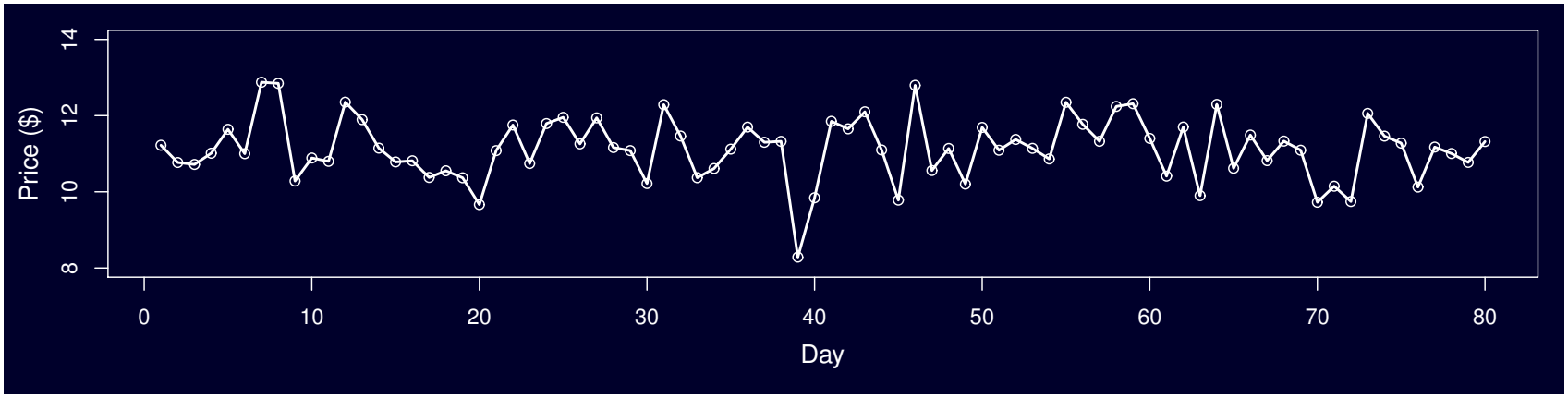
Introduction to Time Series

Week I: Video 2

STAT 485/685, Fall 2020, SFU

Sonja Isberg

What Is a Time Series?

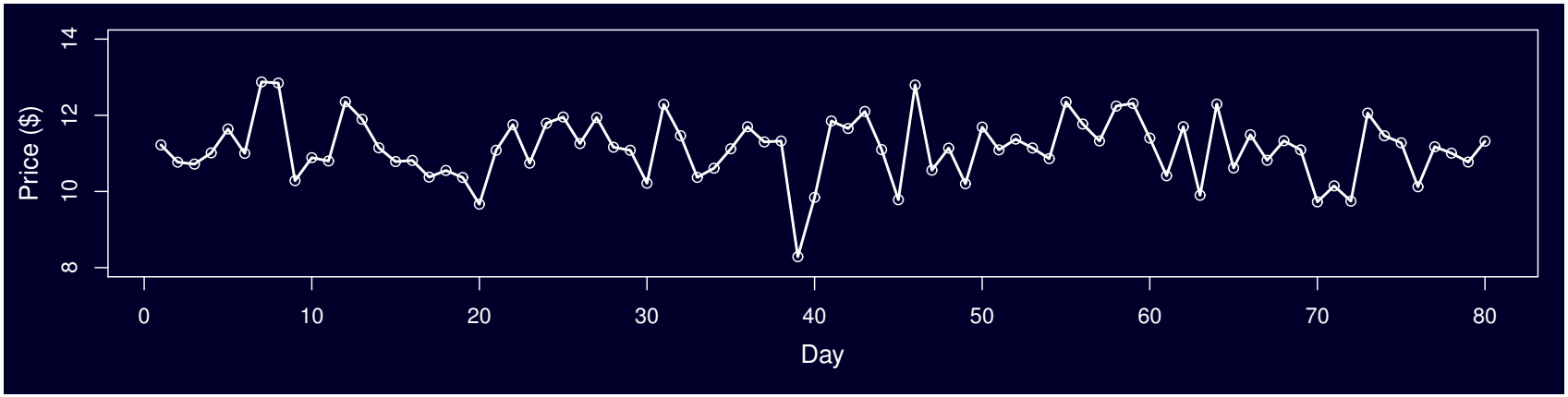


A **time series** is data obtained from observations collected sequentially over time (Cryer & Chan).

