Boone

Sam Tolbert

2024-10-02

```r
knitr::opts_chunk$set(echo = FALSE)
```

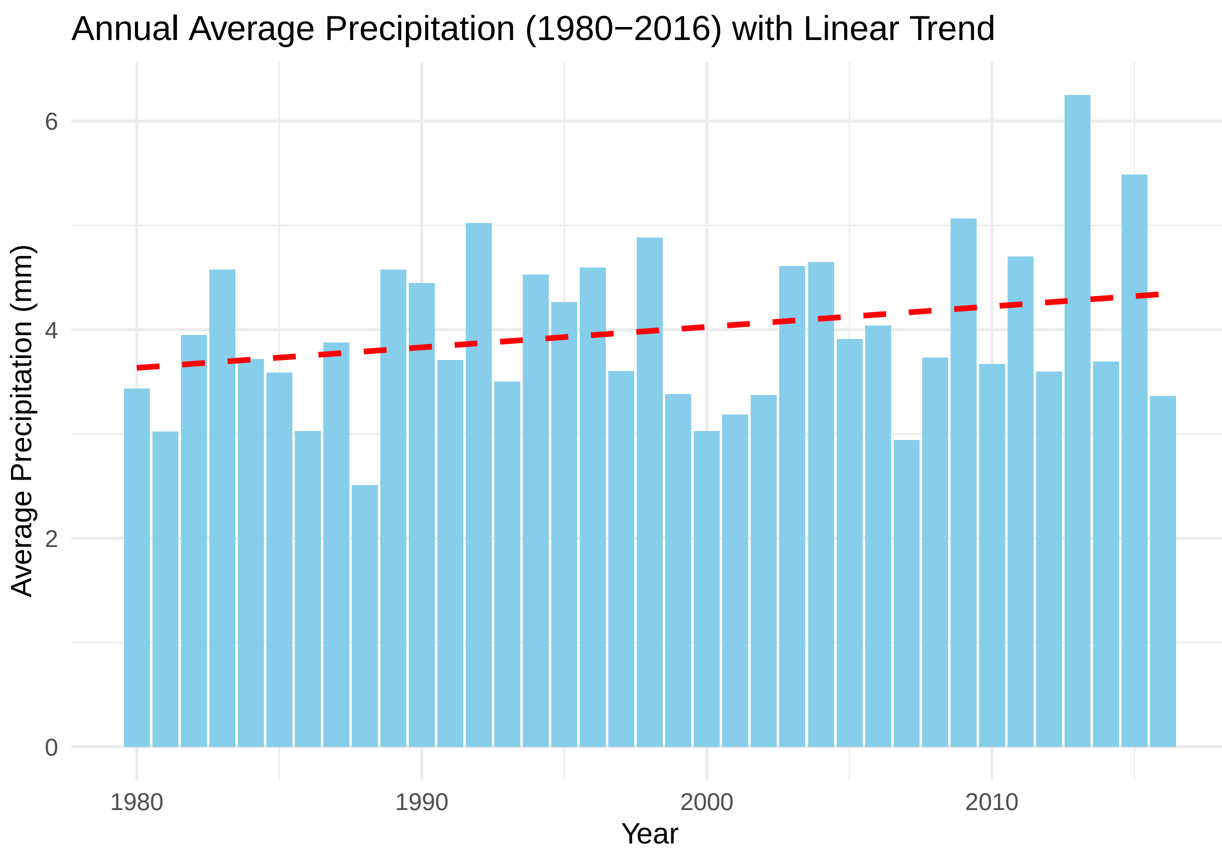
Initial Set-Up

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

*All code was produced in conversation with R Wizard GPT. Prompts included in code where relevant.

