Fire and REcreational Visitation in Yosemite 1999-2019

Elsie Liu, Gaby Antonova

4/9/2022

Fire

Data Wrangling

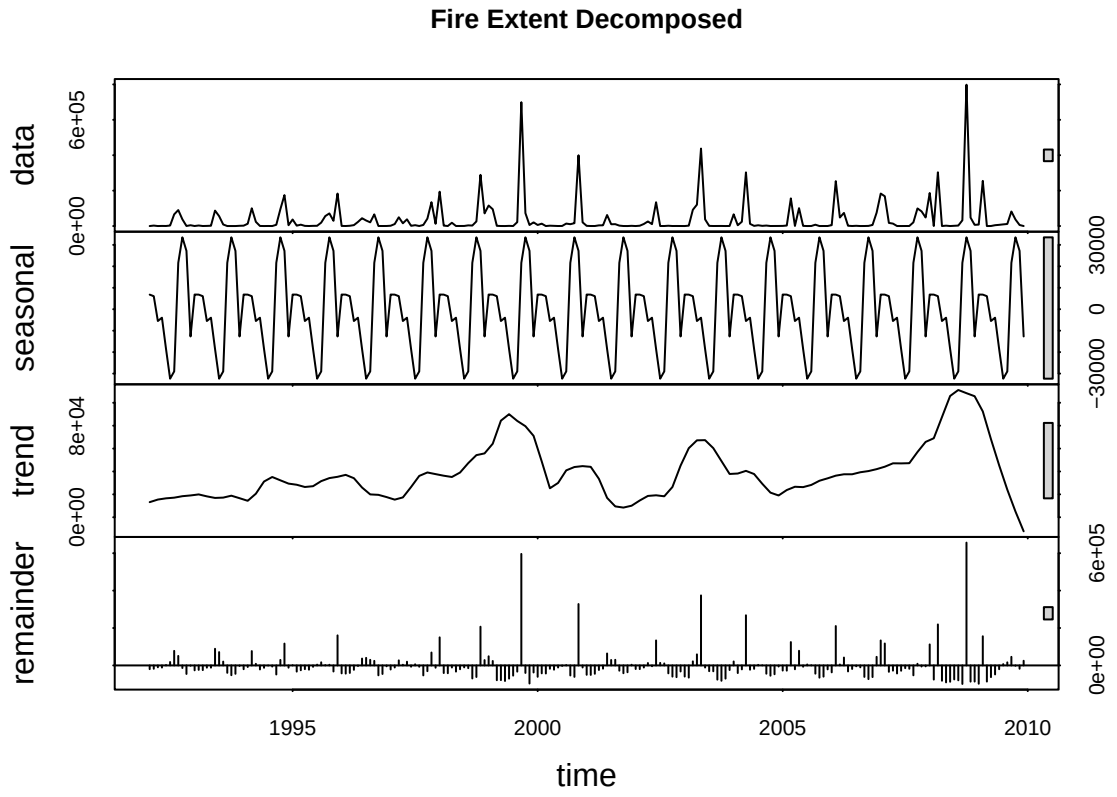
```
fire$Date<- as.Date(fire$ALARM_DATE,format = "%Y/%m/%d")
fire.p <- fire %>%
  select(Date, REPORT_AC) %>%
  mutate(Year= year(Date)) %>%
  mutate(Month= month(Date)) %>%
  filter(Year >1998 & Year < 2020) %>%
  mutate(count = 1) %>%
  drop_na()

fire.sum <- fire.p%>%
  group_by(Year, Month) %>%
  dplyr::summarize(summonth = sum(REPORT_AC),
            Count = sum(count)) %>%
  mutate(Date = ym(paste(Year, Month,sep = "-")))
```

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

Timeseries analysis for fire data

```
library(Kendall)
fire.ts <- ts(fire.sum$summonth, start = c(1992,2,1), frequency = 12)
#decompose
fire.Decomposed <- stl(fire.ts, s.window = "periodic")
plot(fire.Decomposed, main = "Fire Extent Decomposed")
```




```
#Mann Kendall test
fire.trend1 <- Kendall::SeasonalMannKendall(fire.ts)
summary(fire.trend1)
```

```
## Score =  20 , Var(Score) = 8253.333
## denominator =  1817.496
## tau = 0.011, 2-sided pvalue =0.82576
```