This type of data is very common in many fields of study! For example:

- Stock prices over time
- Daily high temperatures in Greenland over the last 50 years
- Number of cases of an infectious disease each month
- Median house prices in San Francisco over the last decade
- Abundance of an animal species over time

What Is a Time Series? (cont'd)

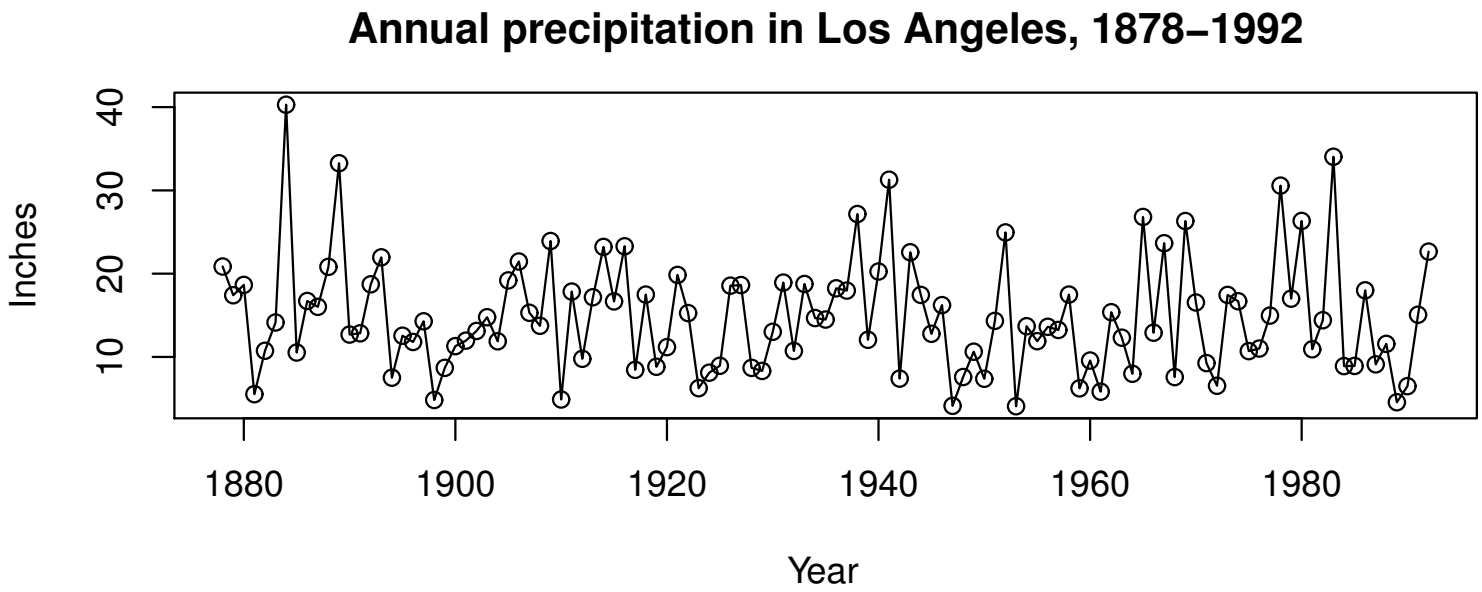


Depending on the problem, some questions we may want to ask are:

- Is there a general trend of the variable increasing/decreasing over time?
- Is there a seasonal effect? (e.g., Do values cycle between high and low at certain intervals?)
- Can we predict what the values will be for future observations?
- Does the value of the variable for one year tell us anything about the value of the variable for the next year?

