

Tweether: a Visualization Tool Displaying Correlation of Weather to Tweets



Fig. 1. Tweether showing the real time mood of Nebraska.

Abstract—As the generation of social media we can instantly express how our day is going, however unknowingly the weather can play a key role in how we are feeling. The weather dictates our lives regardless of what may be happening. The relationship between weather and mood has been immensely studied to show that the weather does play a major factor in regards to our emotions. However the real question is how much it affects us and do we display it on social media for the world to see? Based on the natural correlation between weather and moods we propose Tweether a real time weather and tweet visualization to see how Twitter users in the state of Nebraska are feeling. This visualization displays a current reflection of the emotion in select regions of Nebraska and also predicts the possible emotion in those regions for the next 72 hours in response to the weather forecast. The visualization uses multiple layers to show the connection between the location, weather, and emotion. Aggregating multiple users with similar emotions the design is free of clutter and is straightforward to understand. Limiting the emotions to two categories(positive and negative) creates an aesthetic design that is easily viewable in a 3-dimensional manner.

Index Terms—Time series, weather, clustering, sentiment classification, prediction, line bundling, correlation

1 INTRODUCTION

Weather affects our daily lives, from what we wear, what activities we do, what type of transportation we use, what we eat, or even how we feel. With the increasing accuracy in weather forecasts people can gain an idea on the what type of weather they can expect for upcoming days. Activities are usually planned according to the weather outside (i.e. weddings) and alternative plans must be made in case of inclement weather. How people dress is also affected by weather; when the temperature drops people need to wear coats to stay warm. The economy is also greatly affected by the weather. Certain weather conditions can lower crop yield and cause higher prices in stores. Disastrous weather phenomena such as hurricanes, tornadoes, or even floods can cause devastation in communities resulting in homelessness, death, and destruction. Inclement weather can also cause delays in transportation on roads or for flights. We can also choose to ride our bike to work instead of driving the car if the temperature is warm enough. One thing that is an effect from all these items is how we feel.

- Are you sad that you can't enjoy the outdoors because it's raining?
- Do you love that it's raining so you can bundle up and read your favorite book?
- Do you love the snow because it's close to Christmas?
- Do you hate the winter because you want it to be spring?

These feelings are all brought out by the weather outside. One person can feel positive about a certain type of weather and one person can feel negative. In this work we showcase a novel tool, named Tweether a visualization of real time Twitter and weather data to show the feelings of current users and how their emotions could fluctuate. Having the weather forecast up to 72 hours in the future, the emotions in current regions can be predicted.

We see if the weather has any correlation to the majority of the population and we try to predict the future feelings of given weathers. When thinking about warm and sunny weather we automatically assume majority of the population will be happier in comparison to dreary or cold temperatures. We plan to examine if majority of the population follows this pattern. We will also examine how the overall sentiment changes when we filter tweets so only tweets in regards to weather are displayed.

It is not enough to just determine the correlation between emotion and weather, but a novel visualization is necessary. Our work showcases a 3-dimensional map of Nebraska which highlights select clusters of weather. The correlation of tweets to weather is represented by line bundling. Introducing a clear relationship between weather and tweets the design presents a natural manner of representing correlation.

2 RELATED WORK

Visualizations correlating sentiment and weather is highly sparse; the presence of live visualizations is also non-existent. However, there are works showing the two portions of this work. Clustering of weather data has been done many times in the past. There is also a vast number of visualizations which indicate sentiment of different locations.

2.1 Psychological studies

The natural correlation of weather with emotion has been studied profusely. [1,2,3,4,7] With various factors among the different research the conclusions attained were broad. Humidity, sunshine, and temperature have the greatest effect on mood. [1]

In some research weather and mood, correlations have been debunked. There isn't consistency due to seasons and time spent outside.[2] Having certain emotions regarding the season has strong links to seasonal affective disorder (SAD), where people are depressed in regards to changes of the seasons, which usually occurs during the winter. However, most psychologists believe that the weather has an impact on psychological intentions. [3] When observing serotonin levels in regards to sunshine there were strong relationships to being happy. [4] It has been found that weather may not play a big role in the positive attitude, but the negative attitude can have a correlation with weather.[7]

2.2 Social Media, Weather, and Emotions

When dealing with the correlation of weather with emotion the research is fairly sparse. The works present use Twitter data for their social media feed and some form of weather data. However in comparison to our work the following research was computed on past data and used a 2-dimensional graph implementation to visualize the data.

