Assignment 6: Time Series Analysis

Caroline Watson

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
#verifying working directory
getwd()
```

```
## [1] "/Users/carolinewatson/Documents/Fall 2019/Hydrologic_Data_Analysis/Assignments"
```

```r
#loading packages
suppressMessages(library(tidyverse))
library(lubridate)
```

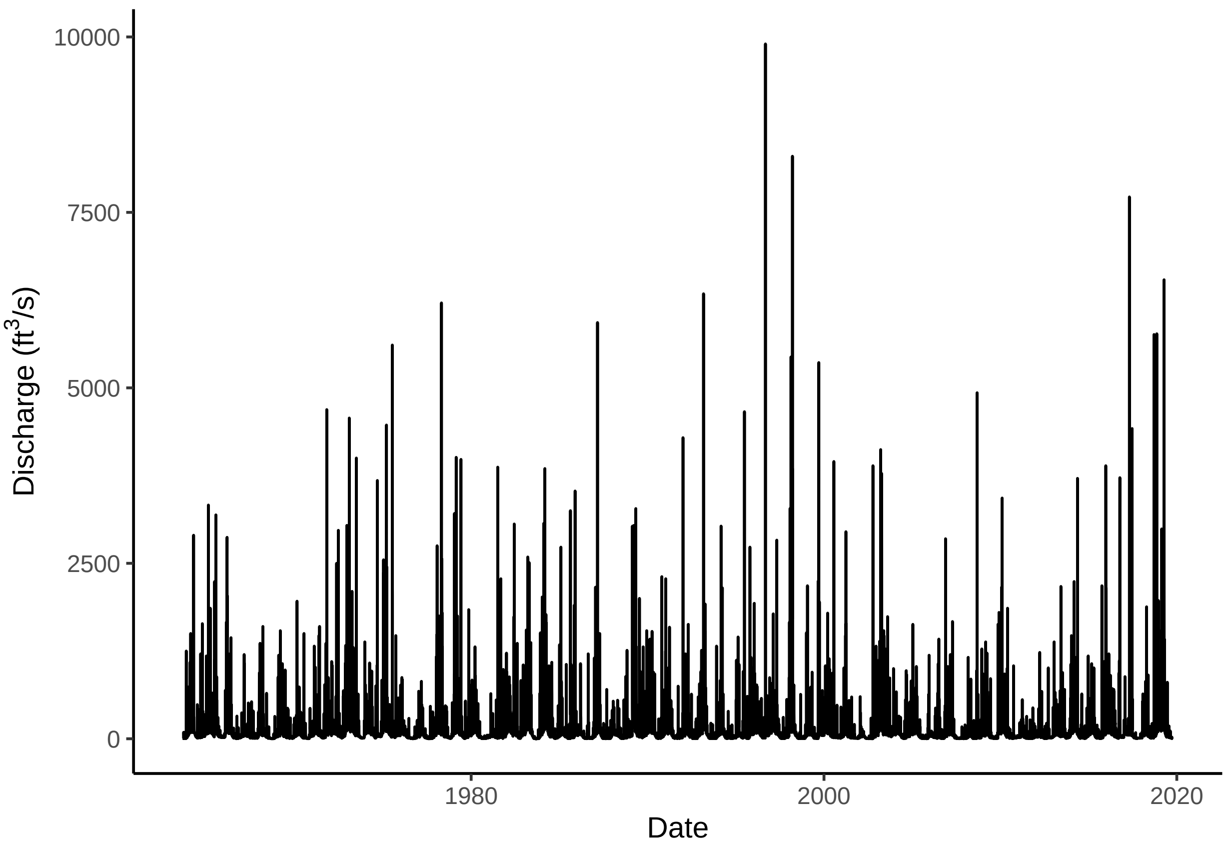
```
##
## Attaching package: 'lubridate'

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

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

#setting ggplot theme
theme_set(theme_classic())

