Time Series Graphics

Shovan Biswas

2020/9/05

Libraries

```
library(knitr)
library(kableExtra)
library(tidyverse)
library(fpp2)
```

Time Series Graphics

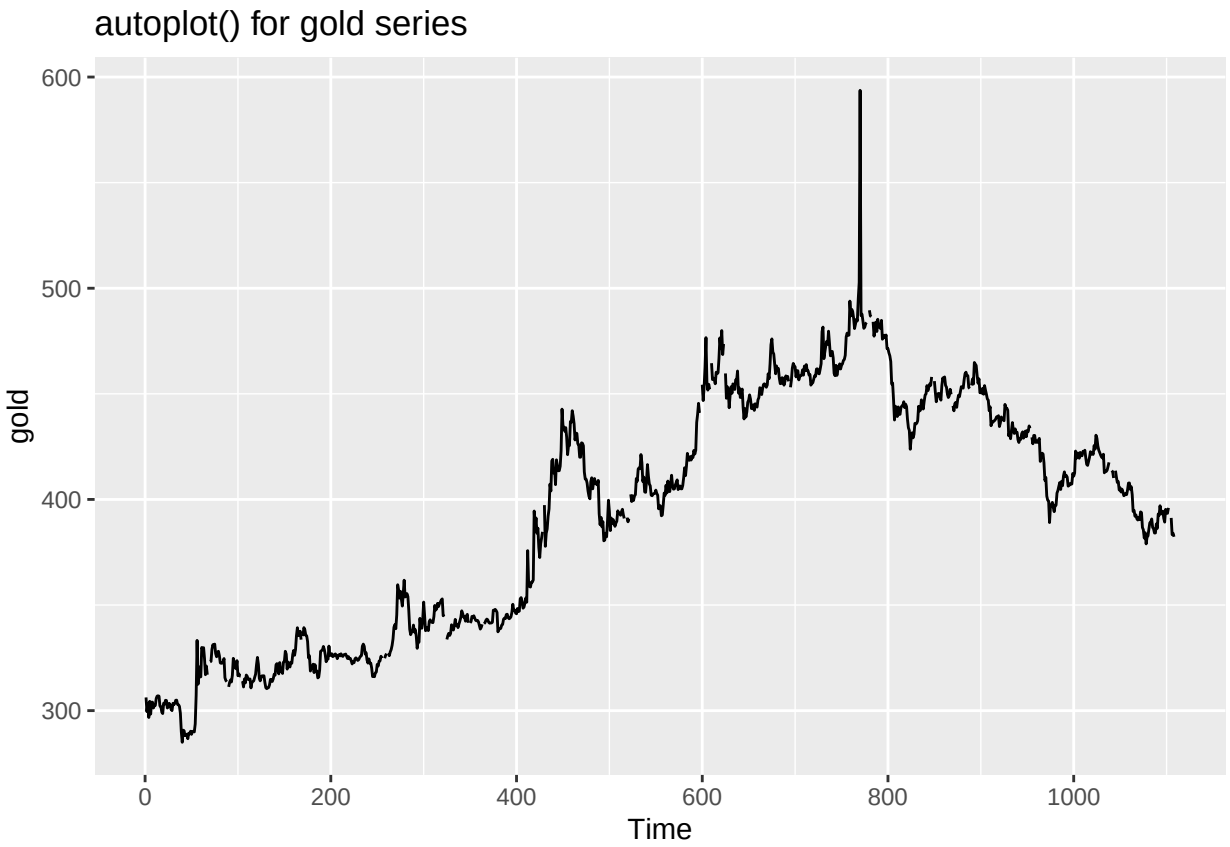
1) Use the help function to explore what the series gold, woolyrnq and gas represent.
    a. Use autoplot() to plot each of these in separate plots.
    b. What is the frequency of each series? Hint: apply the frequency() function.
    c. Use which.max() to spot the outlier in the gold series. Which observation was it?

I explored the help() with gold, woolyrnq and gas. They flashed 3 websites, but I commented them for now. Please uncomment and run.

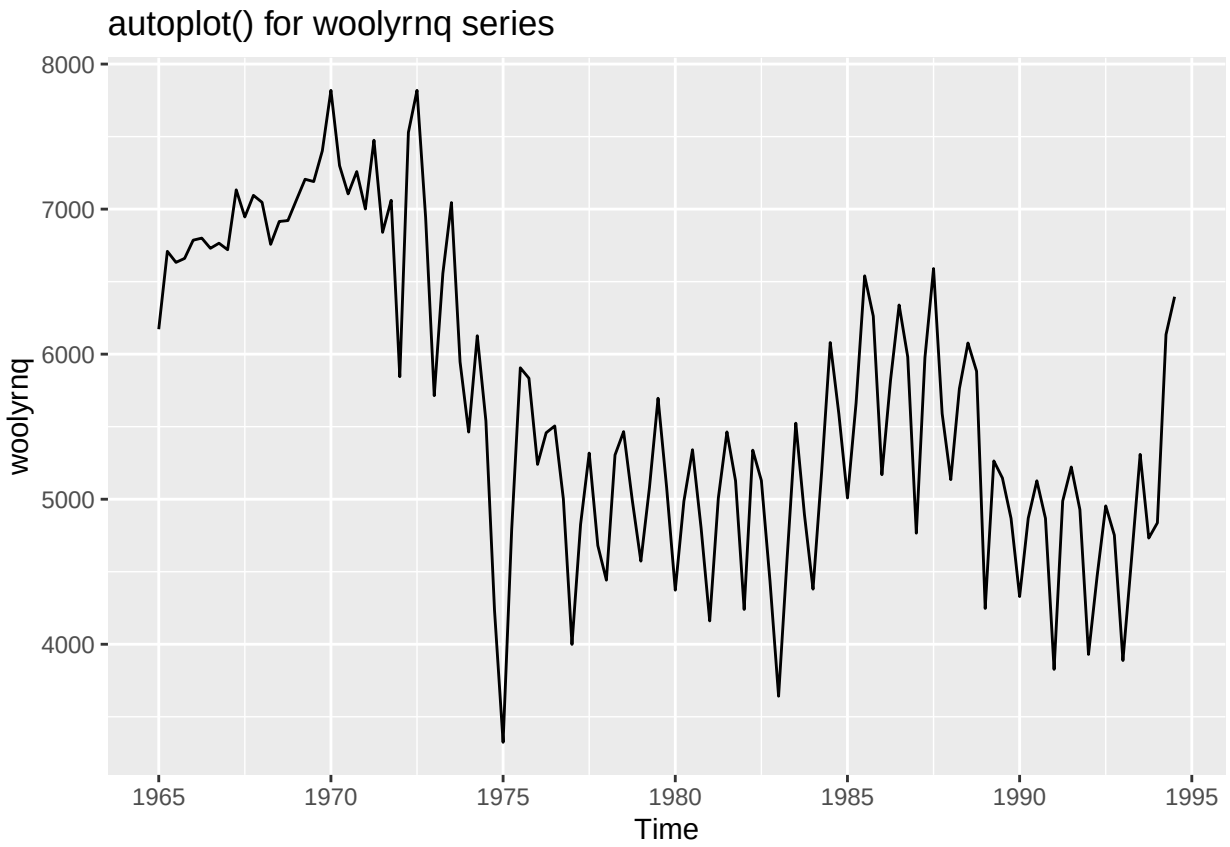
```
# help(gold)
# help(woolyrnq)
# help(gas)
```

1a) Use autoplot() to plot each of these in separate plots.

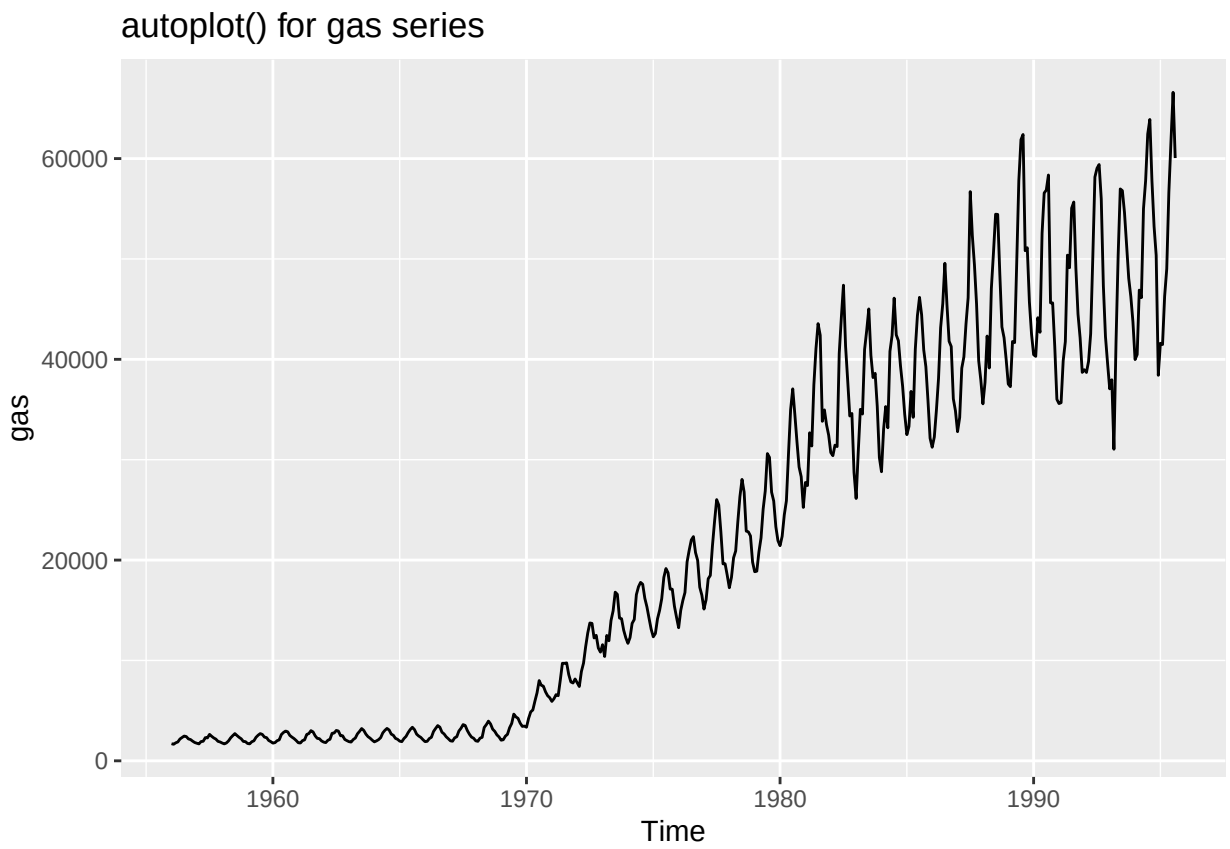
```
autoplot(gold) + ggtitle("autoplot() for gold series")
```




```
autoplot(woolyrnq) + ggtitle("autoplot() for woolyrnq series")
```




```
autoplot(gas) + ggtitle("autoplot() for gas series")
```



1b) What is the frequency of each series? Hint: apply the frequency() function.

```
frequency(gold)
```

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

```
frequency(woolyrnq)
```

```
## [1] 4
```

```
frequency(gas)
```

```
## [1] 12
```

So, gold is annual, woolyrnq is quarterly and gas is monthly.

1c) Use which.max() to spot the outlier in the gold series. Which observation was it?

```
cat('Outlier price is ', gold[which.max(gold)], 'at position ', which.max(gold))
```

```
## Outlier price is  593.7 at position  770
```

2) Download the file tute1.csv from the book website, open it in Excel (or some other spreadsheet application), and review its contents. You should find four columns of information. Columns B through D each contain a quarterly series, labelled Sales, AdBudget and GDP. Sales contains the quarterly sales for a small company over the period 1981-2005. AdBudget is the advertising budget and GDP is the gross domestic product. All series have been adjusted for inflation.

   a. You can read the data into R with the following script:

```
tute1 <- read.csv("./tute1.csv", header = TRUE)
View(tute1)
head(tute1)
```

```
##        X  Sales AdBudget   GDP
## 1 Mar-81 1020.2    659.2 251.8
## 2 Jun-81  889.2    589.0 290.9
## 3 Sep-81  795.0    512.5 290.8
## 4 Dec-81 1003.9    614.1 292.4
## 5 Mar-82 1057.7    647.2 279.1
## 6 Jun-82  944.4    602.0 254.0
```

The code chunk in the book instructs to execute View(). But when the RMD is executed, the View appears for a quick second and disappears. So, I put head() to show the head of tute1.

