MA585-HW1

Ranfei Xu

2022/1/31

1.

Based on Google Trends (https://www.google.com/trends/), think of three search terms each with the following characteristics:

```
ted <- read.csv("E:/MA585/HW585/TED_multiTimeline.csv", header = TRUE, skip = 2)
ski <- read.csv("E:/MA585/HW585/ski_multiTimeline.csv", header = TRUE)
nfl <- read.csv("E:/MA585/HW585/netflix_multiTimeline.csv", header = TRUE, skip = 2)
```

(1) a time series with a clear trend component

The plot below shows the frequency of search term **TED**. We can see that there is a clear downward trend, which indicates that with the prosperity of video platforms, the popularity of ted has gradually declined.

```
# head(ted)
plot.ts(ted$TED...United.States., ylab="Frequency",xlab="Week")
title("Daily Google Search for TED in the Past 5 Years")
```
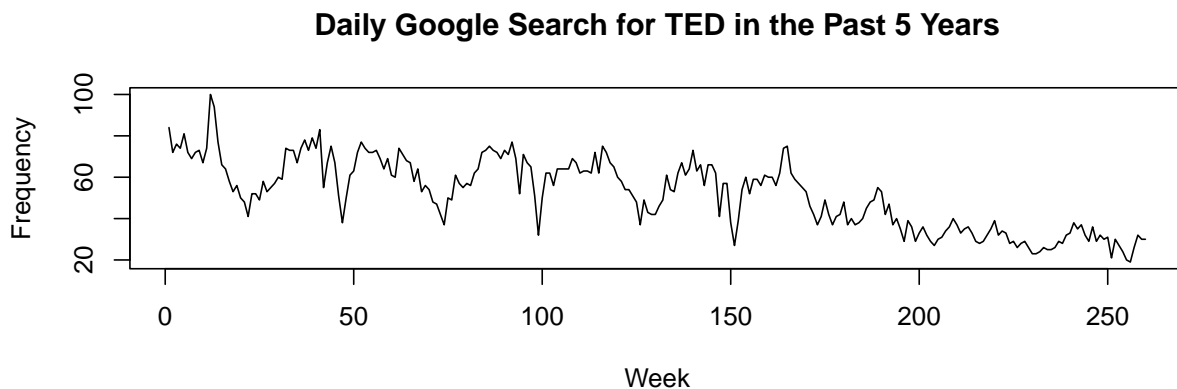

(2) a time series with a clear seasonal component.

The plot below shows the frequency of search term **ski**. We can see that there is a clear seasonal component, which indicates Skiing is only widely participated in winter.

```
# head(ski)
plot.ts(ski$ski...United.States., ylab="Frequency",xlab="Week")
title("Daily Google Search for ski in the Past 5 Years")
```

Daily Google Search for ski in the Past 5 Years
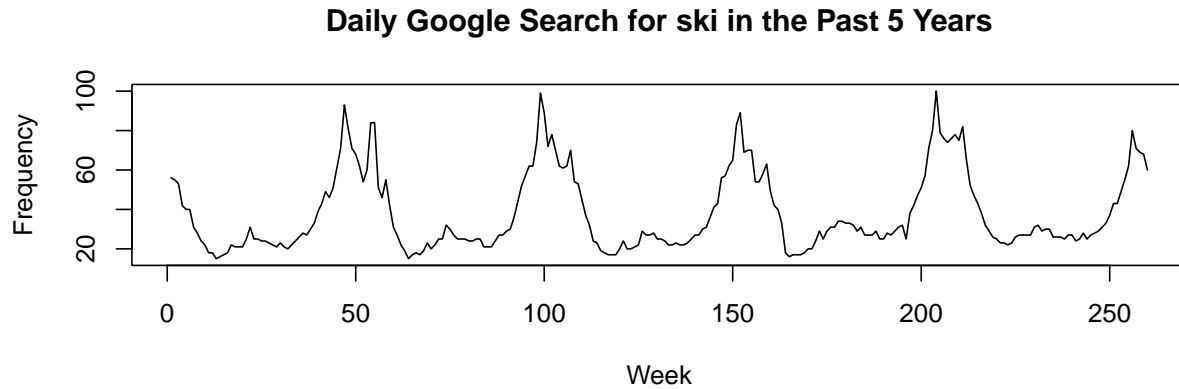
(3) a time series with no clear trend or seasonal pattern and is poorly described by a combination of trend and seasonal components.

