Dapaong Monthly Temperature Averages and Monthly Rainfall Averages

Chris Milesky and Patrick Nielsen

1. In the first report, by Rachel Driehaus, Connor Edwards, Terra Ficke, the conclusion was that the minimum was increasing and the maximum was decreasing. They were dissatisfied with their conclusion that the temperatures were decreasing because of an outlier included in the creation of the model. The second Dapaong Report, by Maria Ruwe, Leah Gillespie, Austin Hardesty, using monthly averages for the data agreed with the conclusion that the minimum temperature was increasing in Dapaong. They did not agree that the maximum temperature was increasing or decreasing. That monthly temperature data is the same we used and our analysis reached some of the same conclusions. We removed more outliers in the than the second report as they used 3 times the interquartile range instead of the standard 1.5 times the interquartile range. This ruthless identification and removal of the temperature outliers meant that we took out many more outliers than the groups previously had and by consequence had a model that fit better. Refer to significant temperature change as half a degree or more.

2.Dapaong is located in the north of Togo, in the Savanes region. It is the northern most city in Togo for which we have data. The rainy season takes place during the summer with the rainiest average month being August. The dry season takes place during the winter and seldom receives any rain having many years of 0 millimeters of rain for years during December and January.

3. We used InterQuartile ranges to identify and remove outliers. Upper Quartile plus 1.5 times InterQuartile Range and Lower Quartile minus 1.5 times the InterQuartile Range formed our boundaries.

In the Maximum data, outliers would be greater 42.886 or less than 24.0077. Five data points were outside of the bounds and removed. These were at months 411, 418, 430, 466, and 468.

For the Minimum data, outliers would reside outsides of the bounds of greater than 27.2524 and less than 17.6105. Thirty nine outliers were identified and removed. The data points to be removed were 1, 12, 13, 36, 60, 61, 65, 66, 72, 73, 85, 96, 97, 108, 120, 121,133,145,169, 180, 181, 193, 204, 205, 216, 217, 228, 259, 283, 292, 396, 408, 409, 531, 543, 544, 570, 591,and 593.

In the rain data the outliers all took place in august which after was further research was the rainiest time in Dapaong so the abnormally high rainfall was left in as it was actual data that high and only took place in the month of August.
Our Data did contain missing rain averages that were replaced with *****'s. We found and removed these points as they would have provided no relevant information to the models and caused errors in Mathematica. The points with ***** existed at all of the raindata table positions: 64, 65, 66, 210, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 370, and 371.

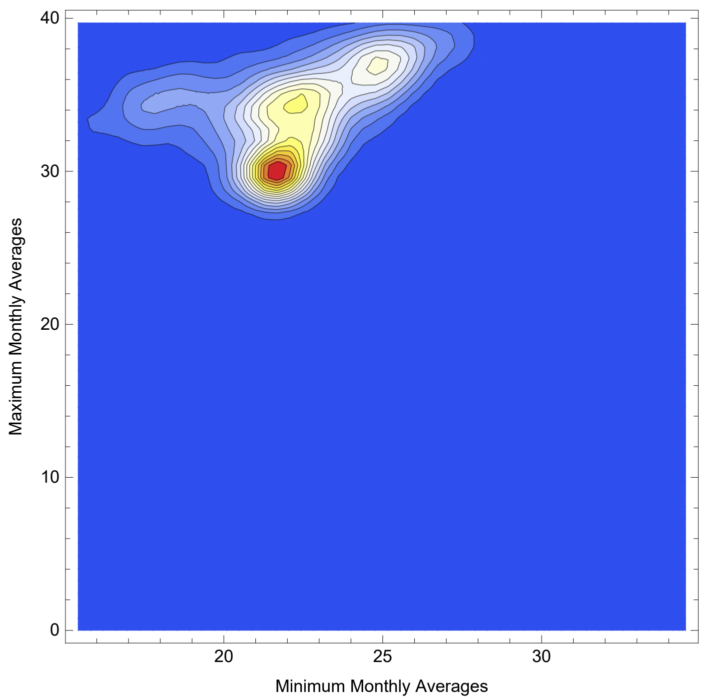
4. The rainfall in Dapaong primarily oscillates during the year between the dry and rainy season. Which would be around our 6 month oscillation term causing the variation between the high rainfall, lower monthly temperature averages to low rainfall, higher monthly averages for temperature. The half year/six month oscillation terms working the best in every model gave a hint that there was some variation in the rain over time.