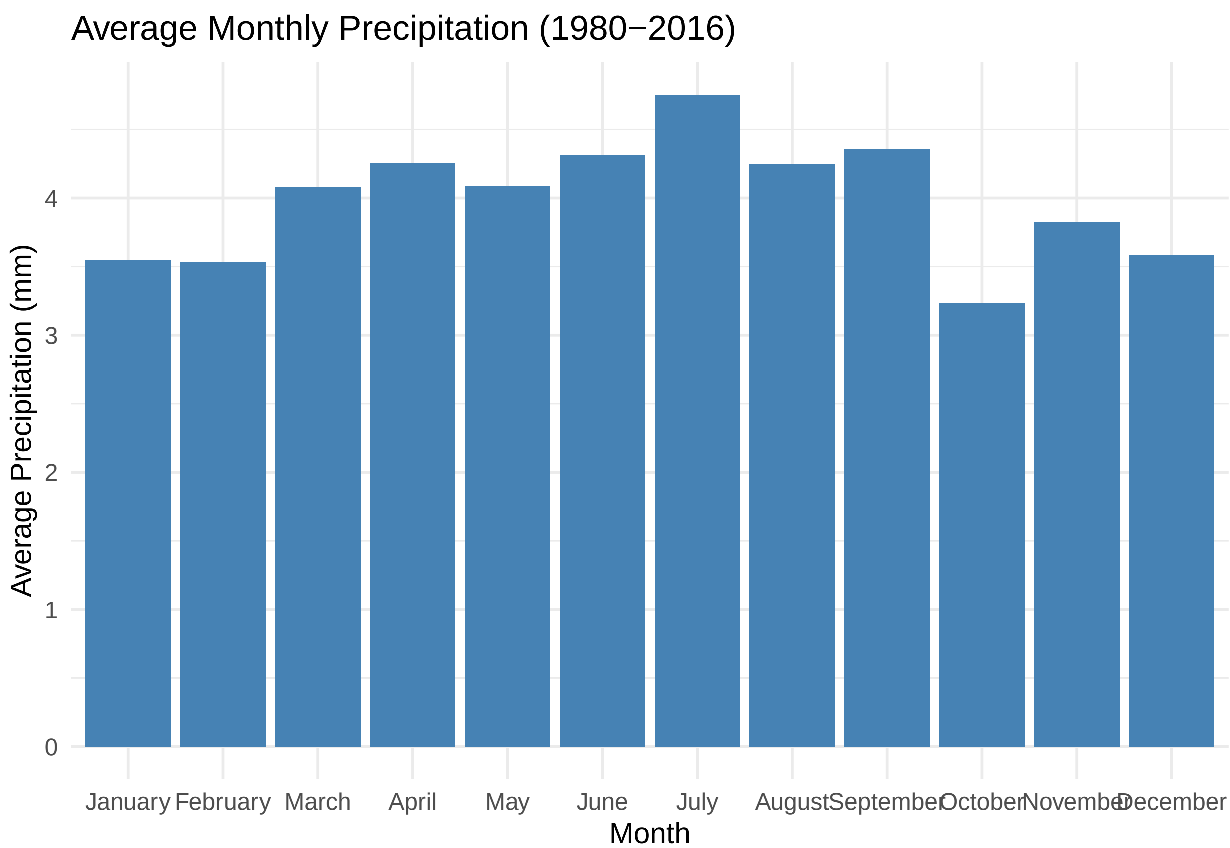
The first analysis we can run is finding large scale trends within the precipitation data. Using the HUC data, we found that a general increase in annual precipitation. We find that from 1980 to 2016 there has been a steady increase in the annual mean preciptation

```
## `geom_smooth()` using formula = 'y ~ x'
```