Visitation data analysis

Data wrangling for visitation data

```
#Pivot Longer to create a tidy dataset
library(tidyverse)

# Gather nutrient data into one column using pivot_longer
Visitation_by_month_gathered <- Visitation_by_Month %>%
  pivot_longer(JAN:DEC, names_to = "Month", values_to = "Recreational_Visits")

#Create a date column

library(lubridate)
Visitation_by_month_gathered$Date <- ymd( paste(Visitation_by_month_gathered$Year, Visitation_by_month_
head(Visitation_by_month_gathered)
```

```
## # A tibble: 6 x 4
##    Year Month Recreational_Visits Date
```

```
##   <dbl> <chr>               <dbl> <date>
## 1  2021 JAN                 67284 2021-01-01
## 2  2021 FEB                128222 2021-02-01
## 3  2021 MAR                159906 2021-03-01
## 4  2021 APR                257158 2021-04-01
## 5  2021 MAY                348006 2021-05-01
## 6  2021 JUN                429980 2021-06-01
```

```
Visitation_by_month_gathered <- filter(Visitation_by_month_gathered, Year > 1998 & Year < 2020 )
Visitation_by_month_gathered$Month <- month(Visitation_by_month_gathered$Date)
```

Data Exploration

```
library(ggplot2)
head(Visitation_by_month_gathered)
```

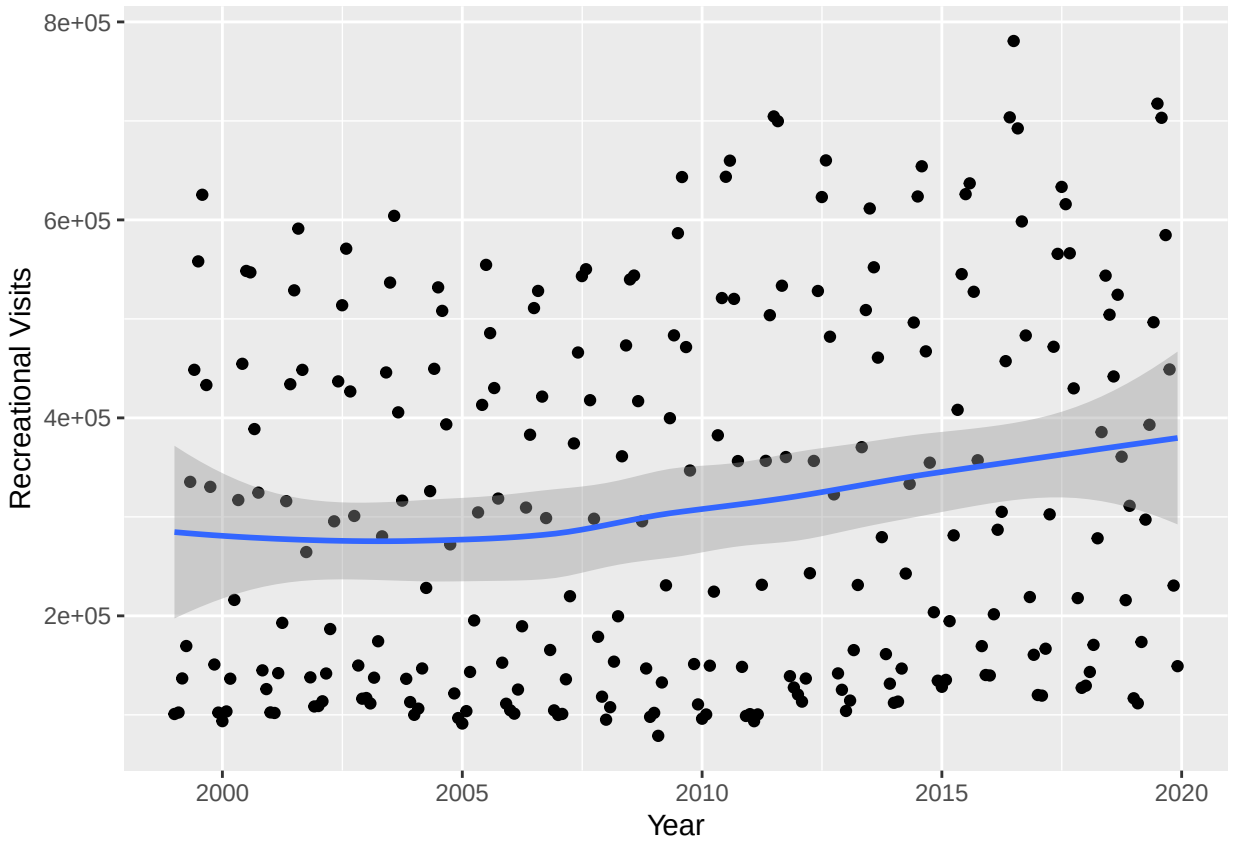
```
## # A tibble: 6 x 4
##    Year Month Recreational_Visits Date
##   <dbl> <dbl>               <dbl> <date>
## 1  2019     1              116746 2019-01-01
## 2  2019     2              111665 2019-02-01
## 3  2019     3              173610 2019-03-01
## 4  2019     4              297207 2019-04-01
## 5  2019     5              393004 2019-05-01
## 6  2019     6              496625 2019-06-01
```

```
summary(Visitation_by_month_gathered$Recreational_Visits)
```