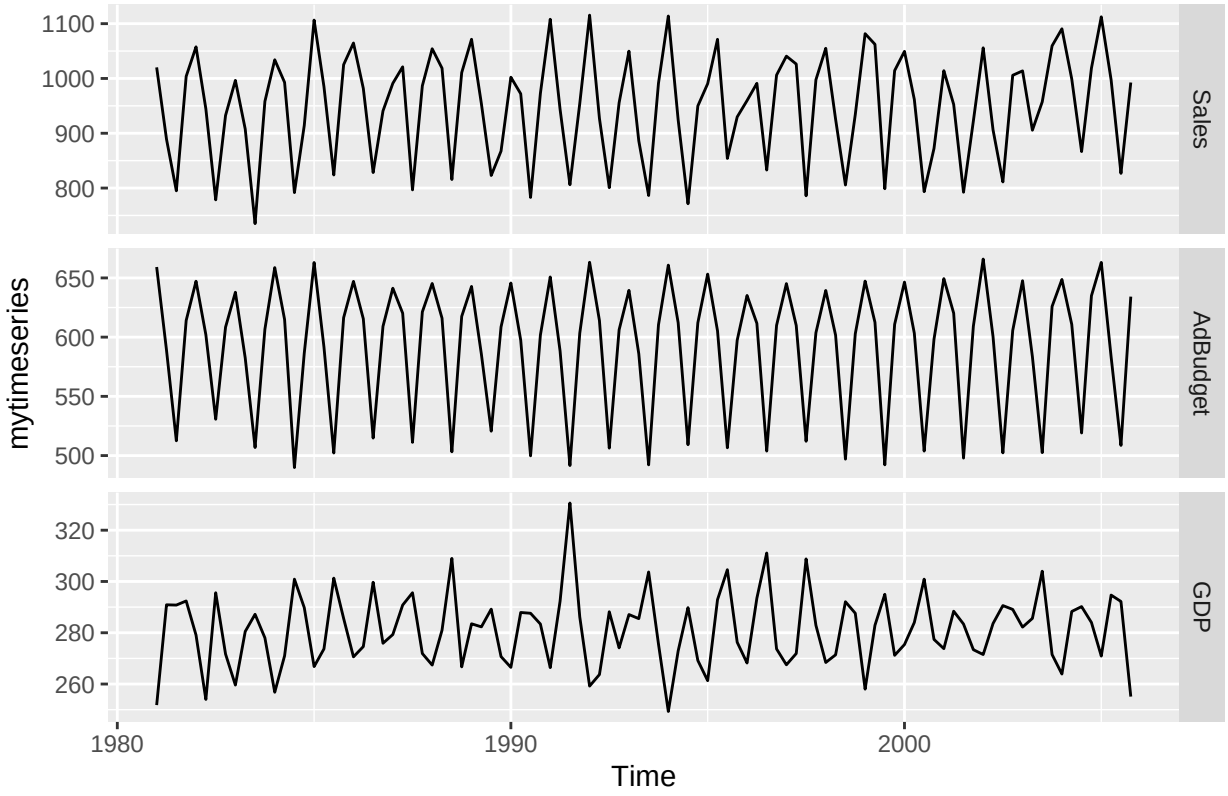
   b. Convert the data to time series

```
mytimeseries <- ts(tute1[,-1], start = 1981, frequency = 4)
```

(The [,-1] removes the first column which contains the quarters as we don't need them now.)

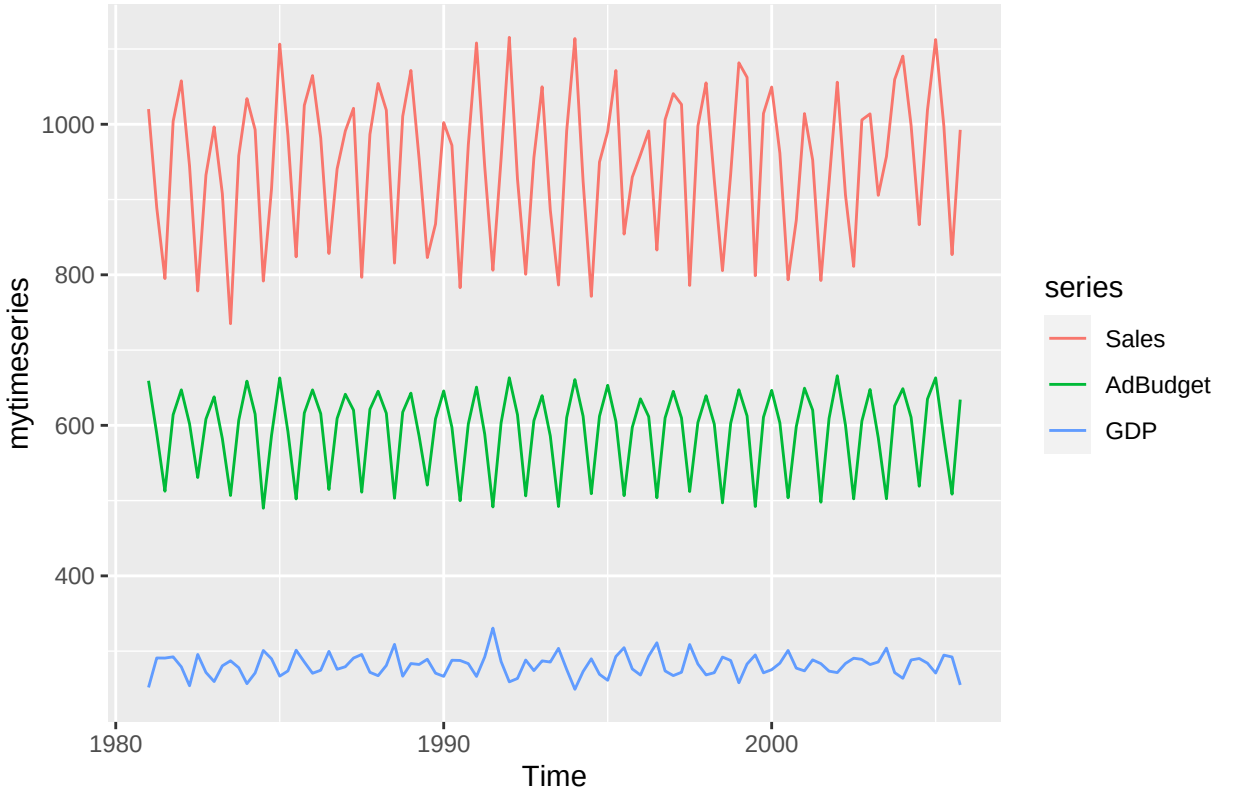
   c. Construct time series plots of each of the three series

```
autoplot(mytimeseries, facets = TRUE)
```



Check what happens when you don't include facets=TRUE.

```
autoplot(mytimeseries)
```



By using facets = TRUE, I got 3 separate vertical axes Sales, AdBudget and GDP. By using facets = FALSE (same as not using facets parameter), I got one vertical axis, with each graph identified by a different color.

3) Download some monthly Australian retail data from the book website. These represent retail sales in various categories for different Australian states, and are stored in a MS-Excel fle.

   a. You can read the data into R with the following script:

```
retaildata <- readxl::read_excel("./retail.xlsx", skip = 1)
```

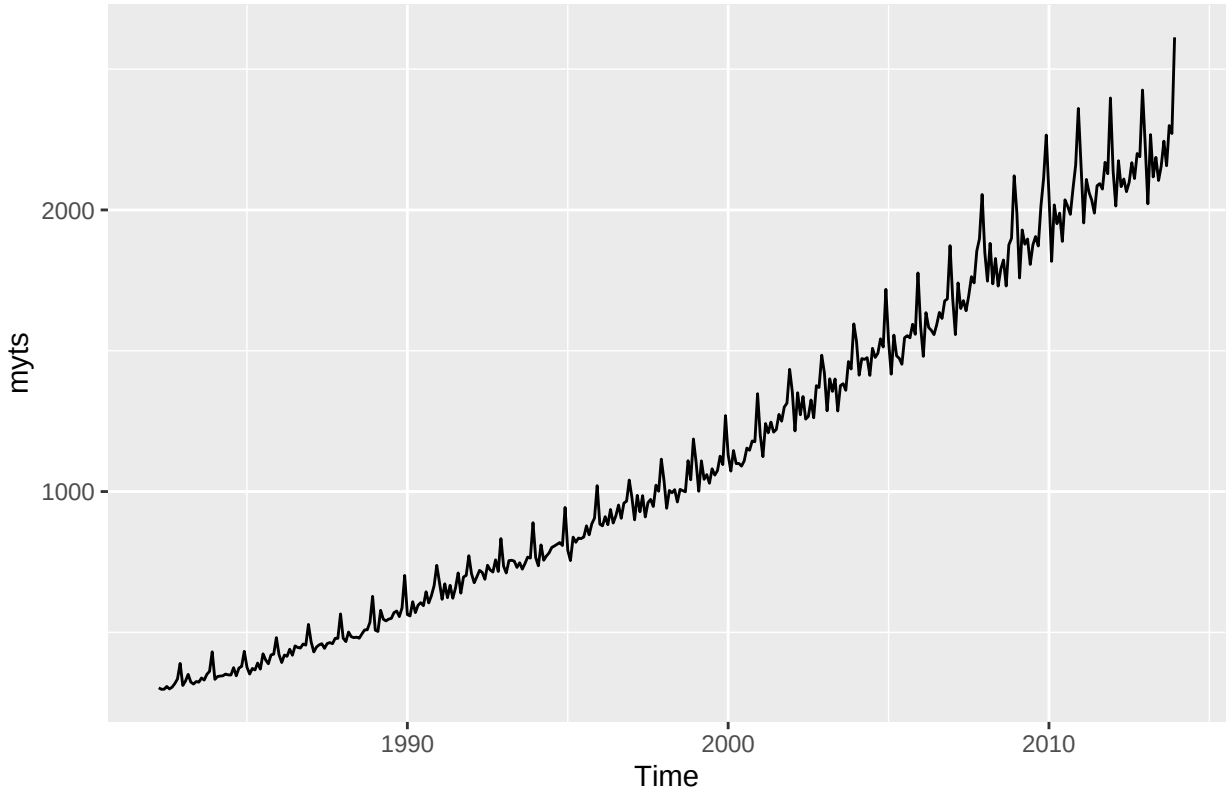
```
The second argument ( skip = 1 ) is required because the Excel sheet has two header rows.
```

   b. Select one of the time series as follows (but replace the column name with your own chosen column):

```
myts <- ts(retaildata[,"A3349335T"], frequency = 12, start = c(1982,4))
```

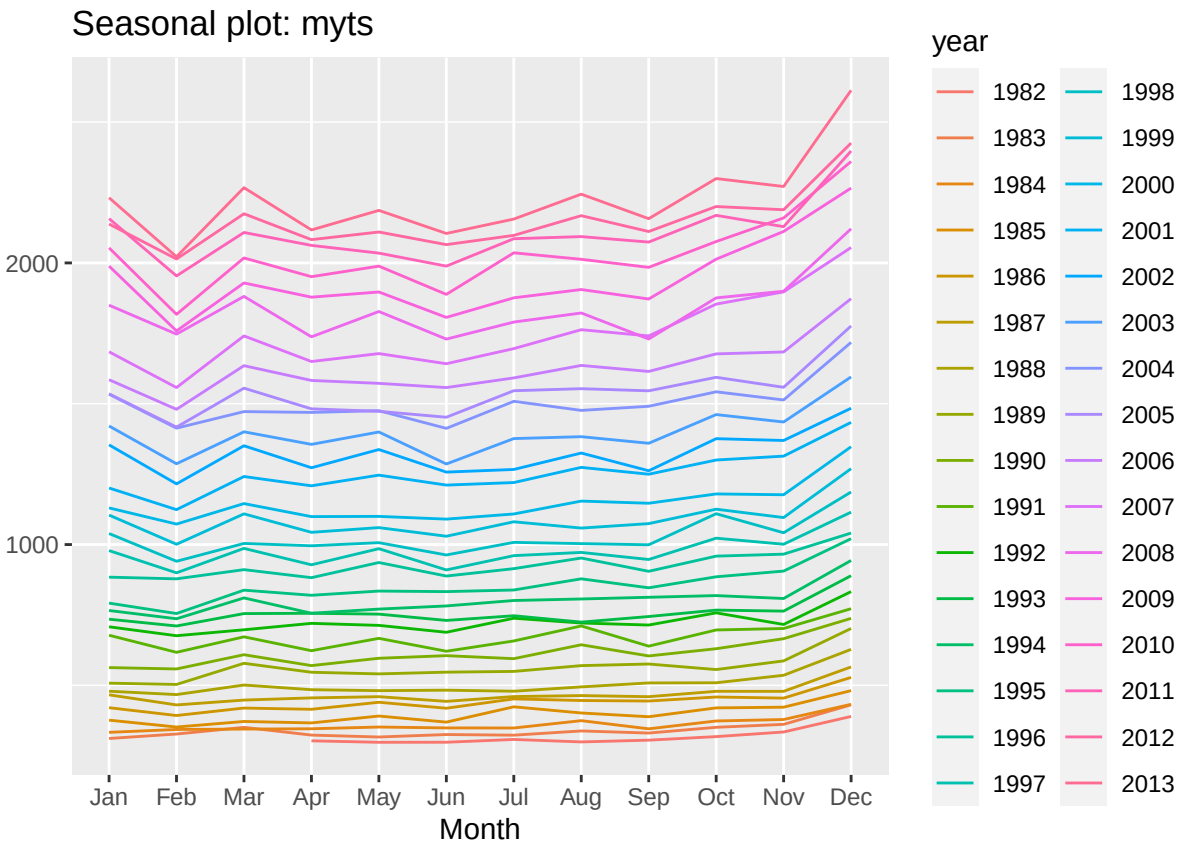
   c. Explore your chosen retail time series using the following functions:
   autoplot(), ggseasonplot(), ggsubseriesplot(), gglagplot(), ggAcf()
   Can you spot any seasonality, cyclicity and trend? What do you learn about the series?

```
autoplot(myts)
```