Work has been done using two to four years of Twitter data and correlating it with meteorological data from NOAA. [3,6] Using urban areas in the United States as the area of interest, the tweets are passed to a sentiment analyzer that has a multi-level process. They first determine keywords which are identified from public events (i.e. entertainment or natural disasters), they then identify the mood state, and finally assign sentiment scores. To correlate the weather with the tweets they use a Generalized Mixed Model to display the non-linear relationship between emotion and weather. Using multiple variables for weather (temperature, temperature change, precipitation, snow depth, wind speed, solar energy, and hail) they determine the connection to hostility-anger, depression-dejection, fatigue-inertia, and sleepiness-freshness. Their results indicate that the warmer temperatures create an angrier atmosphere, lower depression, and less sleepiness, and they determine the influence of temperature to mood is trivial. Their visualization is limited to graphs. [3] Other than using urban areas in USA we see relationships between temperature, humidity, and atmospheric pressure for one month's tweets in the United States and weather data from Weather Underground. Using Linguistic Inquiry and Word Count for sentiment classification they saw a pattern with temperature and emotion of every state in the United States. Using regression analysis they find that the warmer states had a happier mood than the colder states. Their visualization was limited to a bar chart.

These works are limited in visualization and usability. Using past data is useful for our training and testing model however having a live view of what Twitter users feel in Nebraska is what we aim for in this work.

2.3 Sentiment analysis in Social media

Sentiment analysis is a topic that has been studied vastly. There are various methods to detect the sentiment of a sentence, however in re-

gards to tweets we are not using complete sentences. Since tweets are limited to 140 characters the need to express oneself is limited to short meaningful phrases. We find abbreviations, neologisms, acronyms, hashtags, emoticons and URL's throughout most tweets.

Certain features need to be extracted and some need to be filtered out. Filtering URL's, usernames, Twitter special words, and emoticons.[6] Stop words(i.e. a, an, the) are also removed due to not adding any extra sentiment information. For classifying the tweets, a number of different methods are used, however the most prevalent one is Naive Bayes classifier.[6] Emoticons are used as basis for sentiment classification for classifying tweets as positive or negative in the training purposes. [5,6]

Using emoticons for classifying training data is novel, however, most tweets gathered in the live feed had a very low count of emoticons. So using that method was out of question. Filtering of URL's, usernames was used for our method as well. We also added to the filtering process to drop phrases beginning with hashtags.

2.4 Visualizations

2.4.1 Clustering

Clustering of data is a trivial task. However when introducing temporal and time series data the task becomes non-trivial. Weather data needs to be clustered based on a value, proximity, and the changes throughout the given time span. There have been various visualization techniques for time series data. Visualizing time series data using spirals for large data sets is used to better identify periodic structures in data. [9] Using wavelet to transform data along a multiresolution temporal representation to find clusters with similar trends is a useful method for exploring data in a time series fashion.[8] Using smooth data histograms for visualizing clusters in self-organizing maps is a simple method when using 2-dimensional data sets. [10]

The mass majority of clustering visualizations use k-means clustering on the basis of their algorithms.[8,11] We choose to follow this pattern as well. The user's dilemma of what clustering information is needed provides insight as to why k-means is widely chosen; primarily due to it being a simple clustering algorithm.[11]

2.4.2 Bundling

2.5 Predictions

Using the past Twitter data set it is necessary to predict the mood for the next days. Predicting the stock market based on Twitter mood has been studied. [12] The notion that the mood of Twitter users can correlate to stocks immediately is not present, however, the mood is reflected when a few days have passed. Since the general public has strong connections with the outcome of a man-made entity there will be some form of correlation available. In our case, however, the weather is not a man-made entity, so finding a correlation between weather and mood, and then predicting the mood for the future can show zero correlations. We will address this idea later in this paper.

3 IMPLEMENTATION

Implementing Tweether has multiple steps which are composed of individual and interconnected parts. The steps are illustrated in Figure 2. Tweether in its simplest form takes tweets and assigns a sentiment value which is then correlated to the nearest weather cluster. Each tweet is aligned to the map of Nebraska according to its location. The current hour visualization and the prediction visualization have the same user interface.

