Mood and Weather: Feeling the Heat?

Kunwoo Park* **Seonggyu Lee*** **Eunae Kim*** **Minjee Park†** **Juyong Park‡** **Meeyoung Cha‡**

*Division of Web Science and Technology, †Department of Computer Science, ‡Graduate School of Culture Technology
Korea Advanced Institute of Science and Technology, South Korea
{kw.park,shalomeir,raito,friendpe,juyongp,meeyoungcha}@kaist.ac.kr

Abstract

The notion that weather plays a significant role in determining a person's mood is empirically beyond doubt, but quantifying the relationship on a large scale has traditionally been hampered by difficulties in measuring people's sentiments. In this paper we present the relationship between weather and the U.S. residents' sentiments from Twitter data and nationwide meteorological records. We measure how temperature, humidity, and atmospheric pressure correlate with people's moods, demonstrating the potential of large-scale online data in humanities fields such as psychology. We also provide these results with infographics, which can help the public understand scientific results easily.

Introduction

If one were to ask you how you felt on a sunny day, which would be your likelier answer: happy or sad, given other factors being equal? For a long time we have known from our experience that meteorological variables are linked to a wide range of human behaviors and emotions such as mood. Reaching beyond our personal experiences, some academic researches have found correlations between weather and mood, albeit on a relatively small scale (Howarth and Hoffman 1984; Preti 1998). As large-scale data of people's thoughts and sentiments are becoming more available, an ever more accurate assessment of the impact of weather on people's moods is also becoming possible. Given the rise in the attention paid to the significance of mental health in the modern society (Neugebauer 1999), we believe such research has ample avenue for important practical applications as well.

Ever since its birth, the Web has transformed how we study the large-scale behavioral and sentiment patterns of the online public: social media such as Twitter and Facebook especially have made the Web extremely personal, functioning for millions as the primary conduit for expressing emotions and feelings. The open and public nature of the media allows it to function as one giant laboratory, where we can observe people's actions and words on a significant scale (Lazer et al. 2009). Accordingly, analysis of Twitter users to understand their sentiments and predict human behavior has gained popularity in recent years, with applications to predicting stock prices (Bollen, Mao, and Zeng 2011), forecasting German federal elections (Tumasjan et al. 2010), as well as correlating the day of the week and the length of the day with mood (Golder and Macy 2011).

Despite the popularity of Twitter as the preferred source of sentiments in various social, psychological, and behavioral issues, studies on its correlation with weather—an empirically significant factor in determining the mood—have been relatively rare and small in scale. Notable research includes a recent work that showed how weather affects changes in tweeting rates (Kiciman 2012), indicating the existence of self-reporting biases in Twitter. Another piece of work reported that mood is indeed affected by weather, but the analysis was limited to twenty metropolitan areas and a single sentiment dimension (Hannak et al. 2012).

In this paper we overcome these shortcomings by using massive Twitter data, and report the findings on the relationship between weather and mood on a nationwide scale. We analyze the positive and negative sentiments independently in all 50 states of the US. Our main findings are, to be discussed in more detail in the remainder of the paper, that first, people show more positive affects with higher temperature; second, humidity drives people towards negative affects; and finally, in most states people feel more positive on days with higher atmospheric pressure. We also identified interesting outliers. For instance, people in Louisana and Hawaii were happier on days with higher humidity and low atmospheric pressure.

While preliminary, this paper makes contribution in that we test a wide held belief that weather somehow affects people's mood using large-scale empirical data. We identified the relationship between a wide range of meteorological readings and Twitter sentiment, and found that certain weather variables (i.e., temperature, humidity, and atmospheric pressure) indeed have predictive power and that not all states had an equal amount of correlation between weather and mood. These findings provide implications for designing web services that utilize weather information. Given that weather variables are correlated with people's moods, this knowledge can be used to better assist communication between users online. If humidity of an online friend is higher than usual, one can imagine there is some probability that his friend is feeling negative. Even though we did not conduct experiments to show the effects of our visualization, people may have better sense of how their peers are feeling by visualizing weather information of online users as existing works showed the effectiveness of visualization on sentiments in online social media (Cao et al. 2012; Brew et al. 2011).

