Assignment 6: Time Series Analysis

Yixin Wen

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/yixinwen/Box/Duke/2019 Fall/Hydrologic Data Analysis/Hydrologic_Data_Analysis/Assignments
```

```
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
```

```
library(trend)
library(dataRetrieval)
library(forecast)
```

```
## 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
```

```
## Registered S3 methods overwritten by 'forecast':
##   method             from
##   fitted.fracdiff    fracdiff
##   residuals.fracdiff fracdiff
```

```
library(tseries)

theme_set(theme_classic())

library(readr)
ClearCreekDischarge.Monthly <- read_csv("/Users/yixinwen/Box/Duke/2019 Fall/Hydrologic Data Analysis/Hy
```

```
## Parsed with column specification:
## cols(
##   Year = col_double(),
##   Month = col_double(),
##   Discharge = col_double()
## )
```

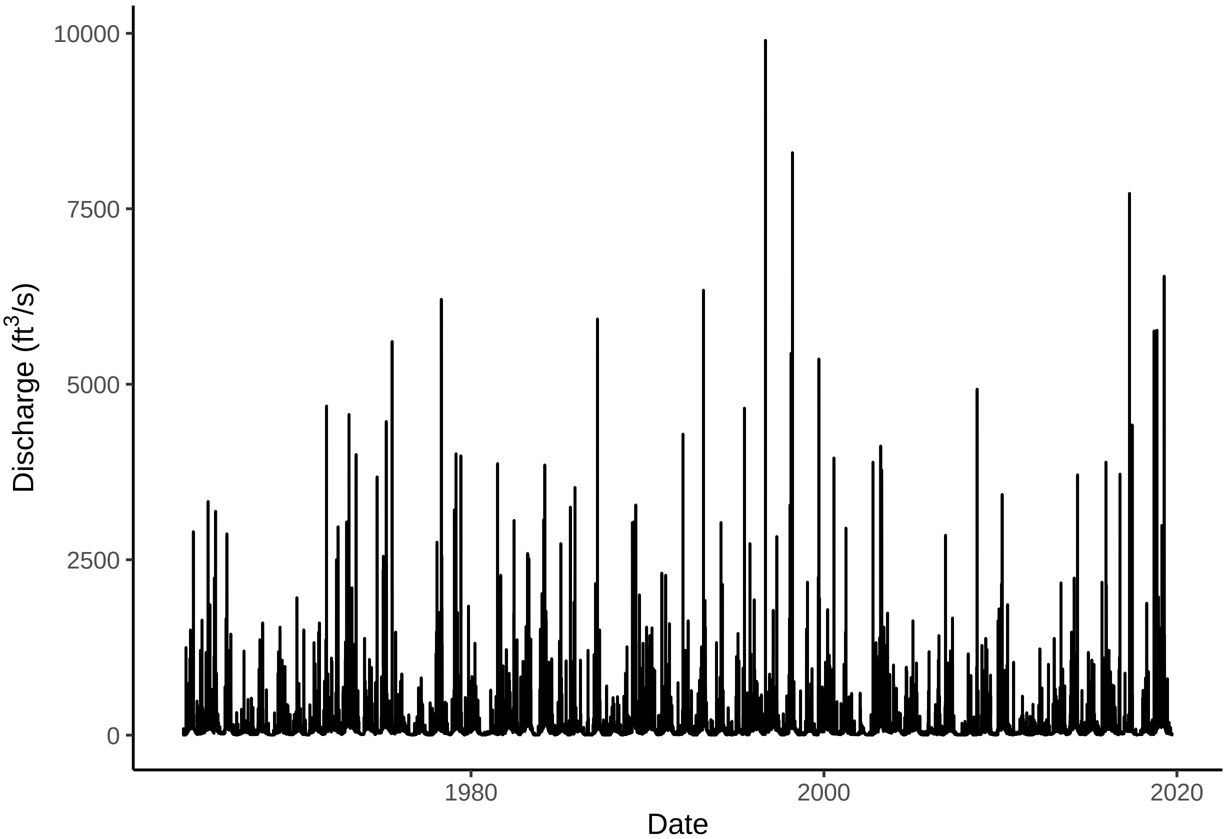
Time Series Decomposition

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.

