

Assignment 8: Time Series Analysis

Andrew Barfield

Fall 2024

OVERVIEW

This exercise accompanies the lessons in Environmental Data Analytics on generalized linear models.

Directions

1. Rename this file `<FirstLast>_A08_TimeSeries.Rmd` (replacing `<FirstLast>` with your first and last name).
2. Change “Student Name” on line 3 (above) with your name.
3. Work through the steps, **creating code and output** that fulfill each instruction.
4. Be sure to **answer the questions** in this assignment document.
5. When you have completed the assignment, **Knit** the text and code into a single PDF file.

Set up

1. Set up your session:
 - Check your working directory
 - Load the tidyverse, lubridate, zoo, and trend packages
 - Set your ggplot theme

```
library(here)
```

```
## here() starts at /home/guest/EDE_Fall2024
```

```
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(lubridate)
library(zoo)
```

```
##
## Attaching package: 'zoo'
##
## The following objects are masked from 'package:base':
##
##      as.Date, as.Date.numeric
```

```
library(trend)
library(dplyr)

here()
```

```
## [1] "/home/guest/EDE_Fall2024"
```

```
mytheme <- theme_classic() +
  theme(legend.position = 'bottom',
        plot.title = element_text(size = 15, color = 'black', hjust = 0.5),
        axis.text = element_text(color = 'black', size = 8))
theme_set(mytheme)
```

2. Import the ten datasets from the Ozone_TimeSeries folder in the Raw data folder. These contain ozone concentrations at Garinger High School in North Carolina from 2010-2019 (the EPA air database only allows downloads for one year at a time). Import these either individually or in bulk and then combine them into a single dataframe named **GaringerOzone** of 3589 observation and 20 variables.

```
#1

EPA.2010 <- read.csv(here("Data/Raw/Ozone_TimeSeries/EPAair_03_GaringerNC2010_raw.csv"),
  stringsAsFactors = TRUE)
EPA.2011 <- read.csv(here("Data/Raw/Ozone_TimeSeries/EPAair_03_GaringerNC2011_raw.csv"),
  stringsAsFactors = TRUE)
EPA.2012 <- read.csv(here("Data/Raw/Ozone_TimeSeries/EPAair_03_GaringerNC2012_raw.csv"),
  stringsAsFactors = TRUE)
EPA.2013 <- read.csv(here("Data/Raw/Ozone_TimeSeries/EPAair_03_GaringerNC2013_raw.csv"),
  stringsAsFactors = TRUE)
EPA.2014 <- read.csv(here("Data/Raw/Ozone_TimeSeries/EPAair_03_GaringerNC2014_raw.csv"),
  stringsAsFactors = TRUE)
EPA.2015 <- read.csv(here("Data/Raw/Ozone_TimeSeries/EPAair_03_GaringerNC2015_raw.csv"),
  stringsAsFactors = TRUE)
EPA.2016 <- read.csv(here("Data/Raw/Ozone_TimeSeries/EPAair_03_GaringerNC2016_raw.csv"),
  stringsAsFactors = TRUE)
EPA.2017 <- read.csv(here("Data/Raw/Ozone_TimeSeries/EPAair_03_GaringerNC2017_raw.csv"),
  stringsAsFactors = TRUE)
EPA.2018 <- read.csv(here("Data/Raw/Ozone_TimeSeries/EPAair_03_GaringerNC2018_raw.csv"),
  stringsAsFactors = TRUE)
EPA.2019 <- read.csv(here("Data/Raw/Ozone_TimeSeries/EPAair_03_GaringerNC2019_raw.csv"),
  stringsAsFactors = TRUE)
```

```
GaringerOzone <- rbind(EPA.2010, EPA.2011, EPA.2012, EPA.2013,
                      EPA.2014, EPA.2015, EPA.2016, EPA.2017,
                      EPA.2018, EPA.2019)
```

