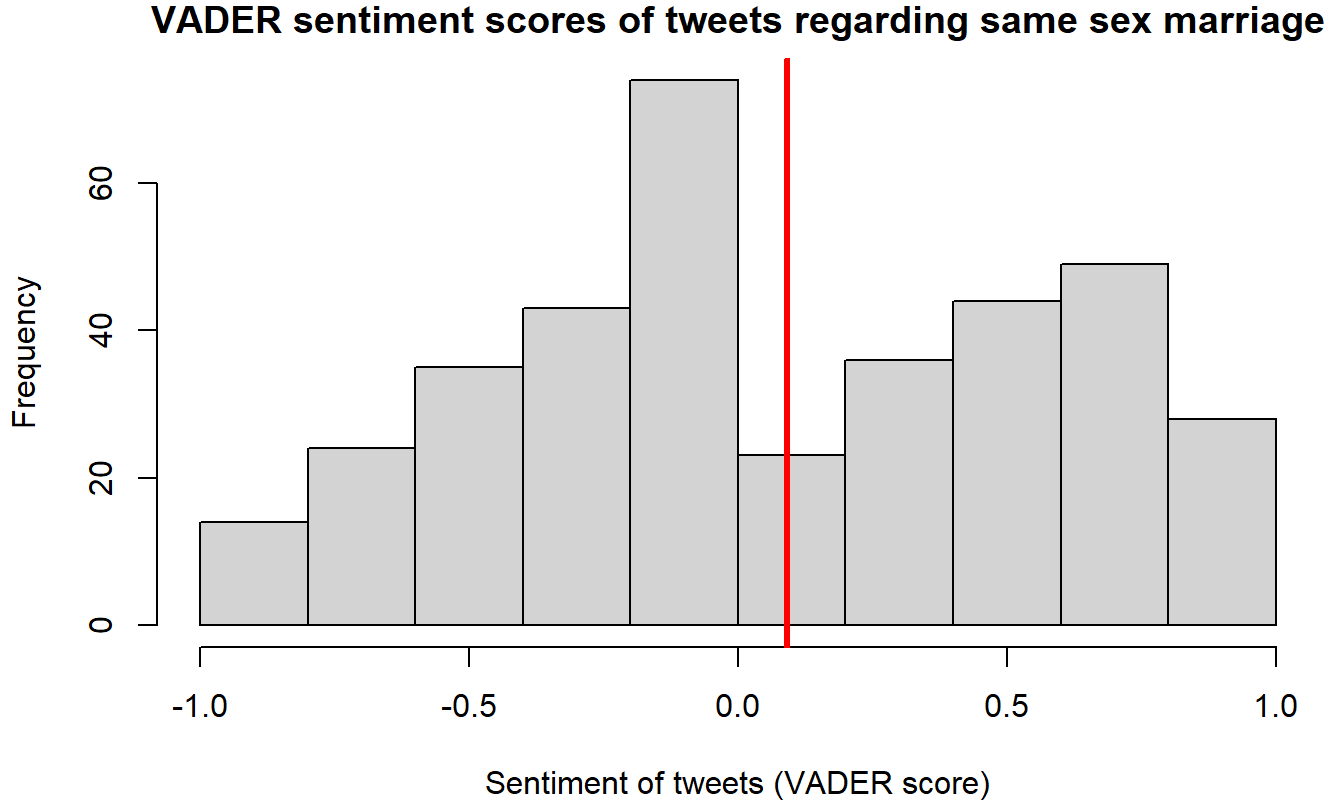
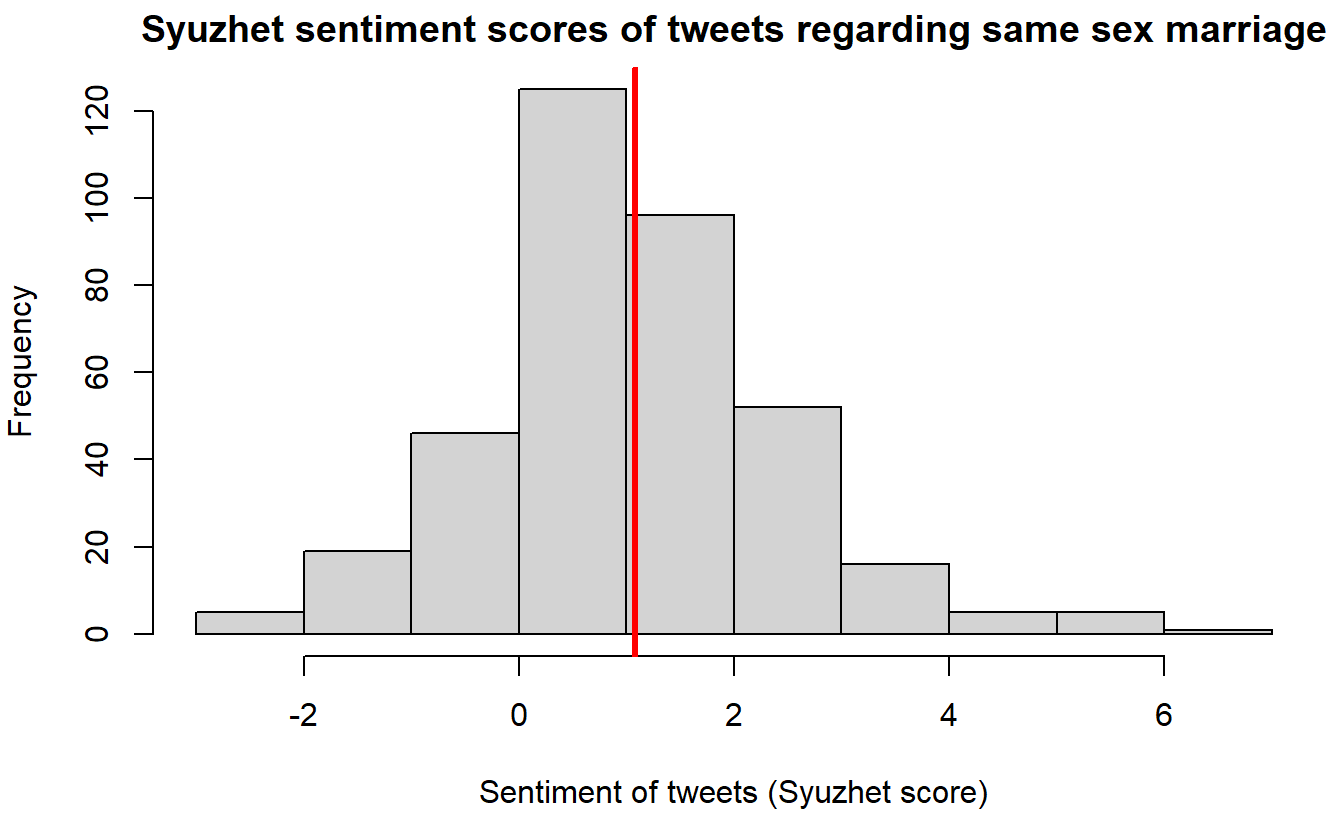
**SAME SEX MARRIAGE**

