

Economic Forecasting

Time series data

Sebastian Fossati

University of Alberta | E493 | 2023


- 1 Time series data
- 2 Time plots
- 3 Seasonal plots
- 4 Seasonal or cyclic?
- 5 Autocorrelation
- 6 White noise
- 7 Your turn




Examples

- daily IBM stock prices
- monthly rainfall
- quarterly Australian beer production
- annual Google profits

Where can we find time series data?

 Statistics Canada / Statistique Canada


Search website 



[Subjects](#) [Data](#) [Analysis](#) [Reference](#) [Geography](#) [Census program](#) [Surveys and statistical programs](#) [About StatCan](#) [Canada.ca](#)

[Home](#) → Data


Data

[What's new on our website](#)

Filter results by 

Sort by  **Apply** 

Keyword(s)



- [Search by vector](#)
- [How to access microdata](#)
- [Customized products and services](#)

Statistics Canada



Where can we find time series data?

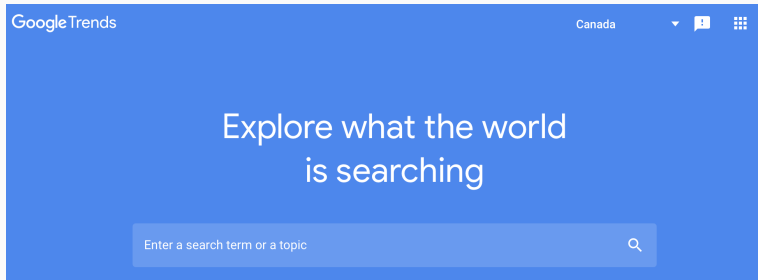


The screenshot shows the FRED website header and main content area. The header includes the FRED logo (Economic Data | St. Louis Fed), the text 'ECONOMIC RESEARCH' and 'FEDERAL RESERVE BANK OF ST. LOUIS', a search bar labeled 'Search FRED', and links for 'REGISTER | SIGN IN'. Below the header is a navigation bar with links: 'FRED® Economic Data', 'Information Services', 'Publications', 'Working Papers', 'Economists', 'About', and 'St. Louis Fed Home'. The main content area features the text 'Download, graph, and track **509,000** US and international time series from **87** sources.' followed by a search input field with the placeholder text 'Search FRED data e.g., gdp, inflation, unemployment' and a search button. Below the search field are links to 'Browse data by Tag, Category, Release, Source, Release Calendar or Get Help'.



FRED

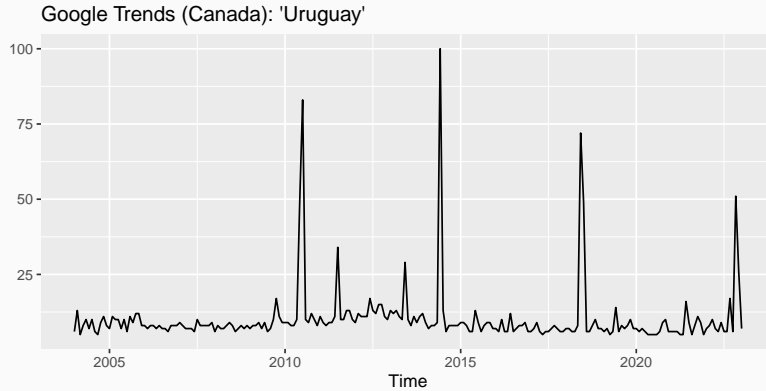
Where can we find time series data?



Google Trends



Your turn



ts objects and ts function

A time series is stored in a ts object in R:

- a list of numbers
- information about times those numbers were recorded



Year	Observation
2012	123
2013	39
2014	78
2015	52

```
y <- ts(c(123,39,78,52), start = 2012)
```


For observations that are more frequent than once per year, add a frequency argument.

