Temperature and Precipitation Changes in Southeast Texas

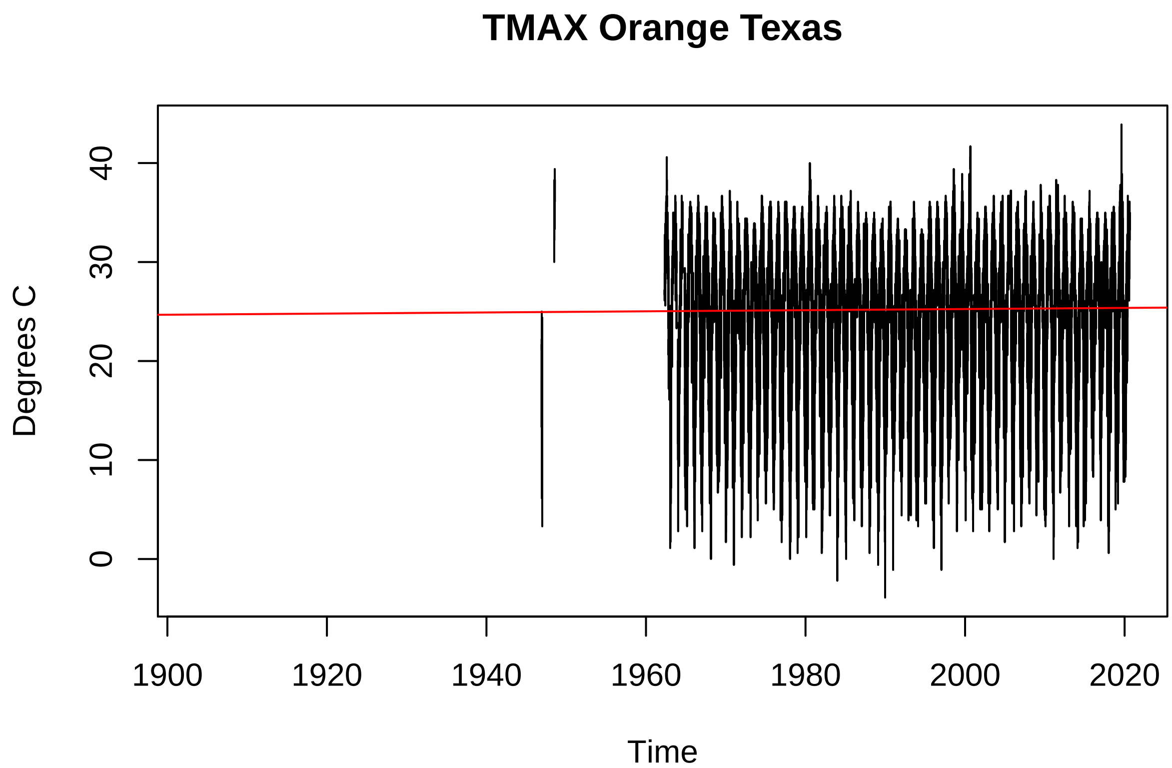
Katy Swiere

9/19/20

In this blog I will be exploring climate change in Southeast Texas, using data from the National Oceanic and Atmospheric Association (NOAA). I'm using Orange, Texas data which is the place I grew up.

First I need to import my data. I do this, and then choose the variable TMAX, which is maximum daily temperature, to focus on.

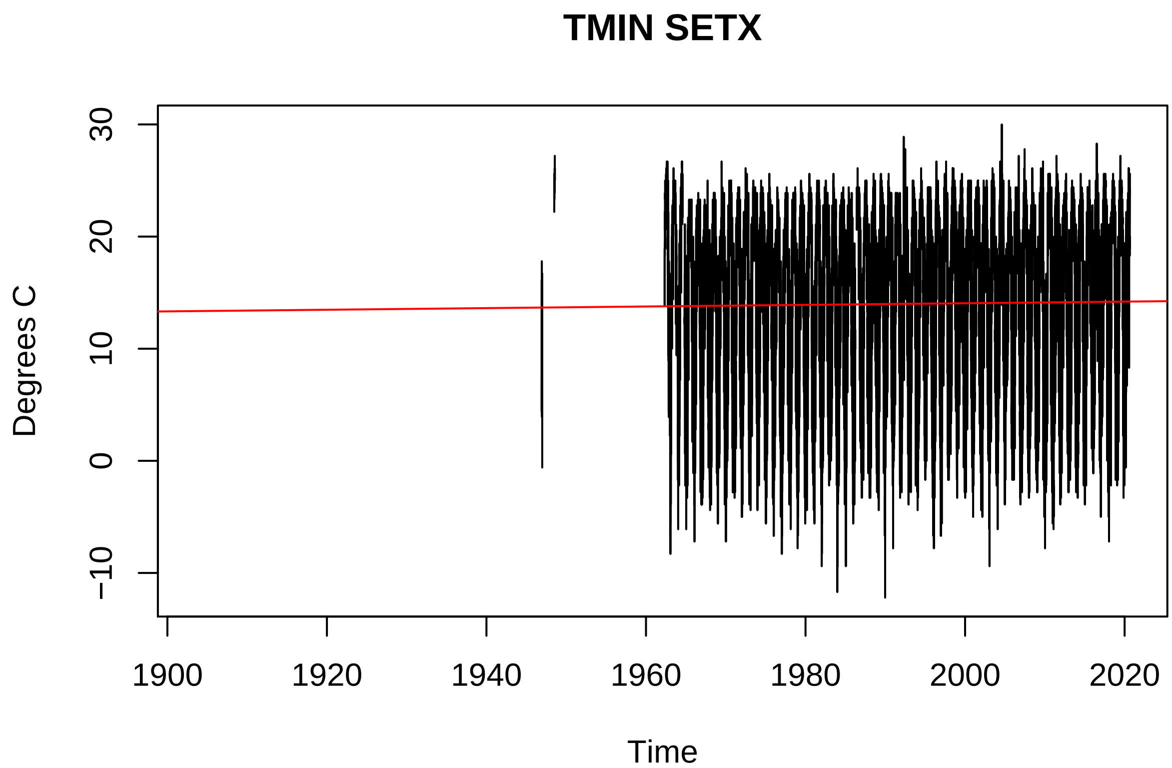
```
climate_data <- read.csv("/home/CAMPUS/kmsa2017/Climate_Change_Narratives/Student_Folders/Swiere/Swiere_


strDates <- as.character(climate_data$DATE) #create object dates as characters
climate_data$NewDate <- as.Date(strDates, "%m/%d/%Y") #turn it into New Date

plot(TMAX~NewDate, climate_data, main = "TMAX Orange Texas", ylab = "Degrees C", xlab = "Time", type =

bestfit <- lm(TMAX~NewDate, climate_data)
abline(coef(bestfit), col = "red")
```



