Assignment 6: Time Series Analysis

Jake Greif

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()
```

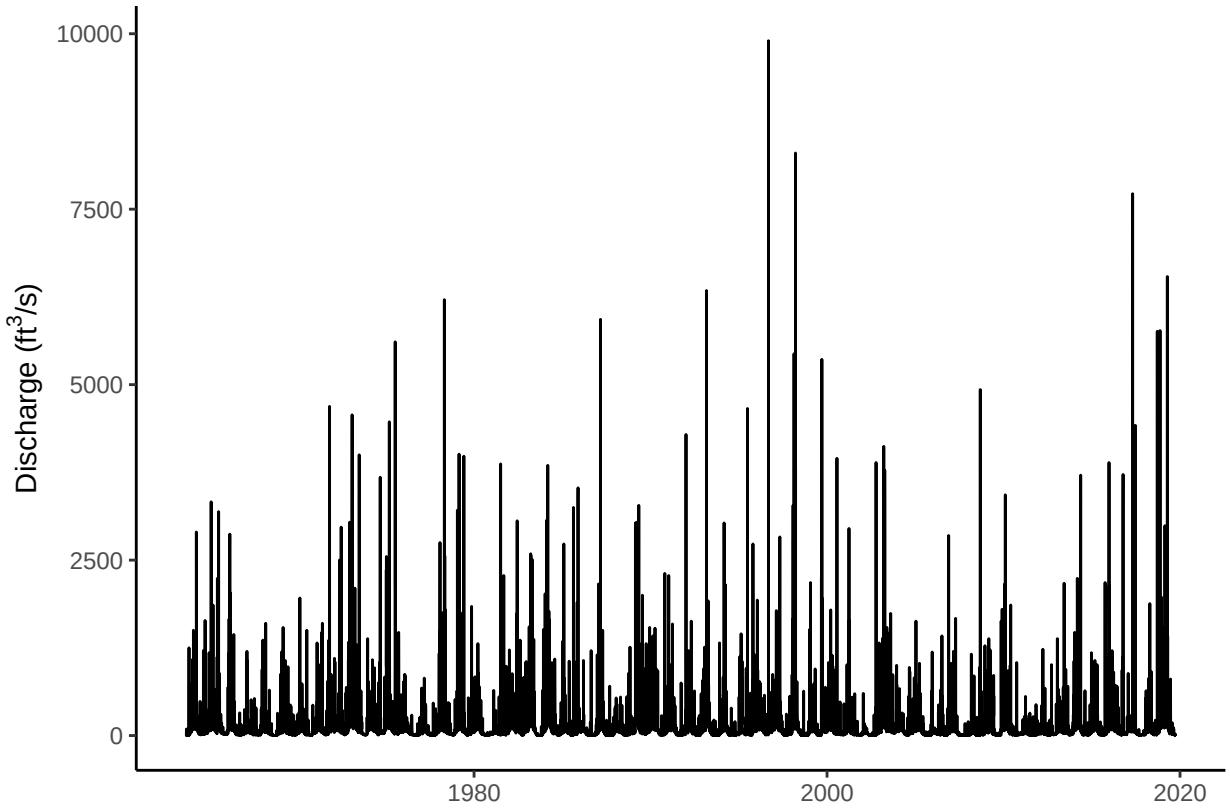
```
## [1] "/Users/jakegreif/Duke/Fall_2019/Hydrologic_Data_Analysis"
```

```
library(tidyverse)
library(lubridate)
library(dataRetrieval)
library(trend)

theme_set(theme_classic())