E.g., monthly data stored as a numerical vector z:

```
y <- ts(z, freq = 12, start = c(2003,1))
```



Time series object: `ts(data, frequency, start)`



Type of data	frequency	start example
--------------	-----------	---------------

Annual		
--------	--	--

Quarterly		
-----------	--	--

Monthly		
---------	--	--

Weekly		
--------	--	--

...		
-----	--	--

ts objects and ts function

Time series object: `ts(data, frequency, start)`

Type of data	frequency	start example
Annual	1	1995
Quarterly		
Monthly		
Weekly		
...		

ts objects and ts function

Time series object: `ts(data, frequency, start)`

Type of data	frequency	start example
Annual	1	1995
Quarterly	4	c(1995,2)
Monthly		
Weekly		
...		

ts objects and ts function

Time series object: `ts(data, frequency, start)`

Type of data	frequency	start example
Annual	1	1995
Quarterly	4	c(1995,2)
Monthly	12	c(1995,9)
Weekly		
...		

ts objects and ts function

Time series object: `ts(data, frequency, start)`

Type of data	frequency	start example
Annual	1	1995
Quarterly	4	c(1995,2)
Monthly	12	c(1995,9)
Weekly	52.18	c(1995,23)
...		



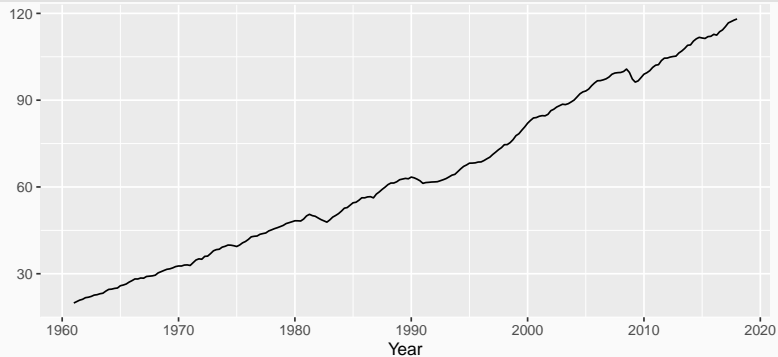


```
# read canada quarterly real gdp data
data <- read.csv("data/NAEXKP01CAQ661S.csv", header = TRUE)
cangdp <- ts(data$NAEXKP01CAQ661S, start = 1961, freq = 4)
# display some observations
window(cangdp, start = c(2001,1), end = c(2005,4))
```

```
##           Qtr1  Qtr2  Qtr3  Qtr4
## 2001 84.43 84.65 84.59 85.11
## 2002 86.37 86.87 87.62 88.11
## 2003 88.60 88.46 88.80 89.41
## 2004 90.05 91.12 92.20 92.86
## 2005 93.18 93.84 95.00 95.94
```

Canadian Real GDP

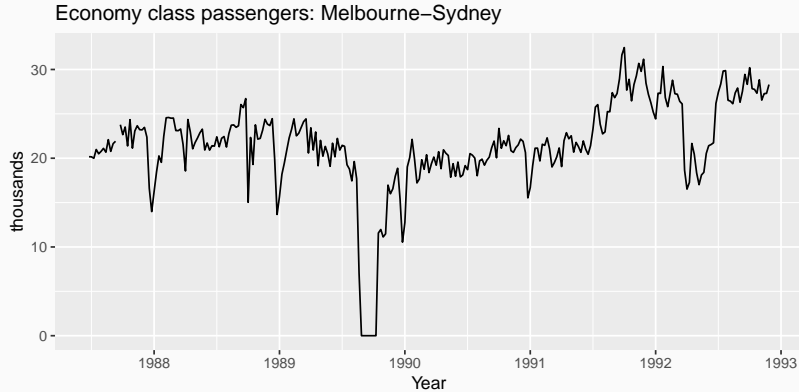
```
# time series plot  
autoplot(cangdp) + xlab("Year") + ylab(" ")
```



Update the Canadian real GDP time series to include the latest observations available (you should be able to get data up to 2022Q3).