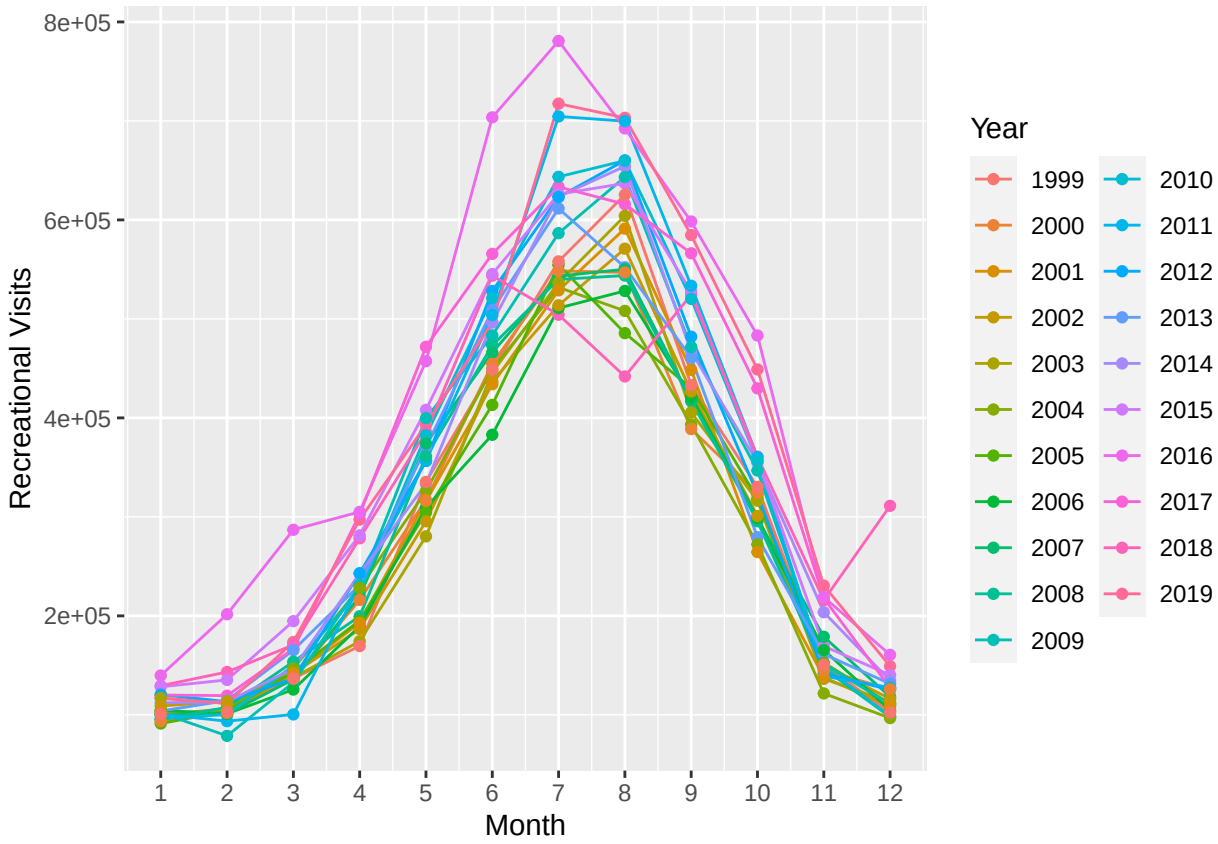
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   78795  136646  291250  312429  471608  780728
```

