Assignment 6: Time Series Analysis

Student Name

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.

```r
getwd()
```

```
## [1] "D:/William/Duke/Study/EOS 722/Hydrologic_Data_Analysis/Assignments/Assignment 6"
```

```r
packages <- c("tidyverse", "lubridate", "trend", "dataRetrieval")
invisible(lapply(packages, library, character.only = T))
```

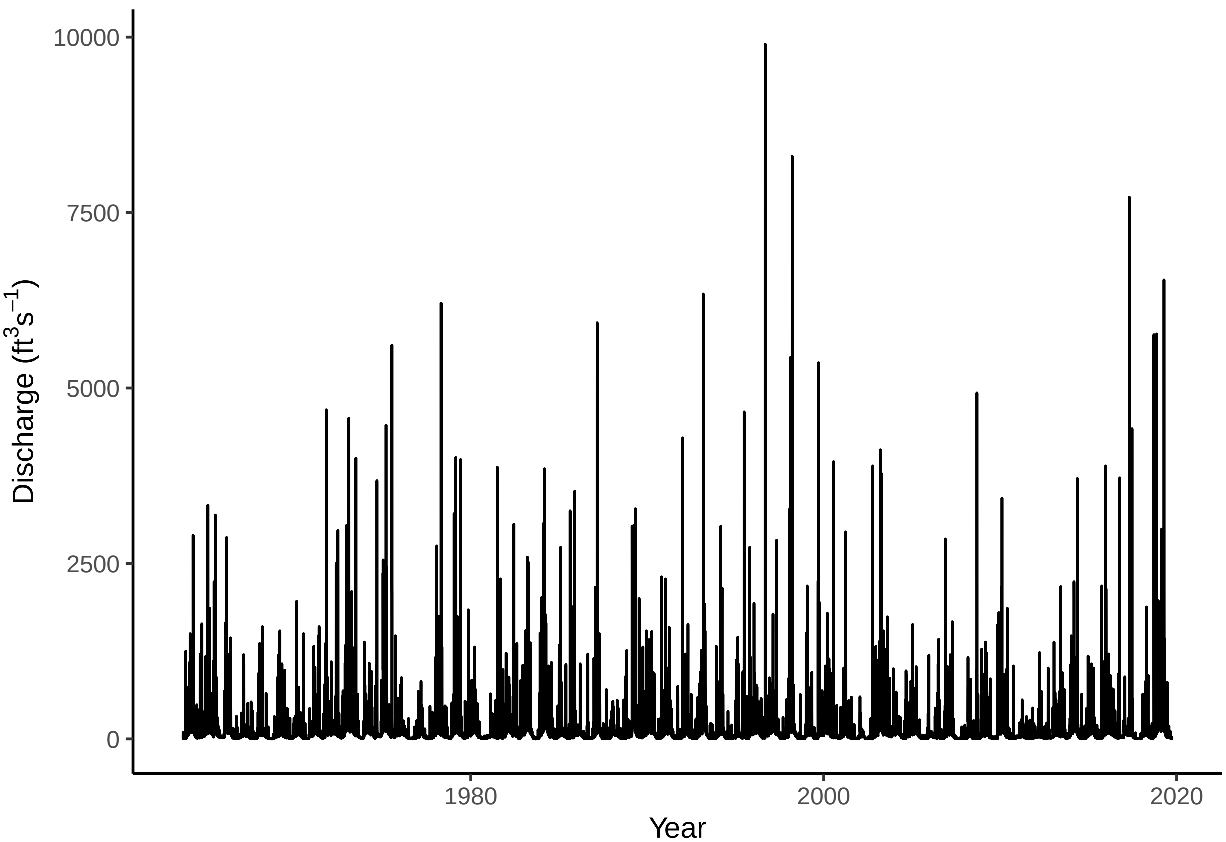
```
## -- Attaching packages --------------------------------------------------- tidyverse 1.2.1 --

## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   1.0.0     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()

##
## Attaching package: 'lubridate'

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

```r
theme_set(theme_classic())