#reading in discharge data for Clear Creek
ClearCreekDischarge.Monthly <- read.csv("../Data/Processed/ClearCreekDischarge.Monthly.csv")
```

Time Series Decomposition

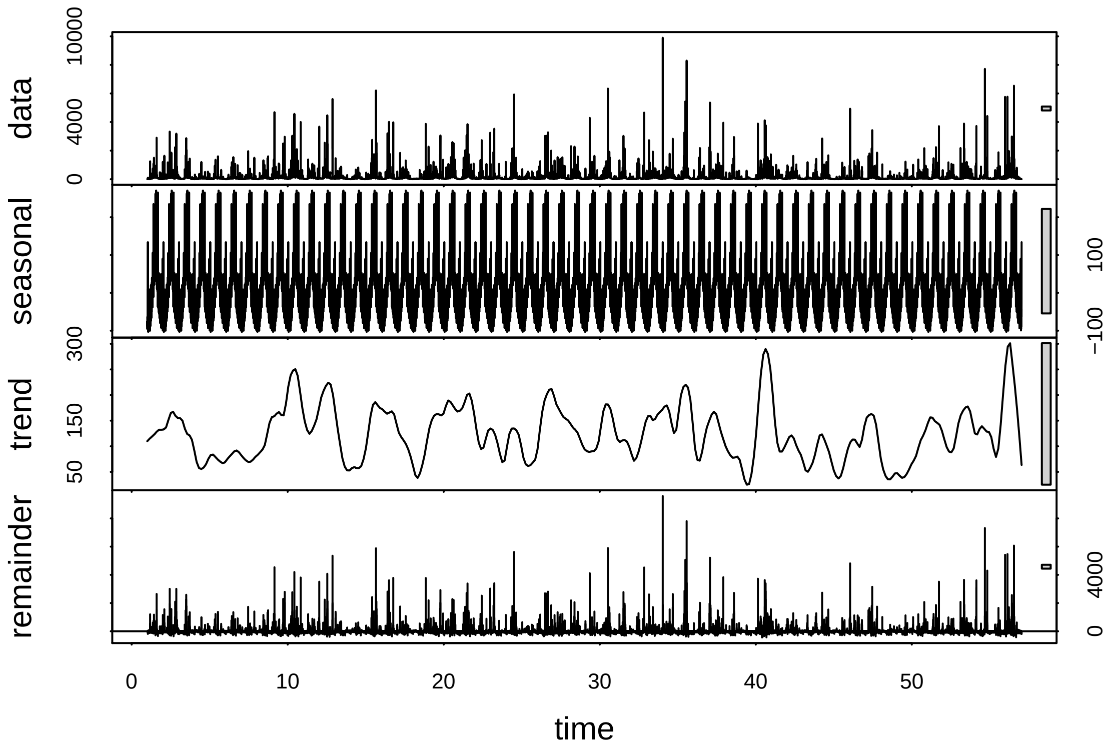
5. Create a new data frame that includes daily mean discharge at the Eno River for all available dates (`siteNumbers = "02085070"`). Rename the columns accordingly.
6. Plot discharge over time with geom_line. Make sure axis labels are formatted appropriately.
7. Create a time series of discharge
8. Decompose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

```
#new data frame with daily mean discharge at Eno River
EnoDischarge.Daily <- readNWISdv(siteNumbers = "02085070",
                     parameterCd = "00060", # discharge (ft3/s)
                     startDate = "",
                     endDate = "")
names(EnoDischarge.Daily)[4:5] <- c("Discharge", "Approval.Code")

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




```
#creating a timeseries of discharge
Eno_ts <- ts(EnoDischarge.Daily[[4]], frequency = 365)

#decomposing timeseries data using 'stl'
Eno_decomposed <- stl(Eno_ts, s.window = "periodic")
```

```r
#visualize decomposed timeseries
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 from Clear Creek discharge is more cyclical than the seasonal component from the Eno. It looks like the seasonal component for the Eno discharge has a lot of noise, or maybe it has larger discharge events and therefore there are more points. Also, the seasonal components from the Eno data peaks when the peaks in the data section peak. Also, the scale for Eno data is from -100 to 100, whereas the scale for Clear Creek seasonal component is from 0 to 400. Therefore, these datasets are not similar in magnitude, when just looking at the seasonal component.

    Trend: The trend components for Clear Creek and Eno River are pretty similar in terms of their patterns. Peaks in the trend components correspond to peaks in the data components. The trend components for Clear Creek have a scale from 50 - 350, where as the trend components for the Eno has a scale from 50 - 300. These scales are more simmilar than the scales for the seasonal components of both Clear Creek and 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.

```r
#time series of monthly discharge from Clear Creek
ClearCreek.timeseries <- ts(ClearCreekDischarge.Monthly$Discharge, frequency = 12,
```

```
                                  start = c(1974, 10), end = c(2019, 10))

#running a Seasonal Mann-Kendall test on monthly discharge data
ClearCreek.trend <- smk.test(ClearCreek.timeseries)

#inspecting results of SMK test
summary(ClearCreek.trend)
```

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: ClearCreek.timeseries
## 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?

    There is no overall significant monotonic trend (positive or negative) in discharge over time at Clear Creek from 1974 to 2019. However, there are monotonic trends in certain months.

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 monthly monotonic trends in discharge over time. These positive trends occur during March and April where the p-value is significant. The trend for March from 1974 to 2019 is positive since the z-score is positive. The trend for April from 1974 to 2019 is also positive and the z-score is also positive. The S value for both March and April are large which indicates the greater magnitude of the trend in March and April.

Reflection

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

    Seasonal component versus the trend component of discharge can be extracted using certain R code. Seasonal timeseries over the years can indicate whether there are monotonic trends in certain months even if there are not monotonic trends every month over the years.

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

For seasonal component, the plot showing the decomposed time series helps to visualize the different components of the time series trend. For the seasonal timeseries over the years, the summary of the Seasonal Mann-Kendall test showed where there was a monotonic trend and in which months.

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

    Hands-on data analysis helped in learning about time series because I was able to better understand time series by applying the R code to the new dataset.

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

    Real-world data was pretty consistant with my expectations from theory.