Greenville Report

Sam Tolbert

2024-10-02

Load Raw Data see broad annual and monthly trends from 1980-2016

```
Greenville <- read.csv("Hydrology/Data/Raw/Greenville_daily_precip_1980-present_HUC_030201030403_dayMet_

Greenville_Data <- Greenville

library(dplyr)
```

```
##
## 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
```

```
library(ggplot2)
library(lubridate)
```

```
##
## Attaching package: 'lubridate'

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

```
# Rename the precipitation column to 'Precipitation in mm'
Greenville_Processed <- Greenville_Data %>%
  rename(Precipitation_mm = Area.Weighted.Mean.Precipitation..mm.per.day.)

# Ensure the 'Date' column is in date format
Greenville_Processed <- Greenville_Processed %>%
  mutate(Date = as.Date(Date))
```

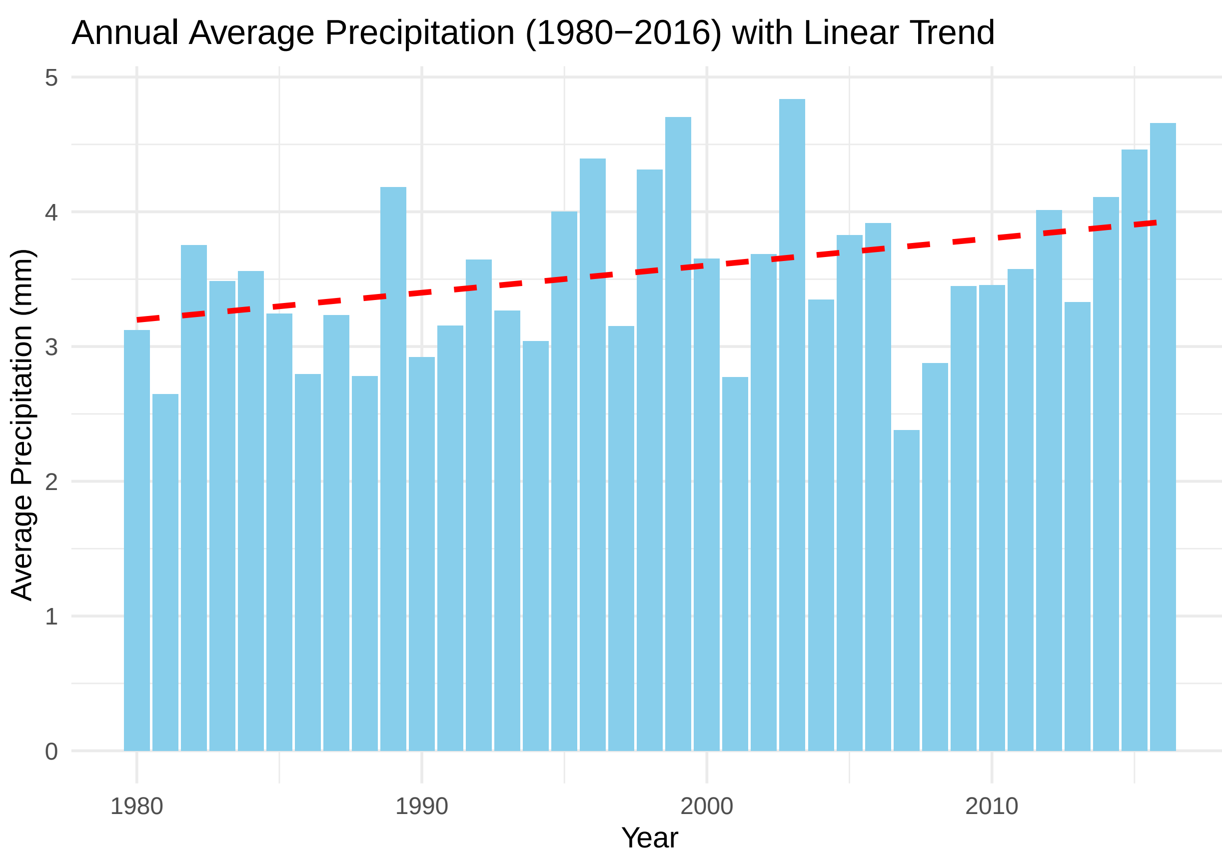
All code was produced in conversation with R Wizard GPT. Prompts included in code where relevant.

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 precipitation.