ClearCreekDischarge.Monthly <-
  read.csv("../../Data/Processed/ClearCreekDischarge.Monthly.csv", header = T)
```

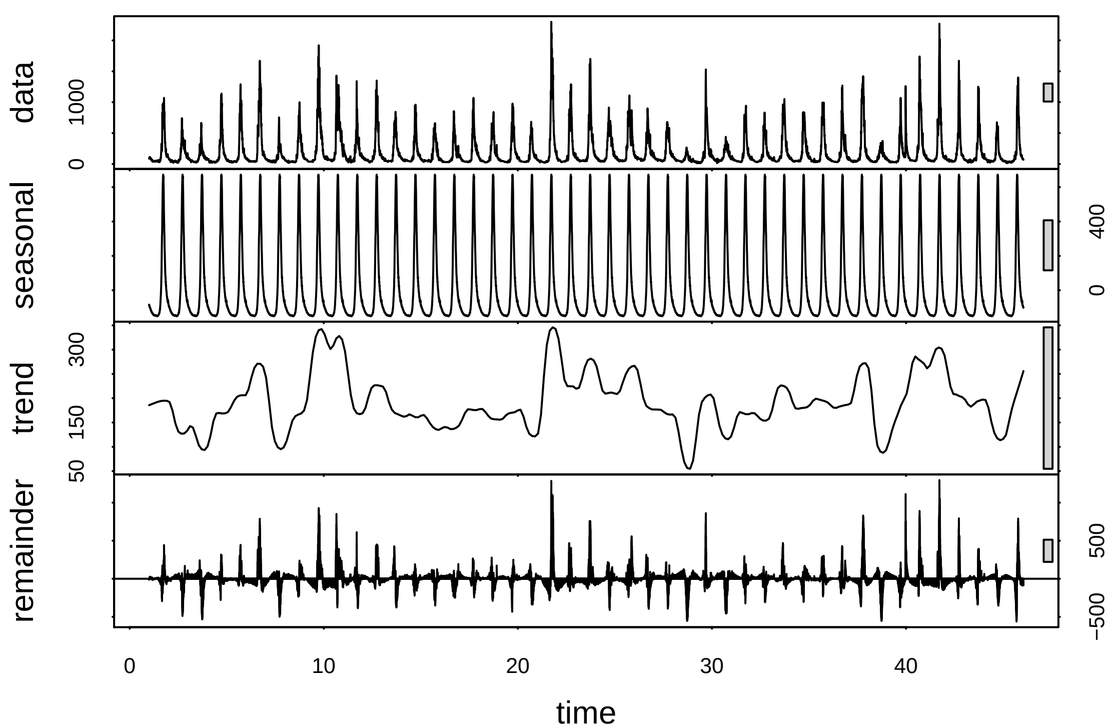
Time Series Decomposition

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.

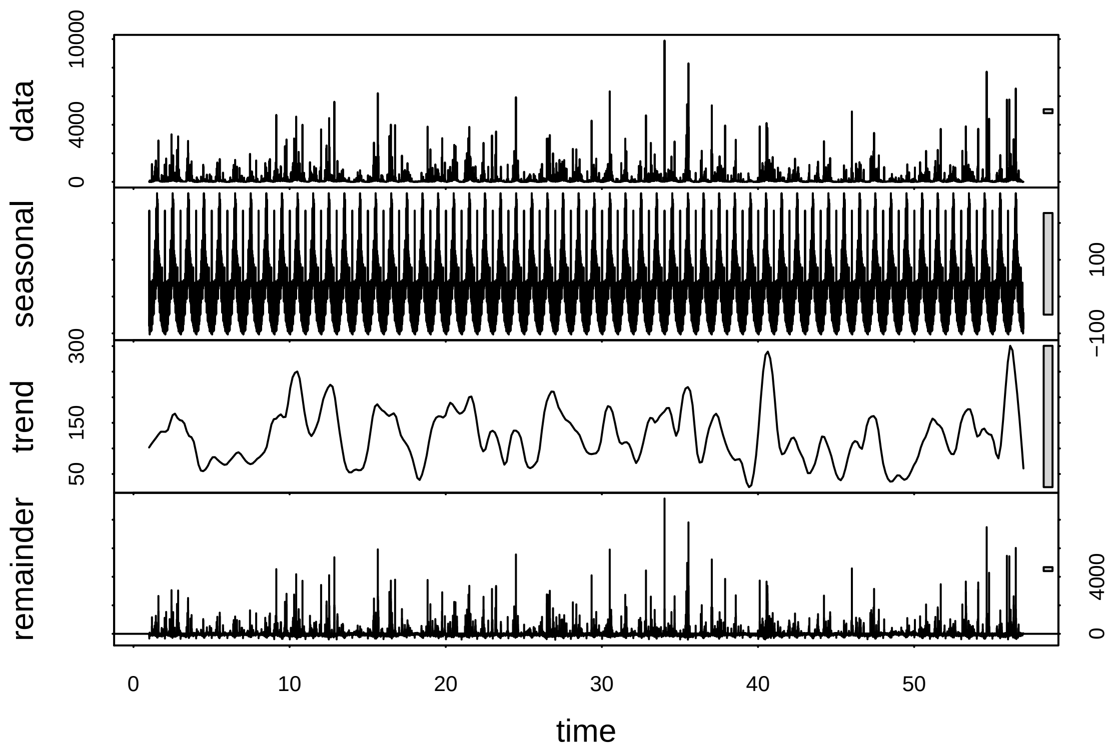
```
# Retrieve Daily Value data
# whatNWISdata(siteNumbers = "02085070")
EnoDischarge <- readNWISdv(siteNumbers = "02085070", parameterCd = "00060",
                           # discharge (ft3/s)
                           startDate = "", endDate = "")
names(EnoDischarge)[4:5] <- c("Discharge", "Approval.Code")