```
Visitation_by_month_plot <- ggplot(Visitation_by_month_gathered, aes(x = Date, y = Recreational_Visits))
  geom_point()+
  geom_smooth()+
  labs(x="Year", colour="Year", y = "Recreational Visits")
print(Visitation_by_month_plot)
```

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




```
visitation_by_month_plot2 <- ggplot(Visitation_by_month_gathered, aes(x = Month, y = Recreational_Visits
  geom_line() +
  geom_point() +
  labs(x="Month", colour="Year", y = "Recreational Visits")+
  scale_x_continuous(breaks = 1:12)
print(visitation_by_month_plot2)
```



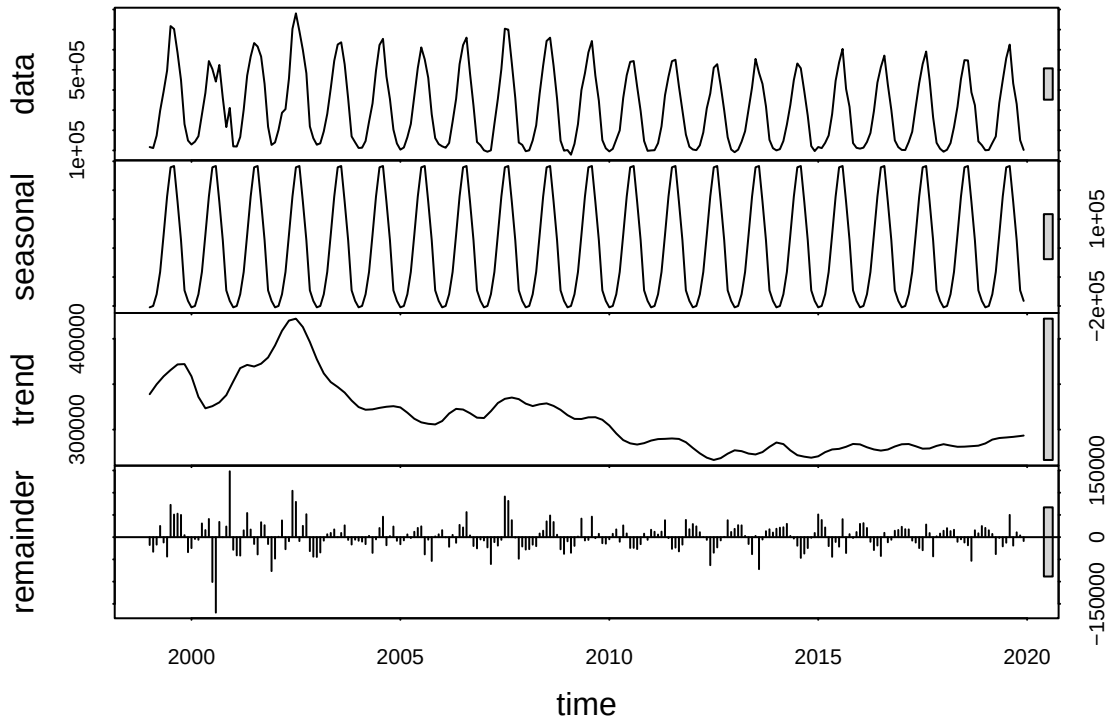
Time Series Analysis of Visitation Data

```
#Question: Has there been an increasing or decreasing visitation trend for Yosemite National Park betwe

Visitation_by_month_gathered_ts <- ts(Visitation_by_month_gathered$Recreational_Visits, start = c(1999,

