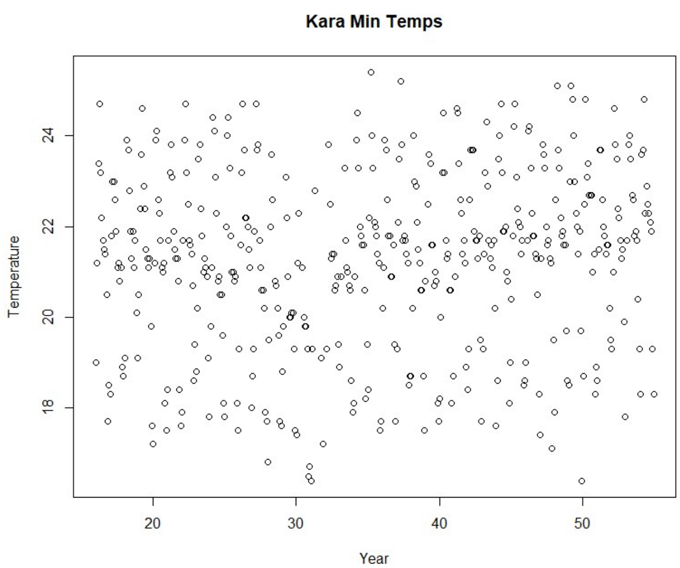
Final City Model: Kara

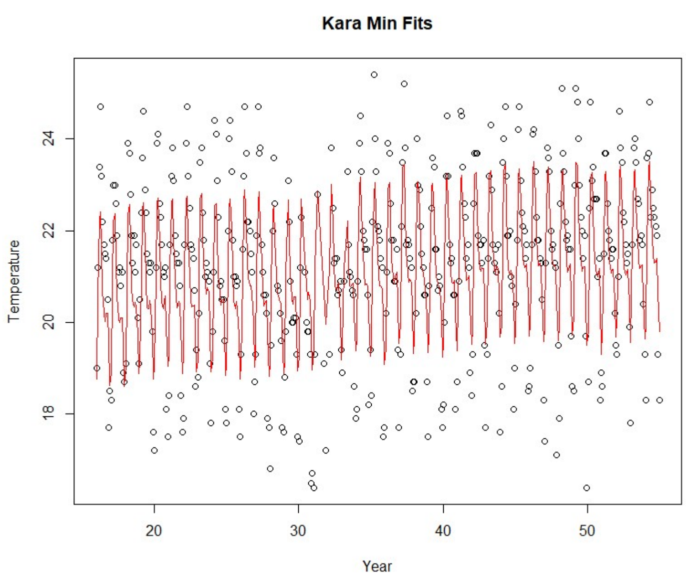
Alyssa Farmer

Minimum Temperature Model

$R^2 = 0.7237$

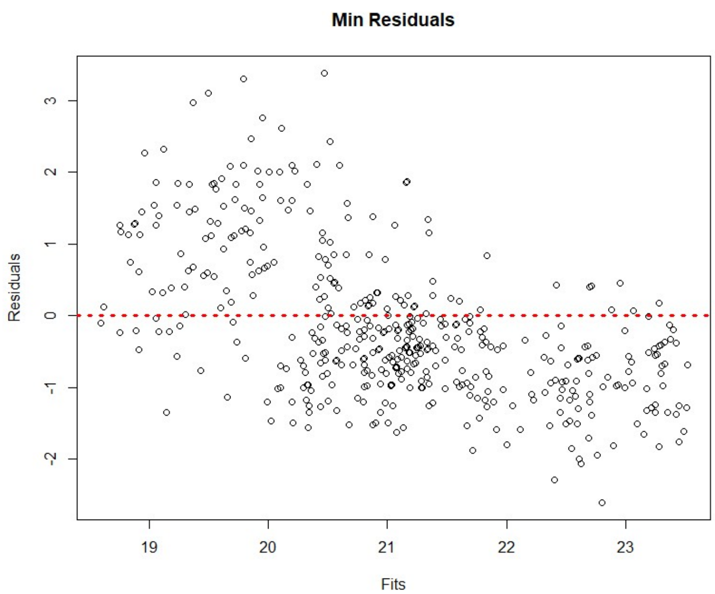
A plot of the Kara minimum temperatures over time can be seen below. Some outliers have been removed, so there are no longer any noticeable questionable values. However, it is also difficult to tell if there is an increase in temperature over time. The final model was fit to the data and can be seen below the plot of just the original data.