While this general trend is useful, in regard to infrastructure, it is vital to know if this precipitation is evenly distributed throughout the year, is concentrated in larger, predictable hurricane systems (Hurricane Season defined by the State of North Carolina as June 1-November 30) or the less predictable, smaller frontal systems throughout the rest of the year. We find that the rainest months occur during Hurricane season, with July, September, and June being the rainiest months.

```
#Average Monthly Precipitation

Boone_Monthly_Averages_AllYears <- Boone_Processed %>%
  filter(year >= 1980 & year <= 2016) %>%
  group_by(month) %>%
  summarize(monthly_avg_precip = mean(Precipitation_mm, na.rm = TRUE))

# Plotting the monthly averages (across all years) using a bar plot
ggplot(Boone_Monthly_Averages_AllYears, aes(x = factor(month), y = monthly_avg_precip)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  labs(title = "Average Monthly Precipitation (1980-2016)",
       x = "Month",
       y = "Average Precipitation (mm)") +
  theme_minimal() +
  scale_x_discrete(labels = month.name)  # Adding month names to the x-axis
```

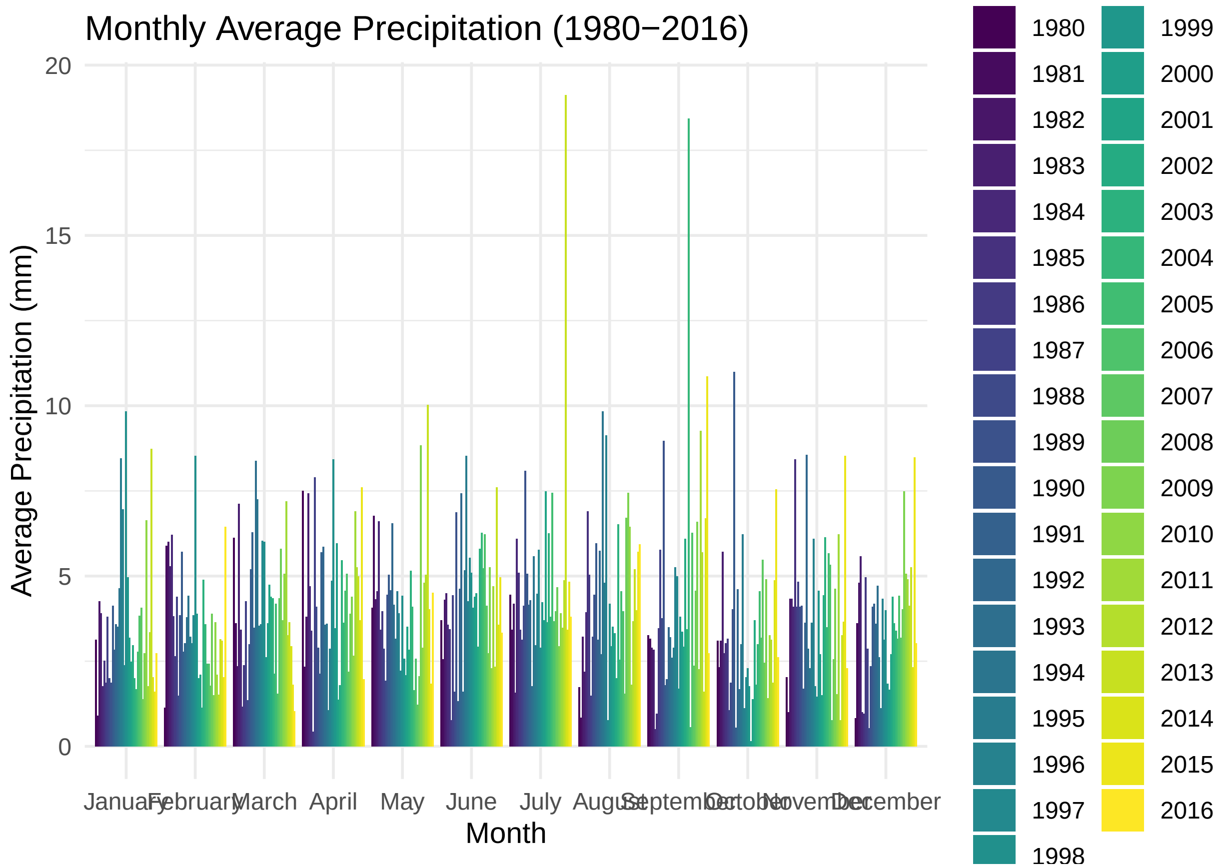


Combining these two analyses we can see the monthly trends from 1980-2016, seeing a steady increase in mean precipation, particularly in the summer months

```r
# Group by year and month, and calculate the mean precipitation for each month
Boone_Monthly_Averages <- Boone_Processed %>%
  filter(year >= 1980 & year <= 2016) %>%
