

Climate Change Twitter Buzz and Recycling

Team Datasaurs 

Katie Friis, Luna Ito-Fisher, William Jurayj, Kendrick Tan

Introduction and Hypothesis

Social media is commonly used to raise awareness about issues such as climate change, often with the intent of spurring action from followers. We sought to investigate whether posts on social media actually lead to more climate conscious behavior. To answer this, we looked at Twitter engagement on the topic of climate change and recycling diversion rates in three U.S. cities--Seattle, Providence, and Buffalo. Specifically, we tested the following hypothesis: **greater Twitter engagement on climate topics leads to more recycling.**

Data

Twitter

We used the Twitter API to query tweets containing the word ‘climate,’ posted between 2013–2019, and geotagged in one of the three cities. We chose to query for “climate” to encompass a multitude of phrases: “climate change,” “climate science,” “climate catastrophe,” etc. To quantify the false positives we were getting from this query, we took a random sample of 332 of our queried tweets and found:

- 81% of the tweets were about climate change being a problem
- 17% of the tweets were not related to climate change
- 2% of the tweets were denying climate change

We opted to accept the 19% false positives rather than miss relevant tweets by using a more specific search.

Outliers in our data included spikes in tweets and engagement across all three cities during August, September, and October 2019. We found this was likely due to the September 2019 international climate strikes, which garnered heavy social media attention. We chose to leave these outliers in our dataset, since they were explainable by the protests and were relevant as social media buzz.

Recycling

We used recycling datasets published by each city and calculated monthly diversion rates as $100 \times (\text{total weight recycled}) / (\text{total waste generated})$. When plotting the diversion rate by month, we noticed a cyclic pattern, with diversion rates spiking in November, December, and January each year (Figure 1).