Wrangle

3. Set your date column as a date class.
4. Wrangle your dataset so that it only contains the columns Date, Daily.Max.8.hour.Ozone.Concentration, and DAILY_AQI_VALUE.
5. Notice there are a few days in each year that are missing ozone concentrations. We want to generate a daily dataset, so we will need to fill in any missing days with NA. Create a new data frame that contains a sequence of dates from 2010-01-01 to 2019-12-31 (hint: `as.data.frame(seq())`). Call this new data frame Days. Rename the column name in Days to "Date".
6. Use a `left_join` to combine the data frames. Specify the correct order of data frames within this function so that the final dimensions are 3652 rows and 3 columns. Call your combined data frame GaringerOzone.

```
# 3
GaringerOzone$Date <- as.Date(GaringerOzone$Date, format = "%m/%d/%Y")

# 4
GaringerOzone.Wrangle <-
  GaringerOzone %>%
  select(Date, Daily.Max.8.hour.Ozone.Concentration, DAILY_AQI_VALUE)

# 5
Days <-
  as.data.frame(seq(as.Date("2010-01-01"), as.Date("2019-12-31"), by = "day"))
colnames(Days)[1] = "Date"

# 6
GaringerOzone <-
  left_join(GaringerOzone.Wrangle, Days)
```

```
## Joining with 'by = join_by(Date)'
```

Visualize

7. Create a line plot depicting ozone concentrations over time. In this case, we will plot actual concentrations in ppm, not AQI values. Format your axes accordingly. Add a smoothed line showing any linear trend of your data. Does your plot suggest a trend in ozone concentration over time?

```
#7
Ozone.Concentrations.Plot <-
  GaringerOzone %>%
  ggplot(aes(x = Date,
             y = Daily.Max.8.hour.Ozone.Concentration)) +
  geom_line() +
  geom_smooth(method = "lm") +
```

```
ggtitle("Ozone Concentrations Over Time") +
ylab("Concentrations (ppm)")
mytheme
```

```
## List of 136
## $ line :List of 6
## ..$ colour : chr "black"
## ..$ linewidth : num 0.5
## ..$ linetype : num 1
## ..$ lineend : chr "butt"
## ..$ arrow : logi FALSE
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_line" "element"
## $ rect :List of 5
## ..$ fill : chr "white"
## ..$ colour : chr "black"
## ..$ linewidth : num 0.5
## ..$ linetype : num 1
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_rect" "element"
## $ text :List of 11
## ..$ family : chr ""
## ..$ face : chr "plain"
## ..$ colour : chr "black"
## ..$ size : num 11
## ..$ hjust : num 0.5
## ..$ vjust : num 0.5
## ..$ angle : num 0
## ..$ lineheight : num 0.9
## ..$ margin : 'margin' num [1:4] 0points 0points 0points 0points
## .. ..- attr(*, "unit")= int 8
## ..$ debug : logi FALSE
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ title : NULL
## $ aspect.ratio : NULL
## $ axis.title : NULL
## $ axis.title.x :List of 11
## ..$ family : NULL
## ..$ face : NULL
## ..$ colour : NULL
## ..$ size : NULL
## ..$ hjust : NULL
## ..$ vjust : num 1
## ..$ angle : NULL
## ..$ lineheight : NULL
## ..$ margin : 'margin' num [1:4] 2.75points 0points 0points 0points
## .. ..- attr(*, "unit")= int 8
## ..$ debug : NULL
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.title.x.top :List of 11
## ..$ family : NULL
```

```

## ..$ face          : NULL
## ..$ colour        : NULL
## ..$ size          : NULL
## ..$ hjust         : NULL
## ..$ vjust         : num 0
## ..$ angle         : NULL
## ..$ lineheight    : NULL
## ..$ margin        : 'margin' num [1:4] 0points 0points 2.75points 0points
## ..- attr(*, "unit")= int 8
## ..$ debug         : NULL
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.title.x.bottom      : NULL
## $ axis.title.y             :List of 11
## ..$ family              : NULL
## ..$ face                : NULL
## ..$ colour              : NULL
## ..$ size                : NULL
## ..$ hjust               : NULL
## ..$ vjust               : num 1
## ..$ angle               : num 90
## ..$ lineheight          : NULL
## ..$ margin              : 'margin' num [1:4] 0points 2.75points 0points 0points
## ..- attr(*, "unit")= int 8
## ..$ debug               : NULL
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.title.y.left       : NULL
## $ axis.title.y.right      :List of 11
## ..$ family              : NULL
## ..$ face                : NULL
## ..$ colour              : NULL
## ..$ size                : NULL
## ..$ hjust               : NULL
## ..$ vjust               : num 1
## ..$ angle               : num -90
## ..$ lineheight          : NULL
## ..$ margin              : 'margin' num [1:4] 0points 0points 0points 2.75points
## ..- attr(*, "unit")= int 8
## ..$ debug               : NULL
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.text               :List of 11
## ..$ family              : NULL
## ..$ face                : NULL
## ..$ colour              : chr "black"
## ..$ size                : num 8
## ..$ hjust               : NULL
## ..$ vjust               : NULL
## ..$ angle               : NULL
## ..$ lineheight          : NULL
## ..$ margin              : NULL
## ..$ debug               : NULL
## ..$ inherit.blank: logi FALSE