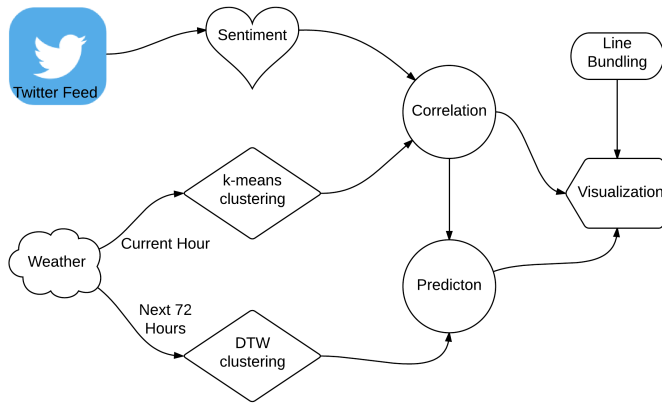


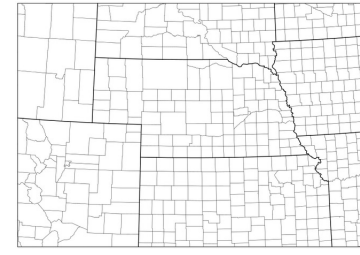
Fig. 2. Steps of Tweether

3.1 Data Set

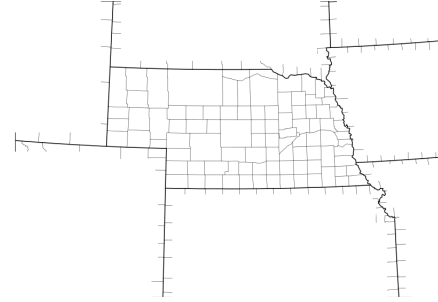
There are two main data groups in this visualization.

The weather data set is provided using the Weather Research and Forecasting (WRF) Model from the Holland Computing Center (HCC). The data is pushed every six hours with a forecast of up to 72 hours in the future. Each WRF file contains multiple variables in regards to weather (i.e. temperature, precipitation, wind speed, etc.). For this visualization, we choose to focus on the TSK variable which is related to the surface skin temperature. Each TSK file is represented via a two-dimensional array that is 34 x 24. The WRF file geographically corresponds to Nebraska and surrounding states as seen in Figure 3.

The Twitter data is the live data feed from Twitter users throughout Nebraska. Only users that have opted-in to use Tweeting With Location feature turned on are used. A geographic filtering process was used to use only select cities in Nebraska. Locations for the tweets are limited to the most populous cities in Nebraska (Omaha, Lincoln, Grand Island, Kearney, Fremont, North Platte, Norfolk, Columbus, and Scottsbluff). Since Nebraska has a fairly low population with most of the land being barren, only the most populous cities were chosen. The Twitter data is stored in JSON format where we need to extract the coordinates of each tweet and the tweet itself. Due to cities being on the border of Nebraska such as Omaha, the Twitter data needs to have a second filter which removes any tweets that don't have a relation with Nebraska.



(a) Map showing surrounding states with counties



(b) Map reduced to show prominence to Nebraska

Fig. 3. Nebraska map before and after

3.2 Clustering

The clustering differs depending on if looking at the current hour or the predicted values for the next 72 hours. For the current hour, the k-means clustering algorithm is used. Since we don't need to worry about the future at the current state nothing additional is added to the algorithm. For predicted moods, we use the dynamic time warping (DTW) algorithm.

The k-means clustering algorithm partitions the matrix of data into five clusters which can be dispersed. In the matrix, the dispersed clusters can create random patterns or create outliers. To remove outliers, they are filtered out creating fewer regions of the same cluster. The filtering process is based on the number of data points in each cluster. If there exists a cluster which has less than one percent of the overall number of clustered elements then this cluster is removed and the remaining data points which belong to a cluster are clustered again. This process is done as long as there exists a cluster which has less than one percent of the overall cluster count.

For the forecasted weather, a modified k-means algorithm was used originally. Since k-means isn't very robust towards outliers due to adding a square weight on the value, we used dynamic time warping to measure the similarity between two hours.

To smooth any randomness, connectivity is used to determine the percentage of surrounding elements with the same cluster value as the central element. If at least five out of 8 neighbors have the same value then this cluster is kept, otherwise the data point is removed. This method is repeated four times to ensure smooth clusters. The number of times to be repeated was chosen due to it creating the most smooth outcome amongst the majority of 73 hours.

