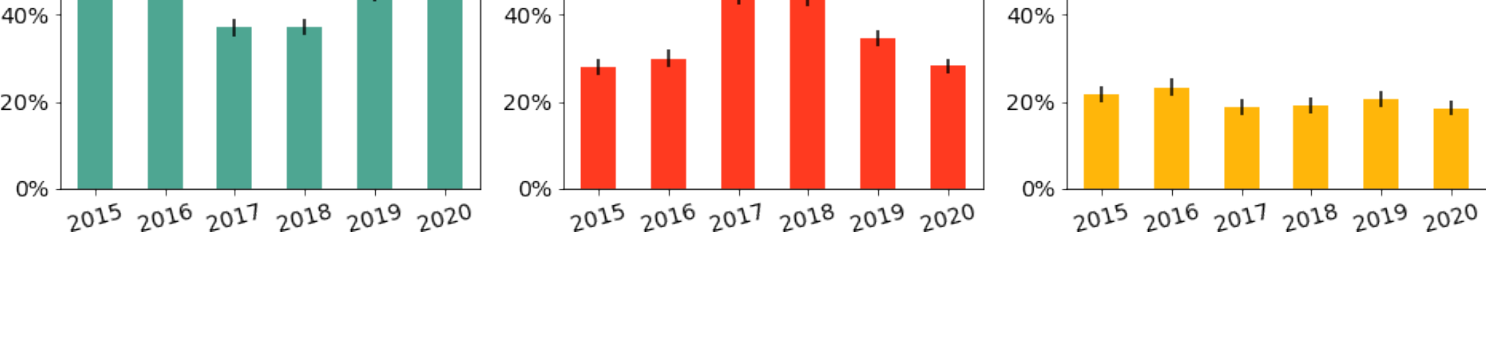
Aggregating the data (each square represents a month (31 days) , and and sentiment interval of 0.05, color represents the proportion) shows furthermore that there is an increase in the quotes with negative sentiment around the end of 2017

todo : find a way to normalize by time period only

Seems like there were a lot of negative quotes end of 2017 We could maybe use the list provided by wikipedia :

https://en.wikipedia.org/wiki/2017_in_the_United_States#October to see if there were quotes about that

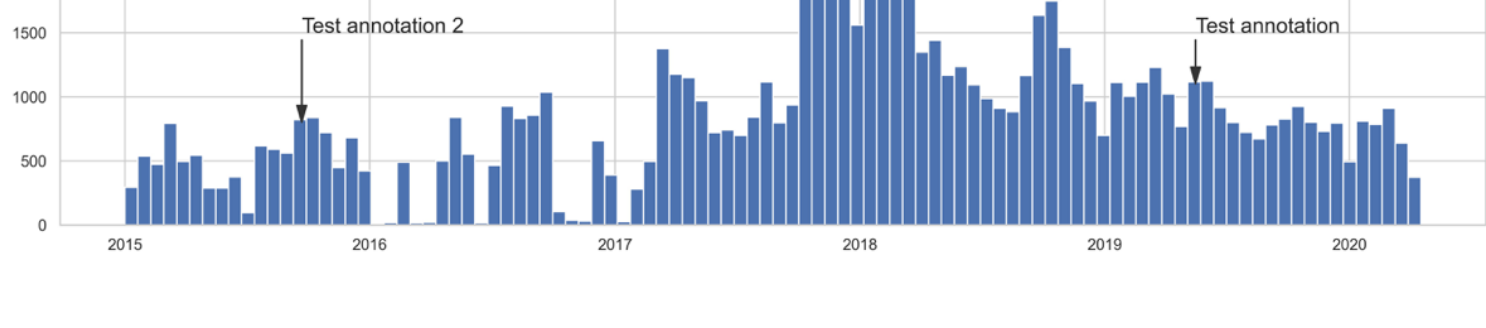
plot with bootstrapped confidence intervals here



Events Summary

to do : NER to recognize events ??