```

```

##   ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.text.x                               :List of 11
##   ..$ family           : NULL
##   ..$ face             : NULL
##   ..$ colour          : NULL
##   ..$ size             : NULL
##   ..$ hjust            : NULL
##   ..$ vjust            : num 1
##   ..$ angle            : NULL
##   ..$ lineheight       : NULL
##   ..$ margin           : 'margin' num [1:4] 2.2points 0points 0points 0points
##   ..- attr(*, "unit")= int 8
##   ..$ debug            : NULL
##   ..$ inherit.blank: logi TRUE
##   ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.text.x.top                               :List of 11
##   ..$ family           : NULL
##   ..$ face             : NULL
##   ..$ colour          : NULL
##   ..$ size             : NULL
##   ..$ hjust            : NULL
##   ..$ vjust            : num 0
##   ..$ angle            : NULL
##   ..$ lineheight       : NULL
##   ..$ margin           : 'margin' num [1:4] 0points 0points 2.2points 0points
##   ..- attr(*, "unit")= int 8
##   ..$ debug            : NULL
##   ..$ inherit.blank: logi TRUE
##   ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.text.x.bottom                           : NULL
## $ axis.text.y                               :List of 11
##   ..$ family           : NULL
##   ..$ face             : NULL
##   ..$ colour          : NULL
##   ..$ size             : NULL
##   ..$ hjust            : num 1
##   ..$ vjust            : NULL
##   ..$ angle            : NULL
##   ..$ lineheight       : NULL
##   ..$ margin           : 'margin' num [1:4] 0points 2.2points 0points 0points
##   ..- attr(*, "unit")= int 8
##   ..$ debug            : NULL
##   ..$ inherit.blank: logi TRUE
##   ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.text.y.left                             : NULL
## $ axis.text.y.right                           :List of 11
##   ..$ family           : NULL
##   ..$ face             : NULL
##   ..$ colour          : NULL
##   ..$ size             : NULL
##   ..$ hjust            : num 0
##   ..$ vjust            : NULL
##   ..$ angle            : NULL
##   ..$ lineheight       : NULL

```

```

## ..$ margin      : 'margin' num [1:4] 0points 0points 0points 2.2points
## .. ..- attr(*, "unit")= int 8
## ..$ debug       : NULL
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.text.theta : NULL
## $ axis.text.r      :List of 11
## ..$ family        : NULL
## ..$ face          : NULL
## ..$ colour        : NULL
## ..$ size          : NULL
## ..$ hjust         : num 0.5
## ..$ vjust         : NULL
## ..$ angle         : NULL
## ..$ lineheight    : NULL
## ..$ margin      : 'margin' num [1:4] 0points 2.2points 0points 2.2points
## .. ..- attr(*, "unit")= int 8
## ..$ debug       : NULL
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.ticks     :List of 6
## ..$ colour       : chr "grey20"
## ..$ linewidth    : NULL
## ..$ linetype     : NULL
## ..$ lineend      : NULL
## ..$ arrow        : logi FALSE
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_line" "element"
## $ axis.ticks.x    : NULL
## $ axis.ticks.x.top : NULL
## $ axis.ticks.x.bottom : NULL
## $ axis.ticks.y    : NULL
## $ axis.ticks.y.left : NULL
## $ axis.ticks.y.right : NULL
## $ axis.ticks.theta : NULL
## $ axis.ticks.r     : NULL
## $ axis.minor.ticks.x.top : NULL
## $ axis.minor.ticks.x.bottom : NULL
## $ axis.minor.ticks.y.left : NULL
## $ axis.minor.ticks.y.right : NULL
## $ axis.minor.ticks.theta : NULL
## $ axis.minor.ticks.r     : NULL
## $ axis.ticks.length     : 'simpleUnit' num 2.75points
## ..- attr(*, "unit")= int 8
## $ axis.ticks.length.x   : NULL
## $ axis.ticks.length.x.top : NULL
## $ axis.ticks.length.x.bottom : NULL
## $ axis.ticks.length.y   : NULL
## $ axis.ticks.length.y.left : NULL
## $ axis.ticks.length.y.right : NULL
## $ axis.ticks.length.theta : NULL
## $ axis.ticks.length.r     : NULL
## $ axis.minor.ticks.length : 'rel' num 0.75
## $ axis.minor.ticks.length.x : NULL

