DOES THE WEATHER INFLUENCE MOOD?

A sentiment-based analysis

IST707 FINAL PROJECT
Ava Nelson
Bhanuteja Sanaka

TABLE OF CONTENTS

INTRODUCTION

Checking the weather forecast is a part of many people's daily routines and can affect many different decisions that they take during the day, from the type of clothing they wear, to the transportation they use, or the food they eat. The popularity of the morning news and the weather apps help demonstrate the significance of the weather events day-to-day. Inclement weather, such as rain, storms, and snow can lead to cancelled plans and delayed events. As a result, people's moods can be well affected by any sudden issues that may arise. For example, sporting event cancellations can cause a loss of money if refunds are not given, in addition to the extra time that it takes to stay updated. People can be saddled with undue stress that would not otherwise be expected. The weather's effect can not only be due to event changes but can also extend to physiological changes.

Seasonal weather changes have been shown to have a well-documented effect on mood and hormones. For example, cold temperatures, in combination with low sunlight, can have a significant effect on thyroid stimulating hormone and serotonin. It is very common for many people in northern states to have a large weight gain during the winter months and also become increasingly more lethargic as the temperature decreases. On the other hand, daily weather changes are more likely to lead to differences in irritability rather than the large hormonal swings normally attributed to seasonality. Since the amount of sun, temperature, and speed of wind can all change within a day or within a week, changes in mood can be daily occurrences. Some people with pre-existing conditions, allergies, or people of old age can be more sensitive to these changes and can further demonstrate the weather's effect on people's mood. Overall, environmental changes, both short term and annual, can lead to behavioral and physical changes in the residents of an area.

People also frequently turn to social media to express their daily victories and frustrations. While many people discuss the events of their day, they could be unaware of how changes in environmental factors affect how they react to any unexpected changes or irritating occurrences. Due to this, people could unknowingly change the tone of their tweets based on differences in weather. In a paper, *Weather Impacts Expressed Sentiment,* written by Baylis et al., the researchers found that exceptionally hot or cold days or days with precipitation have a negative effect on people's sentiment, according to Facebook and Twitter data. Furthermore, the researchers controlled for major events, such as terror attacks and natural disasters to control for extreme sentiments on certain days. As a result, sentiment did not appear to average similarly for each day. Aggregate social media sentiment can be reflective of environmental differences.

ANALYSIS AND MODELS

ABOUT THE DATA- -

The dataset was sourced from kaggle.com (https://www.kaggle.com/grubenm/austin-weather).

This dataset is a collection of weather measurements taken in Austin, Texas, between 2013-2017. The dataset has 1,319 observations consisting of 21 variables. The 21 variables were as follows:

Date: YYYY-MM-DD
TempHighF: Daily High Temperature, in Fahrenheit
TempAvgF: Average Temperature, in Fahrenheit
TempLowF: Daily Low Temperature, in Fahrenheit
DewPointHighF: Daily High Dew Point, in Fahrenheit
DewPointAvgF: Daily Average Dew Point, in Fahrenheit
DewPointLowF: Daily Low Dew Point, in Fahrenheit
HumidityHighPercent: Daily High Humidity Level
HumidityAvgPercent: Average Humidity Level
HumidityLowPercent: Daily Low Humidity Level
SeaLevelPressureHighInches: Maximum Sea Level Pressure
SeaLevelPressureAvgInches: Average Sea Level Pressure
SeaLevelPressureLowInches: Low Sea Level Pressure
VisibilityHighMiles: Maximum Visibility in Miles
VisibilityAvgMiles: Average Visibility in Miles
VisibilityLowMiles: Low Visibility in Miles
WindHighMPH: Daily High Wind Speed, in Miles Per Hour
WindAvgMPH: Average Daily Wind Speed, in Miles Per Hour
WindGustMPH: Daily Wind Gust, in Miles Per Hour
PrecipitationSumInches: Amount of Precipitation, in Inches
Events: Daily Weather Condition.

As the intention of this analysis was to observe how weather influences mood, variables that have a less tangible impact on human observation were removed. The removed variables are those colored in gray above. This reduced the dataframe to 1,319 observations of 7 variables.