# Generate the decomposition
Visitation_by_month_Decomposed <- stl(Visitation_by_month_gathered_ts, s.window = "periodic")

# Visualize the decomposed series.
plot(Visitation_by_month_Decomposed)
```




```
#mostly monotonic trend - visitation appears to be decreasing overall

# Run SMK test

Visitation_by_month_trend1 <- Kendall::SeasonalMannKendall(Visitation_by_month_gathered_ts)

#Answer: The seasonal Mann-Kendall test is the most appropriate because this ozone data has seasonality

# Inspect results
Visitation_by_month_trend1
```

```
## tau = -0.494, 2-sided pvalue =< 2.22e-16
```

```
summary(Visitation_by_month_trend1)
```

```
## Score =  -1244 , Var(Score) = 13160
## denominator =  2520
## tau = -0.494, 2-sided pvalue =< 2.22e-16
```

```
#The p-value us less than 0.05, so we reject the null hypothesis that the data is stationary, therefore

Visitation_by_month_trend2 <- trend::smk.test(Visitation_by_month_gathered_ts)
# Inspect results
Visitation_by_month_trend2
```

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data:  Visitation_by_month_gathered_ts
## z = -10.835, p-value < 2.2e-16
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##     S  varS
```

```
## -1244 13160
```

```
summary(Visitation_by_month_trend2)
```

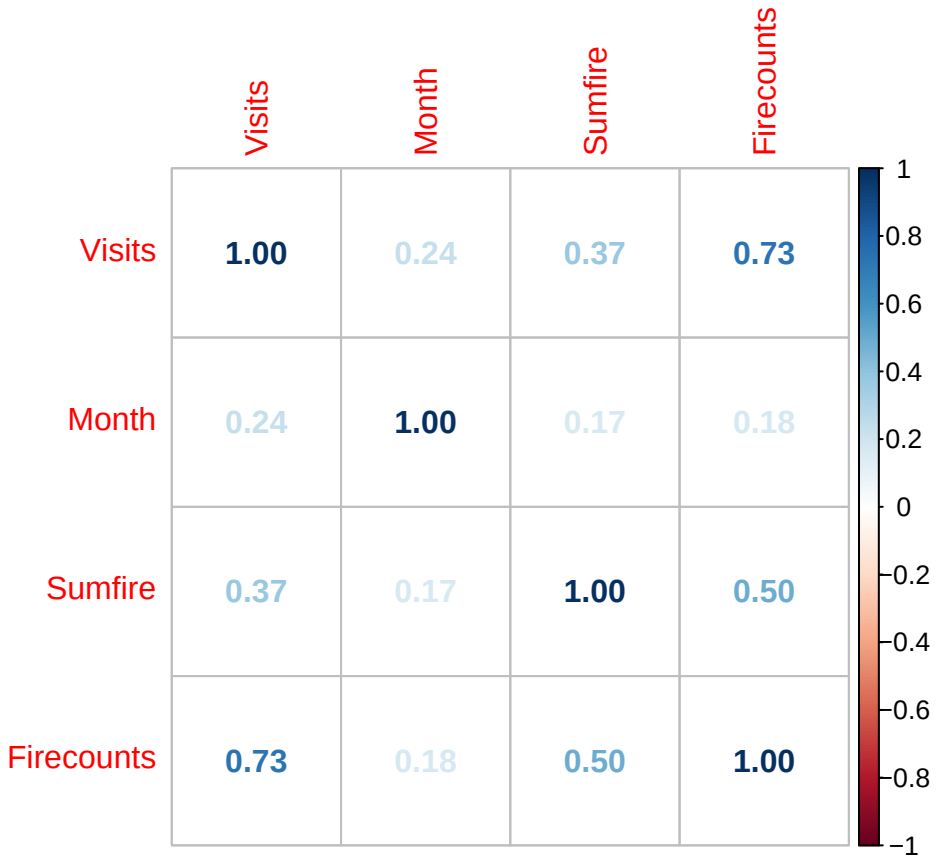
```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: Visitation_by_month_gathered_ts
## alternative hypothesis: two.sided
##
## Statistics for individual seasons
##
## H0
##                        S   varS    tau      z   Pr(>|z|)
## Season 1:   S = 0   -90 1096.7 -0.429 -2.688 0.00719835  **
## Season 2:   S = 0   -70 1096.7 -0.333 -2.084 0.03719769   *
## Season 3:   S = 0   -92 1096.7 -0.438 -2.748 0.00599747  **
## Season 4:   S = 0  -156 1096.7 -0.743 -4.681 2.8614e-06 ***
## Season 5:   S = 0  -112 1096.7 -0.533 -3.352 0.00080271 ***
## Season 6:   S = 0  -132 1096.7 -0.629 -3.956 7.6280e-05 ***
## Season 7:   S = 0   -94 1096.7 -0.448 -2.808 0.00498017  **
## Season 8:   S = 0   -58 1096.7 -0.276 -1.721 0.08521008   .
## Season 9:   S = 0  -122 1096.7 -0.581 -3.654 0.00025836 ***
## Season 10:   S = 0 -100 1096.7 -0.476 -2.989 0.00279439  **
## Season 11:   S = 0 -106 1096.7 -0.505 -3.171 0.00152084  **
## Season 12:   S = 0 -112 1096.7 -0.533 -3.352 0.00080271 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Modeling relationship between visitation and fire