```
# First, we isolate the means of each year:
Greenville_Annual_Averages <- Greenville_Processed %>%
  filter(year >= 1980 & year <= 2016) %>%
  group_by(year) %>%
  summarize(annual_avg_precip = mean(Precipitation_mm, na.rm = TRUE))

# The we plot the annual averages using a bar plot with a linear regression line
ggplot(Greenville_Annual_Averages, aes(x = year, y = annual_avg_precip)) +
  geom_bar(stat = "identity", fill = "skyblue") +  # Bar plot
  geom_smooth(method = "lm", se = FALSE, color = "red", linetype = "dashed") +  # Linear regression lin
  labs(title = "Annual Average Precipitation (1980-2016) with Linear Trend",
       x = "Year",
       y = "Average Precipitation (mm)") +
  theme_minimal()
```

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



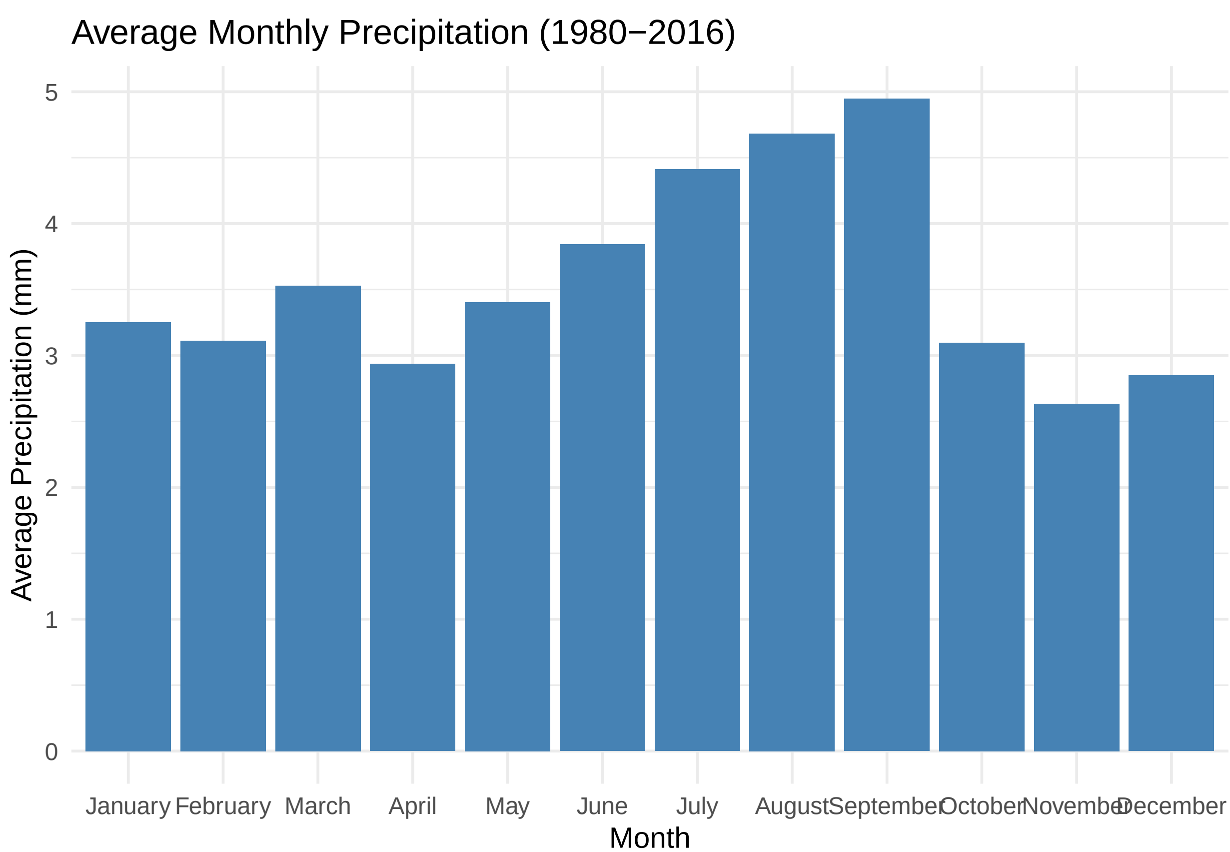
While this general trend is useful, in regard to infrastructure, it is vital to know if this precipitation is evenly distributed throughout the year, is concentrated in larger, predictable hurricane systems (Hurricane

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 September, August, July being the rainiest months.