After this, the daily Heat Index and Wind Chill was calculated and added to the dataframe. A new variable for season was also added. These were chosen to include as these traits are easily observed by humans. The resulting dataset contained the following variables:

Date: YYYY-MM-DD
TempHighF: Daily High Temperature, in Fahrenheit
TempAvgF: Average Temperature, in Fahrenheit
TempLowF: Daily Low Temperature, in Fahrenheit
WindGustMPH: Daily Wind Gust, in Miles Per Hour
HeatIndexAvg: Daily Average Heat Index, in Fahrenheit

HeatIndexHigh: Daily High Heat Index, in Fahrenheit
WindChillAvgF: Daily Wind Chill, in Fahrenheit
WindChillGustF: Daily Wind Chill Gust, in Fahrenheit
PrecipitationSumInches: Amount of Precipitation, in Inches
Events: Daily Weather Condition.
Season: Calendar season observation was made in.

51 observations were selected to use for the sentiment analysis. These observations were selected from the initial 1,319 by choosing the observations that display desired traits. The desired traits included a collection of the hottest days observed (as measured by heat index), the coldest days observed (as measured by wind chill), the stormiest days observed (as measured by precipitation sums), and the outliers as determined by statistical analysis.

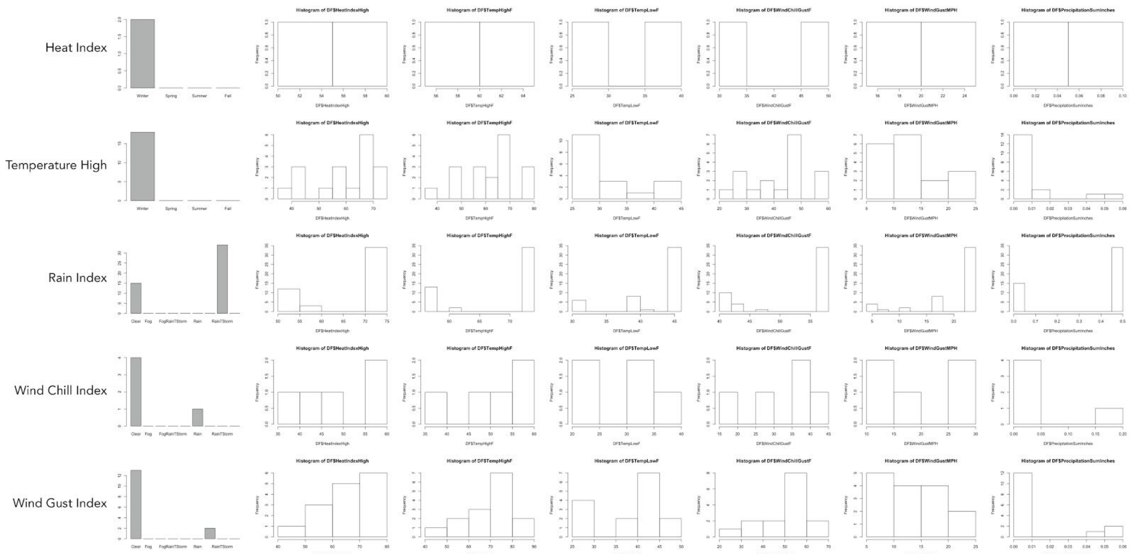
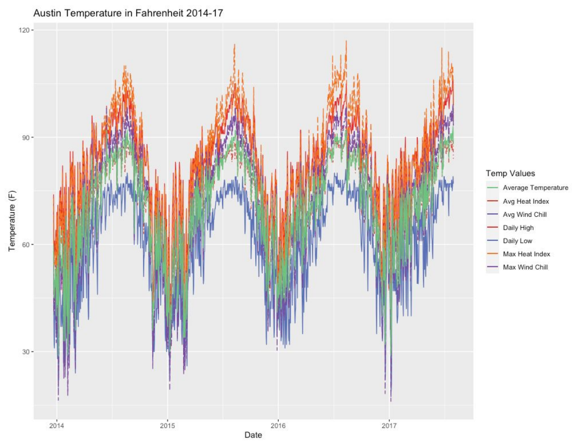


Fig 1.1 Histograms exploring Outliers

In general, the outliers consisted of days occurring during the winter months, often with relatively warm temperatures and clear skies. This is reinforced by knowledge of Austin’s climate and known weather patterns. The outliers also picked up on a rare snow day, and particularly vicious storms. The outliers did not indicate that extremely warm days were out of the ordinary-- again, this is reinforced by existing knowledge of Austin’s climate.

