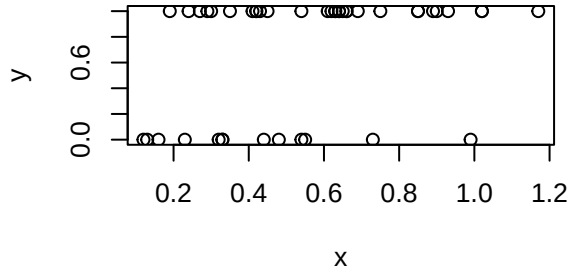
Assignment One - Emily Gill, Amanda Tsui, Mehaer Chhabda, Vaiswi Patel

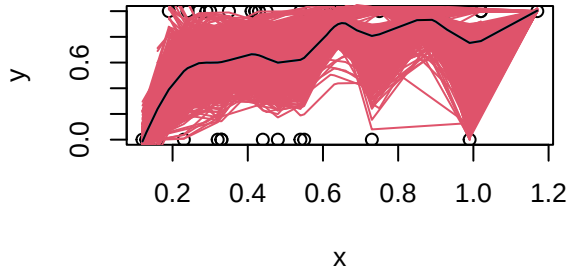
2024-09-25

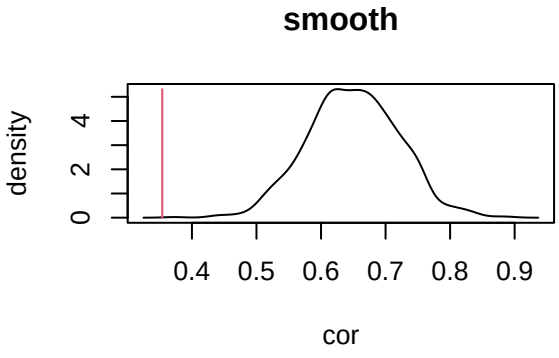
delta.temp and Drought.Count

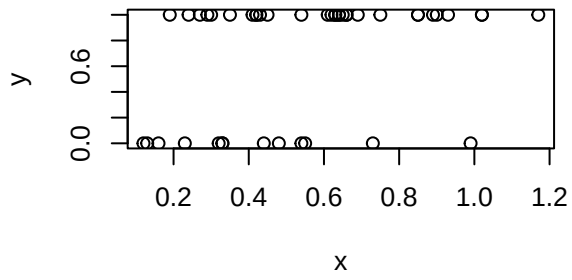
The two variables are best modeled through a distribution. The plots do not represent a function because the observed data is lacking a clear, consistent pattern in the bootmatsmooth plot. It does not follow a straight line, and instead varies. The red shading in the bootmatsmooth plot shows the intense amounts of variability, thus a lack of a clear relationship between the variables.

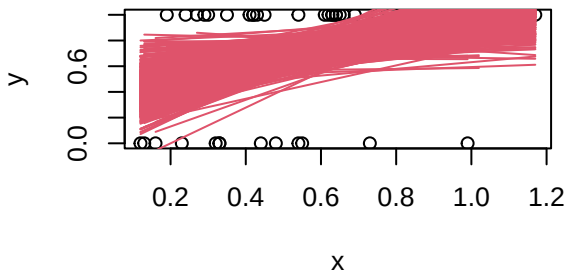
```
temp_drought_smooth <- bootmatsmooth(NOAAGISSWD$delta.temp,NOAAGISSWD$Drought.Count)
temp_drought_lin <- bootmatlin(NOAAGISSWD$delta.temp,NOAAGISSWD$Drought.Count)
```

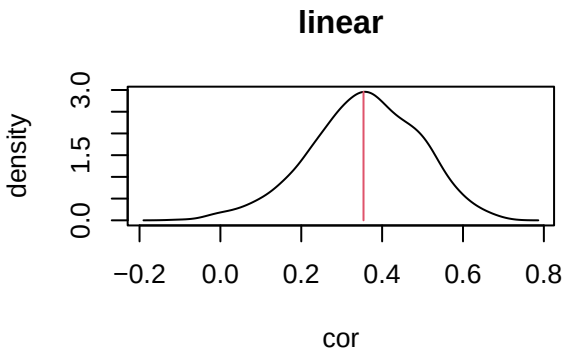








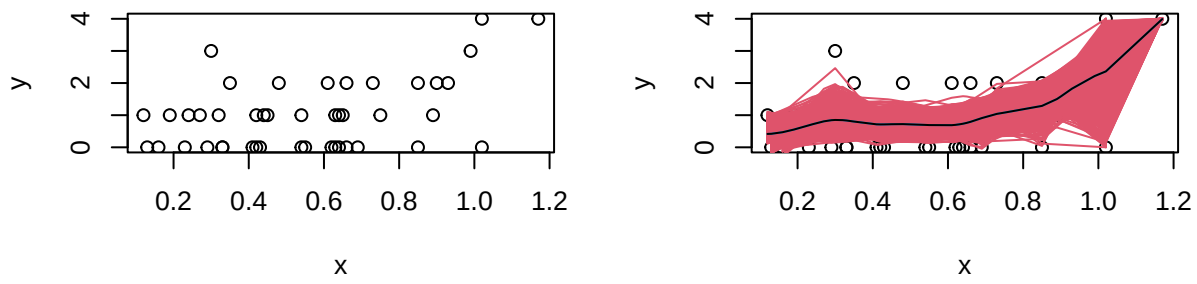


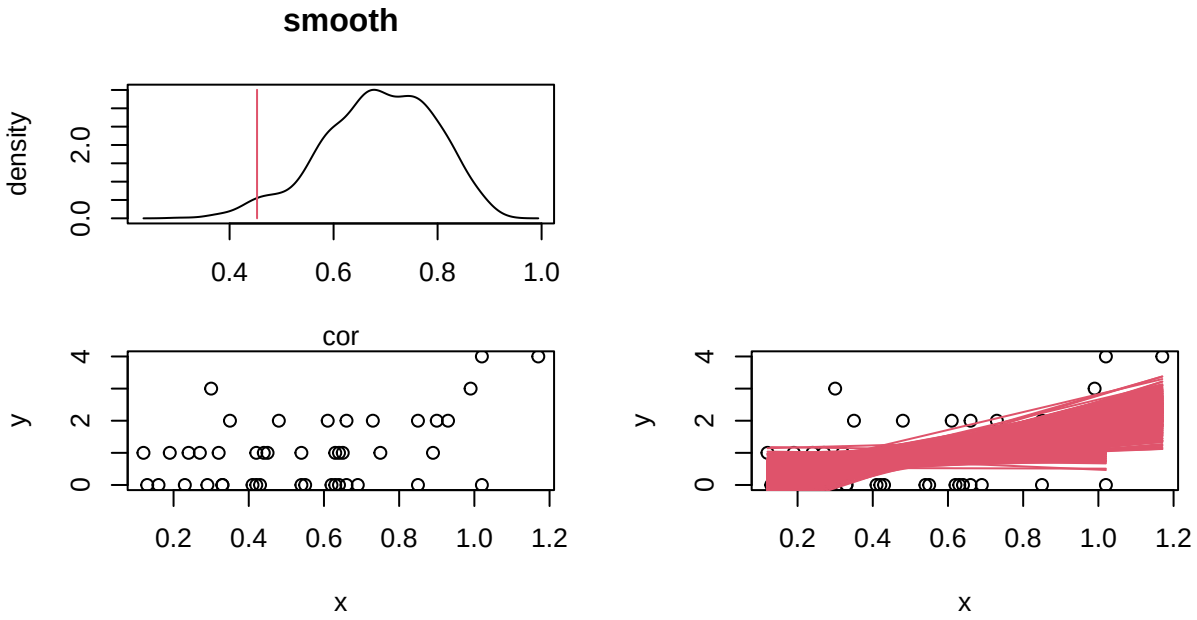


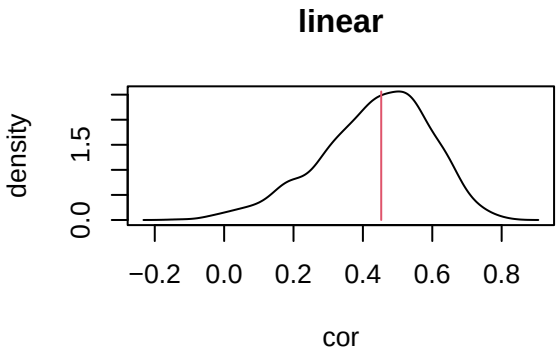
delta.temp and Flooding.Count

The two variables are best modeled through a distribution. The plots do not represent a function because the observed data is lacking a clear, consistent pattern in the bootmatsmooth plot. Though the scatter plots may seem to follow an upwards trend, this is not the case, as the red shading in the bootmatsmooth plot shows a great amount of variability between the variables. Thus, a clear relationship is still lacking.

```
temp_flood_smooth <- bootmatsmooth(NOAAGISSWD$delta.temp,NOAAGISSWD$Flooding.Count)
temp_flood_lin <- bootmatlin(NOAAGISSWD$delta.temp,NOAAGISSWD$Flooding.Count)
```