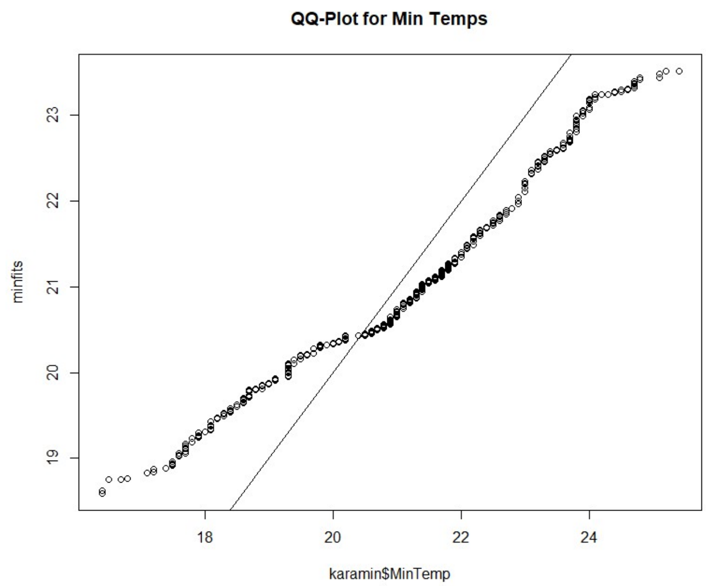




The residual plot for the minimum temperatures does not look very good. Most of the residuals in the earlier years are positive, while most of the residuals in later years are negative. This violates the assumption that the residuals are independently and identically distributed, as the variance is not constant over time. The qq-plot also very obviously shows that the predicted and actual minimum temperatures did not come from the same probability distribution, which is also an issue.

As for how well the model fits the data, the value for R^2 turned out to be 0.7237, which indicates a decently strong relationship between the predicted and actual minimum temperature values. However, just because the relationship is strong, this does not mean that the model fits very well or is consistent.

