

Assignment 1: Aidan Power, Fiona Price, and Emily Burghart

2024-09-18

Setting up the packages and data

Loading in the packages:

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr   1.5.1
## v ggplot2    3.5.1      v tibble    3.2.1
## v lubridate  1.9.3      v tidyr     1.3.1
## v purrr      1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(ggplot2)
library(here)
```

```
## here() starts at /home/guest/Hydrology/Assignment 1
```

```
library(dplyr)
```

Now, we will load in the Eno data and make separate dataframes for the median outflow, mean outflow, and volume.

```
#Eno data:
eno_outflows <- read.csv(
  file = here("./Data Raw/Eno_River_monthly_flow_data_through-2022.csv"),
  stringsAsFactors = TRUE)
#Average of median outflow:
eno_average_median <- eno_outflows %>%
  group_by(year) %>%
  summarise(avg_median_flow = mean(median_flow, na.rm = TRUE))
#Average of mean outflow:
eno_average_mean <- eno_outflows %>%
  group_by(year) %>%
  summarise(avg_mean_flow = mean(mean_flow, na.rm = TRUE))
#Average of volume:
eno_volume_mean <- eno_outflows %>%
  group_by(year) %>%
  summarise(avg_vol = mean(total_volume, na.rm = TRUE))
```

Now, we will load in the Flat data and make separate dataframes for the median outflow, mean outflow, and volume.

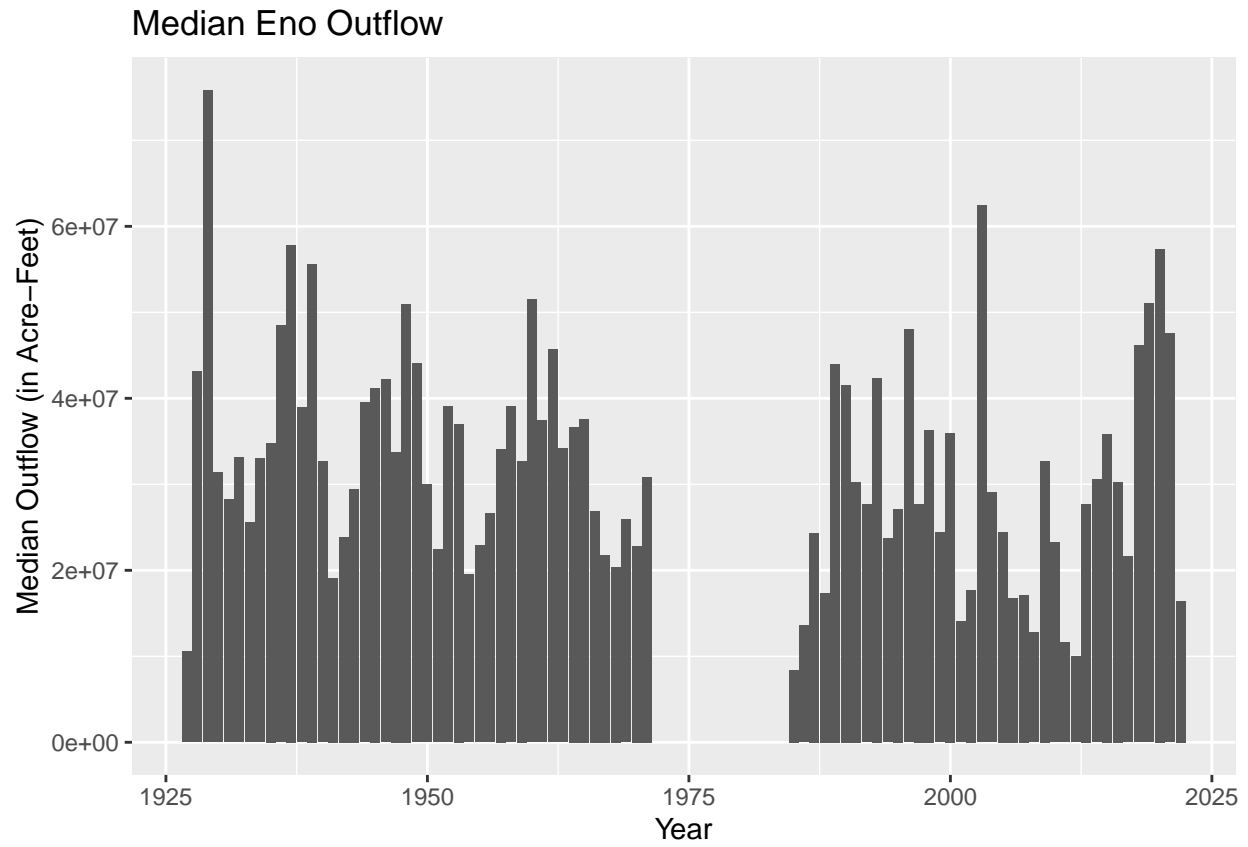
```
#Flats data:
flat_outflows <- read.csv(
  file = here("./Data Raw/Flat_River_monthly_flow_data_through-2022.csv"),
  stringsAsFactors = TRUE)
#Average of median outflow:
flat_average_median <- flat_outflows %>%
  group_by(year) %>%
  summarise(avg_median_flow = mean(median_flow, na.rm = TRUE))
#Average of mean outflow:
flat_average_mean <- flat_outflows %>%
  group_by(year) %>%
  summarise(avg_mean_flow = mean(mean_flow, na.rm = TRUE))
#Average of volume:
flat_volume_mean <- flat_outflows %>%
  group_by(year) %>%
  summarise(avg_vol = mean(total_volume, na.rm = TRUE))
```

Now, we will load in the Neuse data and make separate dataframes for the median inflow, mean inflow, and volume.

```
#Neuse data:
neuse_inflows <- read.csv(
  file = here("./Data Raw/Falls_Lake_outlet_Neuse_River_monthly_flow_data_through-2022.csv"),
  stringsAsFactors = TRUE)
#Average of median inflow:
neuse_average_median <- neuse_inflows %>%
  group_by(year) %>%
  summarise(avg_median_flow = mean(median_flow, na.rm = TRUE))
#Average of mean inflow:
neuse_average_mean <- neuse_inflows %>%
  group_by(year) %>%
  summarise(avg_mean_flow = mean(mean_flow, na.rm = TRUE))
#Average of volume:
neuse_volume_mean <- neuse_inflows %>%
  group_by(year) %>%
  summarise(avg_vol = mean(total_volume, na.rm = TRUE))
```

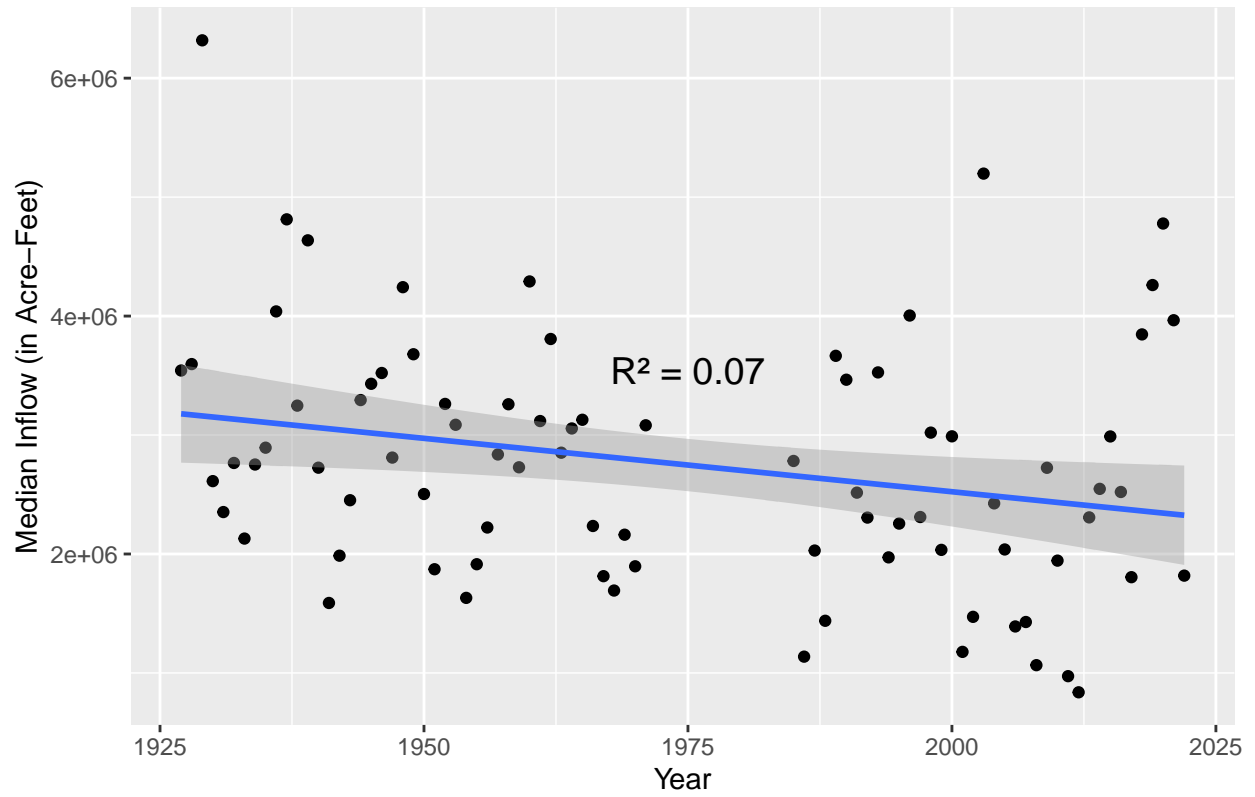
Plots

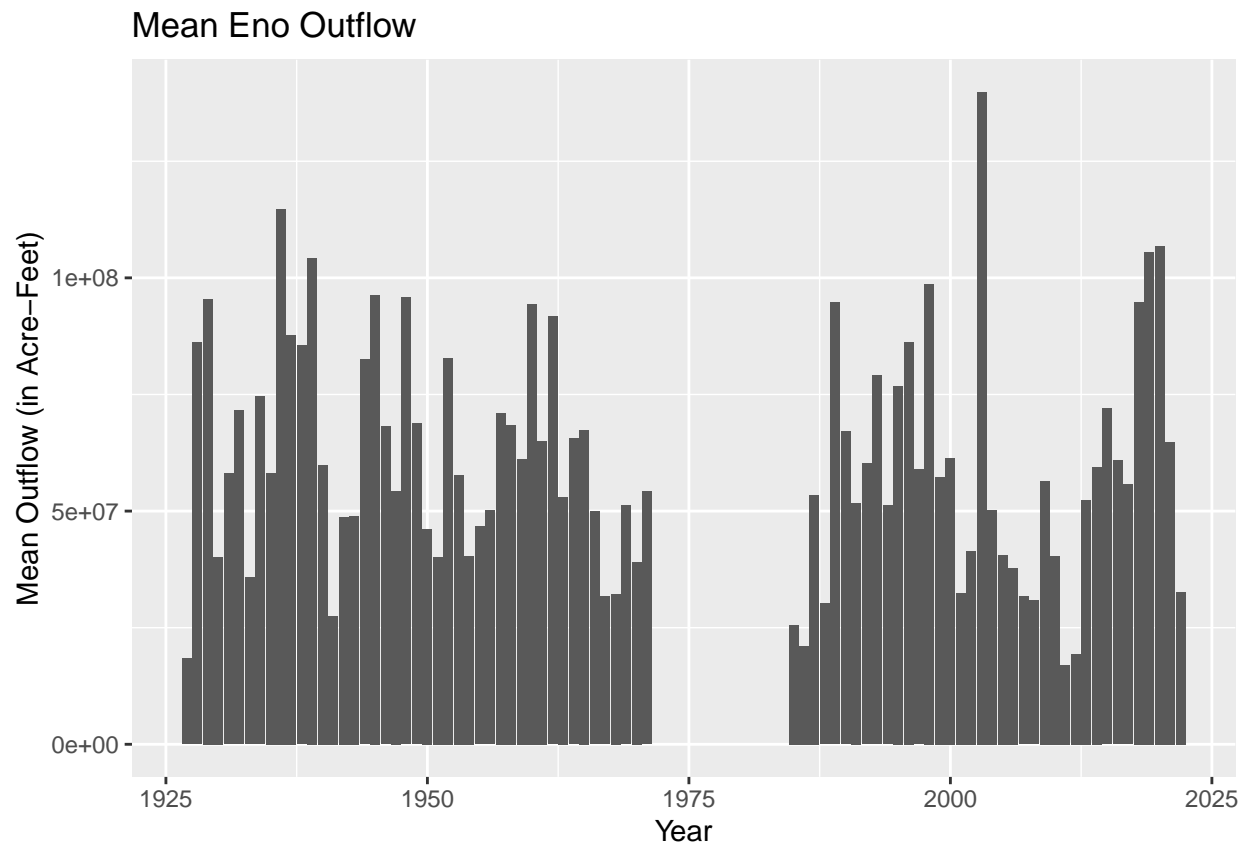
Now, we will start making graphs. First, we will look at Eno. We will start with all years.



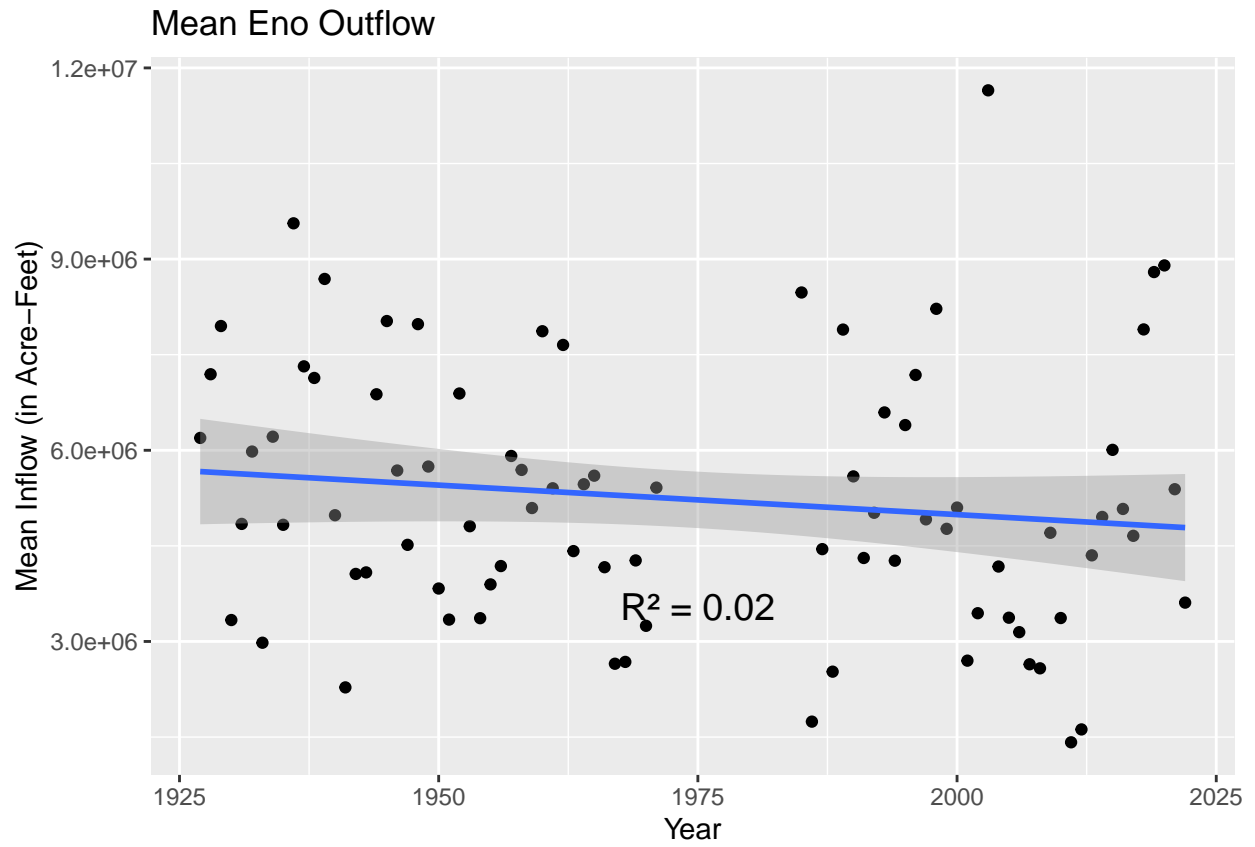
```
## 'geom_smooth()' using formula = 'y ~ x'
```

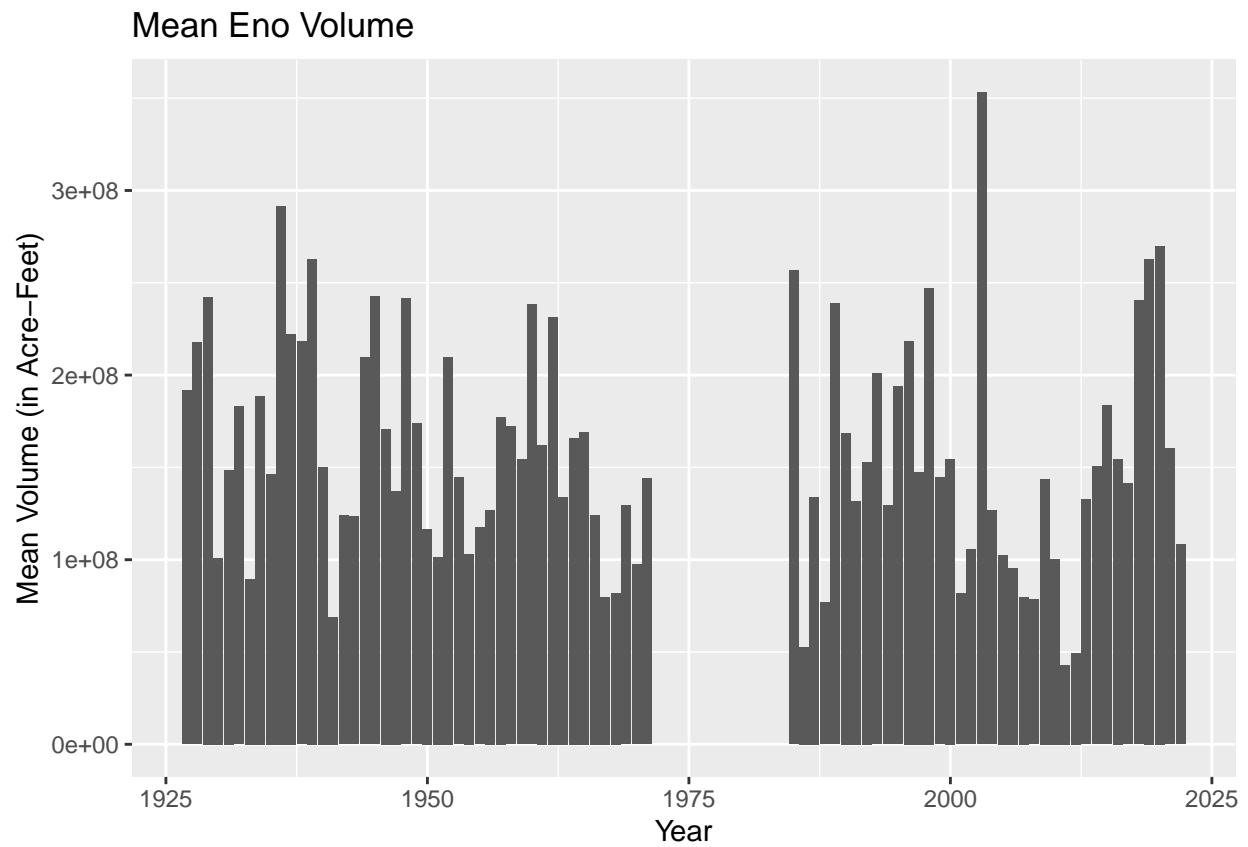
Median Eno Outflow



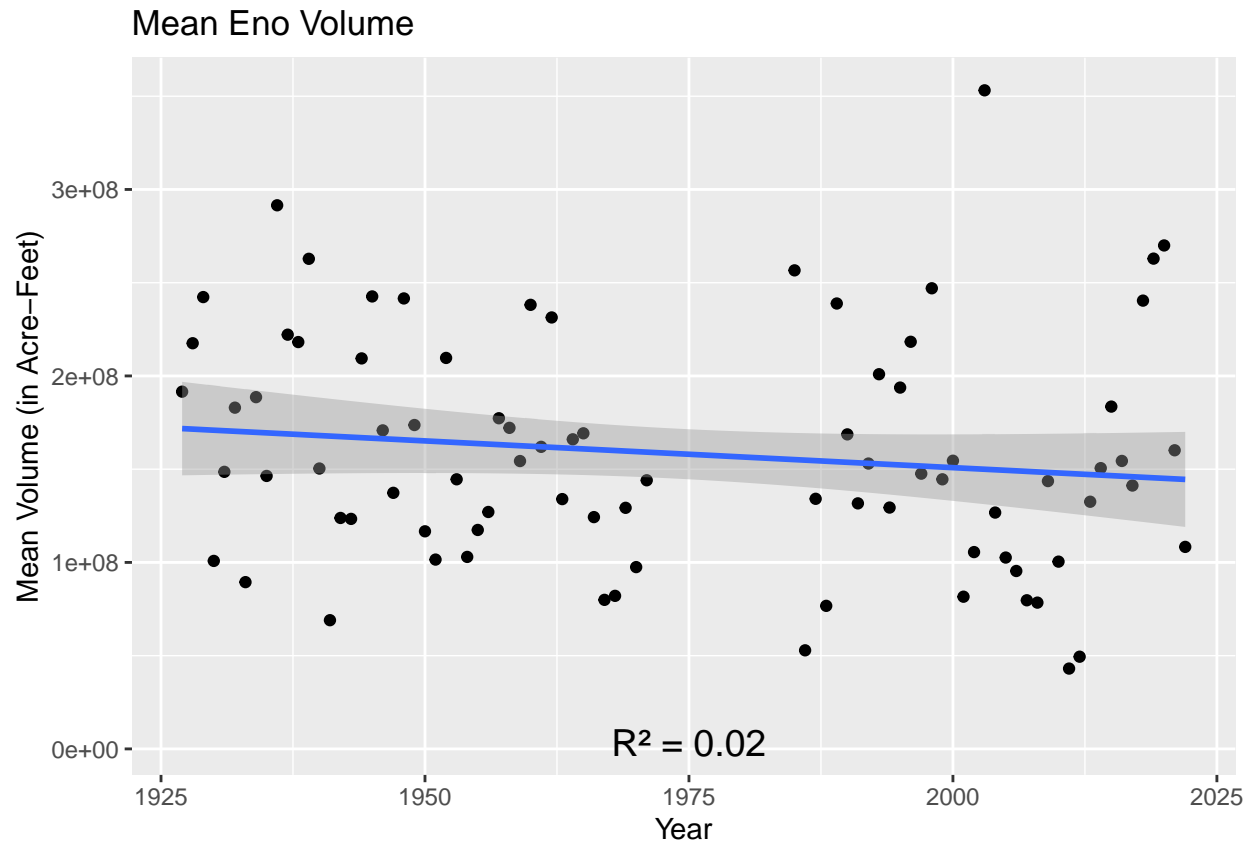


