

# AtakpamePDF

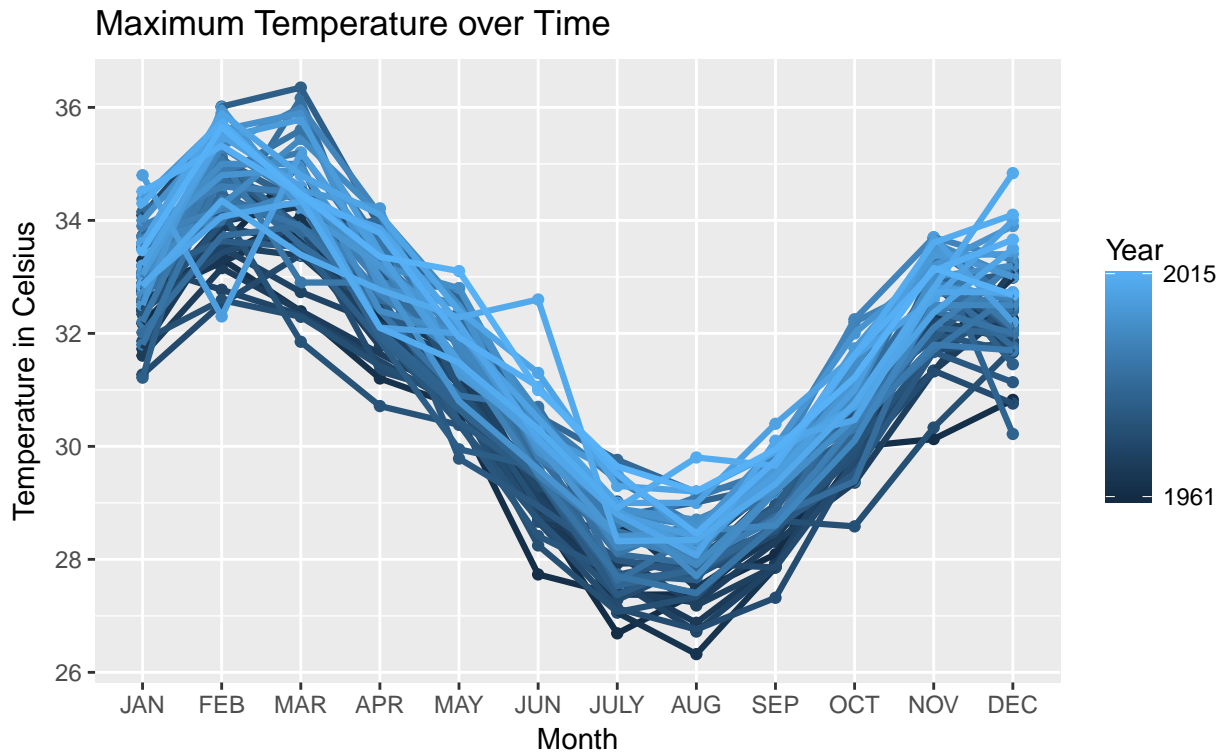
Allison Zembrodt and Lizzy Compton

3/2/2018

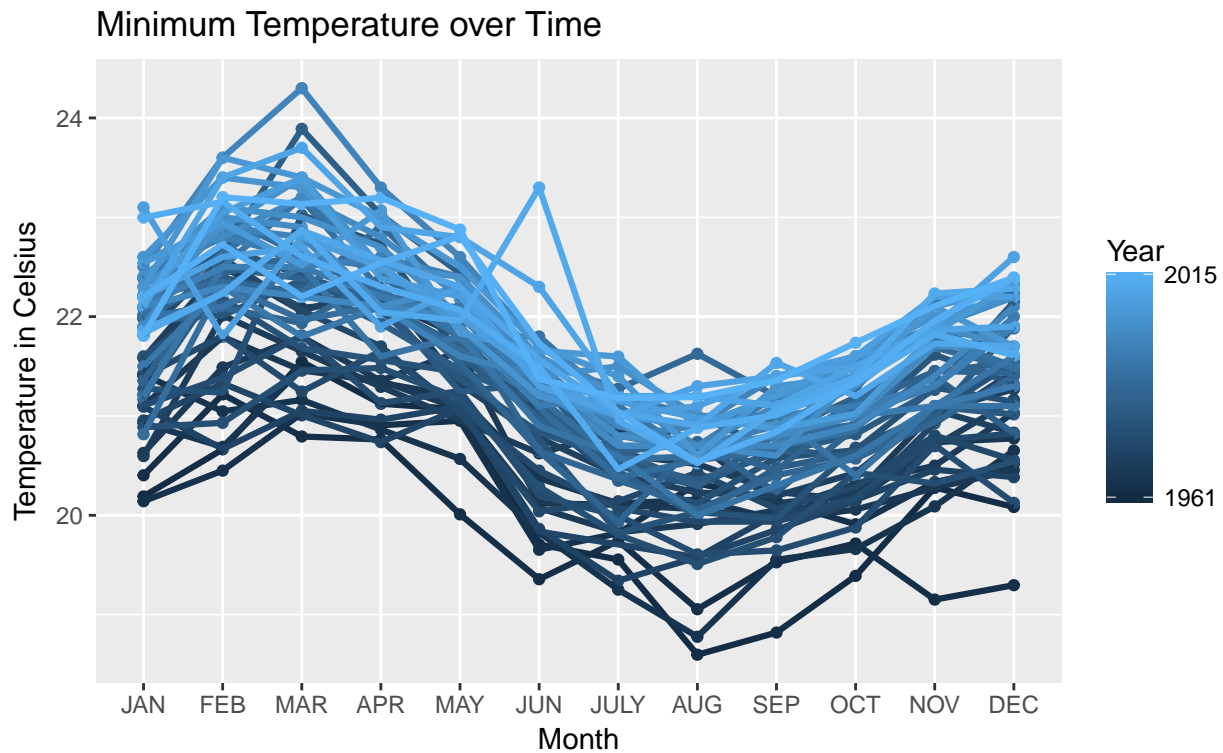
## Previous Report

In the previous report, the group concluded that the best model was the quadratic trend model with a categorical variable for the average minimum temperatures and the linear trend model with a categorical variable for the average maximum temperatures. Overall the average temperatures are rising with the average minimum temperatures rising faster than the average maximum temperatures. We expanded on this report by creating graphs of the temperatures vs. time (month) for each year but having the year lines be on a gradient to show the effect of the temperatures rising over the years. In both the maximum and the minimum temperature graph you can clearly see that the more recent years are higher on the graph indicating rising temperatures. Additionally, the yearly oscillation of the temperatures can be seen. Within the oscillation temperatures are highest in March and lowest in August. In terms of our report, we removed the same outliers that the previous report found.