The plot below shows the frequency of search term **Netflix**. We can see that there is no clear trend or seasonal pattern.

```
# head(nfl)
plot.ts(nfl$Netflix...United.States., ylab="Frequency",xlab="Day")
title("Daily Google Search for netflix in the Past 90 Days")
```

Daily Google Search for netflix in the Past 90 Days
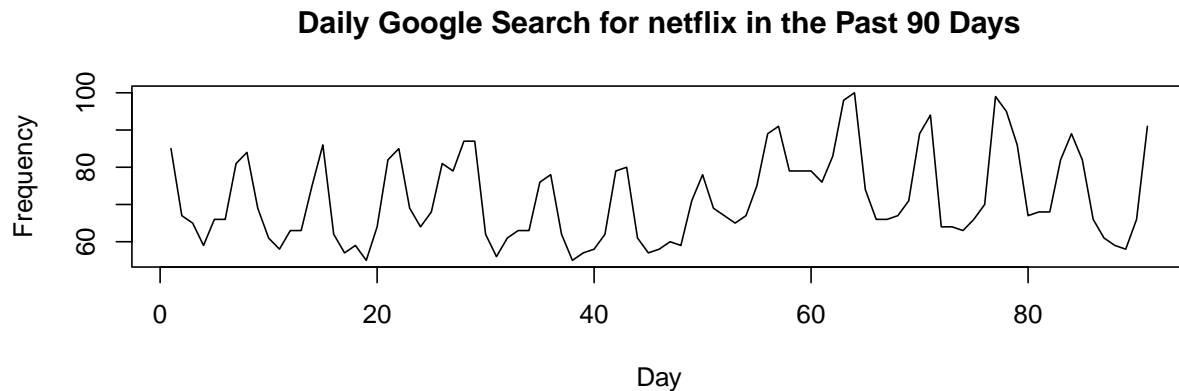
2.

Col.	Description
TPCP	Total precipitation amount for the month (tens of mm)
TSNW	Total snowfall amount for the month (mm)
MMXT	Monthly mean maximum temperature (tenths of degrees of Celsius)
MMNT	Monthly mean temperature (tenths of degrees of Celsius)