  group_by(year, month) %>%
  summarize(monthly_avg_precip = mean(Precipitation_mm, na.rm = TRUE))
```

```
## 'summarise()' has grouped output by 'year'. You can override using the
## '.groups' argument.
```

```r
# Plotting the monthly averages using a bar plot
ggplot(Boone_Monthly_Averages, aes(x = factor(month), y = monthly_avg_precip, fill = factor(year))) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "Monthly Average Precipitation (1980-2016)",
       x = "Month",
       y = "Average Precipitation (mm)") +
  theme_minimal() +
  scale_fill_viridis_d(name = "Year") +
  scale_x_discrete(labels = month.name)  # Adding month names to the x-axis
```



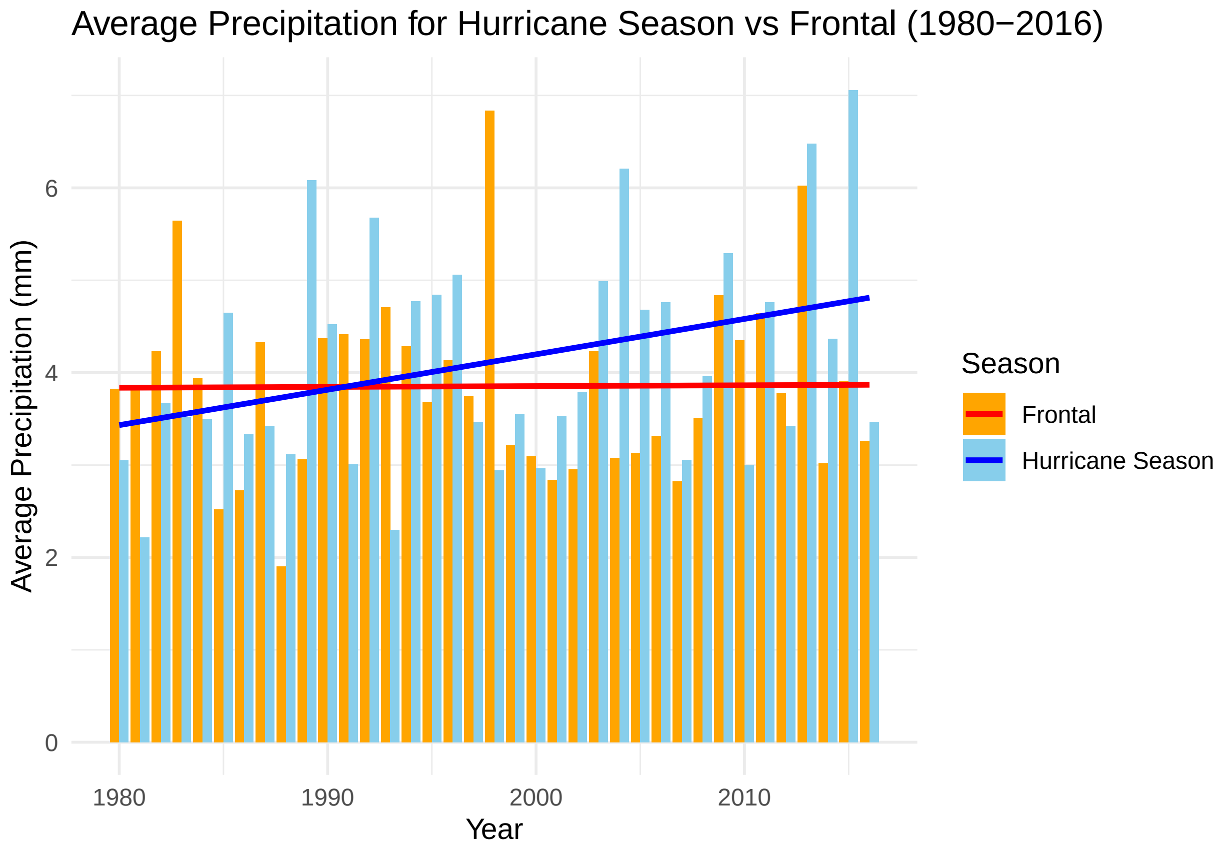
Seeing these trends, we now compare the mean precipiations of Hurricane Season against Frontal systems.

```
#Creating Hurricane Season vs. Frontal Dataframe
Boone_Seasonal <- Boone_Processed %>%
  mutate(Season = case_when(
    (month >= 6 & month <= 11) ~ "Hurricane Season",  # June to November
    TRUE ~ "Frontal"  # December to May
  ))

#Filter data for the years 1980-2016
Boone_Seasonal <- Boone_Seasonal %>%
  filter(year >= 1980 & year <= 2016)

#Grouping by year and season, we calculate the average precipitation for each year and season
Boone_Seasonal_Averages <- Boone_Seasonal %>%
  group_by(year, Season) %>%
  summarize(avg_precip = mean(Precipitation_mm, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'year'. You can override using the
## `.groups` argument.
```