```
library(corrplot)
```

```
## corrplot 0.92 loaded
```

```
yosefire <- inner_join(Visitation_by_month_gathered,fire.sum, by="Date")
yosefire <- yosefire%>%
  select(Recreational_Visits:Count) %>%
  filter(Count<100)
colnames(yosefire)<- c("Visits","Date", "Year","Month","Sumfire","Firecounts")
write.csv(yosefire, "./Yosefire.csv")
yosecorr <- cor(yosefire[,c(1,4:6)])
corrplot(yosecorr,method = "number",insig = "label_sig",
         sig.level = c(.05))
```




```
#pairs(yosefire)

#run linear model
mod.lmfull <- lm(data=yosefire, Visits ~ Year + Month + Sumfire + Firecounts)
summary(mod.lmfull)
```

```
##
## Call:
## lm(formula = Visits ~ Year + Month + Sumfire + Firecounts, data = yosefire)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -336429  -92290    -714   78433  306777
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.921e+06  2.882e+06  -0.666   0.5059
## Year         1.046e+03  1.435e+03   0.729   0.4669
## Month        6.437e+03  2.761e+03   2.332   0.0207 *
## Sumfire     -6.090e-03  1.156e-01  -0.053   0.9580
## Firecounts   6.986e+03  5.350e+02  13.057   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 124600 on 209 degrees of freedom
## Multiple R-squared:  0.5475, Adjusted R-squared:  0.5388
## F-statistic: 63.21 on 4 and 209 DF,  p-value: < 2.2e-16
```