Now I can get a general idea of what it looks like. Then I can start formatting things in useful ways and plotting the data. I've skipped over plotting individual months, since I will do all months together using for loops later.

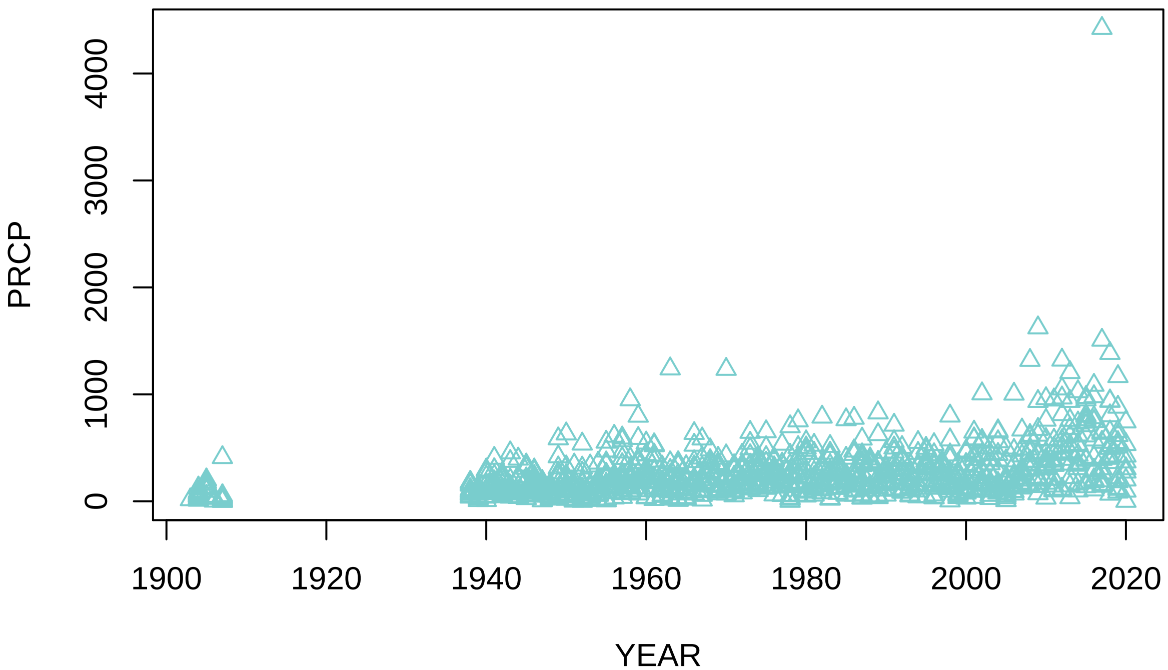
In this next chunk, I'll code new variables into useful formats so that we can use them in our for loops.

```
#tmax
climate_data$Month <- format(as.Date(climate_data$NewDate), format = "%m")
climate_data$Year <- format(as.Date(climate_data$NewDate), format = "%Y")

MonthlyTMAXMean = aggregate(TMAX~Month + Year, climate_data, mean)
MonthlyTMAXMean$YEAR <- as.numeric(MonthlyTMAXMean$Year)
MonthlyTMAXMean$MONTH <- as.numeric(MonthlyTMAXMean$Month)

##tmin
MonthlyTMINMean <- aggregate(TMIN~Month + Year, climate_data, mean)
MonthlyTMINMean$MONTH <- as.numeric(MonthlyTMINMean$Month)
MonthlyTMINMean$YEAR <- as.numeric(MonthlyTMINMean$Year)

#i'll also plot tmin since we did tmax earlier
plot(TMIN~NewDate, climate_data, main = "TMIN SETX", ylab = "Degrees C", xlab = "Time", type = "l")
lm.min <- lm(TMIN~NewDate, climate_data)
abline(lm.min, col = "red")
```



