

# Introduction

AI Nosedal  
University of Toronto

January 8, 2019

A **time series** is a sequence of observations on a variable measured at successive points in time or over successive periods of time. The measurements may be taken every hour, day, week, month, or year, or at any other regular interval. The pattern of the data is an important factor in understanding how the time series has behaved in the past. If such behavior can be expected to continue in the future, we can use the past pattern to guide us in selecting an appropriate forecasting method.

# Time Series Plot

To identify the underlying pattern in the data, a useful first step is to construct a **time series plot**. A time series plot is a graphical representation of the relationship between time and the time series variable; time is on the horizontal axis and the time series values are shown on the vertical axis. Let us review some of the common types of data patterns that can be identified when examining a time series plot.

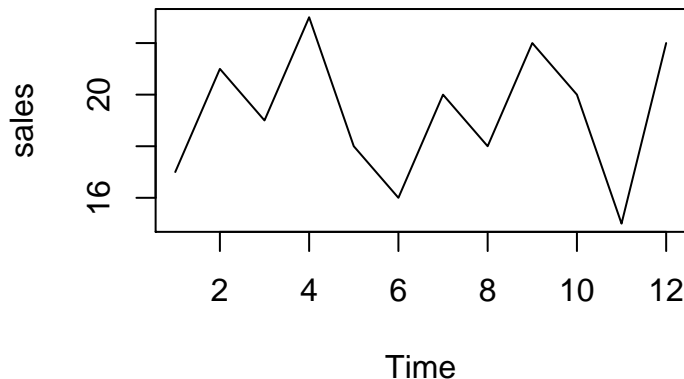
# Horizontal Pattern

A **horizontal pattern** exists when the data fluctuate around a constant mean. To illustrate a time series with a horizontal pattern, consider the 12 weeks of data in Table 1.

Week	Sales (1000s of gallons)
1	17
2	21
3	19
4	23
5	18
6	16
7	20
8	18
9	22
10	20
11	15
12	22

```
sales=c(17,21,19,23,18,16,20,18,22,20,15,22);  
plot.ts(sales, main="Gasoline Sales Time Series Plot");
```

## Gasoline Sales Time Series Plot



# Homework

This course uses R. R is an open-source computing package which has seen a huge growth in popularity in the last few years. R can be downloaded from <https://cran.r-project.org>

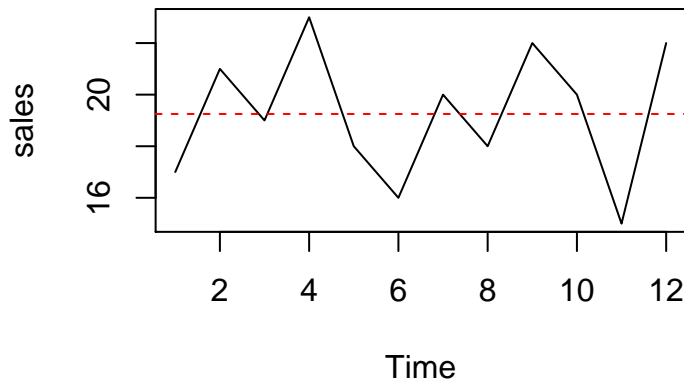
**Please, download R and bring your laptop next time.**

The average value or mean for this time series is 19.25 or 19,250 gallons per week. Note how the data fluctuate around the sample mean of 19,250 gallons. Although random variability is present, we would say that these data follow a horizontal pattern.



```
sales=c(17,21,19,23,18,16,20,18,22,20,15,22);  
plot.ts(sales, main="Gasoline Sales Time Series Plot");  
abline(h=mean(sales),lty=2,col="red");
```

## Gasoline Sales Time Series Plot



The term **stationary time series** is used to denote a time series whose statistical properties are independent of time. In particular this means that

- The process generating the data has a constant mean.
- The variability of the time series is constant over time.

A time series plot for a stationary time series will always exhibit a horizontal pattern. But simply observing a horizontal pattern is not sufficient evidence to conclude that the time series is stationary.

# Trend Pattern

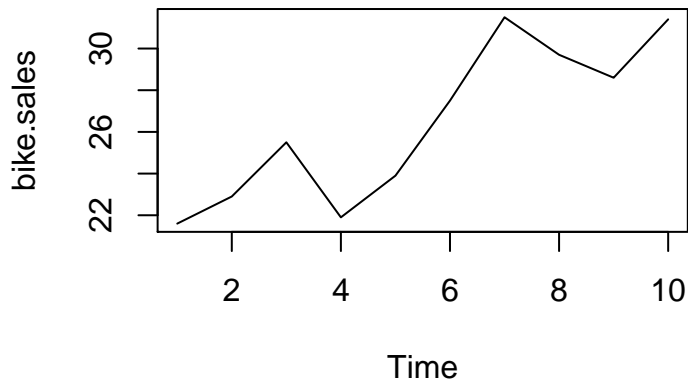
Although time series data generally exhibit random fluctuations, a time series may also show gradual shifts or movements to relatively higher or lower values over a longer period of time. If a time series plot exhibits this type of behavior, we say that a **trend pattern** exists. To illustrate a time series with a trend pattern, consider the time series of bicycle sales for a particular manufacturer over the past 10 years, as shown in Table 2 and Figure 2.