Method

Twitter data

We crawled data for one month (April 2009) and for the US residents. Our data set comprises 38.1 million tweets and 3 million users for whom we could determine the location to correlate with the meteorological data. We inferred the location information of Twitter users by utilizing the location and the timezone fields from their profile information.

Weather data

The weather data were collected via a crawl bot from Weather Underground, a famous group that provides historical as well as real-time weather readings of various locations (*http://wunderground.com*). We averaged the weather variables from multiple weather stations, located inside each state to obtain the representative weather for the state. The weather variables are:

- Temperature: the mean temperature during daytime (°C)
- Dew point: the temperature below which the water vapor in a volume of humid air at a constant barometric pressure will condense into liquid water (°C)
- Humidity: the amount of water vapor in the air (in the ratio of the absolute humidity to the maximum absolute humidity for that temperature and pressure)
- Precipitation: the amount of rain over one day (mm)
- Atmospheric pressure (hPa)
- Wind speed: the mean of wind velocity over a day (km/h)
- Max gust wind speed: the highest wind speed observed in the last 10 minutes over one day (km/h)

Sentiment Data

A single message posted on Twitter (called a "tweet") is limited to 140 characters in length. The affective (sentimental) expressions made by the posters ("tweeters") were analyzed and quantified using Linguistic Inquiry and Word Count (LIWC), a widely-used tool for sentiment analysis. LIWC measures the sentiments in 32 behavioral and psychological dimensions, among which we focus on the positive and the negative affects for this paper. Figure 1 shows the mean positive affects written in Twitter by the residents of each state in April 2009. The color depth indicates the strength of the mood, showing the variations in the positive affects from state to state. An overall pattern is already visible: the positive affect increases as we traverse from North to South, as we show in the inset of Figure 1 (Pearson correlation *r* between latitude and temperature is -0.40 ± 0.09, with Jackknife estimation). Next, we report in more detail the correlations between weather and mood.

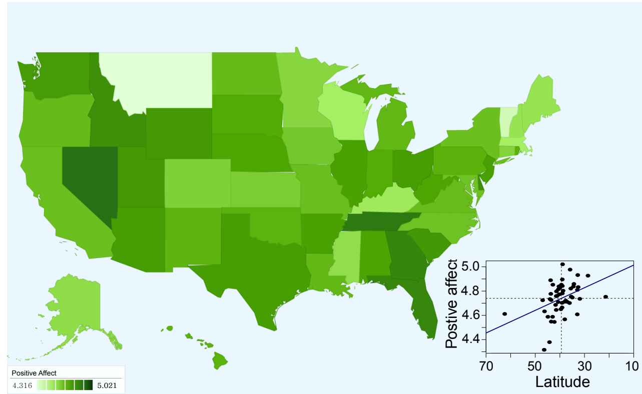


Figure 1: Positive affect across states and latitude

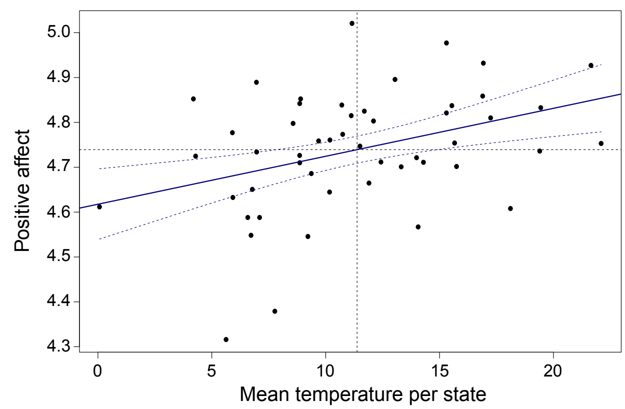


Figure 2: Relationship between the mean temperature and positive affect across 50 states

Analysis Results

We attempted to find weather variables that impact mood through regression analysis. Among the various weather variables, we found three significant factors that impact mood. Except for them, all other weather variables showed negligible correlations with mood ($-0.15 < r < 0.15$). For the three effective factors, we present the correlation using infographic techniques that visualize the relationships in an effective manner (Heer, Bostock, and Ogievetsky 2010).