**Do tweets that use more emotional language in respect of same sex marriage have a higher engagement rate?**

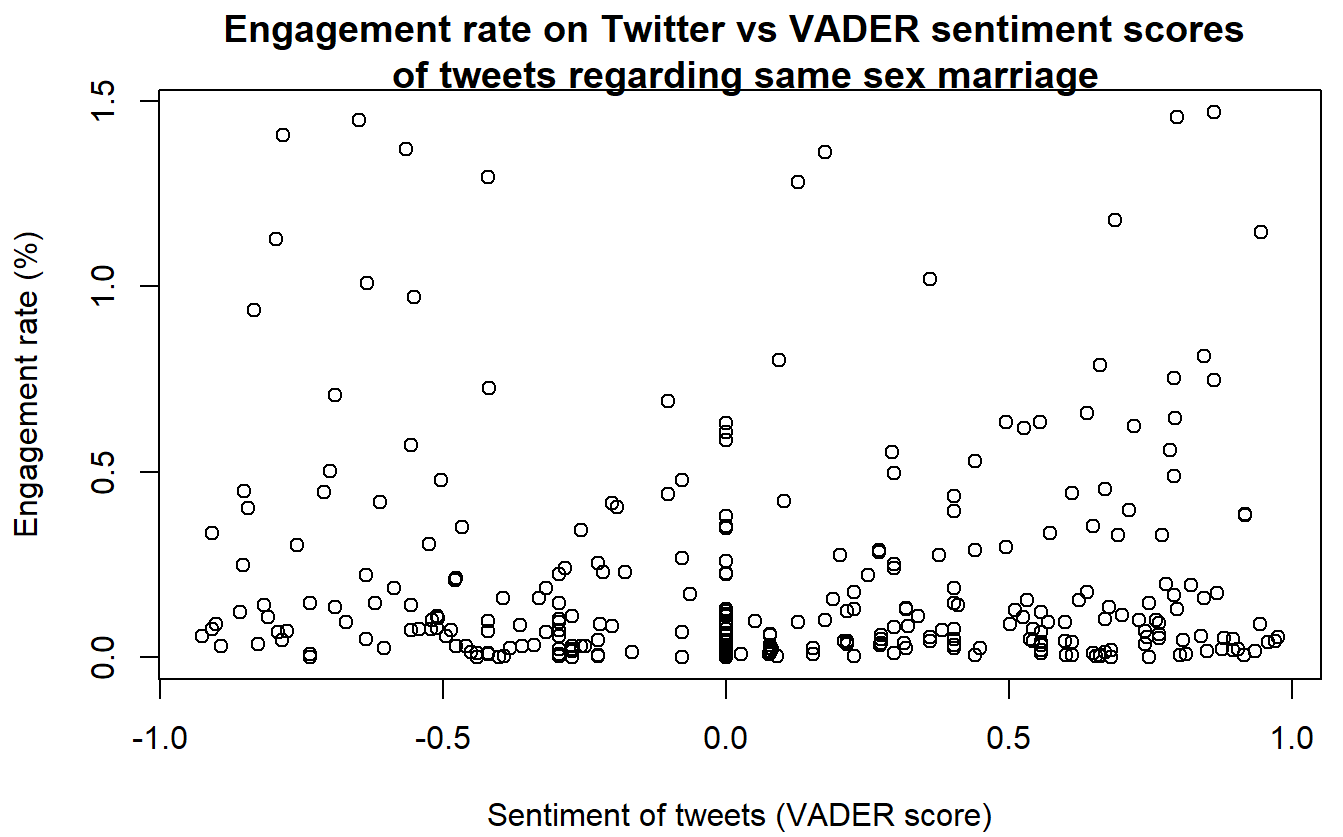
Average engagement rate (for info, in case helpful in discussion): 0.98%, 95% CI [0.41, 1.55]