We express the results with a bar plot with separate linear regression lines for each season

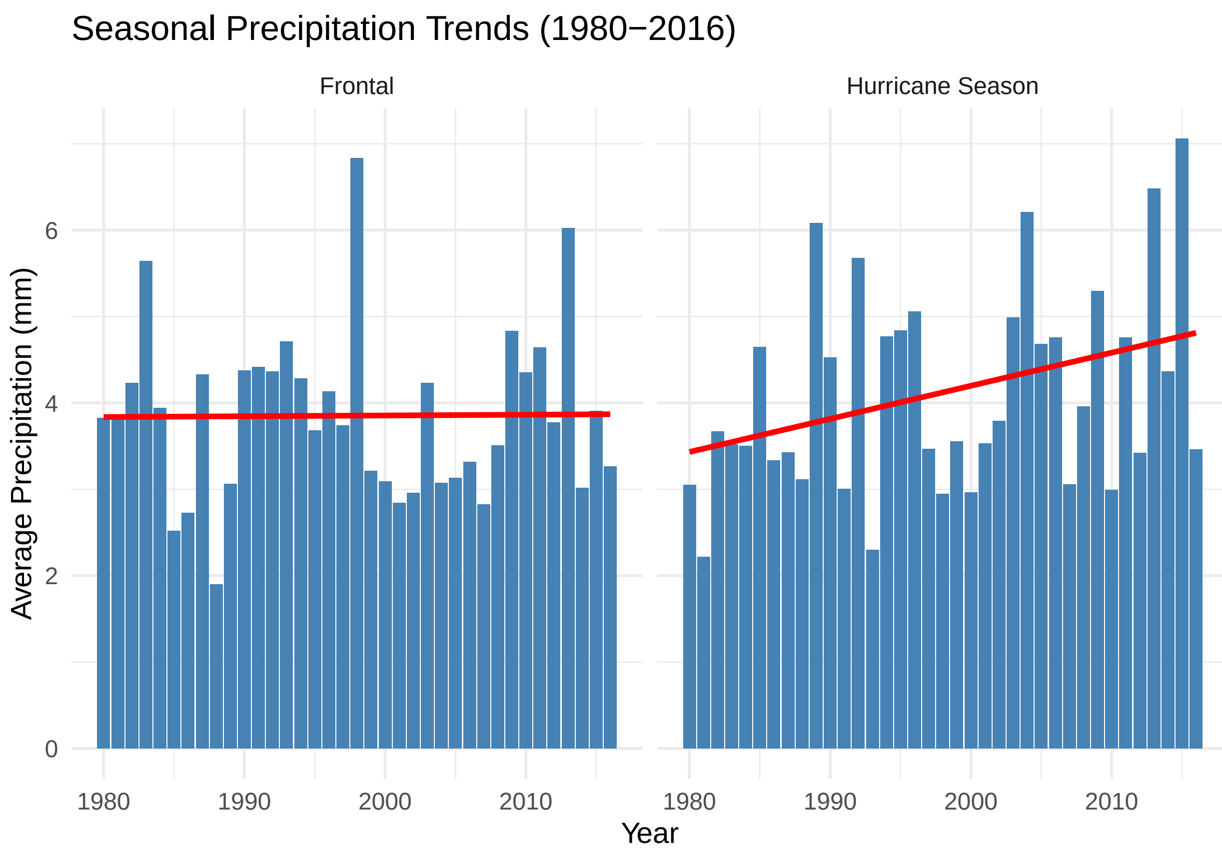
```
## `geom_smooth()` using formula = 'y ~ x'
```