```

```

## $ axis.minor.ticks.length.x.top : NULL
## $ axis.minor.ticks.length.x.bottom: NULL
## $ axis.minor.ticks.length.y : NULL
## $ axis.minor.ticks.length.y.left : NULL
## $ axis.minor.ticks.length.y.right : NULL
## $ axis.minor.ticks.length.theta : NULL
## $ axis.minor.ticks.length.r : NULL
## $ axis.line :List of 6
## ..$ colour : chr "black"
## ..$ linewidth : 'rel' num 1
## ..$ linetype : NULL
## ..$ lineend : NULL
## ..$ arrow : logi FALSE
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_line" "element"
## $ axis.line.x : NULL
## $ axis.line.x.top : NULL
## $ axis.line.x.bottom : NULL
## $ axis.line.y : NULL
## $ axis.line.y.left : NULL
## $ axis.line.y.right : NULL
## $ axis.line.theta : NULL
## $ axis.line.r : NULL
## $ legend.background :List of 5
## ..$ fill : NULL
## ..$ colour : logi NA
## ..$ linewidth : NULL
## ..$ linetype : NULL
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_rect" "element"
## $ legend.margin : 'margin' num [1:4] 5.5points 5.5points 5.5points 5.5points
## ..- attr(*, "unit")= int 8
## $ legend.spacing : 'simpleUnit' num 11points
## ..- attr(*, "unit")= int 8
## $ legend.spacing.x : NULL
## $ legend.spacing.y : NULL
## $ legend.key : NULL
## $ legend.key.size : 'simpleUnit' num 1.2lines
## ..- attr(*, "unit")= int 3
## $ legend.key.height : NULL
## $ legend.key.width : NULL
## $ legend.key.spacing : 'simpleUnit' num 5.5points
## ..- attr(*, "unit")= int 8
## $ legend.key.spacing.x : NULL
## $ legend.key.spacing.y : NULL
## $ legend.frame : NULL
## $ legend.ticks : NULL
## $ legend.ticks.length : 'rel' num 0.2
## $ legend.axis.line : NULL
## $ legend.text :List of 11
## ..$ family : NULL
## ..$ face : NULL
## ..$ colour : NULL
## ..$ size : 'rel' num 0.8