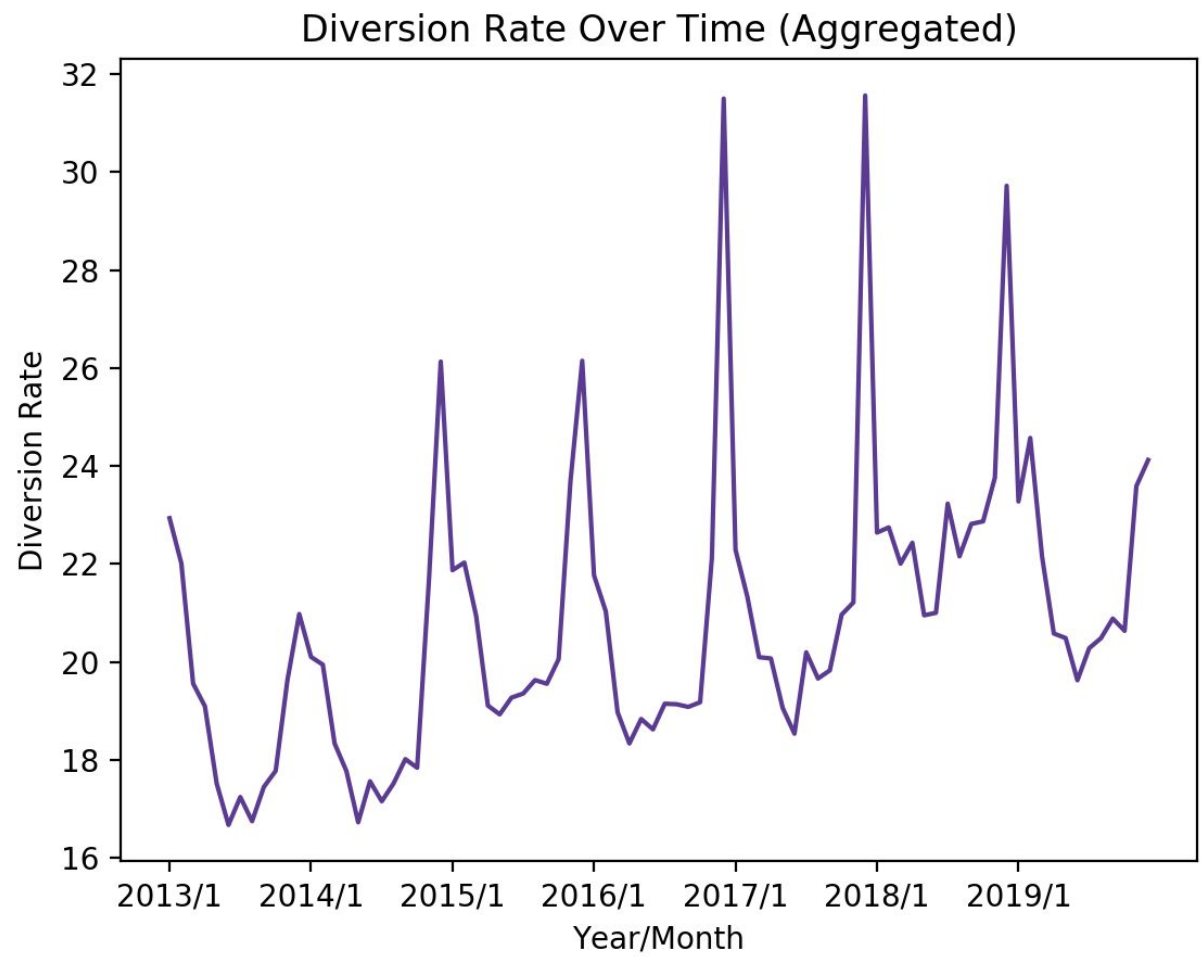


Figure 1. Monthly diversion rates in all three cities from 2013 to 2019.

We found studies showing there is an increased amount of waste and recyclable material during the holidays, which explains this pattern.

Pearson Correlation

Initially, we looked at the Pearson correlation coefficients between recycling diversion rate and each metric we had collected from Twitter (tweets, likes, and retweets), as well as engagement (likes + retweets). Engagement was used because there was a high collinearity between likes and retweets. We computed these coefficients for each city and for the aggregated data (Figure 2).