# Remove Feb 29s in leap years to make frequency of ts constant
allyears <- unique(year(EnoDischarge$Date))
Feb29s <- paste(allyears[leap_year(allyears)], "02", "29", sep = "-")
EnoDischarge <-  filter(EnoDischarge, !(Date %in% as.Date(Feb29s)))

# Plot
ggplot(EnoDischarge, aes(x = Date, y = Discharge))+
  geom_line()+
  labs(x = "Year", y = expression("Discharge (ft"^3*"s"^-1*")"))
```



```r
# Create a time series for Eno River
Eno_ts <- ts(EnoDischarge$Discharge, frequency = 365)

# time series for Clear Creek from L11
ClearCreekDischarge <- readNWISdv(siteNumbers = "06719505",
                        parameterCd = "00060", # discharge (ft3/s)
                        startDate = "",
                        endDate = "")
names(ClearCreekDischarge)[4:5] <- c("Discharge", "Approval.Code")
ClearCreek_ts <- ts(ClearCreekDischarge[[4]], frequency = 365)
ClearCreek_Decomposed <- stl(ClearCreek_ts, s.window = "periodic")

#Decomposition
Eno_Decomposed <- stl(Eno_ts, s.window = "periodic")
plot(ClearCreek_Decomposed)
```




```r
plot(Eno_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: Clear Creek has much higher seasonality than Eno River, as the magnitude of seasonal variation in Clear Creek is about 3 times as large as the bar shown on the right, while that in Eno River is about the same as the bar. Also, the range of y-axis can be used to compare seasonalities.

    Trend: The magnitudes of trends in two rivers are about the same, although they exhibit different trend patterns.

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.

```
# Generate time series
ClearCreekMonthly_ts <- ts(ClearCreekDischarge.Monthly$Discharge, frequency = 12,
                           start = c(1974, 10), end = c(2019, 10))

# Run SMK test
ClearCreekMonthly_test <- smk.test(ClearCreekMonthly_ts)
ClearCreekMonthly_test
```

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
```

```
##
## data:  ClearCreekMonthly_ts
## z = 1.6586, p-value = 0.09719
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##      S   varS
##    590 126102
```

```
summary(ClearCreekMonthly_test)
```

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: ClearCreekMonthly_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
```

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

    Overall the p-value is 0.09719, so we cannot reject the null hypothesis that there is no trend in discharge over time.

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?

    Returned values of summary() function indiate that although there is no overall trend, December and January(season 3 & 4) have positive trend over time, shown by the positive S values, that is statistically significant.

Reflection

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

    Variations in hydrological parameters that tend to have seasonality may be attributed to multiple components: overall trend, seasonality, and random error/residual. Due to the presence of seasonality, linear regression may not be appropriate for detecting overall trend, as it can be masked by seasonality,

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

The contribution of each component can be visualized by the summaizing plot of a decomposed time series. The inadequacy of linear regression in time series can be seen in the graph with `geom_smooth(method = "lm")` plotted in class.

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

    It may not be perplexing to understand changes over time, but actually constructing and analyzing time series model turns out to be relatively complex. It's good to have some practical experience on performing the analysis.

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

    It is as expected that certain hydrological parameters such as discharge could exhibit cycling patterns that are caused by seasons. Also, there could be overall trend due to large-scale shifts, e.g. climate change.