```
## 'geom_smooth()' using formula = 'y ~ x'
```





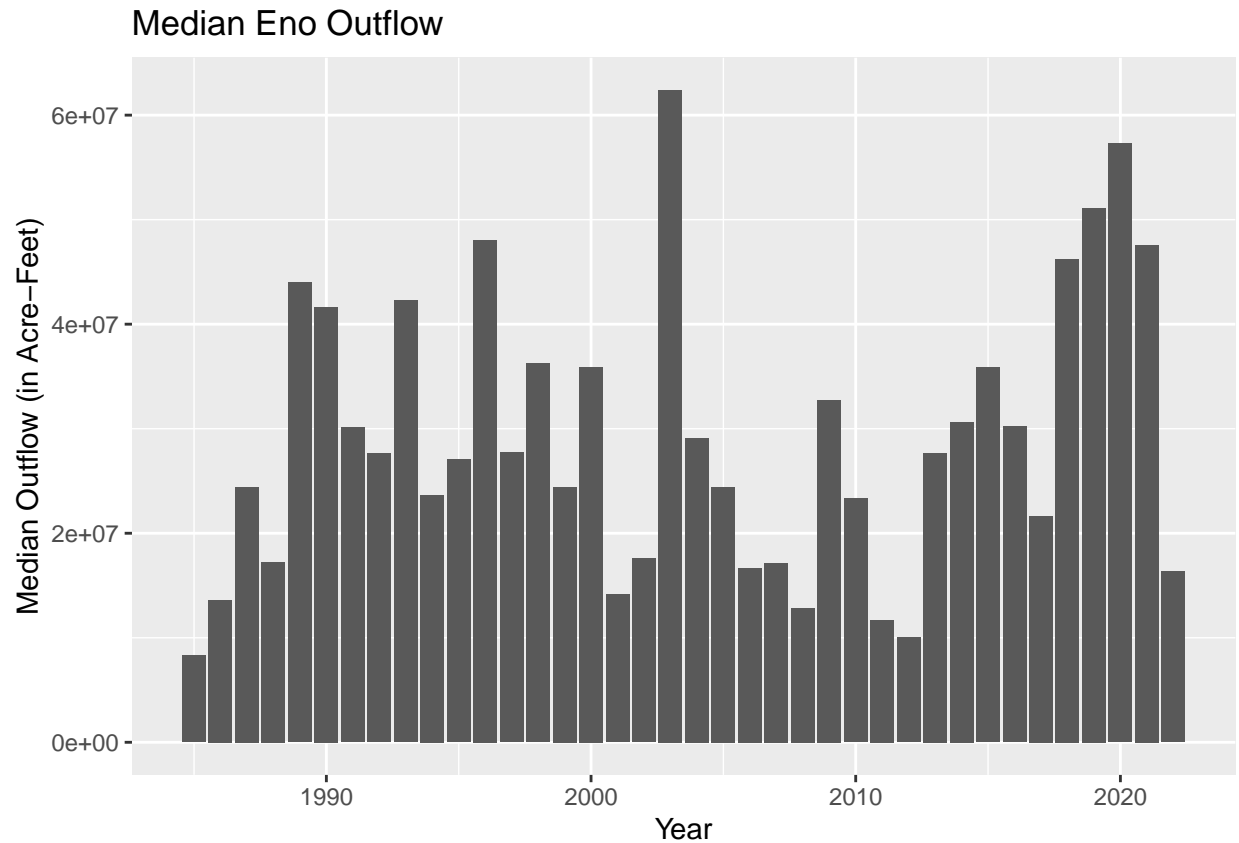
```
## 'geom_smooth()' using formula = 'y ~ x'
```



There is a gap in Eno data from 1971-1985. Now, we will look only at the post-1985 data.

```
#Filter out the data to only include 1985 and beyond.
eno_filtered <- eno_outflows %>%
  filter(year >= 1985)
#Filter out the average median outflow data to only include 1985 and beyond.
eno_avemedian_filtered <- eno_average_median %>%
  filter(year >= 1985)
#Filter out the average mean outflow data to only include 1985 and beyond.
eno_avemean_filtered <- eno_average_mean %>%
  filter(year >= 1985)
#Filter out the average volume data to only include 1985 and beyond.
eno_vol_filtered <- eno_volume_mean %>%
  filter(year >= 1985)

#Median Eno outflow
ggplot(data = (eno_filtered)) +
  geom_col(data = eno_filtered, aes(x = year, y = median_flow)) +
  labs(title = "Median Eno Outflow", x = "Year", y = "Median Outflow (in Acre-Feet)")
```

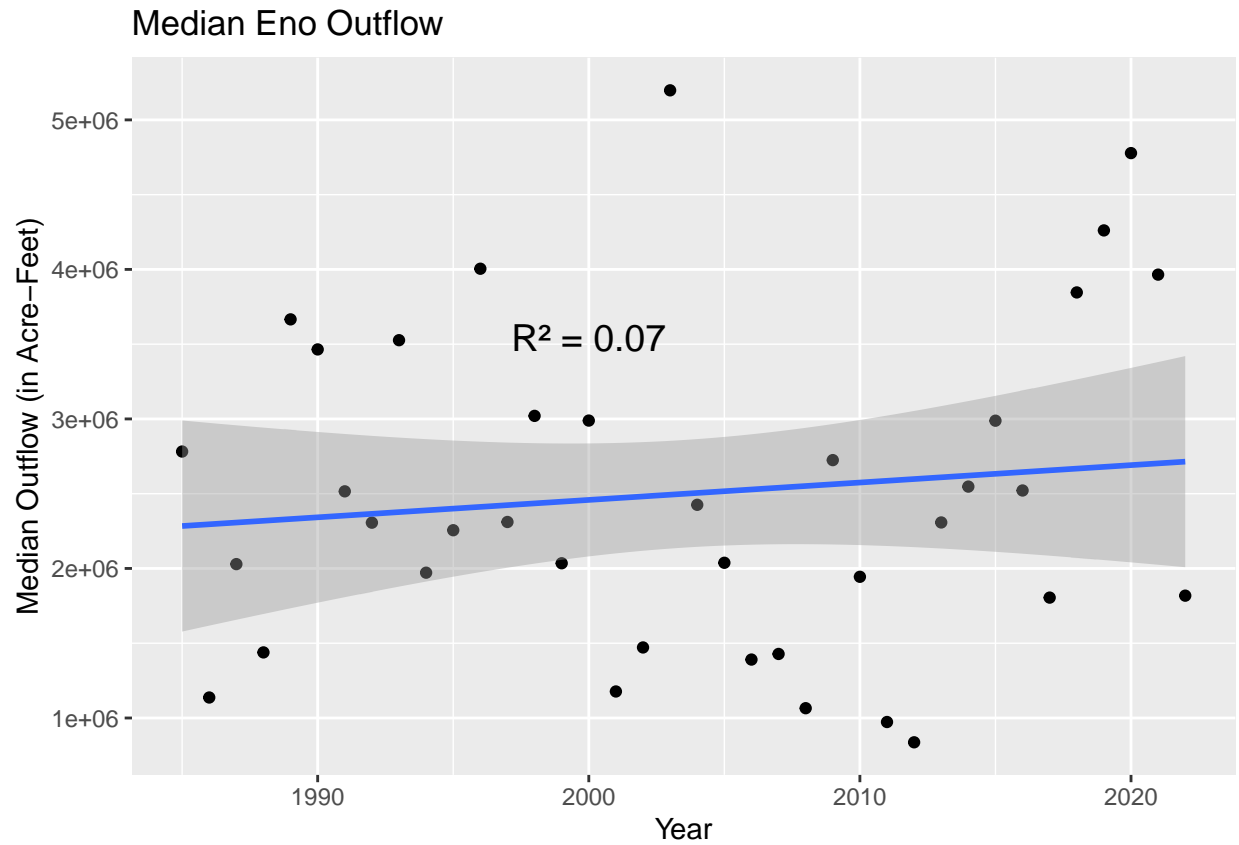
```

eno_med_model_1985 <- lm(year ~ avg_median_flow, data = eno_average_median)
r_squared_4 <- summary(eno_med_model_1985)$r.squared

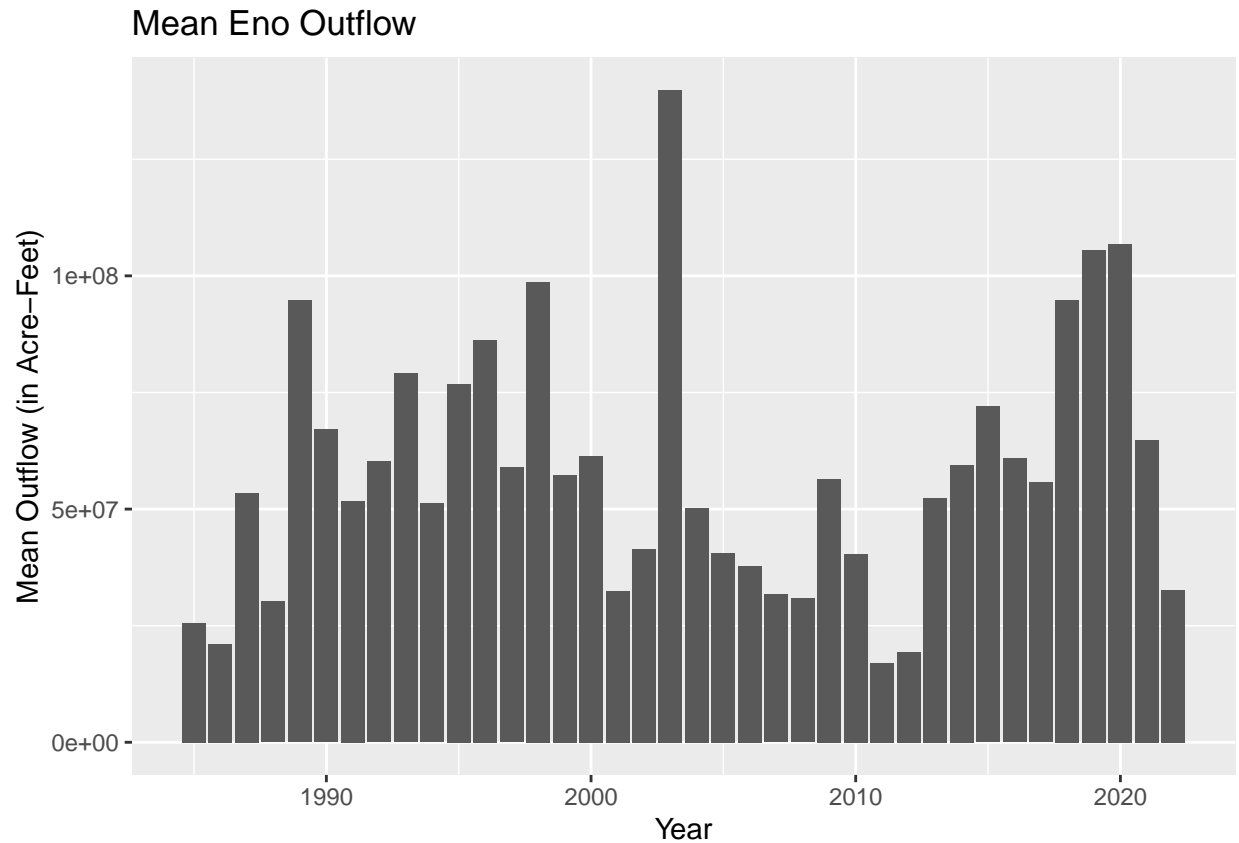
ggplot(data = eno_avemedian_filtered, aes(x = year, y = avg_median_flow)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Median Eno Outflow", x = "Year", y = "Median Outflow (in Acre-Feet)") +
  annotate("text", x = 2000, y = 3542400,
    label = paste("R² =", round(r_squared_4, 2)),
    size = 5)

```

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



```
#Mean Eno outflow  
ggplot(data = eno_filtered) +  
  geom_col(data = eno_filtered, aes(x = year, y = mean_flow)) +  
  labs(title = "Mean Eno Outflow", x = "Year", y = "Mean Outflow (in Acre-Feet)")
```



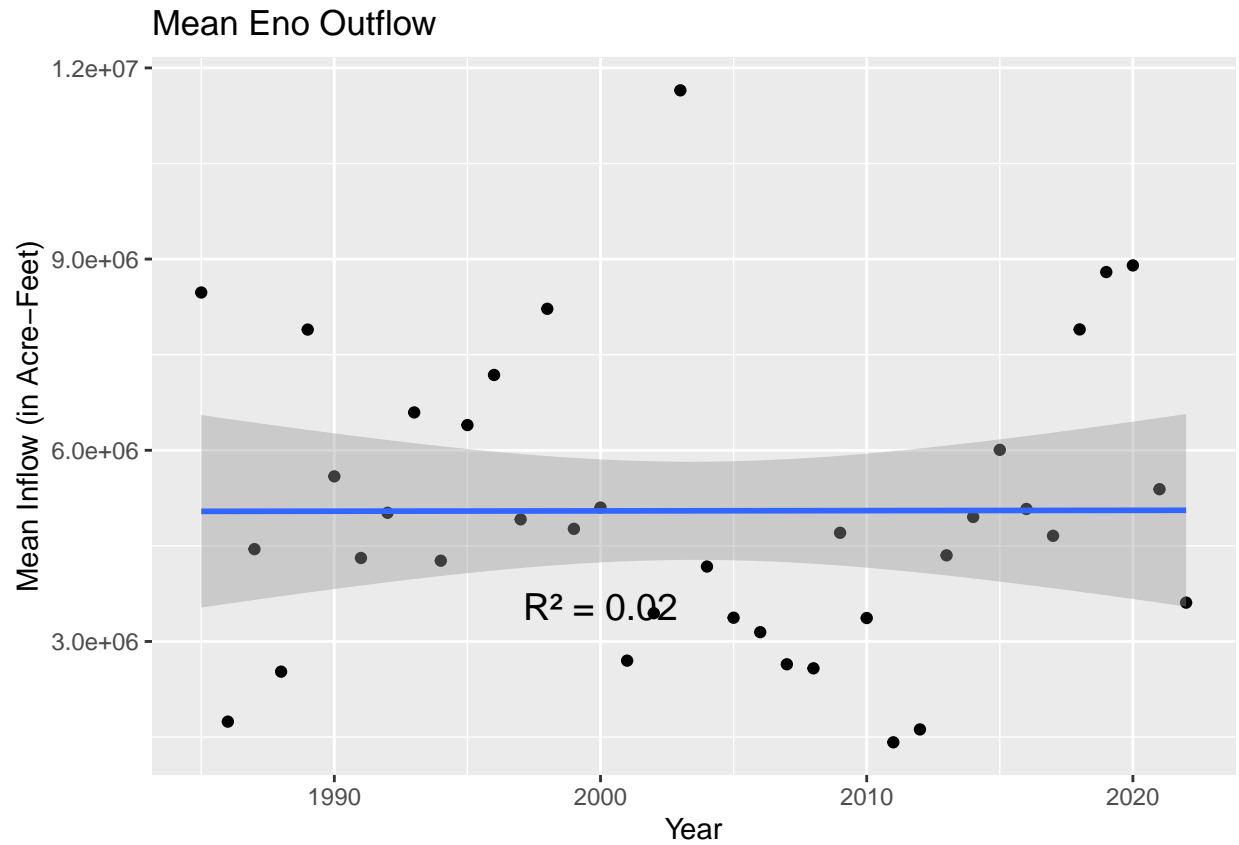
```

eno_mean_model_1985 <- lm(year ~ avg_mean_flow, data = eno_average_mean)
r_squared_5 <- summary(eno_mean_model_1985)$r.squared