After the final variables were selected, exploratory data analysis was performed. Exploration was conducted for both the original 1,319 observations and the 51 observations selected for sentiment analysis.



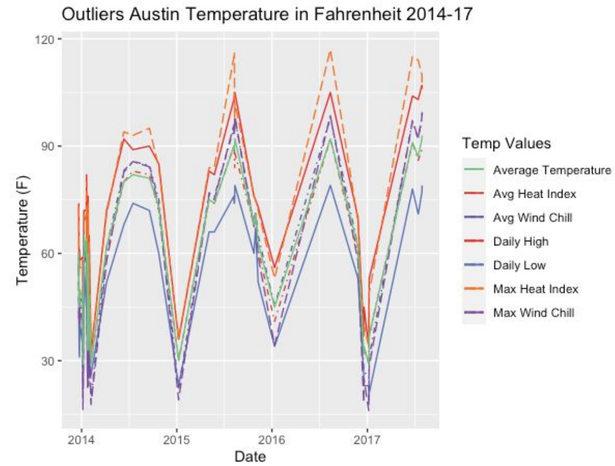
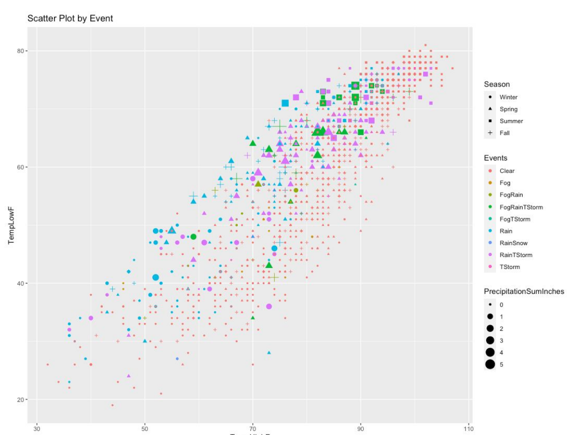


Fig 1.2 (left) Austin Temperature in Fahrenheit Fig 1.3(right) Outliers Austin Temperature in Fahrenheit

First, a time series tracking observed temperature variables was made to visualize the climate traits of the region. Of notice in the rightmost graph is its more jagged track, indicative of the greater level of extremes observed in the data.



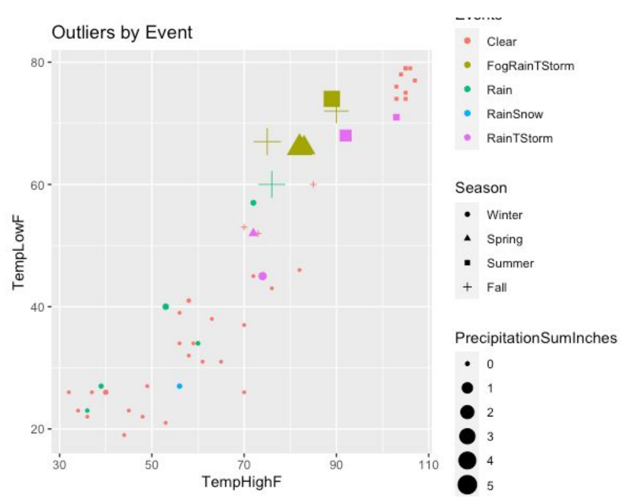


Fig 1.4 (left) Austin's Daily Weather by Event Fig1.5 (right) Outliers Weather by Event

Then, scatterplots were produced to visualize the relationship the data's observations had with selected variables. The rightmost plot showcases the distilled variety of conditions within the selected dataframe, which appears promising for the later sentiment analysis. As the intention is the analysis' ability to measure sentiment, carefully selecting the most promising days for sentiment analysis was a worthwhile time investment.