(1) Plot the time series for each variable and write a short description of the key features of each series.

```
BosClimateData <- read.csv("E:/MA585/HW585/BosClimateData.csv",
header = TRUE)

BosData <- subset(BosClimateData,BosClimateData$STATION_NAME=="BOSTON LOGAN INTERNATIONAL AIRPORT MA US

plot.ts(BosData$TPCP, xlab = "Month")
title("Total precipitation amount for the month (tens of mm)")
```
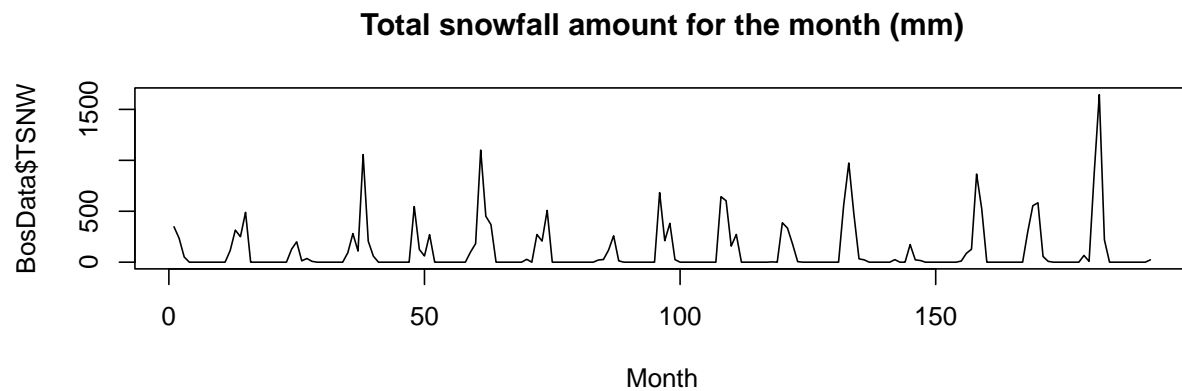
Total precipitation amount for the month (tens of mm)
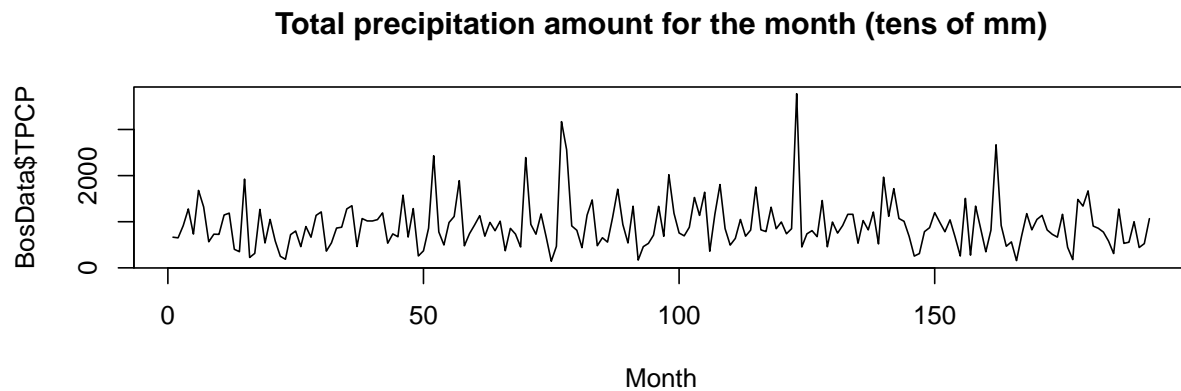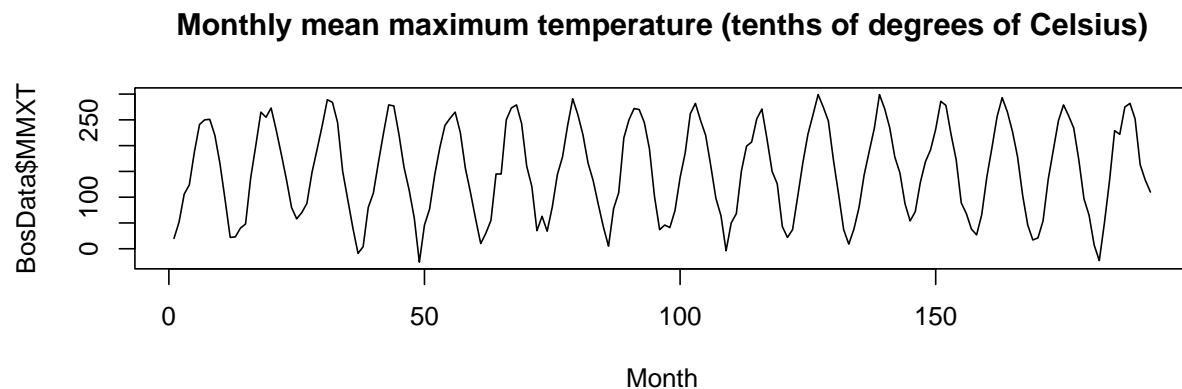

Features: The plot above describes the total precipitation amount for each month during 2020.01 to 2015.12. However there is no clear trend or seasonal pattern.

```
plot.ts(BosData$TSNW, xlab = "Month")
title("Total snowfall amount for the month (mm)")
```
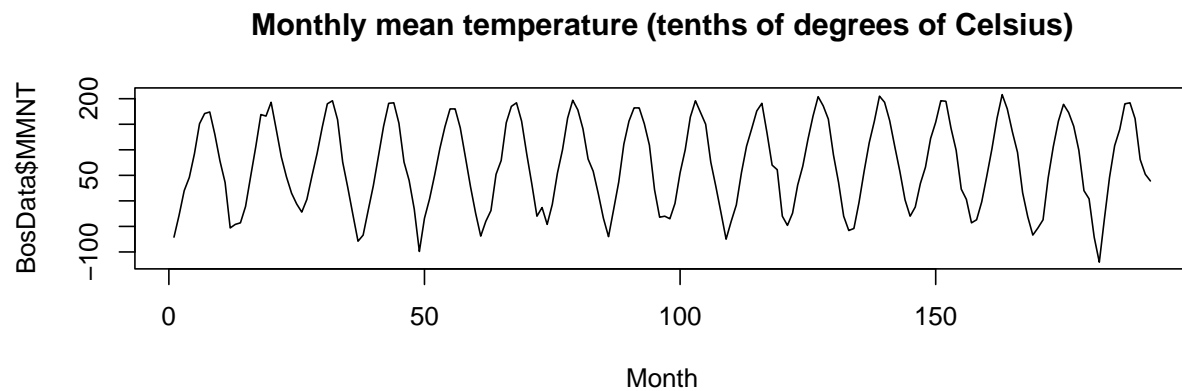
Total snowfall amount for the month (mm)