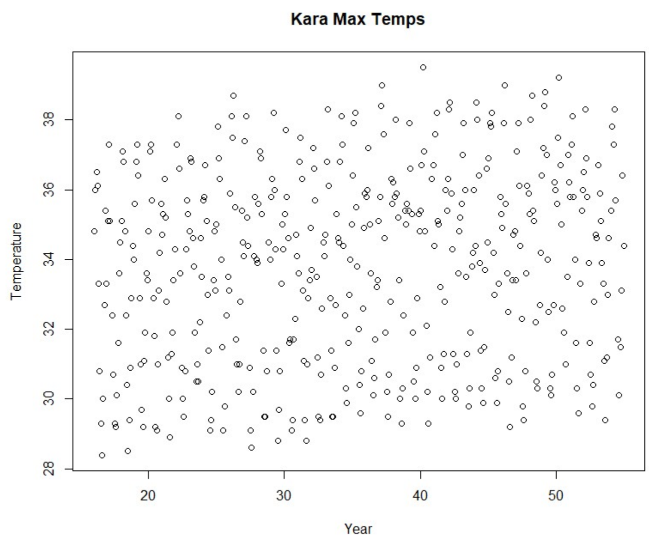


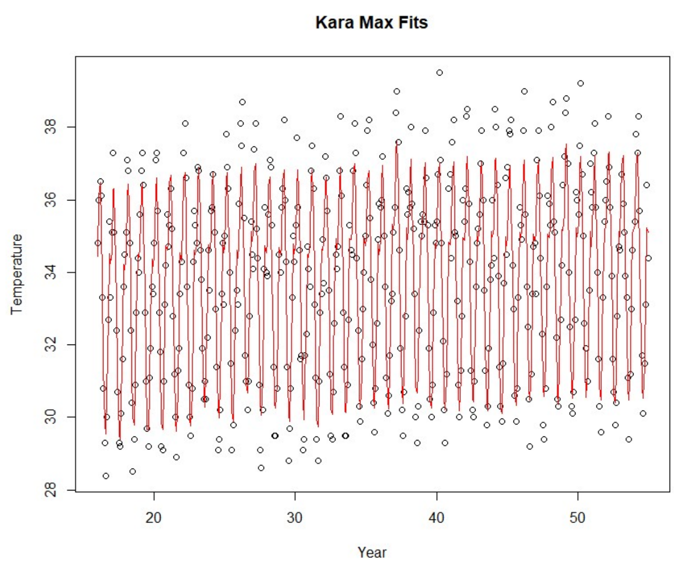

QQ-Plot for Min Temps
minfits
19
20
21
22
23
18
20
22
24
karamin$MinTemp

Maximum Temperature Model

$R^2 = 0.9116$

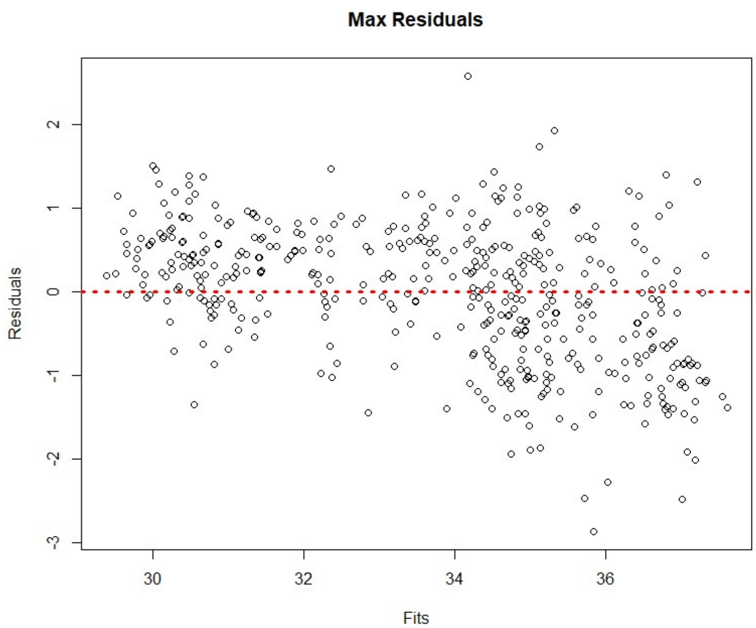
The maximum temperature data looks similar to the minimum temperatures, although there is a slightly more obvious increase in the temperatures over time. However, it is difficult to tell by how much the temperatures are actually increasing. The fitted values against the actual values are plotted again below.

