Mini Project 3: Niamtougou

Adam May & Clayton Frink

Niamtougou is in the Kara region of Togo and is composed of six villages: Niamtougou, Koka, Baga, Ténéga, Yaka, and Agbandé. While Togo is classified as a tropical, sub-Saharan country its length stretches it through six geographic regions, making its climate vary from tropical to savanna. The Kara region is in northern Togo, thus Niamtougou has a tropical climate with an average elevation of 1535 ft. (mini-project 2, Donna Odhiambo & Matthew Gall).

The purpose of this project is to determine if there is a relationship between rainfall and average monthly minimum and maximum temperatures and time, as well as to create a model with rainfall predictions for future years.

Methodology

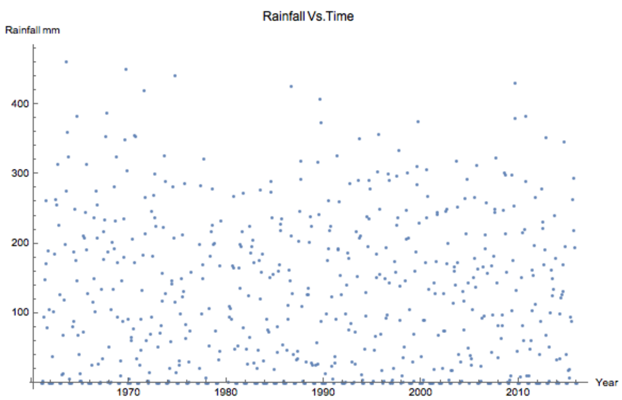
When taking a look at the available data, monthly rainfall measurements were unavailable for numerous months and sometimes almost entire years. We decided to omit both temperature and rainfall data from our models in those respective months for simplicity's sake.

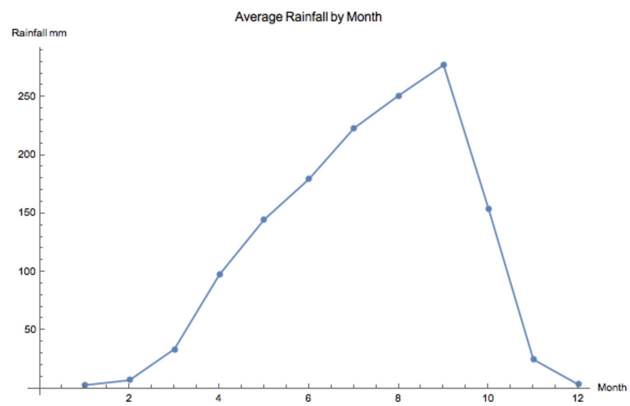
Additionally, we noticed in the period from July to December 1979 both minimum and maximum temperatures are calculated using the Excel average of the same month for the previous eighteen years. Because of this suspect data, we decided to omit these data points in our models.

Although it was never specified, we assume rainfall was measured in millimeters.

Relationship Between Rainfall and Time

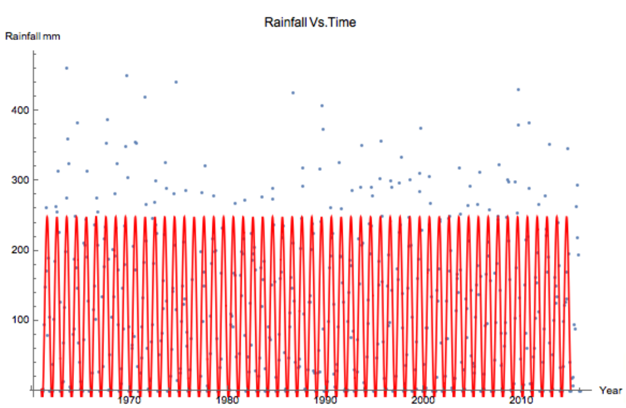
First, we set out to determine a possible relationship between rainfall and time. Below is a scatterplot of rainfall monthly rainfall measurements in Niamtougou over a fifty-five-year period ending in 2015.





Although it is not clear at first glance, we did notice a cyclical relationship between rainfall and time. This is better visualized by observing a plot of average rainfall for each month, which can be seen above.

On average, rainfall in Niamtougou seems to increase at a fixed rate beginning in January each year and reaches its apex in September. This is followed by a sharp decrease in rainfall and somewhat of drought lasting from November to March. We can try to represent this relationship by using sine and cosine terms to explain any annual oscillations as well as a linear term to explain any increase or decrease over time. Below is a graph of our oscillating model.

