# Assignment 6: Time Series Analysis

*Tristen Townsend*

## 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/Tristen/OneDrive - Duke University/Fall 2019/Hydrologic Data Analysis/Hydrologic_Data_Ana
```

```
library(tidyverse)
```

```
## -- Attaching packages ------------------------------- tidyverse 1.2.1 --

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

## -- Conflicts ---------------------------------- tidyverse_conflicts() --
## 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)
library(tseries)
```

```
## Registered S3 method overwritten by 'xts':
##   method     from
##   as.zoo.xts zoo
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```r
theme_set(theme_classic())

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

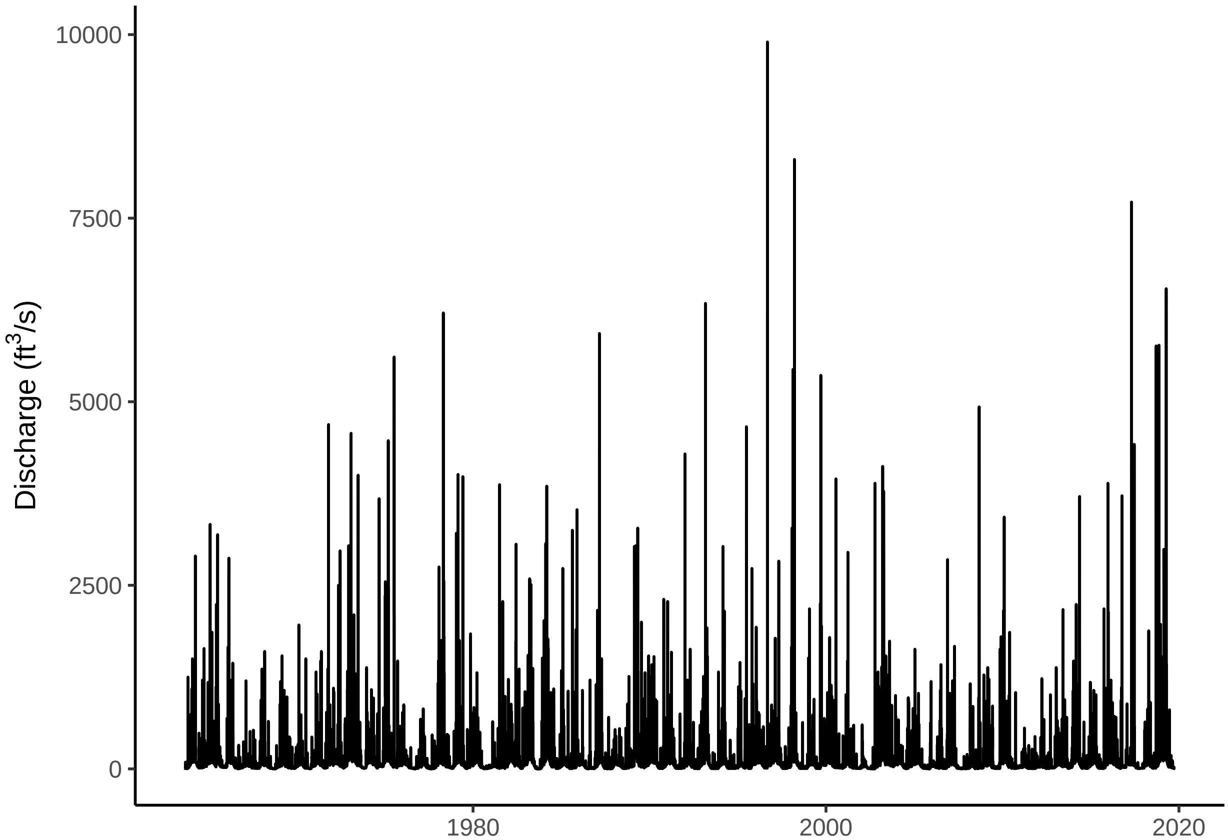
## Time Series Decomposition

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
#Create new dataframe with Eno River discharge
EnoRiverDischarge <- readNWISdv(siteNumbers = "02085070",
                     parameterCd = "00060", # discharge (ft3/s)
                     startDate = "",
                     endDate = "")
names(EnoRiverDischarge)[4:5] <- c("Discharge", "Approval.Code")
class(EnoRiverDischarge$Date)
```