TWITTER DATA

Using the dates from the distilled dataframe, random samples were collected from Twitter and stored in individual dataframes. Each twitter dataset consisted of 100 random tweets from Austin, Texas on the date selected. As the standard twitter API only allows for tweets from the past 7 days, a special API for "Full Archive Search" was acquired, which allows for data collection from the entire history of Twitter.

The text from the tweets was then cleansed of emoji, media, images, and other items incongruous to sentiment analysis. Once the data was cleansed, word clouds of the most negative, neutral, and positive days were generated to visualize the most common sentiment altering language observed.

Before Curse Word Removal

Fig 1.6 most negative day day (2013-12-23) Fig1.7 most neutral day (2014-1-22) Fig1.8 most positive day (2015-1-9)

With the later sentiment analysis, it was noticed that the presence of certain curse words highly influenced the mood of a given tweet. A closer look at individual tweets saw many instances where the use of curse words were for a neutral-positive emphasis, rather than the negative sentiment as given by the lexicon. Due to this discrepancy, the twitter data was also run without the curse words with the hopes of improved models.

After Curse Word Removal

Fig1.9 most negative day day (2013-12-23) Fig1.10 most neutral day (2017-1-4) Fig1.11 most positive day (2015-1-9)

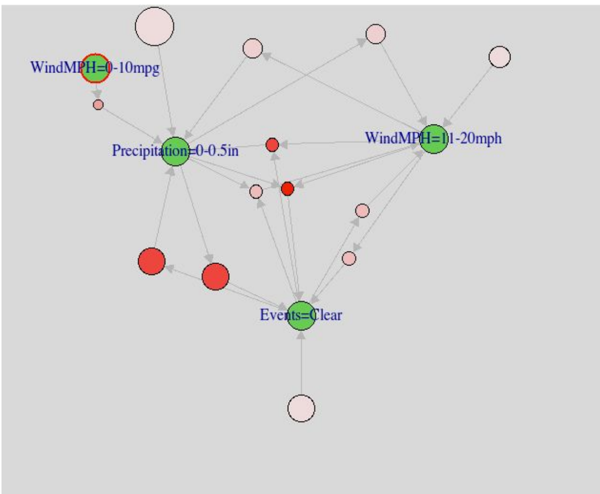
Once generated, sentiment scores were returned to the curated data frame for further analysis.

MODELS- -

ARM, Decision Trees, Naive Bayes, and Sentiment Analysis are used for this analysis.

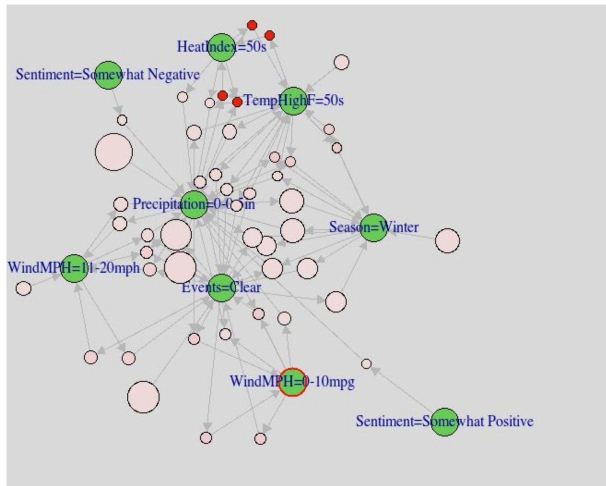
ASSOCIATION RULE MINING (ARM)

Association Rule Mining were run and visualized as follows:




```
    lhs                        rhs                      support   confidence lift     count
[1] {}                      => {WindMPH=21-30mph}       0.4473086 0.4473086  1.000000  590
[2] {}                      => {WindMPH=11-20mph}       0.4761183 0.4761183  1.000000  628
[3] {}                      => {Precipitation=0-0.5in}  0.9332828 0.9332828  1.000000 1231
[4] {WindMPH=21-30mph}      => {Precipitation=0-0.5in}  0.4124337 0.9220339  0.987947  544
[5] {Precipitation=0-0.5in} => {WindMPH=21-30mph}       0.4124337 0.4419171  0.987947  544
[6] {WindMPH=11-20mph}      => {Precipitation=0-0.5in}  0.4617134 0.9697452  1.039069  609
[7] {Precipitation=0-0.5in} => {WindMPH=11-20mph}       0.4617134 0.4947197  1.039069  609
```

