Assignment 6: Time Series Analysis

Lindsay Roth

OVERVIEW

This exercise accompanies the lessons in Hydrologic Data Analysis on time series analysis

Directions

1. Change "Student Name" on line 3 (above) with your name.
2. Work through the steps, **creating code and output** that fulfill each instruction.
3. Be sure to **answer the questions** in this assignment document.
4. When you have completed the assignment, **Knit** the text and code into a single pdf file.
5. After Knitting, submit the completed exercise (pdf file) to the dropbox in Sakai. Add your last name into the file name (e.g., "A06_Salk.html") prior to submission.

The completed exercise is due on 11 October 2019 at 9:00 am.

Setup

1. Verify your working directory is set to the R project file,
2. Load the tidyverse, lubridate, trend, and dataRetrieval packages.
3. Set your ggplot theme (can be theme_classic or something else)
4. Load the ClearCreekDischarge.Monthly.csv file from the processed data folder. Call this data frame ClearCreekDischarge.Monthly.

```
getwd()
```

```
## [1] "/Users/lindsayroth/Documents/MEM 2nd Year/HydroData/Hydrologic_Data_Analysis"
```

```
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------------- tidyv
```

```
## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   1.0.0     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
```

```
## -- Conflicts ------------------------------------------------------------------------- tidyverse_c
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(lubridate)
```

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

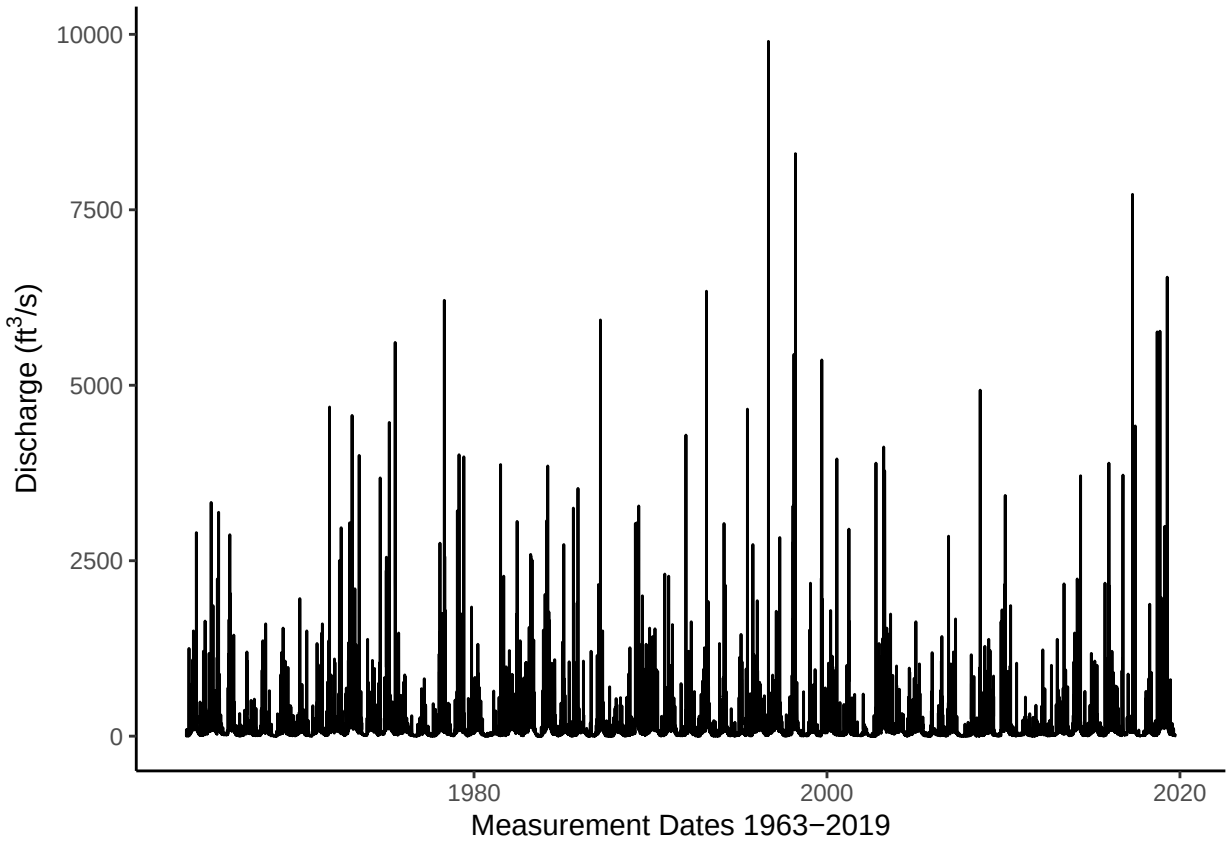
```
## The following object is masked from 'package:base':
##
##     date
```

```r
library(trend)
library(dataRetrieval)

theme_set(theme_classic())

ClearCreekDischarge.Monthly <- read.csv("./Data/Processed/ClearCreekDischarge.Monthly.csv")
```