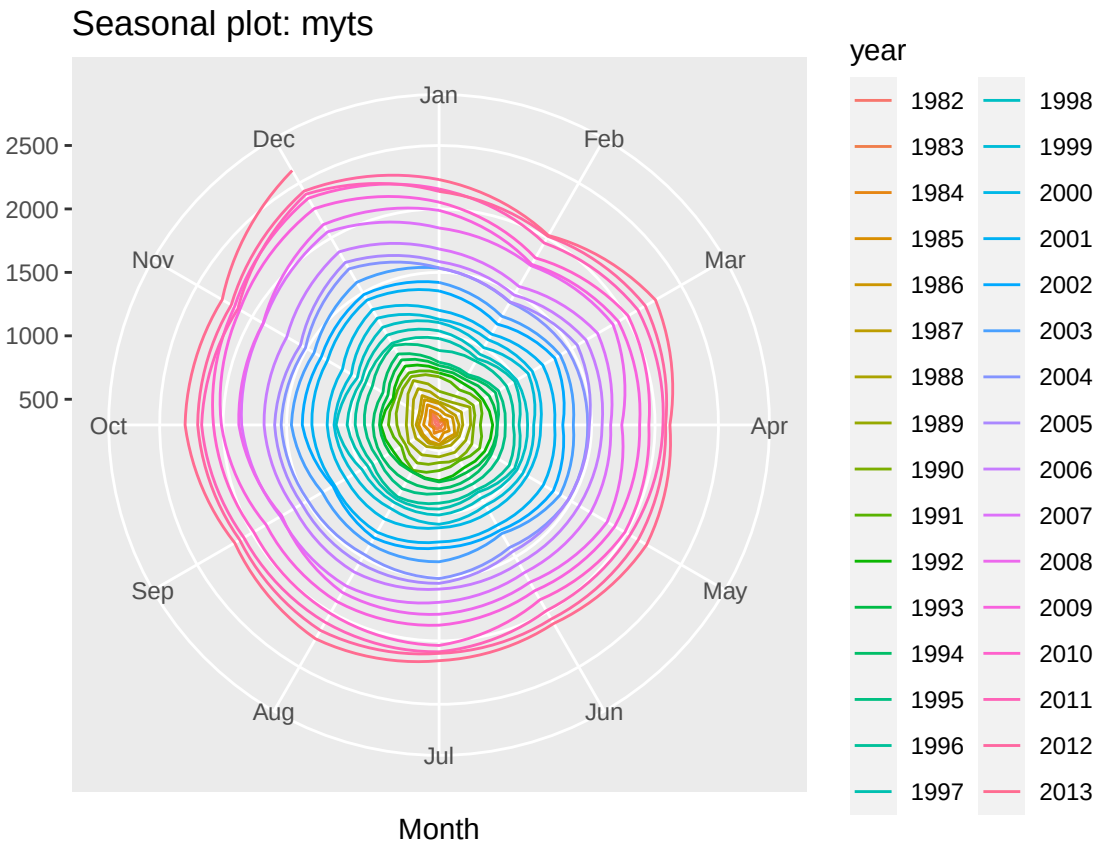
The graph is Seasonal, with a gradual upward Trend.

```
ggseasonplot(myts)
```



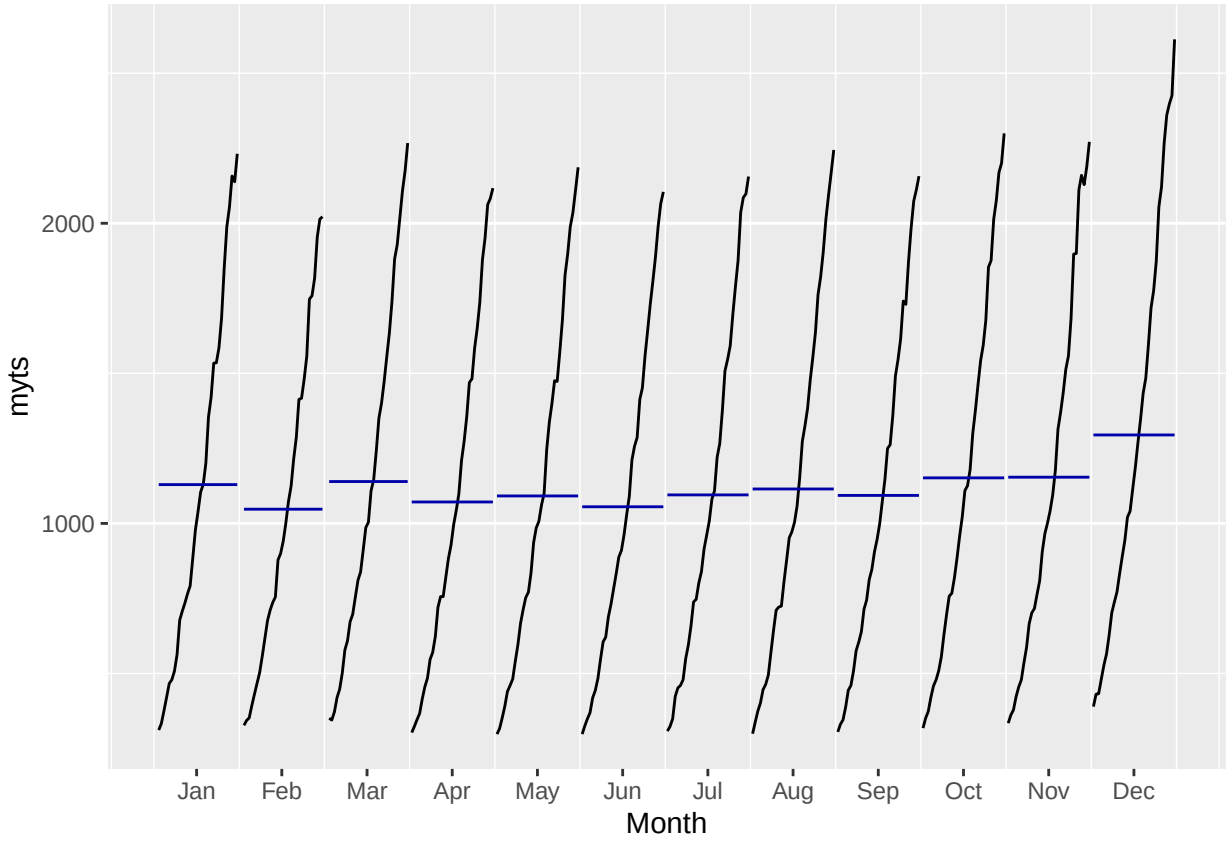
The graph is Seasonal, with very slight Trend. In this situation, a circular graph helps. So, I generated one below, using polar = TRUE.

```
ggseasonplot(myts, polar = TRUE)
```



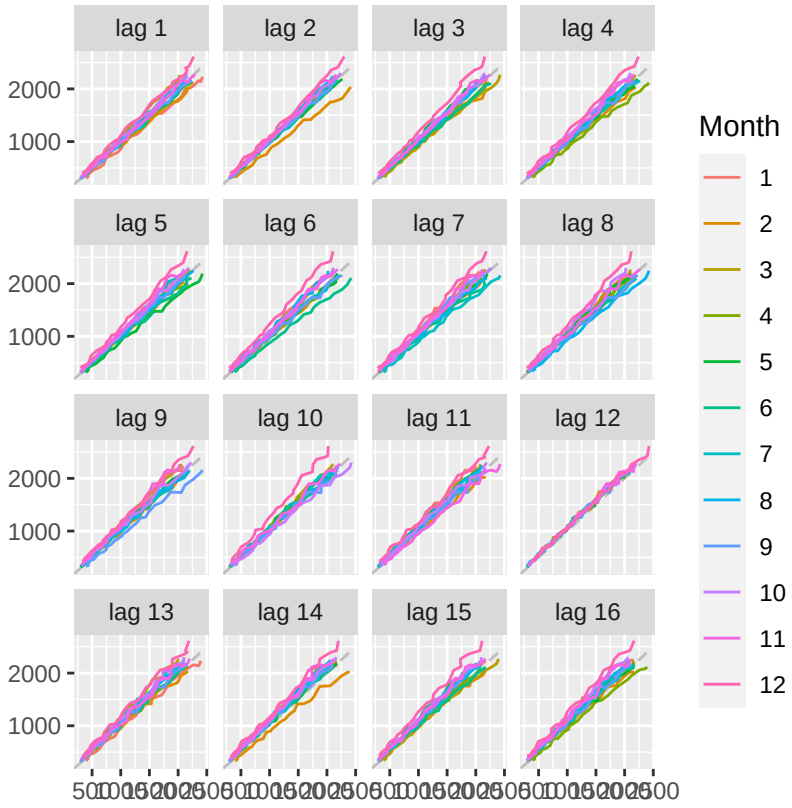
The polar graph makes the Seasonal character obvious.

```
ggsubseriesplot(myts)
```



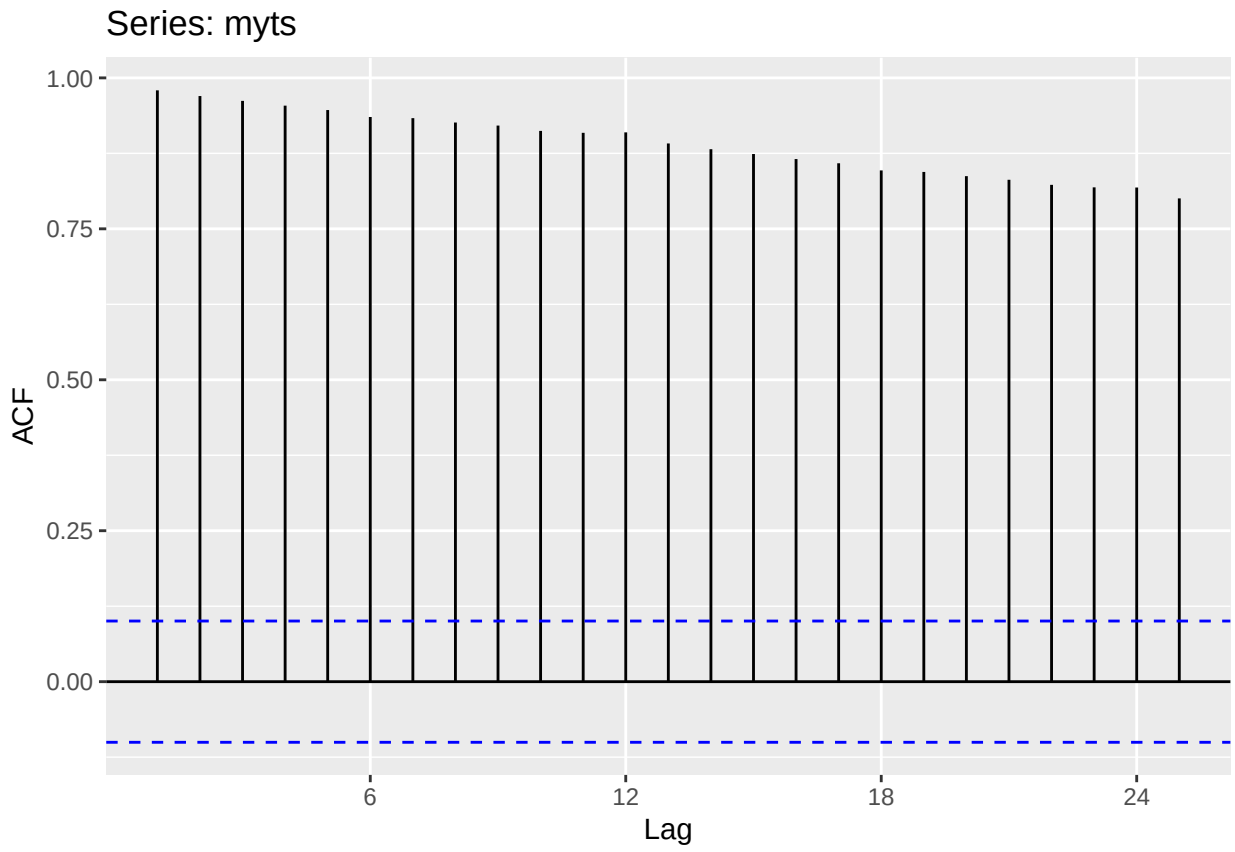
The graph is Seasonal.

```
gglagplot(myts)
```



Shows strong positive correlation. Looks like correlation coefficient is almost 1.

```
ggAcf(myts)
```



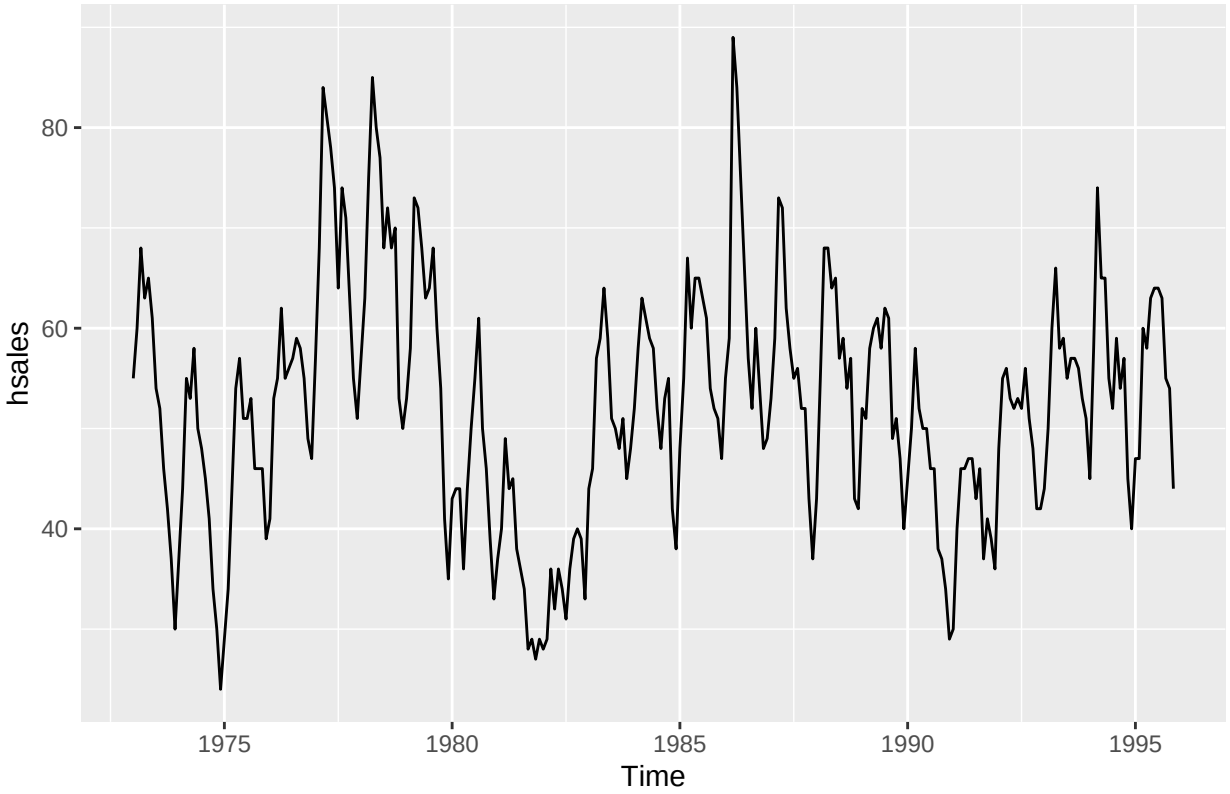
Seasonal and downward Trending.

