Project Proposal

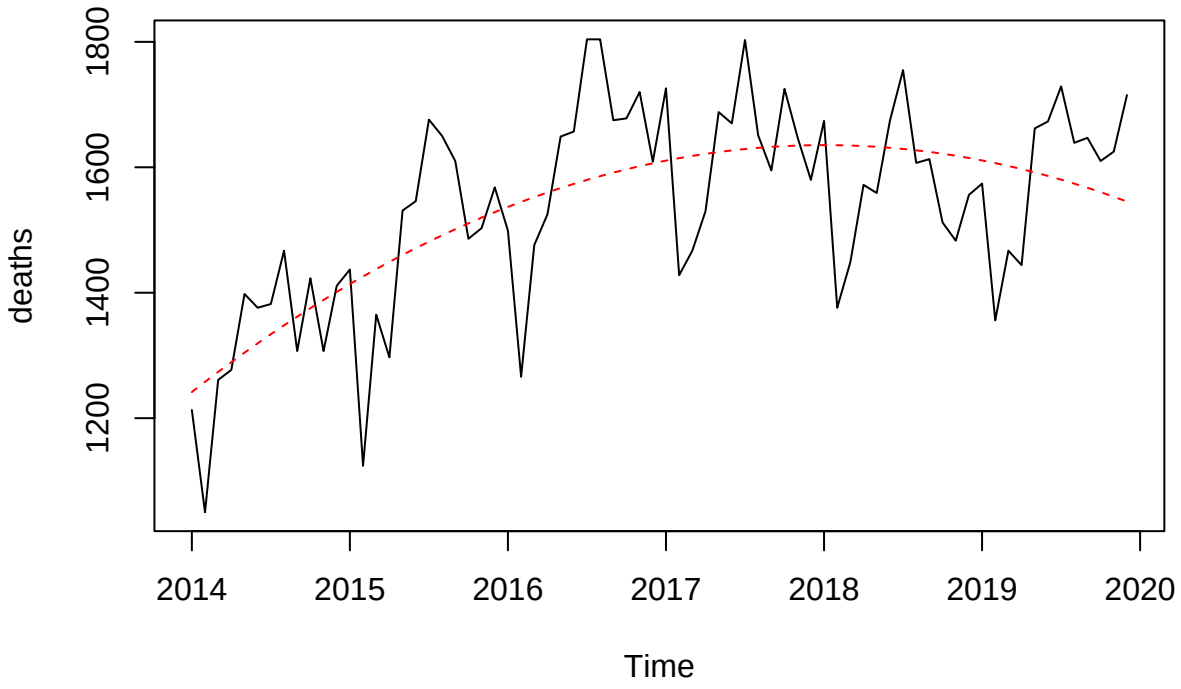
Alex Cluff (cluff.18)