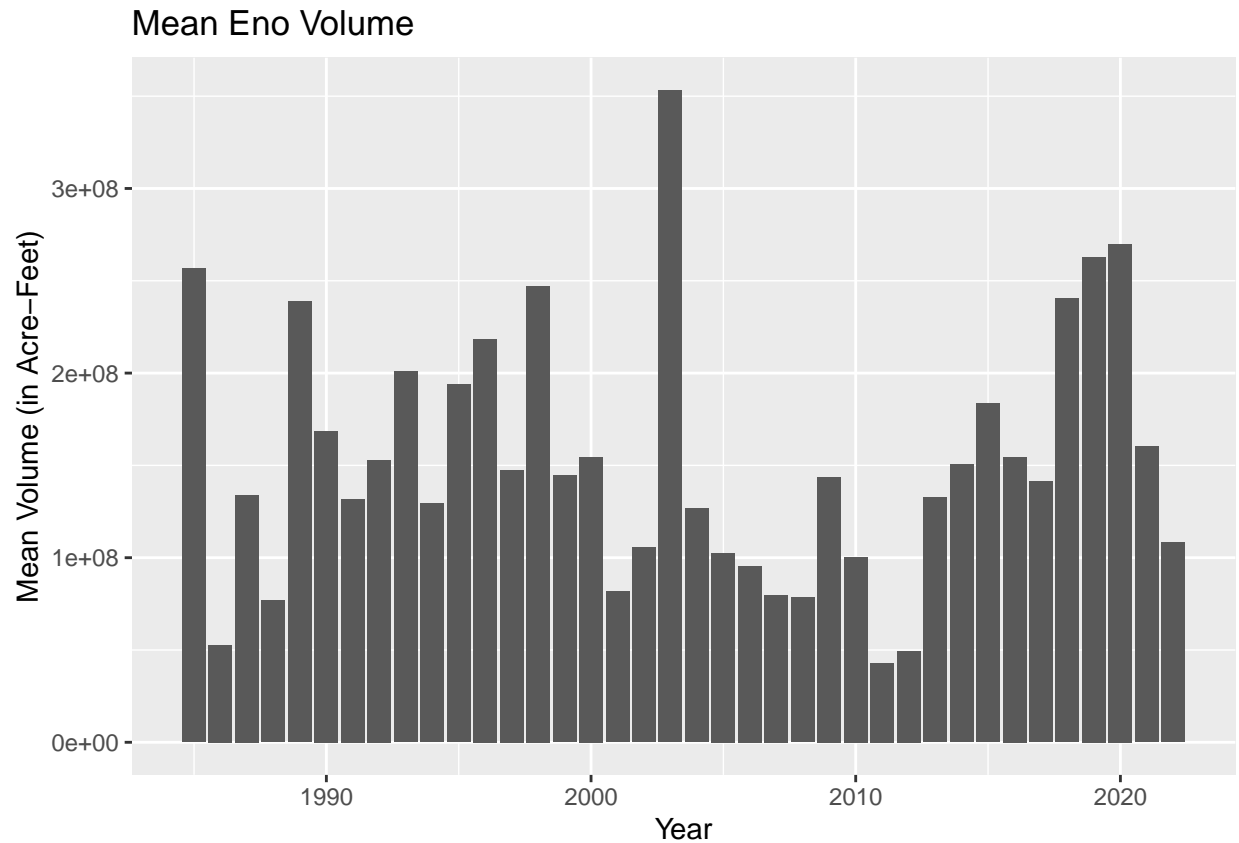
ggplot(data = eno_avemean_filtered, aes(x = year, y = avg_mean_flow)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Mean Eno Outflow", x = "Year", y = "Mean Inflow (in Acre-Feet)") +
  annotate("text", x = 2000, y = 3542400,
    label = paste("R² =", round(r_squared_5, 2)),
    size = 5)

```

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



```
#Mean Eno volume  
ggplot(data = eno_vol_filtered) +  
  geom_col(data = eno_vol_filtered, aes(x = year, y = avg_vol)) +  
  labs(title = "Mean Eno Volume", x = "Year", y = "Mean Volume (in Acre-Feet)")
```



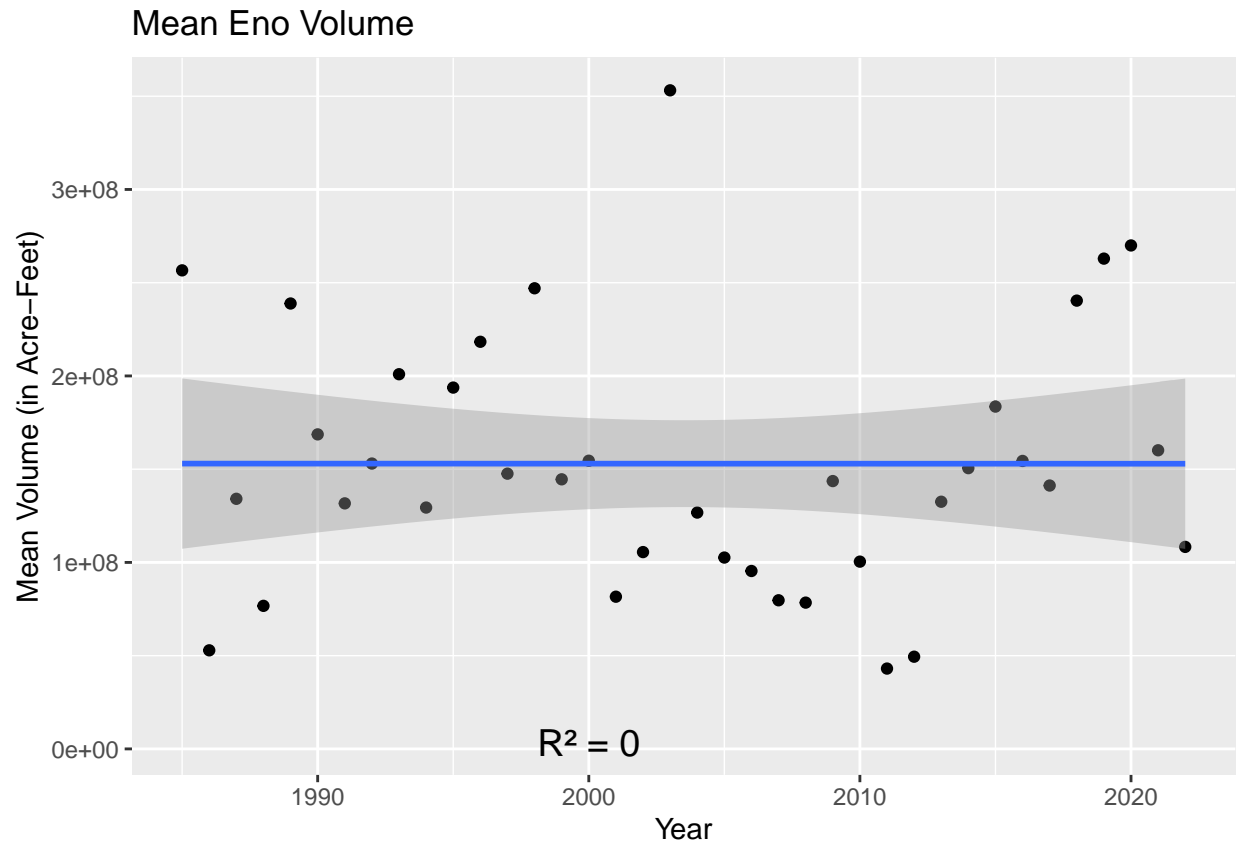
```

eno_vol_model_1985 <- lm(year ~ avg_vol, data = eno_vol_filtered)
r_squared_6 <- summary(eno_vol_model_1985)$r.squared

ggplot(data = eno_vol_filtered, aes(x = year, y = avg_vol)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Mean Eno Volume", x = "Year", y = "Mean Volume (in Acre-Feet)") +
  annotate("text", x = 2000, y = 3542400,
    label = paste("R² =", round(r_squared_6, 2)),
    size = 5)

```

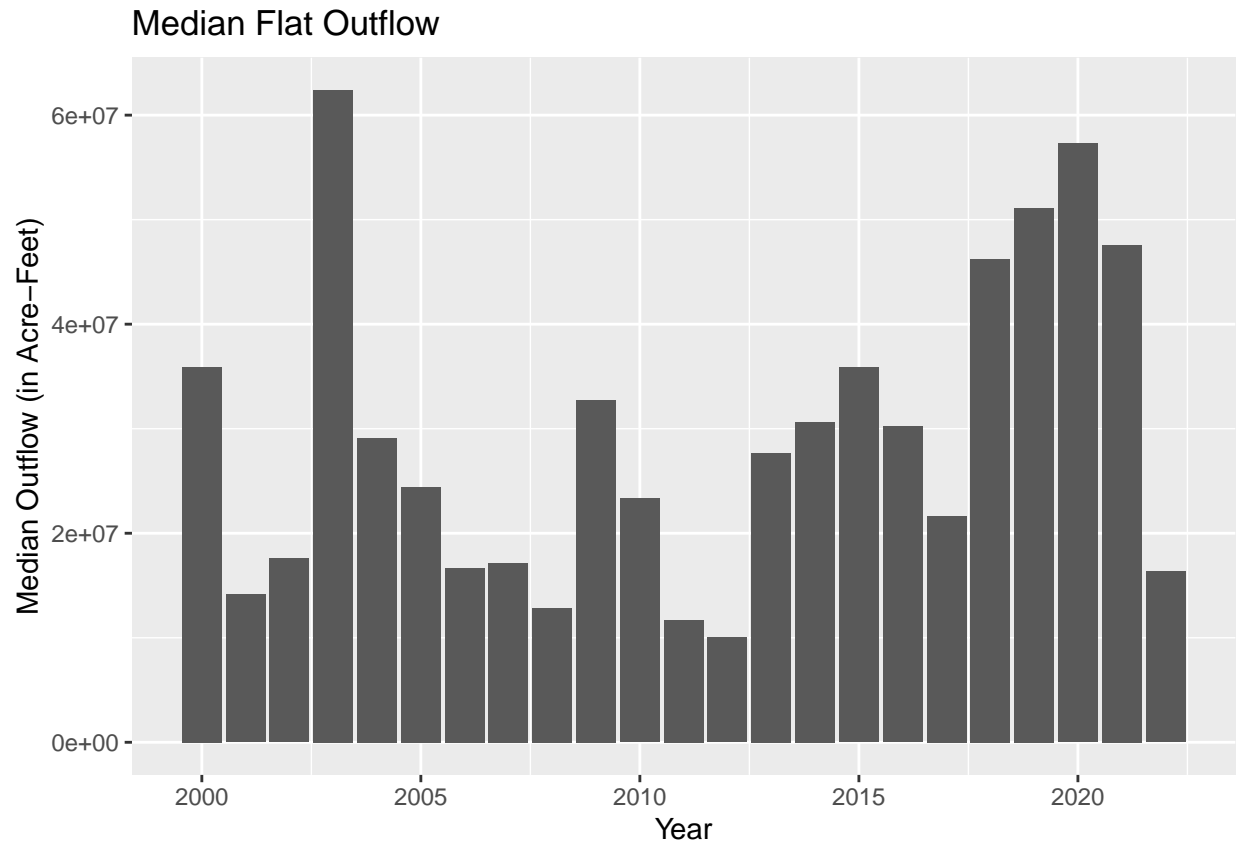
```
## 'geom_smooth()' using formula = 'y ~ x'
```



Lastly, we will look at only years 2000 and beyond to capture recent trends.

```
#Filter out the data to only include 2000 and beyond.
eno_2000 <- eno_outflows %>%
  filter(year >= 2000)
#Filter out the average median outflow data to only include 2000 and beyond.
eno_2000_median <- eno_average_median %>%
  filter(year >= 2000)
#Filter out the average mean outflow data to only include 2000 and beyond.
eno_2000_mean <- eno_average_mean %>%
  filter(year >= 2000)
#Filter out the average volume data to only include 2000 and beyond.
eno_2000_volume <- eno_volume_mean %>%
  filter(year >= 2000)

#Median Eno outflow
ggplot(data = eno_2000) +
  geom_col(data = eno_2000, aes(x = year, y = median_flow)) +
  labs(title = "Median Flat Outflow", x = "Year", y = "Median Outflow (in Acre-Feet)")
```



```

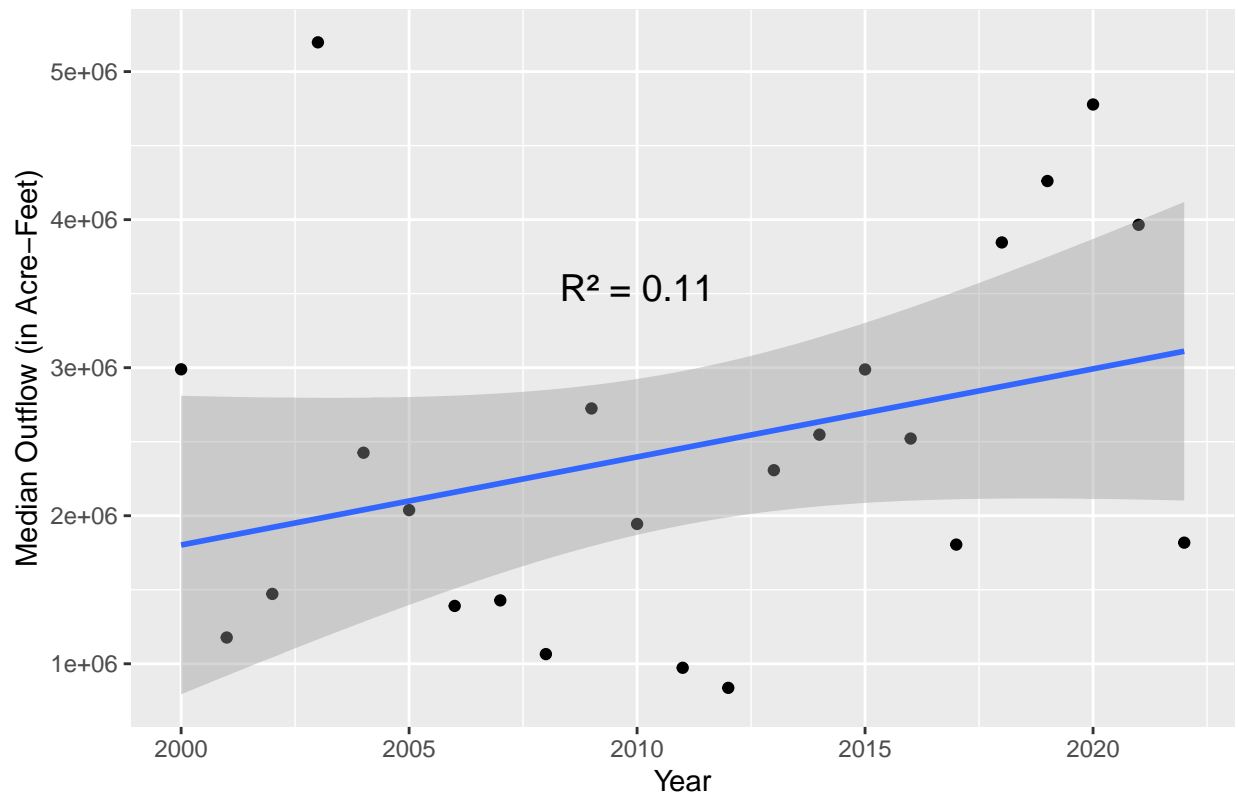
eno_med_model_2000 <- lm(year ~ avg_median_flow, data = eno_2000_median)
r_squared_7 <- summary(eno_med_model_2000)$r.squared

ggplot(data = eno_2000_median, aes(x = year, y = avg_median_flow)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Median Flat Outflow", x = "Year", y = "Median Outflow (in Acre-Feet)") +
  annotate("text", x = 2010, y = 3542400,
    label = paste("R² =", round(r_squared_7, 2)),
    size = 5)