Examples of Time Series

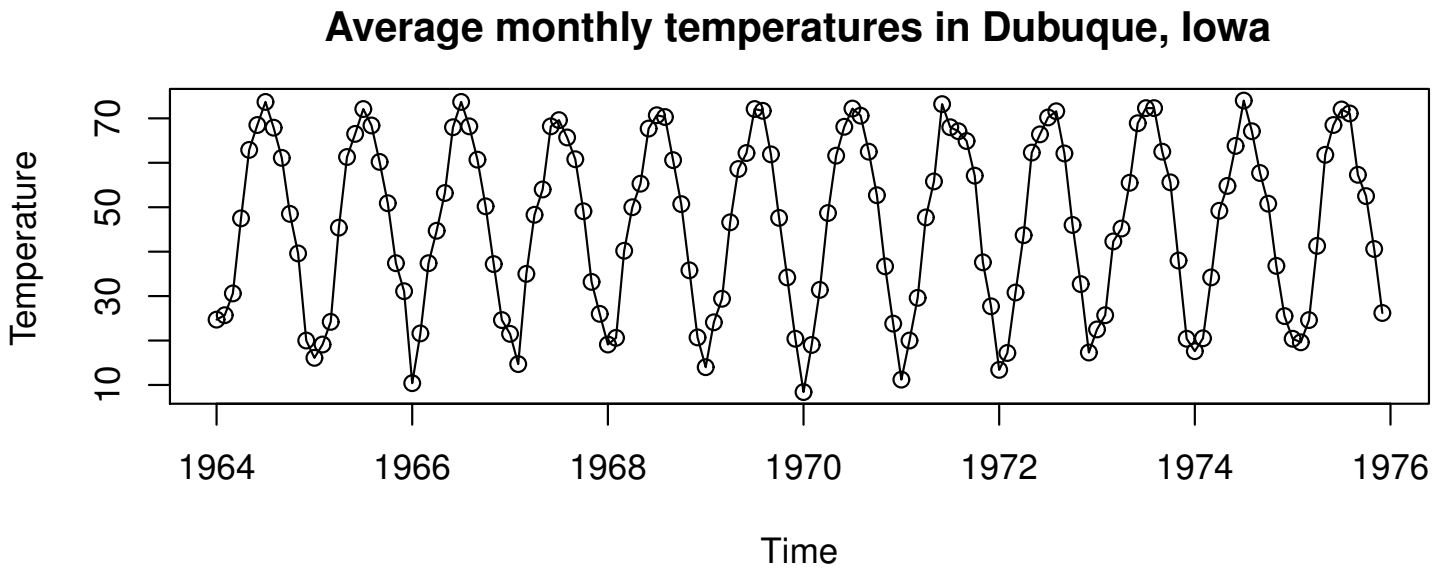
From the dataset "larain" in the TSA package[1]:



Is there any sort of pattern visible here?

Examples of Time Series

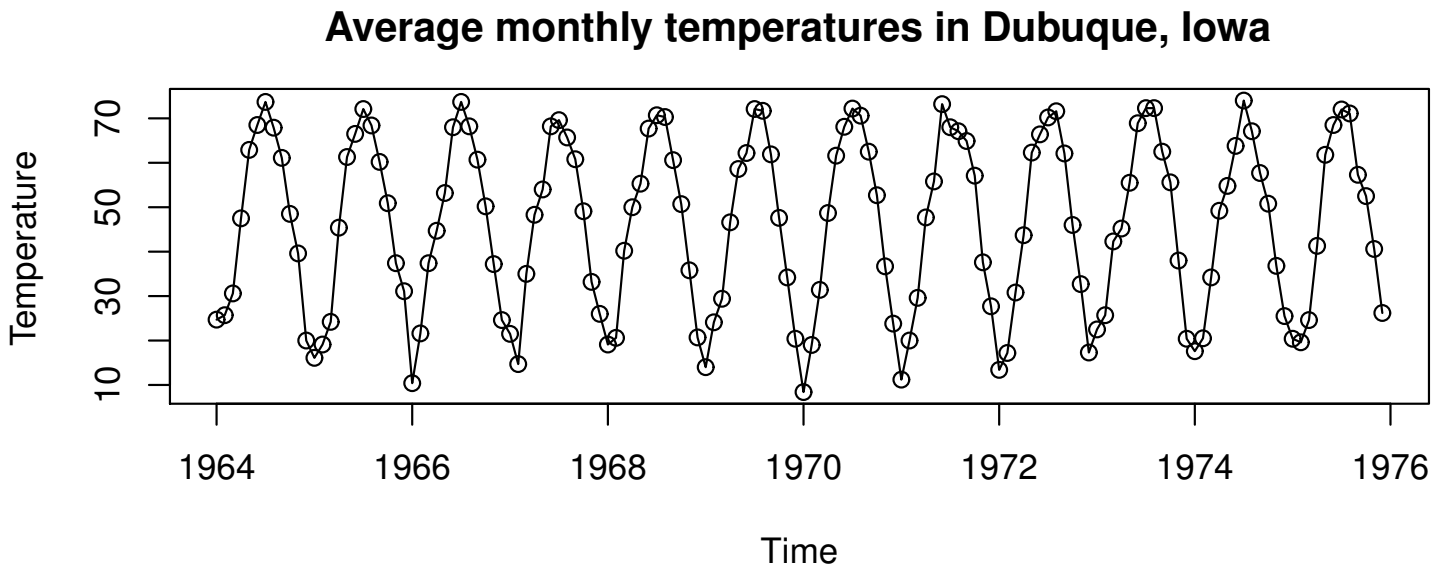
From the dataset "tempdub" in the TSA package[1]:



What pattern do you see here?

Examples of Time Series

From the dataset "tempdub" in the TSA package[1]:

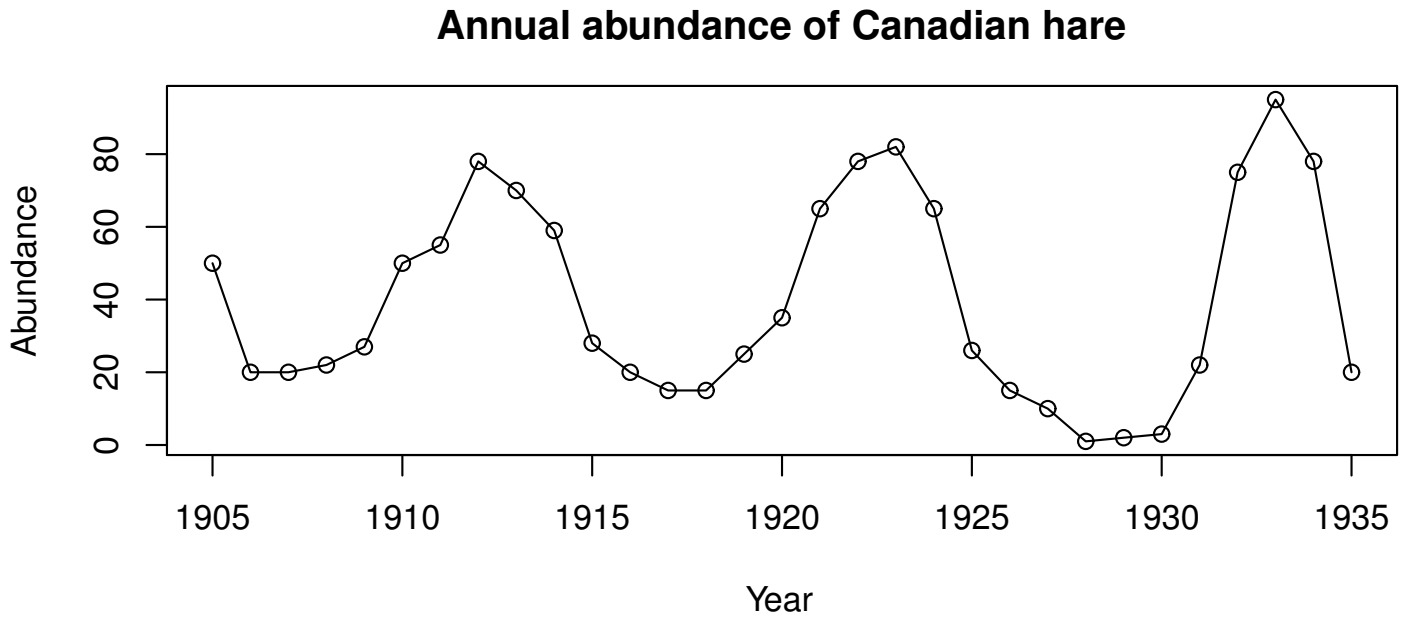


What pattern do you see here?

When observations a certain number of units apart are related in some way, there is said to be **seasonality** in the data. In this example, the seasonality is in monthly values – observations 12 months apart are closely related.