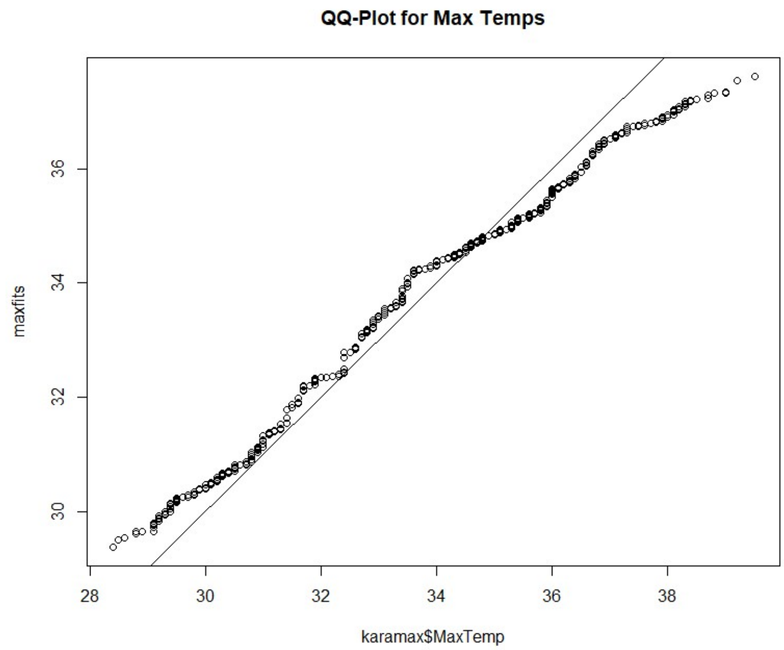




The residuals for the maximum temperatures look slightly better than that of the minimum temperatures, but they are still not great. The variance seems to increase over time. However, there is not an obvious increase or decrease in these, which is an improvement. Looking at the qq-plot, there still seem to be some issues, since the majority of the points are above the line at the beginning, but end up below the line near the end. The points do fit closer to the line than for the minimum temperatures, though, again confirming the improvement.

The R^2 for this model is 0.9116, which is significantly higher than that of the minimum temperature model. This means that 91.16% of the variability in maximum temperatures is explained by our model, which is pretty good.