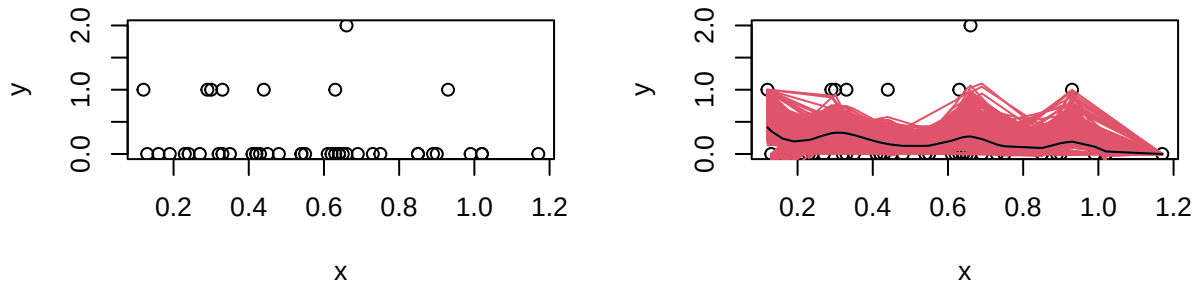


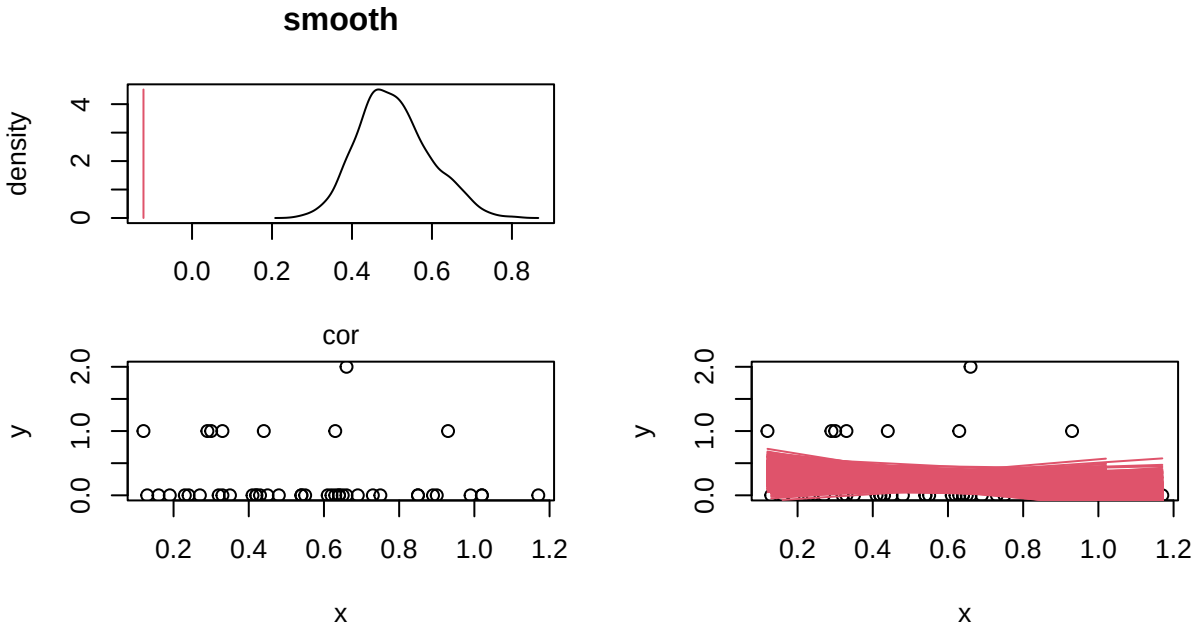


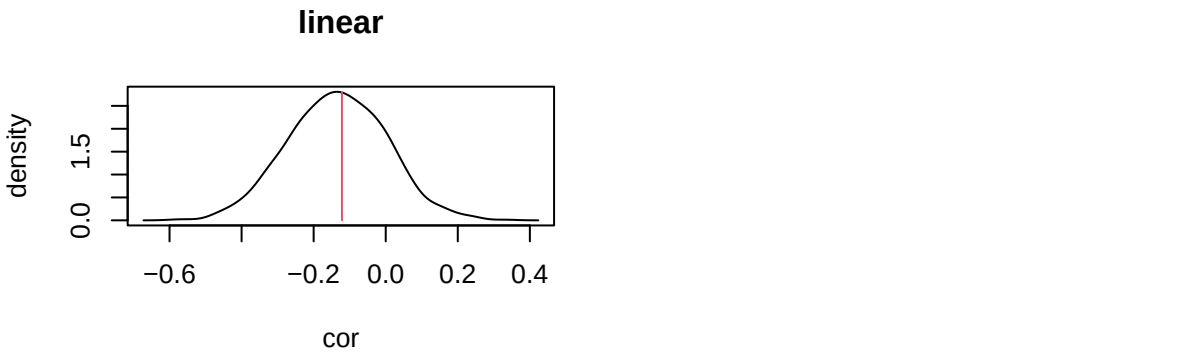
delta.temp and Freeze.Count

The two variables are best modeled through a distribution. There are many outliers in the smoothing plot, which proves that there is a lack of a clear pattern and relationship between the variables. The red shading in the bootmatsmooth plot does not even reach the largest outlier, showcasing why this is a distribution.

```
temp_freeze_smooth <- bootmatsmooth(NOAAGISSWD$delta.temp,NOAAGISSWD$Freeze.Count)
temp_freeze_lin <- bootmatlin(NOAAGISSWD$delta.temp,NOAAGISSWD$Freeze.Count)
```



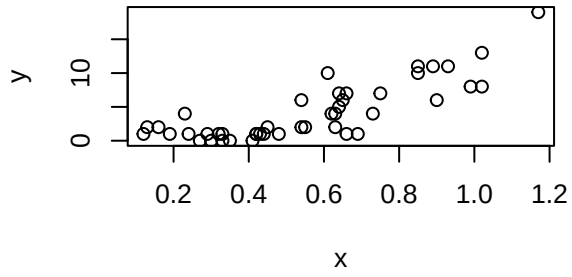


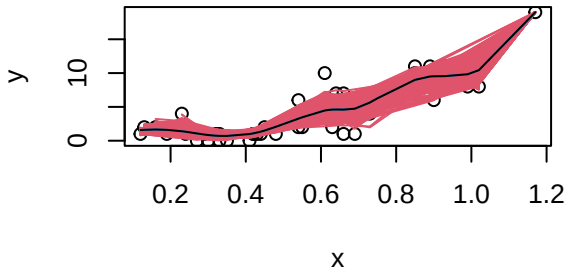


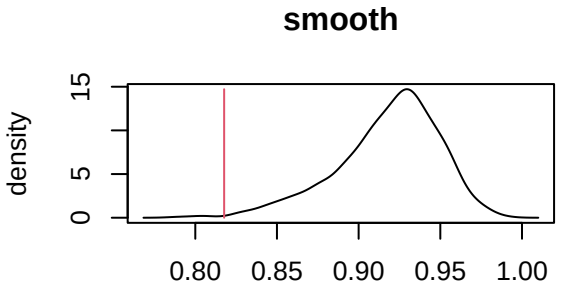
delta.temp and Severe.Storm.Count

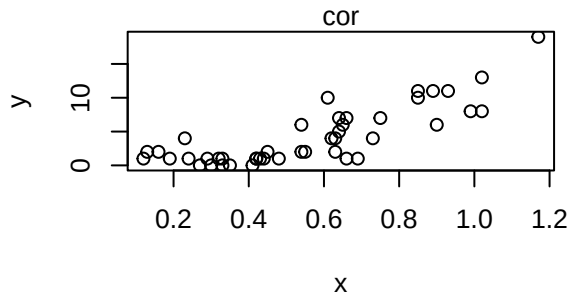
The variables are best showcased through a non-linear function. It is a function because the bootmatsmooth plot showcases a general pattern that trends upwards, however, too much of the data falls out of the red shading, making it best described as a non-linear function. The red shading is very skinny, indicating a relationship between the variables.