Impact of Temperature

We begin by analyzing the effects of the average temperature of a day on the mood extracted from that day. The relationship was analyzed using a simple linear regression model:

$$\text{Temp} = \alpha + \beta \times \text{PositiveAffect},$$

where α and β are the regression coefficients.

Figure 2 shows our result. The x-axis in the upper plot represents the mean temperature of each state for the entire month, and the y-axis represents the score of positive affect in the LIWC category for each state. The best linear fit is given as the solid line, while the dotted lines represent the 90% confidence interval. The figure clearly shows a positive relationship between the two variables (r=0.3752±0.1078, with Jackknife estimation), indicating that people living in a state having higher mean temperature do express more positive affects on Twitter than those living in colder states. Nevada, a warm state bordering California, was the happiest state, while Montana showed the least amount of positive affect in tweets.[1]

Furthermore, we analyze the relationship of temperature and positive sentiments across all states. Figure 3(a) represents the Pearson correlation coefficient r of all states. The x-axis indicate the 50 states in US and the y-axis represents the correlation. Although some states show negative or negligible correlations, a sizable number of states show positive correlations between temperature and positive affect (average of r=0.14). Among them, North Dakota shows the highest correlation between temperature and positive affect (r=0.45). On the other hand, Arizona and New Mexico present the most negative correlations. It implies that many people feel more positive emotions when the day temperature is high, but some people feel less positive affects with high temperature day.

On top of simple representation, we visualize the relationship between temperature and positive affects in North Dakota. There are three objects that indicate variables in our analysis: jar, emoticon, and flame. Each jar indicates one day, each smiley indicates the scale of positive affect, and the size of the flame represents temperature of a day.

Impact of Humidity

The relationship between affects and humidity is also interesting. Compared with temperature, humidity had a stronger impact on the negative affect.

Figure 3(b) shows the Pearson correlation r between humidity and negative affect over the course of 30 days in April, across all states. The x-axis represents the 50 states while the y-axis represents the correlation. Although there are state-to-state variations, the two variables had positive relationship for the majority of states (average of r=0.17). The state of Georgia showed the highest correlation (r=0.59), followed by Alabama (r=0.51). In these states, one could see sudden increase in negative affect on damp days (i.e., elevated humidity levels). Some states like Arizona, however, showed a negative correlation in that an increase in humidity led to a decrease in negative sentiments–perhaps humidity was a welcome factor on a desert landscape.

The infographic at the bottom of the figure depicts how residents in Alabama react to increasing level of humidity. The water level indicates the level of humidity and the facial expression indicates the level of negative affect in Twitter users.

Impact of Atmospheric Pressure

Finally, we also found meaningful positive relationship between the atmospheric pressure and positive sentiments.

Figure 3(c) shows the relationship between atmospheric pressure and positive affect across all states. The x-axis represents all the states and the y-axis represents Pearson's correlation coefficients. The majority of states (76%) exhibited positive affect on days with high atmospheric pressure, yet a few states (8%) showed the opposite trend, most notably Hawaii (r=−0.56). Overall, the mean correlation between the two variables was 0.22. Typically a high atmospheric pressure is associated with clear skies and calm weather, so the result is understandable. However, it is interesting to observe that not all states appear to prefer such weather always, leading us to ponder what the notion of "good" weather means for different geographical regions, which would be an issue to explored in future studies.

We also provide an infographic to effectively visualize the result for Idaho, a state that had the highest correlation. High atmospheric pressure is represented as a white cloud and a vivid rainbow, while low atmospheric pressure is represented as a dark cloud. The numbers are the atmospheric pressure. The smiley represents the level of positive sentiments. We believe this type of infographic can effectively communicate the relationship between weather and mood to the public.