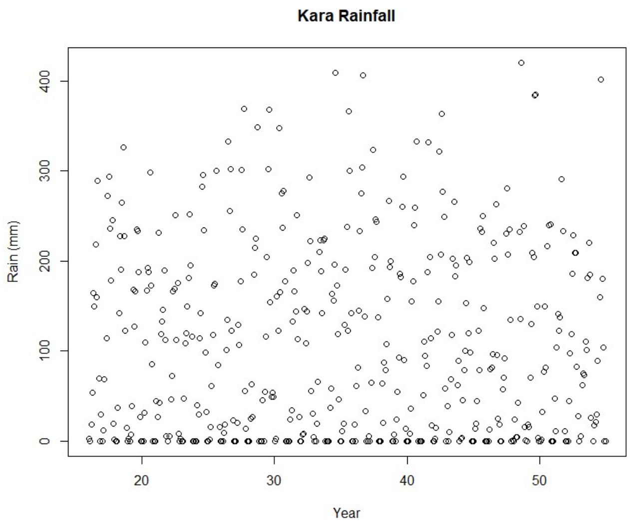


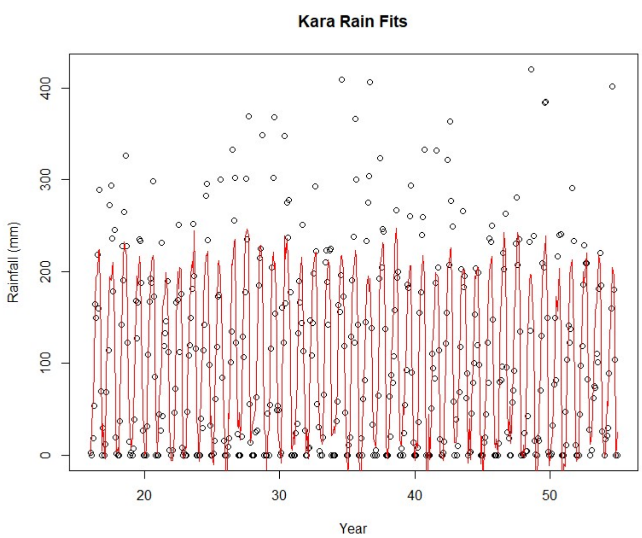

QQ-Plot for Max Temps
maxfits
30
32
34
36
28
30
32
34
36
38
karamax$MaxTemp

Rainfall Model

$R^2 = 0.7742$

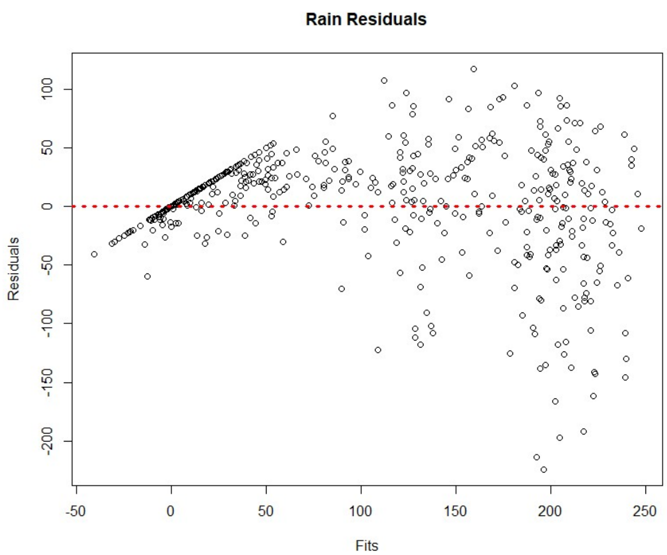
The rainfall data looks extremely scattered without much of a pattern. There are also a lot of zero values, which has a large influence on the final model. As can be seen in the plot of the model against the original data, we can see that many of the rainfall values are not captured by the model, which is a problem.

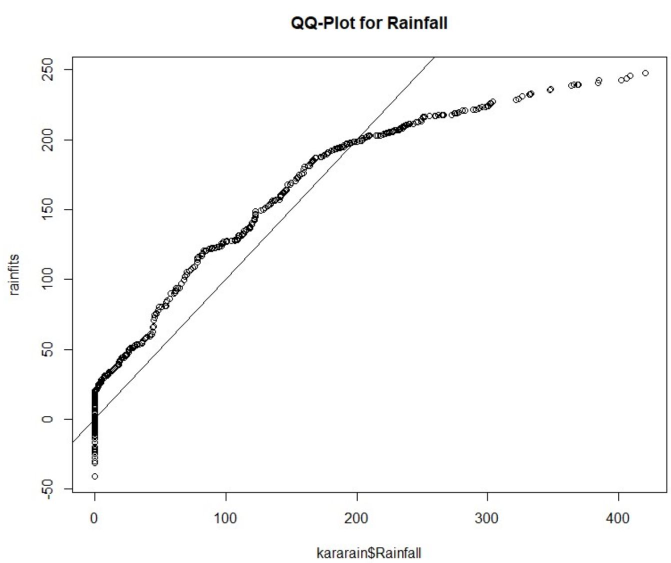




Both the residual plot and the qq-plot for the rainfall model are worse than both of the temperature models. This can most likely be attributed to all of the zero values in the rainfall data. By looking at both of these plots, there are very obviously some problems with normality and non-constant variance. This makes the validity of the final model questionable.

The value of R^2 is 0.7742, which is in between that of the minimum temperatures and the maximum temperatures. This means that a decent amount of the variation in the rainfall data is captured by our model, but it definitely could use some improvement, especially when it comes to the higher rainfall values.





One-Paragraph Summary

The general models for minimum temperatures, maximum temperatures, and rainfall were applied to the Kara data by itself. After examining the R^2 values and evaluating normality assumptions, it can be concluded that adjustments may need to be made to the model. There is evidence that the residuals are not normally distributed, especially for the minimum temperature and rainfall models. The maximum temperature model was the best of the three, with an R^2 value of 0.9116, indicating that 91.116% of the variation in the maximum temperature data can be explained by our model. Similarly, the R^2 values for the minimum temperatures and the rainfall were 0.7237 and 0.7742, respectively. The rainfall model is the most problematic, as many of the values were zero. This caused the model to underestimate many of the values, being unable to capture some of the higher amounts of rainfall. Improvements may be able to be made to these models as more factors that contribute to temperature and rainfall are discovered.