## Maximum Temperature Year Gradient



## Minimum Temperature Year Gradient



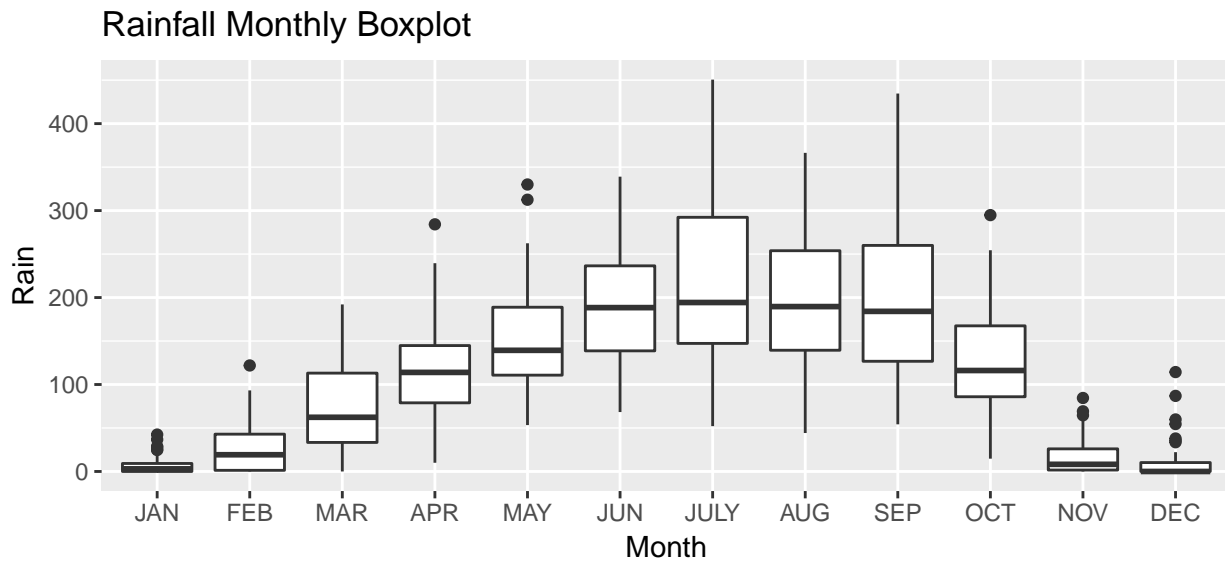
## Background Information

As stated in the previous report, Togo is a coastal, West-African nation slivered between Benin to the east, Ghana to the west, Burkina Faso to the north, and the Gulf of Guinea to the south. Togo gained its independence in 1960 but much unrest followed. In Atakpame specifically, agriculture is prominent and the city is a part of a major trade route through the country. Atakpame, Togo is located at  $7^{\circ}35'N$ ,  $1^{\circ}7'E$ , 402 m (1319 ft). It has a tropical wet and dry/ savanna climate with a pronounced dry season in the low-sun months, no cold season, and wet season is in the high-sun months. According to the Holdridge life zones system of bioclimatic classification Atakpame is situated in or near the tropical dry forest biome (<http://www.atakpame.climatemps.com/>)

## Rainfall

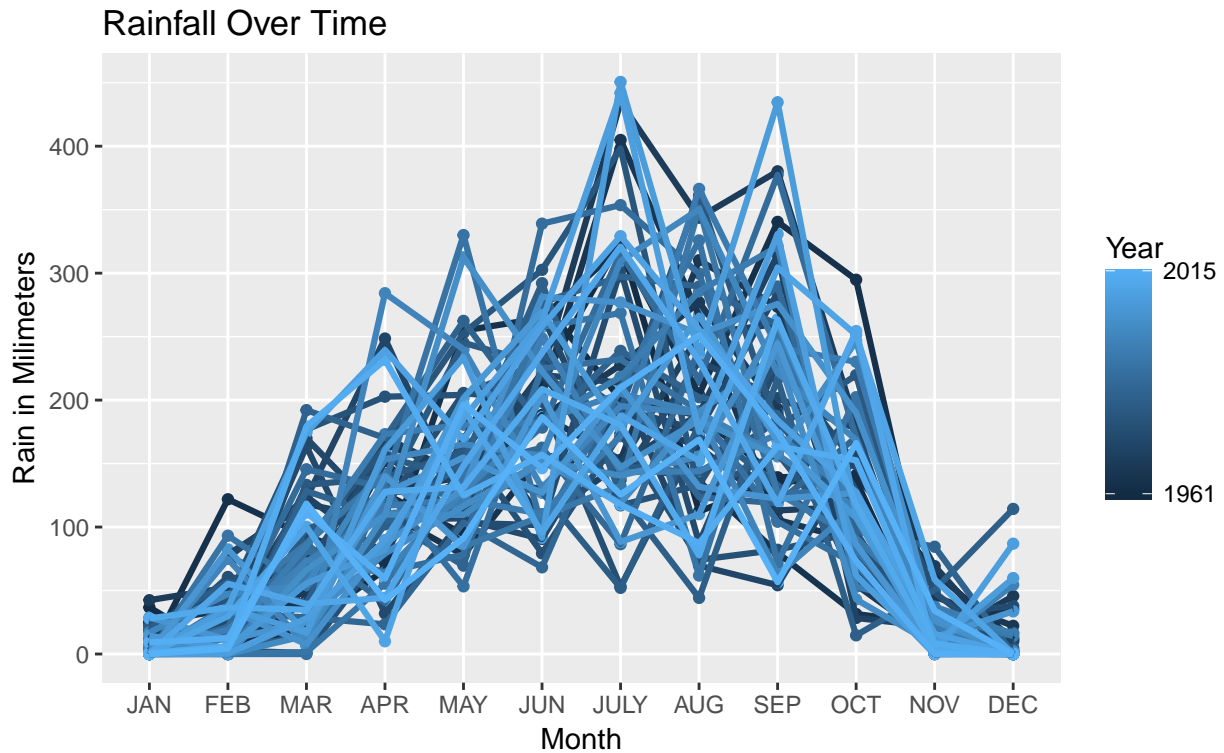
Before we did any modeling on the rain data. We first checked to see if there were any outliers and removed them. The process to remove them included grouping the values by month and removing any that were  $\pm 4$  standard deviations away from the mean. Here is a boxplot to visualize the process. Our outliers were in January 2009 and December 1990.