5. Rainfall(time, temperature) Models

Created a PDF of the minimum and maximum data. Displayed a high probability of certain temperature values.

In[132]:=
```
𝒟 = SmoothKernelDistribution[Transpose[{mintemp, maxtemp}]];
contour = ContourPlot[PDF[𝒟, {x, y}] // Evaluate, {x, Min[mintemp], Max[mintemp]},
  {y, Min[maxtemp], Max[maxtemp]}, PlotRange → All, AxesLabel → {"Minimum", "Maximum"},
  PlotPoints → 50, Contours → 15, ColorFunction → "TemperatureMap",
  FrameLabel → {"Minimum Monthly Averages", "Maximum Monthly Averages"}]
```

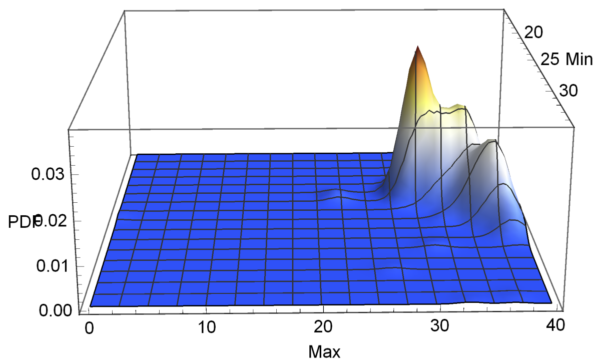
Out[133]=



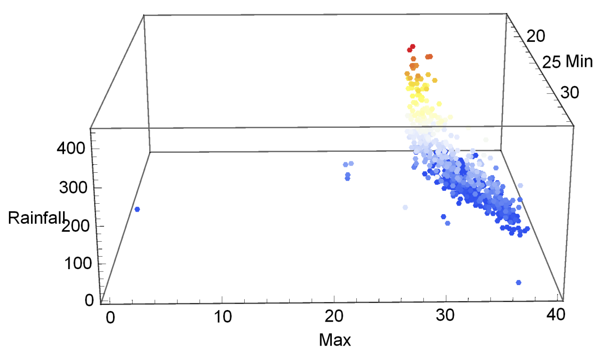
Added in the rain data and then removed the data points that would contain *****'s in the rain value. This 3 coordinate data set was used to create a model of the Rainfall with both Maximum and Minimum temperatures.

```
In[134]:= dens = Plot3D[PDF[𝒟, {x, y}] // Evaluate, {x, Min[mintemp], Max[mintemp]},
   {y, Min[maxtemp], Max[maxtemp]}, PlotRange → All, PlotPoints → 50,
   ColorFunction → "TemperatureMap", AxesLabel → {"Min", "Max", "PDF"}]
threeDmodel = Transpose[{mintemp, maxtemp, rain}];

threeDmodel =
  Delete[threeDmodel, {{64}, {65}, {66}, {210}, {214}, {215}, {216}, {217}, {218}, {219},
    {220}, {221}, {222}, {223}, {224}, {225}, {226}, {227}, {228}, {370}, {371}}];

ListPointPlot3D[threeDmodel, PlotRange → All,
 ColorFunction → "TemperatureMap", AxesLabel → {"Min", "Max", "Rainfall"}]
```