Conclusion

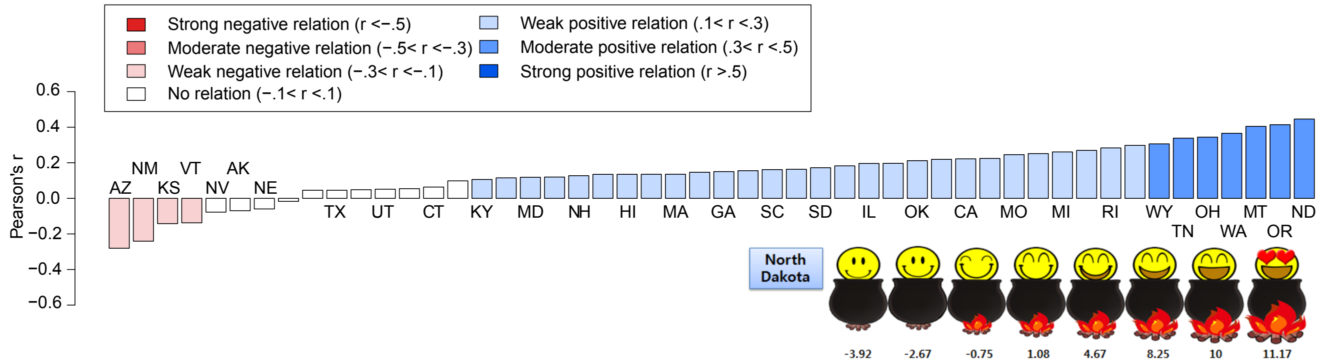
This paper investigated the relationship between various weather variables and the mood of Twitter users. While preliminary, this work is the first to see a direct correlation between the two on a large scale. We found that the temperature and the atmospheric pressure are positively correlated with positive affect, while humidity is positively correlated with negative affect. While these findings largely agree with common experience, we also found interesting exceptions in certain states that suggest further investigation into how meteorological factors may combine with others to influence mood; for instance, our observation that Arizonans preferred humid days and Hawaiians preferred low atmospheric pressure poses interesting questions that merit answers.

We foresee interesting future avenues for taking our research further to enhance our understanding of the factors that influence human mood. While our work clearly demonstrates a correlation between weather and mood, we have conducted only a small portion of possible analyses. First, our observation period was limited to April of 2009, and the possible effect of seasonal changes is not presented in this work. Second, our quantitative analysis was limited to evaluating linear relationships between factors, where in reality more sophisticated multivariate analyses to identify significant factors and remove possible confounding errors via quasi-experimental designs. It would also be interesting to incorporate cultural and economic factors in the analysis by way of examining wide geographical regions that span multiple cultures (e.g., Africa, Asia); we anticipate variations in those factors to show intriguing correlations with mood.

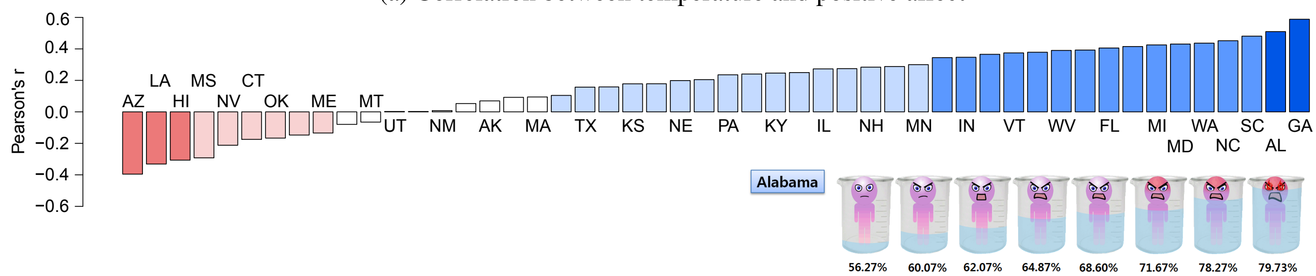
References

Bollen, J.; Mao, H.; and Zeng, X. 2011. Twitter mood predicts the stock market. *Journal of Computational Science* (1):1–8.