```



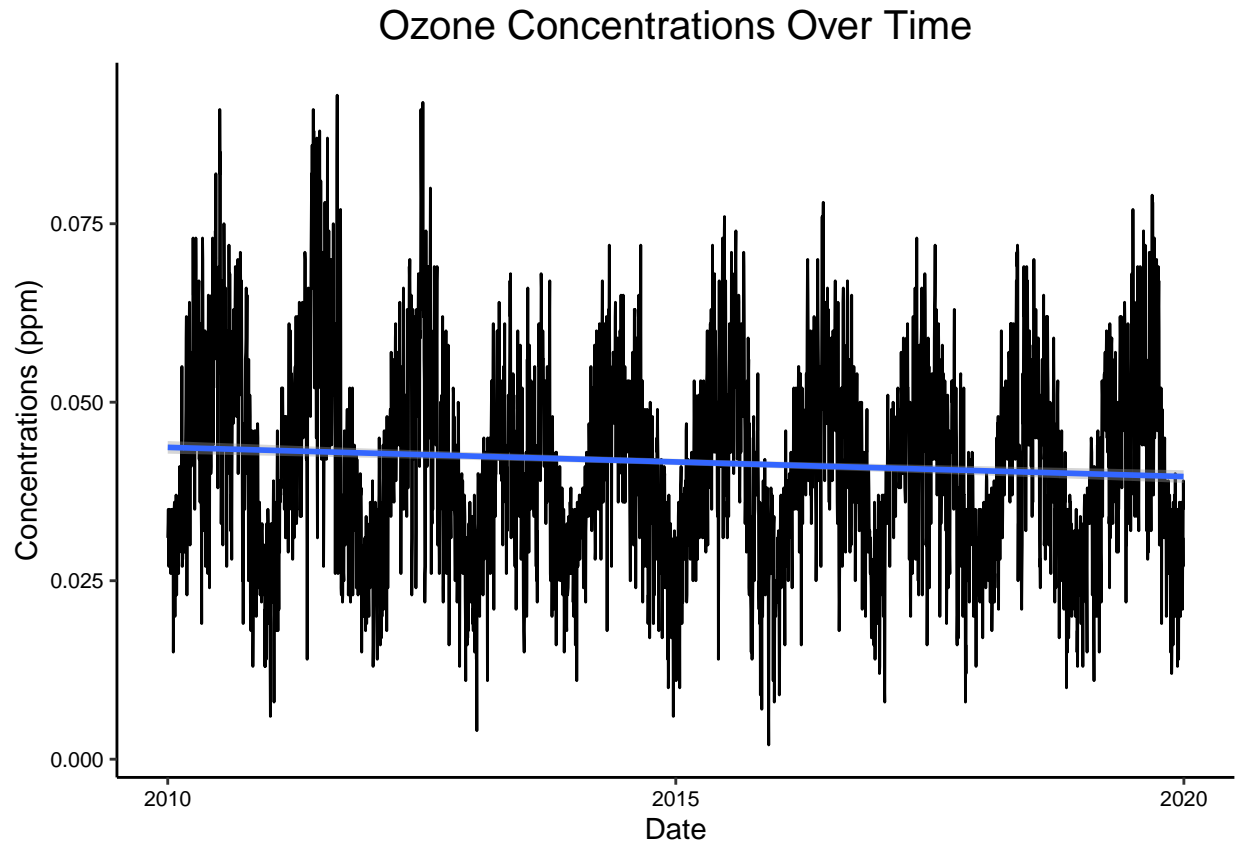
```

## ..$ hjust      : NULL
## ..$ vjust      : NULL
## ..$ angle      : NULL
## ..$ lineheight : NULL
## ..$ margin     : NULL
## ..$ debug      : NULL
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ legend.text.position : NULL
## $ legend.title         :List of 11
## ..$ family            : NULL
## ..$ face              : NULL
## ..$ colour            : NULL
## ..$ size              : NULL
## ..$ hjust             : num 0
## ..$ vjust             : NULL
## ..$ angle            : NULL
## ..$ lineheight       : NULL
## ..$ margin           : NULL
## ..$ debug            : NULL
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ legend.title.position : NULL
## $ legend.position       : chr "bottom"
## $ legend.position.inside : NULL
## $ legend.direction      : NULL
## $ legend.byrow          : NULL
## $ legend.justification  : chr "center"
## $ legend.justification.top : NULL
## $ legend.justification.bottom : NULL
## $ legend.justification.left : NULL
## $ legend.justification.right : NULL
## $ legend.justification.inside : NULL
## $ legend.location       : NULL
## $ legend.box            : NULL
## $ legend.box.just       : NULL
## $ legend.box.margin     : 'margin' num [1:4] 0cm 0cm 0cm 0cm
## ..- attr(*, "unit")= int 1
## $ legend.box.background : list()
## ..- attr(*, "class")= chr [1:2] "element_blank" "element"
## $ legend.box.spacing    : 'simpleUnit' num 11points
## ..- attr(*, "unit")= int 8
## [list output truncated]
## - attr(*, "class")= chr [1:2] "theme" "gg"
## - attr(*, "complete")= logi TRUE
## - attr(*, "validate")= logi TRUE

```

```
print(Ozone.Concentrations.Plot)
```

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



Answer: The plot suggests a slight decrease in ozone concentrations over time, albeit not by very much.

Time Series Analysis

Study question: Have ozone concentrations changed over the 2010s at this station?

8. Use a linear interpolation to fill in missing daily data for ozone concentration. Why didn't we use a piecewise constant or spline interpolation?

```
#8
Ozone.Concentration.Clean <-
  GaringerOzone %>%
  mutate(Daily.Max.8.hour.Ozone.Concentration =
    na.approx(Daily.Max.8.hour.Ozone.Concentration))
```

Answer: We did not use a piecewise constant because there is significant seasonality and oscillations within our data. A piecewise constant is best used when you have little seasonality or variation in your dataset, so if we used it in our case the NA values could have been over or underestimated. We didn't use a spline interpolation because our function showed a linear trend according to our ggplot.