```r
#Average Monthly Precipitation

Greenville_Monthly_Averages_AllYears <- Greenville_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(Greenville_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
```

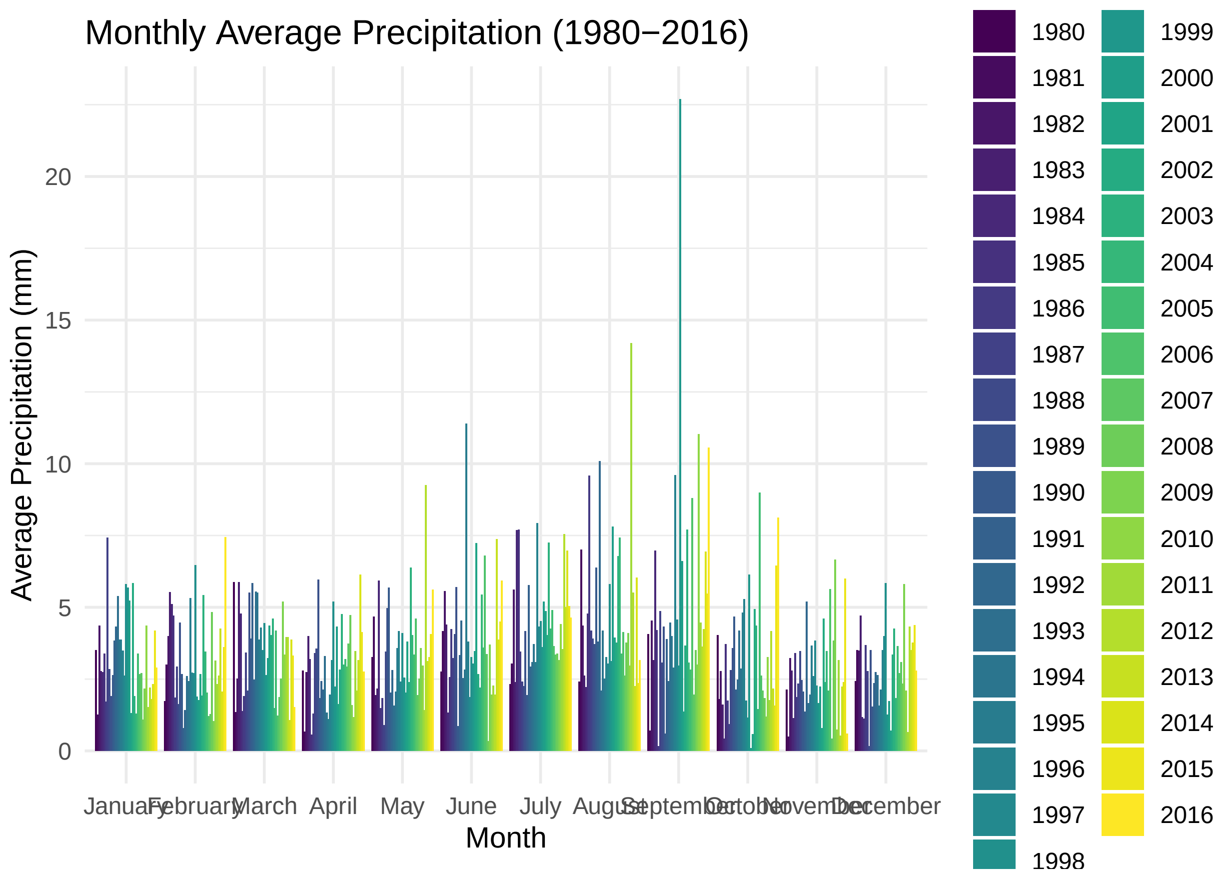


Combining these two analyses we can see the monthly trends from 1980-2016, seeing a steady increase in mean precipation, particularly in the summer months.