```
## [1] "Date"
```

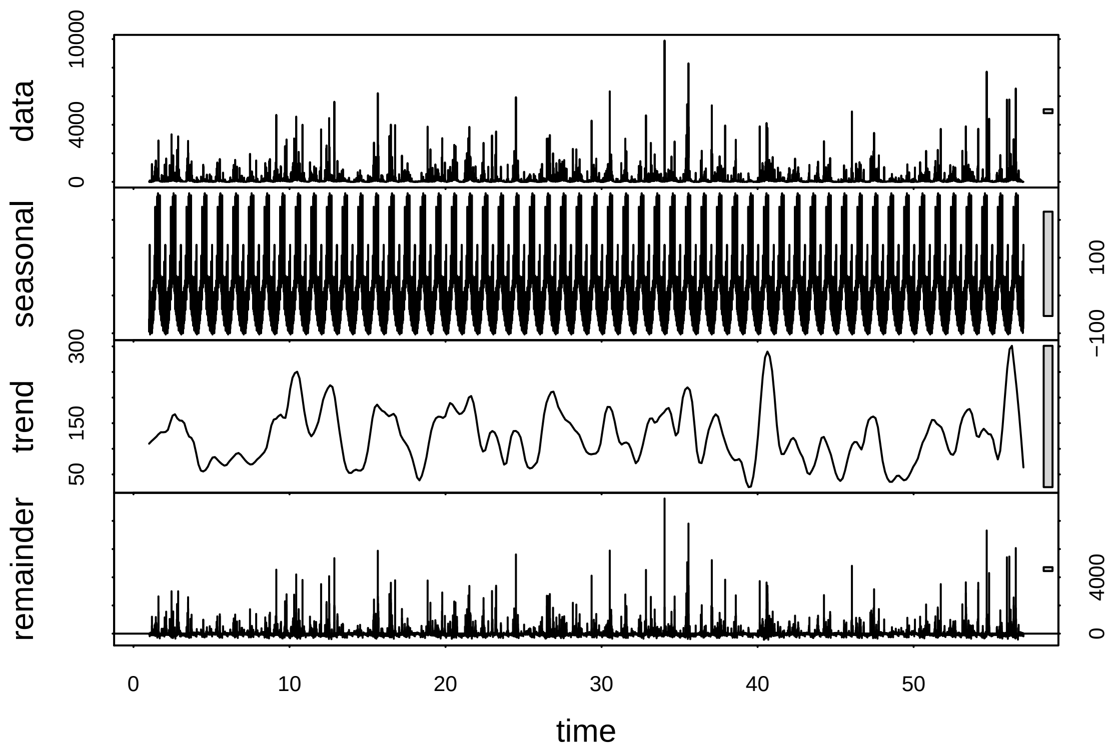
```r
#Plot Eno River discharge
EnoRiverPlot <-
  ggplot(EnoRiverDischarge, aes(x = Date, y = Discharge)) +
  geom_line() +
  labs(x = "", y = expression("Discharge (ft"^3*"/s)")) +
  theme(plot.title = element_text(margin = margin(b = -10), size = 12),
       axis.title.x = element_blank())
print(EnoRiverPlot)
```




```r
#Create time series of discharge data
EnoRiver_ts <- ts(EnoRiverDischarge[[4]], frequency = 365)

#Decompose time series
EnoRiver_Decomposed <- stl(EnoRiver_ts, s.window = "periodic")

#Visualize the decomposed series
plot(EnoRiver_Decomposed)
```




```r
#Extract the components and turn them into data frames
EnoRiver_Components <- as.data.frame(EnoRiver_Decomposed$time.series[,1:3])
EnoRiver_Components <- mutate(EnoRiver_Components,
                        Observed = EnoRiverDischarge$Discharge,
                        Date = EnoRiverDischarge$Date)
```

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 season component for the Eno River is much smaller in magnitude than the Clear Creek discharge dataset.

    Trend: The trend component for the Eno River is more similar to the Clear Creek discharge dataset in magnitude.

## 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.

```r
# Generate time series (smk.test needs ts, not data.frame)
ClearCreek.ts <- ts(ClearCreekDischarge.Monthly$Discharge, frequency = 12,
                        start = c(1974, 10), end = c(2019, 10))
# Run SMK test
```

```r
ClearCreek.trend <- smk.test(ClearCreek.ts)

#Look at results of SMK
ClearCreek.trend #p = 0.09719, Z = 1.6586
```

```
##
##  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
```

```r
summary(ClearCreek.trend)
```

```
##
##  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   35 10449  0.035  0.333 0.739425
## Season 2:   S = 0    4 10450  0.004  0.029 0.976588
## Season 3:   S = 0  204 10450  0.206  1.986 0.047054 *
## Season 4:   S = 0  230 10450  0.232  2.240 0.025081 *
## Season 5:   S = 0  148 10450  0.149  1.438 0.150434
## Season 6:   S = 0   94 10450  0.095  0.910 0.362951
## Season 7:   S = 0  -54 10450 -0.055 -0.518 0.604135
## Season 8:   S = 0  -99 10449 -0.100 -0.959 0.337703
## Season 9:   S = 0  -90 10450 -0.091 -0.871 0.383958
## Season 10:   S = 0  64 11154  0.062  0.597 0.550828
## Season 11:   S = 0  24 10450  0.024  0.225 0.821984
## Season 12:   S = 0  30 10450  0.030  0.284 0.776650
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#Extra test for monotonicity
adf.test(ClearCreek.ts, alternative = "stationary") #p = 0.01
```

```
## Warning in adf.test(ClearCreek.ts, alternative = "stationary"): p-value
## smaller than printed p-value
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  ClearCreek.ts
## Dickey-Fuller = -13.26, Lag order = 8, p-value = 0.01
## alternative hypothesis: stationary
```

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

No, the results of the SMK test do not indicate there is a monotonic trend since the p-value is not less than 0.05. Futhermore, an additional ADF test, also indicates the data is stationary (p-value less than 0.05 which indicates stationarity).

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?

    Yes, there are. They occur in March and April and both are positive trends (p-values lower than 0.05 and positive z-scores).

## Reflection

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

    Time series need to have constant time intervals to be accurate The scale of time you are looking for a change could influence how you interpret your results.

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

    Generation of the time series required inputting the time frequency, so any missing data could make it problematic (since date isn't included). Clear Creek didn't have an overall change in trend (either increasing or decreasing) when we look at things on an annual basis, but there are significant changes in the trends when we look month by month.

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

    Yes, being able to know how to adjust gaps in data is a good skill to understand and practice.

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

    As usual, real-world data is a bit messier to deal with and we have to be mindful of how to adjust for that in our analyses.