9. Create a new data frame called `GaringerOzone.monthly` that contains aggregated data: mean ozone concentrations for each month. In your pipe, you will need to first add columns for year and month

to form the groupings. In a separate line of code, create a new Date column with each month-year combination being set as the first day of the month (this is for graphing purposes only)

```
#9
GaringerOzone.monthly <-
  GaringerOzone %>%
  mutate(year = year(Date), month = month(Date)) %>%
  aggregate(Daily.Max.8.hour.Ozone.Concentration ~ month + year, mean)

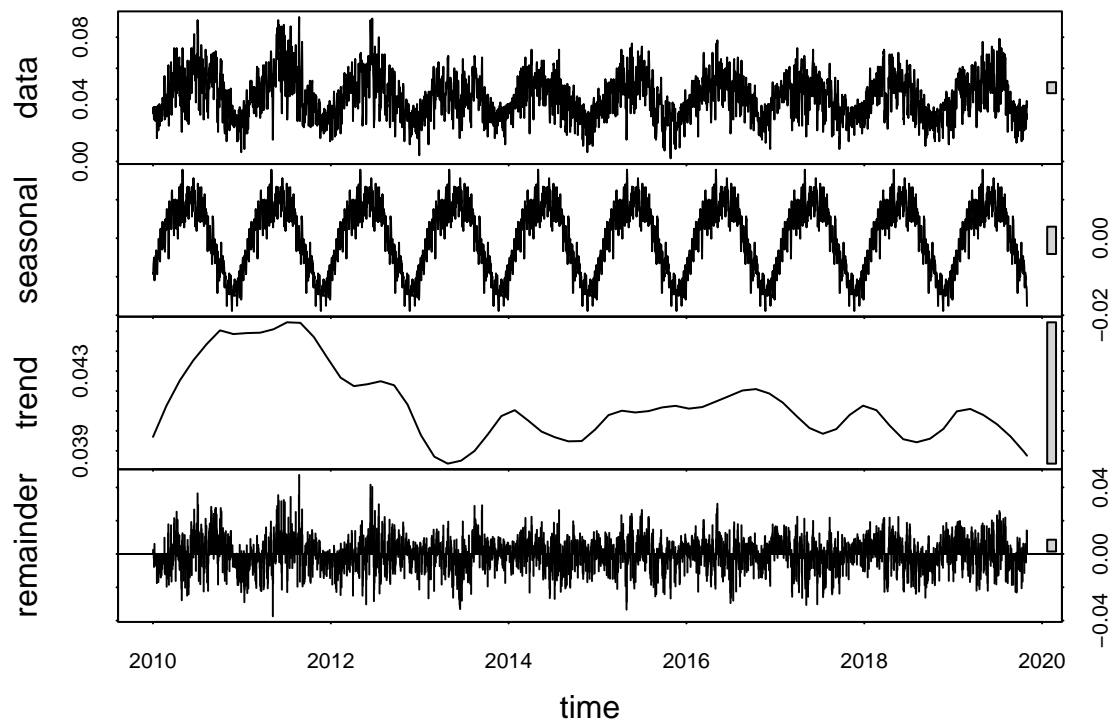
GaringerOzone.monthly$Date <-
  as.Date(paste(GaringerOzone.monthly$year, GaringerOzone.monthly$month,
                1, sep = "-"))
```

10. Generate two time series objects. Name the first `GaringerOzone.daily.ts` and base it on the dataframe of daily observations. Name the second `GaringerOzone.monthly.ts` and base it on the monthly average ozone values. Be sure that each specifies the correct start and end dates and the frequency of the time series.