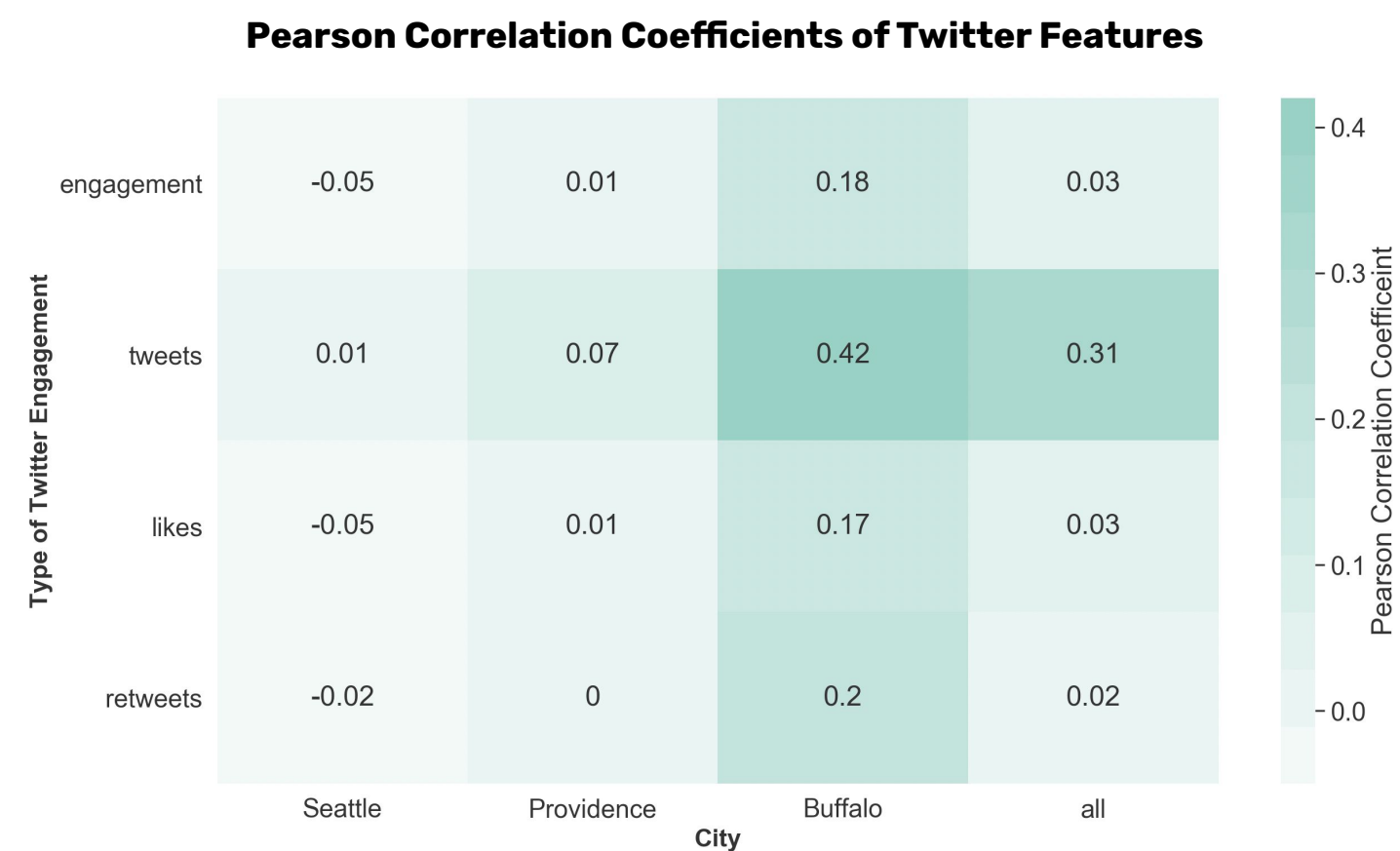


Figure 2. A heatmap of Pearson Correlation Coefficients for Twitter buzz and recycling diversion rate, for each city and Twitter metric.

We then calculated the p -values for the correlation coefficients, and used a significance level of 0.05 to interpret them. Overall, there were stronger correlations in Buffalo, but the only significant correlations were in Buffalo with number of tweets as the Twitter metric, and in the cities aggregated also using number of tweets.

