Group 1: Final Project

IST-718 Su 7:30 p

Alfred Tinsley, Eugene Johnson,

John O’Rourke & Sophea Hummel

Due: 9/24/2022

Abstract

Throughout our analysis, it was determined that states in the southwestern corner of the United States will be most heavily affected by droughts. To combat this as well as prevent these levels of droughts occurring across the United States, we are recommending that the U.S. Governments enact a slow rollout of water preservation tools and methods to shift these states into a more water-conscious mentality.

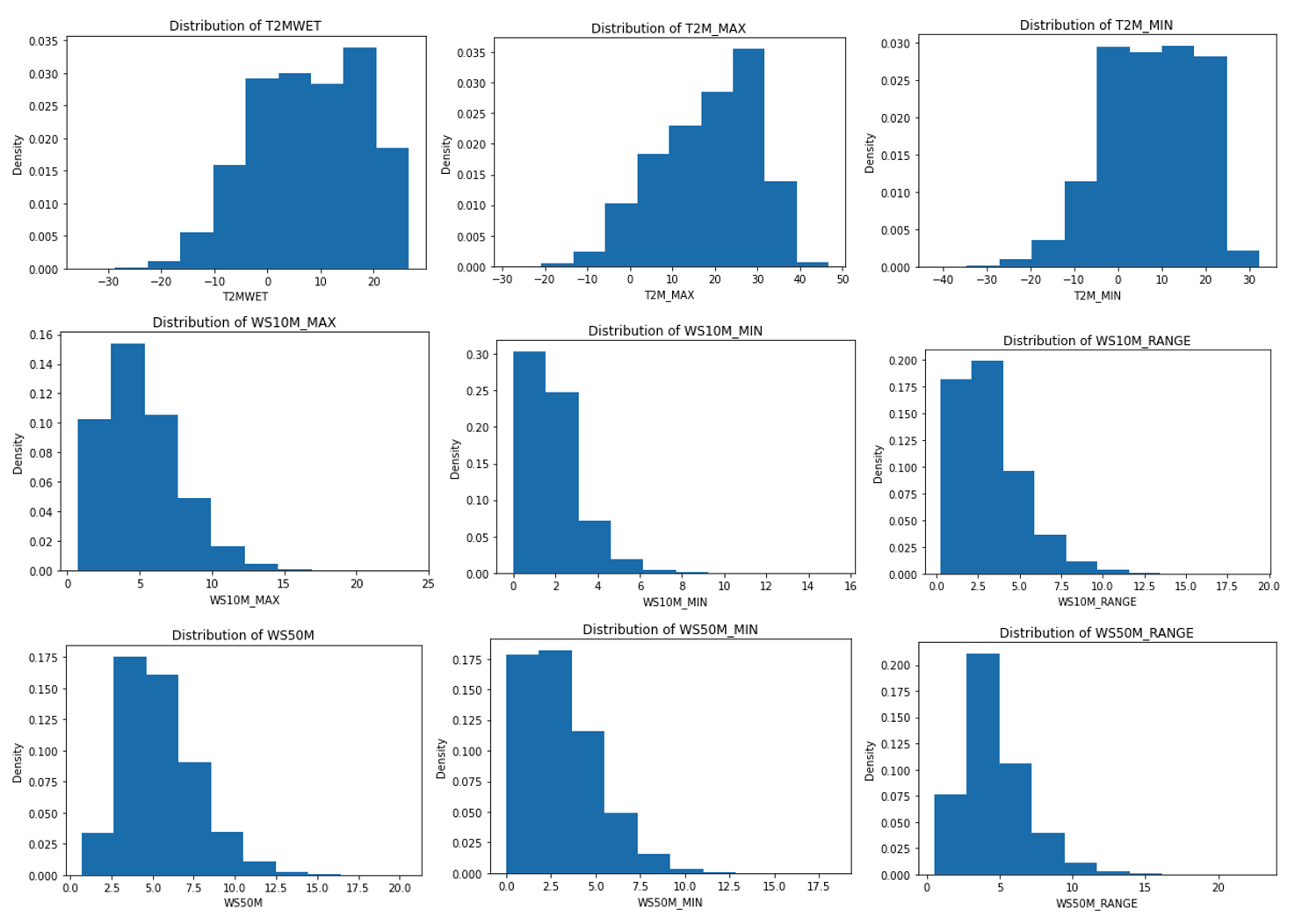
Specification:

Prediction is critical for warning governments and farmer’s to about early droughts. The main problem this project attempts to solve is early drought detection to combat droughts across the United States. The goal of this project is to effectively and efficiently predict where a drought may occur and just how severe it is by developing a forecasting model. This will allow for the government to set up programs that help combat areas affected by severe droughts based on weather and soil conditions observed. We will be using data that has been collected from *Predict droughts using weather & soil data* (Minixhofer, 2021). Below reviews the distribution of scores and utilized dates from the dataset provided by Minixhofer.