Examples of Time Series

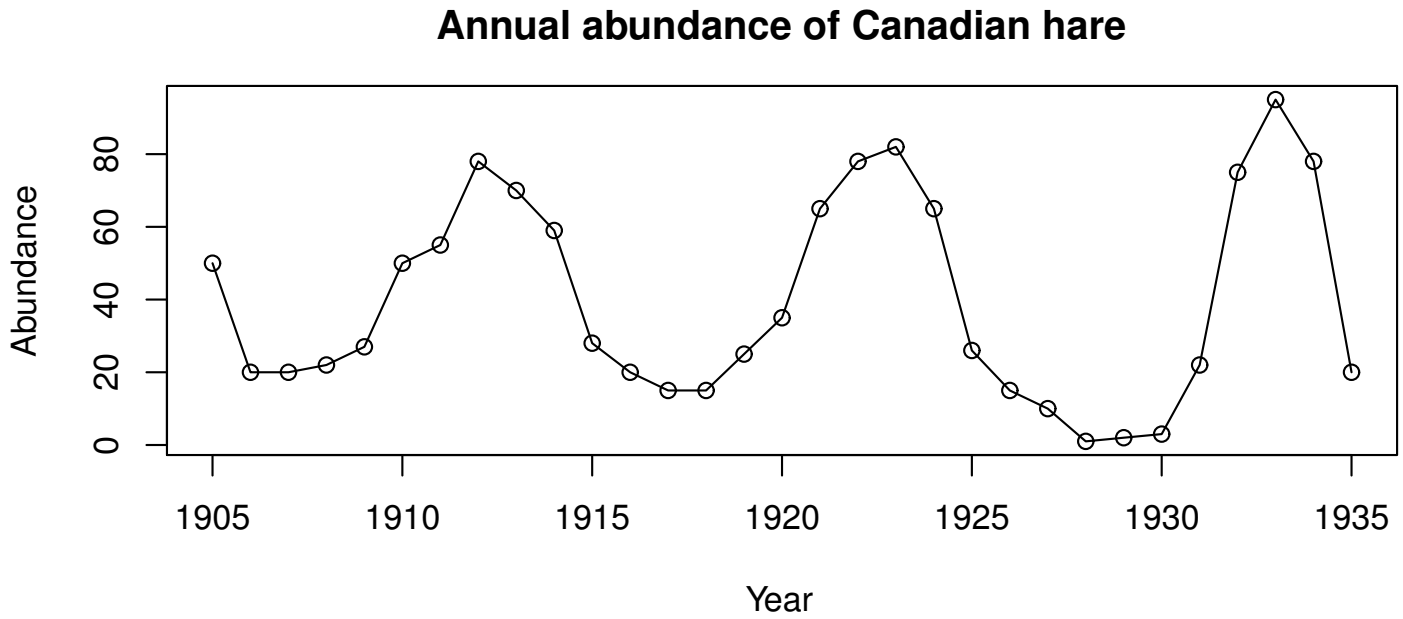
From the dataset "hare" in the TSA package[1]:



What do you see here?

Examples of Time Series

From the dataset "hare" in the TSA package[1]:



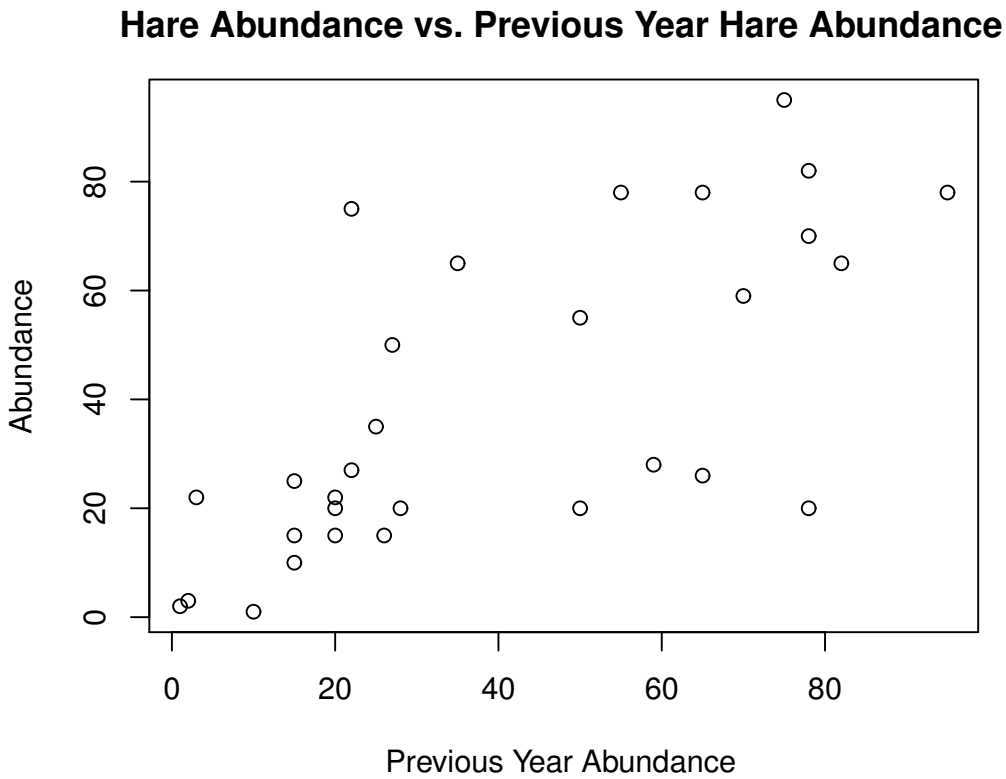
What do you see here?