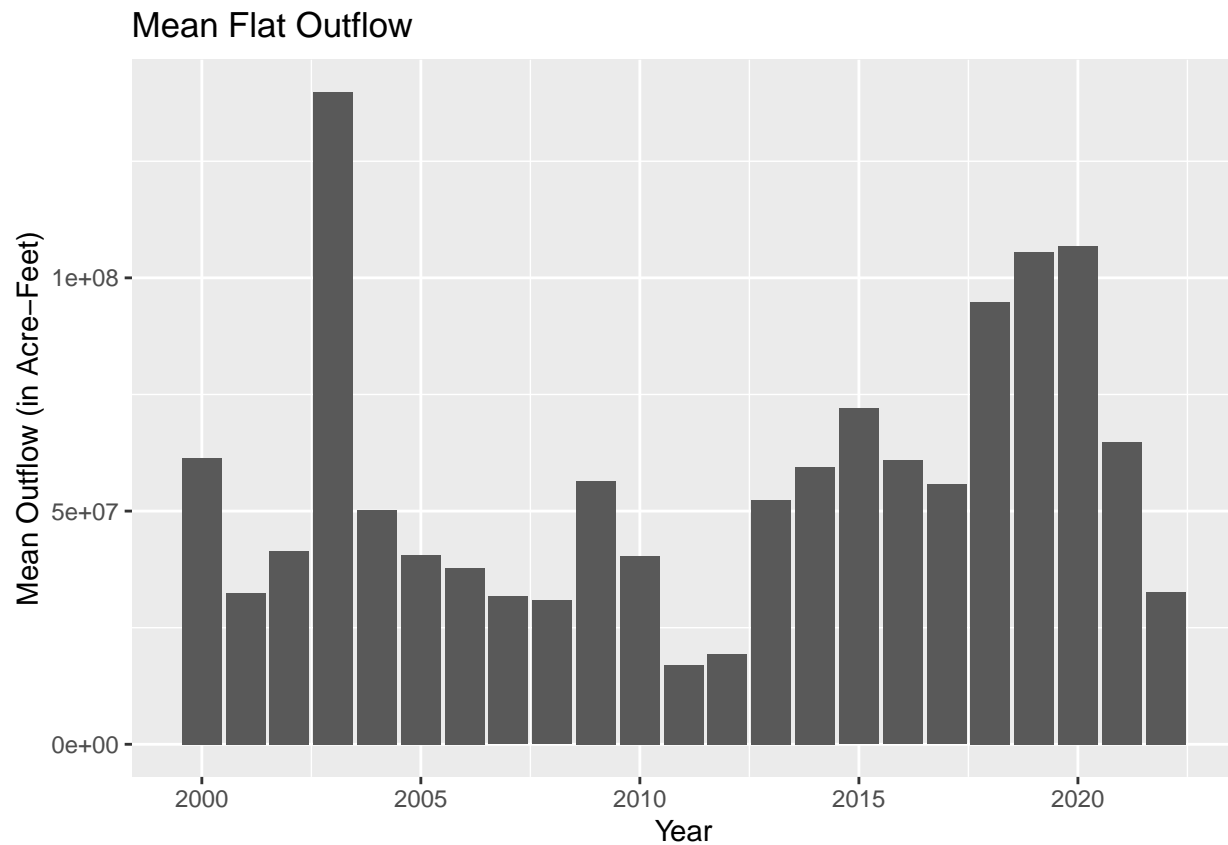
```

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

Median Flat Outflow



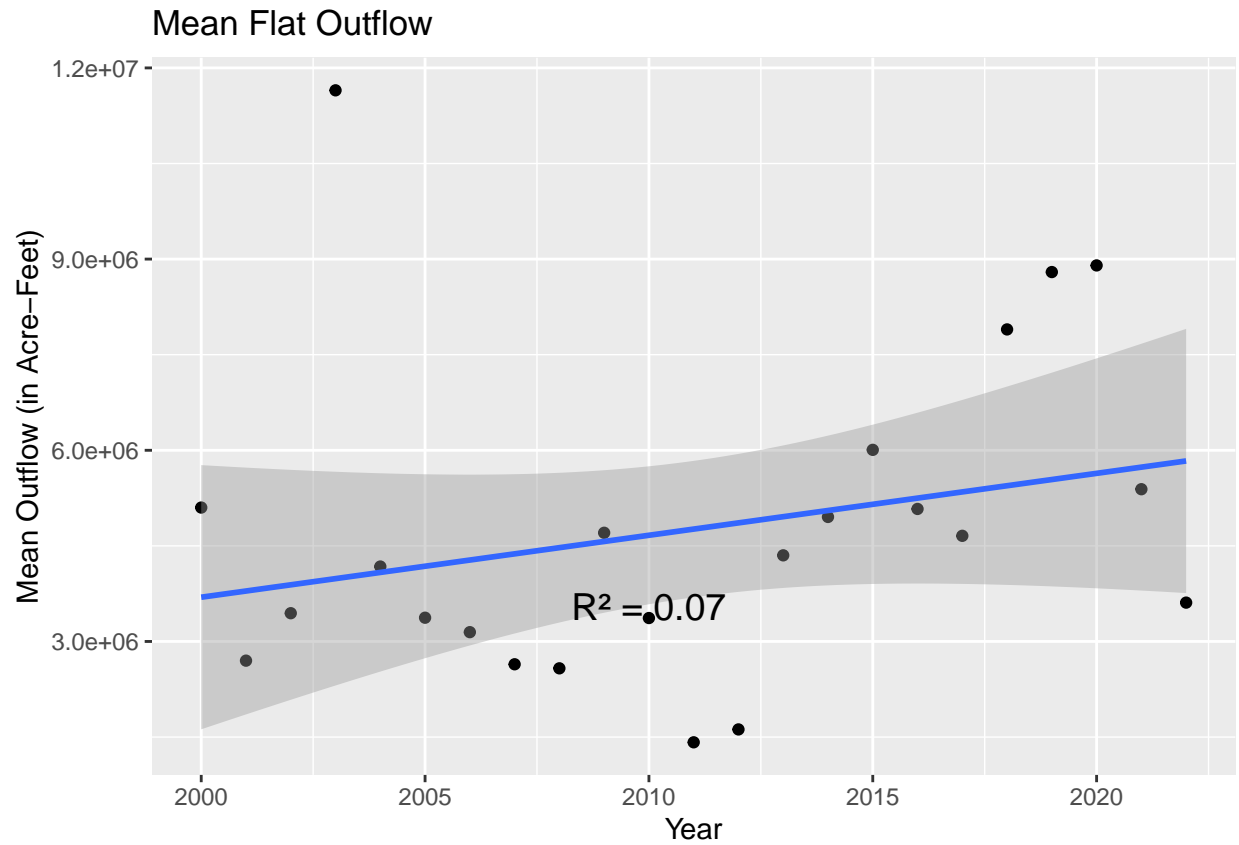
```
#Mean Eno outflow
ggplot(data = eno_2000) +
  geom_col(data = eno_2000, aes(x = year, y = mean_flow)) +
  labs(title = "Mean Flat Outflow", x = "Year", y = "Mean Outflow (in Acre-Feet)")
```

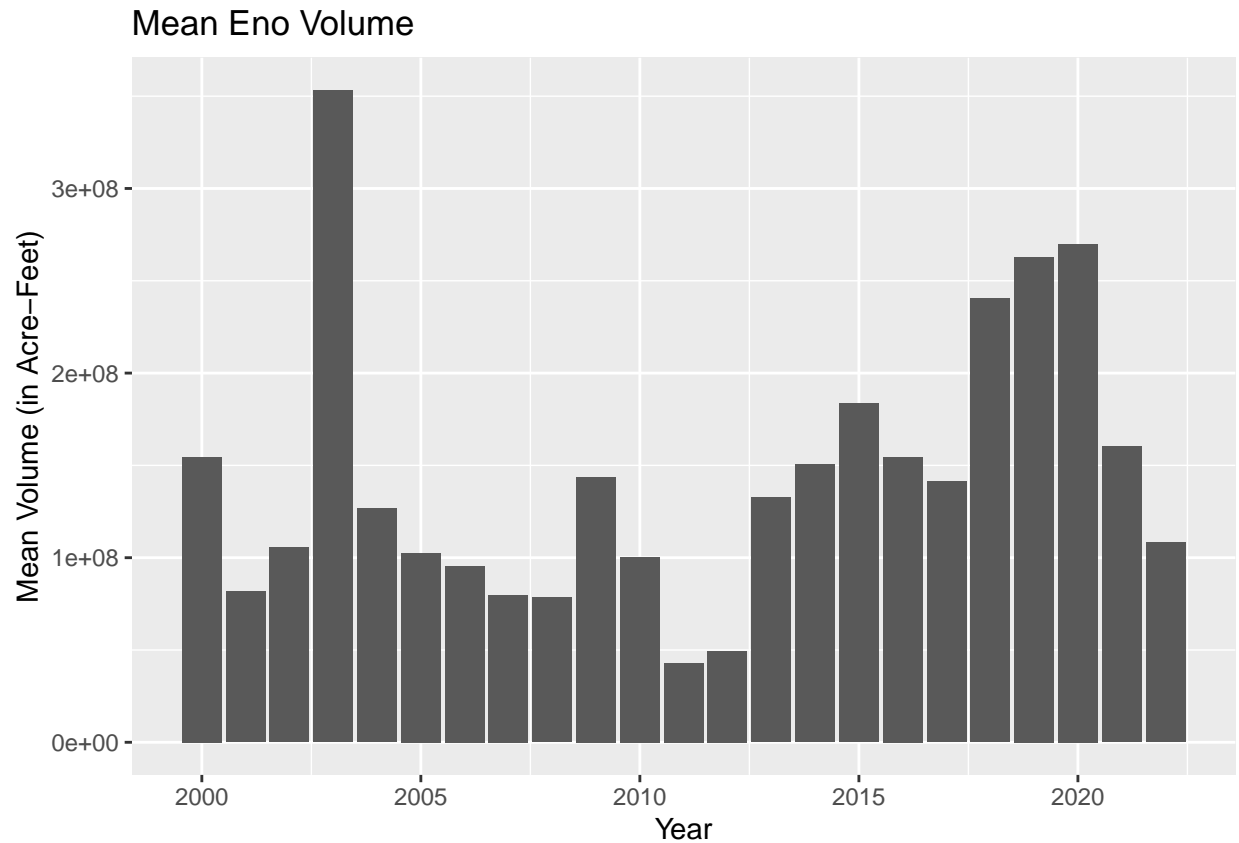
```
eno_mean_model_2000 <- lm(year ~ avg_mean_flow, data = eno_2000_mean)
r_squared_8 <- summary(eno_mean_model_2000)$r.squared
```

```
ggplot(data = eno_2000_mean, aes(x = year, y = avg_mean_flow)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Mean Flat Outflow", x = "Year", y = "Mean Outflow (in Acre-Feet)") +
  annotate("text", x = 2010, y = 3542400,
    label = paste("R² =", round(r_squared_8, 2)),
    size = 5)
```

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



```
#Mean Eno volume  
ggplot(data = eno_2000_volume) +  
  geom_col(data = eno_2000_volume, aes(x = year, y = avg_vol)) +  
  labs(title = "Mean Eno Volume", x = "Year", y = "Mean Volume (in Acre-Feet)")
```



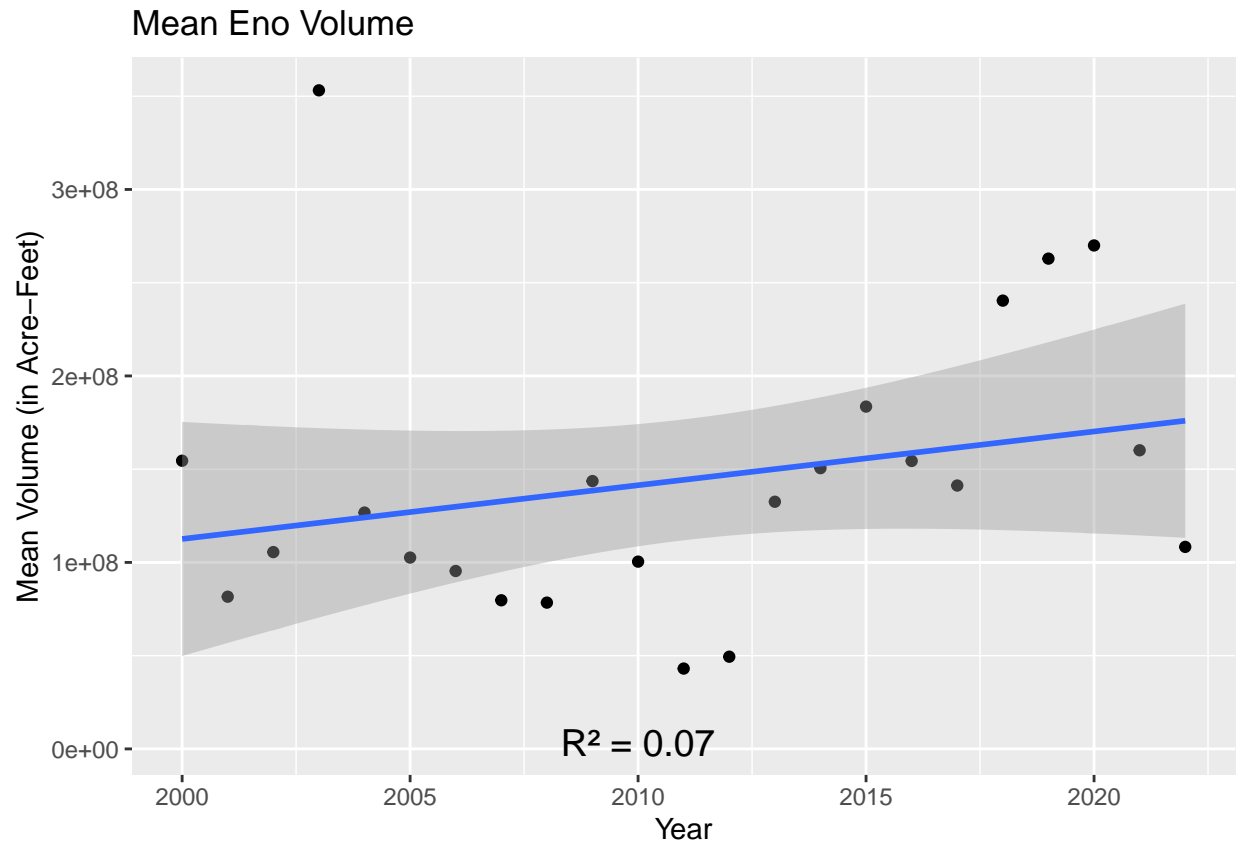
```

eno_vol_model_2000 <- lm(year ~ avg_vol, data = eno_2000_volume)
r_squared_9 <- summary(eno_vol_model_2000)$r.squared

ggplot(data = eno_2000_volume, aes(x = year, y = avg_vol)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Mean Eno Volume", x = "Year", y = "Mean Volume (in Acre-Feet)") +
  annotate("text", x = 2010, y = 3542400,
    label = paste("R² =", round(r_squared_9, 2)),
    size = 5)

```

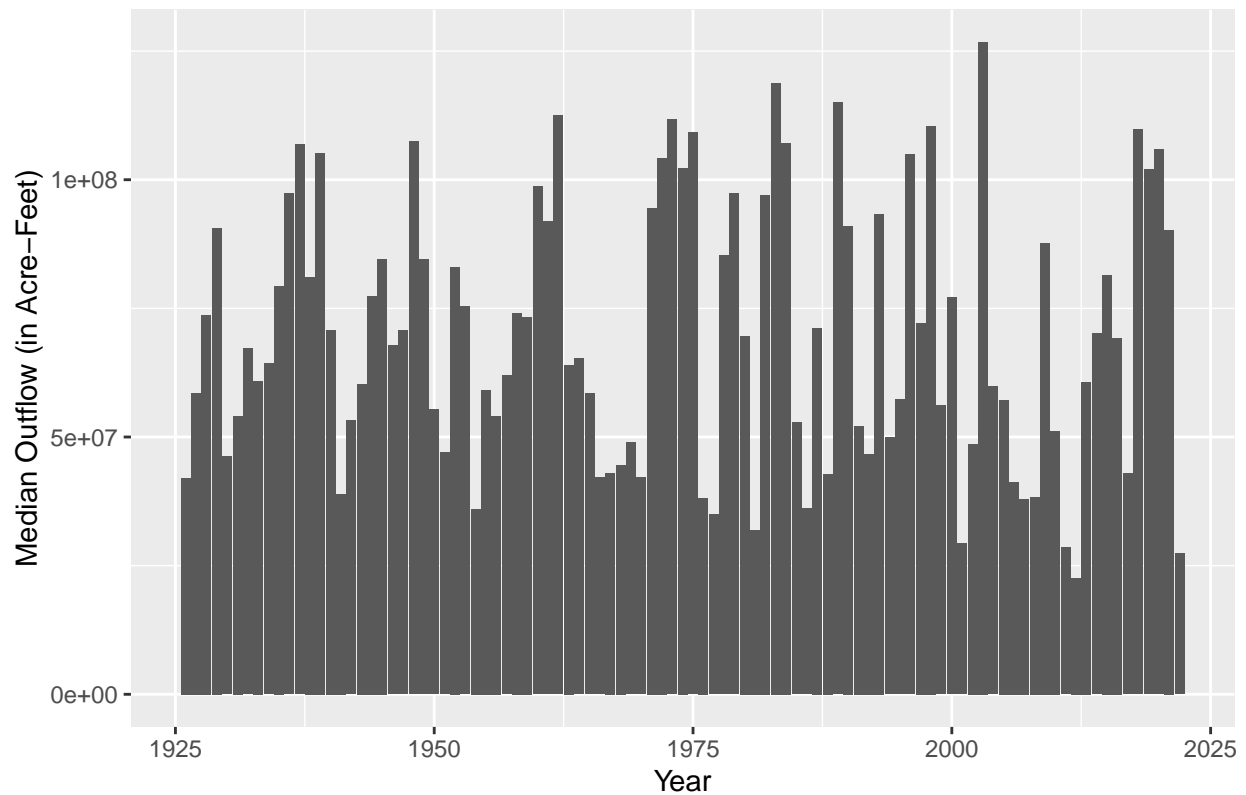
```
## 'geom_smooth()' using formula = 'y ~ x'
```



Now, we will look at Flat data.

```
#Median Flat outflow
ggplot(data = (flat_outflows)) +
  geom_col(data = flat_outflows, aes(x = year, y = median_flow)) +
  labs(title = "Median Flat Outflow", x = "Year", y = "Median Outflow (in Acre-Feet)")
```

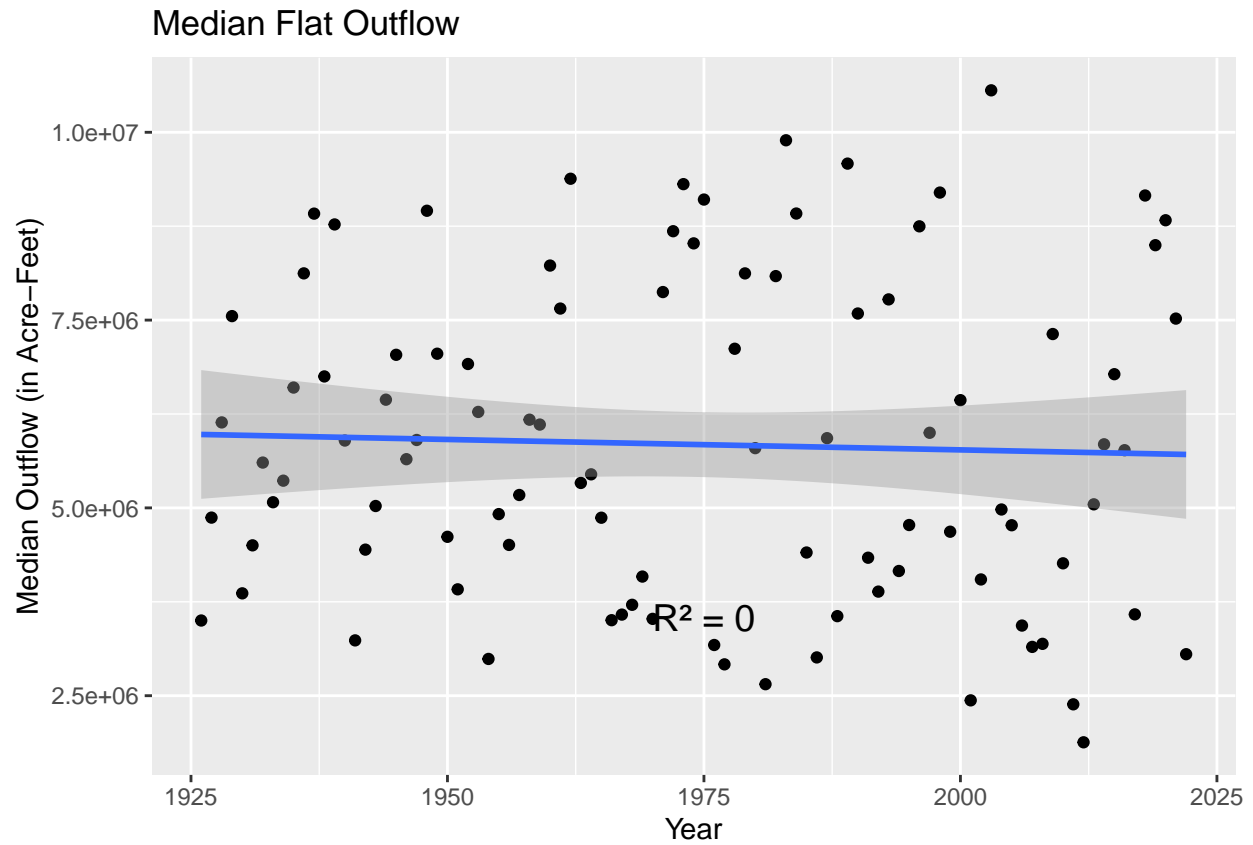
Median Flat Outflow



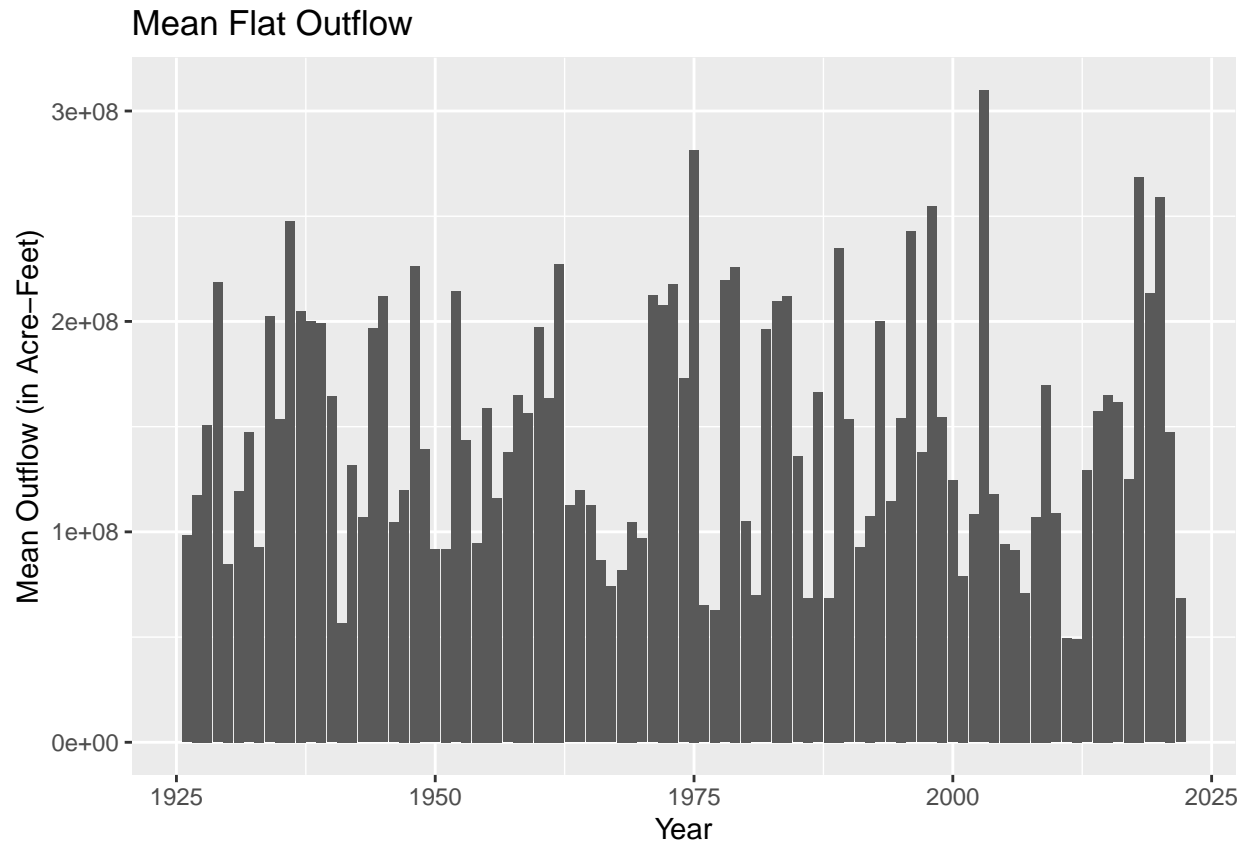
```
flat_med_model <- lm(year ~ avg_median_flow, data = flat_average_median)
flat_r <- summary(flat_med_model)$r.squared