### Outlier Boxplot



Just like the average minimum and average maximum temperatures, the monthly rainfall in Atakpame also follows a yearly oscillation pattern. However this pattern shows maximum rainfall in the summer months of July/August and minimum rainfall in the winter months of December/January. Additionally, there does not appear to be a significant change in rainfall over the years.

### Rainfall Year Gradient



## Rainfall Categorical Model

Next, we wanted to model the Rainfall data using each month, the minimum and maximum temperatures, and the year as our predictors. In doing this, we created a rainfall model with each of the months as categorical variables.

Table 1: Predicting Rain using Months and Temperatures as Categorical Variables

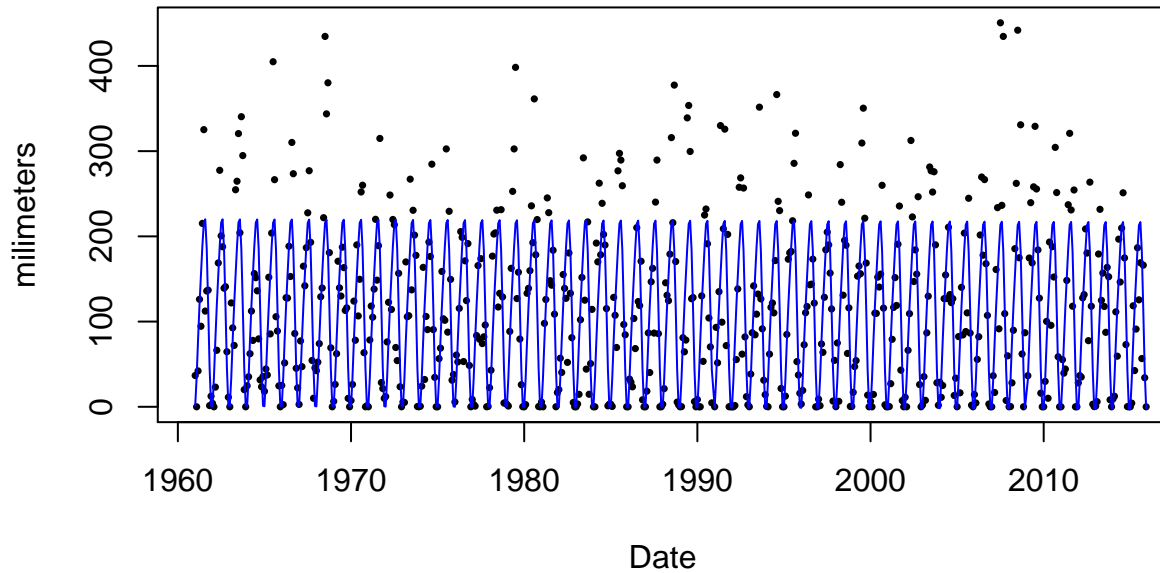
	<i>Dependent variable:</i>
	Rainfall
February	48.455*** (12.857)
March	87.114*** (12.606)
April	97.872*** (12.159)
May	111.362*** (13.228)
June	116.711*** (16.613)
July	113.270*** (20.820)
August	93.084*** (21.367)
September	113.732*** (18.481)
October	74.245*** (14.452)
November	4.133 (11.776)
December	-5.646 (11.849)
Minimum Temperature	12.033* (6.796)
Maximum Temperature	-23.170*** (4.232)
Year	0.277 (0.240)
Constant	-39.048 (387.573)
Observations	645
R <sup>2</sup>	0.643
Adjusted R <sup>2</sup>	0.635
Residual Std. Error	59.868 (df = 630)
F Statistic	81.054*** (df = 14; 630)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

This model uses January as a baseline which is nice, because the rainfall for January is generally very low. The months February, March, April, May, June, July, August, September, and October are significantly larger than the baseline month. Using this categorical model our adjusted R-squared is .629.