Fig2.1 ARM(default)




```
> inspect(sort(rules5b, by="lift", decreasing=TRUE))
    lhs                              rhs                     support   confidence lift count
[1] {}                            => {WindMPH=11-20mph}      0.4313725 0.4313725  1    22
[2] {}                            => {TempHighF=50s}         0.4313725 0.4313725  1    22
[3] {}                            => {Season=Winter}         0.6862745 0.6862745  1    35
[4] {}                            => {Events=Clear}          0.8431373 0.8431373  1    43
[5] {}                            => {Precipitation=0-0.5in} 1.0000000 1.0000000  1    51
[6] {Sentiment=Somewhat Positive} => {Precipitation=0-0.5in} 0.3333333 1.0000000  1    17
>
```

Fig2.2 ARM(with Sentiment)

Association Rules was utilized to aid in determining which days would be isolated for the following sentiment analysis. The days were organized by lift to ensure greatest interest for the analysis. The rules were generated using a confidence of 0.4 and support of 0.3.

The same parameters were also run as ARM with the subsequent gathered sentiment, as shown in Figure 2.2. It did not appear to reveal anything substantial about how sentiment might relate to other variables, so it was disregarded in favor of better models.

DECISION TREE

Decision Tree Analysis was run and visualized as follows:

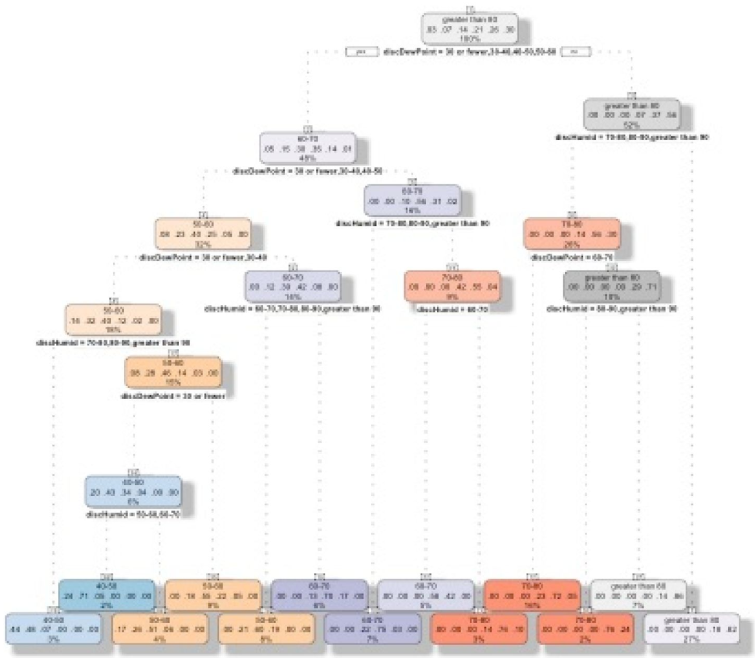


Fig 2.3 Temperature Predictor Tree

The initial decision tree was used to capture days of interest using the branch node extending to the most specific classification. The default parameters resulted in the most accurate model.

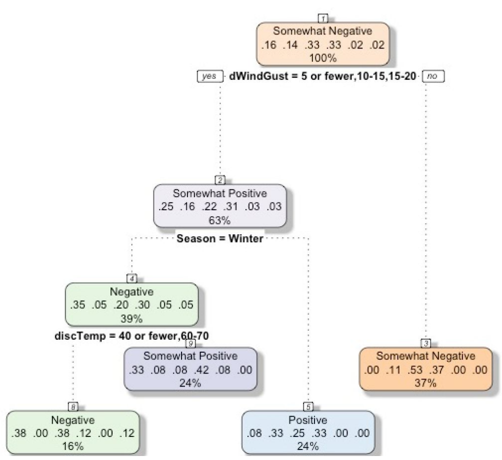


Fig2.4 Sentiment Tree (default)