4/12/2021

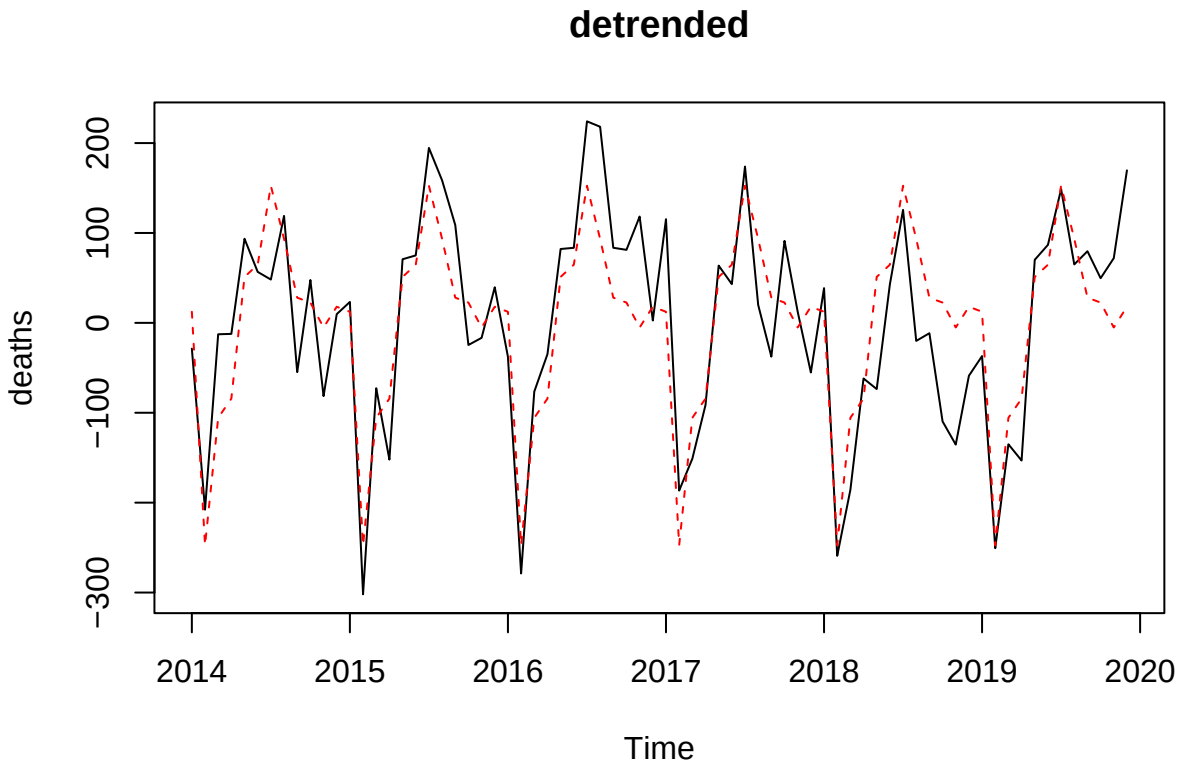
The data is monthly counts of deaths by select causes from 2014-2019 in the US. The data was downloaded from https://catalog.data.gov/dataset/monthly-counts-of-deaths-by-select-causes-2014-2019. I have decided on using the homicide data.

```
library(readr)
deaths.all <- read_csv("Monthly_Counts_of_Deaths_by_Select_Causes__2014-2019.csv")
sorted.deaths <- data.frame(Homicides = dplyr::arrange(deaths.all, Year, Month)$`Assault (Homicide)`)
write_csv(sorted.deaths,"deaths.csv")
```

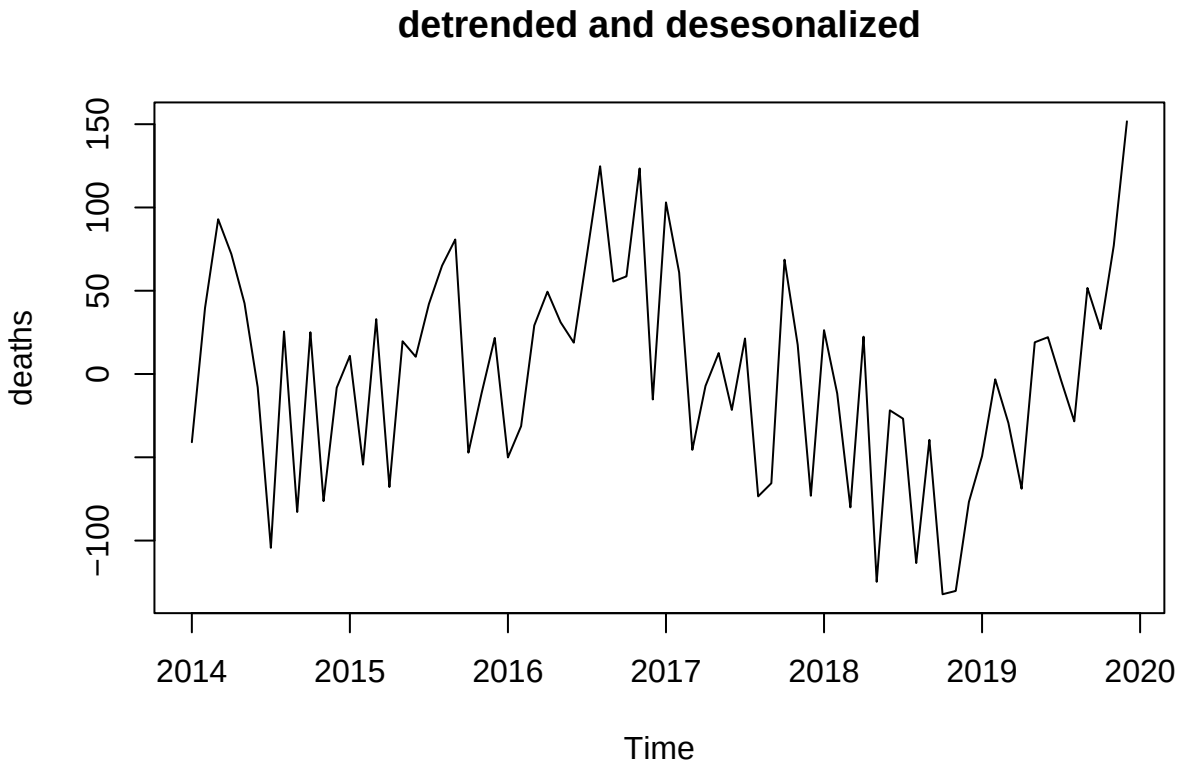
```
deaths <- ts(sorted.deaths, start = 2014, frequency = 12)

t <- time(deaths)
coef <- lm(deaths ~ poly(t, 2, raw = TRUE))$coefficients

trend <- coef[1] + coef[2]*t + coef[3]*t^2
plot(deaths, ylab = "deaths")
lines(trend, col = "red", lty = 2)
```