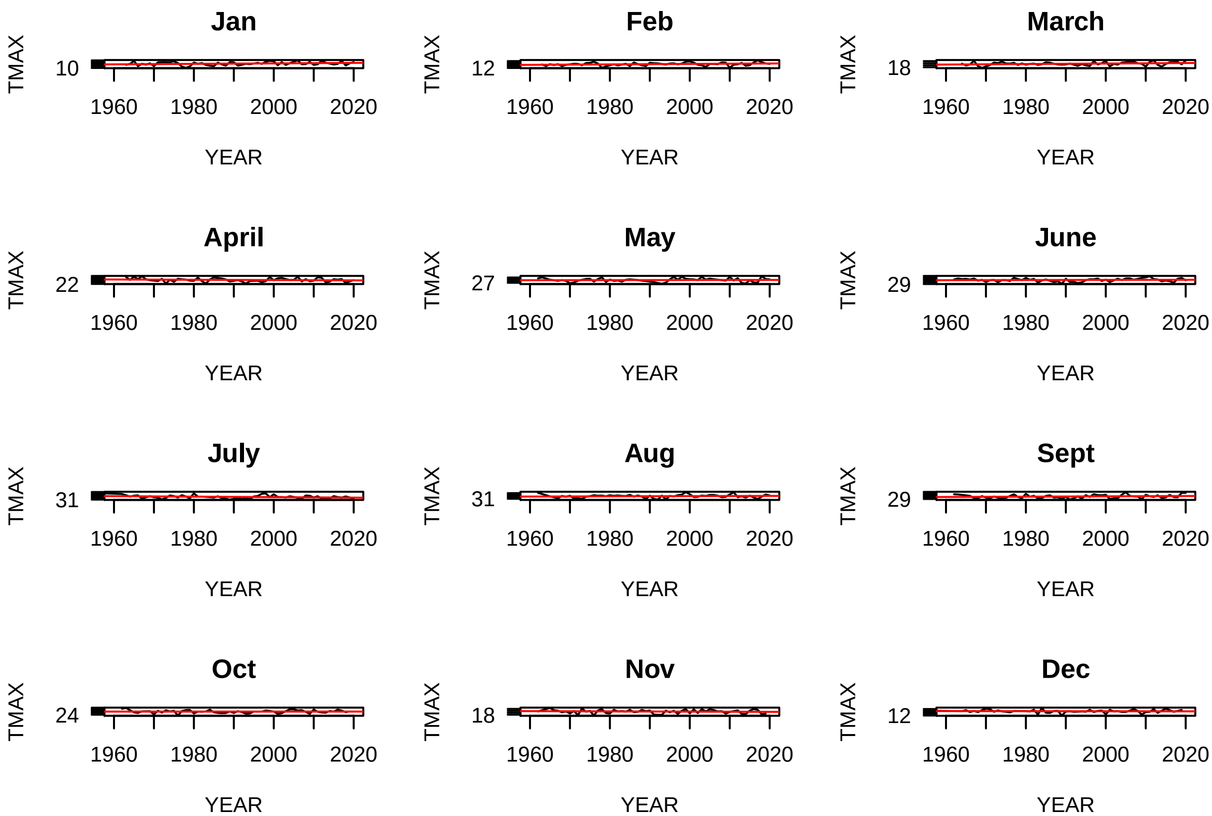
We haven't looked at precipitation, which, in my opinion, is the most interesting since floods are major weather events in the area and can cause high levels of damage!

```
###PRECIPITATION
MonthlyPRCPSum = aggregate(PRCP~ Month + Year, climate_data, sum)

MonthlyPRCPSum$YEAR <- as.numeric(MonthlyPRCPSum$Year)
MonthlyPRCPSum$MONTH <- as.numeric(MonthlyPRCPSum$Month)
plot(PRCP~YEAR, MonthlyPRCPSum, pch = 2, col = "darkslategray3") #does the triangle kind of remind you
```