This graph makes it apparent that Hurricane Season has been steadily increasing in terms of mean precipation in the last 36 years, while frontal precipitation has been staying flat. Combining this with our previous analysis, we can assume that the majority of the annual precipitation increases are due to increased intensities and rainfall of hurricane systems, rather than frontal systems.

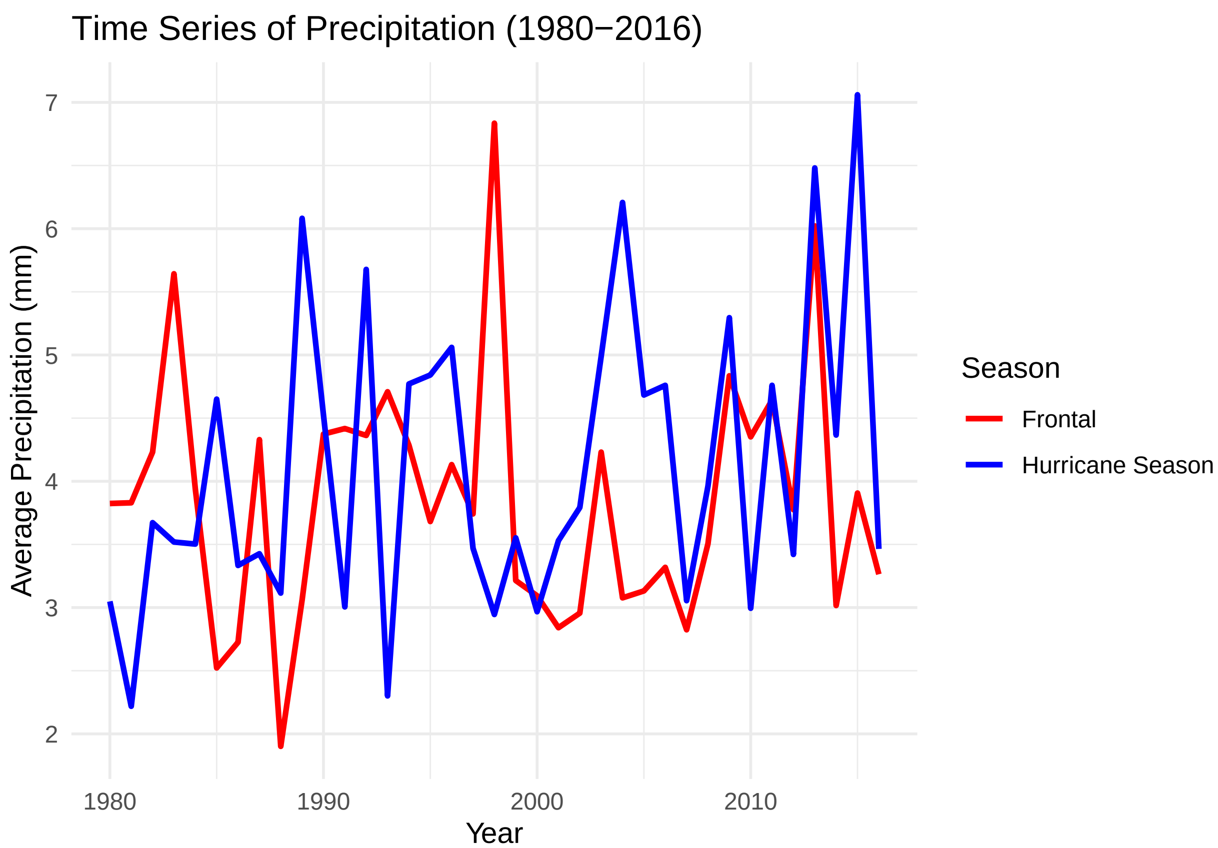
Here are those graphs again, side by side:

```
## ‘geom_smooth()‘ using formula = ’y ~ x’
```