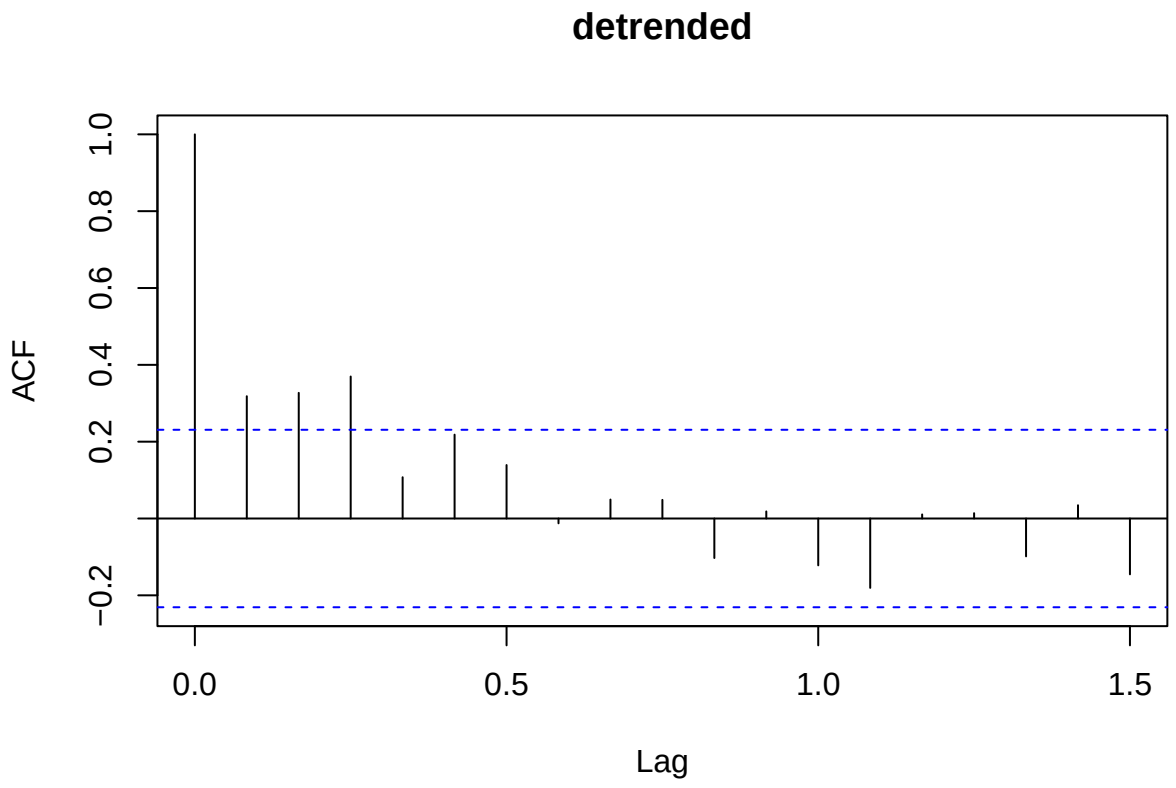

```
detrended <- deaths - trend

M = factor(rep(month.abb, length.out = 72), levels = month.abb)
seasonal.means <- lm(detrended ~ M + 0)$coefficients
seasonality <- ts(rep(seasonal.means, length.out = 72), start = 2014, frequency = 12)


plot(detrended, ylab = "deaths", main = "detrended")
lines(seasonality, col = "red", lty = 2)
```