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

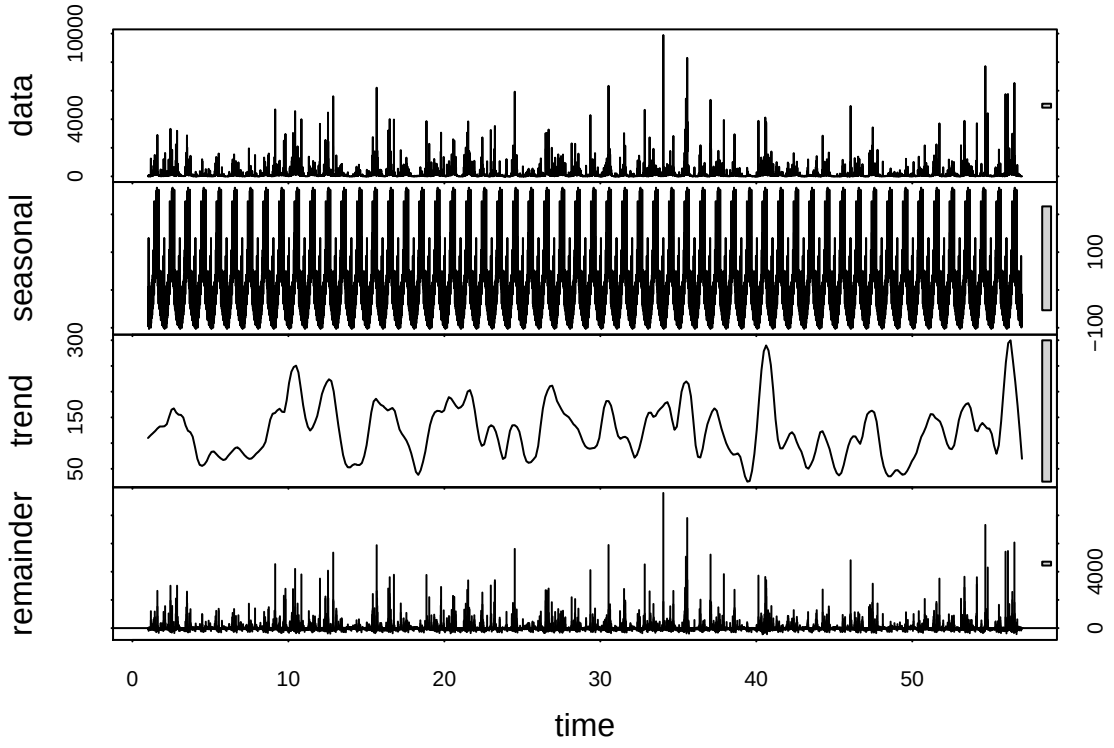
Time Series Decomposition

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
# Import data
EnoDischarge <- readNWISdv(siteNumbers = "02085070",
                     parameterCd = "00060",
                     startDate = "",
                     endDate = "")
names(EnoDischarge)[4:5] <- c("Discharge", "Approval.Code")

# Plot discharge over time
EnoPlot <-
  ggplot(EnoDischarge, aes(x = Date, y = Discharge)) +
  geom_line() +
  labs(x = "", y = expression("Discharge (ft"^3*"/s)"))
print(EnoPlot)
```




```r
# Creat time series
Eno_ts <- ts(EnoDischarge[[4]], frequency = 365)

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

# Visualize the decomposed series.
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: The seasonal component of the Eno is noisier and lesser in magnitude compared to Clear Creek.