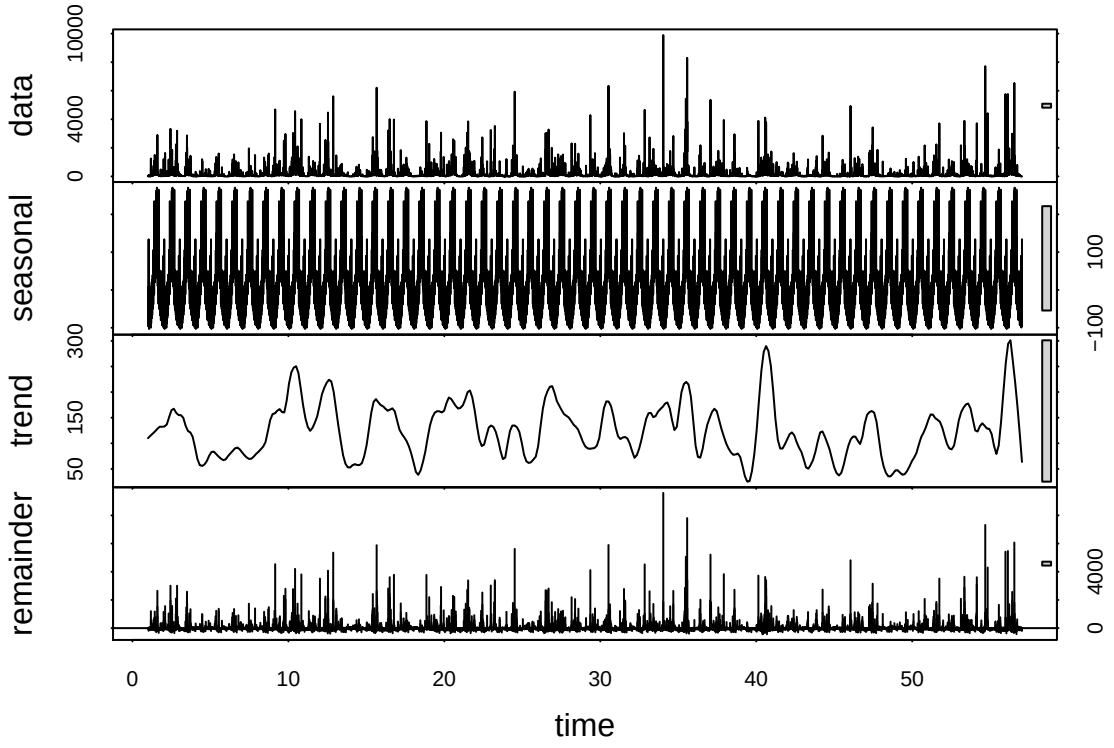
Time Series Decomposition

5. Create a new data frame that includes daily mean discharge at the Eno River for all available dates (`siteNumbers = "02085070"`). Rename the columns accordingly.
6. Plot discharge over time with geom_line. Make sure axis labels are formatted appropriately.
7. Create a time series of discharge
8. Decompose the time series using the `stl` function.
9. Visualize the decomposed time series.

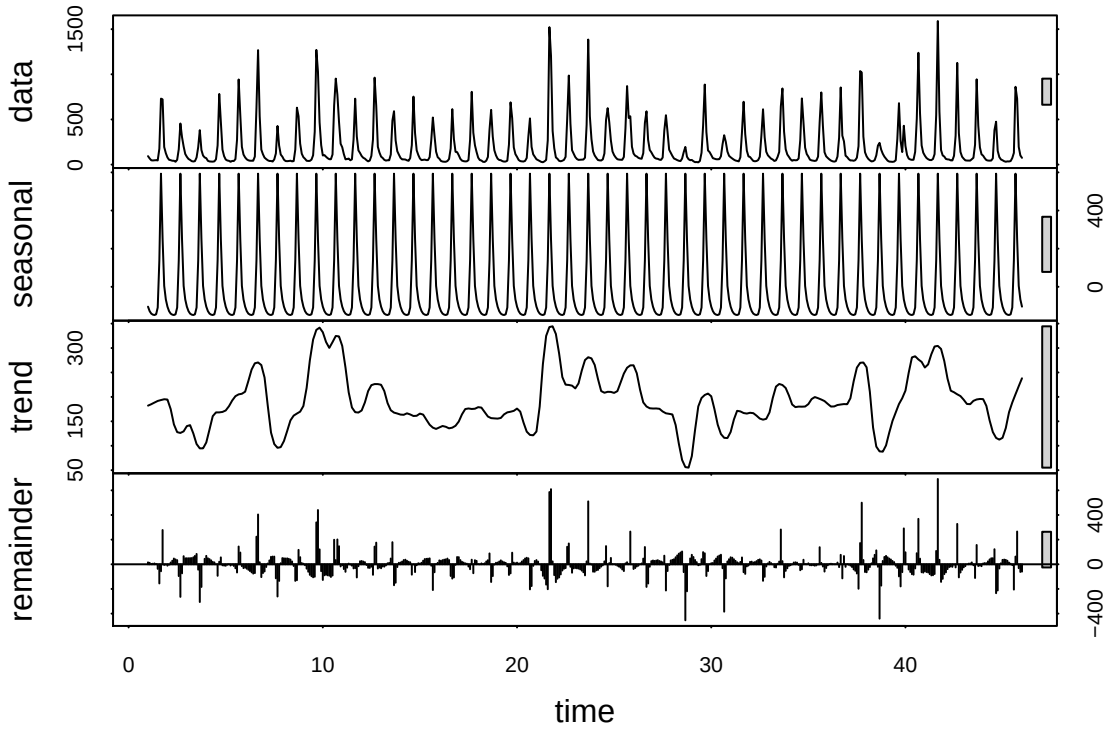
```r
EnoDischarge <- readNWISdv(siteNumbers = "02085070",
                     parameterCd = "00060", # discharge (ft3/s)
                     startDate = "",
                     endDate = "")

names(EnoDischarge)[4:5] <- c("Discharge", "Approval.Code")

EnoDischargePlot <- ggplot(EnoDischarge) +
  geom_line(aes(x = Date, y = Discharge)) +
  labs(x = "Measurement Dates 1963-2019", y = expression("Discharge (ft"^3*"/s)"))
print(EnoDischargePlot)
```




```
EnoDischarge_ts <- ts(EnoDischarge[[4]], frequency = 365)

Eno_Decomposed <- stl(EnoDischarge_ts, s.window = "periodic")

plot(Eno_Decomposed)
```




```
ClearCreekDischarge_ts <- ts(ClearCreekDischarge.Monthly[[3]], frequency = 12)

ClearCreek_Decomposed <- stl(ClearCreekDischarge_ts, s.window = "periodic")

plot(ClearCreek_Decomposed)
```