```
# create a data frame of Eno River
EnoRiverDischarge <- readNWISdv(siteNumbers = "02085070",
                                parameterCd = "00060",
                                startDate = "",
                                endDate = "")
names(EnoRiverDischarge)[4:5] <-c("Discharge","Approval.Code")
class(EnoRiverDischarge$Date)
```

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

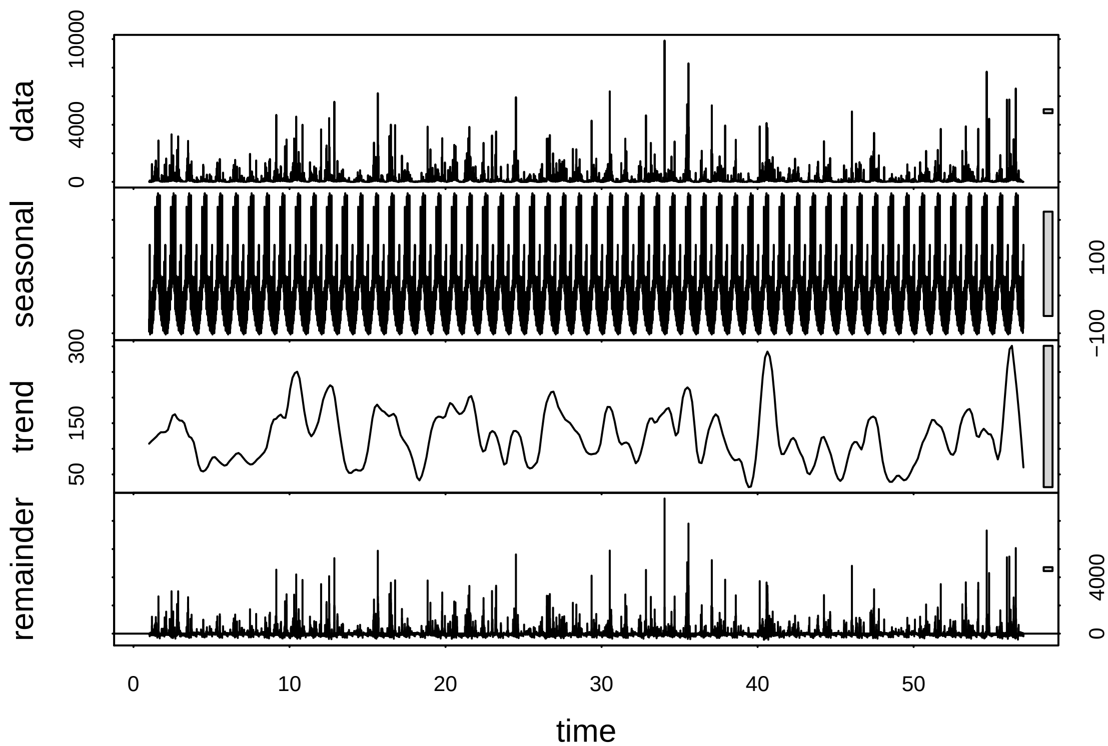
```
# plot discarge over time
EnoRiverDischarge.plot <-
  ggplot(data = EnoRiverDischarge, aes(x = Date, y = Discharge))+
  geom_line()+
  labs(x = "Date", y = expression("Discharge (ft"^3*"/s)"))
print(EnoRiverDischarge.plot)
```




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

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

# visualize the decompozation
plot(EnoRiver_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 range of seasonal change of Eno River discharge is smaller than that of Clear Creek discharge. The Eno River has a more compased seasonal change compared to Clear Creek.

    Trend: The range of trend component of Eno River discharge is similar to that of Clear Creek discharge.

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)

ClearCreekTrend <- smk.test(ClearCreek_ts)

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
```

```r
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?

    since the p-value = 0.09719, which is larger than 0.05, there's no 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?

    June and July, since the z scores are larger than 0, they are positive.

Reflection

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

    1. when p-value is larger than 0.05, there's no significant trend on change; when p-value is smaller than 0.05, there is monotonic trend.
    2. when z score is greater than 0, the trend is positive, and when z score is smaller than 0, the trend is negative.

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

    the Seasonal Mann-Kendall test on ClearCreek can give the results of whether there is monotonic trend on discharge.

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 can let me explore the theory on my own, and it can help me understand it better from examples.

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

    The real-world data may have mistakes in the dataset. Before we do analysis on it, we need to remove the abnormal data, otherwise the results may not be correct.