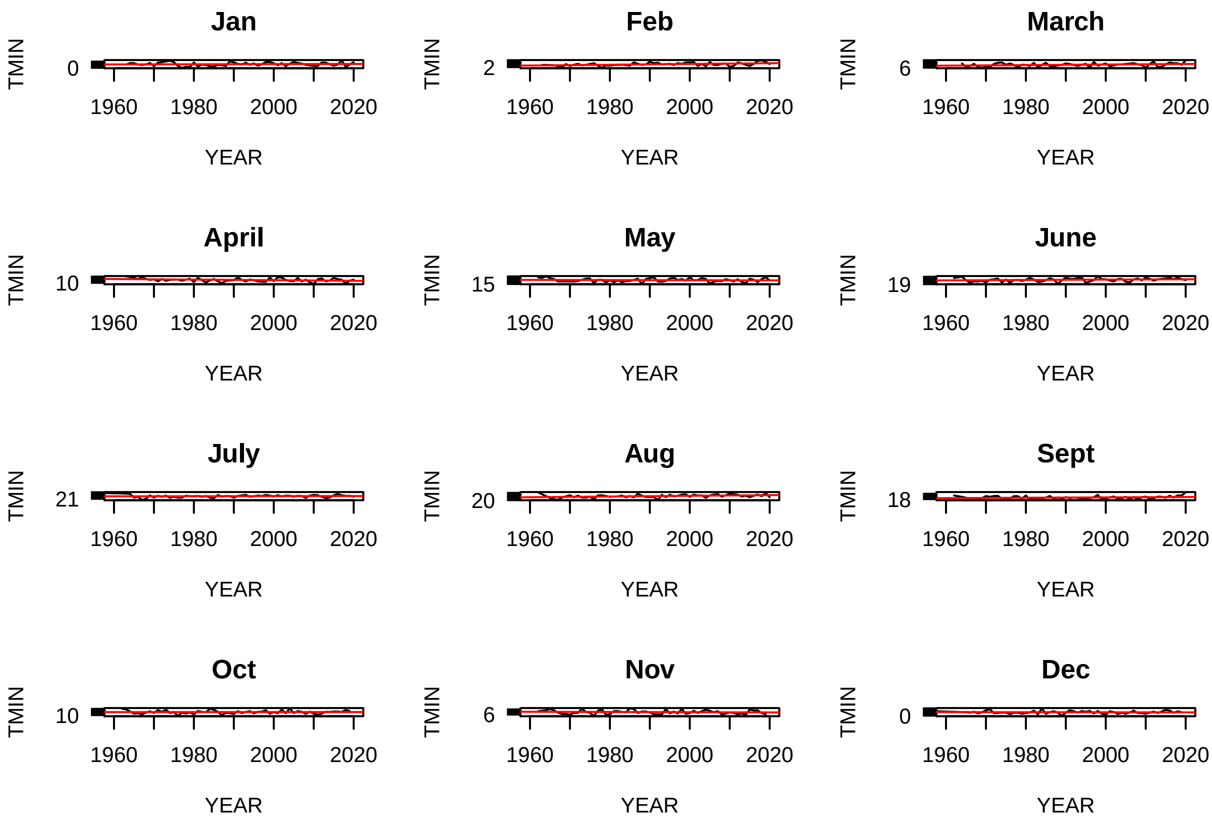


Now that I have all of my data (it very very rarely snows in Southeast Texas, so I won't be using that variable), I want to use for loops to look at all months at once for TMAX, TMIN, and PRCP.

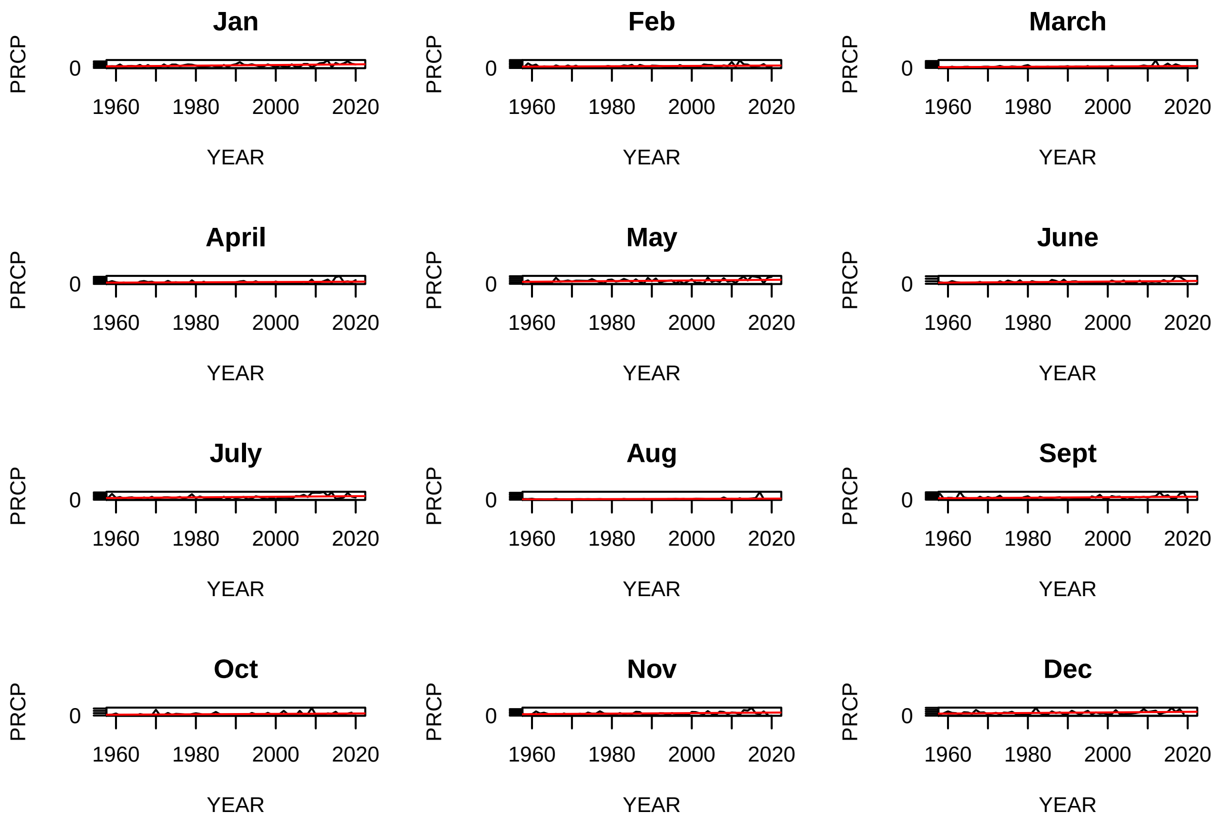
```
#all tmax
Months <- c("Jan", "Feb", "March", "April", "May", "June", "July", "Aug", "Sept", "Oct", "Nov", "Dec")
par(mfrow = c(4,3), mar = c(5,4,3,2) +.1)
TMAXresult <- NA
for(i in 1:12) {
  plot(TMAX~YEAR, data = MonthlyTMAXMean[MonthlyTMAXMean$MONTH == i,], ty = "l", las = 1, xlim = c(1960
  Month.max.lm <- lm(TMAX~YEAR, data = MonthlyTMAXMean[MonthlyTMAXMean$MONTH == i,])
  summary(Month.max.lm)
  abline(coef(Month.max.lm), col = "red")
  TMAXresult <- rbind(TMAXresult, cbind(Months[i], round(coef(Month.max.lm)[2],4), round(summary(Month.
}
```




```r
##plot all tmins at once

MonthlyTMINMean$YEAR = as.numeric(MonthlyTMINMean$Year)
MonthlyTMINMean$MONTH = as.numeric(MonthlyTMINMean$Month)
TMINresult <- NA
par(mfrow = c(4,3), mar = c(5,4,3,2) +.1)
for(i in 1:12) {
  plot(TMIN~YEAR, data = MonthlyTMINMean[MonthlyTMINMean$MONTH == i,], ty = "l", las = 1, xlim = c(1960
  Month.min.lm <- lm(TMIN~YEAR, data = MonthlyTMINMean[MonthlyTMINMean$MONTH == i,])
  summary(Month.min.lm)
  abline(coef(Month.min.lm), col = "red")
  TMINresult <- rbind(TMINresult, cbind(Months[i], round(coef(Month.min.lm)[2],4), round(summary(Month.r
}
```




```
#all prcp at once
par(mfrow = c(4,3), mar = c(5,4,3,2) +.1)
PRCPresult <- NA
for(i in 1:12) {
  plot(PRCP~YEAR, data = MonthlyPRCPSum[MonthlyPRCPSum$MONTH == i,], ty = "l", las = 1, xlim = c(1960,20
  Month.PRCP.lm <- lm(PRCP~YEAR, data = MonthlyPRCPSum[MonthlyPRCPSum$MONTH == i,])
  summary(Month.PRCP.lm)
  abline(coef(Month.PRCP.lm), col = "red")
  PRCPresult <- rbind(PRCPresult, cbind(Months[i], round(coef(Month.PRCP.lm)[2],4), round(summary(Month
}
```