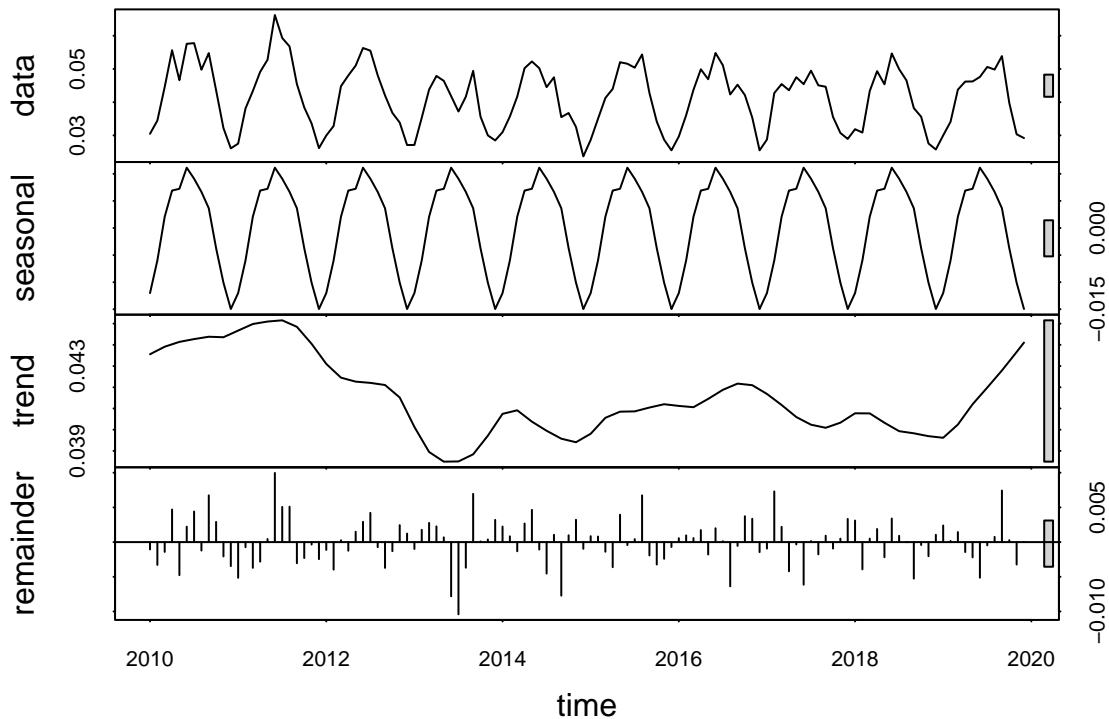
```
#10
GaringerOzone.daily.ts <- ts(Ozone.Concentration.Clean$Daily.Max.8.hour.Ozone.Concentration,
                             start = c(2010,1),
                             frequency = 365)
GaringerOzone.monthly.ts <- ts(GaringerOzone.monthly$Daily.Max.8.hour.Ozone.Concentration,
                               start = c(2010,1),
                               frequency = 12)
```

11. Decompose the daily and the monthly time series objects and plot the components using the `plot()` function.

```
#11
Daily.ts.decompose <- stl(GaringerOzone.daily.ts, s.window = "periodic")
plot(Daily.ts.decompose)
```



```
Monthly.ts.decompose <- stl(GaringerOzone.monthly.ts, s.window = "periodic")
plot(Monthly.ts.decompose)
```