- 1 Time series data
- 2 Time plots
- 3 Seasonal plots
- 4 Seasonal or cyclic?
- 5 Autocorrelation
- 6 White noise
- 7 Your turn

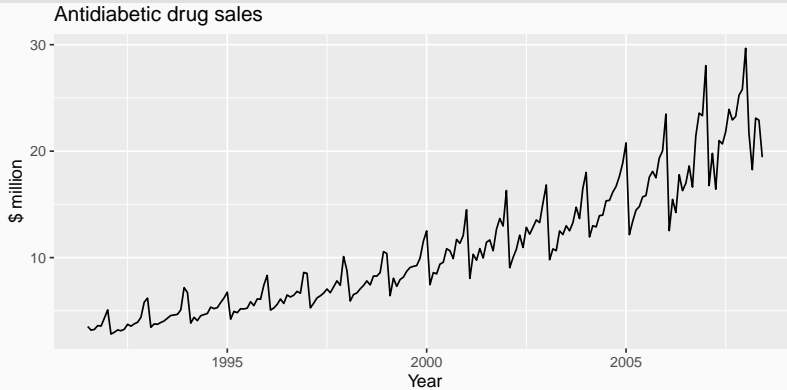
Time plots



Time plots

```
# drug sales
```

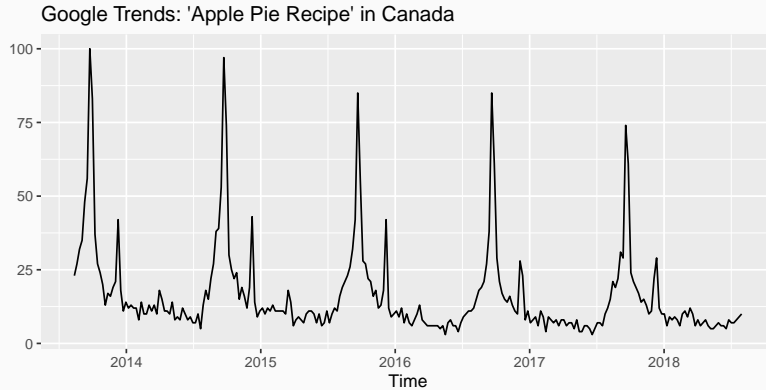
```
autoplot(a10) + ylab("$ million") + xlab("Year") +  
  ggtitle("Antidiabetic drug sales")
```



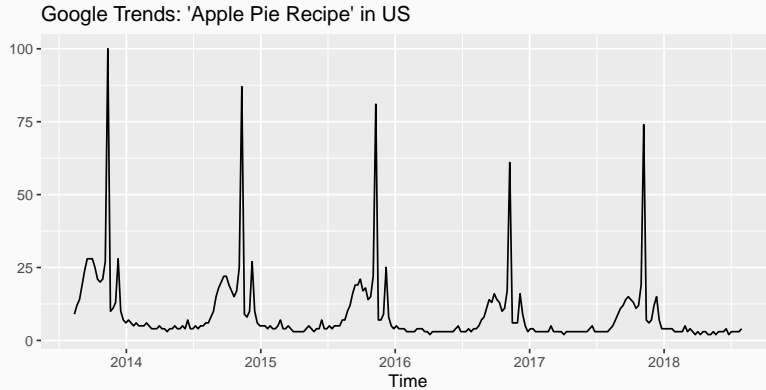
- 1 Time series data
- 2 Time plots
- 3 Seasonal plots**
- 4 Seasonal or cyclic?
- 5 Autocorrelation
- 6 White noise
- 7 Your turn