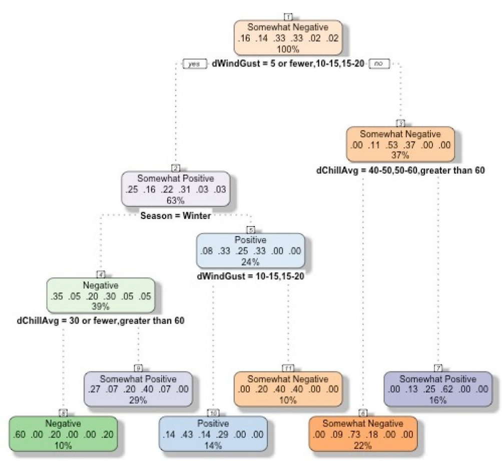


Fig2.5 Sentiment Tree(n=4)

NAIVE BAYES

Naive Bayes was run and visualized as follows:

Before Curse Word Removal

```
> confusionMatrix(NBayePredict2, TestData2$Ser
Confusion Matrix and Statistics

          Reference
Prediction Negative Positive
  Negative        0        3
  Positive        3        7

               Accuracy : 0.5385
                 95% CI : (0.2513, 0.8078)
    No Information Rate : 0.7692
    P-Value [Acc > NIR] : 0.9843

                  Kappa : -0.3

 Mcnemar's Test P-Value : 1.0000
```

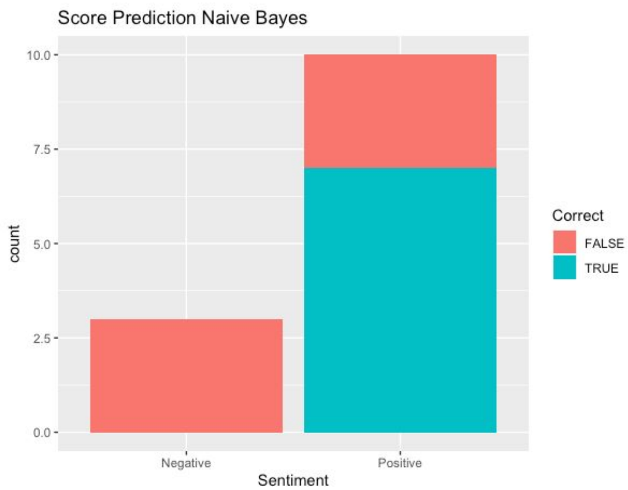


Fig2.6 Naive Bayes Prediction A (Matrix) Fig2.7 Naive Bayes Prediction A (Plot)

Naive Bayes model was created using laplace of 1, for accuracy of 53.8%.

After Curse Word Removal

```
Confusion Matrix and Statistics

          Reference
Prediction Negative Positive
  Negative        0        2
  Positive        3        8

               Accuracy : 0.6154
                 95% CI : (0.3158, 0.8614)
    No Information Rate : 0.7692
    P-Value [Acc > NIR] : 0.943

                  Kappa : -0.2264

 Mcnemar's Test P-Value : 1.000
```

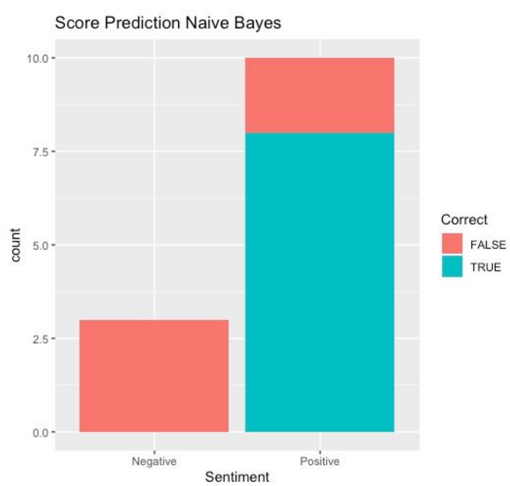


Fig2.8 Naive Bayes Prediction B (Matrix) Fig2.9 Naive Bayes Prediction B (Plot)

Naive Bayes model was created using laplace of 1, for accuracy of 61.5%.