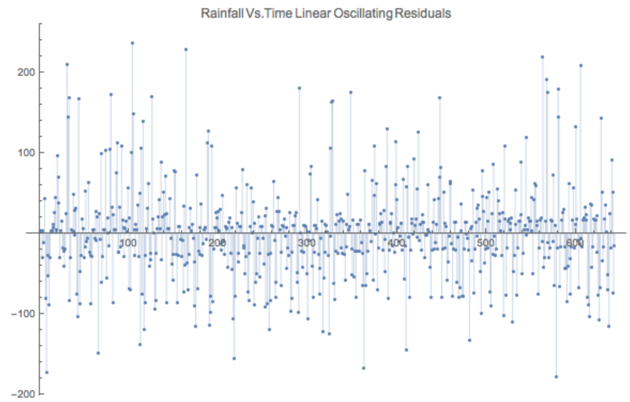


Our model seems to capture predictable seasonal variations in rainfall well, however months with heavy rainfall that occur every twenty or so years are not captured. A parameter table can be seen abovee. The expression for our model is as follows:

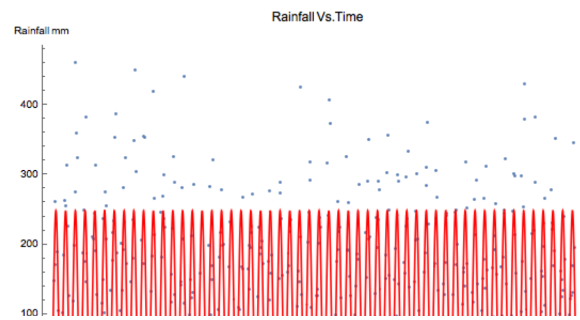
	Estimate	Standard Error	t-Statistic	P-Value
1	638.103	293.402	2.17484	0.030008
Sin[2 π t]	−64.9732	3.32743	−19.5266	6.05965×10^{-67}
Cos[2 π t]	−114.694	3.33475	−34.3937	1.62306×10^{-147}
t	−0.262196	0.147522	−1.77733	0.075989

$$638.103 - 0.262t - 114.694\text{Cos}(2\pi t) - 64.973\text{Sin}(2\pi t)$$

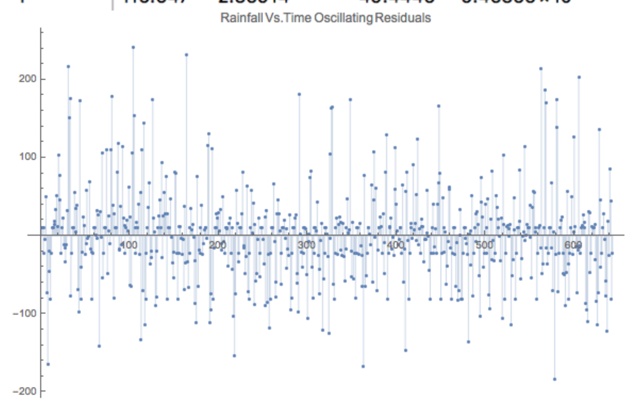
Both the sine and cosine and term’s extremely low p-values confirm our hypothesis of seasonal variation; however, the linear term t’s p-value is just high enough for us to doubt its significance. However, this does show us that Niamtougou has experienced a slight decrease in rainfall over the period being studied. To the right, residual plot can be seen.



Next, we’ll model rainfall over time without the linear t term., a graph of our model as well as its parameter table and residual plot can be seen.



	Estimate	Standard Error	t-Statistic	P-Value
1	116.647	2.35914	49.4445	9.45355×10^{-221}



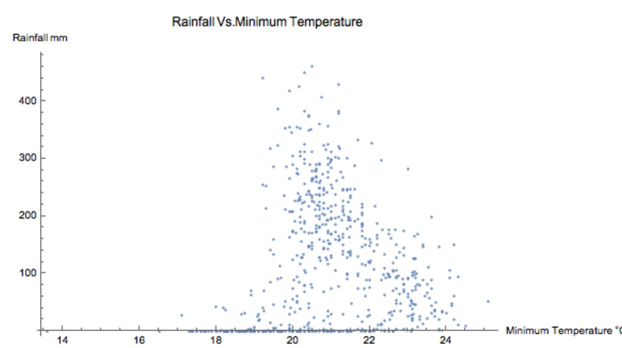
As we expected, both the sine and cosine terms are significant to our model. The expression for our model is as follows:

$$116.647 - 114.677\cos(2\pi t) - 64.8157\sin(2\pi t)$$

In addition, there is a slight wave pattern noticeable in our residuals with a period of about thirty years. This is our best model for predicting rainfall data in Niamtougou over time. Much of the variation of rainfall can be explained by the variable time. Our model has a R-squared value of 0.708 which largely confirms this. Although we cannot conclude whether rainfall has significantly increased or decreased, it has a clear relationship with time.