10. How do the seasonal and trend components of the decomposition compare to the Clear Creek discharge dataset? Are they similar in magnitude?

    Seasonal: The seasonal component has a higher magnitude for Clear Creek has a higher magnitude than the Eno. This makes sense because Clear Creek is monthly data and the Eno discharge data is daily.

    Trend: The trend component has approximately equal magnitude for both Clear Creek and Eno, with Clear Creek's being slightly larger. Both of them have irregular patterns with no clear directional change.

Trend Analysis

Research question: Has there been a monotonic trend in discharge in Clear Creek over the period of study?

11. Generate a time series of monthly discharge in Clear Creek from the ClearCreekDischarge.Monthly data frame. This time series should include just one column (discharge).
12. Run a Seasonal Mann-Kendall test on the monthly discharge data. Inspect the overall trend and the monthly trends.

```
ClearCreek_ts <- ts(ClearCreekDischarge.Monthly[[3]], frequency = 12)

# Run SMK test
ClearCreektrend <- smk.test(ClearCreek_ts)

# Inspect results
ClearCreektrend
```

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data:  ClearCreek_ts
## z = 1.6586, p-value = 0.09719
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##      S   varS
##    590 126102
```

```
summary(ClearCreektrend)
```

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: ClearCreek_ts
## alternative hypothesis: two.sided
##
## Statistics for individual seasons
##
## H0
##                       S  varS    tau      z Pr(>|z|)
## Season 1:   S = 0   64 11154  0.062  0.597 0.550828
## Season 2:   S = 0   24 10450  0.024  0.225 0.821984
## Season 3:   S = 0   30 10450  0.030  0.284 0.776650
## Season 4:   S = 0   35 10449  0.035  0.333 0.739425
## Season 5:   S = 0    4 10450  0.004  0.029 0.976588
## Season 6:   S = 0  204 10450  0.206  1.986 0.047054 *
## Season 7:   S = 0  230 10450  0.232  2.240 0.025081 *
## Season 8:   S = 0  148 10450  0.149  1.438 0.150434
## Season 9:   S = 0   94 10450  0.095  0.910 0.362951
## Season 10:   S = 0 -54 10450 -0.055 -0.518 0.604135
## Season 11:   S = 0 -99 10449 -0.100 -0.959 0.337703
## Season 12:   S = 0 -90 10450 -0.091 -0.871 0.383958
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

13. Is there an overall monotonic trend in discharge over time? If so, is it positive or negative?

    Because the p-value is not less than 0.05, there is not a significant overall monotonic trend in discharge over time. If the p-value would have been below 0.05, the trend would have been positive since the z-score was greater than 0.

14. Are there any monthly monotonic trends in discharge over time? If so, during which months do they occur and are they positive or negative?

    There were two monthly monotonic trends in discharge over time for the months of June and July ($p = 0.047$ and $p = 0.025$, respectively). These trends were both positive because they poth had z-scores greater than 0.

Reflection

15. What are 2-3 conclusions or summary points about time series you learned through your analysis?

    Seasonal fluctuations can vary greatly between sites and just because there are large variations in measurements over time does not mean that there are overall trends.

16. What data, visualizations, and/or models supported your conclusions from 12?

    The seasonal and trend decompositions for the Eno and Clear Creek supported my conclusions.

17. Did hands-on data analysis impact your learning about time series relative to a theory-based lesson? If so, how?

    The breakdown of the seasonal mann-kendall test helped in my understanding of time series analysis.

18. How did the real-world data compare with your expectations from theory?

    The real world data did not differ much from my expectations.