6. Use the following graphics functions: autoplot(), ggseasonplot(), ggsubseriesplot(), gglagplot(), ggAcf() and explore features from the following time series: hsales, usdeaths, bricksq, sunspotarea, gasoline.

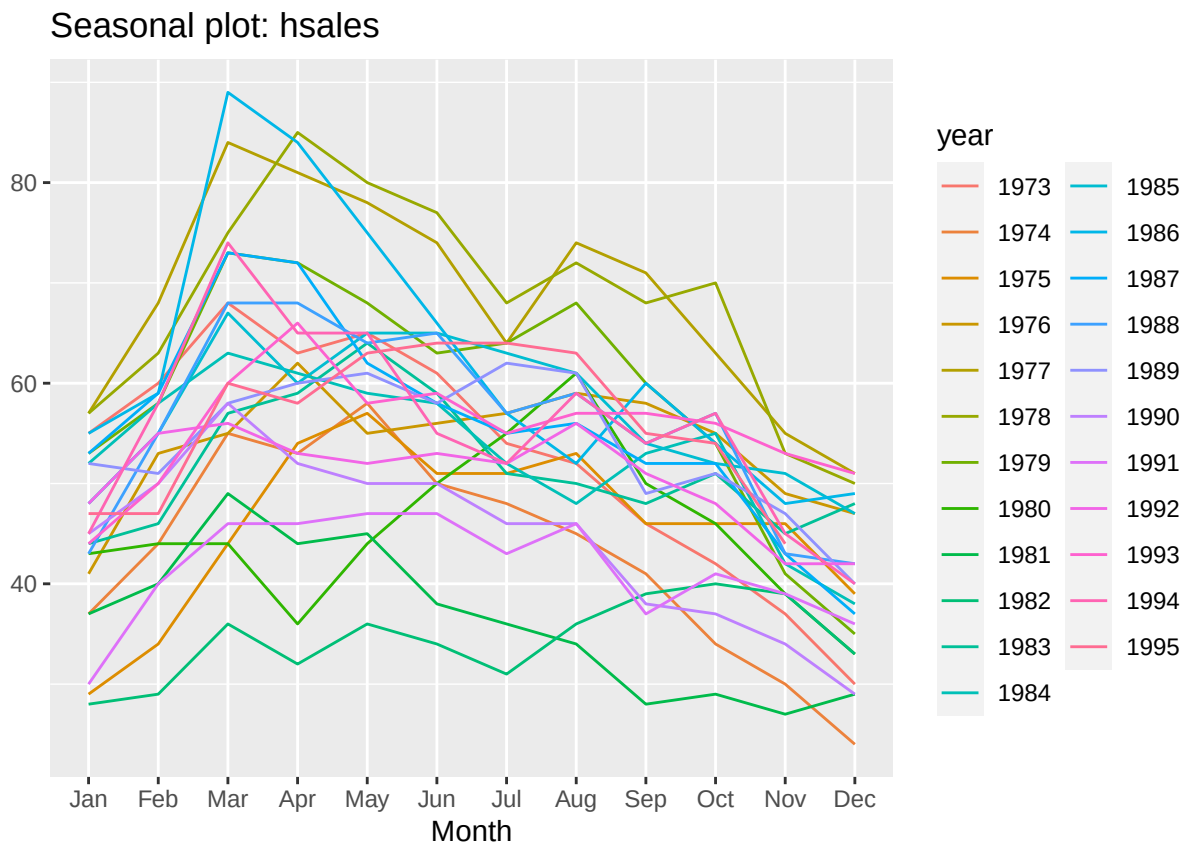
- Can you spot any seasonality, cyclicity and trend?
- What do you learn about the series?

hsales

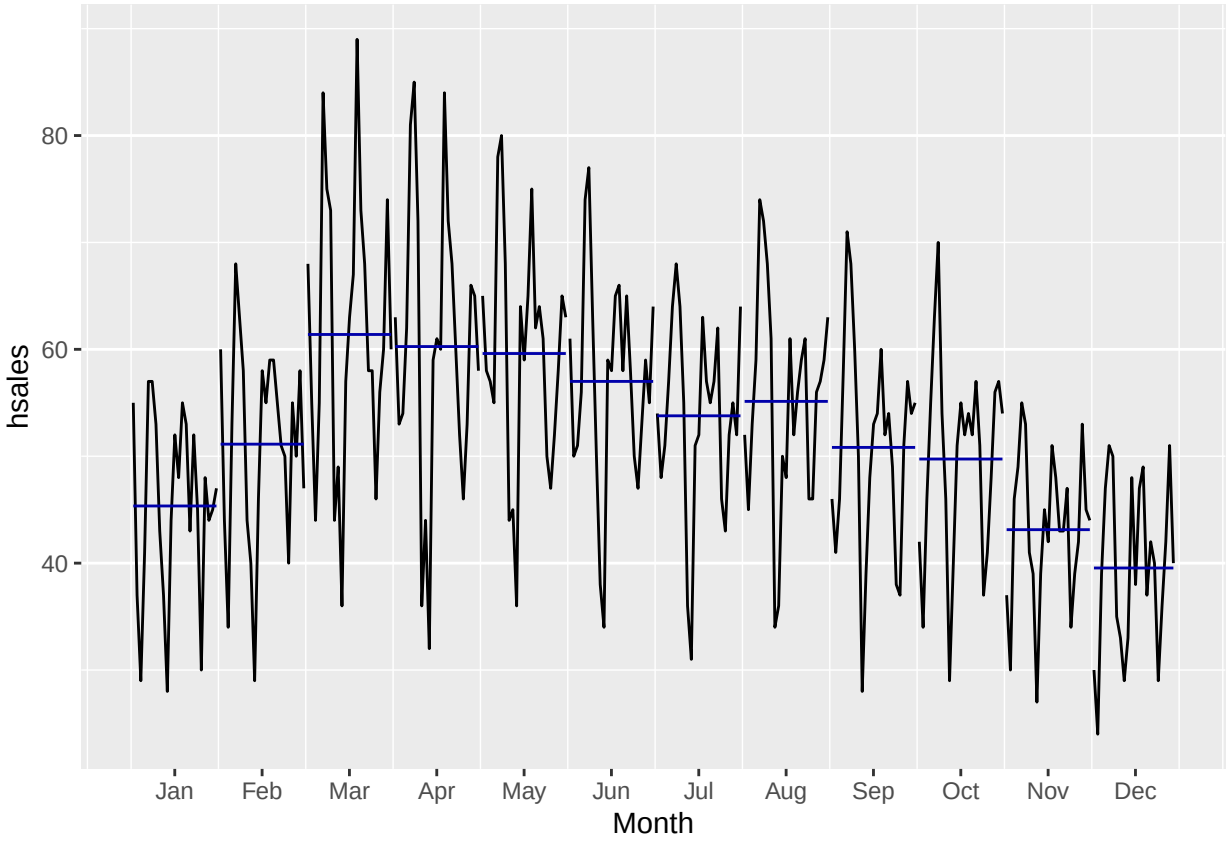
```
autoplot(hsales)
```



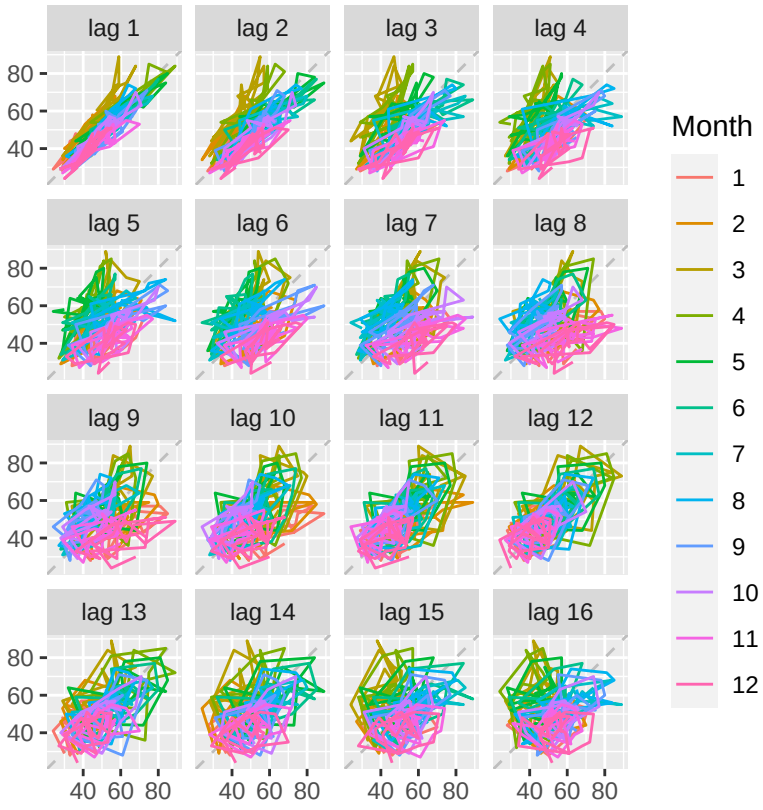

```
ggseasonplot(hsales)
```



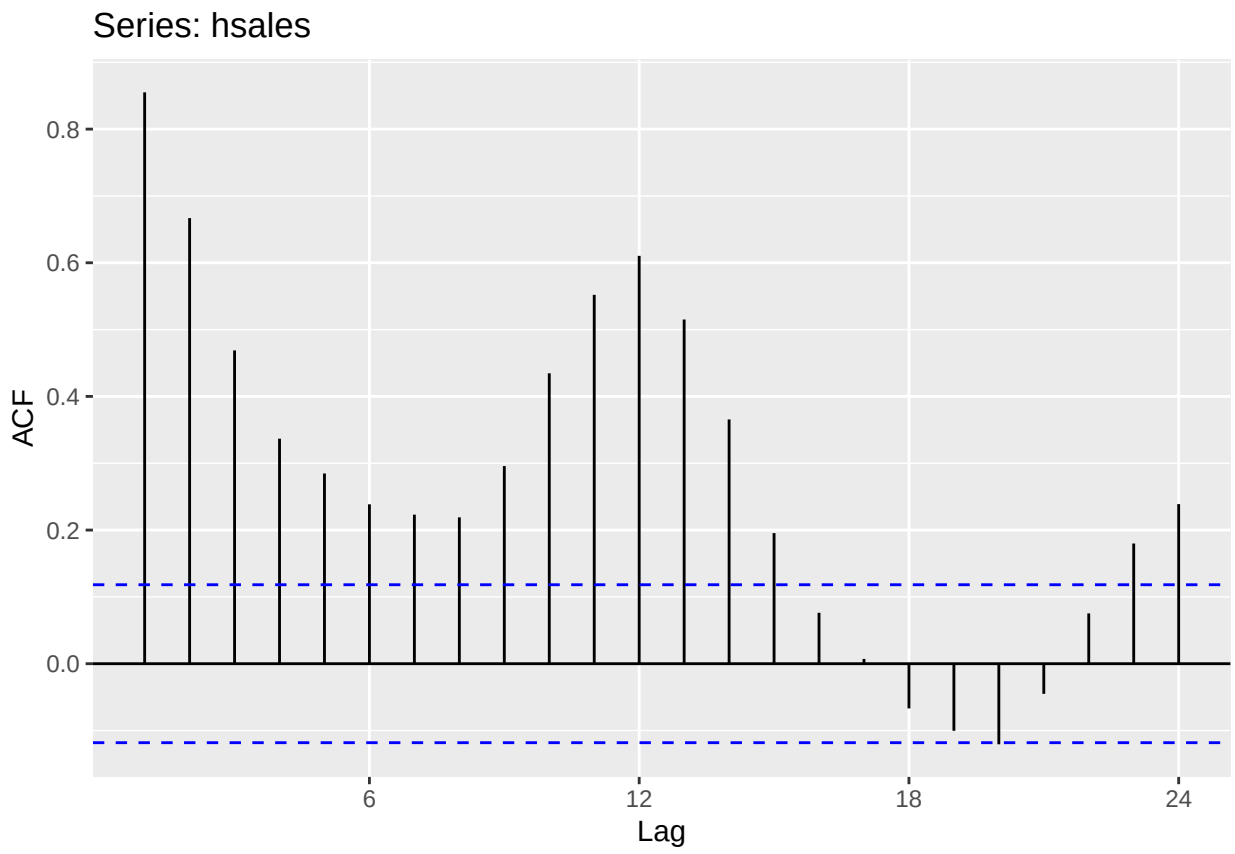

```
ggsubseriesplot(hsales)
```




```
gglagplot(hsales)
```