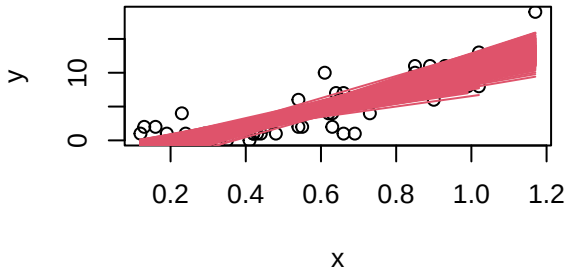
```
temp_severe_smooth <- bootmatsmooth(NOAAGISSWD$delta.temp,NOAAGISSWD$Severe.Storm.Count)
temp_severe_lin <- bootmatlin(NOAAGISSWD$delta.temp,NOAAGISSWD$Severe.Storm.Count)
```

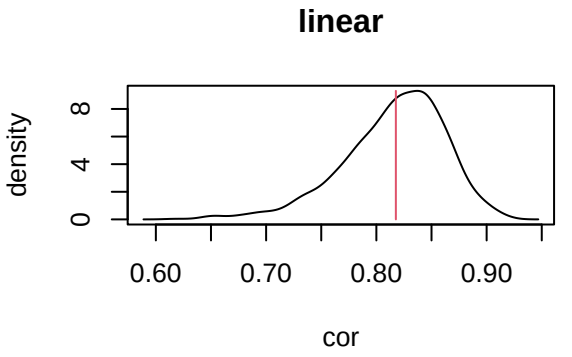








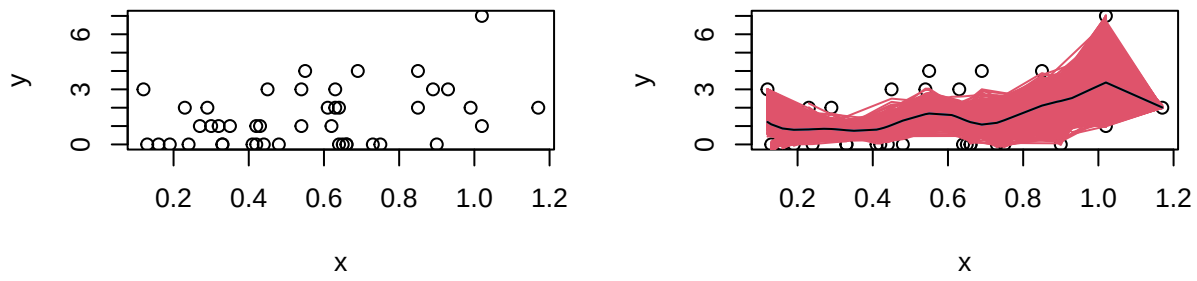


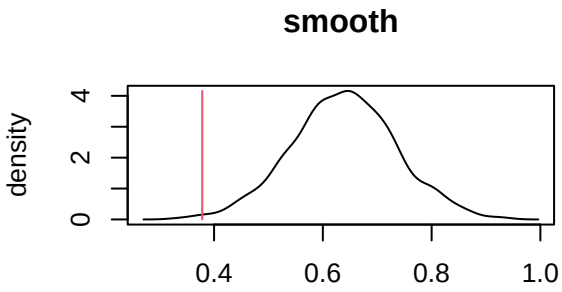


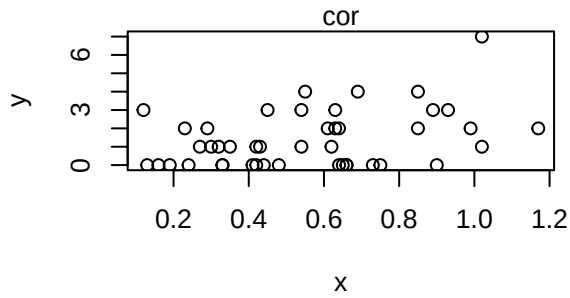
delta.temp and Tropical.Cyclone.Count

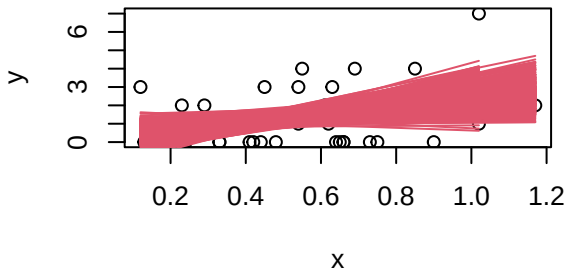
The variables are best showcased as a distribution, as the red shading in the bootmatsmooth plot varies in thickness, representing greater variability. There are also outliers that indicate a lack of a consistent pattern between the variables. The scatterplot indicates a distribution of data, rather then a trend or pattern.

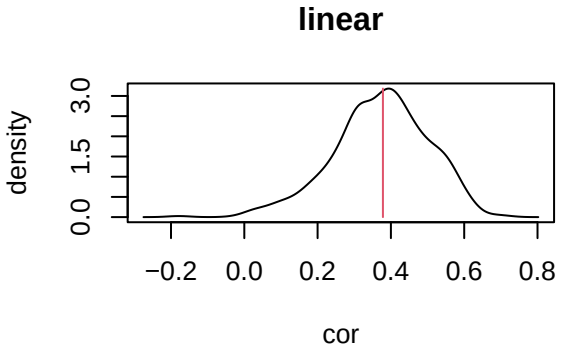
```
temp_cyclone_smooth <- bootmatsmooth(NOAAGISSWD$delta.temp,NOAAGISSWD$Tropical.Cyclone.Count)
temp_cyclone_lin <- bootmatlin(NOAAGISSWD$delta.temp,NOAAGISSWD$Tropical.Cyclone.Count)
```







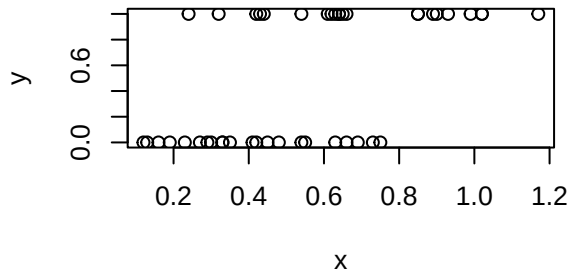


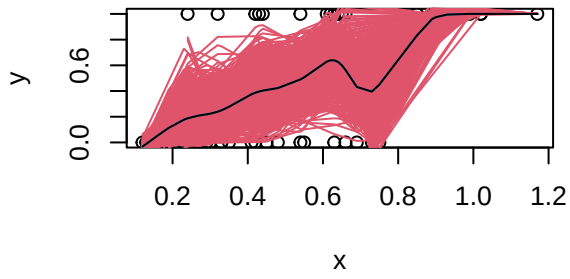


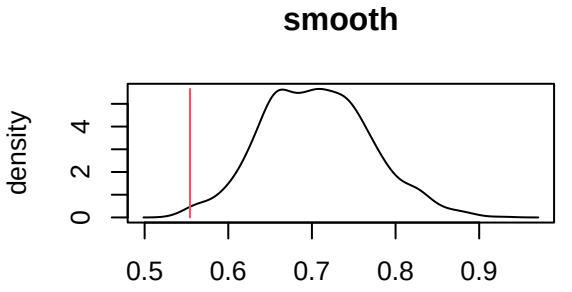
delta.temp and Wildfire.Count

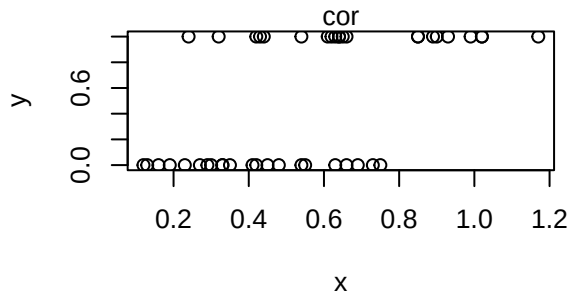
The variables are best modeled with a distribution as the Y values of the data points are either 0 or 1. Additionally, all points fall outside the red region when the Bootstraplin function is used. The bootmatsmooth plot also showcases red shading that is very wide and does not encapsulate all of the data points.