```
step(mod.lmfull)
```

```
## Start:  AIC=5026.53
## Visits ~ Year + Month + Sumfire + Firecounts
##
##              Df  Sum of Sq        RSS    AIC
## - Sumfire     1 4.3092e+07 3.2436e+12 5024.5
## - Year        1 8.2450e+09 3.2518e+12 5025.1
## <none>                     3.2435e+12 5026.5
## - Month       1 8.4394e+10 3.3279e+12 5030.0
## - Firecounts  1 2.6460e+12 5.8895e+12 5152.2
##
## Step:  AIC=5024.53
## Visits ~ Year + Month + Firecounts
##
##              Df  Sum of Sq        RSS    AIC
## - Year        1 8.2028e+09 3.2518e+12 5023.1
## <none>                     3.2436e+12 5024.5
## - Month       1 8.4798e+10 3.3284e+12 5028.1
## - Firecounts  1 3.3966e+12 6.6402e+12 5175.9
##
## Step:  AIC=5023.07
## Visits ~ Month + Firecounts
##
##              Df  Sum of Sq        RSS    AIC
## <none>                     3.2518e+12 5023.1
## - Month       1 8.5198e+10 3.3370e+12 5026.6
## - Firecounts  1 3.5157e+12 6.7675e+12 5177.9
```

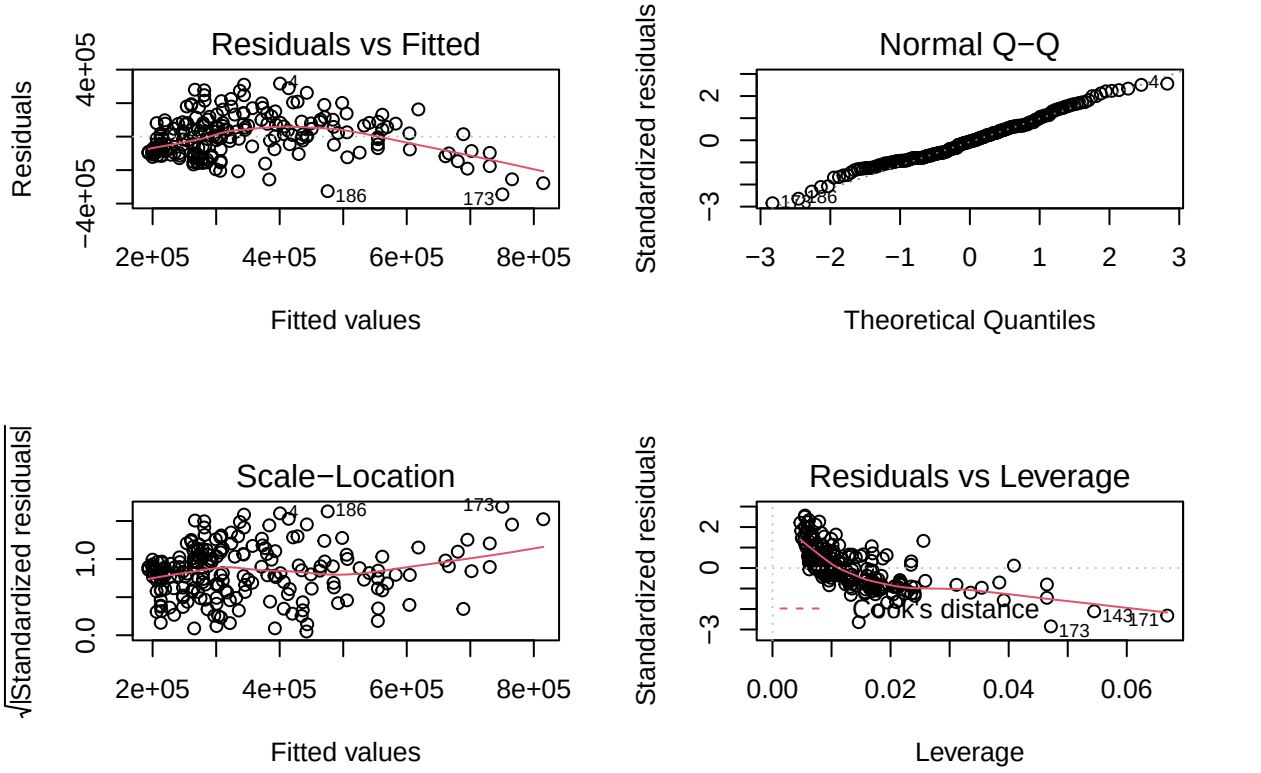
```
##
## Call:
## lm(formula = Visits ~ Month + Firecounts, data = yosefire)
##
## Coefficients:
## (Intercept)        Month   Firecounts
##      180011         6438         7022
```