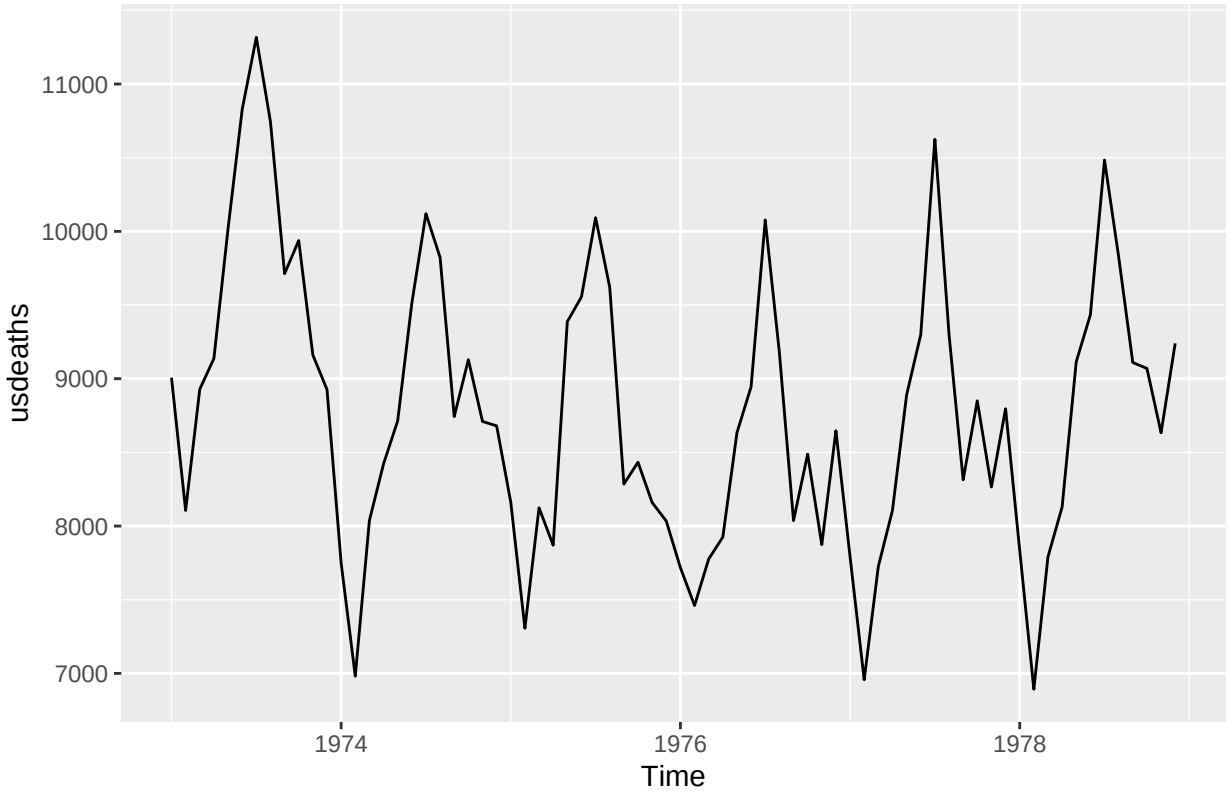

```
ggAcf(hsales)
```



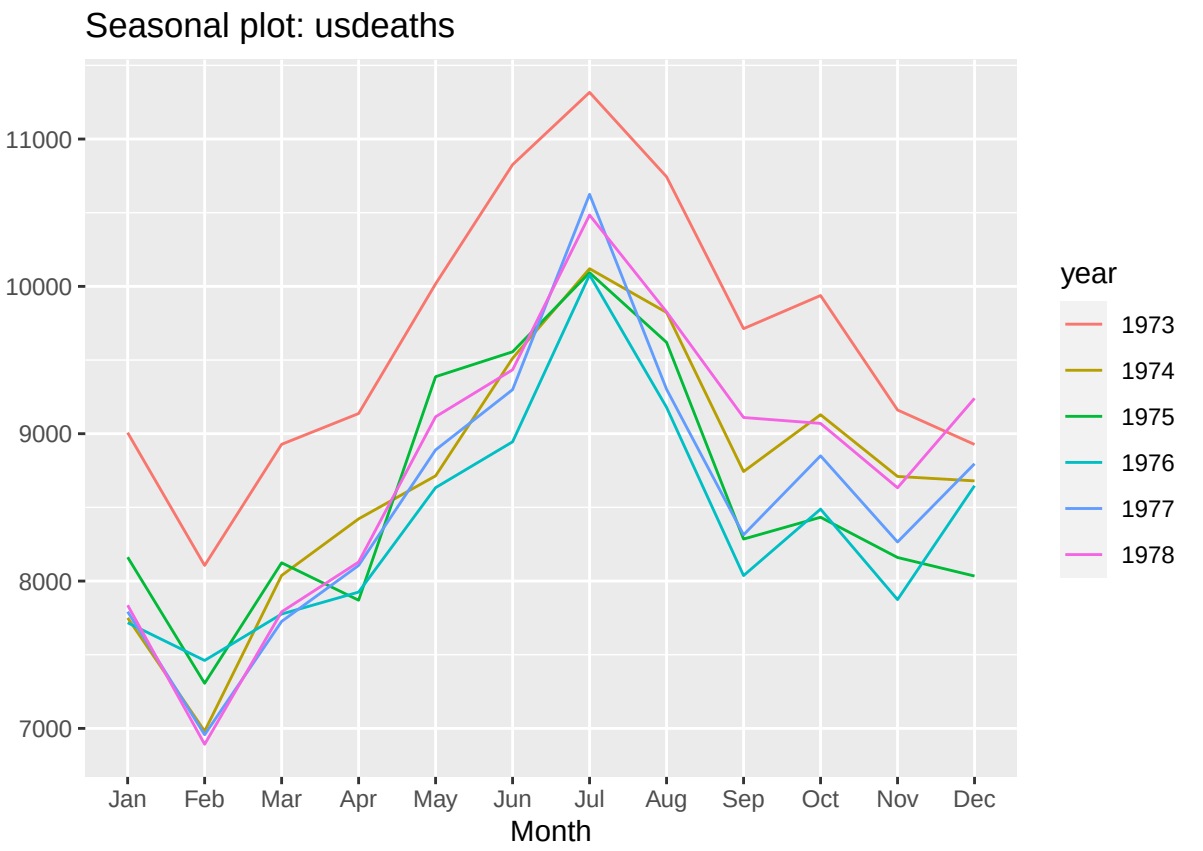
- Can you spot any seasonality, cyclicity and trend?
  Answer: One-family home sales in the US is highest in March. The ACF plot suggests that there is some annual cycle (fluctuating in two years).
- What do you learn about the series?
  Answer: Early Spring (March to May) are busy months, and winter is slow.

usdeaths

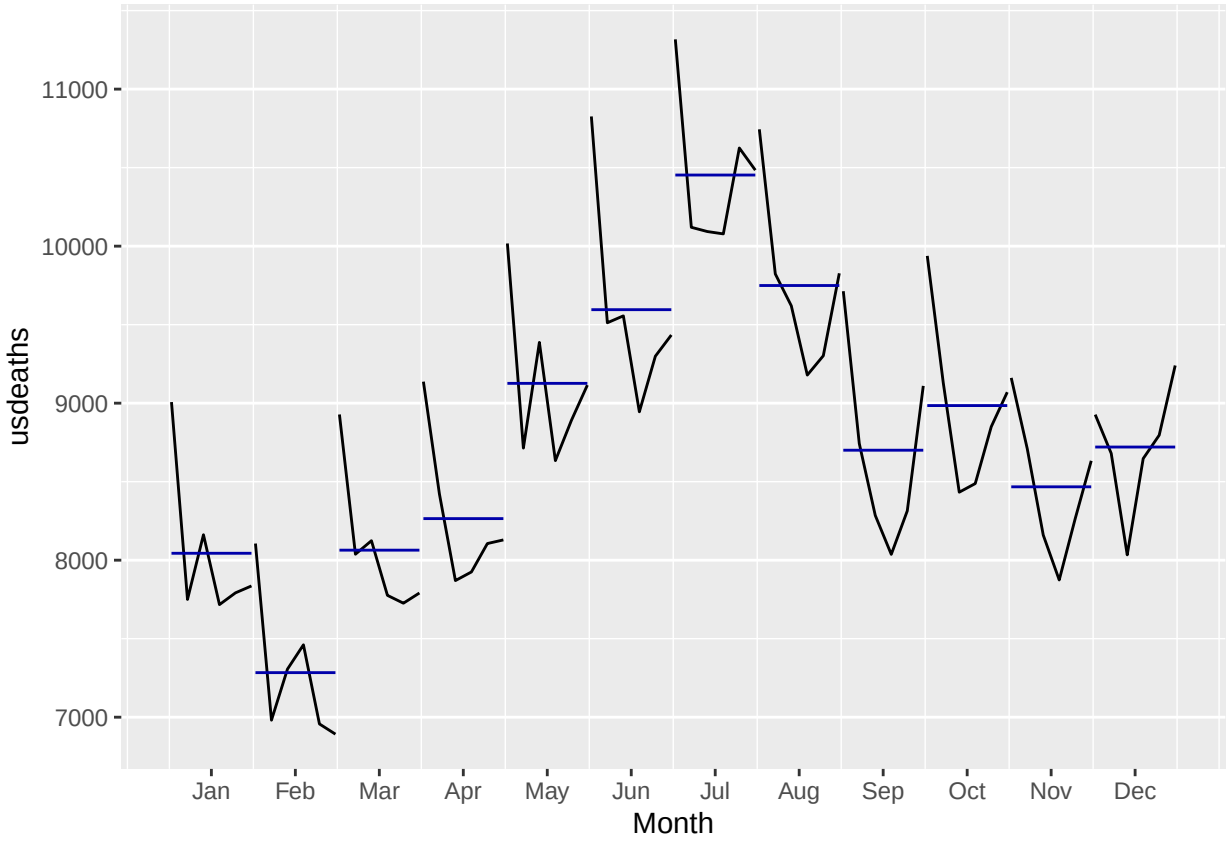
```
autoplot(usdeaths)
```



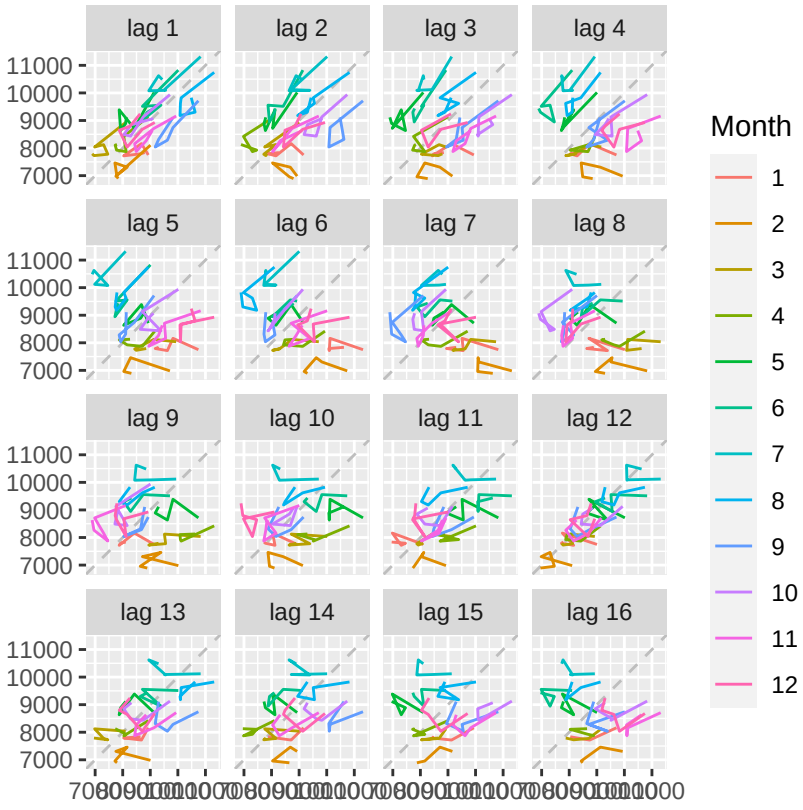

```
ggseasonplot(usdeaths)
```



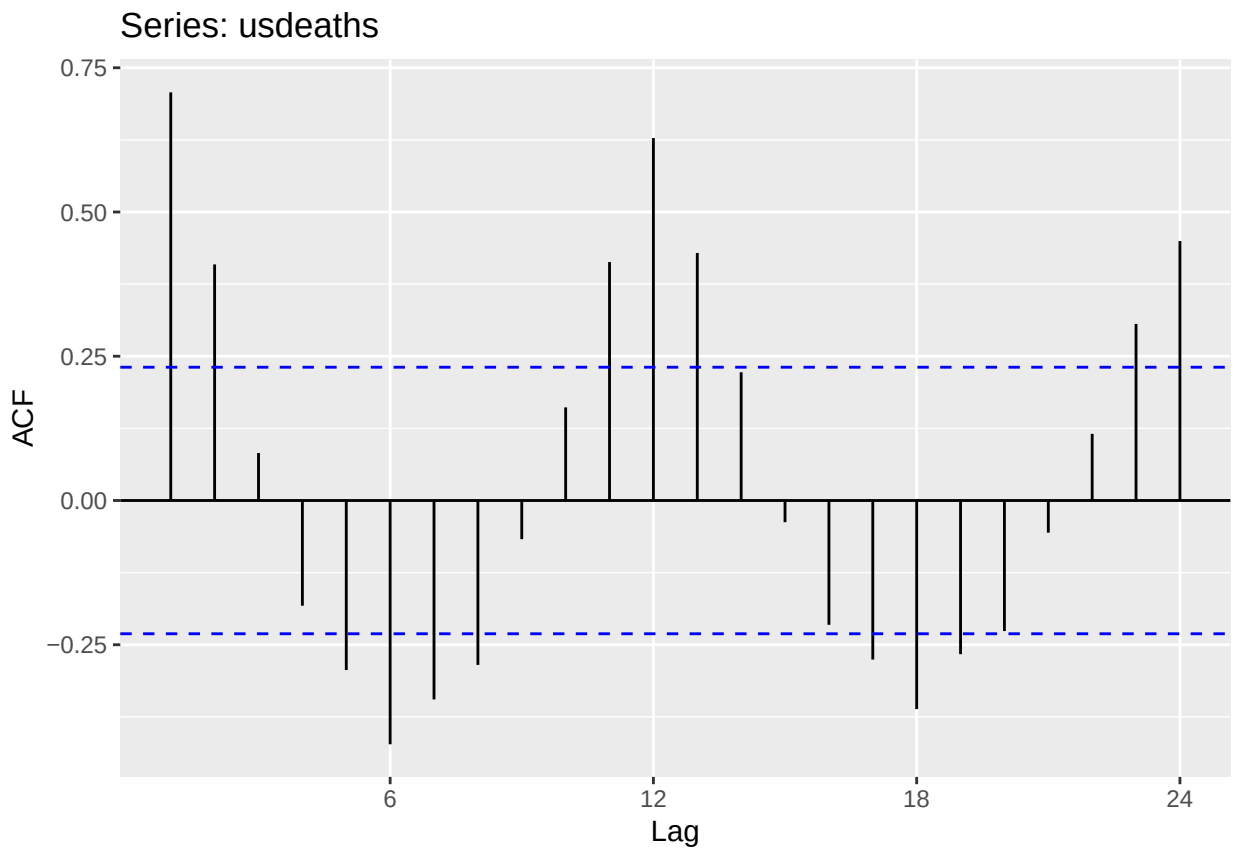

```
ggsubseriesplot(usdeaths)
```




```
gglagplot(usdeaths)
```