Relationship Between Rainfall & Minimum Temperatures

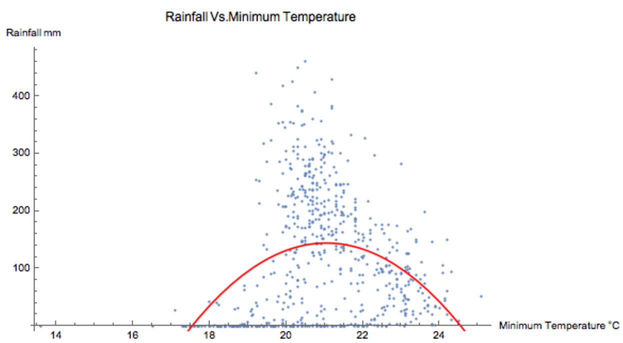
Next, we set out to find a possible relationship between rainfall and minimum temperatures. To the right, a scatterplot of rainfall measurements over minimum temperature can be seen. Initially, we weren’t quite sure what to make of this graph. It seemed that relatively large amounts of rainfall didn’t occur at relatively low and high minimum



temperatures and our scatterplot slightly resembled a normal distribution.

We decided a quadratic model would be best for modeling this behavior. Below is a plot of our model.

Quadratic Model

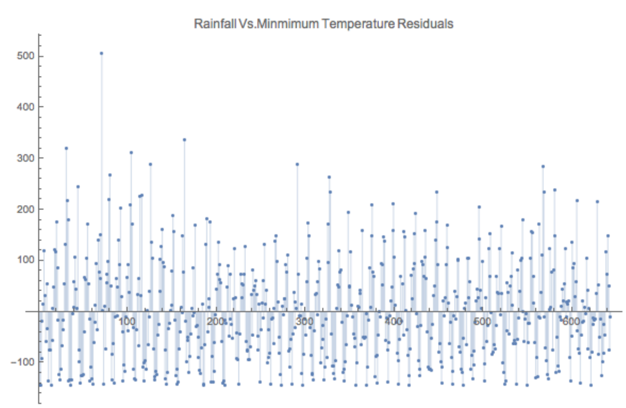


	Estimate	Standard Error	t-Statistic	P-Value
1	−5058.47	479.528	−10.5489	4.27398×10^{-24}
t	494.419	46.141	10.7154	9.36839×10^{-25}
t^2	−11.746	1.10738	−10.607	2.51959×10^{-24}

The expression for our model is:

$$-5{,}058.47 + 494.419t - 11.746t^2$$

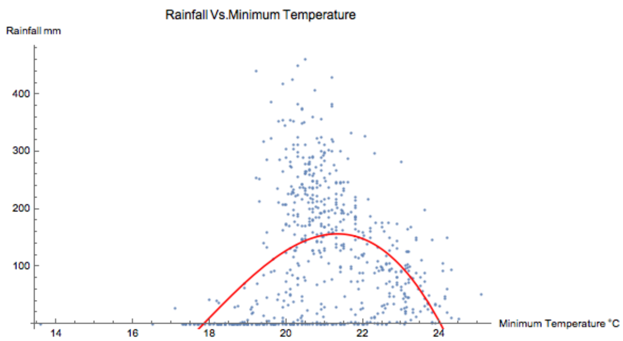
Both the linear t term and quadratic t^2 term have extremely low p-values and significant value to our model. But we still weren't quite sure what connection our model had to the real world. Our model seems to verify our suspicion that months with a lot of precipitation for the most part occur in months with average minimum temperatures and months with either extremely low or high minimum temperatures have little rainfall. Below a plot of the model's residuals can be seen.



It's apparent that the huge underestimation of rainfall early in the time period being studied can be explained by our model's negative values as minimum temperatures become large or small. This could be fixed by the use of a function that approaches zero as temperatures become very large or small. Otherwise, there doesn't appear to be any obvious patterns in our residuals. We do not have much confidence in this model. Our reasoning is that we have many data points where the rainfall equals zero.