ggplot(data = flat_average_median, aes(x = year, y = avg_median_flow)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Median Flat Outflow", x = "Year", y = "Median Outflow (in Acre-Feet)") +
  annotate("text", x = 1975, y = 3542400,
         label = paste("R2 =", round(flat_r, 2)),
         size = 5)
```

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



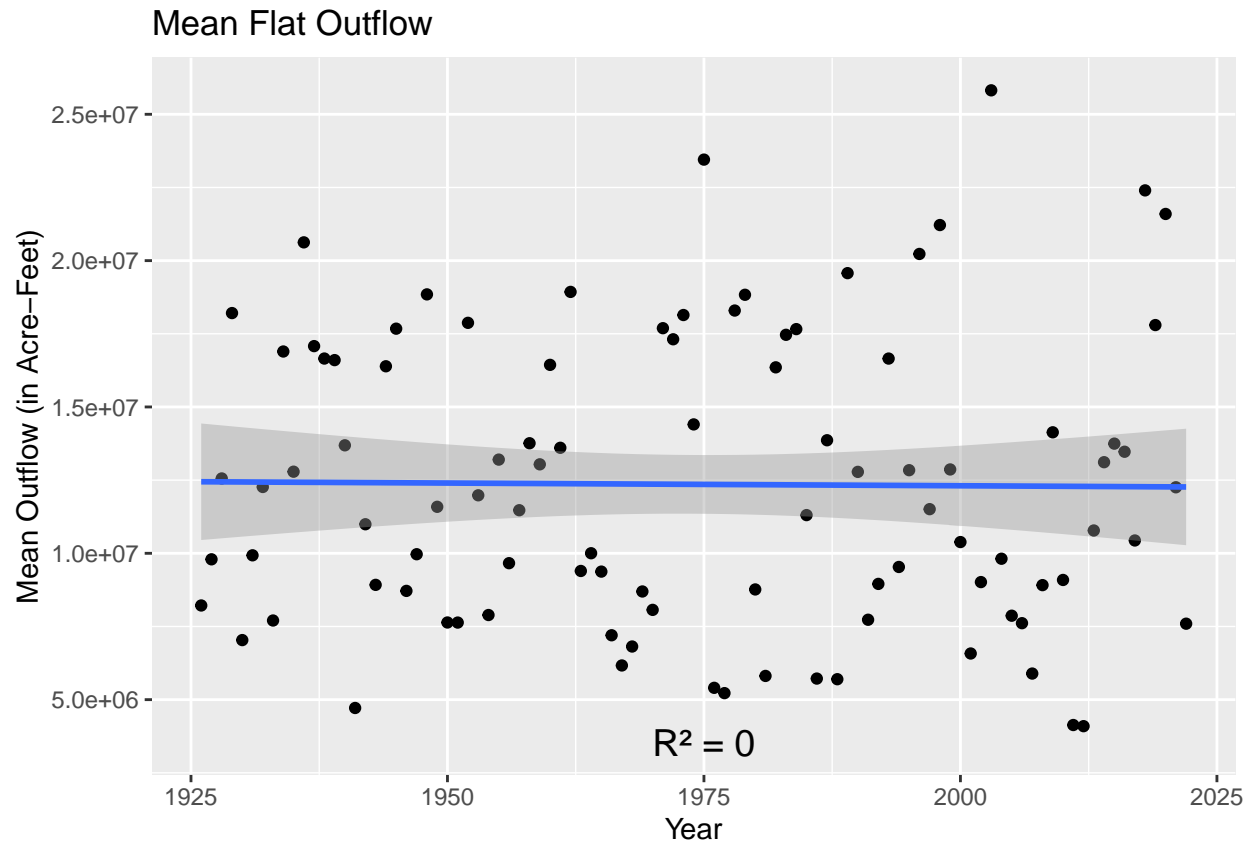
```
#Mean Flat outflow  
ggplot(data = (flat_outflows)) +  
  geom_col(data = flat_outflows, aes(x = year, y = mean_flow)) +  
  labs(title = "Mean Flat Outflow", x = "Year", y = "Mean Outflow (in Acre-Feet)")
```



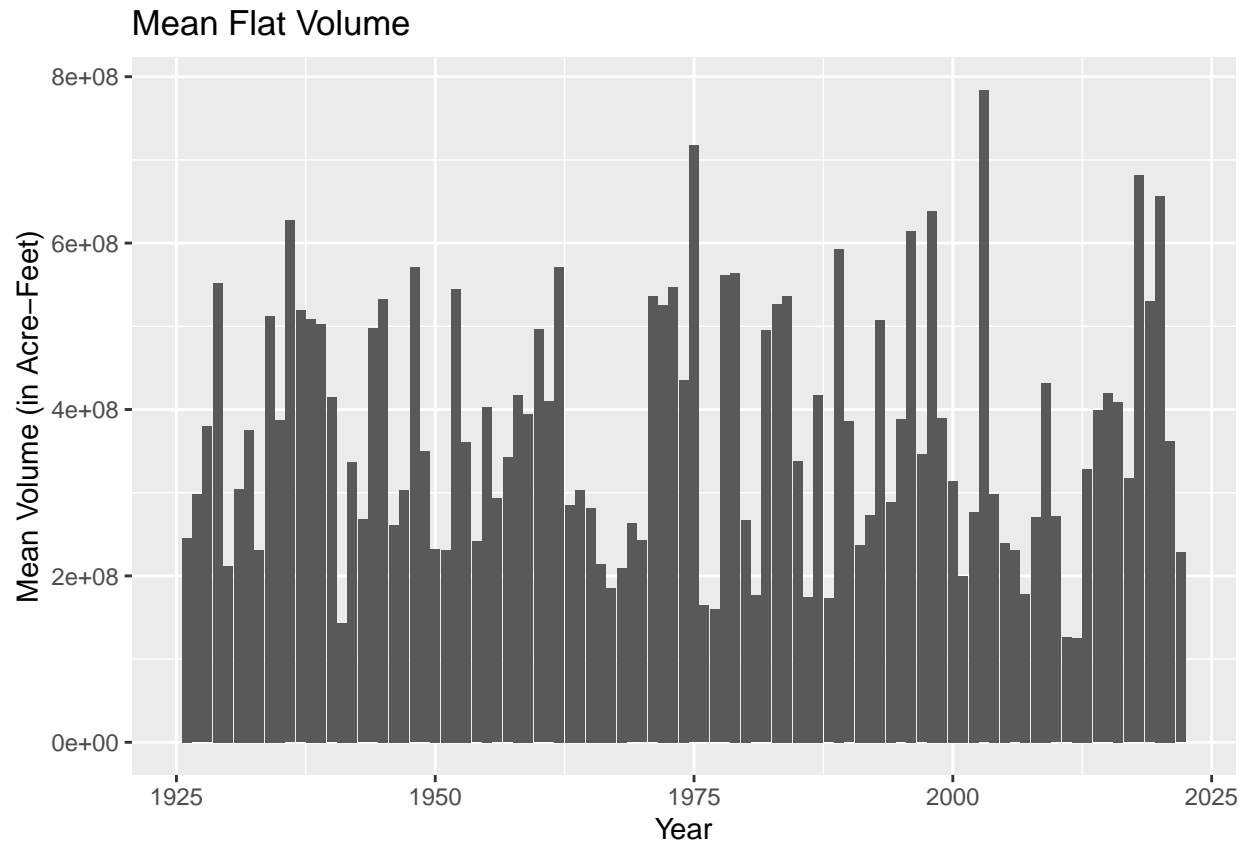
```
flat_mean_model <- lm(year ~ avg_mean_flow, data = flat_average_mean)
flat_r2 <- summary(flat_mean_model)$r.squared
```

```
ggplot(data = flat_average_mean, aes(x = year, y = avg_mean_flow)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Mean Flat Outflow", x = "Year", y = "Mean Outflow (in Acre-Feet)") +
  annotate("text", x = 1975, y = 3542400,
    label = paste("R² =", round(flat_r2, 2)),
    size = 5)
```

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



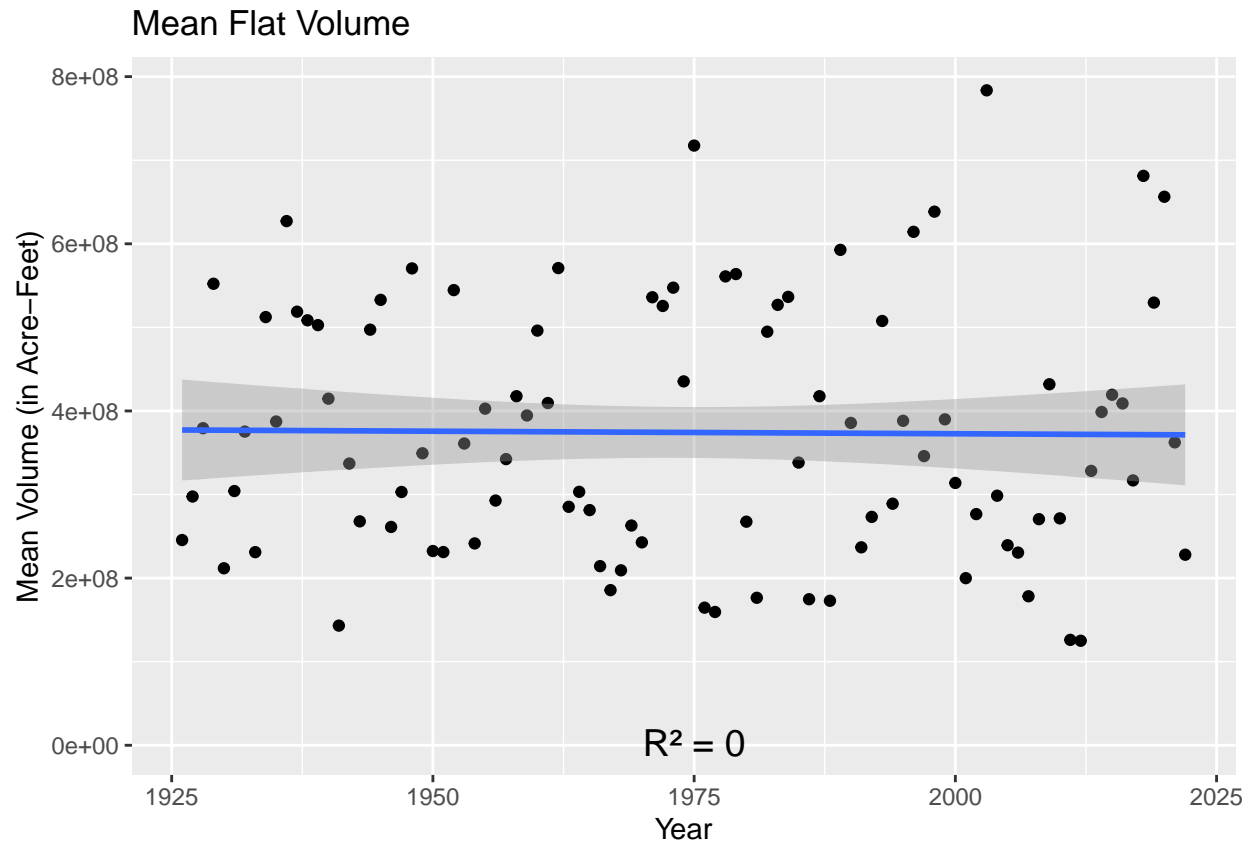
```
#Mean volume  
ggplot(data = flat_volume_mean) +  
  geom_col(data = flat_volume_mean, aes(x = year, y = avg_vol)) +  
  labs(title = "Mean Flat Volume", x = "Year", y = "Mean Volume (in Acre-Feet)")
```

```
flat_vol_model <- lm(year ~ avg_vol, data = flat_volume_mean)
flat_r3 <- summary(flat_vol_model)$r.squared

ggplot(data = flat_volume_mean, aes(x = year, y = avg_vol)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Mean Flat Volume", x = "Year", y = "Mean Volume (in Acre-Feet)") +
  annotate("text", x = 1975, y = 3542400,
         label = paste("R2 =", round(flat_r3, 2)),
         size = 5)
```

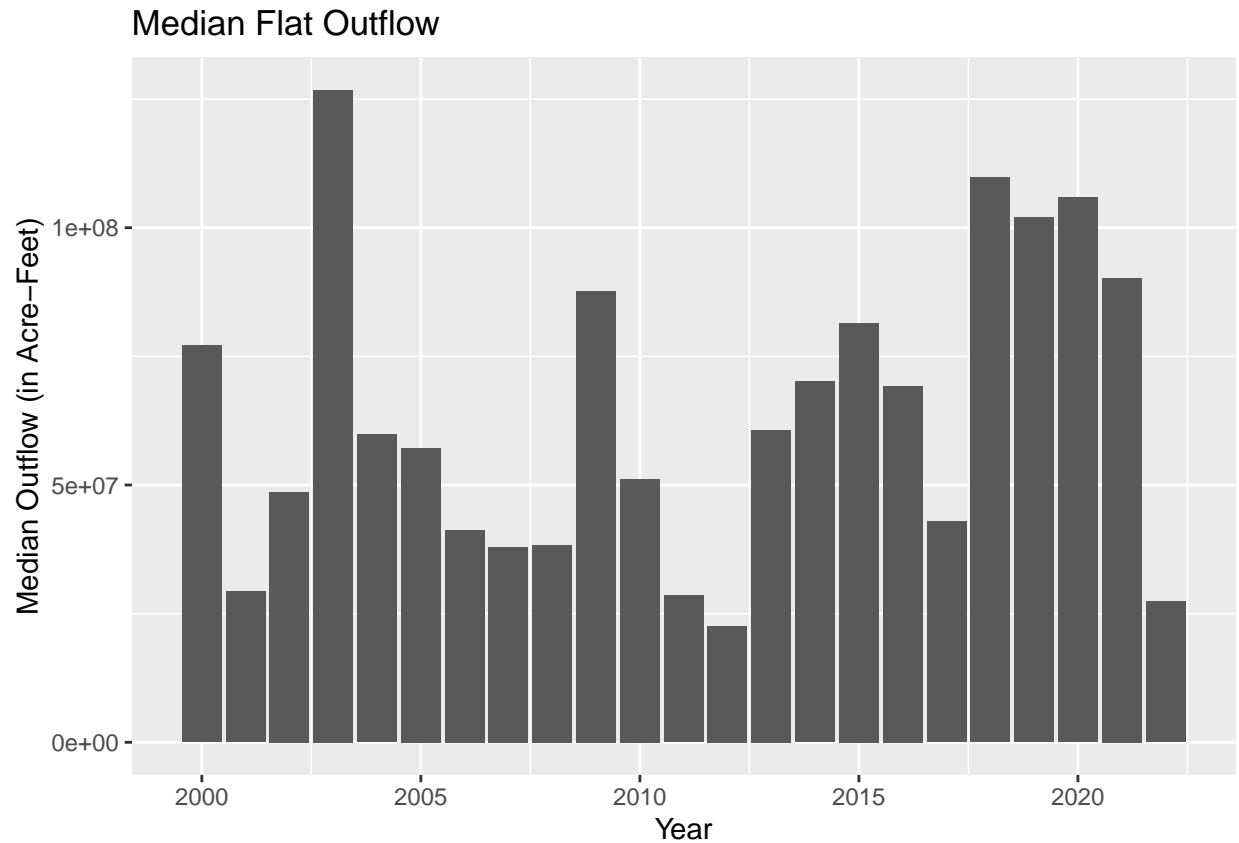
```
## 'geom_smooth()' using formula = 'y ~ x'
```



Now we will look at only years 2000 and beyond to capture recent trends.

```
#Filter out the data to only include 2000 and beyond.
flat_2000 <- flat_outflows %>%
  filter(year >= 2000)
#Filter out the average median outflow data to only include 2000 and beyond.
flat_2000_median <- flat_average_median %>%
  filter(year >= 2000)
#Filter out the average mean outflow data to only include 2000 and beyond.
flat_2000_mean <- flat_average_mean %>%
  filter(year >= 2000)
#Filter out the average volume data to only include 2000 and beyond.
flat_2000_volume <- flat_volume_mean %>%
  filter(year >= 2000)