3.3 Semantic

Each tweet can contain different attributes other than plain sentences. Due to each tweet being limited to 140 characters the majority of users tend to use abbreviations, neologisms (i.e. noob, troll), acronyms, hashtags, emoticons, and URL's. Abbreviations, acronyms, and neologisms are taken into account for training our classifier, however, a few items are filtered from certain tweets. The filtering process removes emoticons, URL's, usernames, and hashtags. In some situations, it is known that hashtags can provide instant insight as to what the users are feeling.[6] However, we feel with most hashtags that we encountered they contained useless text or sentences for tags instead of keywords.

A Naive Bayes classifier is used to determine the sentiment of tweets. Using Robert Plutchik's theory, it states that there are eight basic emotions:

- negative - fear, anger, sadness, depression, disgust
- positive - joy, trust, anticipation, surprise

These emotions are the basic training portion of the classification of tweets. The synonyms for each category are taken into account and this sets up the basic foundation for the tweet classifier.

Other than acronyms we also need to take into account profanity. The use of profanity in social media is very high and it may lead to a positive or negative emotion depending on the situation. To take into account how profanity is used in sentences, tweets were polled to see how these words were used. Using these tweets to train the classifier gave a pretty high accuracy rate in regards to profanity. In the beginning we tried to remove any tweet with profanity, however, this drastically lowered our tweet count. We then tried to remove occurrences of profanity in the tweet and use the remaining words as a judge of emotion. This worked in a few cases however for the most part we believed that profanity gave insight to negative moods so we decided to take into account profanity.

Due to tweets using abbreviations and incomplete sentences, sentiment calculation is a nontrivial task. The classifier is trained using around 10,000 tweets, where each tweet was given a positive and negative score, and there were rare occurrences of duplicate tweet, and it was made sure that there was an equal portion of positive and negative tweets in regards to words where sentiment could go either way.

3.4 Correlation

The correlation between each tweet and the cluster above is represented by a one to one mapping. If there is a natural link then the mapping is pure. In situations where there isn't a direct link to any of the weather clusters the nearest cluster is used.

Other than the natural link between the primary cluster and the tweet we explore similarity of connections to other clusters regardless of the sentiment it may correspond to. This correlation is only available for the current hour. For each tweet, we use the top two relations to other clusters which we deem the secondary and tertiary clusters. Using the Pearson product-moment correlation coefficient, we determine the similarity of each tweet location to other clusters. With a correlation value closer to one we keep the sentiment expressed in the primary cluster, however in the case of a correlation value closer to negative one we swap the sentiment expressed in the primary cluster.

3.5 Line Bundling

3.6 Prediction

Determining the mood of the current locations up to 72 hours in the future is a nontrivial task. This prediction technique is completely based on the current hour and the previous day. We choose not to use data from earlier times since the trends today are definitely not the same a year ago let alone a month ago. We again use Pearson product-moment correlation coefficient to determine the correlation between the weather of the forecasted hour and the current hour. Then we use the same algorithm to determine the correlation of the forecasted hour to the previous day. With these two values, we determine sentiment value. In this situation, the current hours value has a slight higher impact on the forecasted hour than the previous days value. This has to do with keeping the most current sentiments relevant. Using these values, we can easily predict the sentiment of the select locations.

4 VISUALIZATION

The novelty of this work is pictured through the 3-dimensional visualization. The visualization is comprised of multiple layers where there is a natural visual connection between the layers. Many different methods could represent this data correlation however, we have created a novel visualization. The layers are connected via line bundles to dictate how the majority of the Twitter users in Nebraska are feeling.

4.1 Layers

The visualization is represented by three main layers. The base layer contains the map of Nebraska and neighboring states. We have chosen to show the counties in Nebraska to gain insight on the population distribution in the state, however we choose to keep the other states plain to indicate emphasis of correlation to Nebraska as shown in Figure 3.

On the map of Nebraska, there are circles which indicate the area that Twitter users are posting from. The radius of the circle corresponds to the amount of people posting tweets in the vicinity. Using the WRF data set for the longitude and latitude values for the map, we obtained the intermediate values via interpolation. For each tweet location, the euclidean distance is calculated to get the nearest longitude and latitude location.

The middle region contains the line bundles which starts from the Twitter user location and ends at the weather cluster they correspond to above depending on the sentiment. The topmost layer is the weather clusters.

The middle region contains the line bundles from the Twitter user location to the weather cluster they correspond to above depending on the sentiment. The topmost layer is the weather clusters.