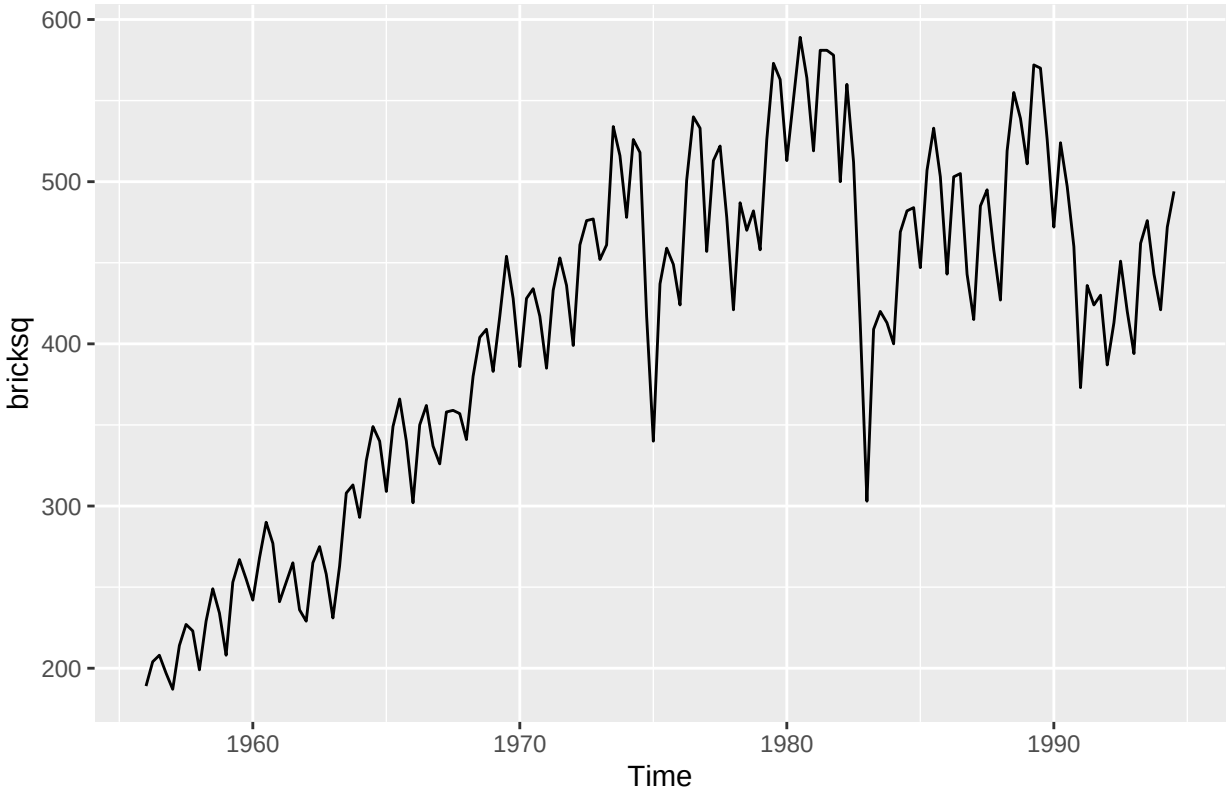

```
ggAcf(usdeaths)
```



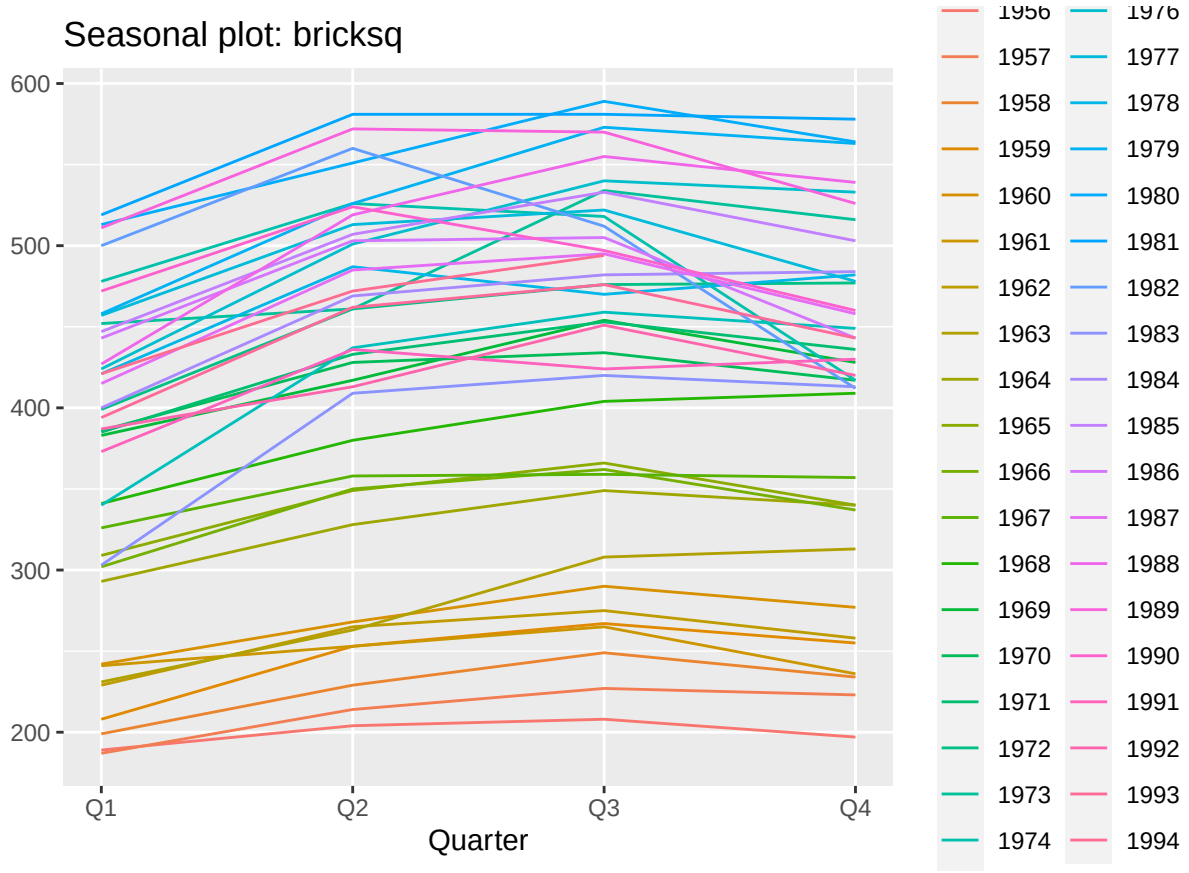
- Can you spot any seasonality, cyclicity and trend?
  Answer: Data is a seasonal.
- What do you learn about the series?
  Answer: Accidental deaths are highest in July.