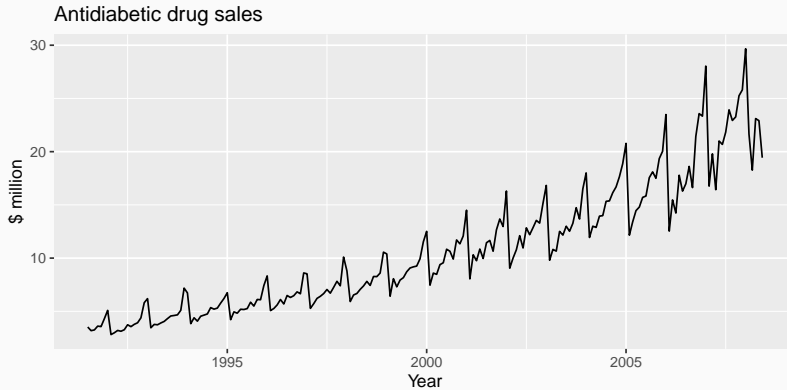
Seasonal data



Seasonal data



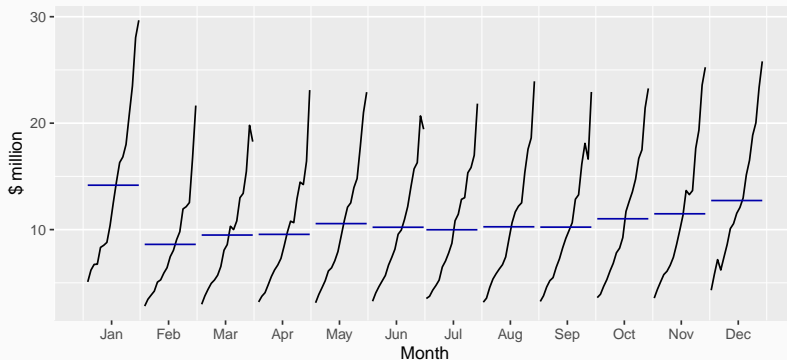
Seasonal data



Seasonal subseries plots

```
# seasonal plot  
ggsubseriesplot(a10) + ylab("$ million") +  
  ggtitle("Subseries plot: antidiabetic drug sales")
```

Subseries plot: antidiabetic drug sales



Remarks:

- data for each season collected together in time plot as separate time series
- enables the underlying seasonal pattern to be seen clearly, and changes in seasonality over time to be visualized
- in R: `ggsubseriesplot()`
- see also `ggseasonplot()`

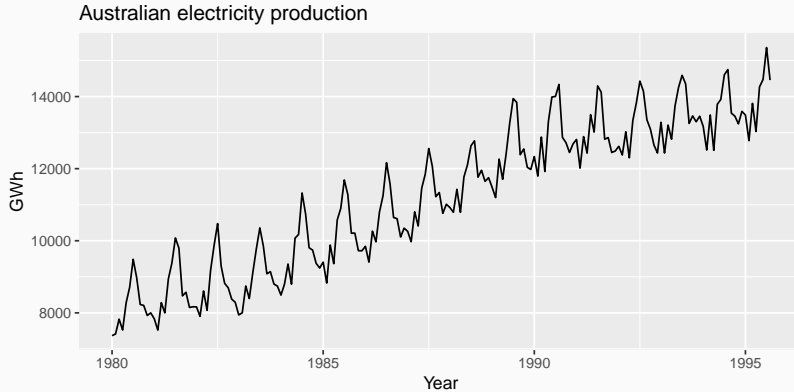
- 1 Time series data
- 2 Time plots
- 3 Seasonal plots
- 4 Seasonal or cyclic?**
- 5 Autocorrelation
- 6 White noise
- 7 Your turn

Trend: pattern exists when there is a long-term increase or decrease in the data

Seasonal: pattern exists when a series is influenced by seasonal factors (eg, the quarter of the year, the month, or day of the week)

Cyclic: pattern exists when data exhibit rises and falls that are *not of fixed period* (duration usually of at least 2 years)