Frequency of sentiment scores with average plotted as a red vertical line:

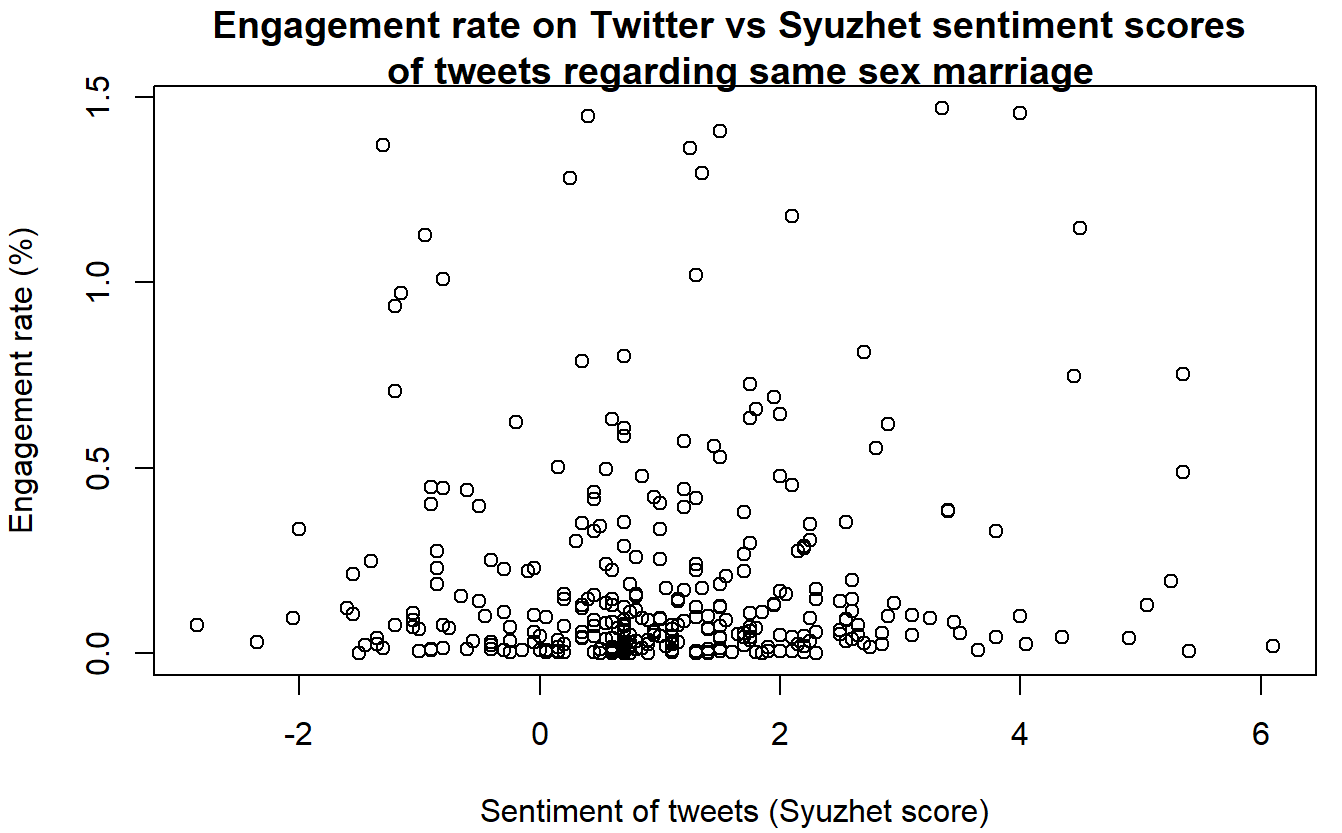




Scatter plots showing engagement rates vs sentiment scores to look for any observable relationships (NOTE: Just for the plots, but not the statistical analysis, data was filtered so that the upper limit of the 95% confidence interval was used as the max engagement rate. This is because there are some large outliers that otherwise make it difficult to see what is going on.):