|  |  |
| --- | --- |

Observation:

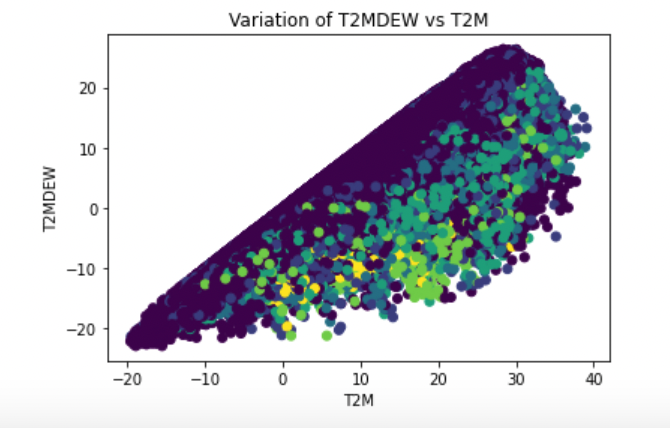
Within our analysis of the time series data, we saw that the numeric attributes of the dataset are filled with wide distributions of values with a significant amount of data classified as outliers. Below are samples of the histograms and boxplots generated during our exploration of the data. Between the below various plots, one conclusion drawn was that there was a vast degree of records provided by the data that made it difficult to produce quality models. Another observation was that the attributes had a strong correlation to predicting the score.



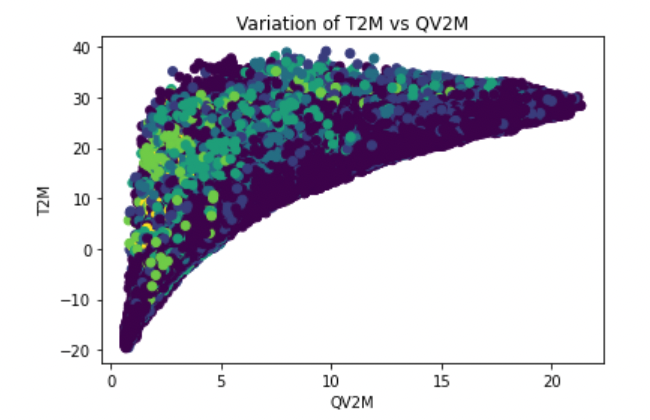
The categorical attributes of the data are evenly distributed seeing as these attributes are for classifying the year, month, and day the data was recorded on. Below show the bar charts for our distribution of attributes for our dataset showing the outliers within the data.

|  |  |
| --- | --- |
|  |  |

Below demonstrates the relationship between the dew point and temperature at two meters. The color of the dots represents the drought score of that correlational relationship.