## Rainfall Sines and Cosines

We also wanted to try adding sines and cosines into our model to account for the periodicity. We went up to adding 13 terms in the model, but found that, while some were significant, what they added to the model was minimal. So we ended up going with the simpler model with only 5 terms.

## Rainfall Model with Sines and Cosines



The model appears to not fit the entirety of the data. It seems like the zeroes are pulling the data down.

## Residuals of Rainfall Model

While, we commented that we did not feel that the model fit the data, the residuals seem to be fairly evenly distributed. There is no pattern amongst them.

## Rainfall Model with Sines and Cosines Residuals

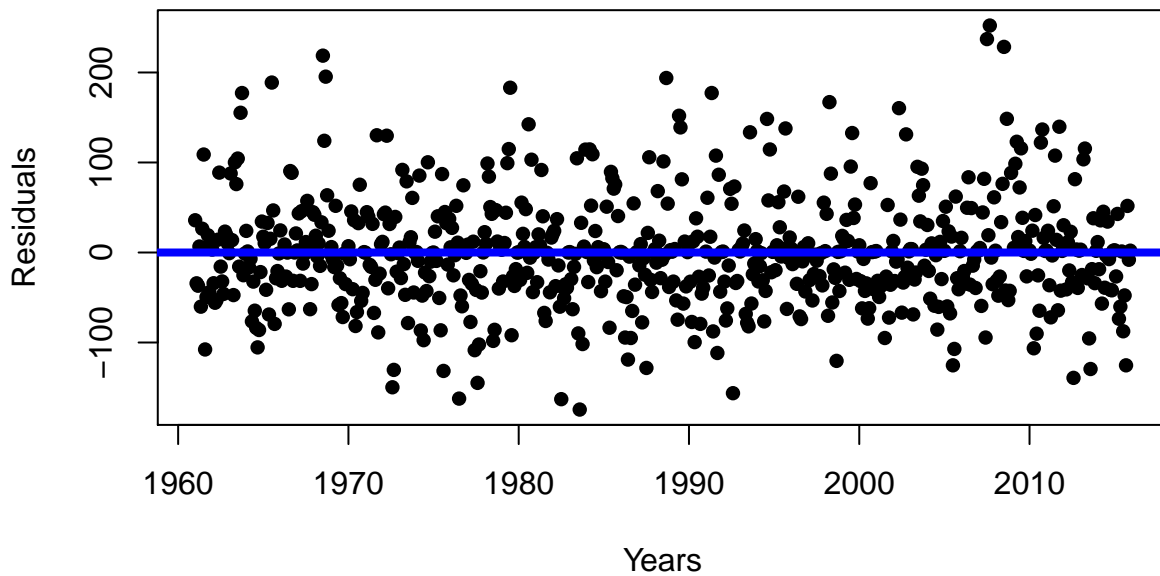


Table 2: Predicting Rain using Sines and Cosines

	<i>Dependent variable:</i>
	Rainfall
Sine 2pi * Long Date	0.814 (3.401)
Cosine 2pi * Long Date	-107.752*** (3.409)
Sine 4pi * Long Date	18.504*** (3.404)
Cosine 4pi * Long Date	-5.003 (3.406)
Year	-0.062 (0.152)
Constant	234.244 (301.515)
Observations	658
R <sup>2</sup>	0.612
Adjusted R <sup>2</sup>	0.609
Residual Std. Error	61.759 (df = 652)
F Statistic	205.947*** (df = 5; 652)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