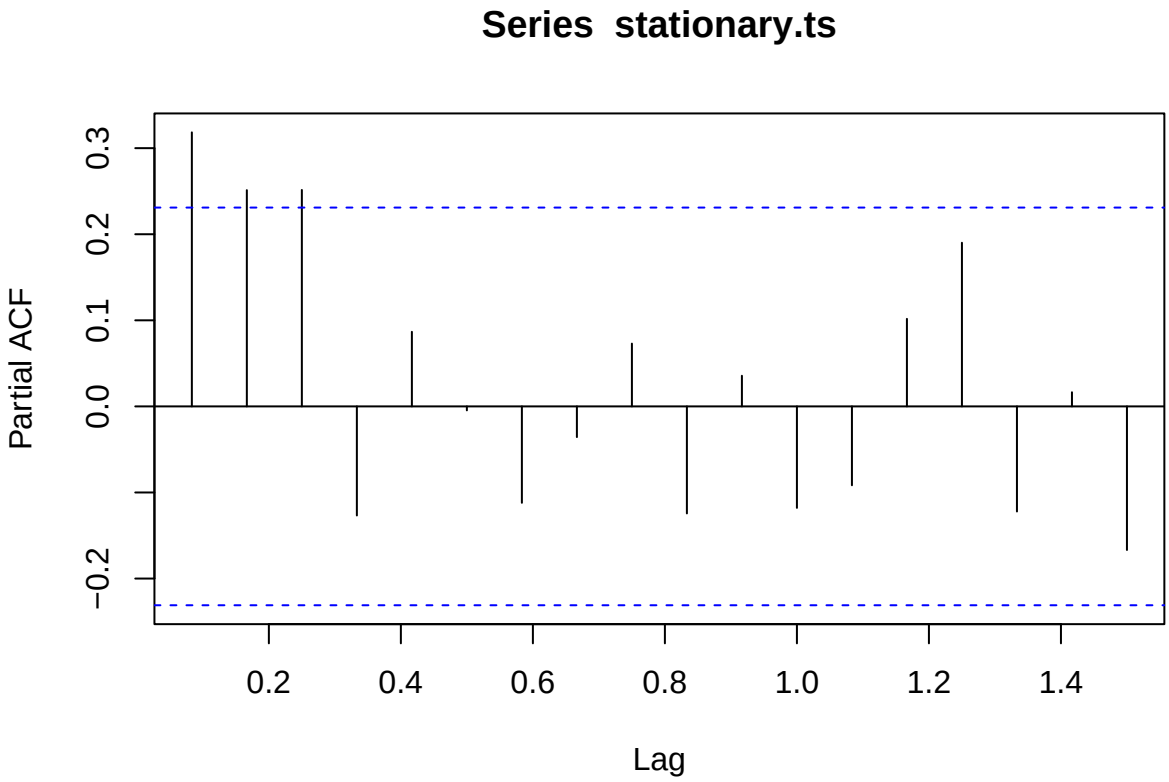

```
stationary.ts <- detrended - seasonality

plot(stationary.ts, ylab = "deaths", main = "detrended and desesonalized")
```




```
acf(stationary.ts)
```




```
pacf(stationary.ts)
```


Series stationary.ts
Partial ACF
0.3
0.2
0.1
0.0
−0.2
0.2
0.4
0.6
0.8
1.0
1.2
1.4
Lag