```
mod.lmfin <- update(mod.lmfull,.~.-Year-Sumfire)
summary(mod.lmfin)
```

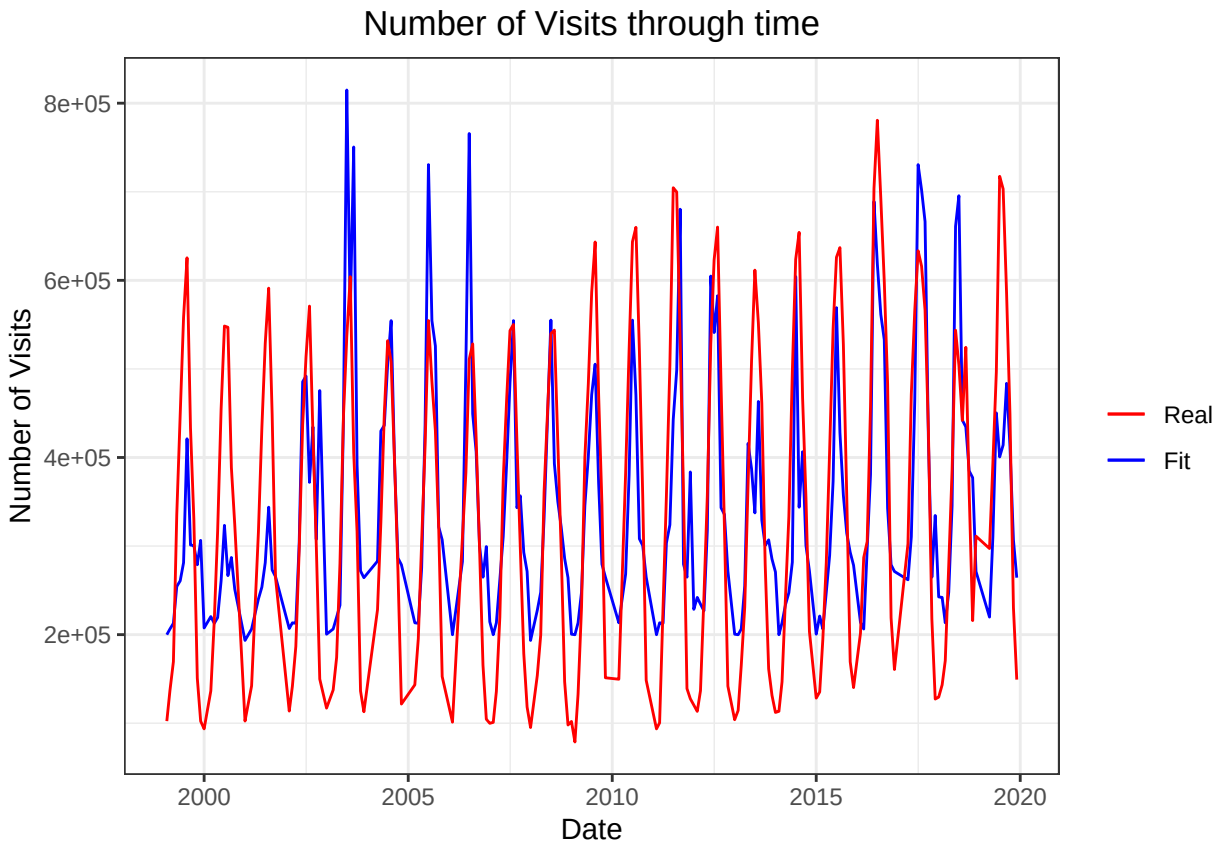
```
##
## Call:
## lm(formula = Visits ~ Month + Firecounts, data = yosefire)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -344926  -96364   -7274   78680  316843
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 180010.9    20683.7   8.703 9.32e-16 ***
## Month         6438.5     2738.4   2.351   0.0196 *
## Firecounts    7021.6      464.9  15.104  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 124100 on 211 degrees of freedom
## Multiple R-squared:  0.5463, Adjusted R-squared:  0.542
## F-statistic:   127 on 2 and 211 DF,  p-value: < 2.2e-16
```

```
par(mfrow=c(2,2))
plot(mod.lmfin)
```




```
#plot model fit
coef <- mod.lmfin$coefficients
yosefire$fit <- coef[1]+ coef[2]*yosefire$Month + coef[3]*yosefire$Firecounts
ggplot() +
  geom_line(data = yosefire, aes(x = Date, y = fit, color="Fit")) +
  geom_line(data = yosefire, aes(x = Date, y = Visits, color = "Real")) +
  theme_bw() +
  ggtitle("Number of Visits through time") +
  theme(plot.title = element_text(hjust = 0.5)) +
  xlab("Date") +
  ylab("Number of Visits")+
  scale_color_manual(name = "", values = c("Real" = "red", "Fit" = "blue"))
```


Number of Visits through time
8e+05
6e+05
4e+05
2e+05
Number of Visits
2000
2005
2010
2015
2020
Date
Real
Fit