Now that I've plotted them all, it will be helpful to see the actual numbers behind the graph to determine if any of the trends are statistically significant.

```
##Results
Results <- data.frame(Month = TMINresult[,1],
                      TMINSlope = TMINresult[,2],
                      TMIN_P = as.numeric(TMINresult[,3]),
                      TMINRsq = TMINresult[,4],
                      TMAXSlope = TMAXresult[, 2],
                      TMAX_P = as.numeric(TMAXresult[,3]),
                      TMAXRsq = TMAXresult[,4],
                      PRCPSlope = PRCPresult[,2],
                      PRCP_P = as.numeric(PRCPresult[,3]),
                      PRCPsq = PRCPresult[,4])

Results$starTMIN <- ifelse(Results$TMIN_P <= .001, "***",
                           ifelse(Results$TMIN_P <= .01, "**",
                            ifelse(Results$TMIN_P <= .05, "*", "NS")))
Results$starTMAX <- ifelse(Results$TMAX_P <= .001, "***",
                                   ifelse(Results$TMAX_P <= .01, "**",
                                          ifelse(Results$TMAX_P <= .05, "*", "NS")))
Results$starPRCP <- ifelse(Results$PRCP_P <= .001, "***",
                           ifelse(Results$PRCP_P <= .01, "**",
                                  ifelse(Results$PRCP_P <= .05, "*", "NS")))

Results$TMINSlope = paste(Results$TMINSlope, Results$starTMIN)
Results$TMAXSlope = paste(Results$TMAXSlope, Results$starTMAX)
```

```r
Results$PRCPSlope = paste(Results$PRCPSlope, Results$starPRCP)

Results <- Results[c(2:13), c(1,2,3,4,5,6,7,8,9,10)]

colnames(Results) = c("Month", "Slope TMIN", "P value", "R^2", "Slope TMAX", "P value", "R^2", "Slope Pl
library(htmlTable)
htmlTable(Results)
```

Month

Slope TMIN

P value

R^2

Slope TMAX

P value

R^2

Slope PRCP

P Value

R^2

2

Jan

0.0098 NS

0.5948

0.005

0.0313 NS

0.0908

0.052

4.0429 ***

0

0.344

3

Feb

0.0548 **

0.0012

0.173

0.0331 NS

0.0524

0.066

2.1265 **

0.0012

0.118

4

March

0.0364 *

0.0463

0.073

0.025 NS

0.0661

0.062

3.0203 ***

0

0.203

5

April

-0.0375 *

0.0281

0.085

-0.0108 NS

0.3061

0.019

2.2526 **

0.0032

0.099

6

May

-0.0063 NS

0.569

0.006

0.0022 NS

0.7872

0.001

3.3353 ***

0

0.21

7

June

0.0101 NS

0.1501

0.036

0.0042 NS

0.5747

0.006

5.2733 ***

0

0.282

8

July

0.0026 NS

0.6667

0.003

-0.0158 *

0.0445

0.069

4.0397 ***

0

0.218

9

Aug

0.019 **

0.0052

0.134

0.0073 NS

0.4269

0.012

6.8938 **

0.0012

0.121

10

Sept

0.021 NS

0.108

0.045

0.0114 NS

0.2332

0.025

3.9259 ***

6e-04

0.136

11

Oct

0.0011 NS

0.9505

0

2e-04 NS

0.9818

0

4.3246 ***

5e-04

0.14

12

Nov

-0.0119 NS

0.5019

0.008

-0.0125 NS

0.4284

0.011

3.1742 ***

0

0.219

13

Dec

-0.0073 NS

0.689

0.003

-0.0048 NS

0.7356

0.002

2.8844 ***

0

0.218

Amazing! Now we have a lovely table that lays out all of the analysis I just did.

Our monthly average minimum temperatures (TMIN) only have a statistically significant trend in the months of February, March, April, and August, with February and August being the strongest. We can only reject the null hypothesis in these 4 months. All of the effects are slight, and April's is actually negative (meaning there is a trend of lower minimum temperatures in April).