We can also see this incrasing dominance of Hurricane systems, especially since the year 2000 in the followng Time Series Analysis:

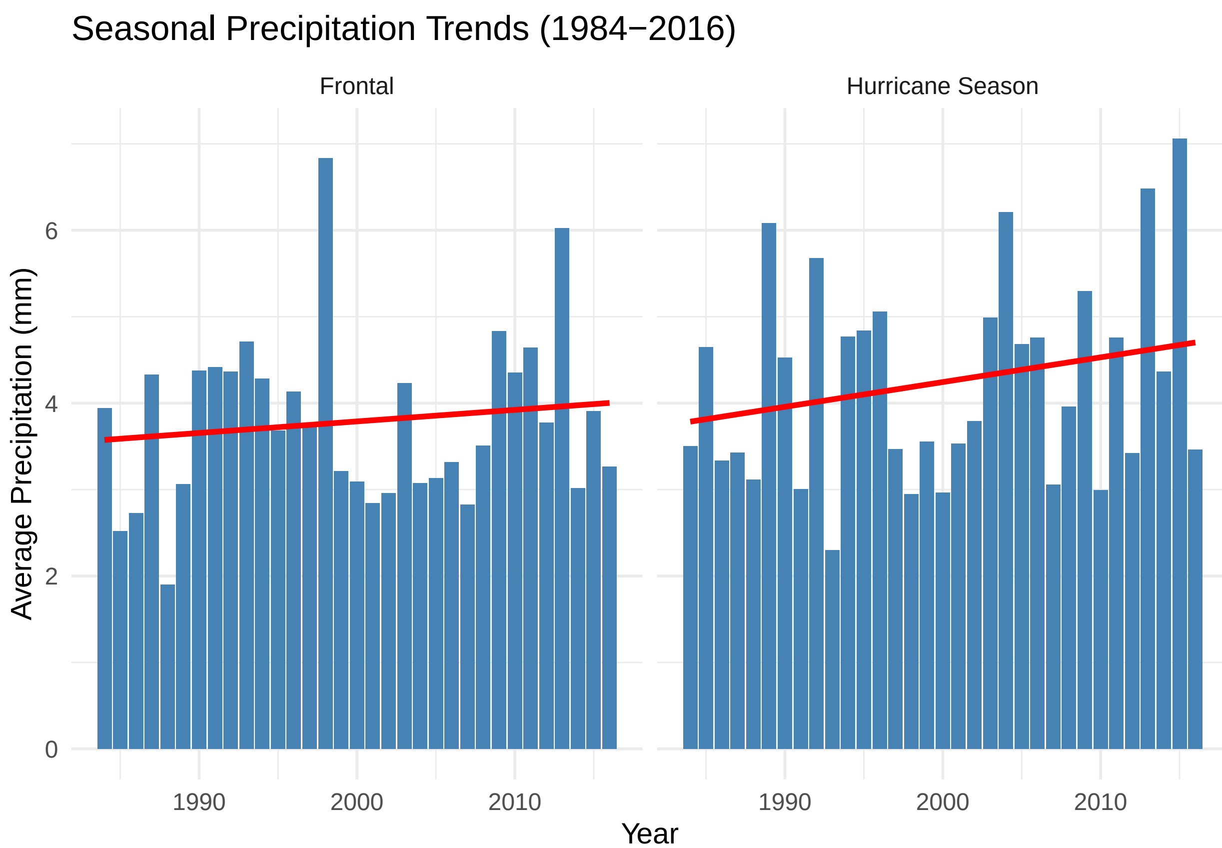
```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```



These increasing trends become even more apparent when we give equal weight to pre-2000 and post-2000 date, filtering data to give a 16 year year window pre/post 2000.