bricksq

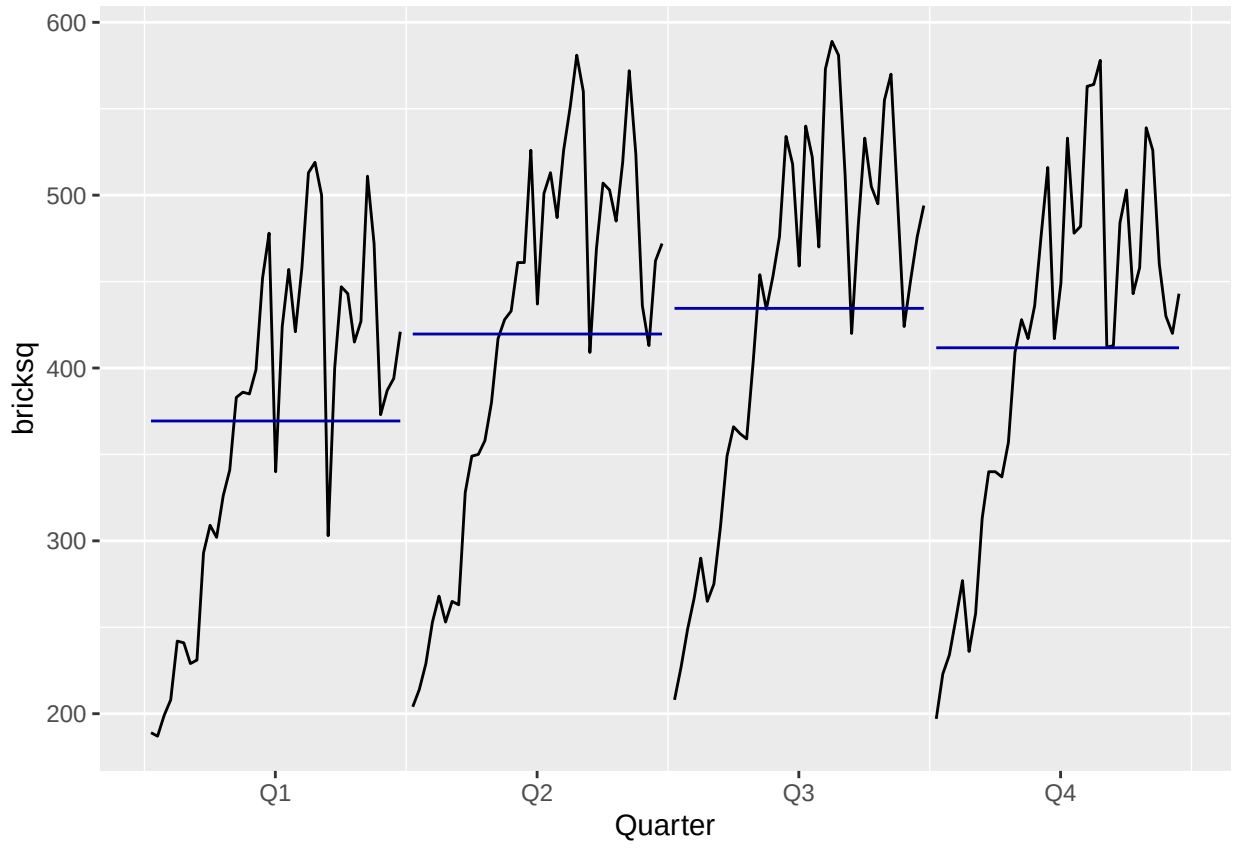
```
autoplot(bricksq)
```



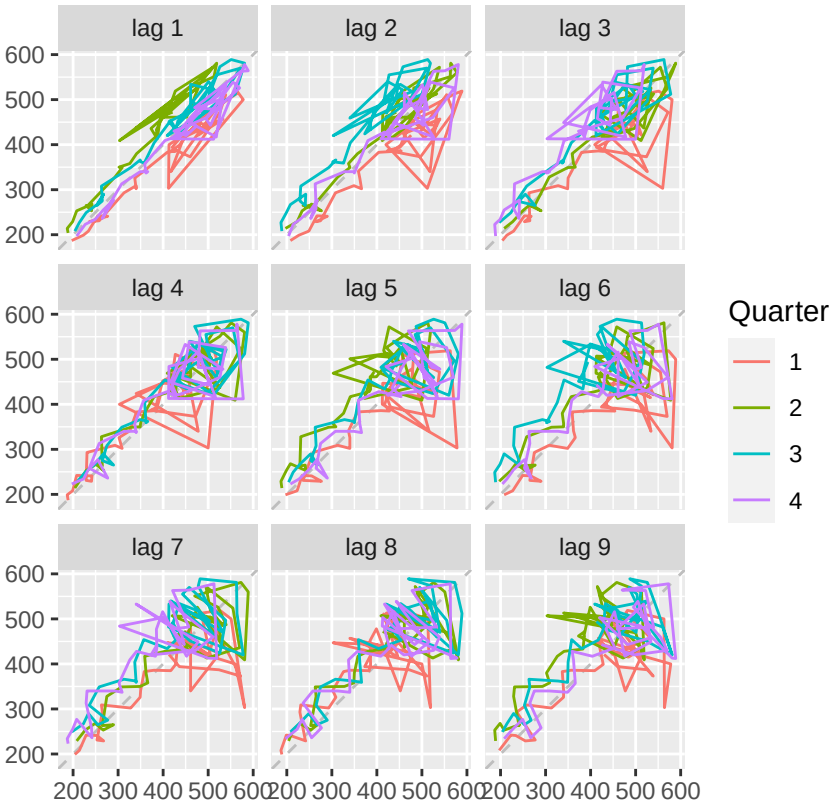

```
ggseasonplot(bricksq)
```



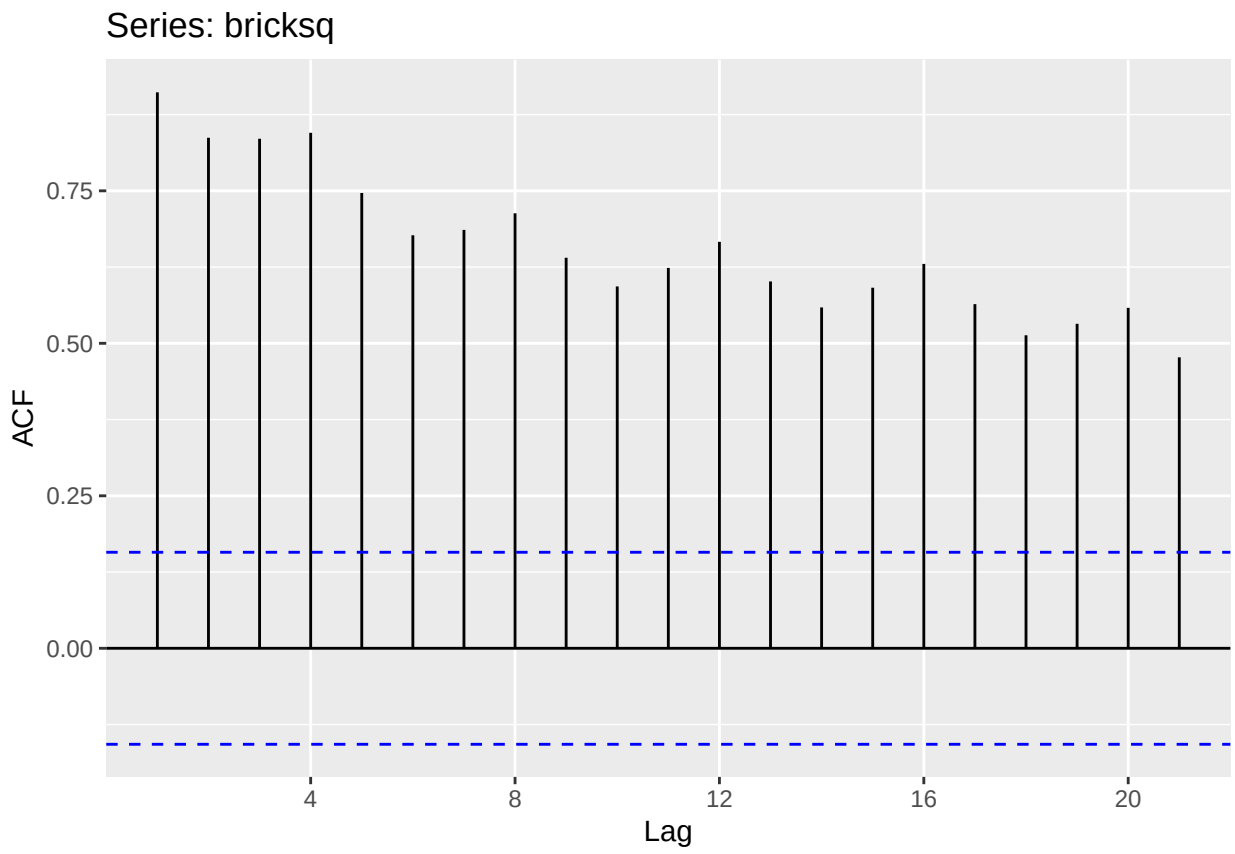

```
ggsubseriesplot(bricksq)
```




```
gglagplot(bricksq)
```



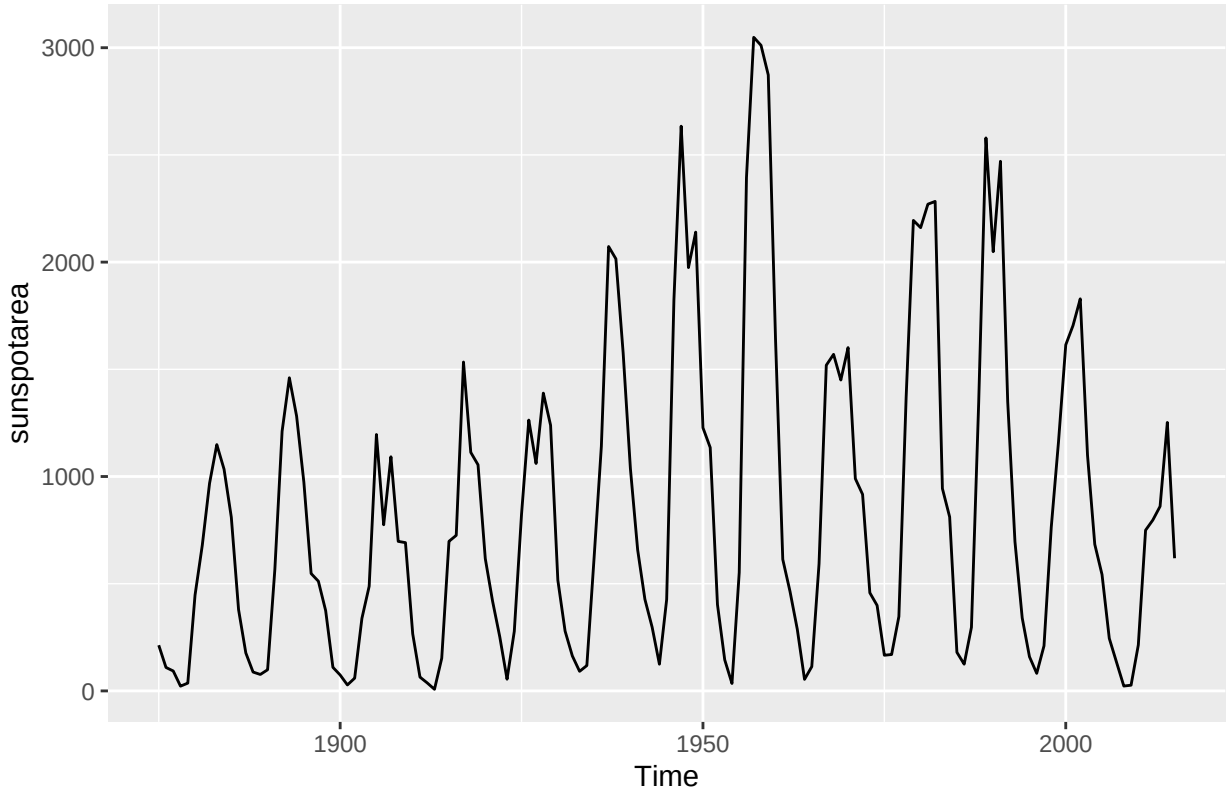

```
ggAcf(bricksq)
```



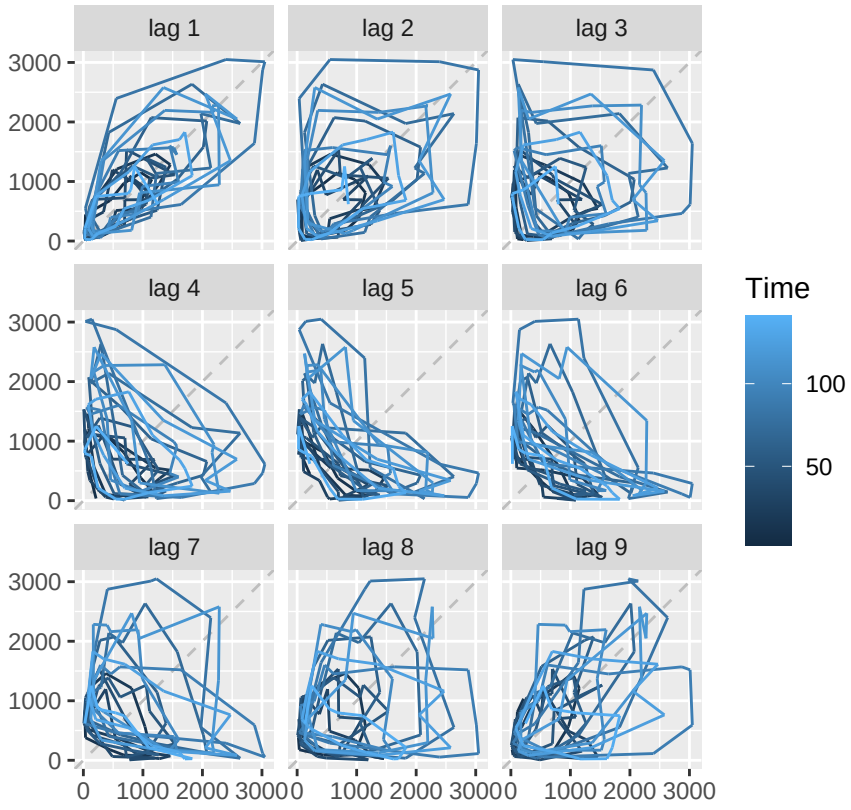
- Can you spot any seasonality, cyclicity and trend?
  Answer: Trending upward until about the 1980's, and then plumets.
- What do you learn about the series?
  Answer: First quarter is a slow relative to other quarters.

sunspotarea

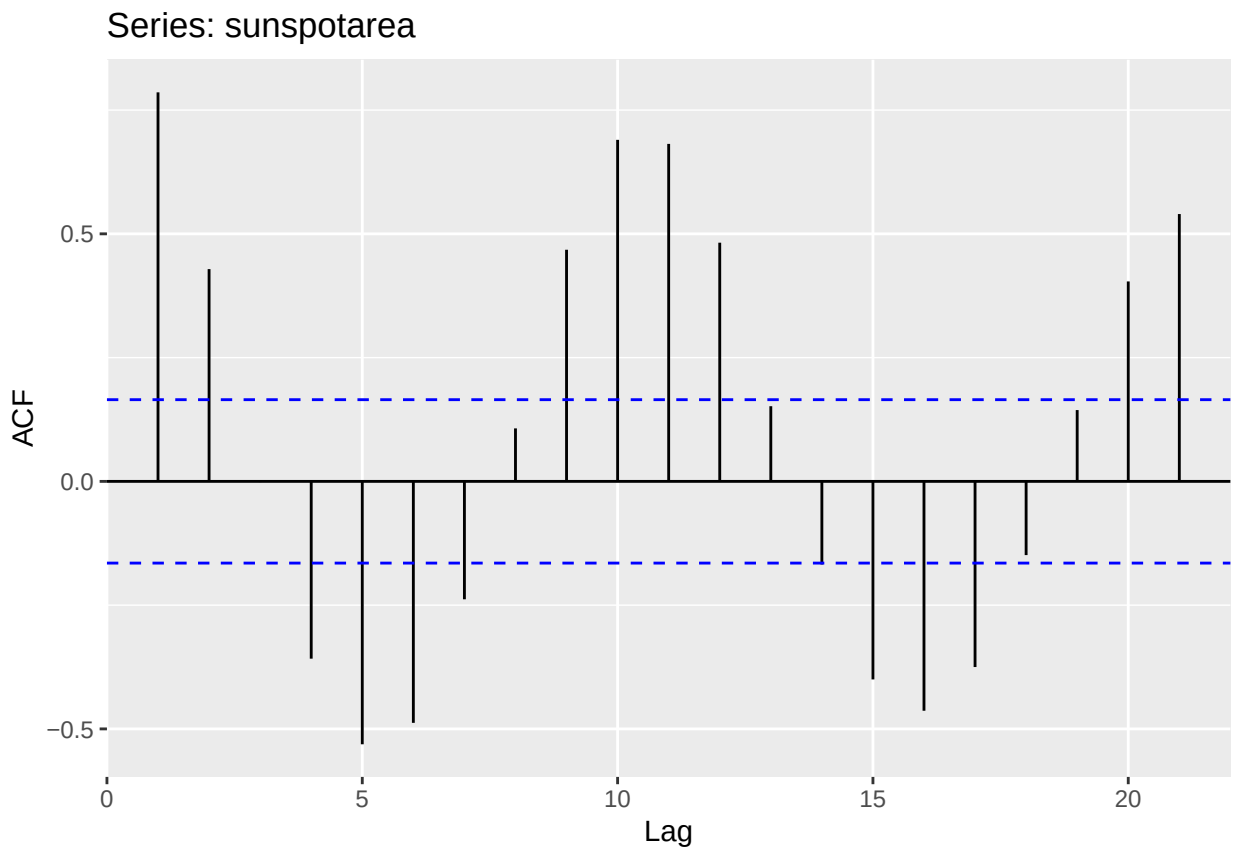
```
autoplot(sunspotarea)
```




```
#ggseasonplot(sunspotarea)      ## Doesn't fulfil condition for Seasonality
#ggsubseriesplot(sunspotarea)   ## Doesn't fulfil condition for Seasonality
gglagplot(sunspotarea)
```