Out[134]=



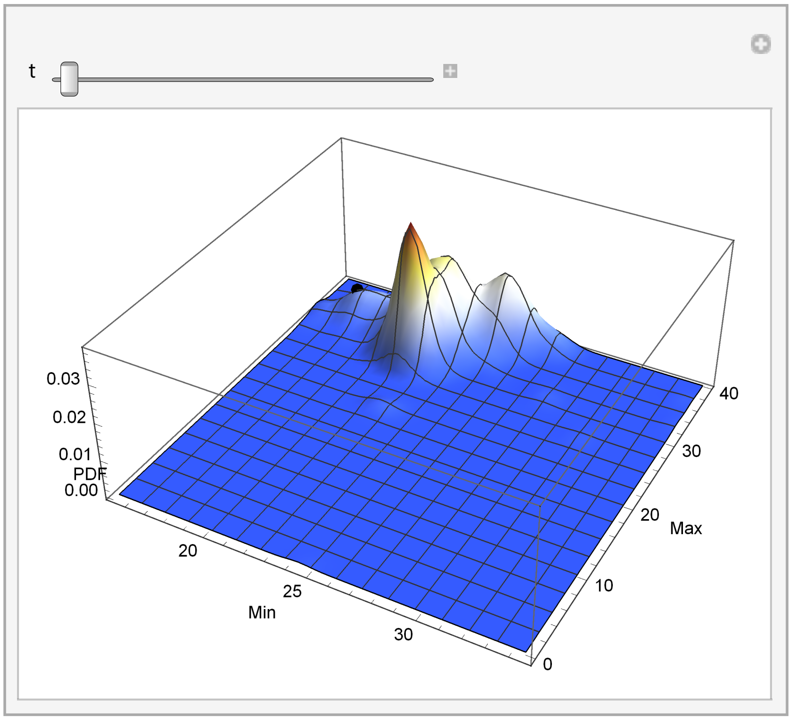
Out[137]=



Scaled the rain to be used as a plot point for future manipulation.

In[138]:=
```
scalerain = rain * .03 / Max[rain];
Manipulate[
 p1 = ListPointPlot3D[{{mintemp[[t]], maxtemp[[t]], scalerain[[t]]}
    , {mintemp[[t]], maxtemp[[t]], PDF[𝒟, {mintemp[[t]], maxtemp[[t]]}] + scalerain[[t]]}},
   PlotRange → All, PlotStyle → {Black, PointSize[0.02 + scalerain[[t]]]}];
 Show[dens, p1],
 {{t, 1}, 1, Length[rain], 1}
]
```

Out[139]=



Linear models for the Minimum temperature, Maximum temperature, and the rainfall data. The periods are one year(1/12), six months(2/12), or three months(4/12).

In[140]:=
```
minmodel = LinearModelFit[datamin, {1, x,
    Sin[2 Pi (1 x / 12)], Cos[2 Pi (1 x / 12)],
    Sin[2 Pi (2 x / 12)], Cos[2 Pi (2 x / 12)]}, x];
minmodel["ParameterConfidenceIntervals"]
minmodel["ParameterTable"]
minmodel["RSquared"]
ListPlot[minmodel["FitResiduals"], Filling → Axis,
 PlotLabel → "Residuals Plot for Min Data", Axes → True, AxesLabel → {"\!\(\*
StyleBox[\"Months\",\nFontWeight->\"Plain\"]\)", "Temp"}]
p1 = ListPlot[datamin];
p2 = Plot[minmodel[x], {x, 0, 660}];
Show[p1, p2, Frame → True, PlotLabel → "Report 3 Minimum graph",
 FrameLabel → {"Months since January 1961", "Monthly Minimum Average Temperature"}]
```