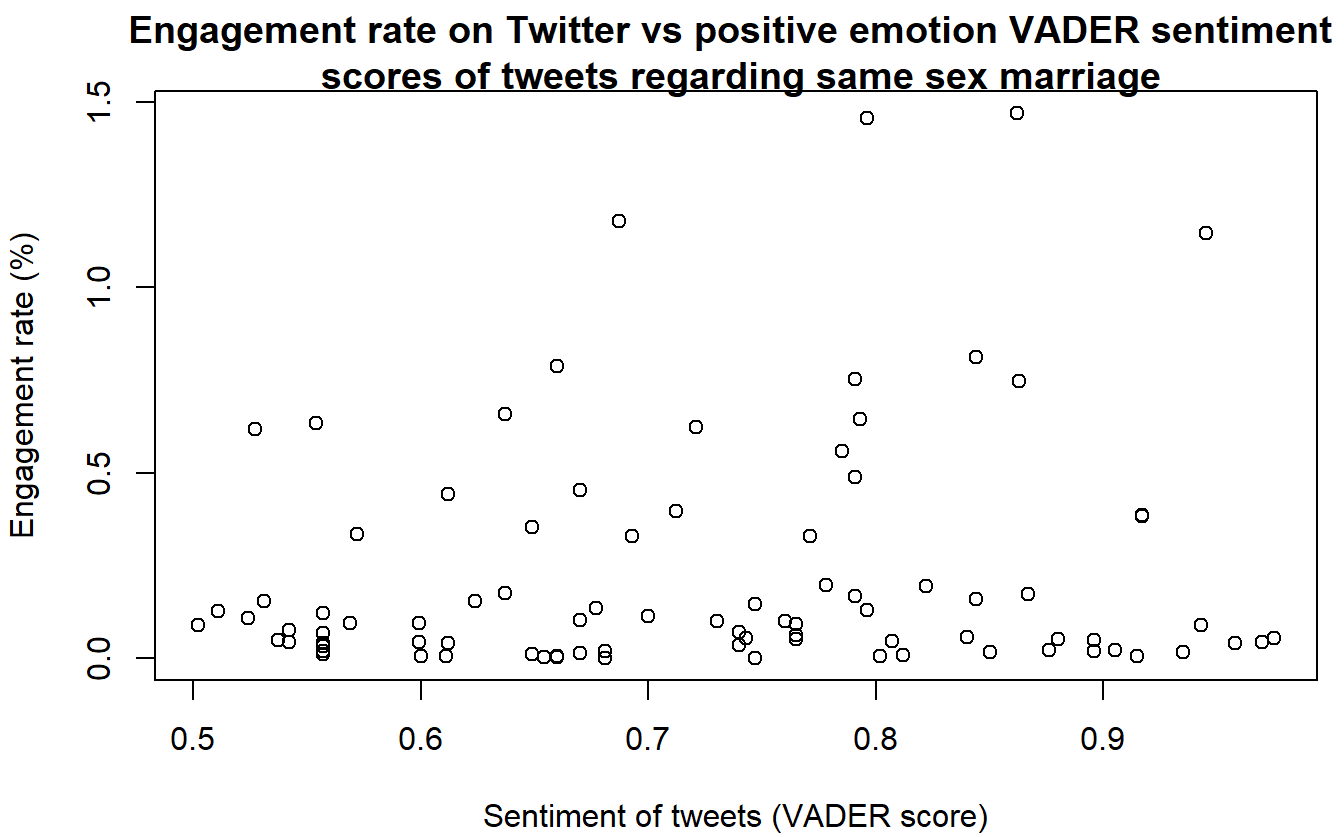
Simple linear regression was used to test if VADER sentiment scores of Tweets significantly predicted the engagement rate and no statistical significance was found (β = 0.00715, 95% CI [-0.00203, 0.01632], p = 0.12651, adjusted R2 = 0.00363). Assumptions of the linear regression model were checked.