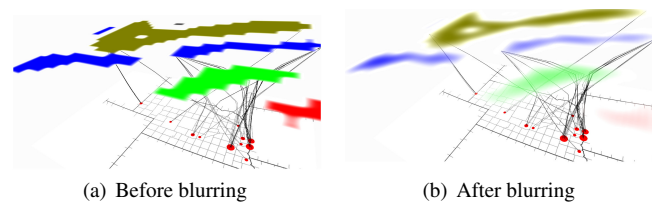


Fig. 4. Clustering of the weather. In this hour we see five different clusters. The blue cluster is seen in three different locations.

4.2 Cloud Representation

Each cluster is represented by a different color. Clusters which aren't connected of the same color indicate the same conditions in other locations of Nebraska as seen in Figure 4.

Once the cluster for each hour was attained we needed two edge points for the data to link up to. These two edge points correspond to positive and negative emotions. We had different ideas on how the points should be picked. At first we were thinking to give half of one cluster to the positive emotion and the other half to negative emotions. This method seemed like a good idea at the beginning however in the visualization the output looked cluttered.(Figure 5) The second

method we tried split the cluster in half and used the center points of the two half's. This was also a good idea until we came into situations where the cluster was represented by a very small amount of data points causing the visualization to look awkward. (Figure 5) We finally decided on a method where for each cluster the positive tweets would map to top rightmost point in the cluster and the negative tweets would map to the bottom leftmost point. This reduced visual clutter and helped in situations where the cluster size is fairly small. (Figure 5)

In our visualization, the clusters appear as clouds. Initially, we had the user space and the weather space flipped, however, we believe that the way we present the data is novel and more intuitive. Each cluster is blurred at the boundaries to make the graphical interface cleaner and aesthetically pleasing as seen in Figure 4(b).

4.3 Line Bundling

Each bundle is composed of multiple lines where each line represents one tweet. Every bundle represents positive or negative sentiment. The sentiment is represented in two ways. Depending on which corner the bundle is linked to indicates the sentiment, also the more intuitive notion is the color of the bundle. Naturally we associate the positive sentiments to the green bundles and the negative sentiments to the red bundles.

The opacity of each bundle represents the intensity of the correlation between the cluster and the tweets. The bolder the line the stronger the relationship between the two points is. The primary cluster has a stronger hue in comparison to the secondary and tertiary clusters. Both of these links will have a staggered decrease of intensity of color. Thus, we state that the more prominent the link between the tweet and the cluster the higher the opacity of the line.

5 FINE TUNING

5.1 Scalability of Line Bundles

6 CASE STUDIES

For our work, we have two goals. We first want to determine if there is any correlation between weather and mood. Secondly we would like to determine if our predicted mood is correct.

6.1 Tweet weather correlation

We use two data sets to show our visualization. Using weather and twitter data from two back to back weekends we see if there is any relationship between the temperature and the overall mood of people. We are lucky that the days we chose show warmer changes in weather. Being on the brink of Spring we predict that the overall Tweets for the warmer weekend will have a more positive sentiment in comparison to the colder weekend. This weather pattern also can show any correlation we find for Seasonal Affective Disorder.

6.1.1 All Tweets

Of all the cities, we see that Omaha and Lincoln have the most tweets, so we will use these two cities to display our results. The weather patterns of the two weekends are shown in Figure 6(a). As seen in the figure, we choose to look at times where most of the population is awake (i.e. noon to 11 PM). From Figure 6(b) we can see the tweets and their sentiment for select days through the two weeks, where more emphasis is placed on the first week and less on the later. Each date has a positive and negative count, where we see that the pattern throughout is that the positive tweets outweigh the negative tweets. Another pattern that we see is that as the temperature increases the number of tweets increase however we see that more people tweet on Saturday than on Sunday. We think that this has to do with a number of factors; people go to church on Sunday where they choose to be among family, they do work around the house, or they don't have any eventful things which they feel they need to share.

We do not see that with the change of temperature that there is a clear trend of more positive tweets versus negative tweets. The number of positive tweets is at least half the amount of negative tweets. The only item that we could see which may be weather related is that the amount of negative tweets is percentage wise far more in the colder temperatures in comparison to the warmer temperatures. In lieu of not finding any clear results, we look at tweets only with words related to weather to see if we find any trends.