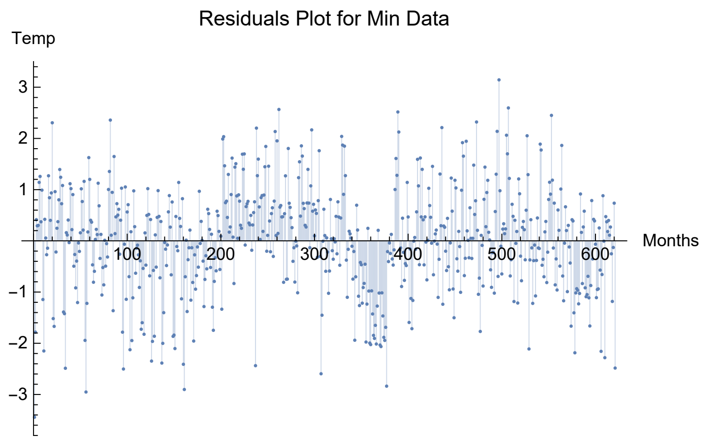
Out[140]=
```
{{21.0984, 21.4389}, {0.00278415, 0.00366205}, {1.15042, 1.38062},
 {-1.40215, -1.16296}, {-1.14181, -0.908904}, {-0.943989, -0.707542}}
```

Out[141]=

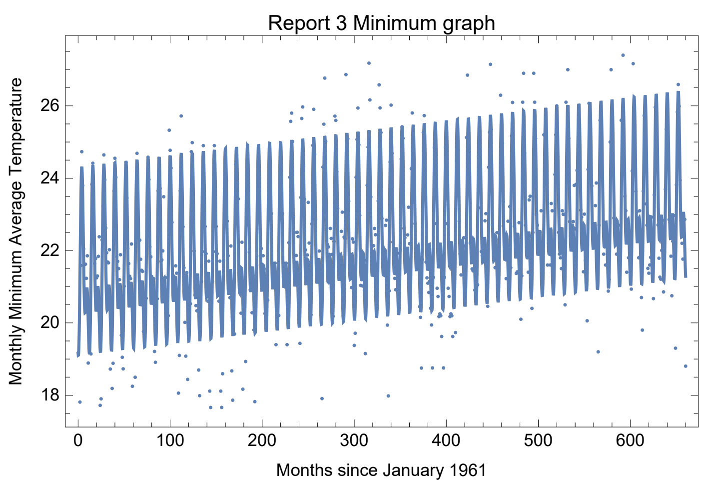
	Estimate	Standard Error	t-Statistic	P-Value
1	21.2686	0.0866865	245.351	$1.026871875359\times10^{-615}$
x	0.0032231	0.000223519	14.4198	8.10412×10^{-41}
$\mathrm{Sin}\left[\frac{\pi x}{6}\right]$	1.26552	0.0586111	21.5918	2.19548×10^{-77}
$\mathrm{Cos}\left[\frac{\pi x}{6}\right]$	−1.28256	0.0608976	−21.0608	1.49419×10^{-74}
$\mathrm{Sin}\left[\frac{\pi x}{3}\right]$	−1.02536	0.0592986	−17.2914	6.88197×10^{-55}
$\mathrm{Cos}\left[\frac{\pi x}{3}\right]$	−0.825765	0.0602005	−13.7169	1.49184×10^{-37}

Out[142]= 0.707355

Out[143]=



Out[146]=



In the Minimum model, we removed the three month oscillation terms as they had parameters of confidence that included zero. The residuals model appears random besides a primarily positive section between 200 months and 350 months. With a R-Squared of 0.707355 This model is better than that of the previous group who had a minimum model R-Squared of 0.46. The Minimum model appears significantly increasing.

In[147]:=
```
maxmodel = LinearModelFit[datamax, {1, x,
    Sin[2 Pi (1 x / 12)], Cos[2 Pi (1 x / 12)],
    Sin[2 Pi (2 x / 12)], Cos[2 Pi (2 x / 12)]}, x];
maxmodel["ParameterConfidenceIntervals"]
maxmodel["ParameterTable"]
maxmodel["RSquared"]
ListPlot[maxmodel["FitResiduals"], Filling → Axis,
 PlotLabel → "Residuals Plot for Max Data", Axes → True, AxesLabel → {"\!\(\*
StyleBox[\"Months\",\nFontWeight->\"Plain\"]\)", "Temp"}]
p1 = ListPlot[datamax];
p2 = Plot[maxmodel[x], {x, 0, 660}];
Show[p1, p2, Frame → True, PlotLabel → "Report 3 Maximum graph",
 FrameLabel → {"Months since January 1961", "Monthly Maximum Average Temperature"}]
maxmodel[1]
maxmodel[661]
```