```r
# Group by year and month, and calculate the mean precipitation for each month
Greenville_Monthly_Averages <- Greenville_Processed %>%
  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.
```

```
# Plotting the monthly averages using a bar plot
ggplot(Greenville_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 precipiation of Hurricane Season against Frontal systems.

```
#Creating Hurricane Season vs. Frontal Dataframe
Greenville_Seasonal <- Greenville_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
Greenville_Seasonal <- Greenville_Seasonal %>%
  filter(year >= 1980 & year <= 2016)
```

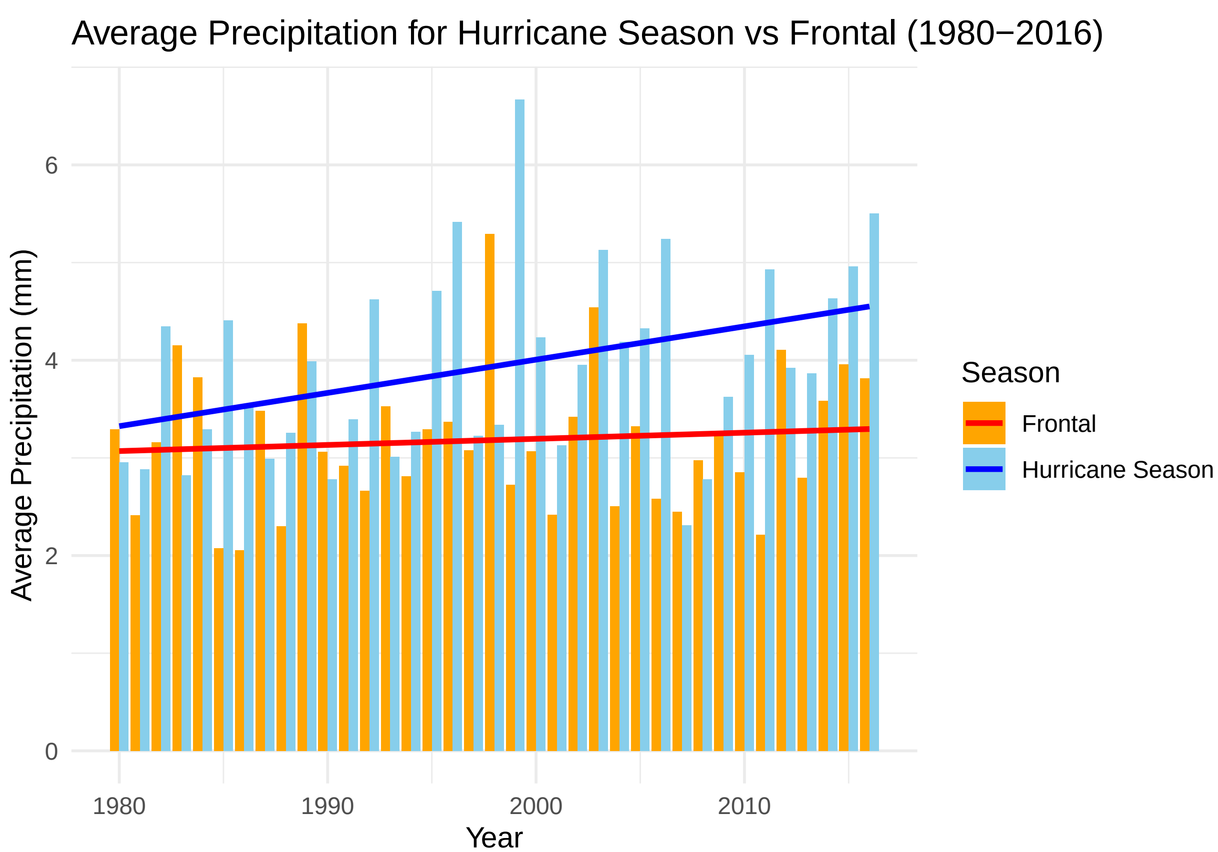
```r
#Grouping by year and season, we calculate the average precipitation for each year and season
Greenville_Seasonal_Averages <- Greenville_Seasonal %>%
  group_by(year, Season) %>%
  summarize(avg_precip = mean(Precipitation_mm, na.rm = TRUE))
```

```
## `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.