Logarithmic Model

Just for curiosity's sake, we decided to add a logarithmic term to our model and see what would happen. Below is the plot for our model.

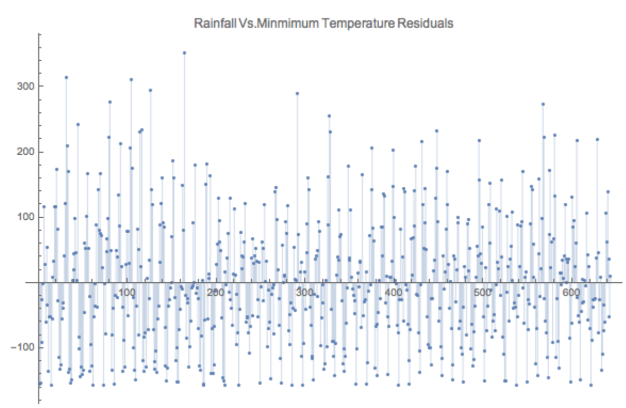


	Estimate	Standard Error	t-Statistic	P-Value
1	52 857.8	10 042.3	5.26352	1.93229×10^{-7}
t	4443.57	685.493	6.4823	1.80727×10^{-10}
t^2	−61.1968	8.63301	−7.0887	3.5944×10^{-12}
Log[t]	−39 096.5	6771.71	−5.7735	1.2119×10^{-8}

The expression for our model is:

$$52{,}857 + 4{,}4443.57t - 61.1968t^2 - 39{,}096.5\ln(t)$$

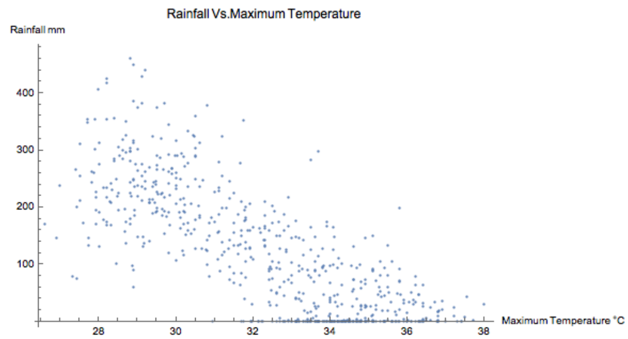
The extremely low p-value of the natural log term was very surprising to us. We believe it's significance can be explained by the higher absolute maximum of our model, its slightly leftward skew, and almost linear behavior before and after the absolute maximum. Below is a plot of the model's residuals.



The residuals lack of large negative values and plethora of large positive values can be explained by our model's failure to predict very large rainfall measurements. Otherwise, no obvious pattern is apparent.

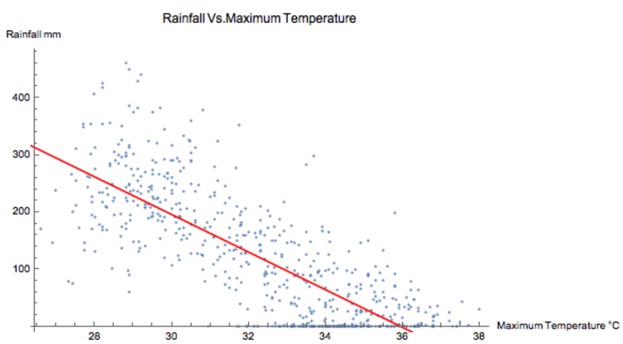
Relationship Between Rainfall & Maximum Temperatures

Our next task was to model rainfall as a function of average maximum temperature. Below is a scatterplot of rainfall for maximum temperature values. It's easy to see a negative linear relationship, which is what we fist sought out to model.



Linear Model

Below is graph of our linear rainfall model.