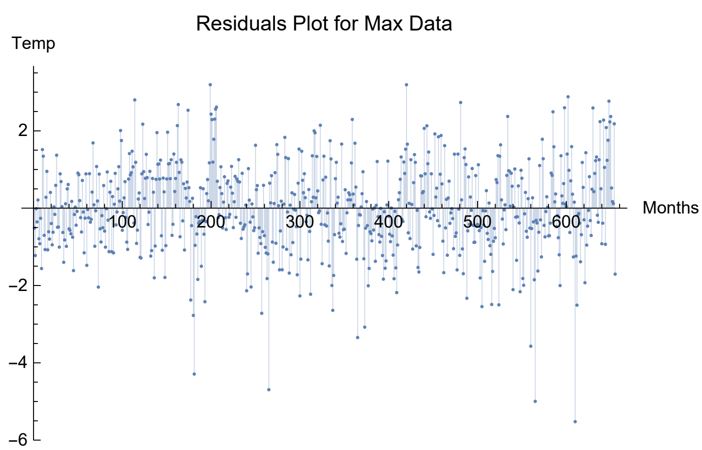
Out[148]= {{33.8047, 34.1665}, {-0.00129971, -0.000350193}, {2.92179, 3.17886}, {1.02711, 1.28298}, {-1.73624, -1.47957}, {-0.691522, -0.43525}}

Out[149]=

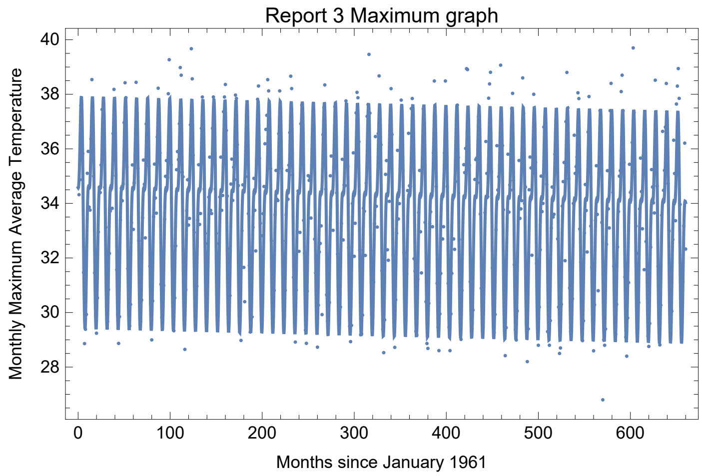
	Estimate	Standard Error	t-Statistic	P-Value
1	33.9856	0.0921158	368.944	$2.768903858514 \times 10^{-756}$
x	-0.000824954	0.000241778	-3.41204	0.000684845
$\text{Sin}\left[\frac{\pi x}{6}\right]$	3.05033	0.0654583	46.5995	3.11977×10^{-209}
$\text{Cos}\left[\frac{\pi x}{6}\right]$	1.15504	0.0651538	17.728	1.2108×10^{-57}
$\text{Sin}\left[\frac{\pi x}{3}\right]$	-1.6079	0.0653565	-24.6021	6.30514×10^{-95}
$\text{Cos}\left[\frac{\pi x}{3}\right]$	-0.563386	0.0652547	-8.63364	4.58602×10^{-17}

Out[150]= 0.831605

Out[151]=



Out[154]=



Out[155]= 34.8361

Out[156]= 34.2916

In the Maximum model, we removed the three month oscillation terms as the Cosine had parameters of confidence that included zero. The residuals model appears random. With a R-Squared of 0.831605 This model fits the data far better than the previous groups cubic model with oscillation terms, which had a R-Squared of 0.6359. Using the model to predict a value 55 years forward from January 1961 we can look if there was a significant decrease or increase in the model's predicted temperature.
There was a decreasing of 0.5444 which means it is possible there is a significant decrease in maximum temperatures of Dapaong.