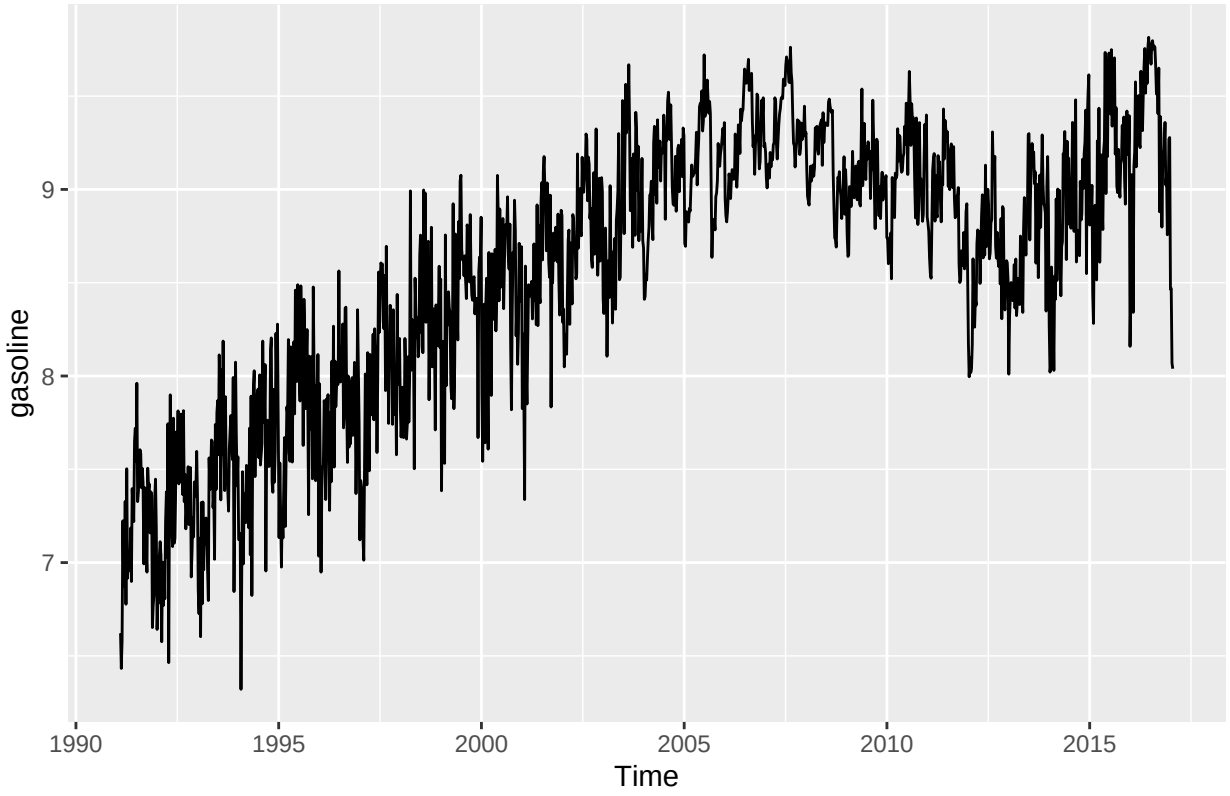

```
ggAcf(sunspotarea)
```



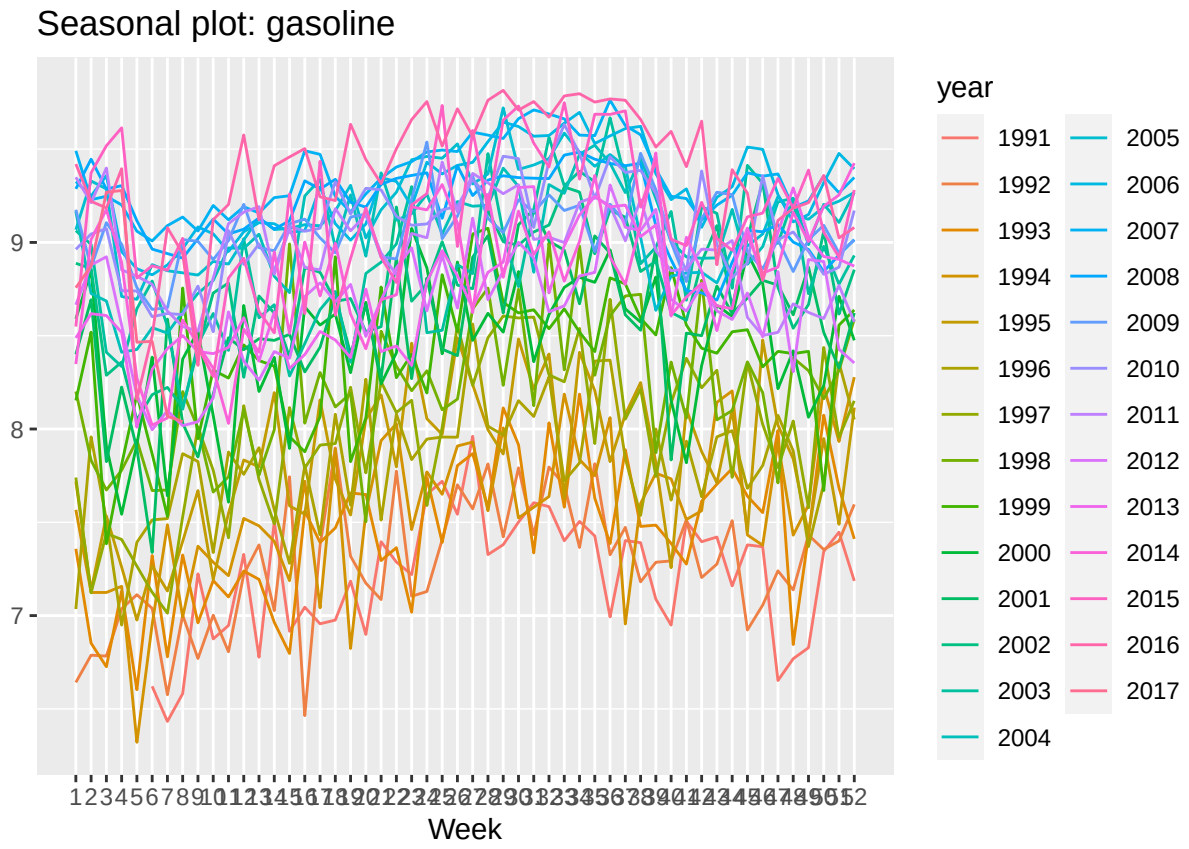
- Can you spot any seasonality, cyclicity and trend?
  Answer: There appears to be a cycle of about a decade.
- What do you learn about the series?

gasoline

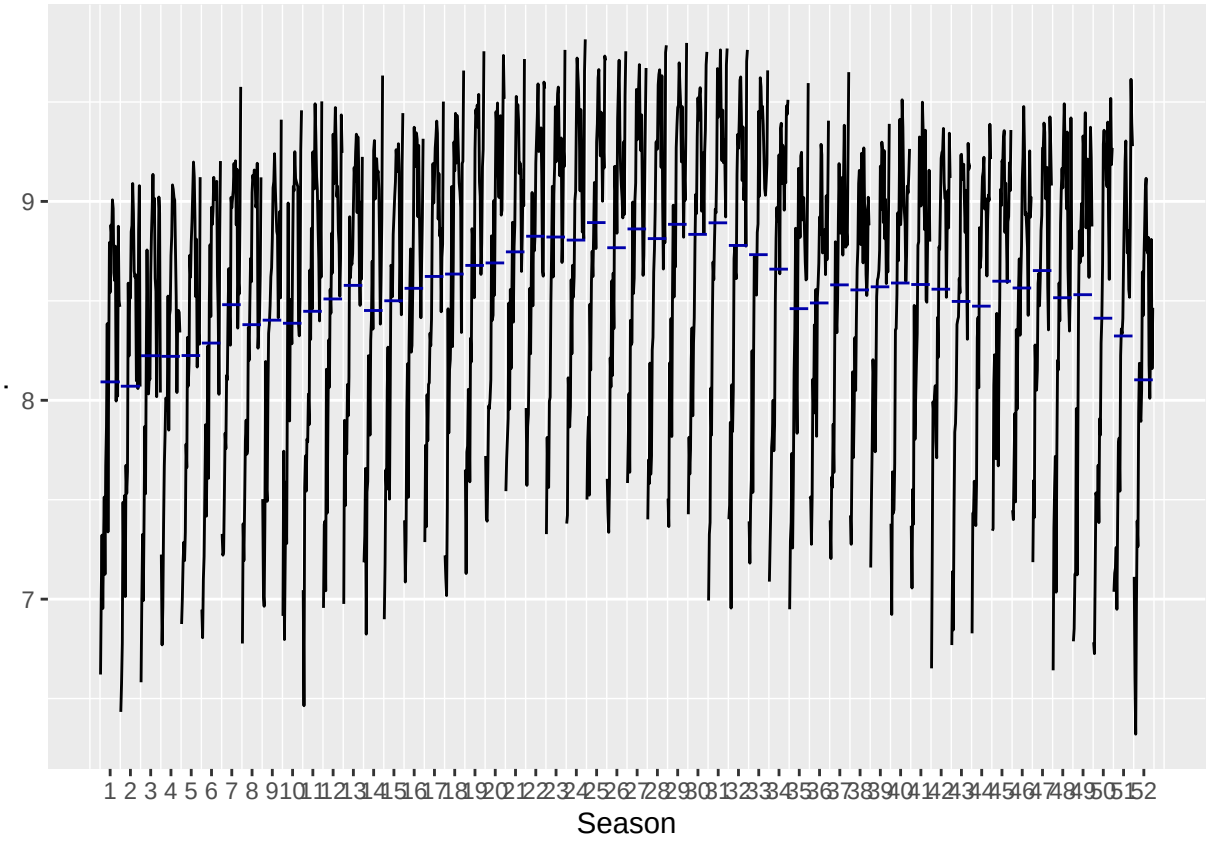
```
autoplot(gasoline)
```



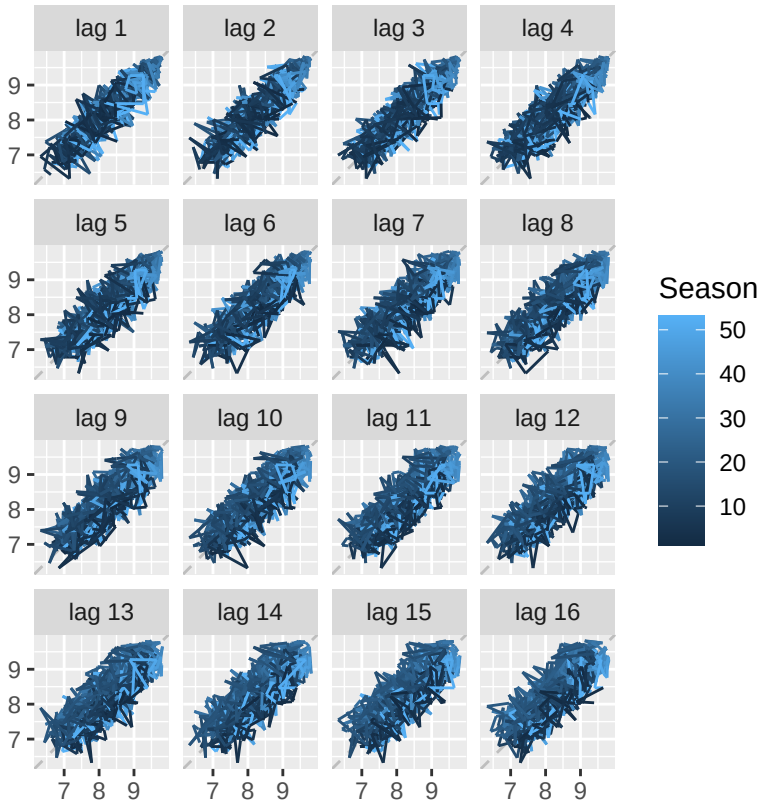

```
ggseasonplot(gasoline)
```



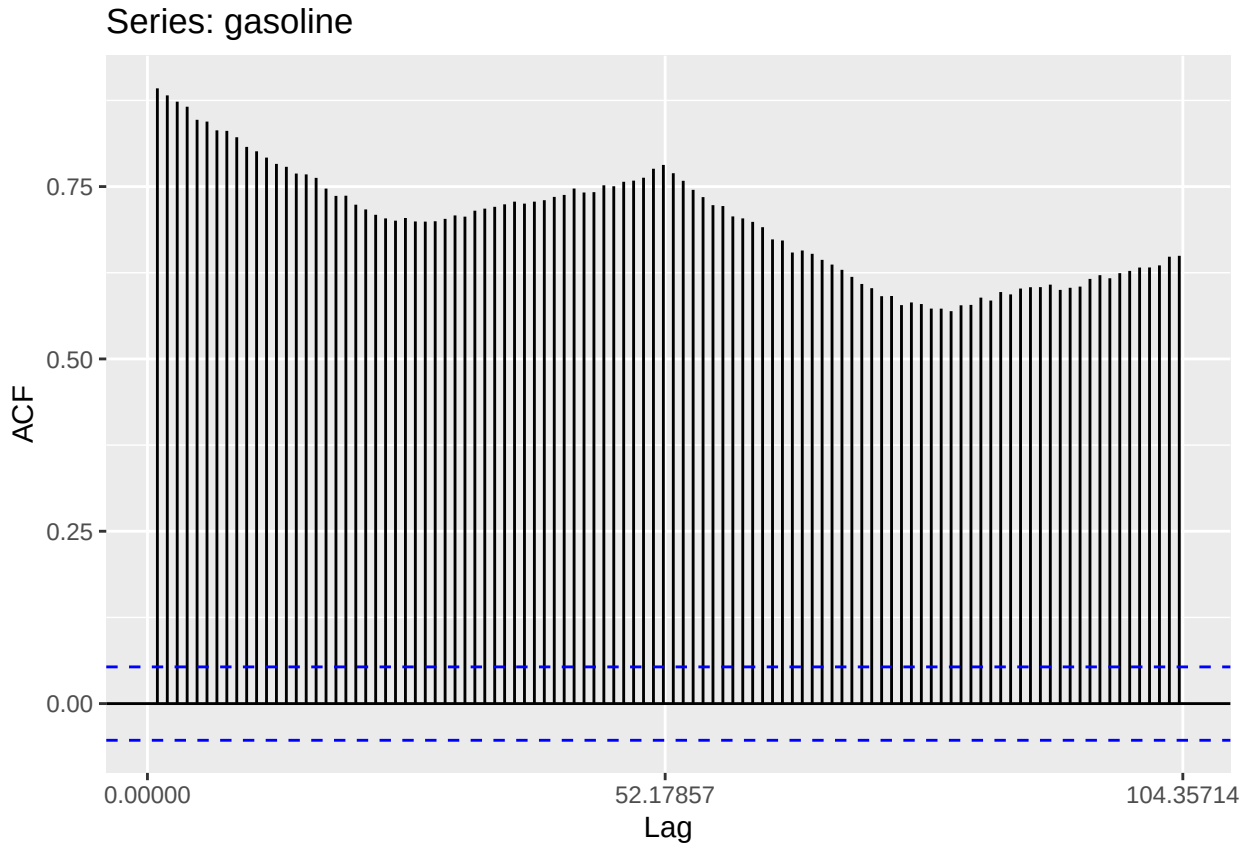

```
gasoline %>%
  as.vector()%>%
  ts(., frequency=52) %>%
  ggsubseriesplot()
```




```
gglagplot(gasoline)
```




```
ggAcf(gasoline)
```



- Can you spot any seasonality, cyclicity and trend?
  Answer: There is a trend and some seasonality to the data.
- What do you learn about the series?
  Answer: The trend of the gasoline availability has been on the rise, specially during the summer.

Marker: 624-01