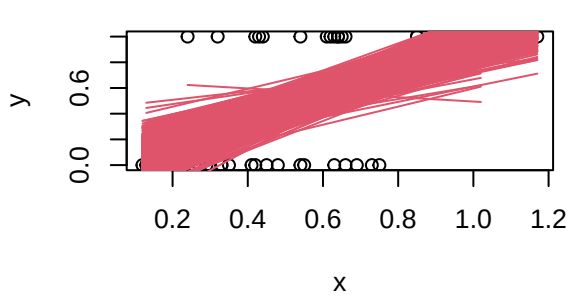
```
temp_wildfire_smooth <- bootmatsmooth(NOAAGISSWD$delta.temp,NOAAGISSWD$Wildfire.Count)
temp_wildfire_lin <- bootmatlin(NOAAGISSWD$delta.temp,NOAAGISSWD$Wildfire.Count)
```

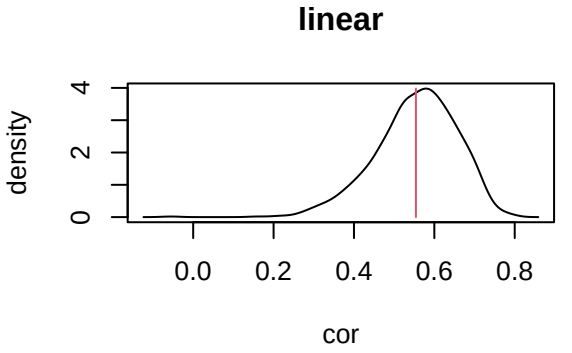








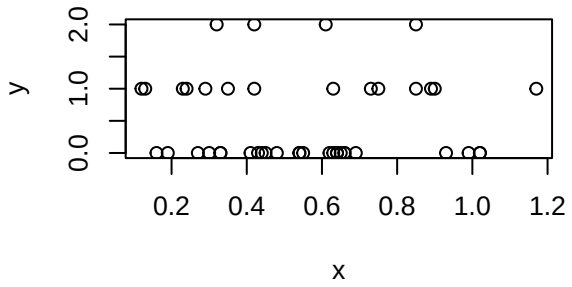


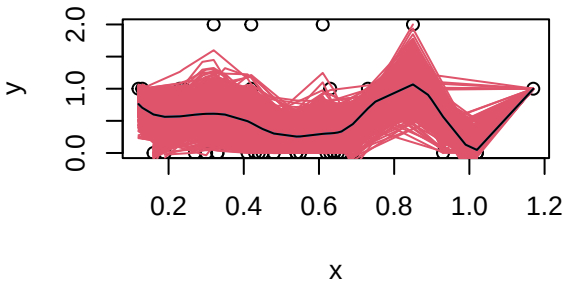


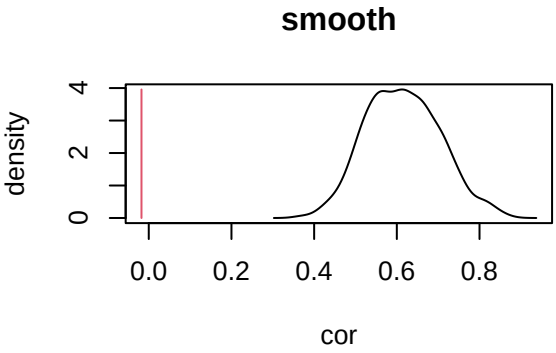
delta.temp and Winter.Storm.Count

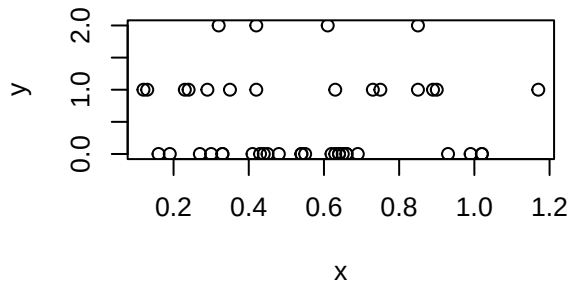
The two variables are best modeled with a distribution function. There seems to be some overlap at the 0.4 value on the x axis. The red shading in the bootmatsmooth plot does not encapusalte all of the data. There are too many outliers in varying positions to be able to consider this a function. There is too much variability between the data to indicate a clear pattern.

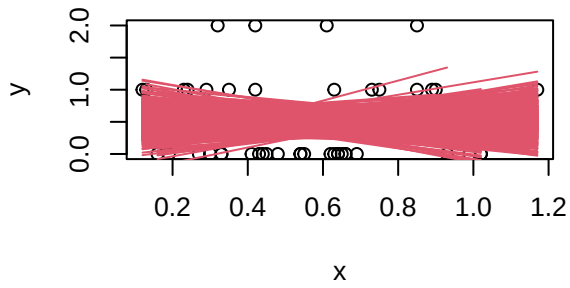
```
temp_winter_smooth <- bootmatsmooth(NOAAGISSWD$delta.temp,NOAAGISSWD$Winter.Storm.Count)
temp_winter_lin <- bootmatlin(NOAAGISSWD$delta.temp,NOAAGISSWD$Winter.Storm.Count)
```

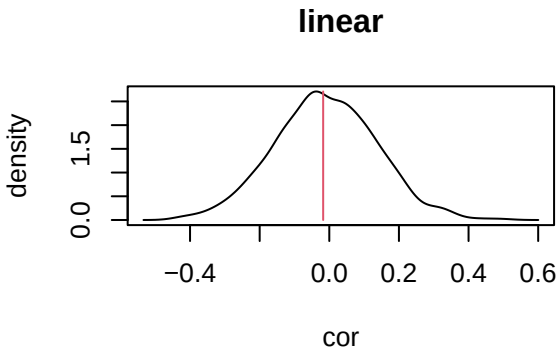








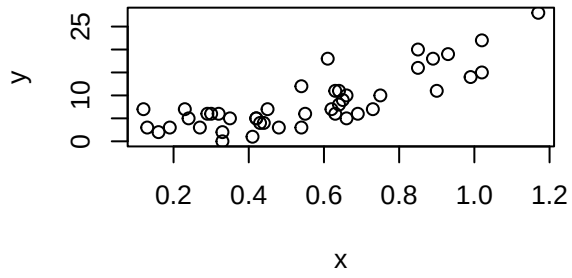


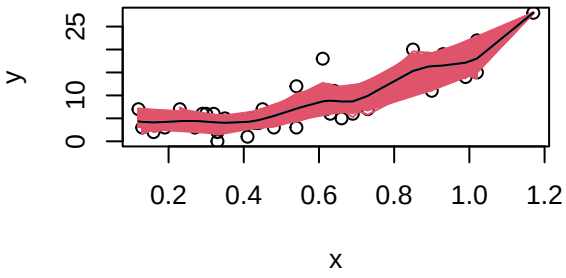


delta.temp and All.Disasters.Count

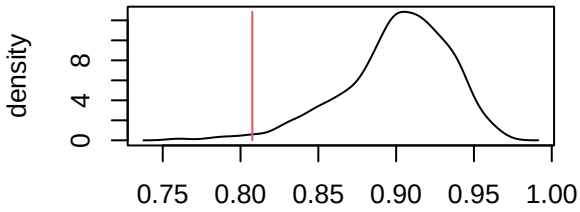
The two variables are best modeled with a non-linear function. Although this function may look like a linear function because of its higher correlation coefficient, which is above 0.8, it is in fact non-linear because the red region in Bootmatsmooth strays away from the black line quite a bit. Additionally, many points lie outside the red region in Bootmatlin, therefore, it would be wrong to assume that this function is linear.