```r
ggplot(Greenville_Seasonal_Averages, aes(x = year, y = avg_precip, fill = Season)) +
  geom_bar(stat = "identity", position = "dodge") +  # Bar plot for both seasons side-by-side
  geom_smooth(method = "lm", aes(color = Season), se = FALSE) +  # Add separate linear regression lines
  labs(title = "Average Precipitation for Hurricane Season vs Frontal (1980-2016)",
       x = "Year",
       y = "Average Precipitation (mm)") +
  theme_minimal() +
  scale_fill_manual(values = c("Hurricane Season" = "skyblue", "Frontal" = "orange")) +
  scale_color_manual(values = c("Hurricane Season" = "blue", "Frontal" = "red"))
```

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

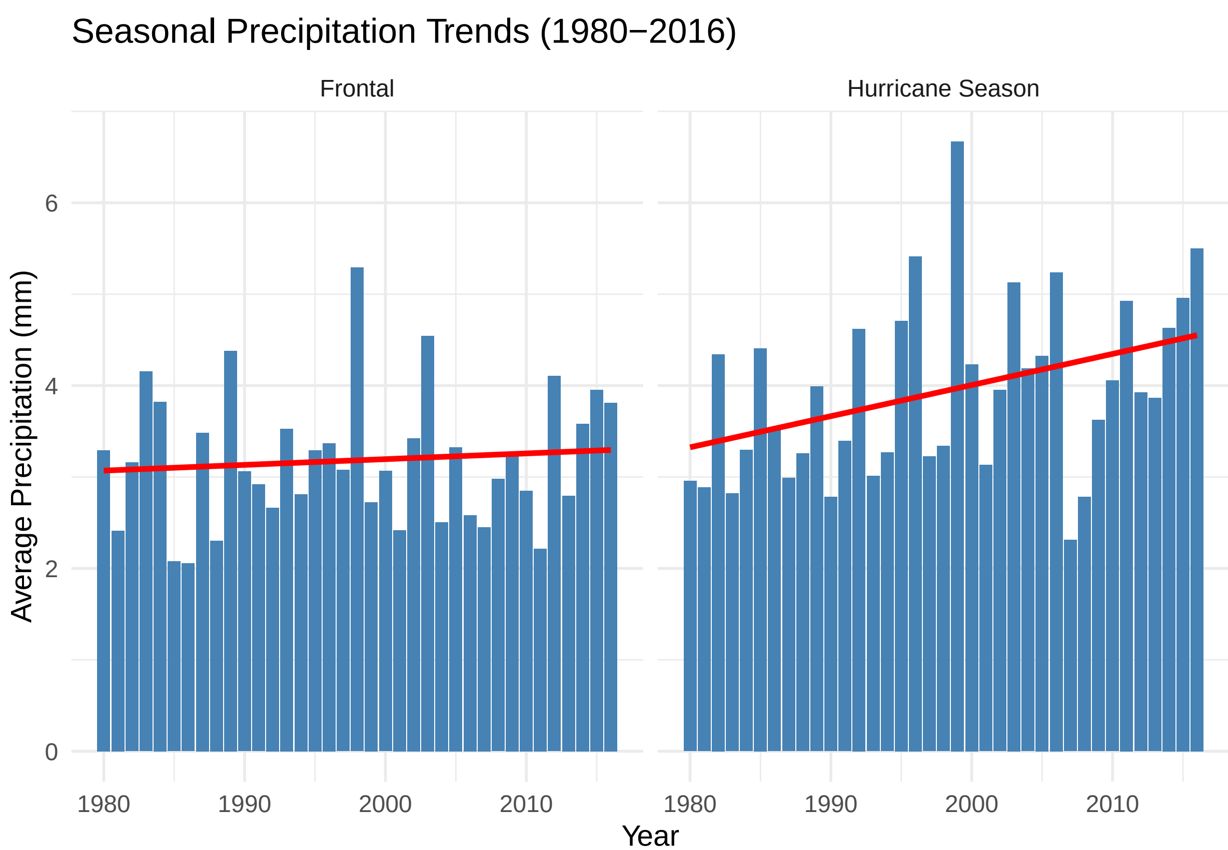


This graph makes it apparent that Hurricane Season has been steadily increasing in terms of mean precipation in the last 36 years, while frontal precipitation has been staying flat. Combining this with our previous analysis, we can assume that the majority of the annual precipitation increases are due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

Here are those graphs again, side by side:

```
ggplot(Greenville_Seasonal_Averages, 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 (1980-2016)",
       x = "Year",
       y = "Average Precipitation (mm)") +
  theme_minimal()
```

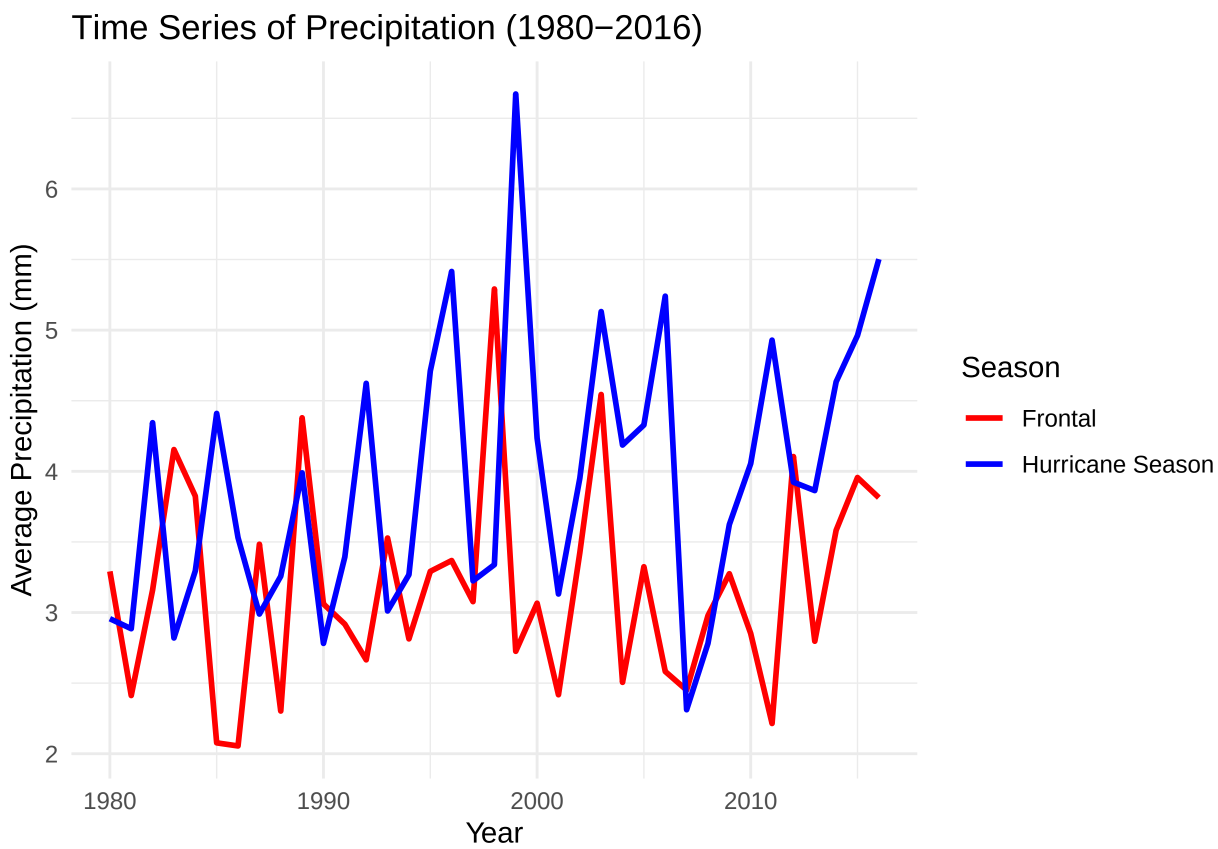
```
## `geom_smooth()` using formula = 'y ~ x'
```



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

```
ggplot(Greenville_Seasonal_Averages, aes(x = year, y = avg_precip, color = Season, group = Season)) +
  geom_line(size = 1) +
  labs(title = "Time Series of Precipitation (1980-2016)",
       x = "Year",
       y = "Average Precipitation (mm)") +
  theme_minimal() +
  scale_color_manual(values = c("Hurricane Season" = "blue", "Frontal" = "red"))
```

```
## 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.
```



Thus, we know that mean precipitation is increasing year after year, and almost entirely attributable to predictable, but heavy, hurricane systems.