SENTIMENT ANALYSIS

Sentiment Analysis was run and visualized as follows:

With Curse Words

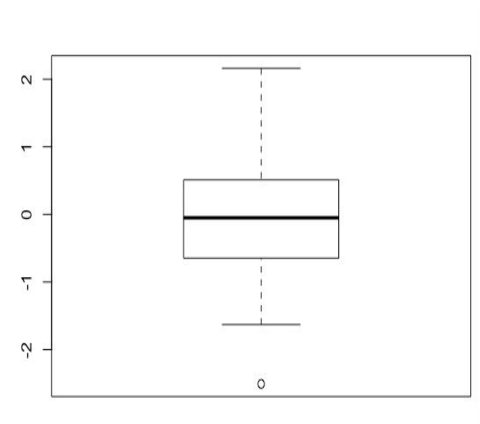


Fig2.10 Sentiment Analysis Boxplot (original)

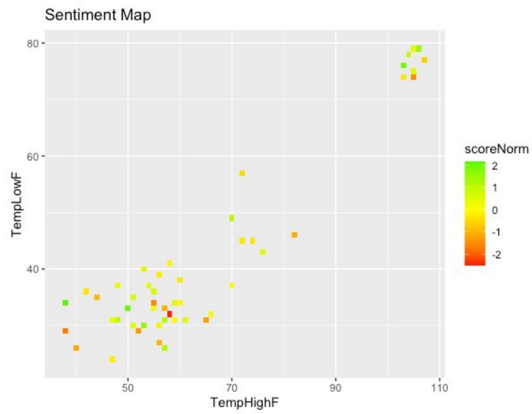


Fig2.11 Sentiment Analysis Map (original)

Without Curse Words

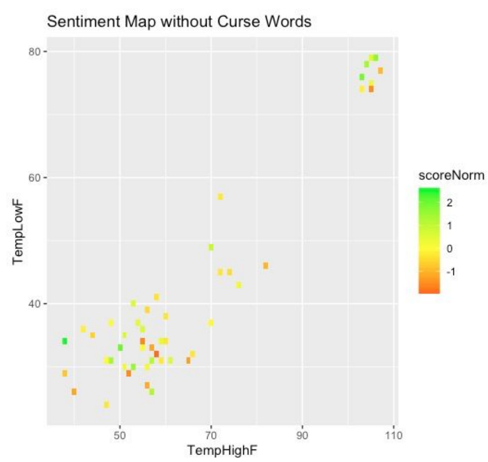


Fig2.12 Sentiment Analysis Boxplot (no cursing)

Fig2.13 Sentiment Analysis Map (no cursing)

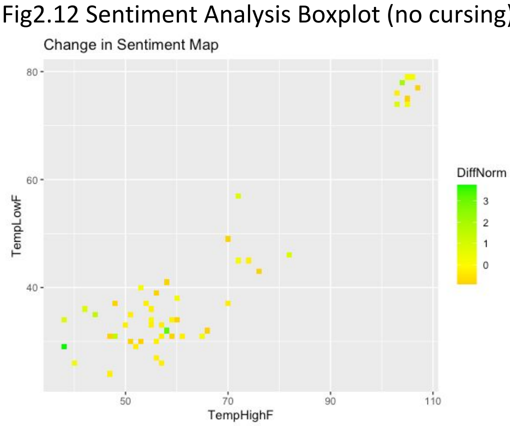


Fig2.14 Sentiment Analysis Map (Sentiment Shift)

RESULTS

Prior to modeling the mood of the Twitter data collected, each day needed an overall score to relate weather metrics to the tweet, which was created through a sentiment analysis. The lexicon used to complete the sentiment analysis scored any curse words as the most negative score of -5. When tweeting, people frequently curse without necessarily expressing any negative emotion, which can result in confounded scores. For example, curse words can be often used as modifiers to the following word.

Due to this added noise, sentiment analysis was completed twice to understand the differences in daily sentiment. The curse words present in the Twitter data resulted in a negative shift in the overall sentiment of the tweets than with the curse words removed. The removal of curse words resulted in the most negative day increasing it's score by the largest value. This trend continued for many of the most negative days. On the other hand, days previously evaluated as neutral or positive did not change the normalized values significantly.