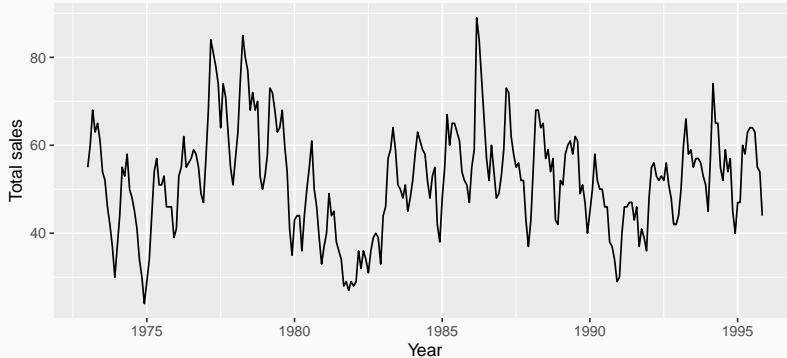
Time series patterns



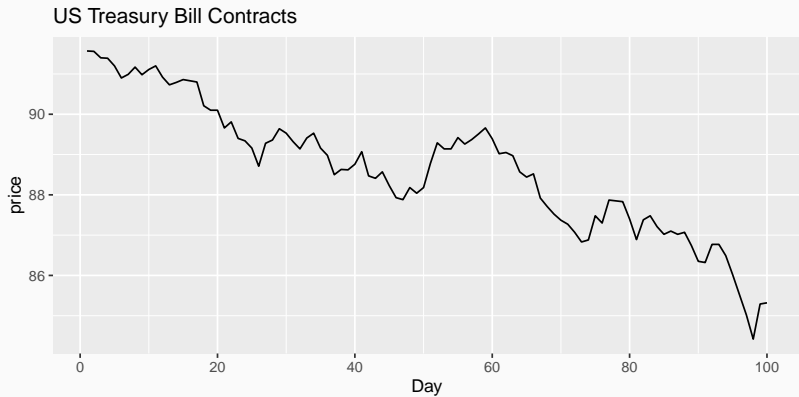
Time series patterns



Sales of new one-family houses, USA



Time series patterns



Seasonal or cyclic?

Differences between seasonal and cyclic patterns:

- seasonal pattern constant length; cyclic pattern variable length
- average length of cycle longer than length of seasonal pattern
- magnitude of cycle more variable than magnitude of seasonal pattern



Seasonal or cyclic?

Differences between seasonal and cyclic patterns:

- seasonal pattern constant length; cyclic pattern variable length
- average length of cycle longer than length of seasonal pattern
- magnitude of cycle more variable than magnitude of seasonal pattern

The timing of peaks and troughs is predictable with seasonal data, but unpredictable in the long term with cyclic data.

- 1 Time series data
- 2 Time plots
- 3 Seasonal plots
- 4 Seasonal or cyclic?
- 5 Autocorrelation**
- 6 White noise
- 7 Your turn

Covariance and **correlation**: measure extent of **linear relationship** between two variables (y and x).

Autocovariance and **autocorrelation**: measure linear relationship between **lagged values** of a time series y .

We measure the relationship between:

- y_t and y_{t-1}
- y_t and y_{t-2}
- y_t and y_{t-3}
- etc.



We denote the sample autocovariance at lag k by c_k and the sample autocorrelation at lag k by r_k . Define

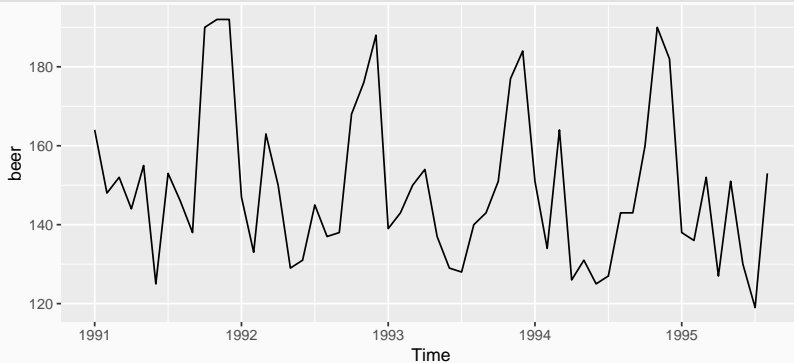
$$c_k = \frac{1}{T} \sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y})$$

and $r_k = c_k / c_0$

- r_1 indicates how successive values of y relate to each other
- r_2 indicates how values two periods apart relate to each other
- etc.

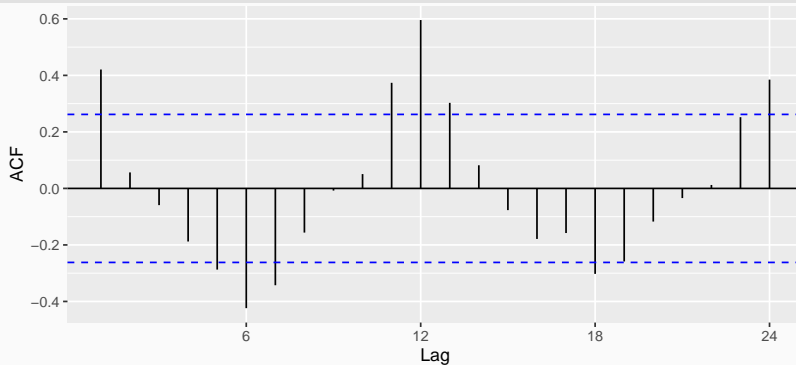
Autocorrelation

```
# plor beer data  
autoplot(beer)
```