## Predicting Rain with Temperature

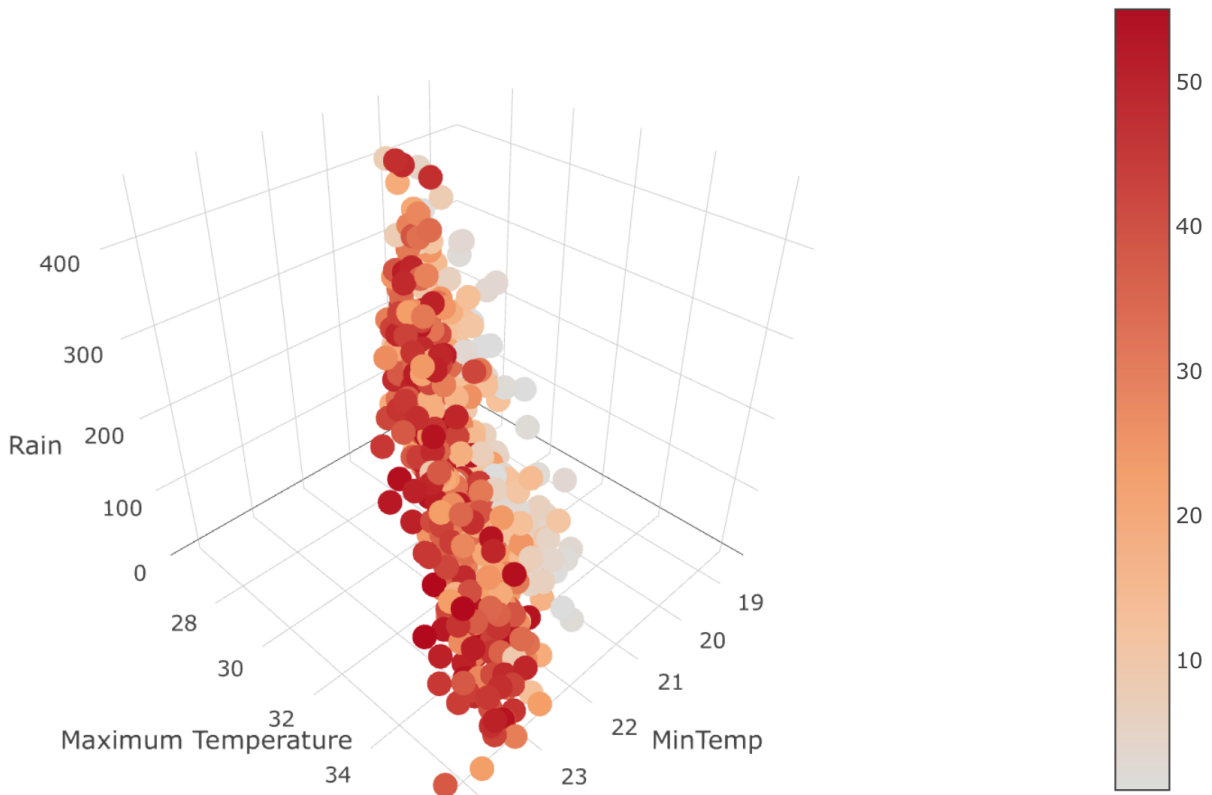
We wanted to see if temperature had an effect on rainfall. To do this, we expanded on the previous model, by adding in a term for minimum temperature and a term for maximum temperature. The results show that only the maximum temperature was significant and overall the r-squared improved minimally from the previous model. We ended up just keeping the sin and cosines terms with 2 and 4 pi. We tried adding terms all the way up to 12 pi, but while some were minimally significant they only added .001 to the adjusted R-squared. For simplicities sake, we left them out.