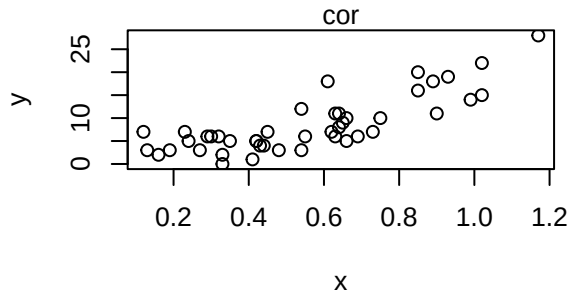
```
temp_all_smooth <- bootmatsmooth(NOAAGISSWD$delta.temp,NOAAGISSWD$All.Disasters.Count)
temp_all_lin <- bootmatlin(NOAAGISSWD$delta.temp,NOAAGISSWD$All.Disasters.Count)
```

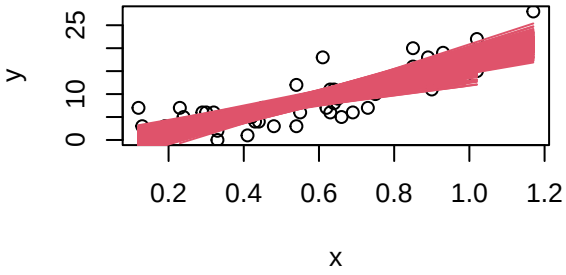




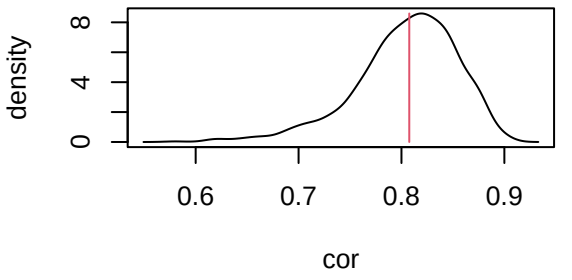
smooth







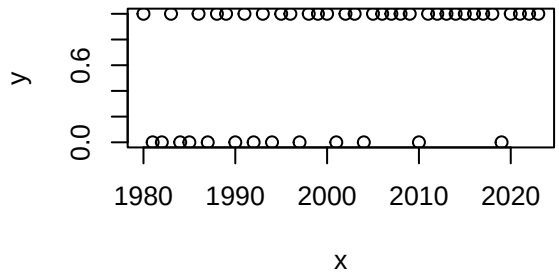
linear

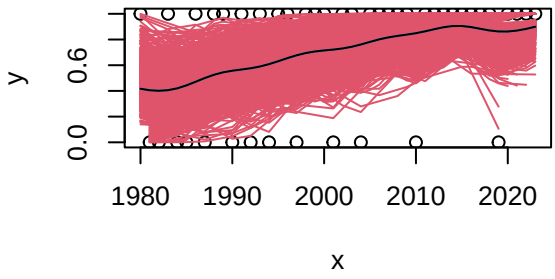


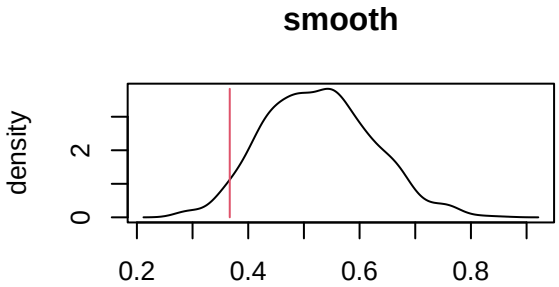
Year and Drought.Count

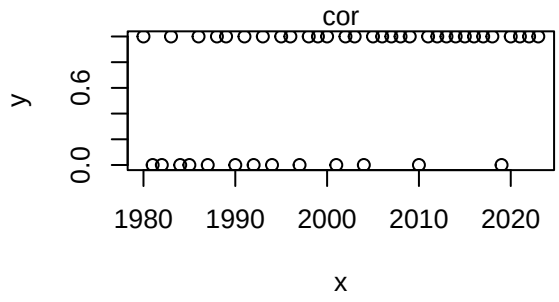
The two variables are best modeled with a distribution because the scatter plot shows that y has two values (0 and 1), indicating that it is not a linear model. The red shading in the bootmatsmooth plot does not significantly touch the data points and is very wide, indicating great amounts of variability. Only very few data points are touched by the red shading in the bootmatlin plot, indicating a lack of relationship between the variables.

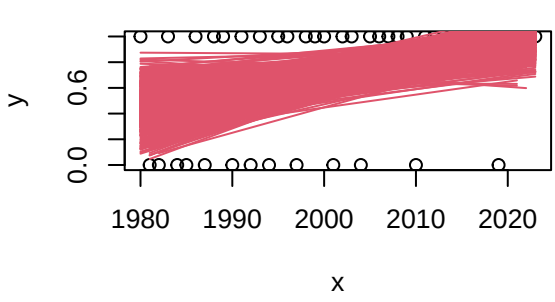
```
year_drought_smooth <- bootmatsmooth(NOAAGISSWD$Year,NOAAGISSWD$Drought.Count)
year_drought_lin <- bootmatlin(NOAAGISSWD$Year,NOAAGISSWD$Drought.Count)
```

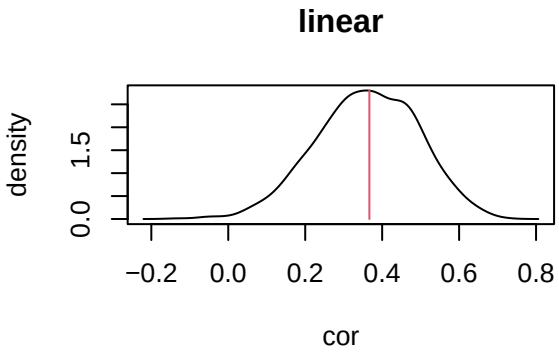








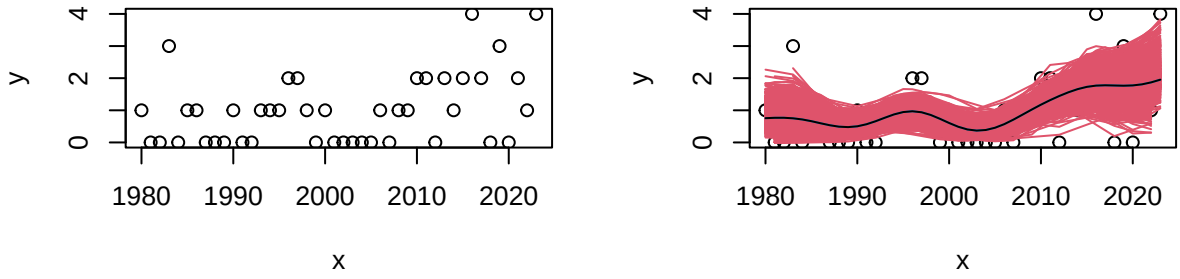


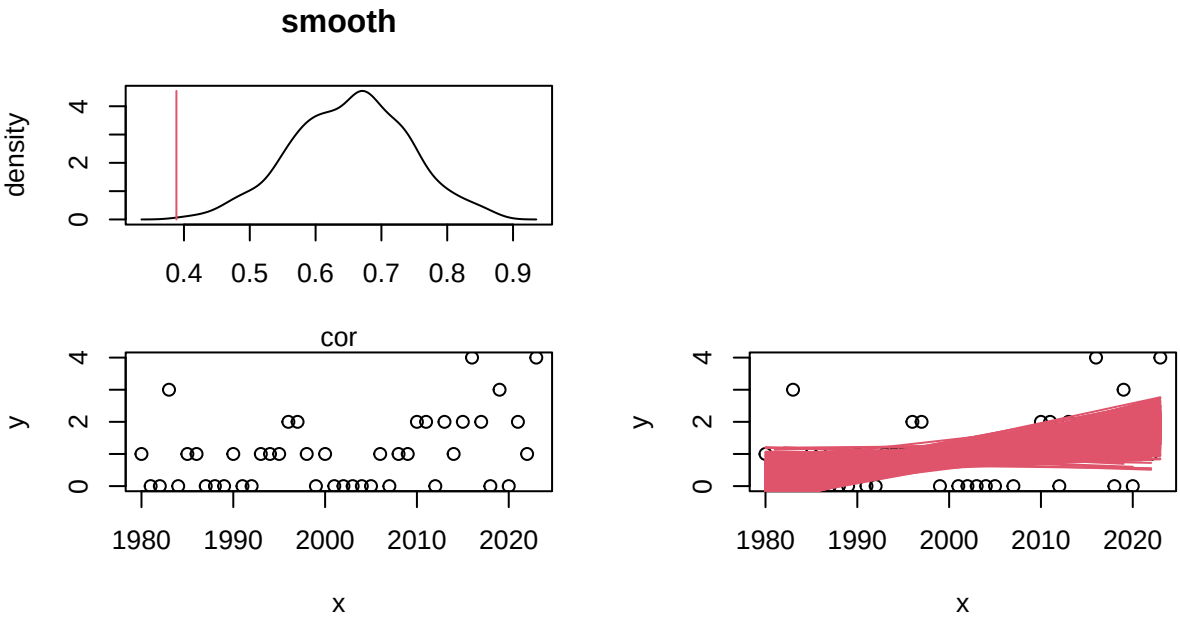


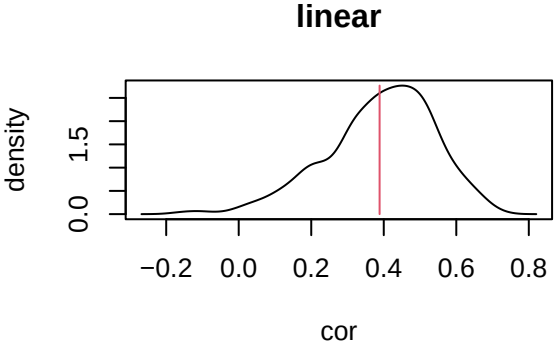
Year and Flooding.Count

The two variables are best modeled with a non-linear function. The smoothing plot correlation offers a better fit than the linear one, reinforcing its non-linear nature. There appears to be a relationship between the variables that slightly trends upwards, however, it is not a linear relationship.

```
year_flood_smooth <- bootmatsmooth(NOAAGISSWD$Year,NOAAGISSWD$Flooding.Count)
year_flood_lin <- bootmatlin(NOAAGISSWD$Year,NOAAGISSWD$Flooding.Count)
```