Neighbouring values are very closely related. The hare abundance for any given year gives you a pretty good idea of what the hare abundance will be the following year.

In other words, hare abundance is highly **positively correlated** between consecutive years.

Hare Example Continued

One way we can investigate this correlation is by plotting each year's hare abundance against the previous year's hare abundance:



We see a moderate positive correlation between the values, confirming our earlier observation that consecutive values are related.

Why Is Time Series Analysis Interesting?

In each of the above plots, we have some number of observations n over time. Whey can't we just fit a linear regression, or a similar statistical model?

Why Is Time Series Analysis Interesting?

In each of the above plots, we have some number of observations n over time. Whey can't we just fit a linear regression, or a similar statistical model?

One of the most common assumptions for the usual statistical models is that the observations must be *independent*.

- Since the observations come from repeated measurements of the same variable over time, they *cannot* be assumed to be independent!

So, we need some way to model the variable Y_t over time t, taking this dependence into account.

What Will We Learn?

1. **Key Ideas:**

   - **Intro and fundamental concepts (Ch. 1-2):** Means, variances and correlations of time series, and the concept of stationarity.
   - **Estimating trends (Ch. 3):** Temporarily ignoring the rest of the variability in the series, how do we just estimate the mean trend? Can a regression be useful here?
   - **Models for stationary time series (Ch. 4):** These are the bulk of the key time series models that will be useful to us: MA, AR and ARMA models.
   - **Models for non-stationary time series (Ch. 5):** How do we put together the ideas in Ch. 3 & 4, to create a very general model?

What Will We Learn? (cont'd)

2. **Building a Model:**
   - **Model specification (Ch. 6):** How do we choose between the different models that we know?
   - **Parameter estimation (Ch. 7):** Now that we've chosen a model, there will be parameters whose values are unknown. How do we estimate these parameter values?
   - **Model diagnostics (Ch. 8):** How good is our chosen model? Should we be using a different model?
3. **Forecasting (Ch. 9):** One of our main goals in building a model is to be able to forecast future values. This chapter will teach us how to do this.
4. Other topics, as time permits.

Final Comments

In this course, we will be covering both theory and methods of time series. Assignments will often be based on hands-on problems that need to be solved using the R package "TSA".

Installation of R, RStudio and the TSA package is covered in the Canvas course module "Week 1".

We won't be going over R code in the lecture videos. Examples will be covered in the weekly tutorials, and detailed code is also given in the textbook underneath each figure and table. Attending the tutorials will be useful for helping you complete the assignment questions.

Final Comments

In this course, we will be covering both theory and methods of time series. Assignments will often be based on hands-on problems that need to be solved using the R package "TSA".

Installation of R, RStudio and the TSA package is covered in the Canvas course module "Week 1".

We won't be going over R code in the lecture videos. Examples will be covered in the weekly tutorials, and detailed code is also given in the textbook underneath each figure and table. Attending the tutorials will be useful for helping you complete the assignment questions.

Please make sure to complete the brief quiz linked below the video.

Coming Up Next in STAT 485/685: Fundamental concepts for Time Series!

References

[1] Chan, K. S., & Ripley, B. (2020). TSA: Time Series Analysis. R package version 1.2.1.

[2] Cryer, J. D., & Chan, K. S. (2008). *Time series analysis: with applications in R.* Springer Science and Business Media.