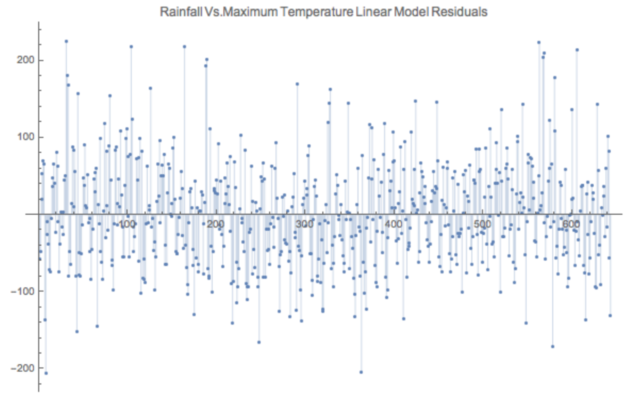
	Estimate	Standard Error	t-Statistic	P-Value
1	1182.18	33.4979	35.291	2.19403×10^{-152}
t	−32.8775	1.02945	−31.9369	1.19188×10^{-134}

The expression for our model is:

$$1182.18 - 32.878t$$

The linear relationship indicates a consistent decrease in rainfall as average maximum temperature increases. We think this could be due to warmer temperatures evaporating more moisture, but we'll leave that up to the climatologists. There also seems to be substantial variation in rainfall for any particular temperature, but as a whole rainfall decreases as temperature rises.

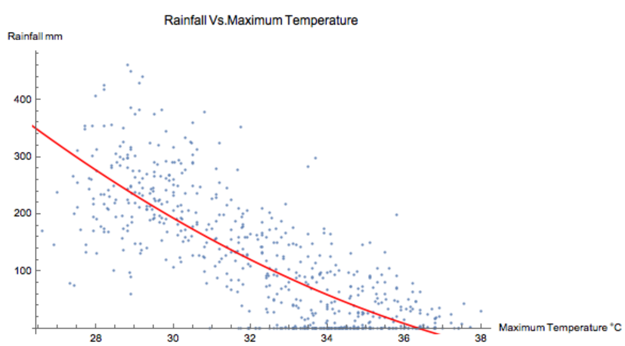
The parameter table's extremely low p-value for the linear t term indicates how necessary it is for our model. Below a residual plot for our model.



There does appear to be a slight wave-like pattern in the residuals. We'll add a quadratic t^2 to see if it makes any difference.

Quadratic Model

Below is graph of our quadratic rainfall model.



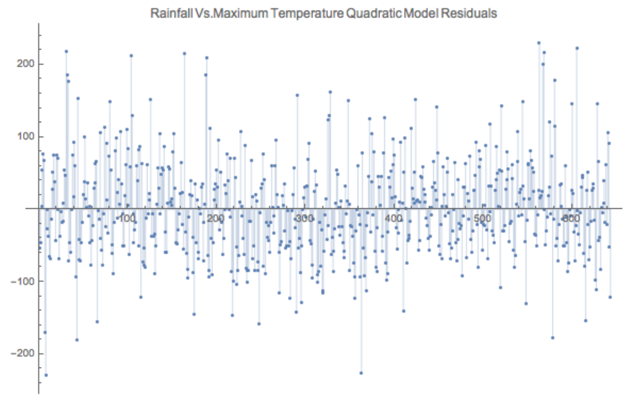
	Estimate	Standard Error	t-Statistic	P-Value
1	2599.11	418.234	6.21449	9.28118×10^{-10}
t	−121.444	26.0794	−4.65671	3.90881×10^{-6}
t^2	1.37466	0.404473	3.39865	0.000719237

For the most part, it looks the same as our linear model with some slight curvature.

The expression for our model is:

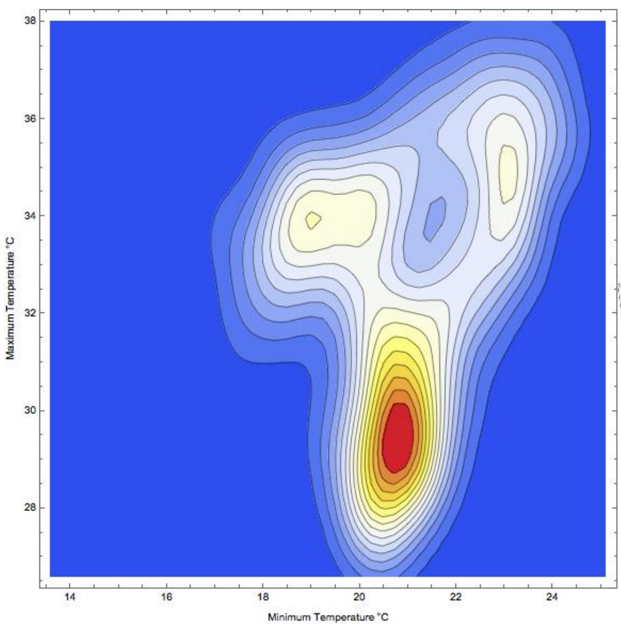
$$2{,}599.11 - 121.444t + 1.375t^2$$

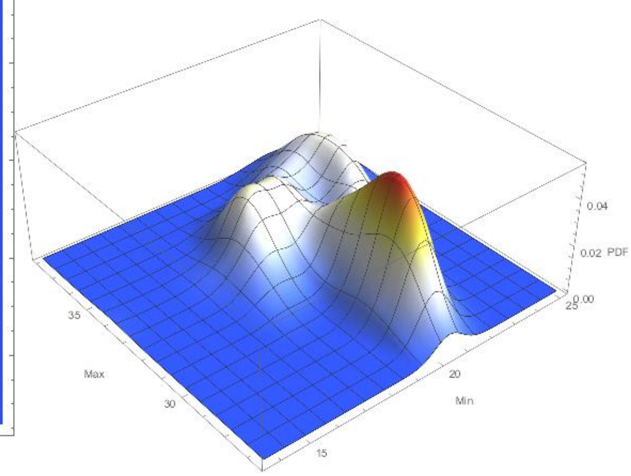
The quadratic t^2 term's is low enough for us to not doubt its meaningful contribution to the model. A plot of our residuals can be seen below.



No pattern is evident in our residuals. This quadratic model is our best attempt to represent the relationship between rainfall and average maximum temperatures. Another model we considered trying is exponential decay.

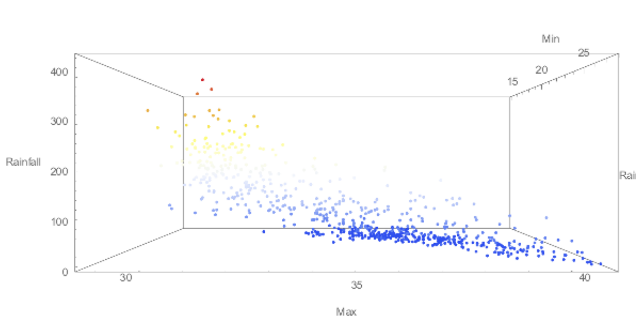
More Interesting Graphics

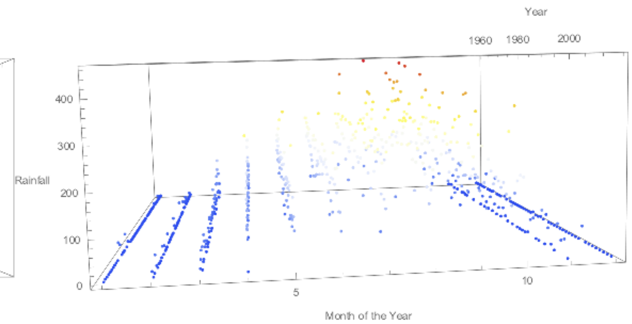




1: Contour plot of temperatures. Our data is centered around 21°C minimum temperature and 29°C maximum temperature.

2: 3D PDF plot of rainfall, minimum temperature, and maximum temperature.





3: 3D PDF plot of rainfall, minimum temperature, and maximum temperature.

4: 3D plot of rainfall by month over time. Rainfall peaks in September.

Possible Outliers

Year	Month	Minimum Temperature	Maximum Temperature	Rainfall
1966	November	13.6 °C		
1963	July			460.6 mm
1969	February		38 °C	
1969	August			450.6 mm

Additional Questions for the Togolese

1. How did they calculate average monthly minimum and maximum temperatures?
2. Are we correct in assuming they measured rainfall in millimeters?
3. What happened from July to December 1979?

Conclusions

We sought out to determine possible relationships between rainfall, average monthly minimum and maximum temperatures, and time in order to better predict future rainfall in Niamtougou. We had to throw out a number of data points because of missing or inaccurate measurements. Some variables were difficult to find relationships, particularly rainfall and average minimum temperature, and other variables had obvious relationships like rainfall and average maximum temperature and rainfall and time.

Our final model will not include average minimum temperature and will instead only include the explanatory variables average maximum temperature and time. We will combine these two models for rainfall by taking a simple average. The equation below is the result, where *t* represents time in years and *m* represents average maximum temperature in degrees Celsius.

$$Rainfall\ (mm) = \frac{((116.647 - 114.677Cos(2\pi t) - 64.816Sin(2\pi t) + (2{,}599.11 - 121.444m + 1.375m^2)\ )}{2}$$

We believe this model does a fair job combining the relationship rainfall has with both average maximum temperature and time. Although it is a crude model, we think it does an effective job.