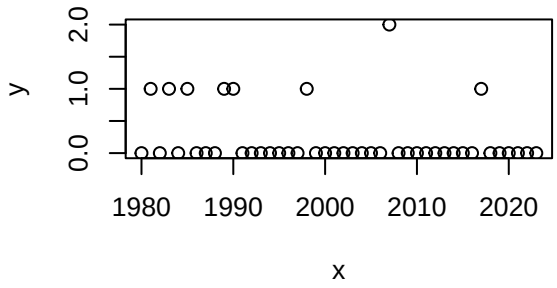


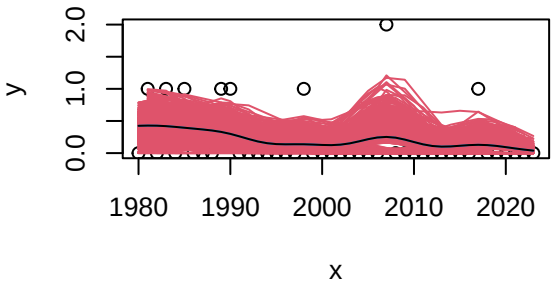


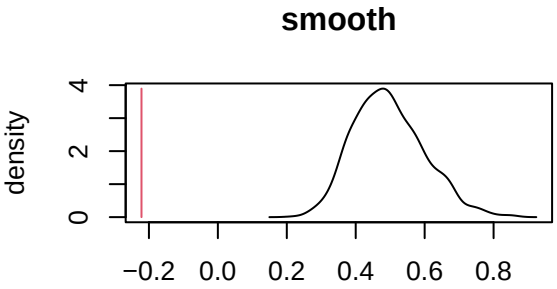
Year and Freeze.Count

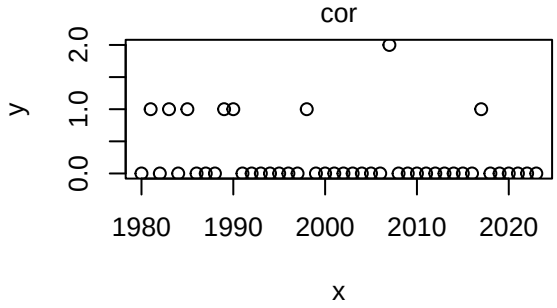
The two variables are best modeled with a distribution as the y values consist mostly of ones and zeroes, with one outlier. The plots display multiple y values for each x, which means it does not meet the criteria of a function. Although the smooth curve fits the data better than the linear model, it still represents a correlation rather than a true function.

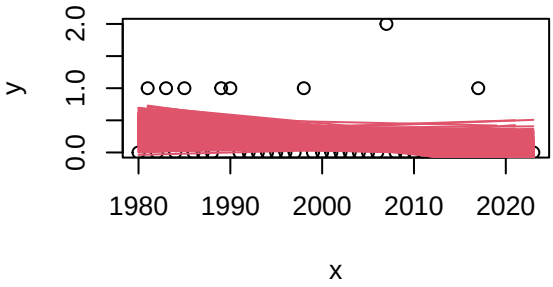
```
year_freeze_smooth <- bootmatsmooth(NOAAGISSWD$Year,NOAAGISSWD$Freeze.Count)
year_freeze_lin <- bootmatlin(NOAAGISSWD$Year,NOAAGISSWD$Freeze.Count)
```

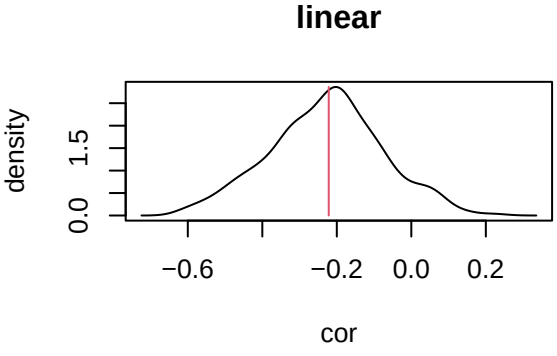








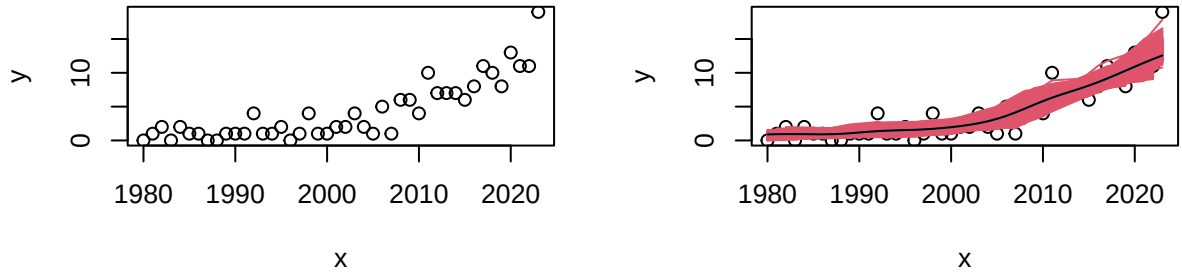


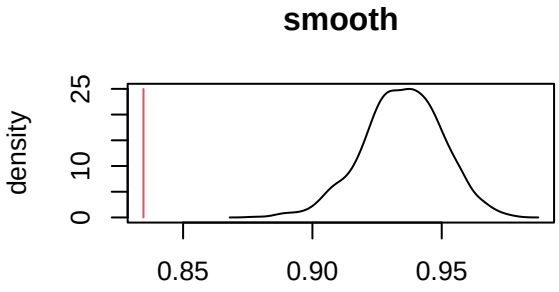


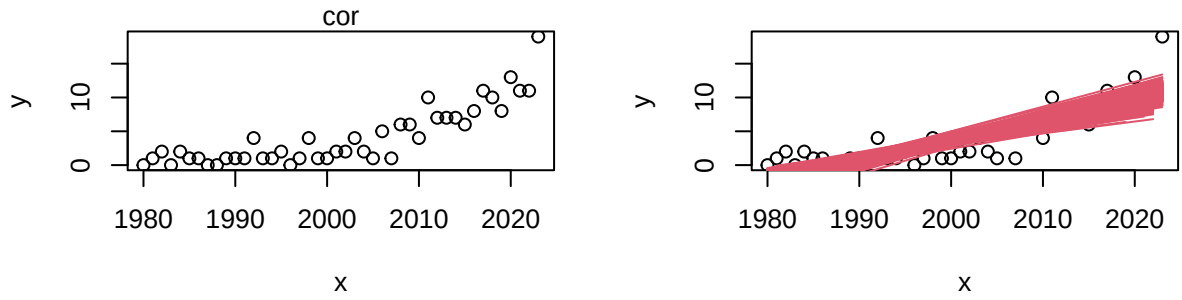
Year and Severe.Storm.Count

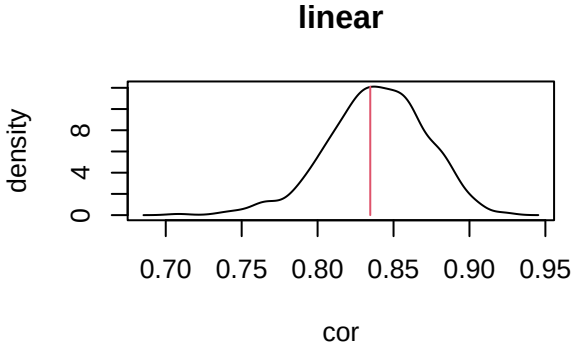
The two variables are best modeled with a non-linear function. The red shading in the bootmatsmooth plot does a better job at encapsulating the most data points comapred to the bootmatlin plot. It also has a skinnier red shading region, indicating that it does a better job at showcasing the variability and correlation between the data.

```
year_severe_smooth <- bootmatsmooth(NOAAGISSWD$Year,NOAAGISSWD$Severe.Storm.Count)
year_severe_lin <- bootmatlin(NOAAGISSWD$Year,NOAAGISSWD$Severe.Storm.Count)
```