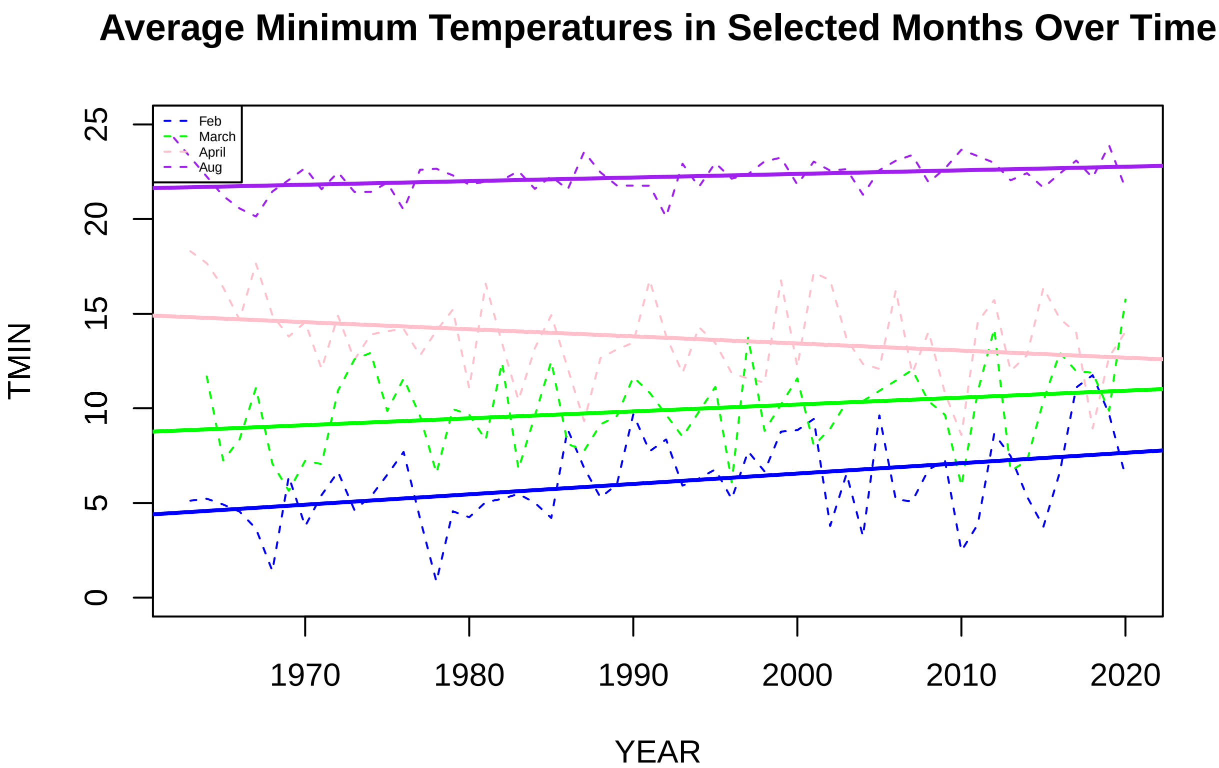
```
.0548+.0364+.019-.0375
```

```
## [1] 0.0727
```

It seems though that overall the increases in minimum temperature slightly outweigh April's decrease. This makes sense, since our first graph showed a slight up trend, though it was not statistically significant. I'm also not sure if this is the best/correct way to determine if the increases in 3 months outweigh the decreases of one month, but at least it's a start!

Here's a plot including TMINs for the statistically significant months:

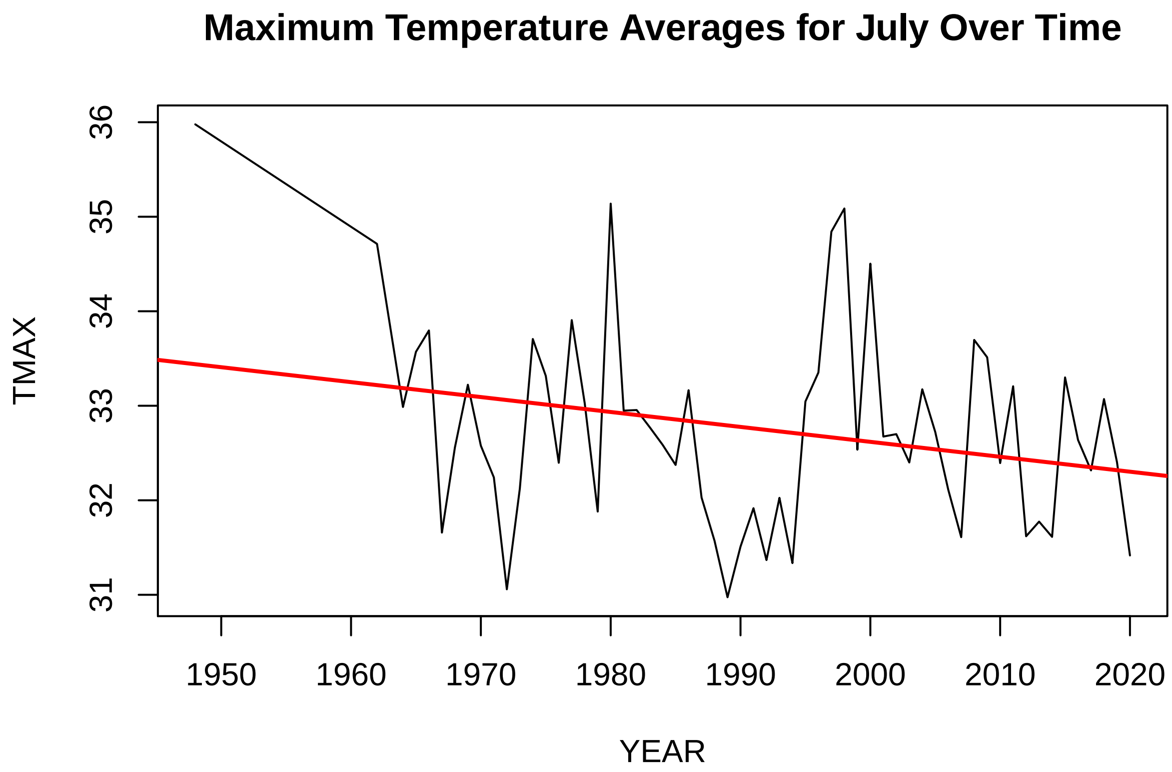
```
plot(TMIN~YEAR, MonthlyTMINMean[MonthlyTMINMean$MONTH == 4,], main = "Average Minimum Temperatures in Se
lines(TMIN~YEAR, MonthlyTMINMean[MonthlyTMINMean$MONTH == 2,], col = "blue", lty=2)
lines(TMIN~YEAR, MonthlyTMINMean[MonthlyTMINMean$MONTH == 8,], col = "purple", lty=2)
lines(TMIN~YEAR, MonthlyTMINMean[MonthlyTMINMean$MONTH == 3,], col = "green", lty=2)
abline(coef(lm(TMIN~YEAR, MonthlyTMINMean[MonthlyTMINMean$MONTH == 2,])), col = "blue", lwd = 2)
abline(coef(lm(TMIN~YEAR, MonthlyTMINMean[MonthlyTMINMean$MONTH == 4,])), col = "pink", lwd = 2)
abline(coef(lm(TMIN~YEAR, MonthlyTMINMean[MonthlyTMINMean$MONTH == 8,])), col = "purple", lwd = 2)
abline(coef(lm(TMIN~YEAR, MonthlyTMINMean[MonthlyTMINMean$MONTH == 3,])), col = "green", lwd = 2)
legend("topleft", legend=c("Feb", "March", "April", "Aug"),
       col=c("blue", "green", "pink", "purple"), lty=2, cex=0.4)
```