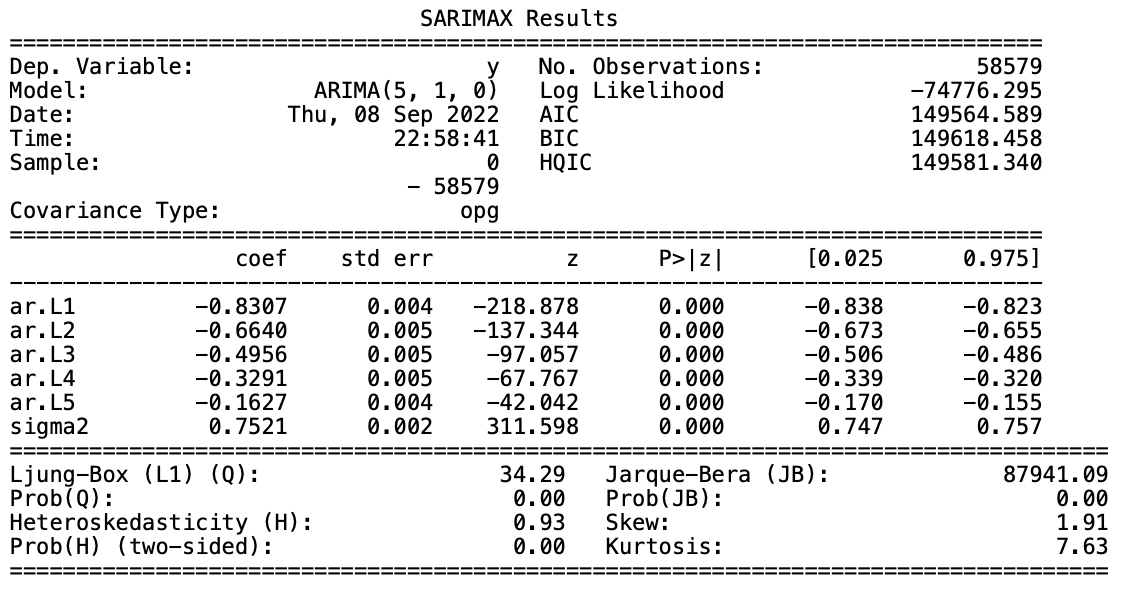


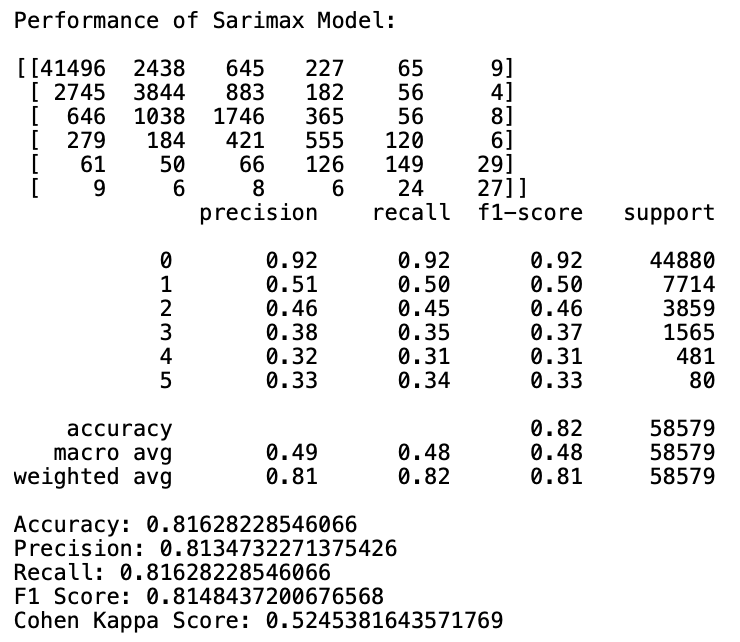
Below demonstrates the relationship between the temperature and humidity at two meters. The color of the dots represents the drought score of that correlational relationship.

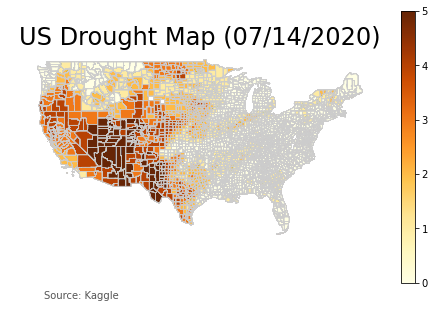


Analysis:

Utilizing random forest and decision tree modeling techniques we effectively utilized our historical data to build a statistical model for predictions through regression and classification. By using these modeling techniques it enhanced our understanding of the data as well as predict possible values. The forecasting models and geographical mapping of where droughts will occur have been completed. SARIMAX predictive model was chosen to output our prediction for geo mapping. These modeling techniques and visualization gave us an accurate representation of the location for potential droughts as well as its severity.







Recommendation:

The continental United States has, like the rest of the world, gone through a dramatic climate shift in recent years. Unfortunately, combatting droughts is not a simple undertaking and requires time, funding and manpower to effectively reverse. With the above analysis, we were able to develop a proper plan of action that the U.S. Government could effectively battle climate change.

Our predictive analysis shows that the states that will be most-heavily impacted fall into the south western states as well as the mid-south states. These states were typically on the dryer side of the scale when compared against the eastern half of the United States but in recent years, these states gradually fell into the extreme and exceptional drought range. The midwest and eastern half of the United States are not currently experiencing anything higher than an abnormally dry or moderate drought rating. This is most likely due to the harsh winters the midwest and north east have as well as the more tropical climate that the south east has nearly year round.

With this information presented, it is recommended that the United States Government focus their efforts on the states affected by the harshest droughts being those states in the south west of the country. By focusing in these areas, through the usage of drought combat measures for both industrial and residential use-cases, droughts can be prevented. These methods are water-conscious methods such porous asphalt/concrete that allows for rain water to seep through the material and into the soil rather than pooling in a street, usage of rain barrels, specially developed soil that retains water longer, and even drought-preventative fixtures within homes and businesses to avoid wasteful usage. Once these efforts have been completed in the states most affected by current metrics, this plan can slowly shift and cover states to the north of California and east of Texas until eventually these measures are in place at every given county in the country.

This rollout of services will take time but with the above measures taken, we can help combat current droughts as well as prevent severe droughts from occurring.

References

Minixhofer, C. (2021, March 3). *Predict droughts using weather & soil data*. Kaggle. Retrieved September 21, 2022, from https://www.kaggle.com/datasets