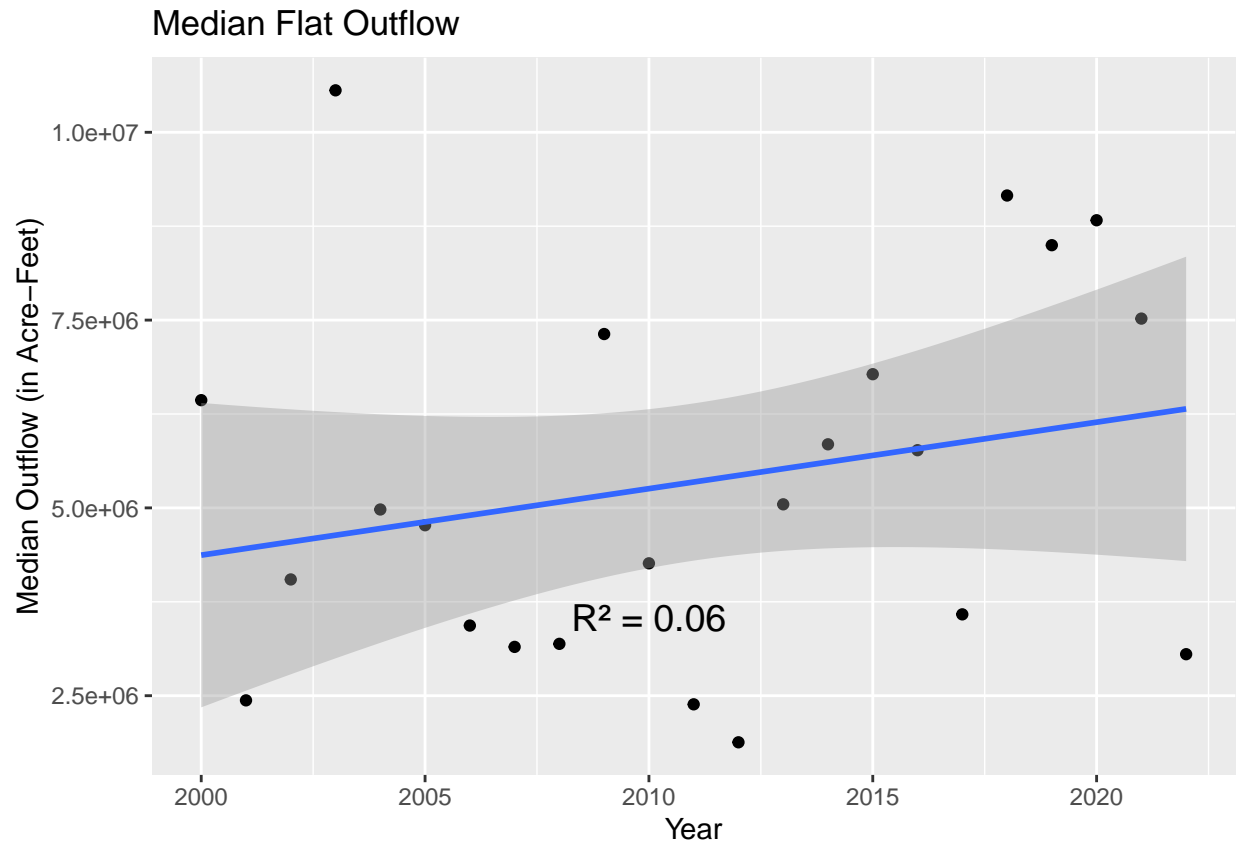
#Median Flat outflow
ggplot(data = flat_2000) +
  geom_col(data = flat_2000, aes(x = year, y = median_flow)) +
  labs(title = "Median Flat Outflow", x = "Year", y = "Median Outflow (in Acre-Feet)")
```



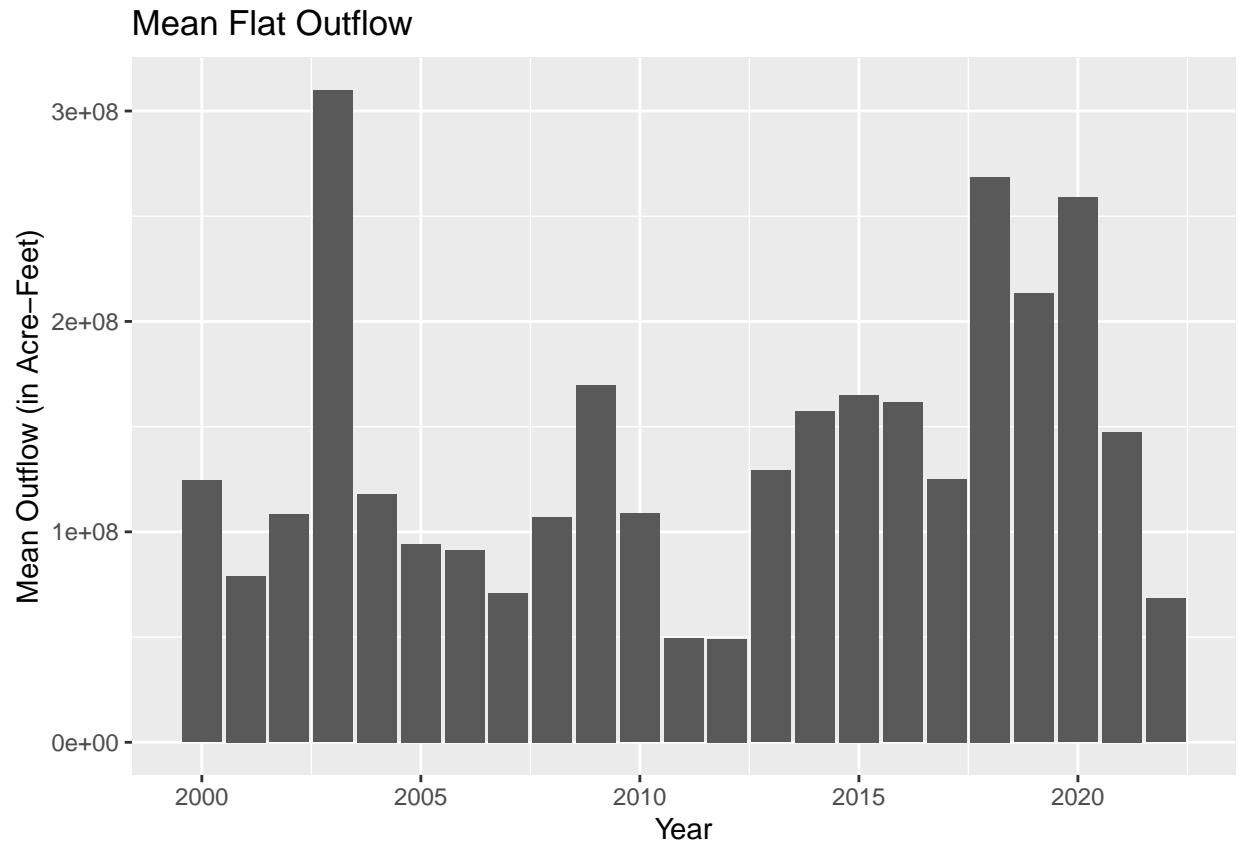
```
flat_med_model_2000 <- lm(year ~ avg_median_flow, data = flat_2000_median)
flat_r4 <- summary(flat_med_model_2000)$r.squared

ggplot(data = flat_2000_median, aes(x = year, y = avg_median_flow)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Median Flat Outflow", x = "Year", y = "Median Outflow (in Acre-Feet)") +
  annotate("text", x = 2010, y = 3542400,
    label = paste("R² =", round(flat_r4, 2)),
    size = 5)
```

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



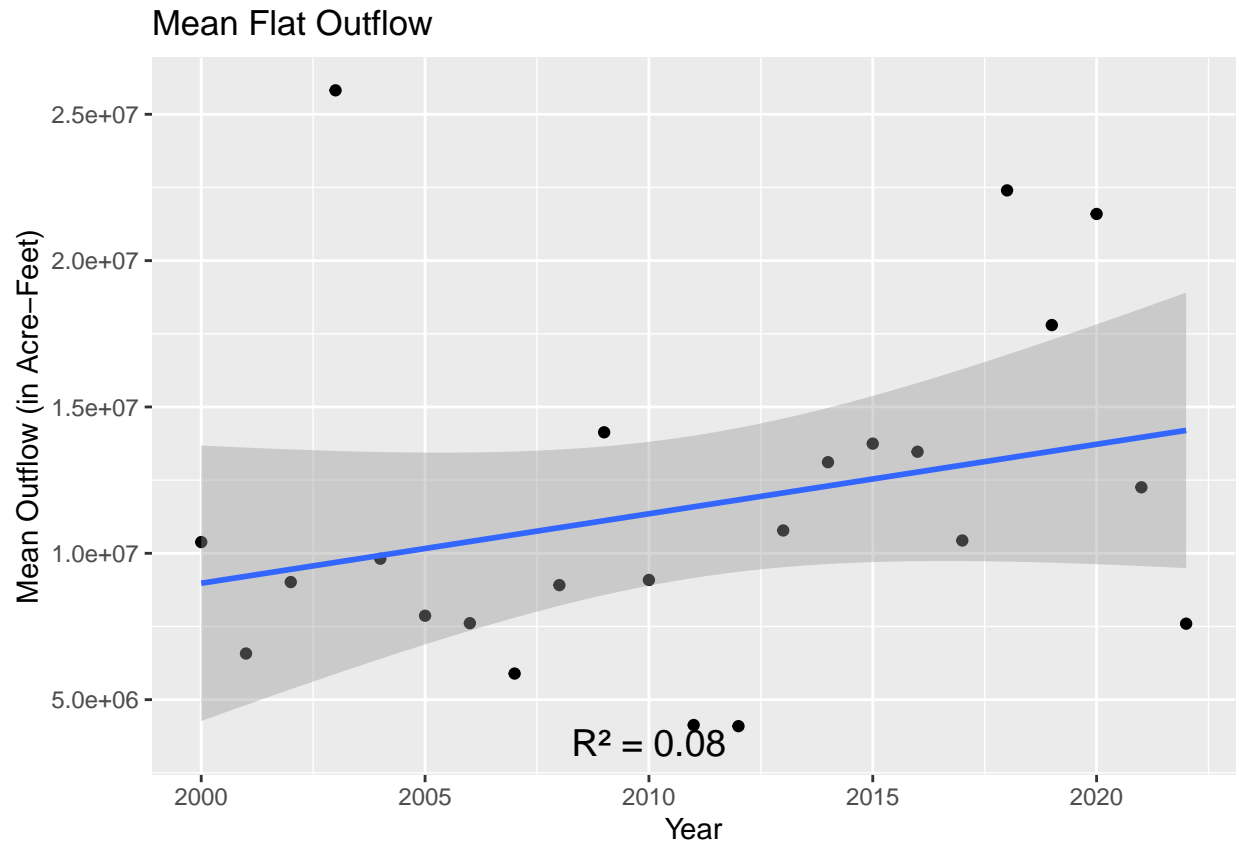
```
#Mean Flat outflow  
ggplot(data = flat_2000) +  
  geom_col(data = flat_2000, aes(x = year, y = mean_flow)) +  
  labs(title = "Mean Flat Outflow", x = "Year", y = "Mean Outflow (in Acre-Feet)")
```



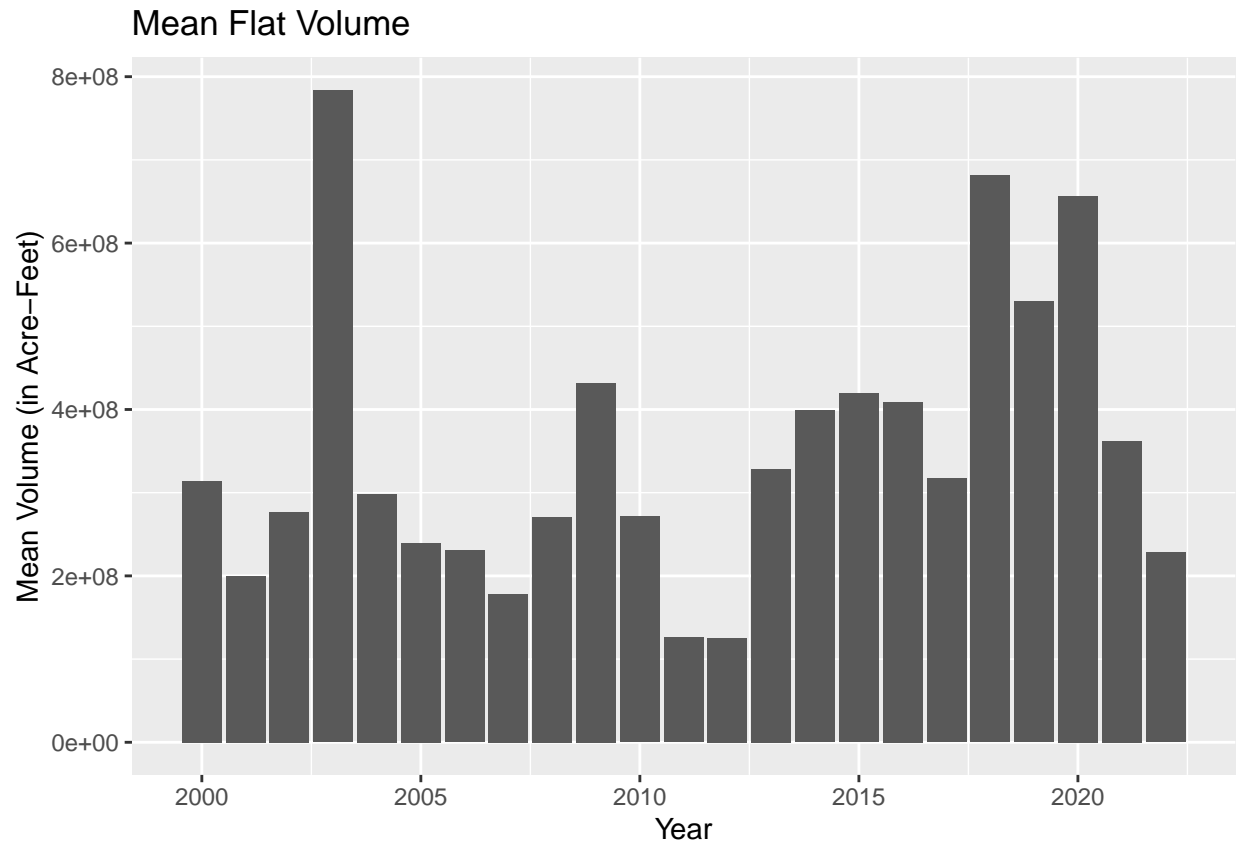
```
flat_mean_model_2000 <- lm(year ~ avg_mean_flow, data = flat_2000_mean)
flat_r5 <- summary(flat_mean_model_2000)$r.squared
```

```
ggplot(data = flat_2000_mean, aes(x = year, y = avg_mean_flow)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Mean Flat Outflow", x = "Year", y = "Mean Outflow (in Acre-Feet)") +
  annotate("text", x = 2010, y = 3542400,
    label = paste("R² =", round(flat_r5, 2)),
    size = 5)
```

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



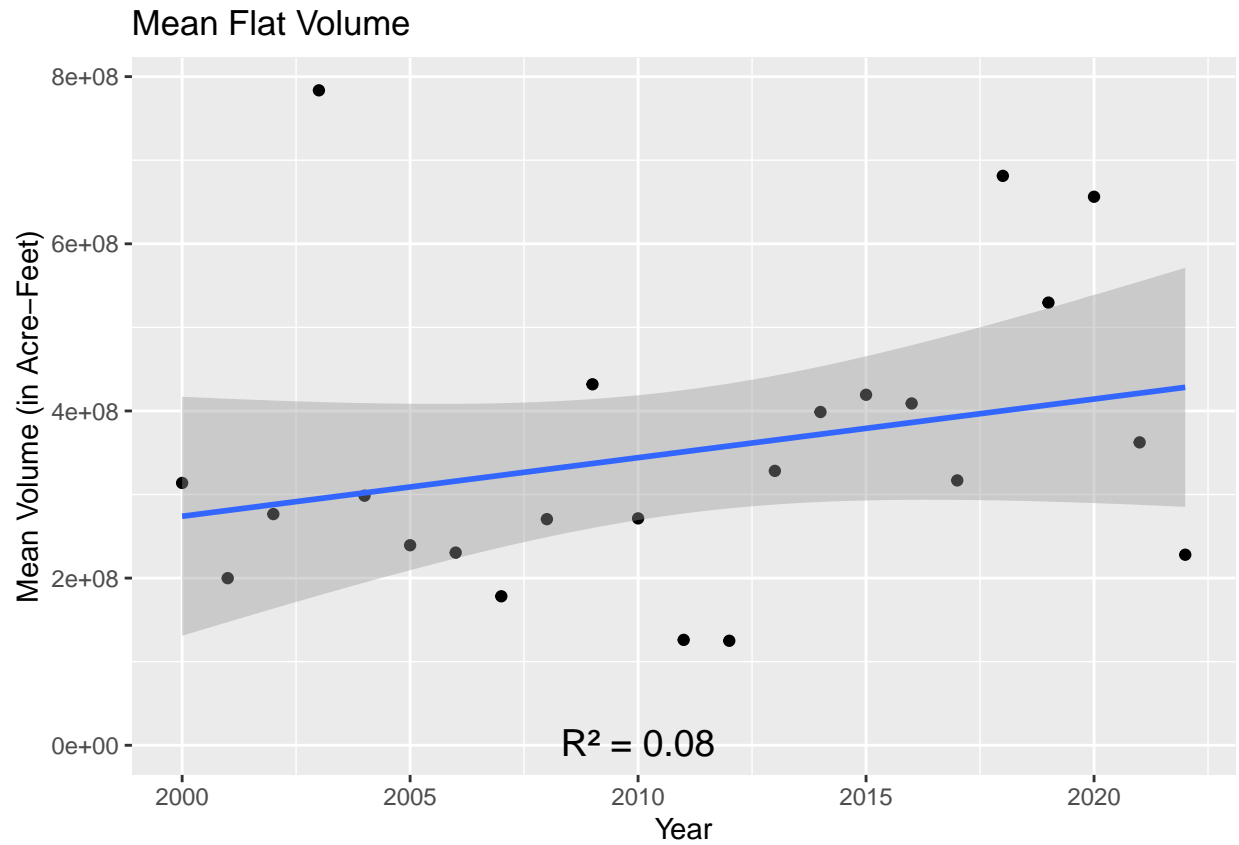
```
#Mean Flat volume  
ggplot(data = flat_2000_volume) +  
  geom_col(data = flat_2000_volume, aes(x = year, y = avg_vol)) +  
  labs(title = "Mean Flat Volume", x = "Year", y = "Mean Volume (in Acre-Feet)")
```