12. Run a monotonic trend analysis for the monthly Ozone series. In this case the seasonal Mann-Kendall is most appropriate; why is this?

```
#12
Monthly.Trend <- Kendall::SeasonalMannKendall(GaringerOzone.monthly.ts)
summary(Monthly.Trend)

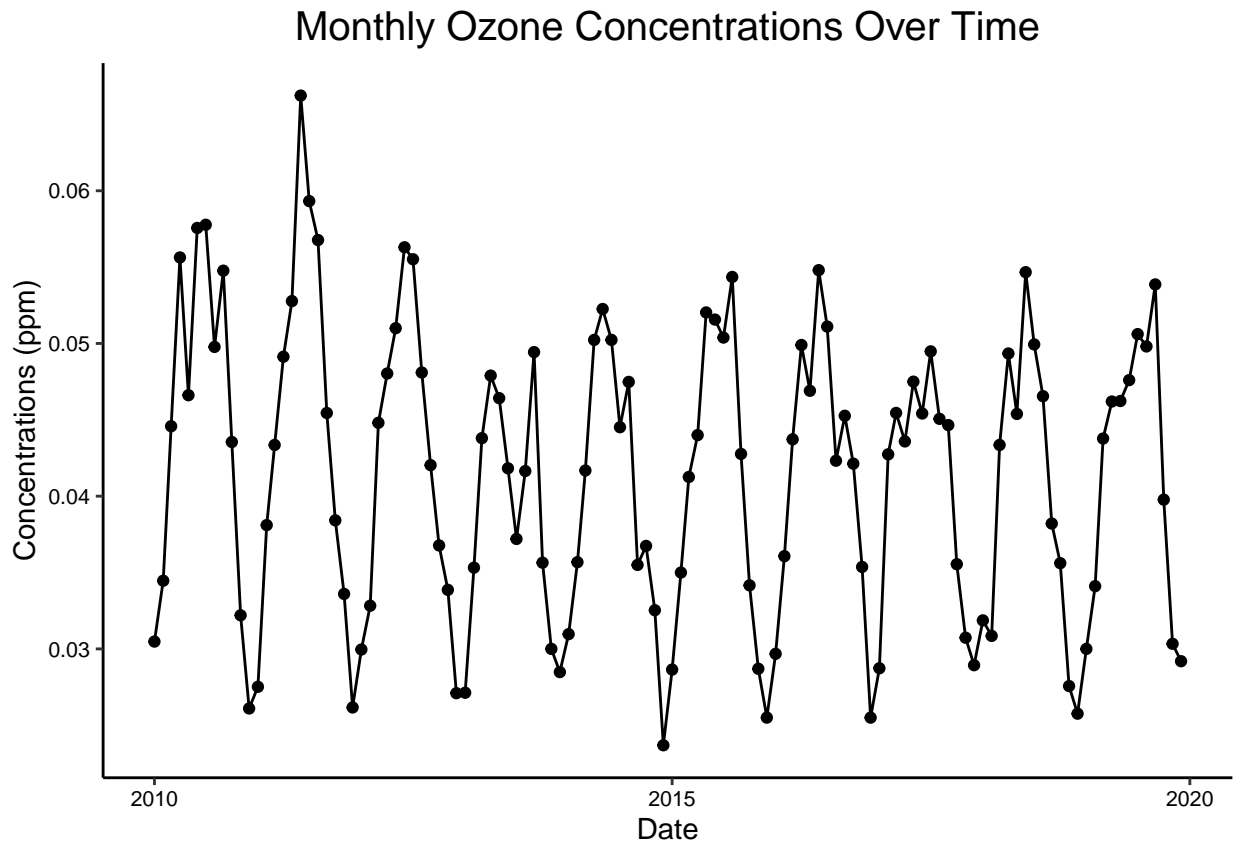
## Score = -88 , Var(Score) = 1498
## denominator = 538.9944
## tau = -0.163, 2-sided pvalue =0.022986
```

Answer: The seasonal Mann-Kendall trend analysis is most appropriate because our data experiences trends in seasonality (repeated patterns of short-term variation). All other monotonic trend analyses are used for datasets that do not experience any seasonality.

13. Create a plot depicting mean monthly ozone concentrations over time, with both a `geom_point` and a `geom_line` layer. Edit your axis labels accordingly.

```
# 13
Monthly.Plot <-
  GaringerOzone.monthly %>%
  ggplot(aes(x = Date,
             y = Daily.Max.8.hour.Ozone.Concentration)) +
  geom_line() +
```

```
geom_point() +
ylab("Concentrations (ppm)") +
ggtitle("Monthly Ozone Concentrations Over Time") +
mytheme
print(Monthly.Plot)
```



14. To accompany your graph, summarize your results in context of the research question. Include output from the statistical test in parentheses at the end of your sentence. Feel free to use multiple sentences in your interpretation.

Answer: The results depict that ozone concentrations have decreased at this particular station throughout the 2010's ($\tau = -0.163$, $p\text{-value} = 0.023$).

15. Subtract the seasonal component from the `GaringerOzone.monthly.ts`. Hint: Look at how we extracted the series components for the `EnoDischarge` on the lesson Rmd file.
16. Run the Mann Kendall test on the non-seasonal Ozone monthly series. Compare the results with the ones obtained with the Seasonal Mann Kendall on the complete series.

```
#15
Monthly.Nonseasonal <- GaringerOzone.monthly.ts - Monthly.ts.decompose$time.series

#16
Monthly.Trend2 <- Kendall::MannKendall(Monthly.Nonseasonal)
summary(Monthly.Trend2)
```

```
## Score = -520 , Var(Score) = 5205498
## denominator = 64619
## tau = -0.00805, 2-sided pvalue =0.82005
```

Answer: The results of the non-seasonal Ozone monthly series reveal that there is less of a negative trend ($\tau = -0.008$, $p\text{-value} = 0.82$) of Ozone concentrations over time when compared to the Seasonal Mann Kendall test. In fact, the slight negative trend represented by the τ value in the non-seasonal test is not statistically significant, meaning we failed to reject the null hypothesis of the variable being stationary over time.