In[158]:=
```
rainmodel = LinearModelFit[newraindata, {1, x,
    Sin[2 Pi (1 x / 12)], Cos[2 Pi (1 x / 12)],
    Sin[2 Pi (2 x / 12)], Cos[2 Pi (2 x / 12)]}, x];
rainmodel["ParameterConfidenceIntervals"]
rainmodel["ParameterTable"]
rainmodel["RSquared"]
ListPlot[rainmodel["FitResiduals"], Filling → Axis,
 PlotLabel → "Residuals Plot for Rain Data", Axes → True, AxesLabel → {"\!\(\*
StyleBox[\"Months\",\nFontWeight->\"Plain\"]\)", "Rain"}]
p1 = ListPlot[newraindata];
p2 = Plot[rainmodel[x], {x, 0, 660}];
Show[p1, p2, Frame → True, PlotLabel → "Report 3 Minimum graph",
 FrameLabel → {"Months since January 1961", "Monthly Average Rainfall"}]
```

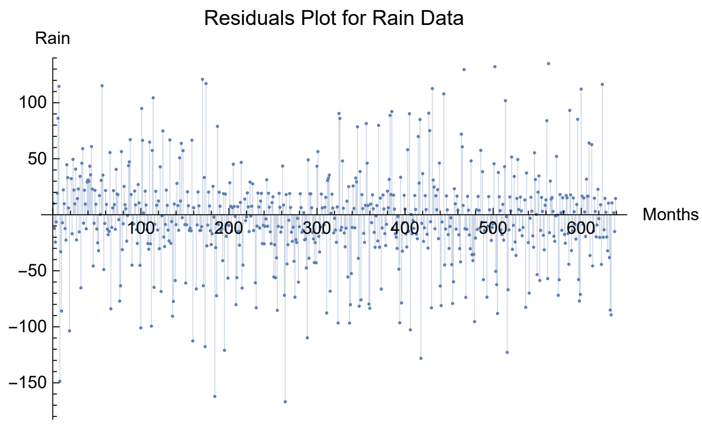
Out[159]= {{75.5153, 90.9628}, {-0.00807527, 0.0319877}, {-82.7632, -71.9153}, {-92.5582, -81.6599}, {29.3524, 40.2264}, {-23.9496, -13.0767}}

Out[160]=

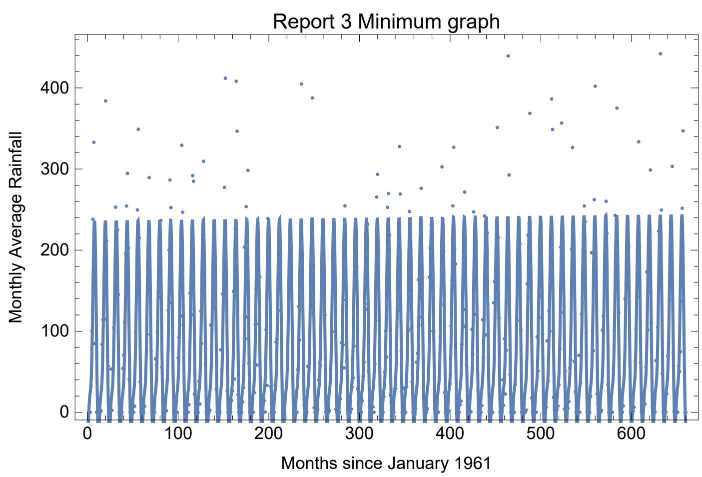
	Estimate	Standard Error	t-Statistic	P-Value
1	83.2391	3.93322	21.1631	1.39844×10^{-75}
x	0.0119562	0.0102008	1.17209	0.241603
$\mathrm{Sin}\left[\frac{\pi x}{6}\right]$	−77.3392	2.76207	−28.0005	7.27936×10^{-113}
$\mathrm{Cos}\left[\frac{\pi x}{6}\right]$	−87.109	2.7749	−31.3918	3.74351×10^{-131}
$\mathrm{Sin}\left[\frac{\pi x}{3}\right]$	34.7894	2.76873	12.5651	1.73858×10^{-32}
$\mathrm{Cos}\left[\frac{\pi x}{3}\right]$	−18.5131	2.76845	−6.68718	5.00417×10^{-11}