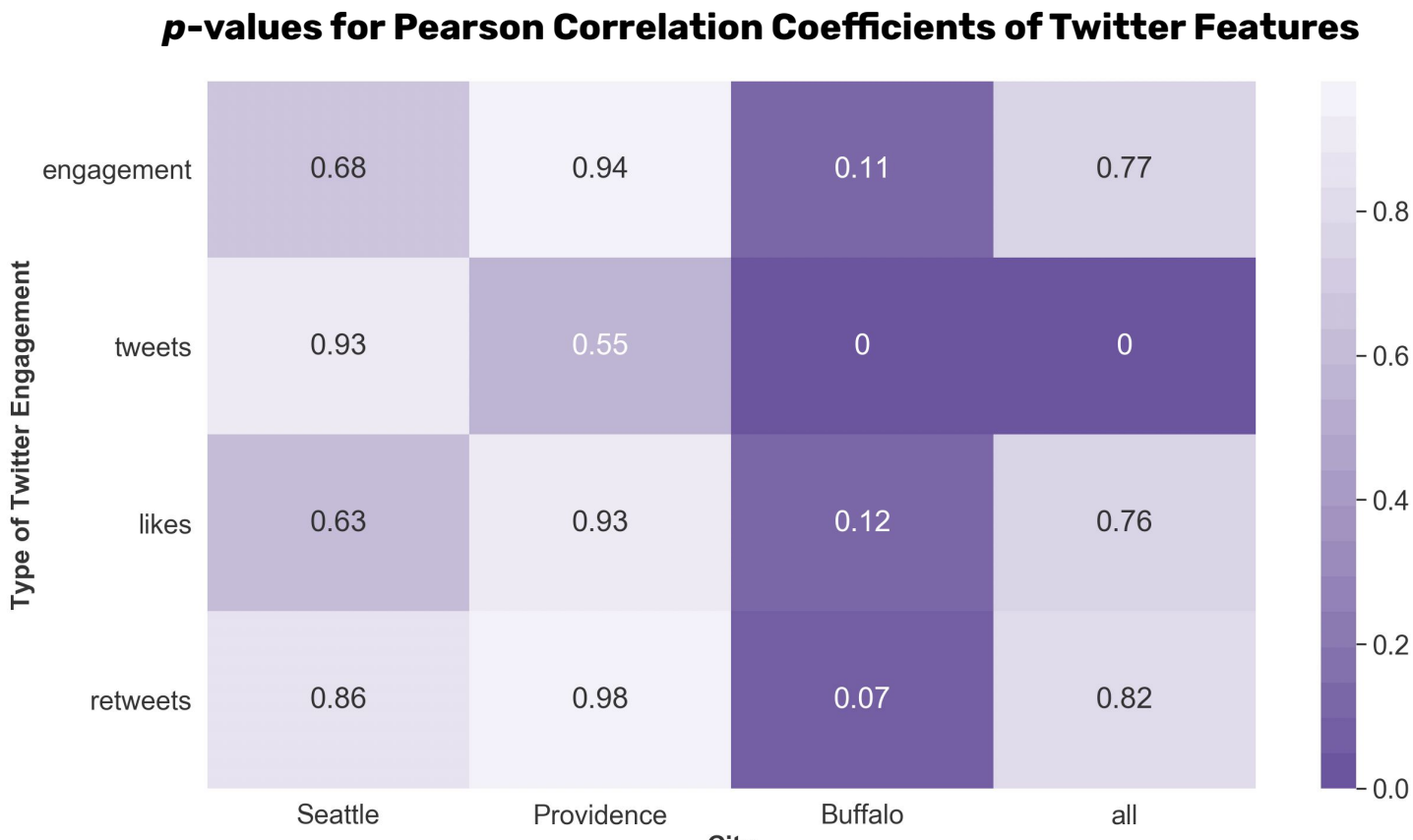


Figure 3. A heatmap of p -values for Twitter buzz and recycling diversion rate, for each city and Twitter metric.

Conclusions

After introducing additional variables, the significant relationships we saw during the first analysis proved to be insignificant. This leads us to conclude that our hypothesis does not hold, and that **Twitter engagement on climate topics does not lead to more recycling.** Rather, in all the cities, the holiday season seems to be a bigger driver of recycling than Twitter engagement on climate topics.

Buffalo initially showed significant correlations with the number of tweets, but this went away once we added the campaign and holiday dummy variables. This leads us to believe that the initial correlation was mostly reflecting the effect of the city’s recycling initiative, rather than showing a relationship between Buffalo’s recycling practices and Twitter activity.