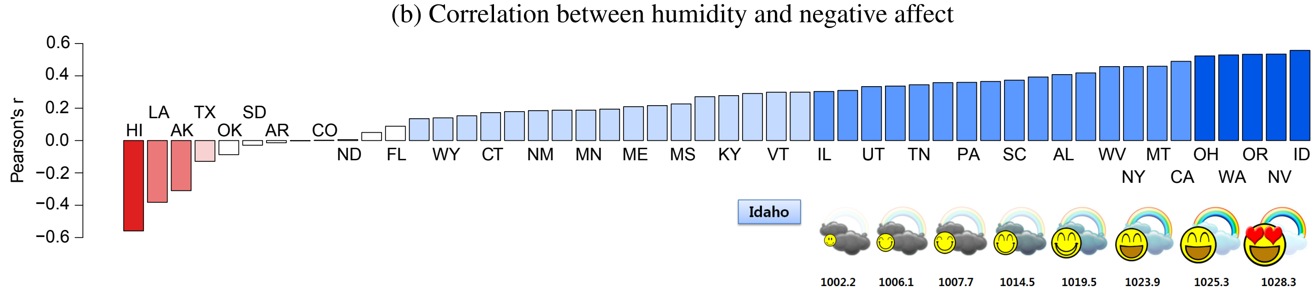
[1]While Nevada was the happiest state, it was at the same time the saddest state; it had high score for negative affects as well.



(a) Correlation between temperature and positive affect



(b) Correlation between humidity and negative affect



(c) Correlation between atmospheric pressure and positive affect

Figure 3: Correlation between weather variables and sentiments across all states in US and their infographic

Brew, A.; Greene, D.; Archambault, D.; and Cunningham, P. 2011. Deriving Insights from National Happiness Indices. In *IEEE International Conference on Data Mining Workshops*.

Cao, N.; Lin, Y.-R.; Sun, X.; Lazer, D.; Liu, S.; and Qu, H. 2012. Whisper: Tracing the Spatiotemporal Process of Information Diffusion in Real Time. *IEEE Transactions on Visualization and Computer Graphics* 18(12):2649–2658.

Golder, S. A., and Macy, M. W. 2011. Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. *Science* (6051):1878–1881.

Hannak, A.; Anderson, E.; Barrett, L. F.; Lehmann, S.; Mislove, A.; and Riedewald, M. 2012. Tweetin' in the Rain: Exploring Societal-scale Effects of Weather on Mood. In *ICWSM*.

Heer, J.; Bostock, M.; and Ogievetsky, V. 2010. A tour through the visualization zoo. *Communications of the ACM*.

Howarth, E., and Hoffman, M. S. 1984. A multidimensional approach to the relationship between mood and weather. *British Journal of Psychology* 75(1):15–23.

Kiciman, E. 2012. OMG, I Have to Tweet That! A Study of Factors that Influence Tweet Rates. In *ICWSM*.

Lazer, D.; Pentland, A.; Adamic, L.; Aral, S.; lászló Barabási, A.; Brewer, D.; Christakis, N.; Contractor, N.; Fowler, J.; Gutmann, M.; Jebara, T.; King, G.; Macy, M.; Roy, D.; and Alstyne, M. V. 2009. Computational Social Science. *Science* 323(6):721–723.

Neugebauer, R. 1999. Mind matters: the importance of mental disorders in public health's 21st century mission. *American Journal of Public Health* (9):1309–1311.

Preti, A. 1998. The influence of climate on suicidal behaviour in Italy. *Psychiatry Research* 78:9–19.

Tumasjan, A.; Sprenger, T. O.; Sandner, P. G.; and Welpe, I. M. 2010. Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. In *ICWSM*.