Our monthly average maximum temperatures only show one statistically significant trend, which is in July. We can reject the null hypothesis for July, but no other months. It shows that maximum temperatures in the month of July are actually decreasing slightly, which I found very interesting. Unfortunately, our significant data with TMIN and TMAX doesn't overlap on months, which means I can't make a cool graph with both TMIN and TMAX for a given month... but I'll do this later for all months anyway because it's fun...

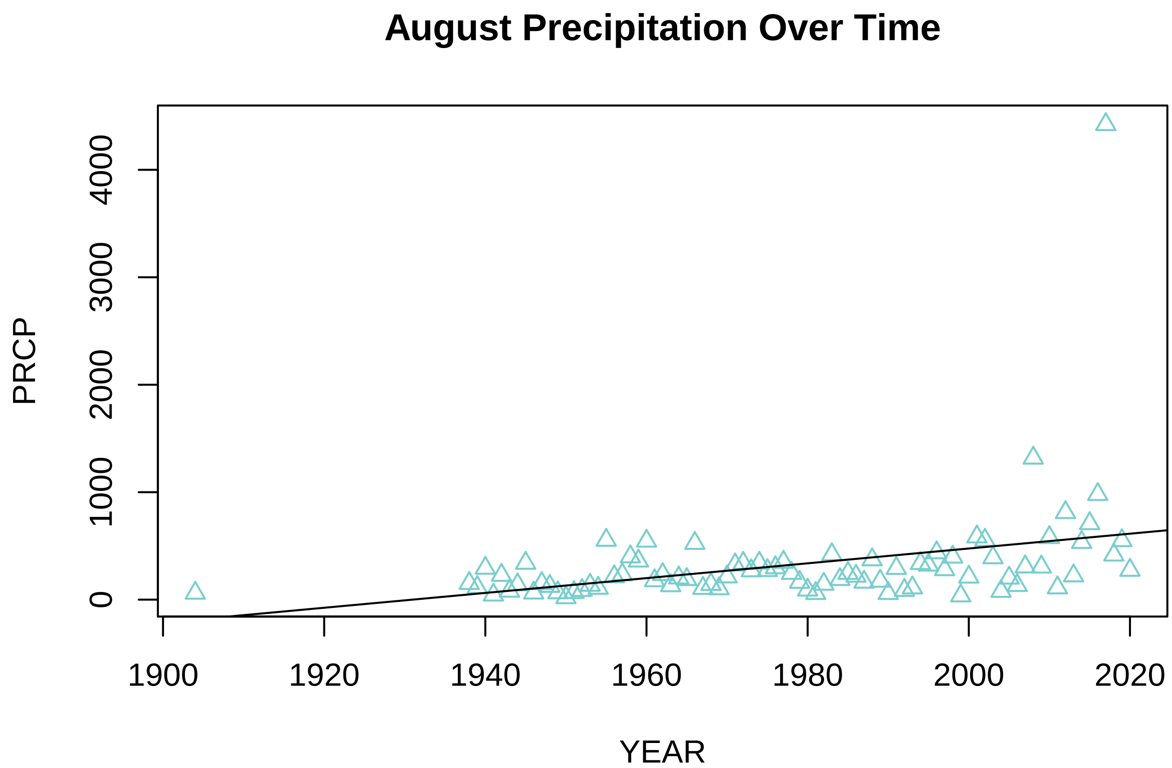
Our TMAX graph for July is messy and shows graphically what we knew from our chart, which is that the trend is small. This is a reminder that just because we designate something as "statistically significant" doesn't necessarily mean it is significant enough for other purposes (such as city planning or policymaking). The trend is there - that is all that our P value tells us; it does not tell us whether the trend is important or not for decision making.

```
plot(TMAX~YEAR, MonthlyTMAXMean[MonthlyTMAXMean$MONTH == 7,], main = "Maximum Temperature Averages for J
abline(coef(lm(TMAX~YEAR, MonthlyTMAXMean[MonthlyTMAXMean$MONTH == 7,])), col = "red", lwd = 2)
```