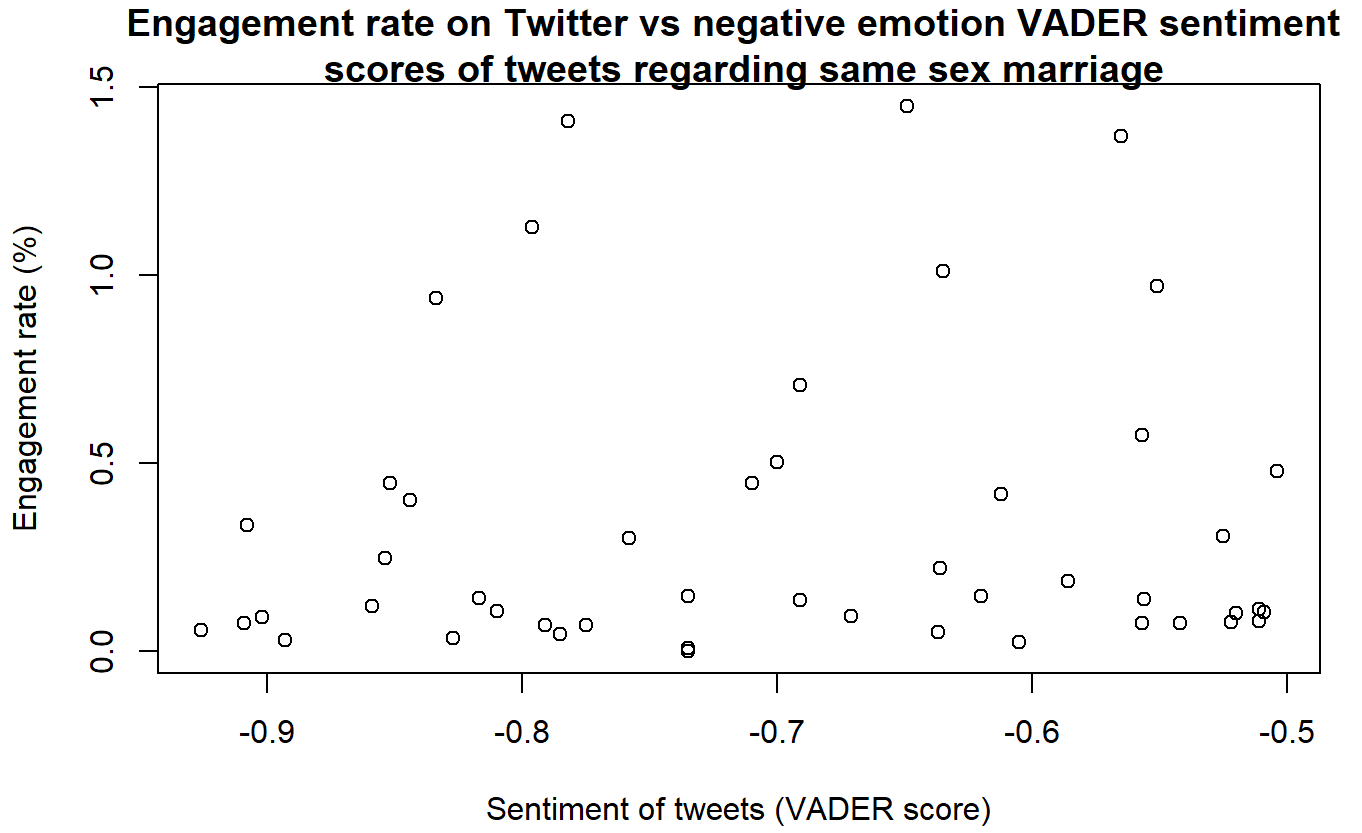
Simple linear regression was used to test if Syuzhet sentiment scores of Tweets significantly predicted the engagement rate and no statistical significance was found (β = 0.00981, 95% CI [-0.01538, 0.03501], p = 0.444, adjusted R2 = -0.00112). Assumptions of the linear regression model were checked.

We also separately looked at the relationship between engagement rates and positive and negative sentiment scores to check if there is a relationship that was not observable when the entire dataset was used (NOTE: Just for the plots, but not the statistical analysis, data was filtered so that the upper limit of the 95% confidence interval was used as the max engagement rate. This is because there are some large outliers that otherwise make it difficult to see what is going on.):