With the curse words present in the tweets, the most negative day of the dataset was an outlier with a scaled z-score of -2.51. This day captured an argument between two Twitter users, which included many of the curse words. After curse word removal, the most positive day of the dataset became an outlier with a z-score of 2.51. Graphically, the coldest days appear to have the largest variance in sentiment. In addition, the coldest days have lower sentiments on average. Constrastingly, the hottest days had a more positive sentiment on average.

Limitations of the sentiment analysis outside the use of curse words include the use of sarcasm and irony in tweets. Tokenizing each word removes much of the context behind the relationship of each of the words. As a result, text mining can lead to some sentiment inaccuracies that would otherwise be captured by a human.

After the completion of the sentiment analysis, the association rules mining and decision tree were executed again with the addition of the normalized and discretized sentiment variable. As a result, the decision trees predicting the sentiment of the day support the conclusion that the weather can be indicative of the mood of tweets on the day. The association rules did not result in any associations between the sentiment and weather metrics. The model likely couldn't determine any strong rules due to the small sample size.

The decision tree model found that people, on average, tweeted with more positive sentiments on warmer days, with a higher likelihood of negative sentiment on colder days. A surprising finding from the decision tree was the relationship between wind gust speed and sentiment. Conceptually, a negative relationship between larger wind gusts and sentiment was assumed, but the decision tree showed that higher wind gusts led to a more positive sentiment. It is possible that since higher wind gusts are frequently correlated to storms, Twitter users could be more positive in the hopes of a storm in a drier climate like Texas. Historically, from 2013 to 2017, Texas was afflicted by a serious drought, which frequently affected over 50% of the land, which could explain the initially unintuitive results.

The Naive Bayes model with curse words included resulted in an accuracy rating of 53.9% while the model after the curse word removal had a much higher accuracy rating of 61.5%. This is likely due to the amount of noise that curse words add to the data. Since they can be used both emphatically and negatively, it is difficult for the lexicon to provide an accurate score. In both Naive Bayes models, the prediction resulted in predicting the negative days incorrectly. In contrast, the model, with curse words still included, predicted the positive days correct 70% of the time. The model, with curse words removed, predicted the positive days correctly 72% of the occurrences. This pattern could have occurred due to the different context that negative tweets contain, resulting in less accurate sentiment and models. Each model had a low number of negative days included in the test data set despite the presence of a higher number of negative days overall. With a larger sample size, the models could have predicted the negative days correctly more often.

CONCLUSIONS

This analysis concludes the presence of an association between the perceived mood of people and a given day's weather. While further analysis is required to determine causality from the correlation, this initial foray into mood prediction from weather metrics is promising.

The greatest impact on mood was seen in the more extreme weather metrics-- the hottest days, coldest days, and windiest days. This is reinforced by outside knowledge of the climate in the region, as these conditions appear to result in greater variety of emotion from casual observation. The presence of this wide range of mood in the polar ends of the year could also be correlated to the presence of certain calendar holidays, or other non-weather related events occurring that have direct impact on mood.

The presence of expletives in collected tweets skewed the observed negative sentiment to being more negative, which in turn impacted the predictive power of the chosen model. As expletives are extremely dependent on the context in which they are used, it is challenging for the model to quantify when an expletive is used in a positive setting (as emphasis, for example), or in a negative setting. Once expletives were removed from the lexicon, clarity of the model was improved and the predictive power increased. As the samples used for the model were small in number, it is concluded that the model is good.

Room for improvement can be given to a desire for a larger sample size from Twitter. This can be obtained by purchasing a premium API for the Full Archive search features, and could lend itself to more robust modeling. Additionally, further research into the nuance of context words are used would assist any future models in interpreting the true sentiment of a given tweet.

Data sourced from other social media platforms would also be useful for further study, as Twitter's demographic is much younger than what is seen on sites like Facebook. Diversifying the demographics to perform sentiment on may improve a model's ability to interpret sentiment as well.

Lastly, sourcing methods to return sentiment from image based posts (e.g. memes, emoji use) would also provide needed context in returning a more robust model.

In conclusion, as the model presented in this analysis returned promising results, it appears likely that an improved predictive model would be able to predict mood with a higher degree of accuracy. Writing more context into the model will help it more nimbly navigate the dynamic linguistic trends seen on social media. In the meantime, however, current results support the analysis query.