6.1.2 Weather related tweets

Other than using all the tweets we collected we filter the collected tweets for weather-related terms. (i.e. snow, sunny, warm, cold, rain, etc.) We wanted to see if there was a stronger correlation in this situation compared to using all the tweets collected. In this situation, we are unfortunately limited to a few tweets. As we see in Figure 6(c) we see that the number of tweets drastically reduces to at most slightly less than 160 tweets for one day.

Fortunately, we are able to see a pattern with weather and positive tweets. There is a spike of tweets when there is warmer weather. This pattern is seen in both sets of tweet data sets. However, we are now able to confirm that the correlation is due to weather. As the weather increased slightly on March 1st we see a spike in amount of tweets overall(Figure 6(b)) and also when the tweets had weather-related words(Figure 6(c)). We can also see that in Figure 6(c) that with the warmer weather the number of tweets in relation to the weather increased by approximately twice the amount. We still see the trend of more positive tweets in comparison to negative tweets when using the filtered data set.

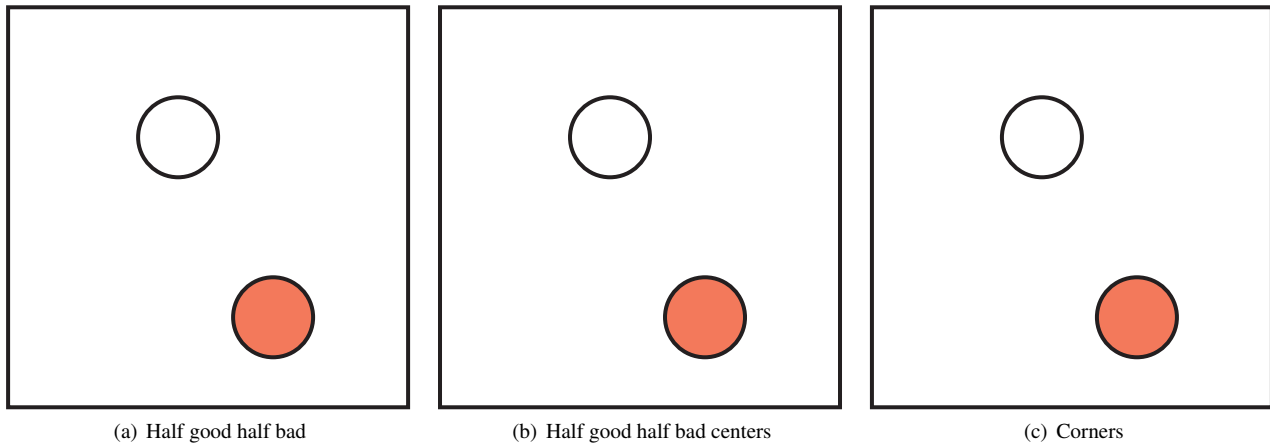


Fig. 5. Determining the best position for indicating negative and positive tweets in the cloud clusters.

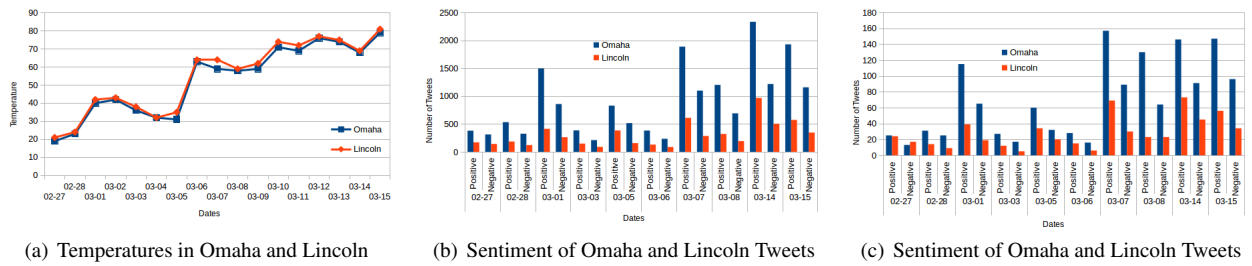


Fig. 6. The comparison of weather to sentiment.

6.2 Predicted?

Using the same data set we determine if the predicted value was close.