```
flat_vol_model_2000 <- lm(year ~ avg_vol, data = flat_2000_volume)
flat_r6 <- summary(flat_vol_model_2000)$r.squared
```

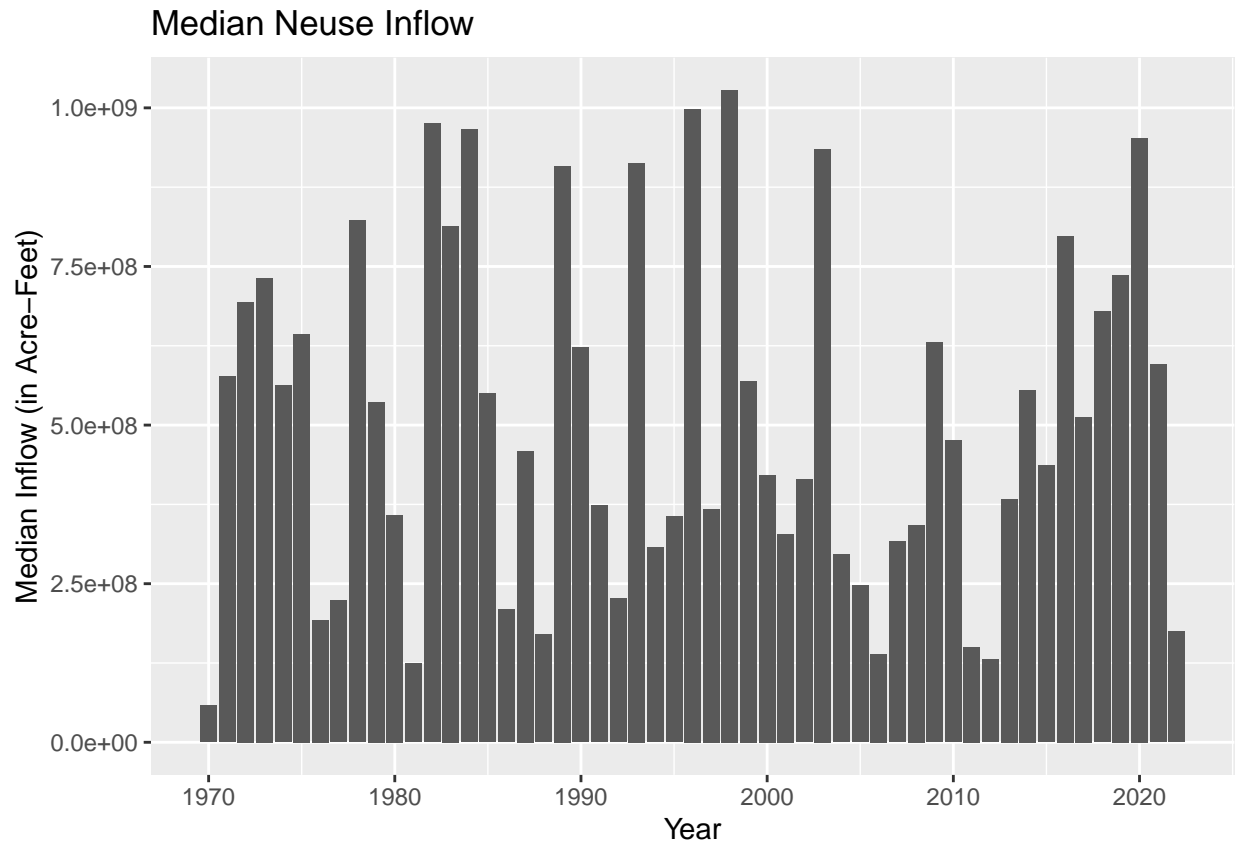
```
ggplot(data = flat_2000_volume, aes(x = year, y = avg_vol)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Mean Flat Volume", x = "Year", y = "Mean Volume (in Acre-Feet)") +
  annotate("text", x = 2010, y = 3542400,
    label = paste("R² =", round(flat_r6, 2)),
    size = 5)
```

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



Finally, we will look at Neuse inflows.

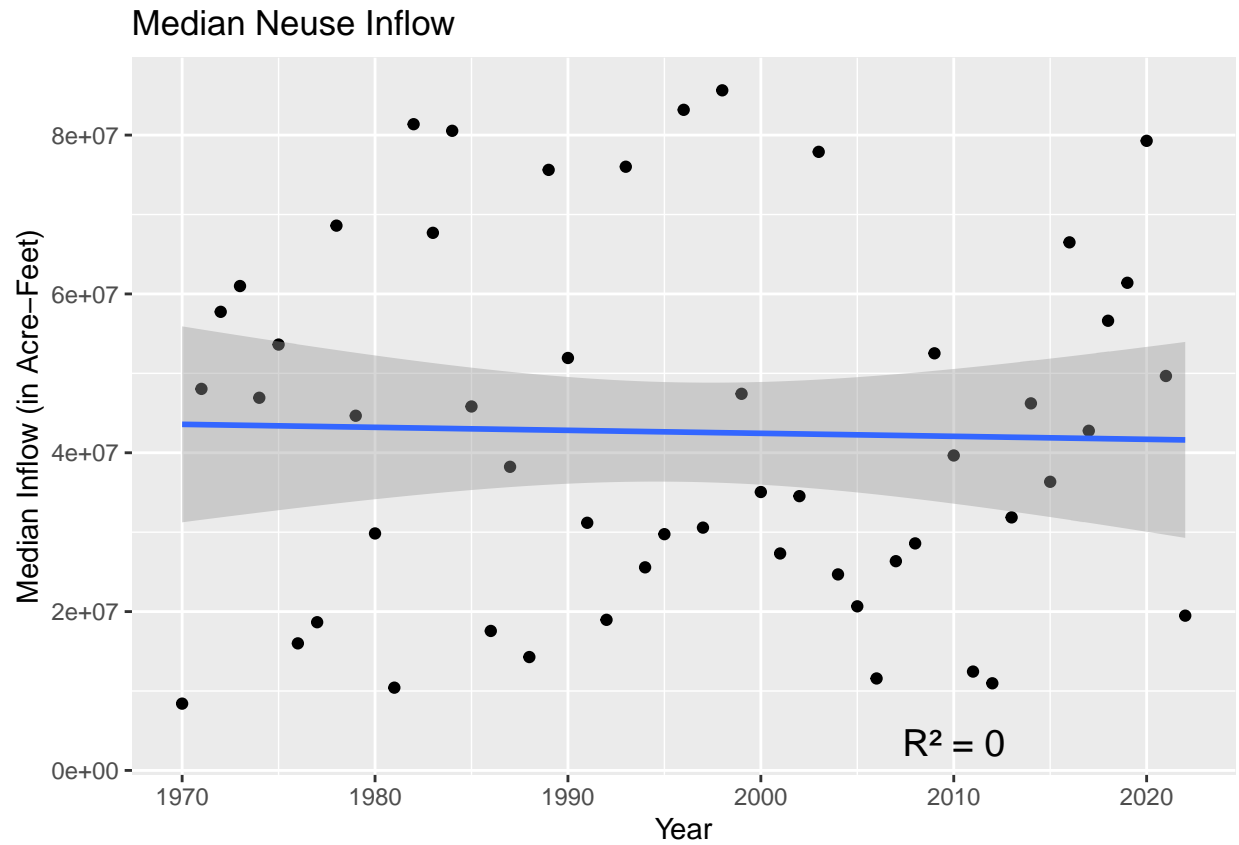
```
#Median Neuse inflow
ggplot(data = (neuse_inflows)) +
  geom_col(data = neuse_inflows, aes(x = year, y = median_flow)) +
  labs(title = "Median Neuse Inflow", x = "Year", y = "Median Inflow (in Acre-Feet)")
```

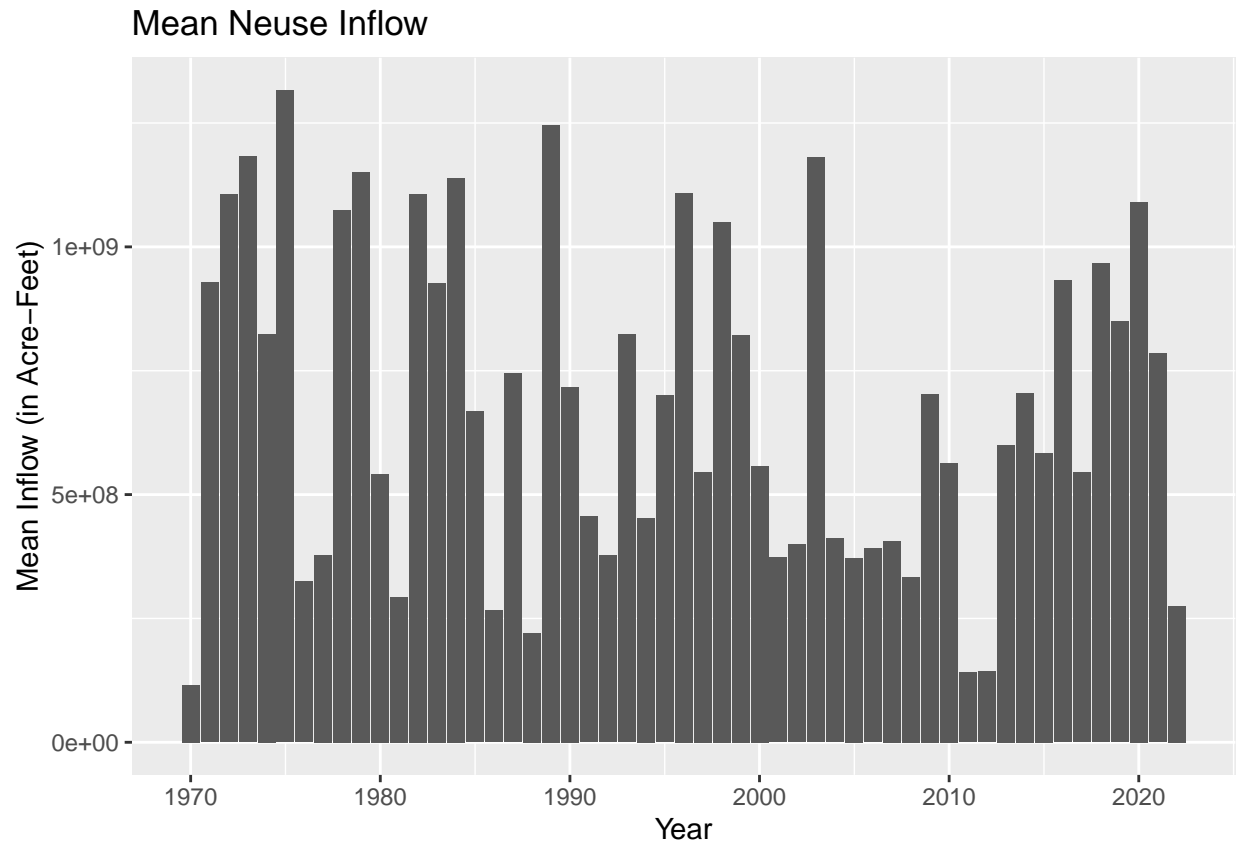
```
neuse_med_model <- lm(year ~ avg_median_flow, data = neuse_average_median)
neuse_r <- summary(neuse_med_model)$r.squared

ggplot(data = neuse_average_median, aes(x = year, y = avg_median_flow)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Median Neuse Inflow", x = "Year", y = "Median Inflow (in Acre-Feet)") +
  annotate("text", x = 2010, y = 3542400,
         label = paste("R2 =", round(neuse_r, 2)),
         size = 5)
```

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



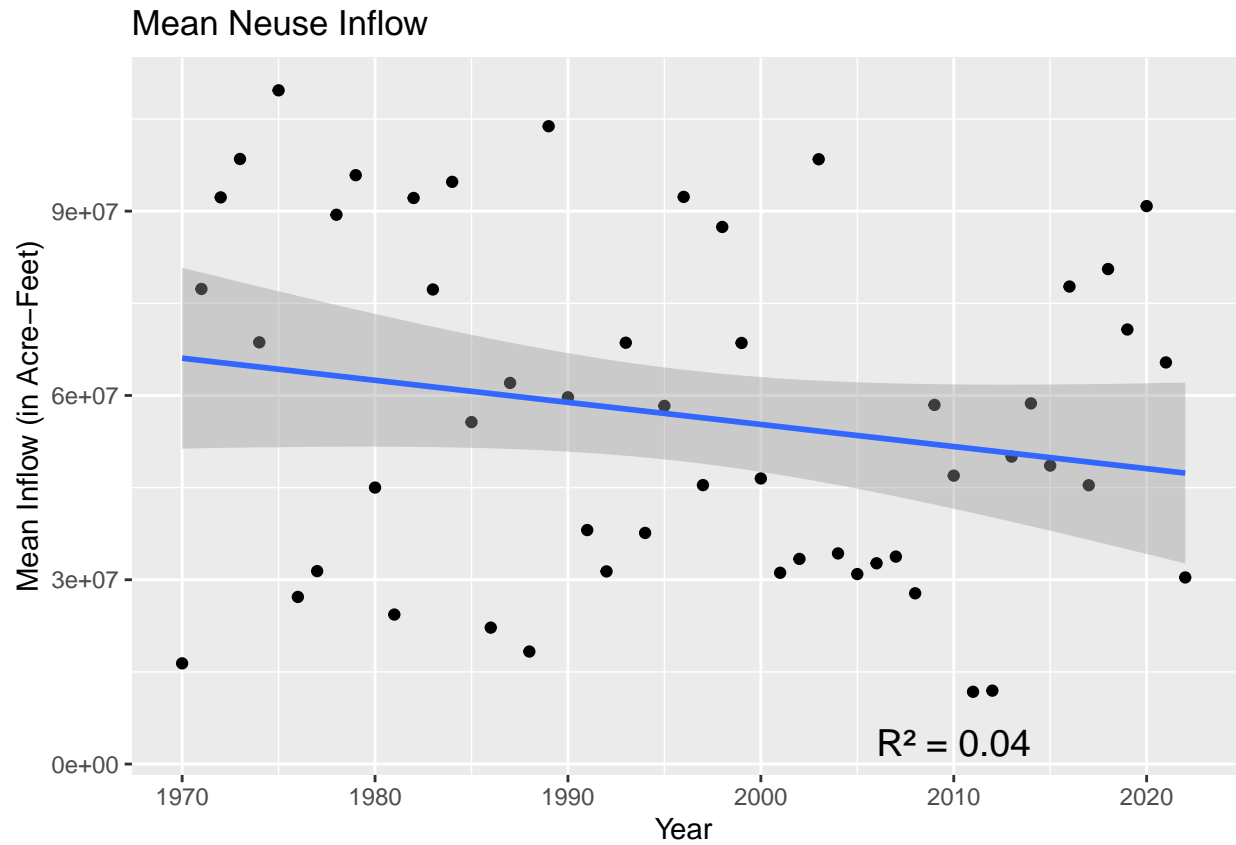
```
#Mean Neuse inflow  
ggplot(data = neuse_inflows) +  
  geom_col(data = neuse_inflows, aes(x = year, y = mean_flow)) +  
  labs(title = "Mean Neuse Inflow", x = "Year", y = "Mean Inflow (in Acre-Feet)")
```



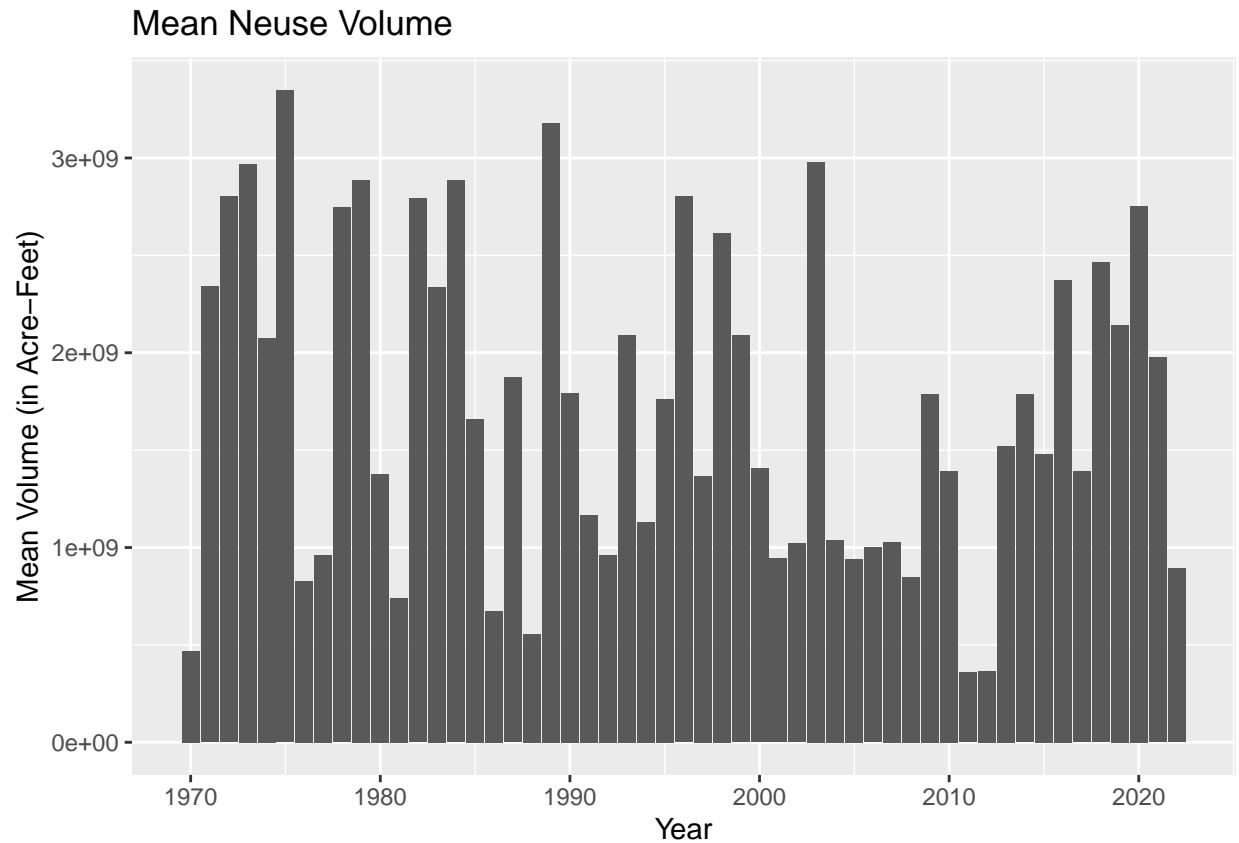
```
neuse_mean_model <- lm(year ~ avg_mean_flow, data = neuse_average_mean)
neuse_r2 <- summary(neuse_mean_model)$r.squared

ggplot(data = neuse_average_mean, aes(x = year, y = avg_mean_flow)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Mean Neuse Inflow", x = "Year", y = "Mean Inflow (in Acre-Feet)") +
  annotate("text", x = 2010, y = 3542400,
    label = paste("R2 =", round(neuse_r2, 2)),
    size = 5)
```

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



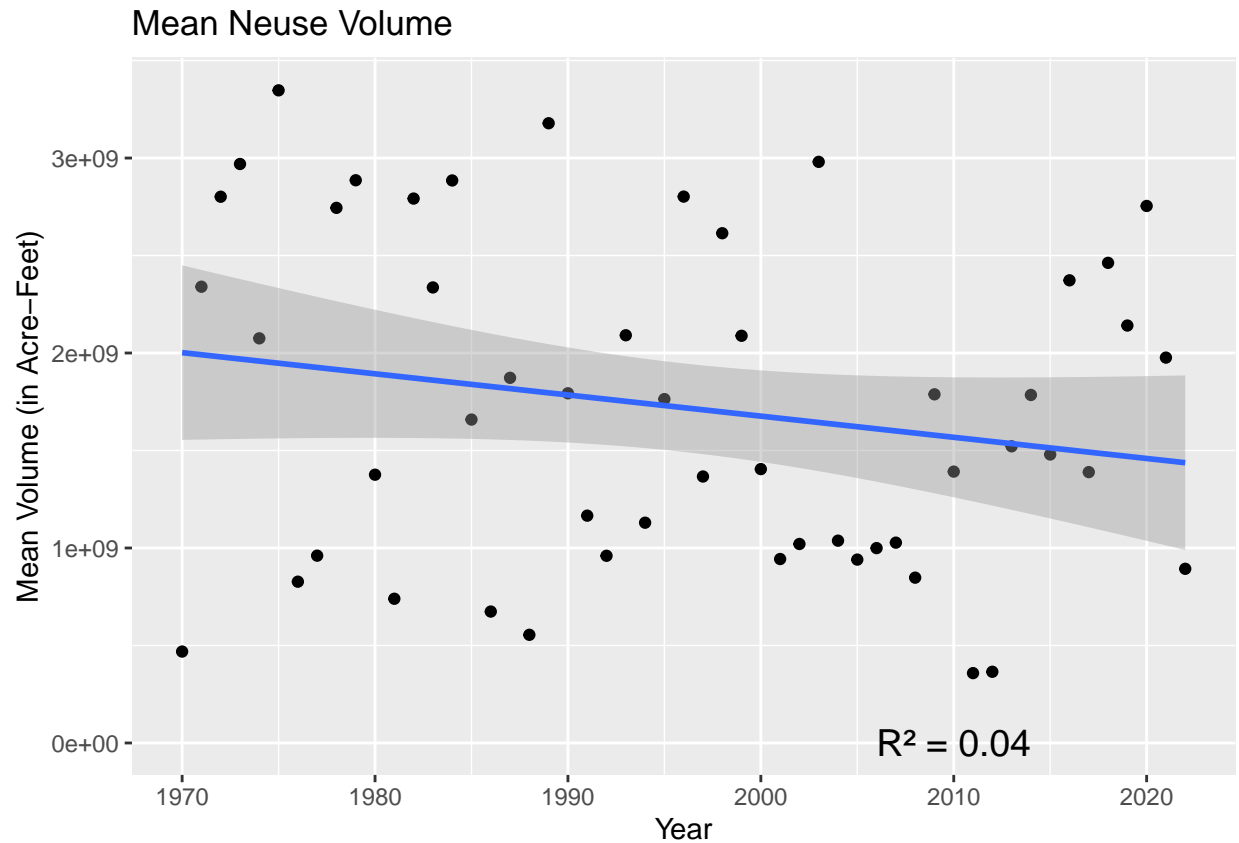
```
#Mean volume  
ggplot(data = neuse_volume_mean) +  
  geom_col(data = neuse_volume_mean, aes(x = year, y = avg_vol)) +  
  labs(title = "Mean Neuse Volume", x = "Year", y = "Mean Volume (in Acre-Feet)")
```



```
neuse_vol_model <- lm(year ~ avg_vol, data = neuse_volume_mean)
neuse_r3 <- summary(neuse_vol_model)$r.squared
```