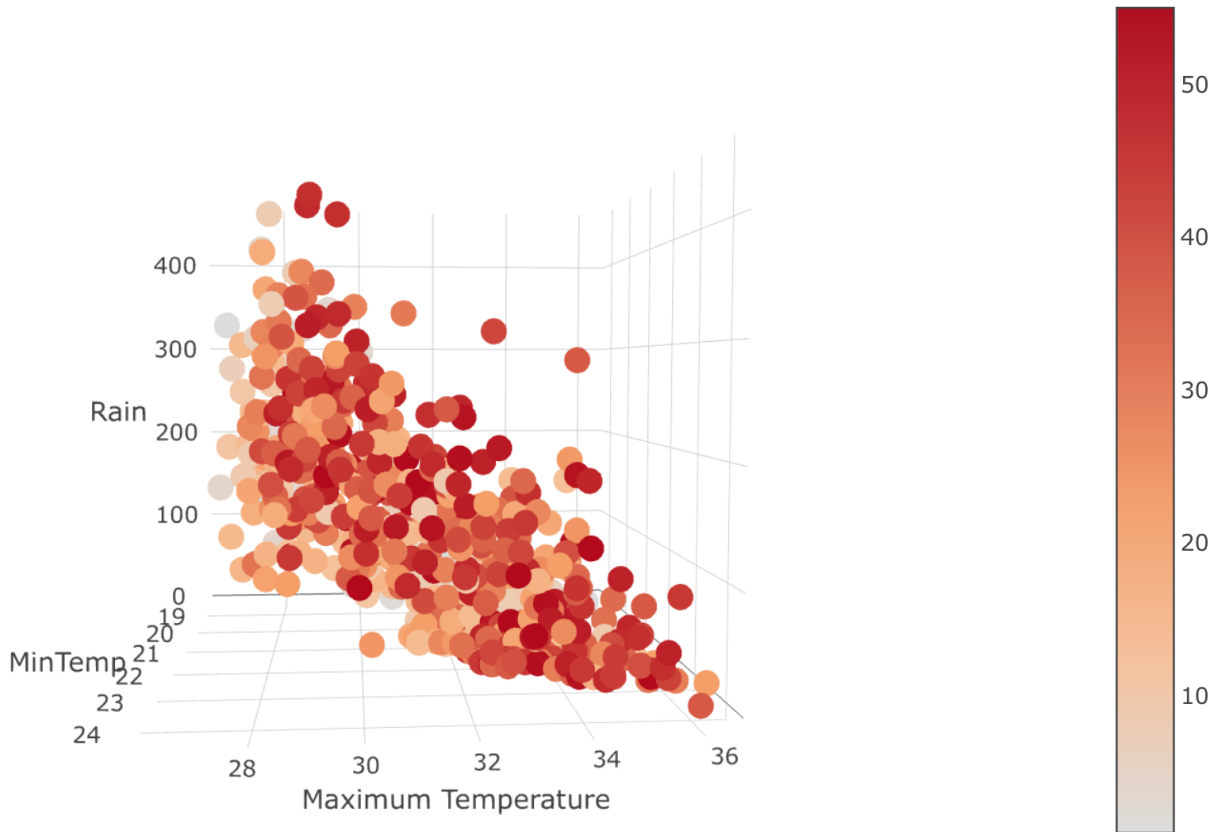
Table 3: Predicting Rain using Sines and Cosines and Temperature

	<i>Dependent variable:</i>
	Rainfall
Sine 2pi * Long Date	22.370*** (5.664)
Cosine 2pi * Long Date	-54.782*** (9.245)
Sine 4pi * Long Date	16.817*** (3.385)
Cosine 4pi * Long Date	-15.139*** (3.801)
Maximum Temperature	10.737 (6.762)
Minimum Temperature	-22.620*** (3.803)
Long Date	0.298 (0.240)
Constant	0.676 (390.469)
Observations	645
R <sup>2</sup>	0.633
Adjusted R <sup>2</sup>	0.629
Residual Std. Error	60.364 (df = 637)
F Statistic	156.985*** (df = 7; 637)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

# 3D Model of Minimum Temperature, Maximum Temperature and Rainfall

Since we are modeling the rainfall using minimum and maximum temperatures, we wanted to build a 3D model. Using a package called plotly, we were able to easily build a model and add a gradient of color to represent the year. The lighter the color, the earlier the time. If you view this report using the html format, you can move and zoom in around the points. In terms of time, there does NOT appear to be a relationship with rainfall. The colors on the rainfall axis are evenly distributed. However, there does seem to be relationships between temperature and rainfall. When rainfall is high, the temperatures are low. We also see a relationship between color(year) and temperature. When the color is lighter (earlier years), the temperatures seem to be lower. This is consistent to what was found in the previous report, as it shows and increase in temperature over time.





## Conclusion

All of our models are very similar in their ability to predict rainfall. For each of our models, our adjusted R-squares were around .63 with little variation. Since they were all so close, choosing a model was a little difficult. We chose to go with the model using sines, cosines and temperatures. This model was in the middle in terms of simplicity and R-squared value. From looking at all the models, there does not appear to be a significant change in rainfall over time, however there does seem to be a periodicity within the year. It would be helpful to know more information on the outliers we removed, like why there was a value of 100 millimeters of rain in December, when that appears to be the dry season. Is that value a human error, or just a unnatural month of rain? It would also be nice to know how the data was collected and if it was collected consistently.