## Table 2

Year	Sales (1000s )
1	21.6
2	22.9
3	25.5
4	21.9
5	23.9
6	27.5
7	31.5
8	29.7
9	28.6
10	31.4

```
bike.sales=  
c(21.6,22.9,25.5,21.9,23.9,27.5,31.5,29.7,28.6,31.4);  
plot.ts(bike.sales, main="Bicycle Sales Time Series Plot");
```

## Bicycle Sales Time Series Plot



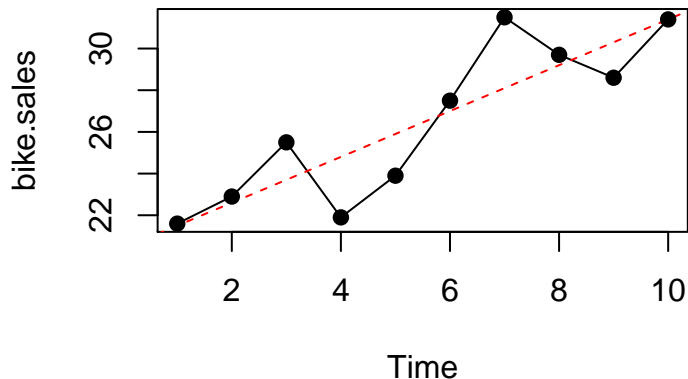
Visual inspection of the time series plot shows some up and down movement over the past 10 years, but the time series seems to have a systematically increasing upward trend.

The trend for the bicycle sales time series appears to be linear and increasing over time, but sometimes a trend can be described better by other types of patterns.



```
bike.sales=  
c(21.6,22.9,25.5,21.9,23.9,27.5,31.5,29.7,28.6,31.4);  
plot.ts(bike.sales, main="Bicycle Sales Time Series Plot");  
year=seq(1,10,by=1);  
trend=lm(bike.sales~year);  
abline(trend$coeff,lty=2,col="red");  
points(year,bike.sales,pch=19);
```

## Bicycle Sales Time Series Plot



# Seasonal Pattern

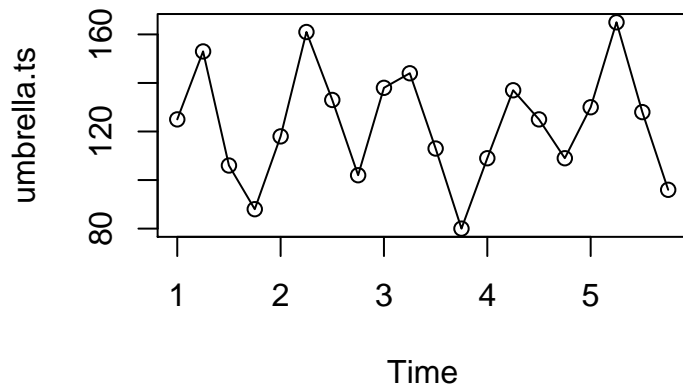
The trend of a time series can be identified by analyzing multiyear movements in historical data. Seasonal patterns are recognized by seeing the same repeating patterns over successive periods of time. For example, a manufacturer of snow removal equipment expects low sales activity in the spring and summer months, with peak sales in the fall and winter months. Not surprisingly, the pattern for a time series plot that exhibits a repeating pattern over a one-year period due to season influences is called a **seasonal pattern**. While we generally think of seasonal movement in a time series as occurring within one year, time series data can also exhibit seasonal patterns of less than one year in duration. As an example of a seasonal pattern, consider the number of umbrellas sold at a clothing store over the past five years. Table 3 shows the time series and Figure 3 show the corresponding time series plot.

Table 3

Year	Quarter	Sales	Year	Quarter	Sales
1	1	125	4	1	109
1	2	153	4	2	137
1	3	106	4	3	125
1	4	88	4	4	109
2	1	118	5	1	130
2	2	161	5	2	165
2	3	133	5	3	128
2	4	102	5	4	96
3	1	138			
3	2	144			
3	3	113			
3	4	80			

```
umbrella.sales=  
c(125,153,106,88,118,161,133,102,138,144,113,80,  
109,137,125,109,130,165,128,96);  
umbrella.ts=ts(umbrella.sales,frequency=4,start=c(1,1));  
plot.ts(umbrella.ts, main="Umbrella Sales Time Series Plot");  
points(umbrella.ts);
```

## Umbrella Sales Time Series Plot



The time series plot does not indicate any long-term trend in sales. In fact, unless you look carefully at the data, you might conclude that the data follow a horizontal pattern. But closer inspection of the time series plot reveals a regular pattern in the data. That is, the first and third quarters have moderate sales, the second quarter has the highest sales, and the fourth quarter tends to have the lowest sales volume. Thus, we would conclude that a quarterly seasonal pattern is present.

# Trend and Seasonal Pattern

Some time series include a combination of a trend and seasonal pattern. For instance, the data in Table 4 and the corresponding time series plot in Figure 4 show television sales for a particular manufacturer over the past four years. Clearly, an increasing trend is present. But Figure 4 also indicates that sales are lowest in the second quarter of each year and increase in quarters 3 and 4. Thus, we conclude that a seasonal pattern also exists for television sales. In such cases we need to use a forecasting method that has the capability to deal with both trend and seasonality.



Table 3

Year	Quarter	Sales (1000s)	Year	Quarter	Sales (1000s)
1	1	4.8	3	1	6.0
1	2	4.1	3	2	5.6
1	3	6.0	3	3	7.5
1	4	6.5	3	4	7.8
2	1	5.8	4	1	6.3
2	2	5.2	4	2	5.9
2	3	6.8	4	3	8.0
2	4	7.4	4	4	8.4

```
tv.sales=  
c(4.8,4.1,6.0,6.5,5.8,5.2,6.8,7.4,  
6.0,5.6,7.5,7.8,6.3,5.9,8.0,8.4);  
tv.ts=ts(tv.sales,frequency=4,start=c(1,1));  
plot.ts(tv.ts, main="Quarterly TV Sales Time Series Plot");  
points(tv.ts);
```

## Quarterly TV Sales Time Series Plot