* VADER

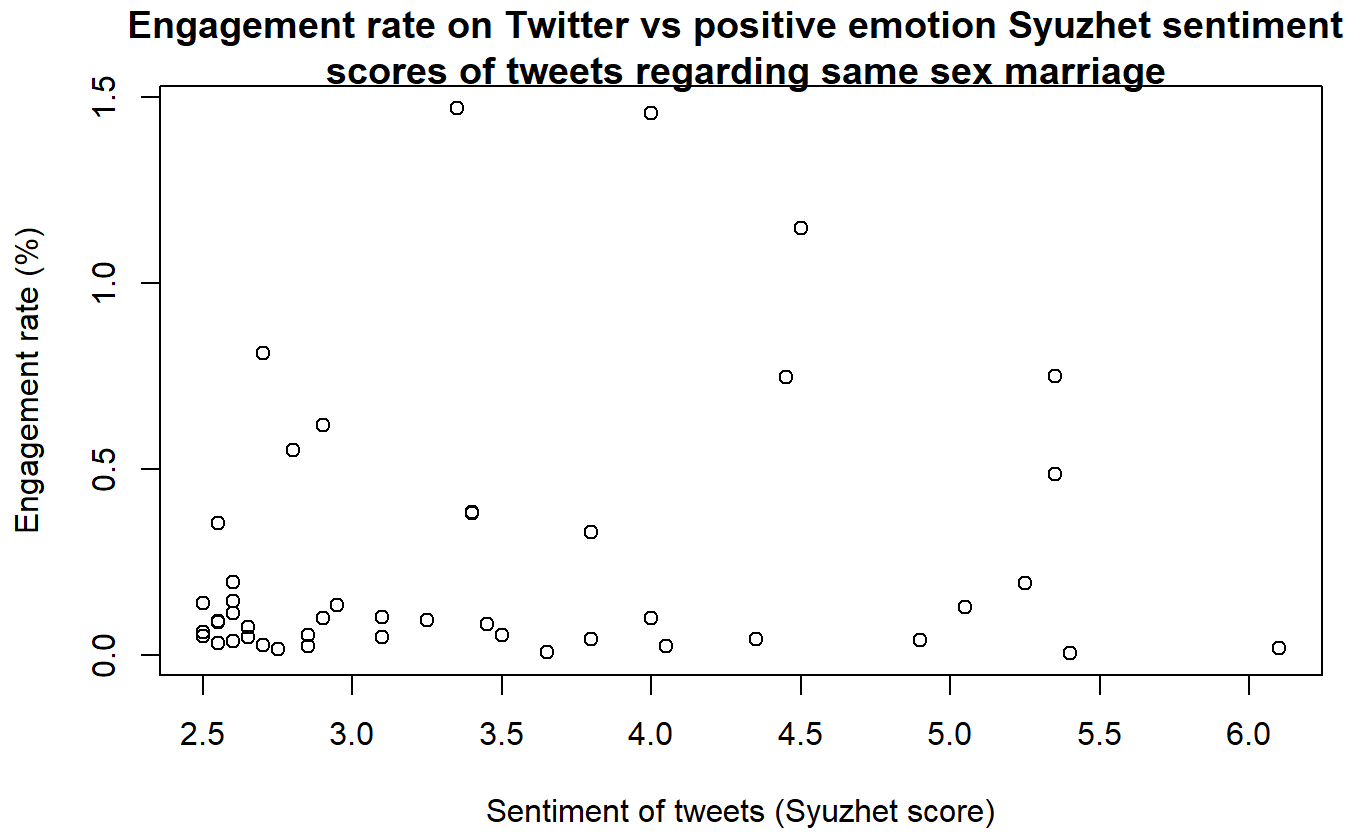


Simple linear regression was used to test if positive VADER sentiment scores of Tweets significantly predicted the engagement rate and no statistical significance was found (β = -0.00018, 95% CI [-0.00271, 0.00236], p = 0.891, adjusted R2 = -0.01001). Assumptions of the linear regression model were checked.



Simple linear regression was used to test if negative VADER sentiment scores of Tweets significantly predicted the engagement rate and no statistical significance was found (β = 0.00116, 95% CI [-0.04291, 0.04524], p = 0.958, adjusted R2 = -0.01918). Assumptions of the linear regression model were checked.