Out[161]= 0.758574

Out[162]=



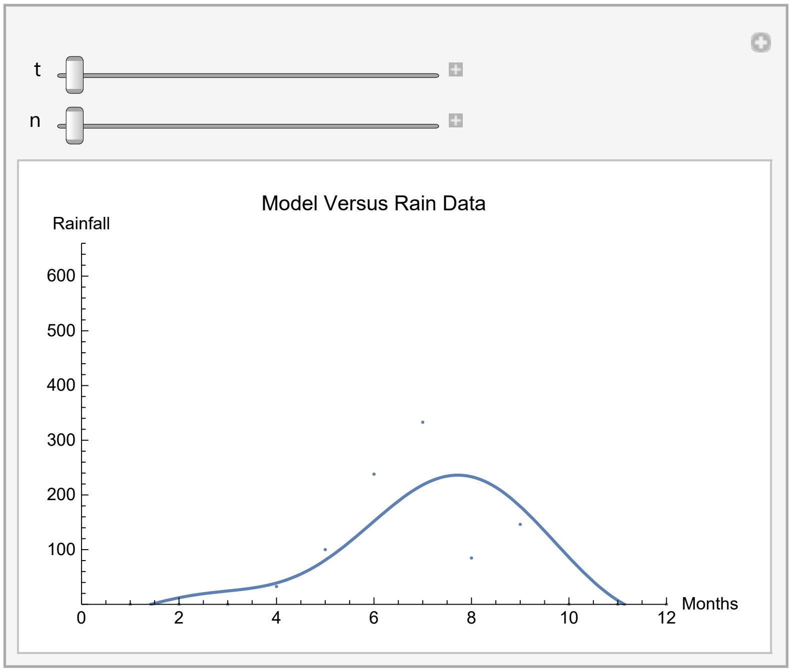
Out[165]=



The rain model the R-Squared was 0.758574 the highest out of any of our models. The oscillation terms were for a quarter year oscillation were again removed as the Cosine contained zero as a potential parameter. The rain model seems also be inconclusive as to whether or not it is significantly increasing or decreasing. The coefficient on the x term is small, 0.0119, and does not provide much slope to conclude either way.

In[166]:=
```
raindataplot = ListPlot[newraindata];
Manipulate[
 rainmodelplot = Plot[rainmodel[x], {x, t * 12, (t + n) * 12},
   PlotRange → {{t * 12, (t + n) * 12}, {Min[newraindata], Max[newraindata]}},
   PlotLabel → "Model Versus Rain Data", AxesLabel → {"Months", "Rainfall"}];
 Show[rainmodelplot, raindataplot],
 {{t, 0}, 0, 50, 1},
 {{n, 1}, 1, 5, 1}]
```

Out[167]=



This manipulate displays the rain model overlaid with the rain data so show how the model fits the data.

---

8. Moving forward we would like to find ways to fix the issues of the missing data in rainfall. The missing year and a quarter in a row did not help our ability to model the data accurately. Is the rainfall in the month of August just so from the mean rainfall for the year that is should be considered an outlier? Was that a mistake of our calculations of the InterQuartile Ranges? It would be interesting to hear back from the Togolese meteorologist on some of the previous mini reports before preceding to see what they think of all this and discuss with the the implication that global climate change has on a nation such as Togo. Have the Togolese meteorologist also worked with the data and we could compare our models with their own?