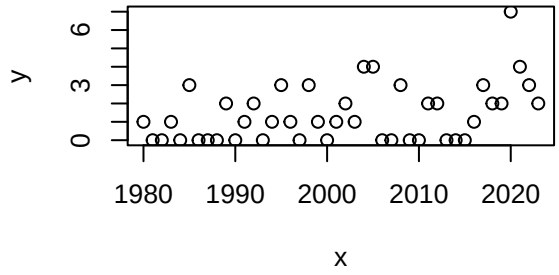


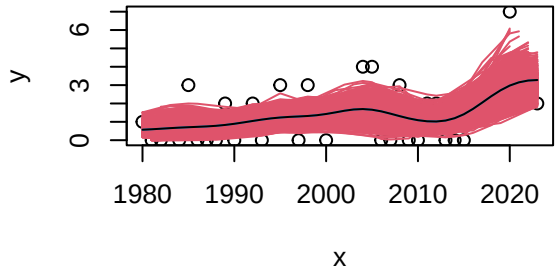


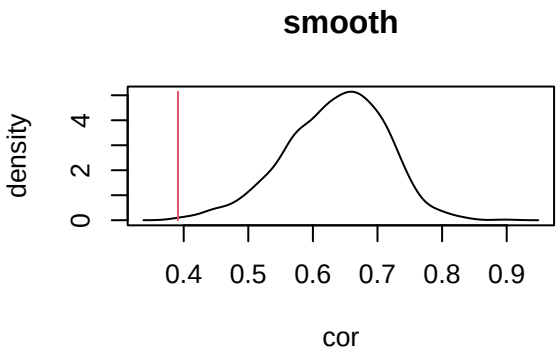
Year and Tropical.Cyclone.Count

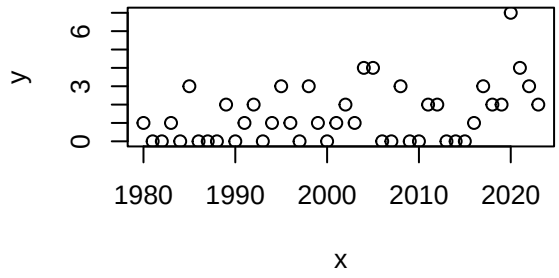
The two variables are best modeled by a non-linear function. Despite the fact that the scatter plot appears to be a linear function, as demonstrated by the minimally upward slope, the majority of the data points within the bootmatlin graph fall outside of the red shaded area. Therefore, it would be better portrayed by the smooth fit. We can see this demonstrated through the smooth fit peaking between 0.6 and 0.7, indicating that a nonlinear function better suits the data points.

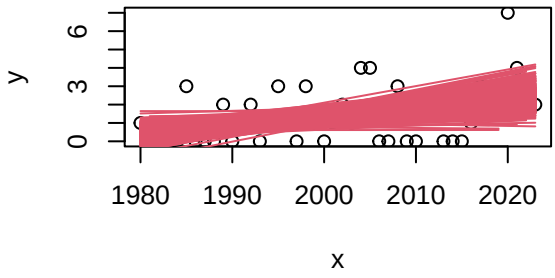
```
year_cyclone_smooth <- bootmatsmooth(NOAAGISSWD$Year,NOAAGISSWD$Tropical.Cyclone.Count)
year_cyclone_lin <- bootmatlin(NOAAGISSWD$Year,NOAAGISSWD$Tropical.Cyclone.Count)
```

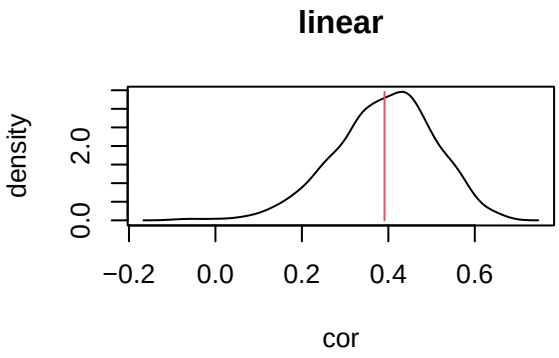








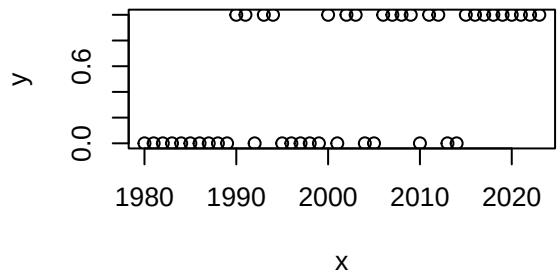


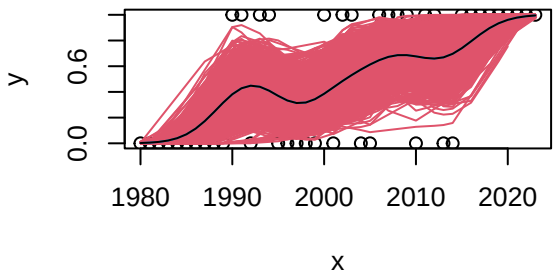


Year and Wild.Fire.Count

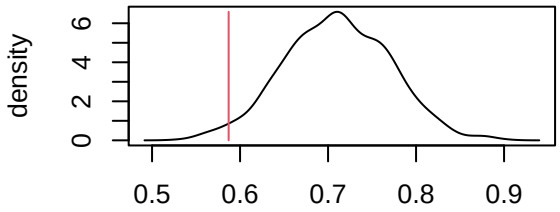
The two variables are best modeled by a distribution because the majority of the data points are either zero or one. Furthermore, a majority of the data points fall outside of the red shaded area in both the bootmatlin and bootmatsmooth plots, thus indicating that this is a distribution. The red shading is extremly wide, making it clear that there is no pattern between the variables and data, thus showcasing a lack of a relationship.

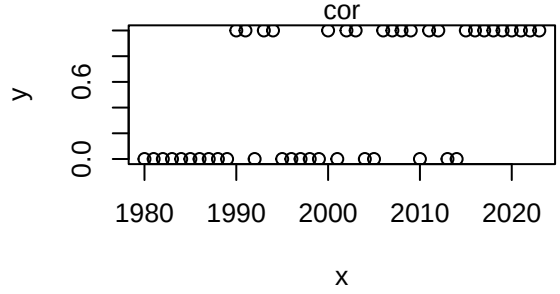
```
year_wildfire_smooth <- bootmatsmooth(NOAAGISSWD$Year,NOAAGISSWD$Wildfire.Count)
year_wildfire_lin <- bootmatlin(NOAAGISSWD$Year,NOAAGISSWD$Wildfire.Count)
```