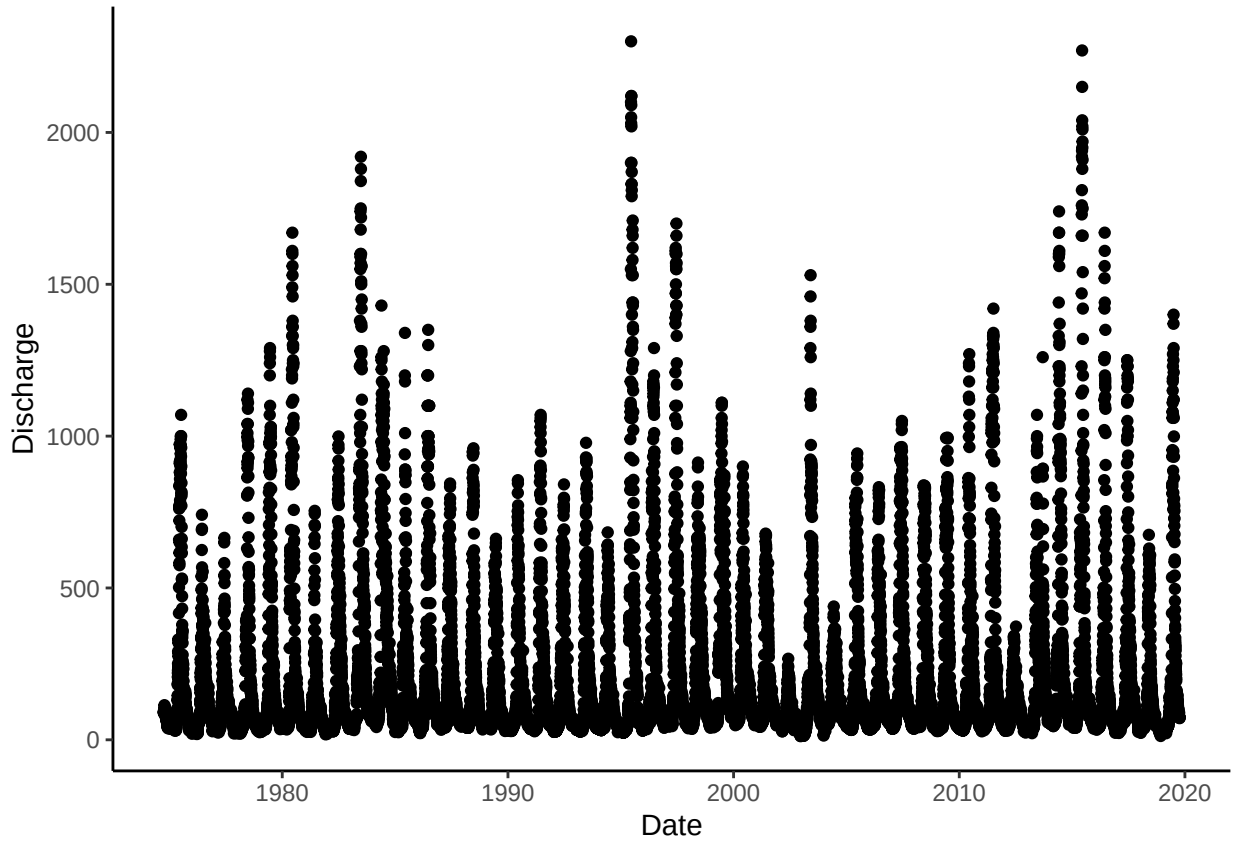
    Trend: The trend component of the Eno is lower in magnitude compared to Clear Creek, but Clear Creek trend seems to be more smooth (less intense changes) than the Eno.

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.

```
# Plot, observe data
ClearCreekDischarge <- readNWISdv(siteNumbers = "06719505",
                     parameterCd = "00060", # discharge (ft3/s)
                     startDate = "",
                     endDate = "")
names(ClearCreekDischarge)[4:5] <- c("Discharge", "Approval.Code")

ggplot(ClearCreekDischarge, aes(x = Date, y = Discharge)) +
  geom_point()
```




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

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

# Inspect results
CCtrend
```

```
##
##  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(CCtrend)
```

```
##
##  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?

    There is not a significant overall monotonic 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?

    There are posotive monotonic trends in discharge in the months of June and July.

Reflection

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

    Time series contain a lot of useful information, but they require high quality data. It is possible to run trend analyses in the abscence of high quality data, but it's important to carefully choose the proper interpolation techniques and trend analyses because certain types of statistical tests are only appropriate for specific types of data.

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

    The assumptions for each of the monotonic trend tests are all slightly different and require careful consideration of the data before using them. The 'challenges' of time series analysis layed out at the beginning of lesson 11 also highlight the difficulty of analyzing time series.

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

    Yes, because the hands-on data analysis allowed me to see how the statistical tests can be applied. Without the hands-on experience, I would have a much more difficult time understanding how to apply the statistical tests to data in the future.

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

    I assumed that there would be more significant trends in discharge over time or from month to month based on what I've learned about the impacts of climate change on the hydrologic cycle.