7 DISCUSSION

There are a few items in regards to tweets that enforces a limit as to what we can assess. One drawback is that with a state population as small as Nebraska the chance of a user with their location enabled is minute. Another limitation in some situations is tweet length since this may cause the complete expression of a user's feelings to be curbed. Other than Twitters limits we need to examine the suppressed feelings of users. Since Twitter is a social media hub, the need to express oneself in 140 characters may place a mask to what the user may be feeling. As proposed before there is no distinct way to determine the true feelings of a person.[3] Due to this phenomena we proposed the usage of only weather related tweets, due to there being a low chance of users masking their emotions. Even with this we saw that the majority of users use their tweets in a positive manner, so to expect a negative tweet is cynical in some sense.

8 FUTURE WORK

In the future, we plan to make a cleaner looking visualization. Currently, there seems to be some rigidness to the design where corners could be smoothed have a clearer indication of what places link up to where.

This work as of now only sees the relation to temperature, but this study could be extended to be applied to precipitation, humidity, wind speed, or any other weather variables. There may be a significant correlation between rain amount and negative tweets. As we approach warmer climates we can conduct research regarding what the overall sentiment of people to rain, storms, and strong wind conditions.

We also believe that we can determine if there is a better correlation of weather to other variables, possibly ones which are directly impacted by weather such as road accidents, calamities from natural disasters, or even medical episodes like the flu. Since there is no real way to determine if a person's mood is truly affected by weather we think using some other variables can provide more concrete correlation results.

9 CONCLUSION

In this paper, we introduced Tweether, a visualization tool for displaying the correlation of weather to tweets. Other than making a novel visualization to showcase correlation between weather and sentiment our main goal was to determine if there was any correlation between usage of social media and weather patterns. Using two data sets, we were able to determine that there is some correlation with the amount of tweets related to weather when it is warm in comparison to the colder weather.

Visualizations showcasing social media sentiment has been done many times before, and similarly clustering of weather has been implemented numerous times before. However, there have been few to none works out there which try to find the correlation between the two.

We realized that this work can be expanded to apply to many different scenarios, but we chose to apply it to a question that has been asked for centuries. ...

Our prediction technique worked for...

The line bundling implemented in this visualization is an entity on its own and should be applied to many different works. ...

ACKNOWLEDGMENTS

The authors wish to thank A, B, C. This work was supported in part by a grant from XYZ.

REFERENCES

- [1] J. Bollen, H. Mao, and X. Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1–8, 2011.
- [2] J. J. Denissen, L. Butalid, L. Penke, and M. A. Van Aken. The effects of weather on daily mood: a multilevel approach. *Emotion*, 8(5):662, 2008.
- [3] A. Hannak, E. Anderson, L. F. Barrett, S. Lehmann, A. Mislove, and M. Riedewald. Tweetin' in the rain: Exploring societal-scale effects of weather on mood. In *ICWSM*, 2012.
- [4] E. Howarth and M. S. Hoffman. A multidimensional approach to the relationship between mood and weather. *British Journal of Psychology*, 75(1):15–23, 1984.
- [5] A. K. Jain. Data clustering: 50 years beyond k-means. *Pattern recognition letters*, 31(8):651–666, 2010.
- [6] M. C. Keller, B. L. Fredrickson, O. Ybarra, S. Côté, K. Johnson, J. Mikels, A. Conway, and T. Wager. A warm heart and a clear head the contingent effects of weather on mood and cognition. *Psychological Science*, 16(9):724–731, 2005.
- [7] G. Lambert, C. Reid, D. Kaye, G. Jennings, and M. Esler. Effect of sunlight and season on serotonin turnover in the brain. *The Lancet*, 360(9348):1840–1842, 2002.
- [8] J. Li, X. Wang, and E. Hovy. What a nasty day: Exploring mood-weather relationship from twitter. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pages 1309–1318. ACM, 2014.
- [9] A. Pak and P. Paroubek. Twitter as a corpus for sentiment analysis and opinion mining. In *LREC*, volume 10, pages 1320–1326, 2010.
- [10] E. Pampalk, A. Rauber, and D. Merkl. *Using smoothed data histograms for cluster visualization in self-organizing maps*. Springer, 2002.
- [11] M. Weber, M. Alexa, and W. Müller. Visualizing time-series on spirals. In *Information Visualization, IEEE Symposium on*, pages 7–7. IEEE Computer Society, 2001.
- [12] J. Woodring and H.-W. Shen. Multiscale time activity data exploration via temporal clustering visualization spreadsheet. *Visualization and Computer Graphics, IEEE Transactions on*, 15(1):123–137, 2009.