Autocorrelation

```
# plot autocorrelations  
ggAcf(beer) + ggtitle("")
```



Remarks:

- r_4 is higher than for the other lags due to **the seasonal pattern in the data**: the peaks tend to be **4 quarters** apart
- r_2 is more negative than for the other lags because troughs tend to be 2 quarters behind peaks
- together, the autocorrelations at lags 1, 2, ..., make up the *autocorrelation* or ACF
- and the plot is known as a **correlogram**

Remarks:

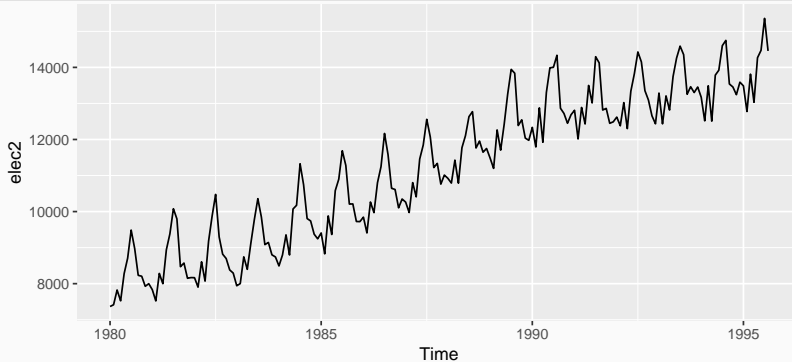
- when data have a trend, the autocorrelations for small lags tend to be large and positive
- when data are seasonal, the autocorrelations will be larger at the seasonal lags (ie, at multiples of the seasonal frequency)
- when data are trended and seasonal, you see a combination of these effects

Aus monthly electricity production

```
# electricity data
```

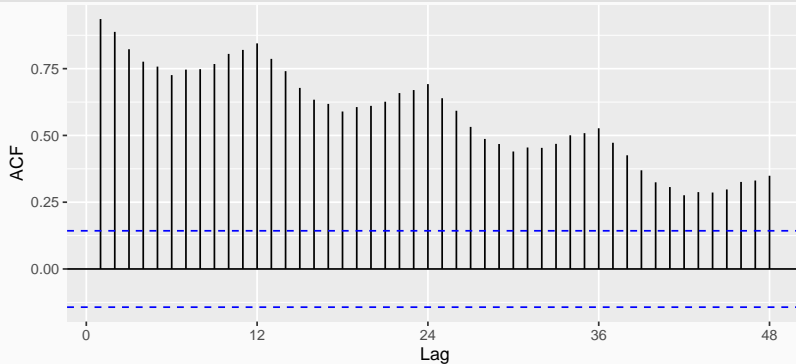
```
elec2 <- window(elec, start = 1980)
```

```
autoplot(elec2)
```



Aus monthly electricity production

```
# plot autocorrelations  
ggAcf(elec2, lag.max = 48) + ggtitle("")
```



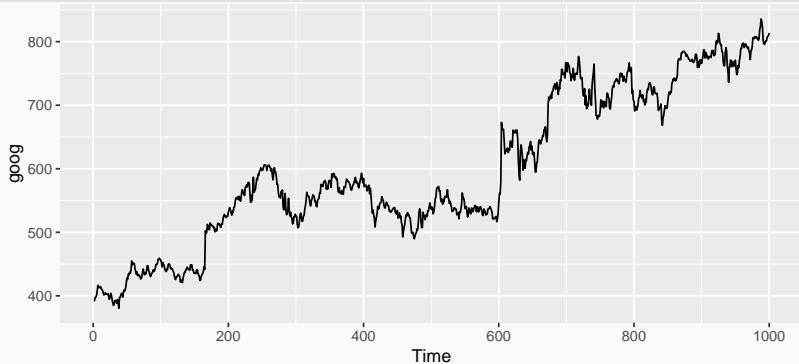
Time plot shows clear trend and seasonality.