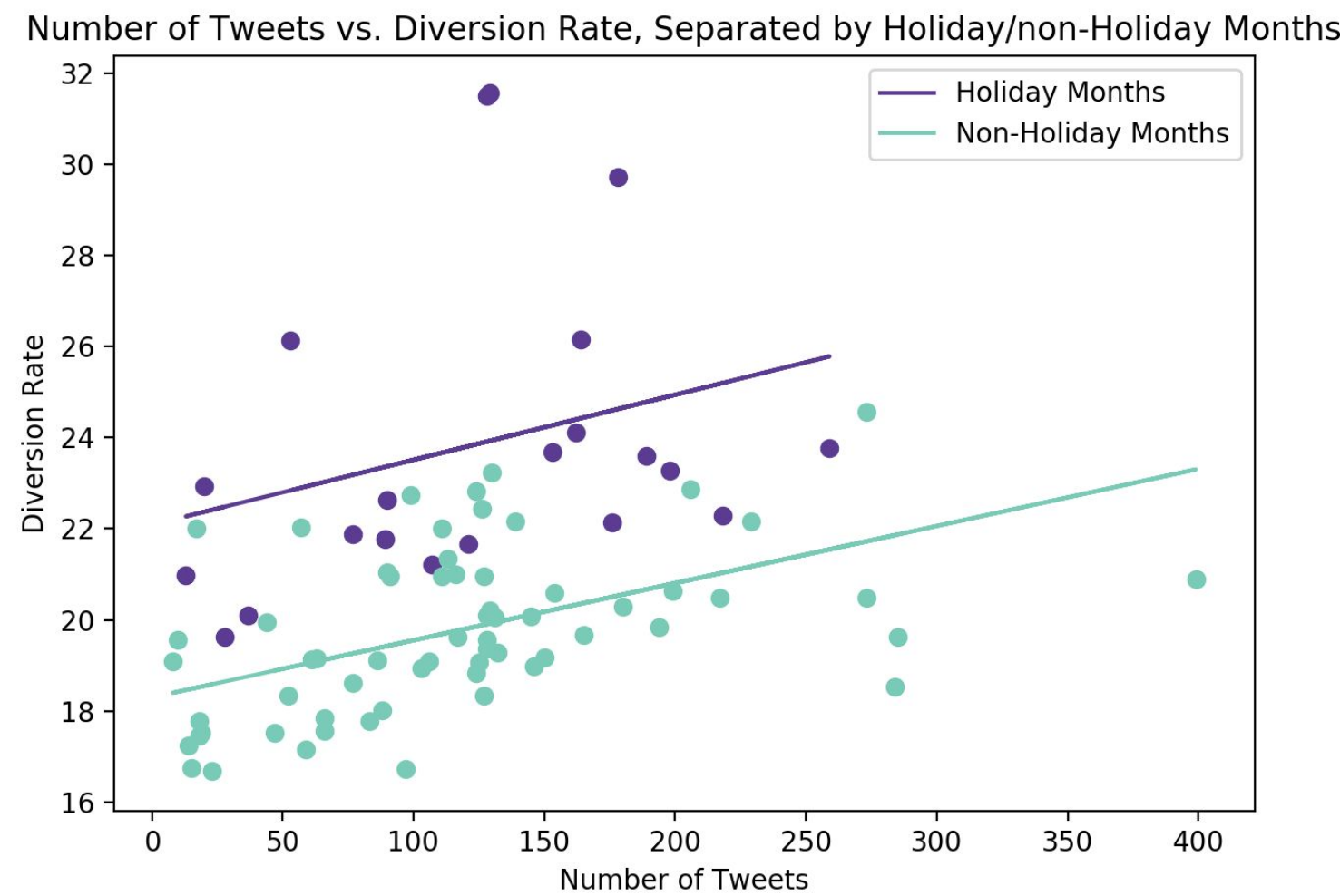
Our findings are relevant for people and groups trying to encourage more climate conscious behavior. Our findings suggest that posting on Twitter is not an effective method of encouraging recycling, and that efforts should be focused on other mediums.

Multiple Linear Regression

We proceeded to investigate these two significant relationships using multiple linear regression, seeking to understand what factors may have lead to their significance.

Aggregated cities: Recycling vs. Number of tweets

We looked into the significant correlation between recycling and number of tweets in the aggregated cities and considered additional variables that may be important to the relationship. From our initial data exploration, we knew that the recycling rates spiked during holiday seasons (Figure 1), so we introduced a dummy variable for whether or not it was the holiday season (Nov, Dec, Jan).



After controlling for holidays, there was little correlation between diversion rate and number of tweets, whereas holidays showed a much stronger correlation.:

Aggregate	Coefficient	p -value
Holidays	3.9479	0.000
Number of Tweets	0.0132	0.001
Engagement	-8.701e-05	0.294
Intercept	18.3204	0.000

Buffalo: Recycling vs. Number of tweets

It seemed unusual that Buffalo across the board showed more significant relationships, so we researched other factors that may have led to these differences. We found that since 2015, the city of Buffalo has been running the *34 and More* campaign to bring local recycling rates above the national average of 34%. To account for this, we created a dummy variable for whether or not there was a recycling campaign going on.

Buffalo	Coefficient	p -value
Holidays	6.5074	0.000
Campaign	7.4568	0.000
Number of Tweets	0.0902	0.329
Engagement	0.0007	0.808
Intercept	8.8866	0.000

After adding the campaign and holiday season variables, we found no significant correlation between diversion rate and Buffalo’s Twitter metrics. This is consistent with the findings for the other cities.

Limitations

- Three cities are not representative of the country, let alone the world. Further research across more locations would be needed to conclude our hypothesis universally.
- Only 1–3% of all tweets are geotagged, meaning there are many more tweets about climate change that we did not query. And, people look at tweets that are posted by people outside of their own location. If recycling data could be found for a comprehensive set of cities, the geotag aspect could be removed to fix this issue.
- Querying for “climate” is not exhaustive of the Twitter conversation on the topic. Querying for additional keywords may include more relevant tweets, as well as encompass terms that in the past were more common like “global warming.”