* Syuzhet

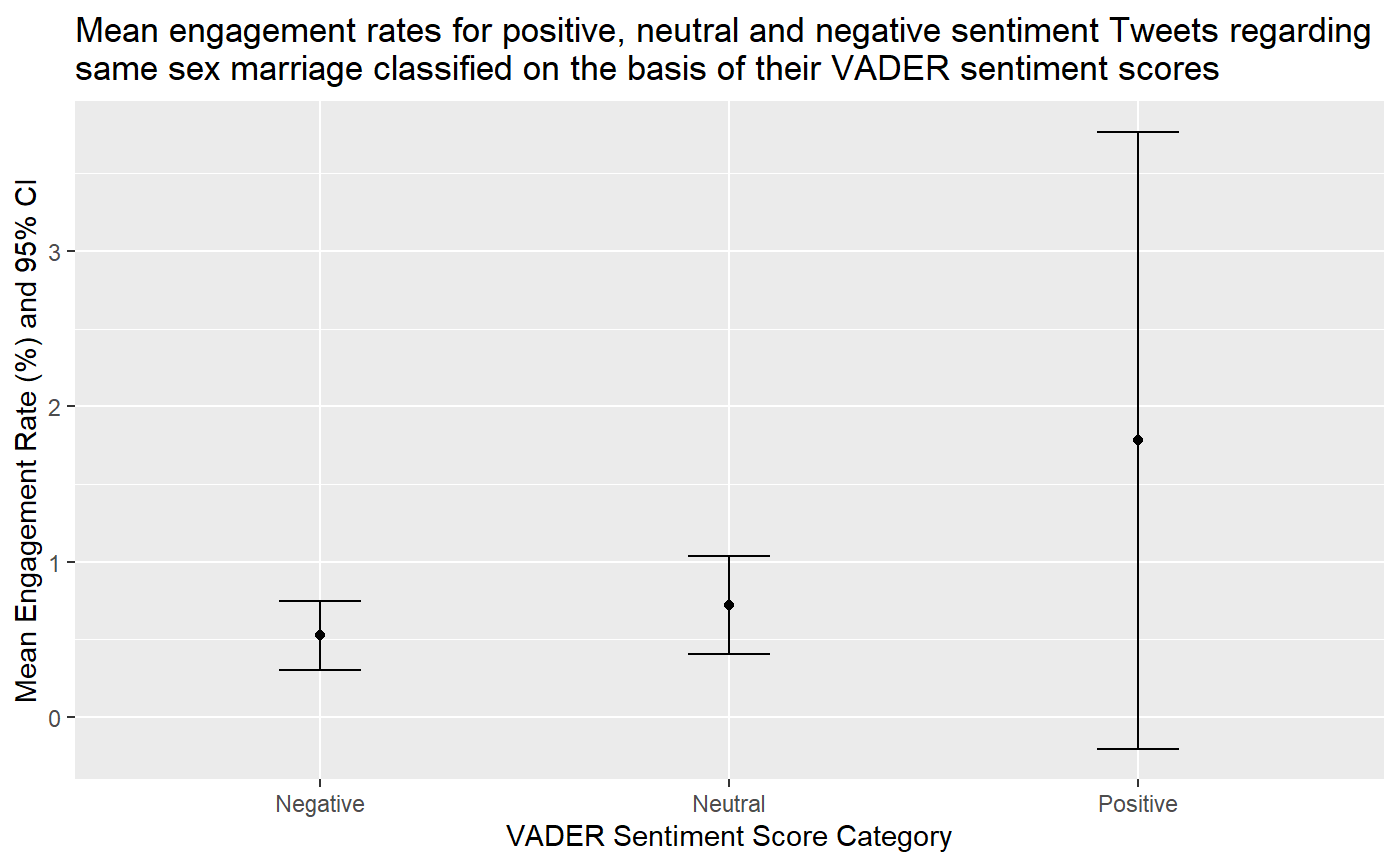


Simple linear regression was used to test if positive Syuzhet sentiment scores of Tweets significantly predicted the engagement rate and no statistical significance was found (β = -0.01916, 95% CI [-0.13155, 0.09323], p = 0.733, adjusted R2 = -0.01761). Assumptions of the linear regression model were checked.

No plots or analysis for negative emotion Syuzhet scores as there is only one data point.

**Is there a difference between engagement rates of tweets containing positive and negative emotional language?**

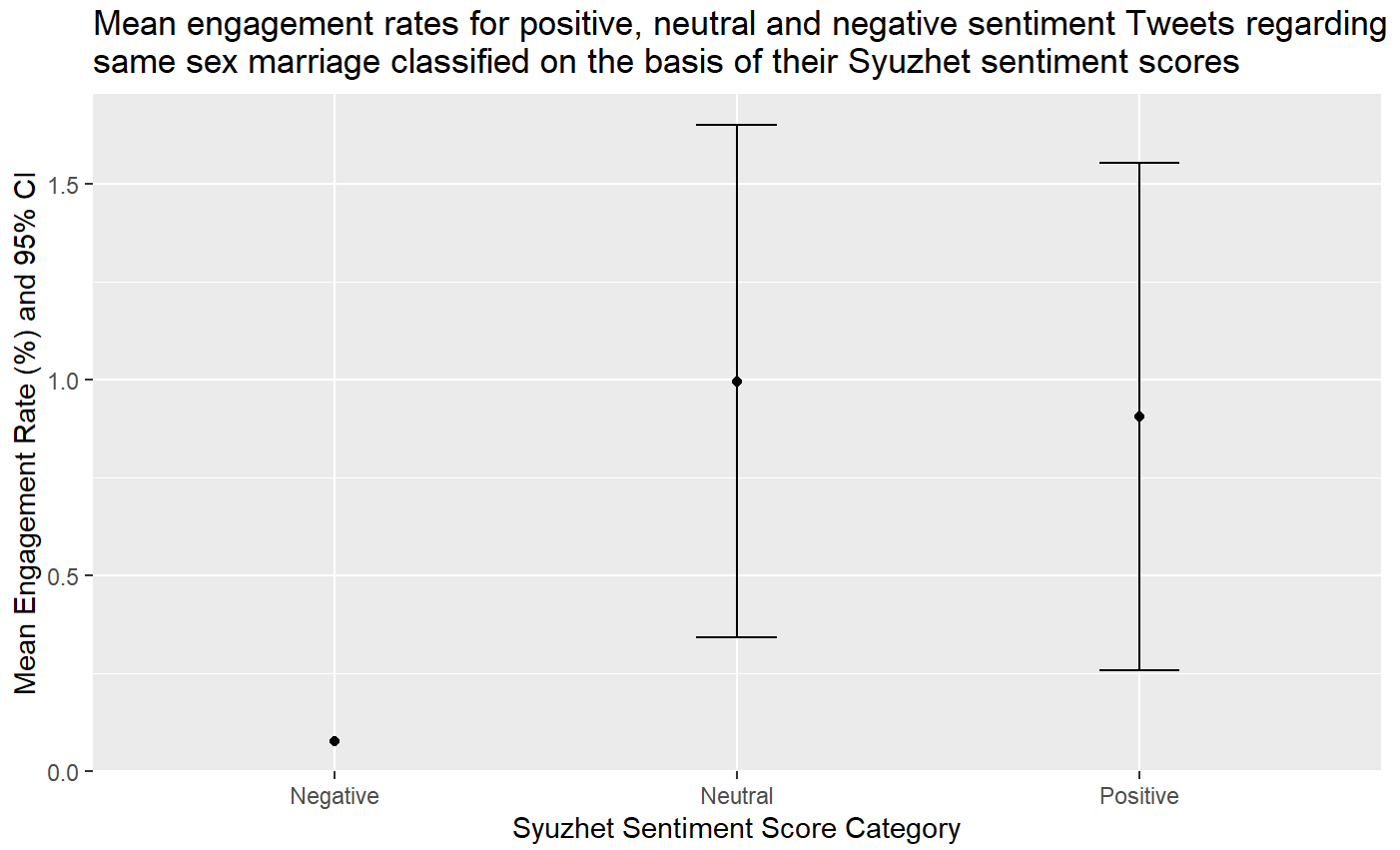
Plots of mean engagement rate showing 95% CI for three categories of sentiment (positive, neutral and negative):