Features: The plot above describes the total snowfall amount for each month during 2020.01 to 2015.12. We can see a clear seasonal component that indicates Snowfall is only concentrated at certain times of the year.

```
plot.ts(BosData$MMXT, xlab = "Month")
title("Monthly mean maximum temperature (tenths of degrees of Celsius)")
```

Monthly mean maximum temperature (tenths of degrees of Celsius)

Features: The plot above describes the monthly mean maximum temperature during 2020.01 to 2015.12. We can see a clear seasonal component with continuity, which indicates the monthly mean maximum temperature in Boston changes stable, continuous and noticeable.

```
plot.ts(BosData$MMNT, xlab = "Month")
title("Monthly mean temperature (tenths of degrees of Celsius)")
```

Monthly mean temperature (tenths of degrees of Celsius)

Features: The plot above describes the monthly mean temperature during 2020.01 to 2015.12. We can see a clear seasonal component with continuity, just as the Monthly mean maximum temperature plot.

(2) What proportion of months had total snowfall greater than a foot?

```
# 1 foot = 304.8 mm
TSNW   <- BosData$TSNW
l.TSNW <- length(TSNW)
s.TSNW <- length(subset(BosData$TSNW, BosData$TSNW >= 304.8))
(s.TSNW/l.TSNW)
```

```
## [1] 0.125
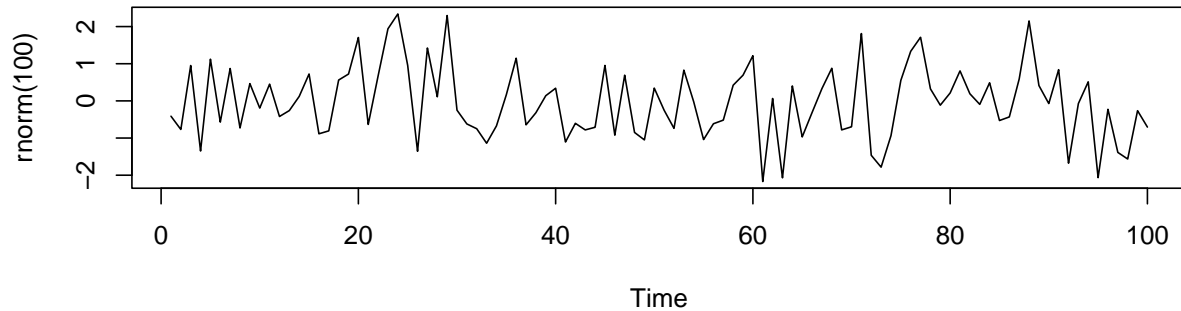```

(3) What proportion of months mean maximum temperature exceeded 80F?

```
80F = 26.67 Celsius,
MMXT   <- BosData$MMXT
l.MMXT <- length(MMXT)
s.MMXT <- length(subset(BosData$MMXT, BosData$MMXT >= 266.7))
(s.MMXT/l.MMXT)
```
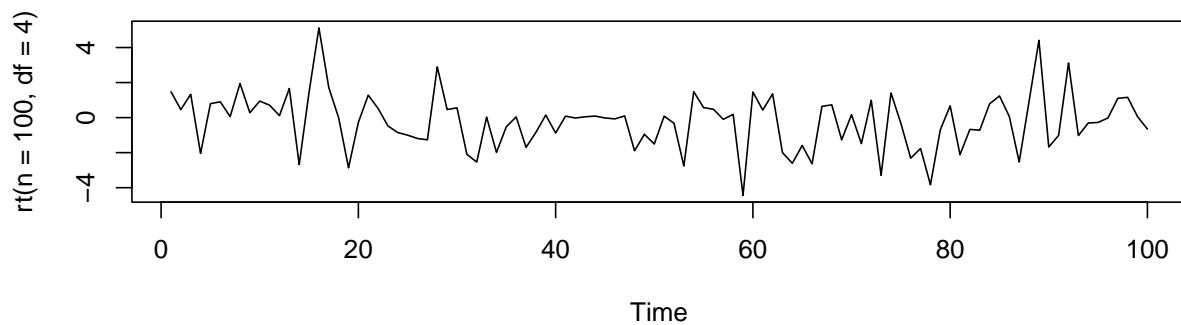
```
[1] 0.1145833
```

# 3. Simulation

(1)

```
plot.ts(rnorm(100))
```


Answer: After repeat the simulation several times, there are no obvious patterns in the plots, so I think it looks random.

(2)

```
plot.ts(rt(n=100, df=4))
```


Answer: The amplitude is small, which means the variance is relatively small at most of the "Time", thus might indicates the data generating process may not be Gaussian.