```
ggplot(data = neuse_volume_mean, aes(x = year, y = avg_vol)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Mean Neuse Volume", x = "Year", y = "Mean Volume (in Acre-Feet)") +
  annotate("text", x = 2010, y = 3542400,
    label = paste("R2 =", round(neuse_r3, 2)),
    size = 5)
```

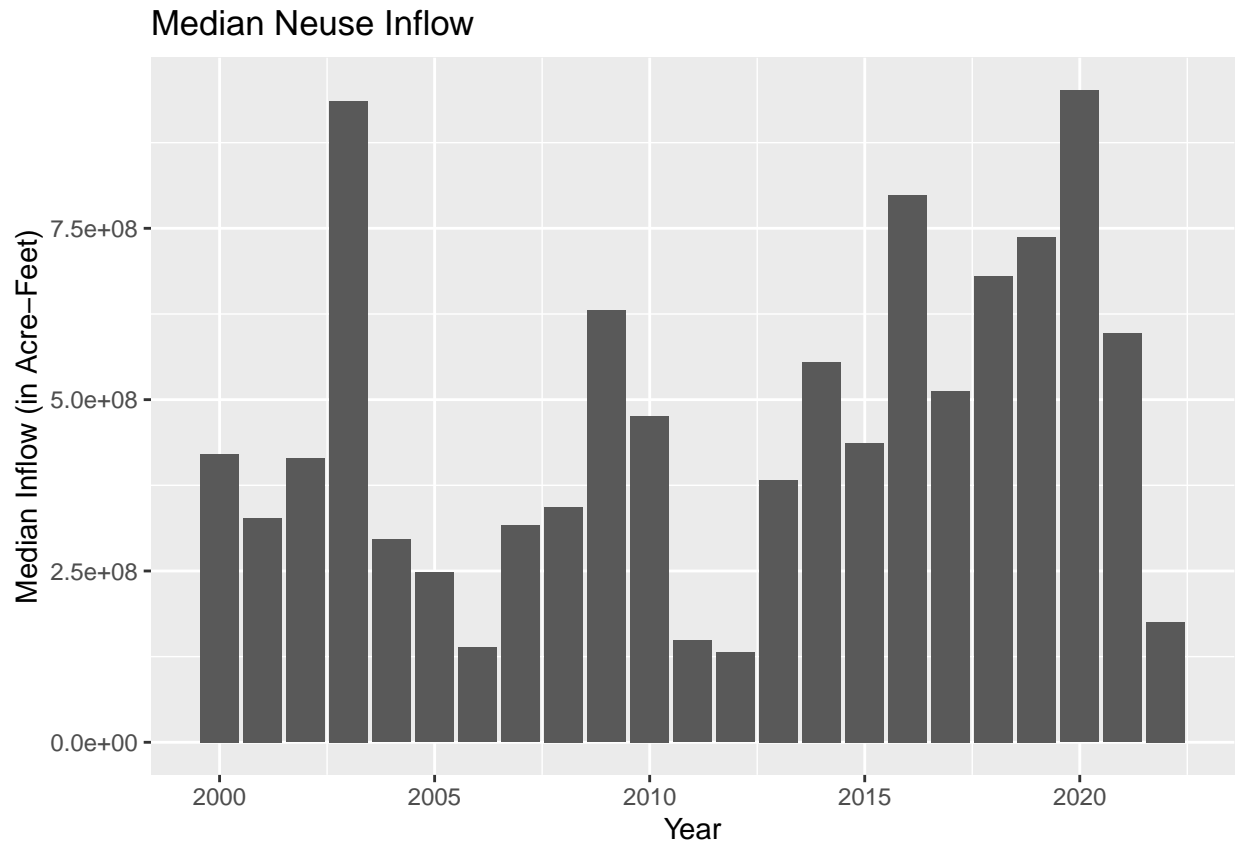
```
## 'geom_smooth()' using formula = 'y ~ x'
```



And now with only 2000 and beyond:

```
#Filter out the data to only include 2000 and beyond.
neuse_2000 <- neuse_inflows %>%
  filter(year >= 2000)
#Filter out the average median outflow data to only include 2000 and beyond.
neuse_2000_median <- neuse_average_median %>%
  filter(year >= 2000)
#Filter out the average mean outflow data to only include 2000 and beyond.
neuse_2000_mean <- neuse_average_mean %>%
  filter(year >= 2000)
#Filter out the average volume data to only include 2000 and beyond.
neuse_2000_volume <- neuse_volume_mean %>%
  filter(year >= 2000)

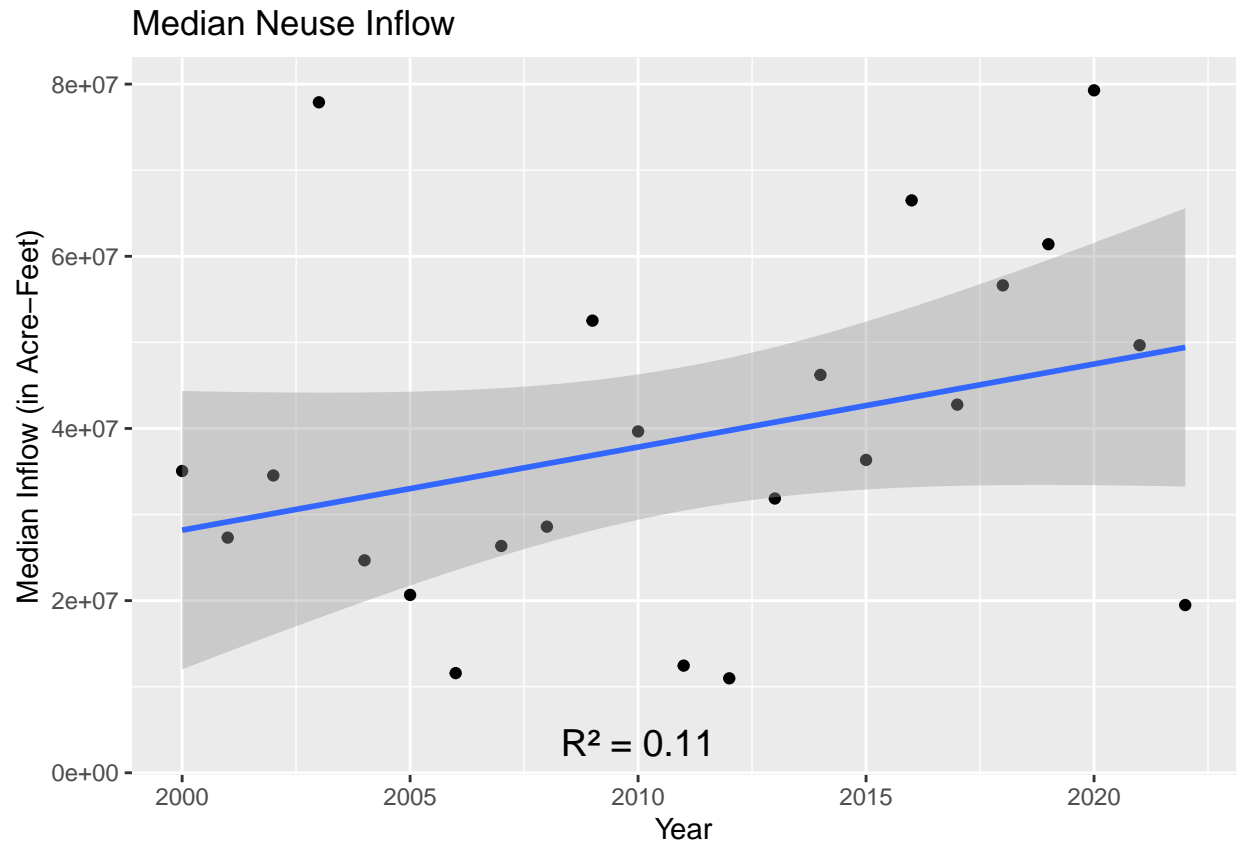
#Median Neuse inflow
ggplot(data = (neuse_2000)) +
  geom_col(data = neuse_2000, aes(x = year, y = median_flow)) +
  labs(title = "Median Neuse Inflow", x = "Year", y = "Median Inflow (in Acre-Feet)")
```



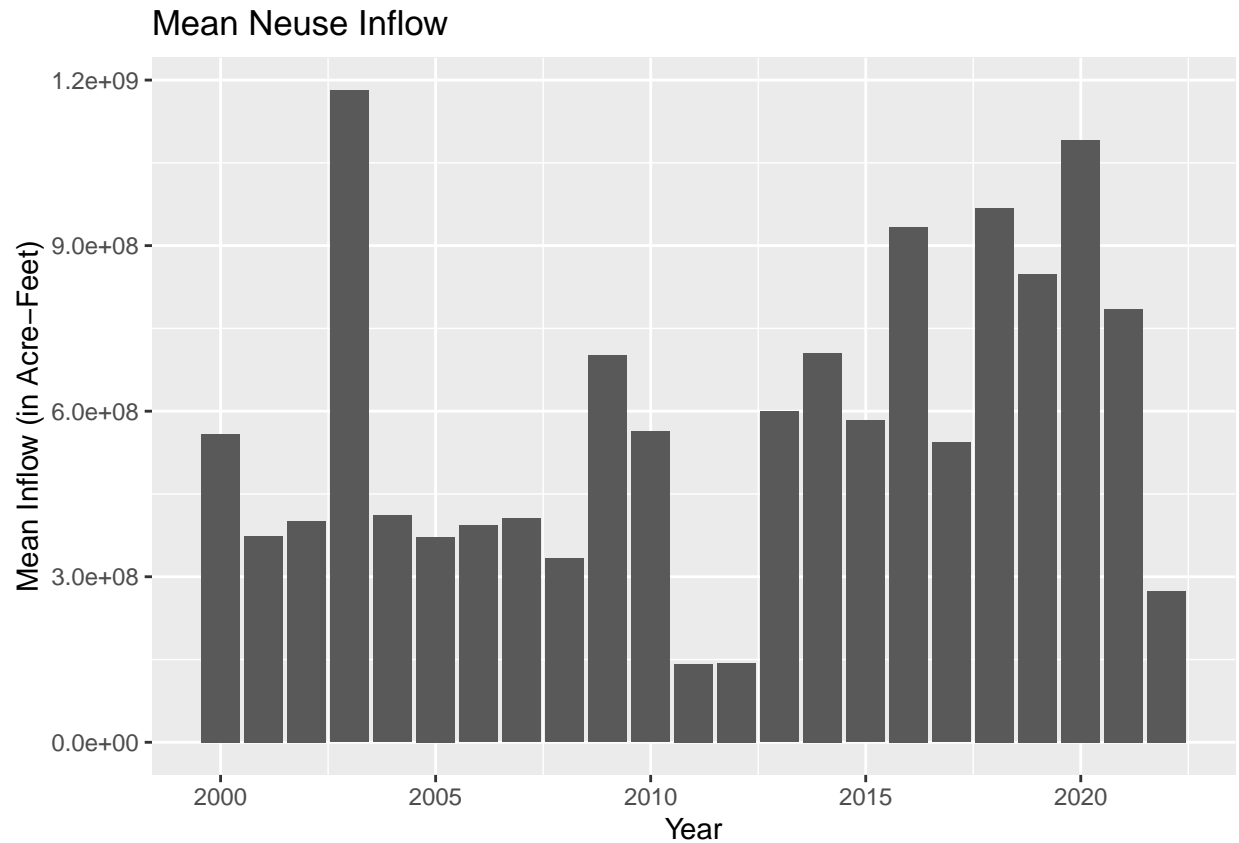
```
neuse_med_model_2000 <- lm(year ~ avg_median_flow, data = neuse_2000_median)
neuse_r4 <- summary(neuse_med_model_2000)$r.squared

ggplot(data = neuse_2000_median, aes(x = year, y = avg_median_flow)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Median Neuse Inflow", x = "Year", y = "Median Inflow (in Acre-Feet)") +
  annotate("text", x = 2010, y = 3542400,
    label = paste("R² =", round(neuse_r4, 2)),
    size = 5)
```

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



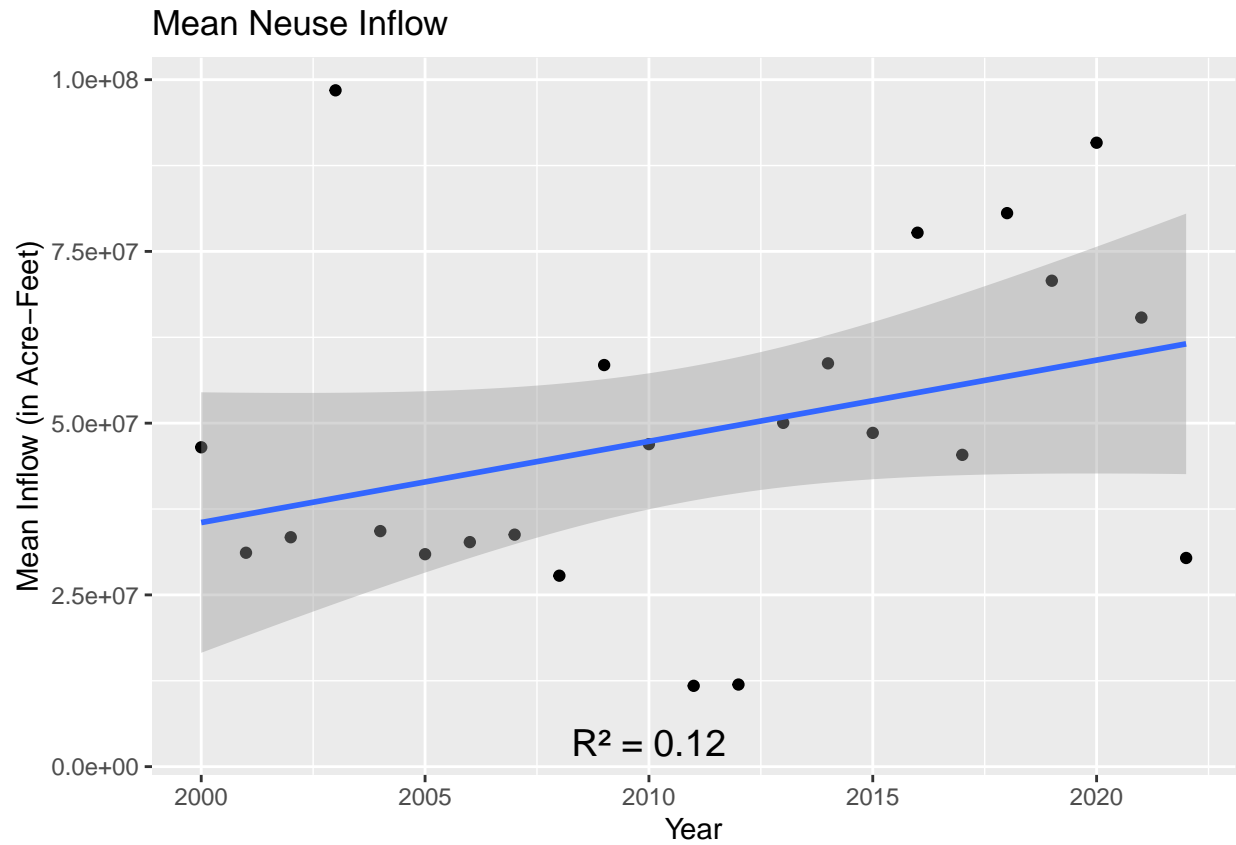
```
#Mean Neuse inflow
ggplot(data = neuse_2000) +
  geom_col(data = neuse_2000, aes(x = year, y = mean_flow)) +
  labs(title = "Mean Neuse Inflow", x = "Year", y = "Mean Inflow (in Acre-Feet)")
```

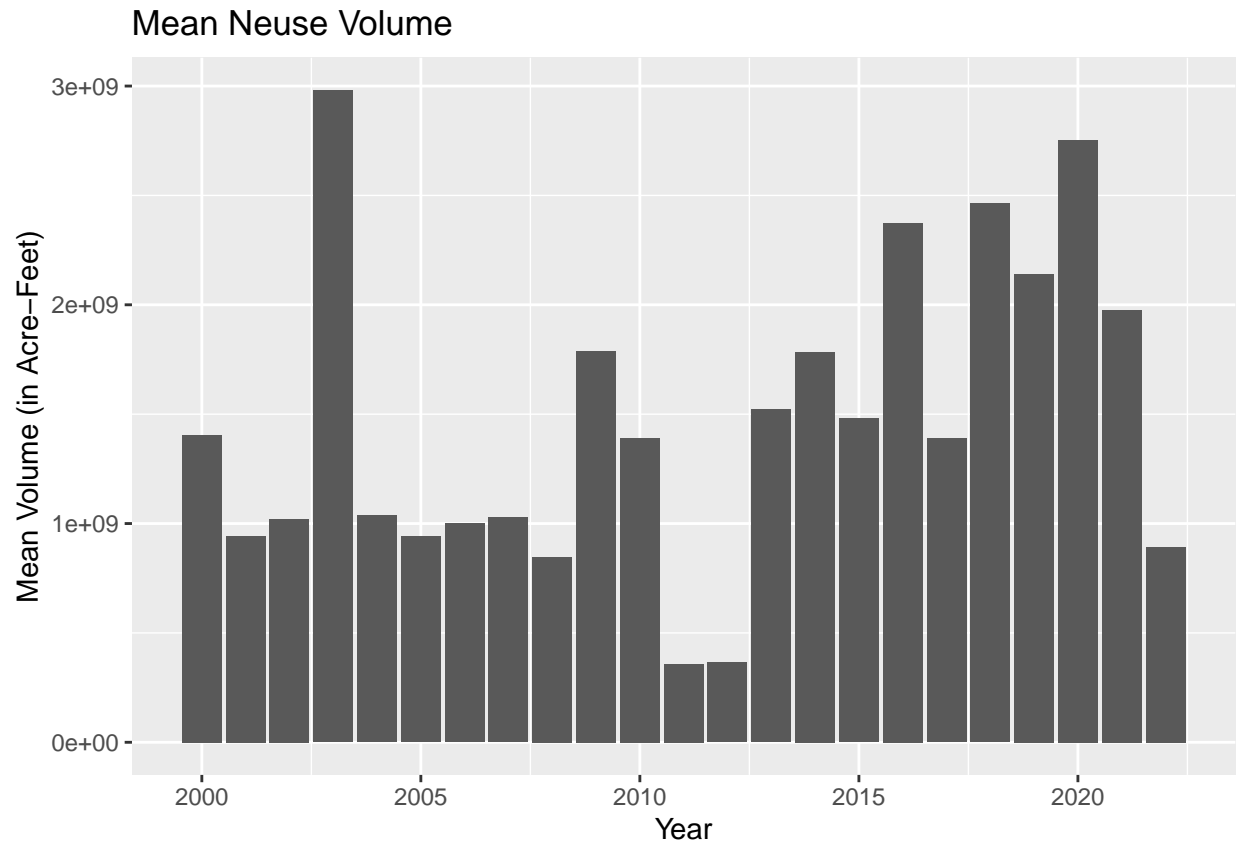
```
neuse_mean_model_2000 <- lm(year ~ avg_mean_flow, data = neuse_2000_mean)
neuse_r5 <- summary(neuse_mean_model_2000)$r.squared

ggplot(data = neuse_2000_mean, aes(x = year, y = avg_mean_flow)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Mean Neuse Inflow", x = "Year", y = "Mean Inflow (in Acre-Feet)") +
  annotate("text", x = 2010, y = 3542400,
    label = paste("R² =", round(neuse_r5, 2)),
    size = 5)
```

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



```
#Mean Neuse volume  
ggplot(data = neuse_2000_volume) +  
  geom_col(data = neuse_2000_volume, aes(x = year, y = avg_vol)) +  
  labs(title = "Mean Neuse Volume", x = "Year", y = "Mean Volume (in Acre-Feet)")
```



```
neuse_vol_model_2000 <- lm(year ~ avg_vol, data = neuse_2000_volume)
neuse_r6 <- summary(neuse_vol_model_2000)$r.squared

ggplot(data = neuse_2000_volume, aes(x = year, y = avg_vol)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Mean Neuse Volume", x = "Year", y = "Mean Volume (in Acre-Feet)") +
  annotate("text", x = 2010, y = 3542400,
    label = paste("R² =", round(neuse_r6, 2)),
    size = 5)
```

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