The same features are reflected in the ACF.

- the slowly decaying ACF indicates trend
- the ACF peaks at lags 12, 24, 36, ..., indicate seasonality of length 12

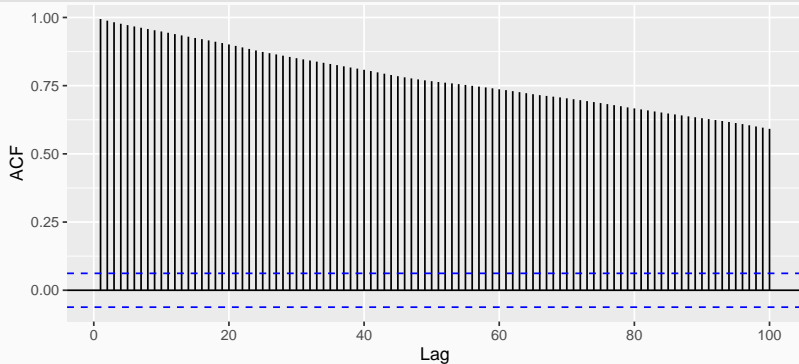
Google stock price

```
# plot data  
autoplot(goog)
```



Google stock price

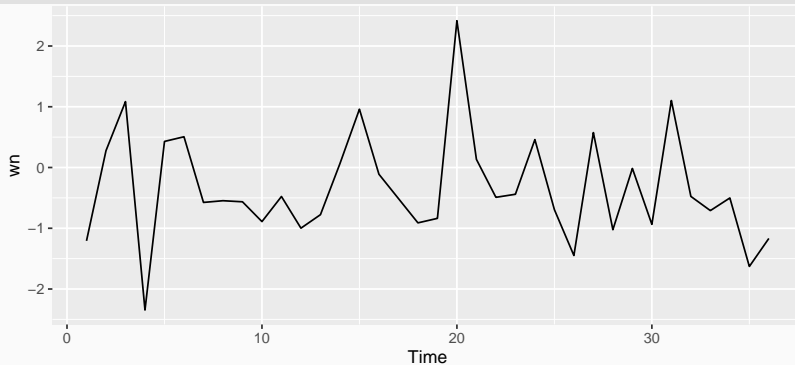
```
# plot autocorrelations  
ggAcf(goog, lag.max = 100) + ggtitle("")
```



- 1 Time series data
- 2 Time plots
- 3 Seasonal plots
- 4 Seasonal or cyclic?
- 5 Autocorrelation
- 6 White noise**
- 7 Your turn

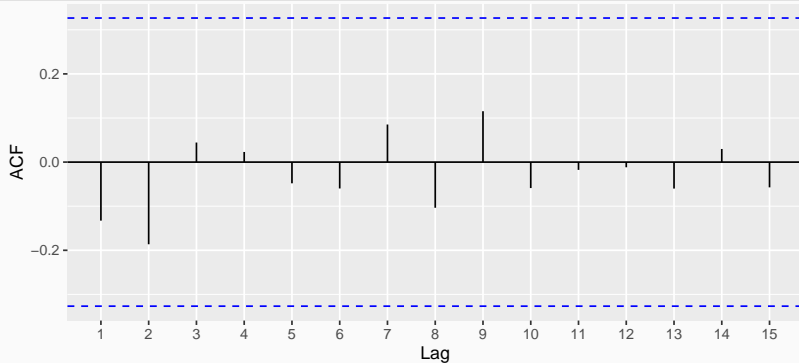
Example: White noise

```
# simulate white noise process  
wn <- ts(rnorm(36))  
autoplot(wn)
```



Example: White noise

```
# plot autocorrelations  
ggAcf(wn) + ggtitle("")
```

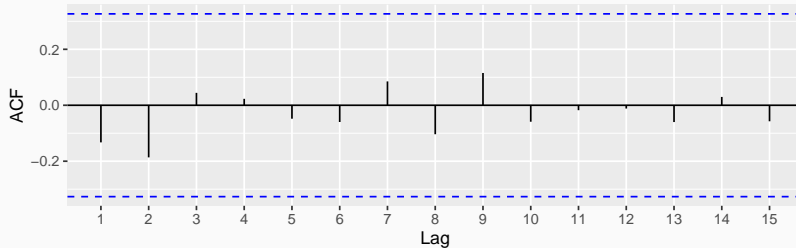


Sampling distribution of autocorrelations

Sampling distribution of r_k for white noise data is asymptotically $N(0, 1/T)$.