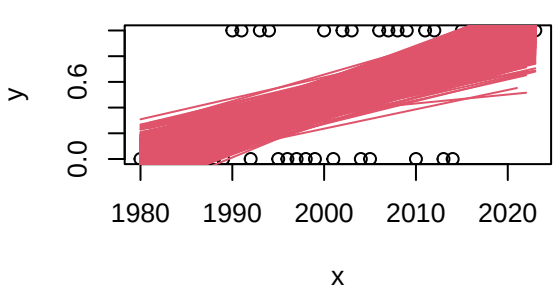




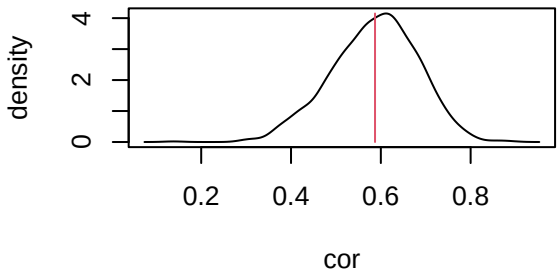
smooth







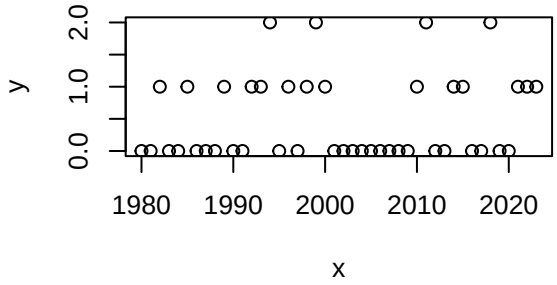
linear

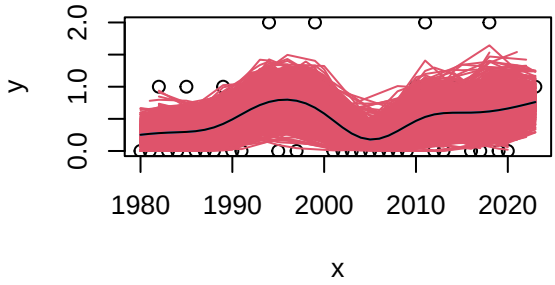


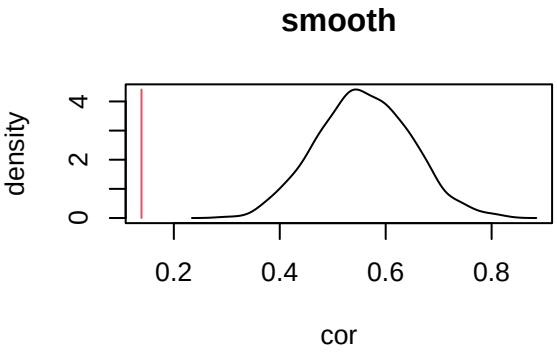
Year and Winter.Storm.Count

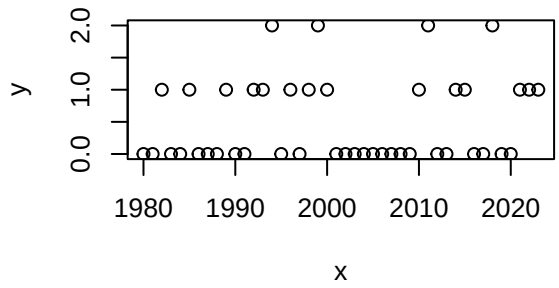
The two variables are best modeled with a distribution as there are at most two values, with 4 outliers. Furthermore, nearly all of the data points fall outside of the red shaded region in the bootmatlin and bootmatsmooth plots and do not follow a clear pattern. The red shading is very wide in the bootmatsmooth plot, showcasing that there is great variability between the variables and no clear relationship.

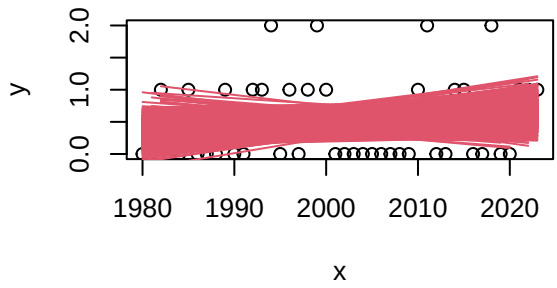
```
year_winter_smooth <- bootmatsmooth(NOAAGISSWD$Year,NOAAGISSWD$Winter.Storm.Count)
year_winter_lin <- bootmatlin(NOAAGISSWD$Year,NOAAGISSWD$Winter.Storm.Count)
```

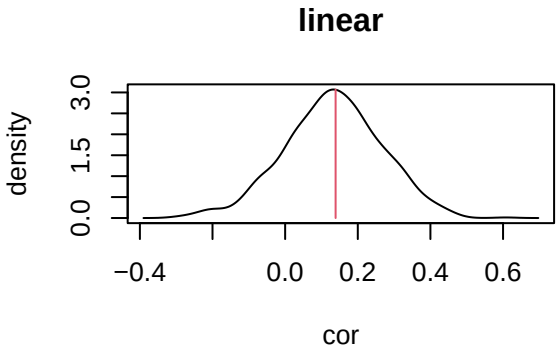








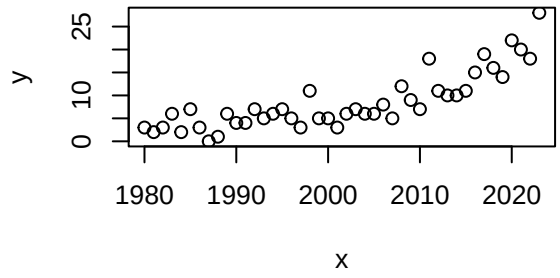


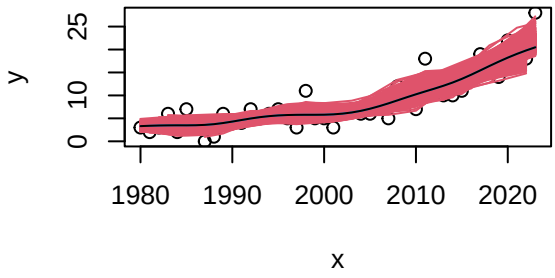


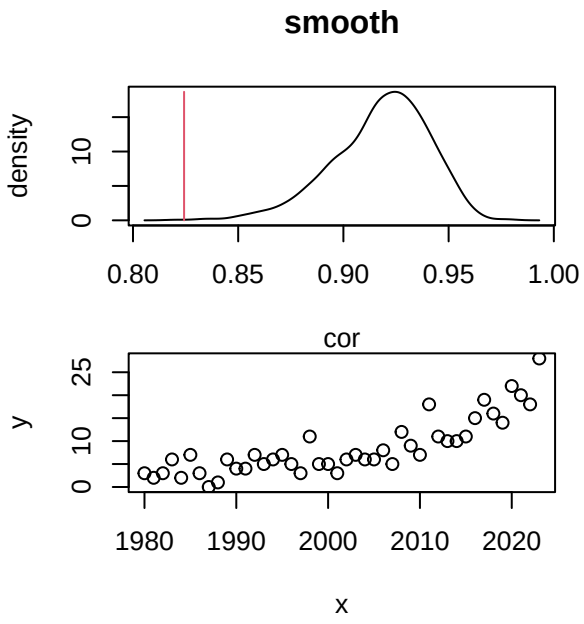
Year and All.Disasters.Count

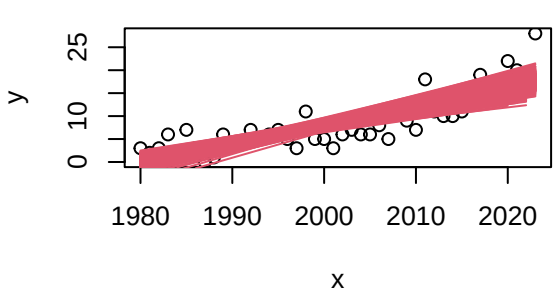
The variables are best modeled with a non-linear function. While it appears to look like a linear function, given by the straight line in the bootmatlin graph, upon further inspection, there is a slight curve in the bootmatsmooth graph, making this graph nonlinear. Furthermore, the smooth fit has a higher correlation, indicating that the relationship is more complex than a linear function.

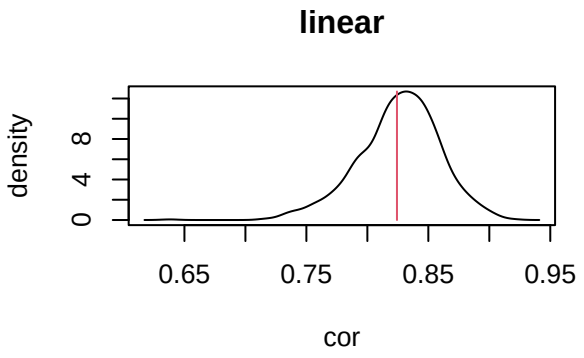
```
year_all_smooth <- bootmatsmooth(NOAAGISSWD$Year,NOAAGISSWD$All.Disasters.Count)
year_all_lin <- bootmatlin(NOAAGISSWD$Year,NOAAGISSWD$All.Disasters.Count)
```


y
0
10
25
1980
1990
2000
2010
2020
x


y
0
10
25
1980
1990
2000
2010
2020
x


smooth
density
0
10
0.80
0.85
0.90
0.95
1.00
cor
y
0
10
25
1980
1990
2000
2010
2020
x


y
0
10
25
1980
1990
2000
2010
2020
x


linear
density
0
4
8
0.65
0.75
0.85
0.95
cor