Precipitation was the most interesting and statistically revelant data. All of the trends are significant, meaning we can reject the null hypothesis for all months for the precipitation variable. They are also positive, meaning rainfall has increased over time in all months. The highest of these are especially from June to October, which makes perfect sense as it coincides with North Atlantic Hurricane Season. I would like to find a dataset with tropical cylones to plot as a supplement for this data.

```
plot(PRCP~YEAR, MonthlyPRCPSum[MonthlyPRCPSum$MONTH == 8,], pch = 2, col = "darkslategray3", main = "Aug
abline(coef(lm(PRCP~YEAR, MonthlyPRCPSum[MonthlyPRCPSum$MONTH == 8,])))
```



It looks like there's one outlier that may be skewing the data, what is it?

```
max(MonthlyPRCPSum$PRCP, na.rm = T)
```

```
## [1] 4422.4
```

```
MonthlyPRCPSum$Year[MonthlyPRCPSum$PRCP == 4422.4]
```

```
## [1] "2017"
```

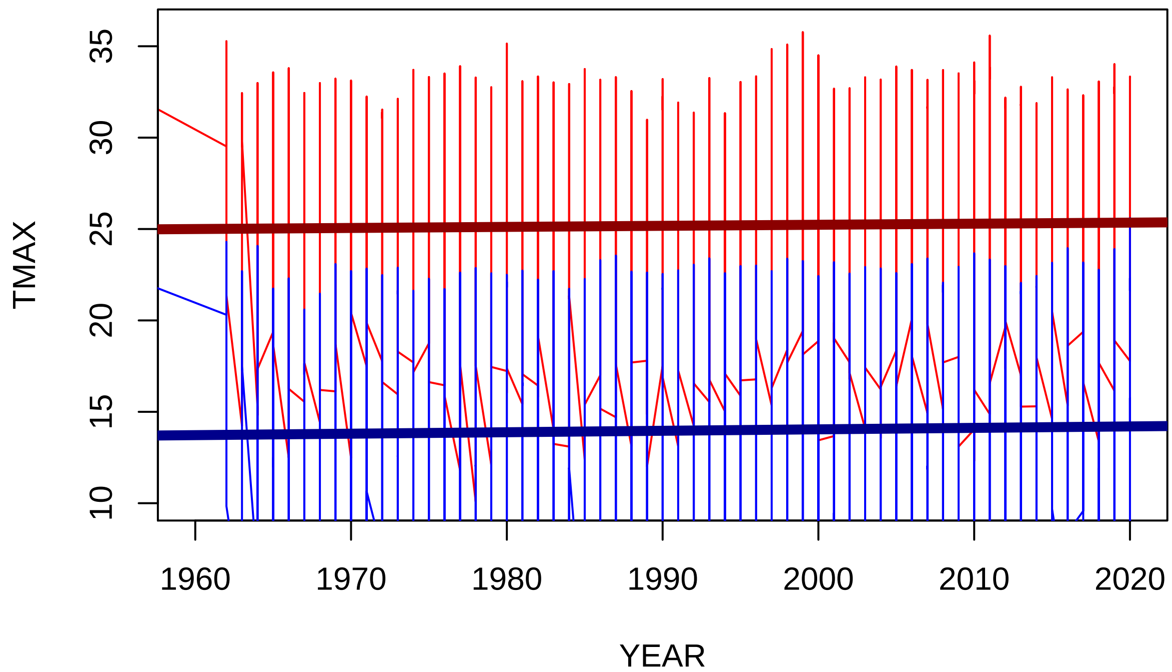
So it's from August 2017, which was when Hurricane Harvey hit.

What happens when we excluse this "100 year storm"?

plot(PRCP~YEAR, MonthlyPRCPSum[MonthlyPRCPSum$YEAR != 2017,], pch = 2, col = "darkslategray3", main = "August Precipitation Over Time, excluding 2017") #this doesn't work yet but I'm going to figure it out eventually, just maybe not in time for this assignment submission! this also isn't just august.. i think I need to use an & symbol maybe? abline(coef(lm(PRCP~YEAR, MonthlyPRCPSum[MonthlyPRCPSum$YEAR != 2017])))

Can we plot min and max graphs together? This is just for fun :)

```
plot(TMAX~YEAR, MonthlyTMAXMean, type = "l", col = "red", xlim = c(1960,2020))
lines(TMIN~YEAR, MonthlyTMINMean, col = "blue", xlim = c(1960,2020))
abline(coef(lm(TMAX~YEAR, MonthlyTMAXMean),), col = "darkred", lwd = 5)
abline(coef(lm(TMIN~YEAR, MonthlyTMINMean),), col = "darkblue", lwd = 5)
```



In conclusion, the climate data for Southeast Texas does not by itself show a warming trend, but does show increased rainfall over time. This highlights the importance of what can happen when you cherry pick data, and also shows why “climate change” is a more accurate term than “global warming.” The globe overall is warming, but that does not mean that every single place is getting hotter! These data could easily be used by deniers and perhaps partially explains the prevalence of them in the region. . . but we know better!