- 95% of all r_k for white noise should lie within $\pm 1.96/\sqrt{T}$
- if this is not the case, the series may not WN
- common to plot lines at $\pm 1.96/\sqrt{T}$ when plotting ACF (these are the *critical values*)

Autocorrelation



Example:

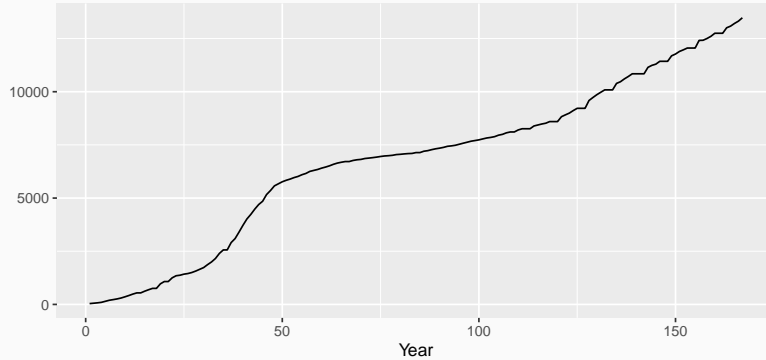
$T = 36$ and so critical values at $\pm 1.96 / \sqrt{36} = \pm 0.327$.

All autocorrelation coefficients lie within these limits, the data appear to be white noise.

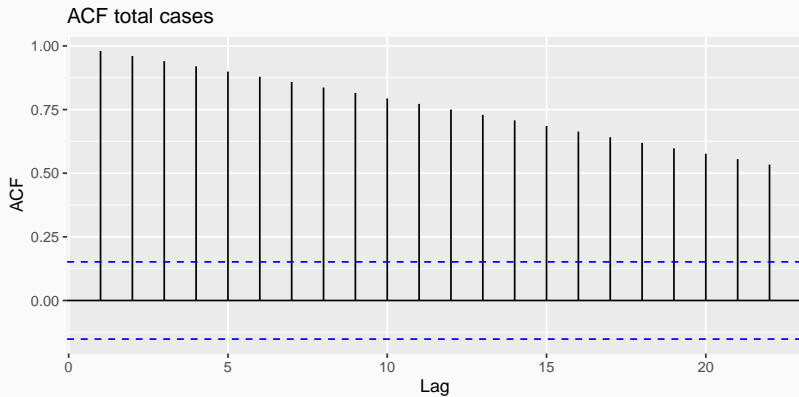
- 1 Time series data
- 2 Time plots
- 3 Seasonal plots
- 4 Seasonal or cyclic?
- 5 Autocorrelation
- 6 White noise
- 7 Your turn**

COVID-19 in Alberta

Total cases of COVID-19 in Alberta since March 15, 2020

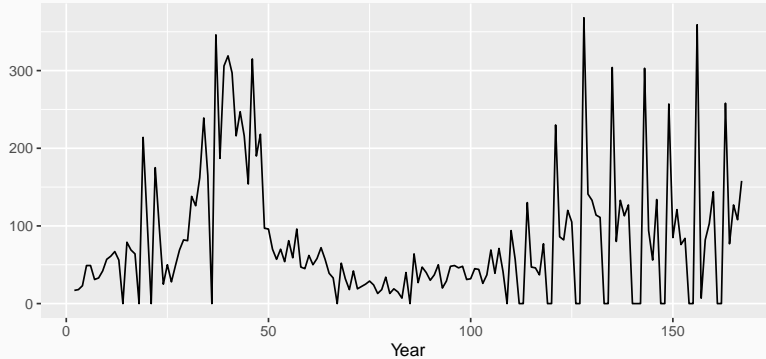


COVID-19 in Alberta



COVID-19 in Alberta

New cases of COVID-19 in Alberta since March 15, 2020



COVID-19 in Alberta