Here we use the additional dataset (wikipedia historical US events page) do show what events are mainly talked about - todo : make sure this is relevant : show proof that quotes are mostly US related

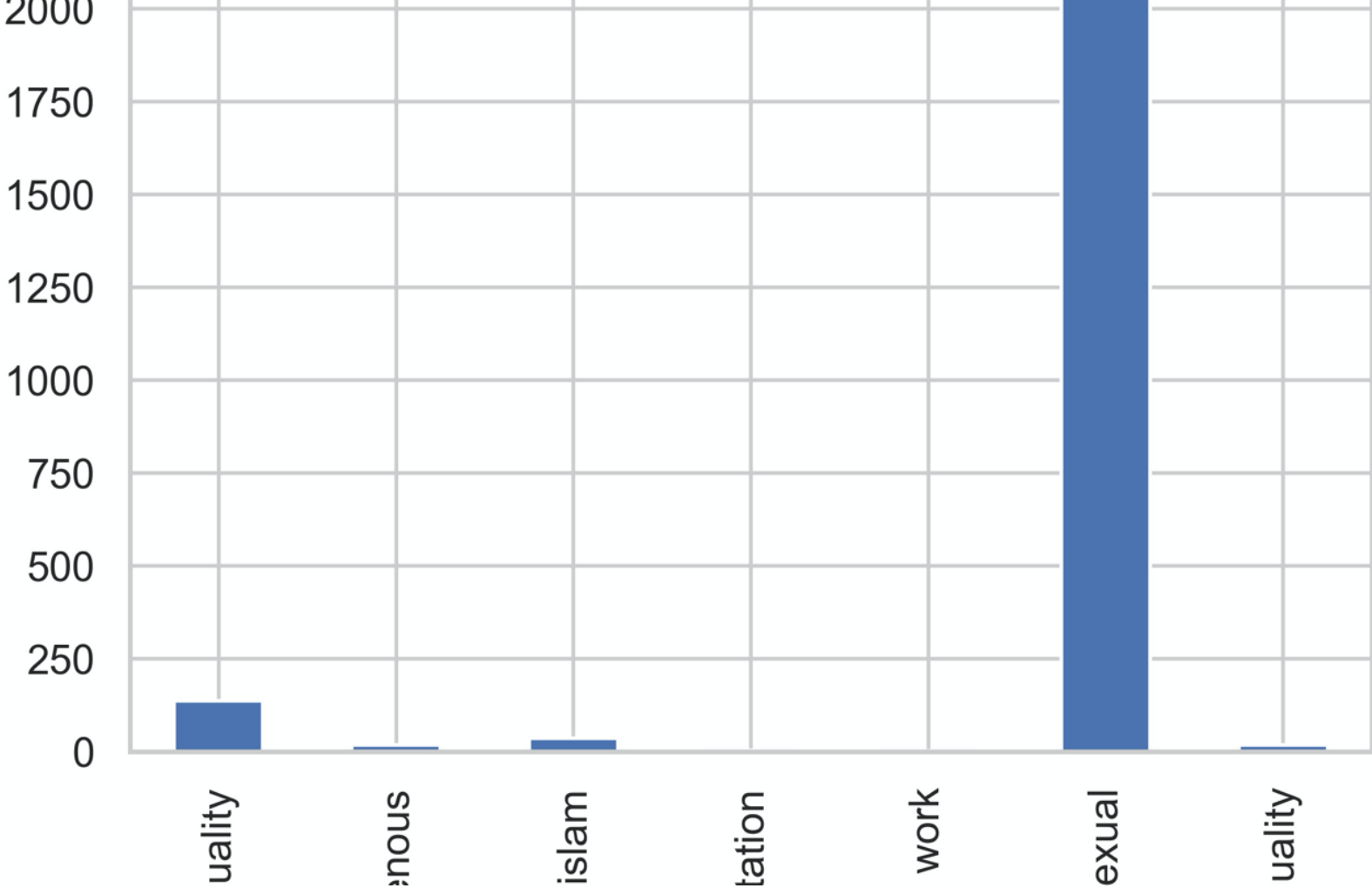


Candidate events are :