Negative = scores from -1 to -0.5; neutral = scores from -0.4 to 0.4; positive = scores from 0.5 to 1.

To determine which statistical test to perform, data was tested for normality. Engagement rates of Tweets with negative and positive VADER sentiment scores were found to have non-normal distribution using Shapiro-Wilk's test (W = 0.60914, p-value = 9.618e-11; W = 0.14867, p-value < 2.2e-16; respectively).

Wilcoxon rank-sum test was carried out to test if the mean engagement rates of Tweets with negative and positive VADER sentiment scores are significantly different from each other and no significant difference was found at 5% significance level (W = 2293, p-value = 0.1238).



Negative = scores from -5 to -2.5; neutral = scores from -2.4 to 2.4; positive = scores from 2.5 to 5.

In this case, there was only one negative sentiment Tweet, hence why there is no CI bar.

To determine which statistical test to perform, data was tested for normality. Engagement rates of Tweets with positive Syuzhet sentiment scores were found to have non-normal distribution using Shapiro-Wilk's test (W = 0.40057, p-value = 2.936e-13). There was only one data point for negative Syuzhet sentiment score so no test was carried out.

Wilcoxon rank-sum test was carried out to test if the mean engagement rates of Tweets with negative and positive Syuzhet sentiment scores are significantly different from each other and no significant difference was found at 5% significance level (W = 33, p-value = 0.6709).

—-----------------------------------------------------------------------

**Exploratory Analysis**

Note: Retrieved 22/01/09 20:30, 1219 entries before filtering

Top 5 tweets according to engagement rate:

1. *@magi\_jay It used to be almost taboo in Republican circles to look to any other country for an example - witness the faux fainting when Justice Kennedy cited how other countries were addressing same sex marriage. But hey, this year's CPAC is in Hungary to kiss Orban's ring...*2 retweets, 1 followers, 200% engagement rate   
   VADER score 0.710, Syuzhet score 1.20  
   Not a famous person, not a lot of followers, not informative.
2. *@fahad\_s\_ali My entire Iranian family voted yes on the same sex marriage survey, my anglo LNP voting inlaws voted no. This ridiculous assumption that middle eastern people are automatically homophobic is very effectively weaponised by the Zionist movement.*  
   18 retweets, 134 follower, 13.43% engagement   
   VADER score 0.791, Syuzhet score 1.25  
   Reply to a user with over 4k followers.
3. *@Aijay\_Nwoye @blaq\_ko @ojay\_obin Sex is as important as any other thing in marriage. Same way the woman feels protection, care is very important in marriage, same way is sex to men. You know this, I know this. No communication will cure an intentional offence*3 retweets, 25 followers, 12% engagement rate   
   VADER score 0.401, Syuzhet score 3.45  
   Not about same sex marriage.
4. *“if legal ang same sex marriage sa future, I would marry you for fun” @notbanana07*  
   1 retweets, 9 followers, 11.11% engagement rate   
   VADER score 0.128. Syuzhet score 2.30  
   Not a famous person, not a lot of followers, not informative.
5. *Friendly reminder: @RebeccaforReal view on same sex marriage 👇* [*https://t.co/jsGmo0FBnl*](https://t.co/jsGmo0FBnl)16 retweets, 148 followers, 10.81% engagement rate   
   VADER score 0.494, Syuzhet score 1.45  
   It mentions a twitter user with over 16k followers.

Interestingly, the top 4 tweets with lowest engagement rates are all from the same user - The Economist. All have positive sentiment but low engagement rates (low rates are not just due to time of posting and data retrieval).