```
Boone_Seasonal_Averages_Modified <- Boone_Seasonal_Averages %>%
  filter(year >= 1984 & year <= 2016) #creating a data set from 1984-2016

# Plot with separate regression lines for each season (1984-2016)
ggplot(Boone_Seasonal_Averages_Modified, aes(x = year, y = avg_precip)) +
  geom_bar(stat = "identity", fill = "steelblue") +  # Bar plot for average precipitation
  geom_smooth(method = "lm", se = FALSE, color = "red") +  # Add linear regression line
  facet_wrap(~ Season) +  # Facet by season ("Hurricane Season" and "Frontal")
  labs(title = "Seasonal Precipitation Trends (1984-2016)",
       x = "Year",
       y = "Average Precipitation (mm)") +
  theme_minimal()
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



This leads us to begin to want to track the temporal changes, not just the seasonal. How have intensities altered intra season since date started being collected?

```
# Load necessary libraries
library(dplyr)
library(lubridate)

# Assuming the data is already loaded into Boone

# Create new columns for year if not already present
Boone_Data <- Boone_Data %>%
  mutate(Date = as.Date(Date))  # Ensure the Date column is in Date format

# Function to calculate Fa (Return Period) using the Hazen Equation
# n is the rank and y is the total number of observations
calc_fa <- function(n, y) {
  return(100 * (2 * n - 1) / (2 * y))
}

# Filter for data between 1980-01-01 and 2000-12-31 (daily resolution)
Boone_filtered_19802000 <- Boone_Data %>%
  filter(Date >= as.Date("1980-01-01") & Date <= as.Date("2000-12-31"))

# Rank the precipitation data for 1980-2000 in descending order (highest to lowest)
Boone_Hazen_Mutate_19802000 <- Boone_filtered_19802000 %>%
  arrange(desc(`Area.Weighted.Mean.Precipitation..mm.per.day.`)) %>%
```

```r
  mutate(
    rank = row_number(),  # Assign ranks where 1 is the highest precipitation
    Fa = calc_fa(rank, n())  # Calculate Fa using Hazen Equation
  ) %>%
  mutate(likelihood_event = 100 / Fa)  # Add likelihood column

# Filter for data between 2000-01-01 and 2016-12-31 (daily resolution)
Boone_filtered_20002016 <- Boone_Data %>%
  filter(Date >= as.Date("2000-01-01") & Date <= as.Date("2016-12-31"))

# Rank the precipitation data for 2000-2016 in descending order (highest to lowest)
Boone_Hazen_Mutate_20002016 <- Boone_filtered_20002016 %>%
  arrange(desc(`Area.Weighted.Mean.Precipitation..mm.per.day.`)) %>%
  mutate(
    rank = row_number(),  # Assign ranks where 1 is the highest precipitation
    Fa = calc_fa(rank, n())  # Calculate Fa using Hazen Equation
  ) %>%
  mutate(likelihood_event = 100 / Fa)  # Add likelihood column
```

heatmap at the bottom because it looks cool but I don't think we ultimately need?

```
## 'summarise()' has grouped output by 'year'. You can override using the
## '.groups' argument.
```