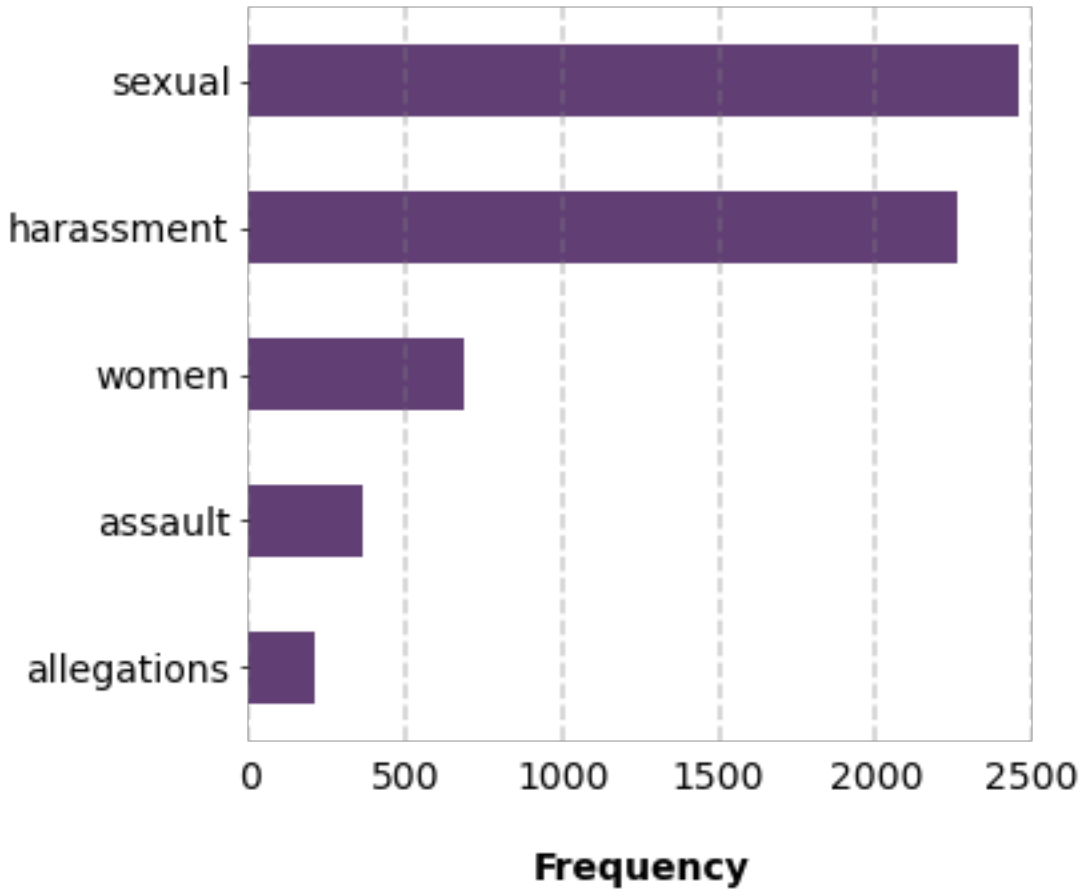
- “Tough talk on immigration: Shortly after taking office in January, President Trump sought to make good on his “America First” campaign promise by imposing a series of contentious travel bans on citizens from several Muslim-majority nations.”
- “Federal district courts struck down implementation of the bans, though a Supreme Court ruling in December 2017 reversed the lower courts’ decisions, allowing the administration to fully implement the bans.”
- “Trump also continued to promote his election campaign idea of a border wall with Mexico that he says will help quell illegal immigration from Mexico and points south.”
- https://www.history.com/topics/21st-century/2017-events#section_1

Are peaks in sentiment related to political or social events ?

To further investigate what happened around 2017, to try and find what even this corresponds to, keywords are used to put the quotes into different categories. Surprisingly enough almost all of the quotes go into the “sexual” category. This suggests there might have been a sexual harassment scandal around this time period.



Here are the 5 most commons words in the quotes selected from the heatmap. This confirm our hypothesis and allows us to look into more details.



Let’s look at some of the quotes categorized to see if we can find what they are talking about :

- quote 1 : todo : add quotes here
- quote 2 :
- quote 3 :

While these might not be entirely representative they seem to be pointing at

Methods

Tools

We discuss what tools were used for the analyses above and the validity or each of the analyses

Tests

Comparing Sentiment scores to a human

Categorizing quotes

Keywords

Tags: timeline with sentiment subcategories plot

