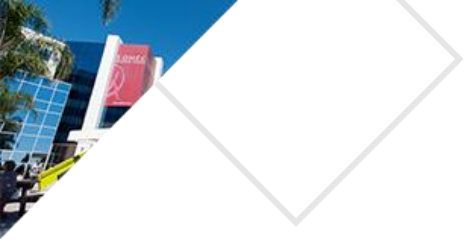
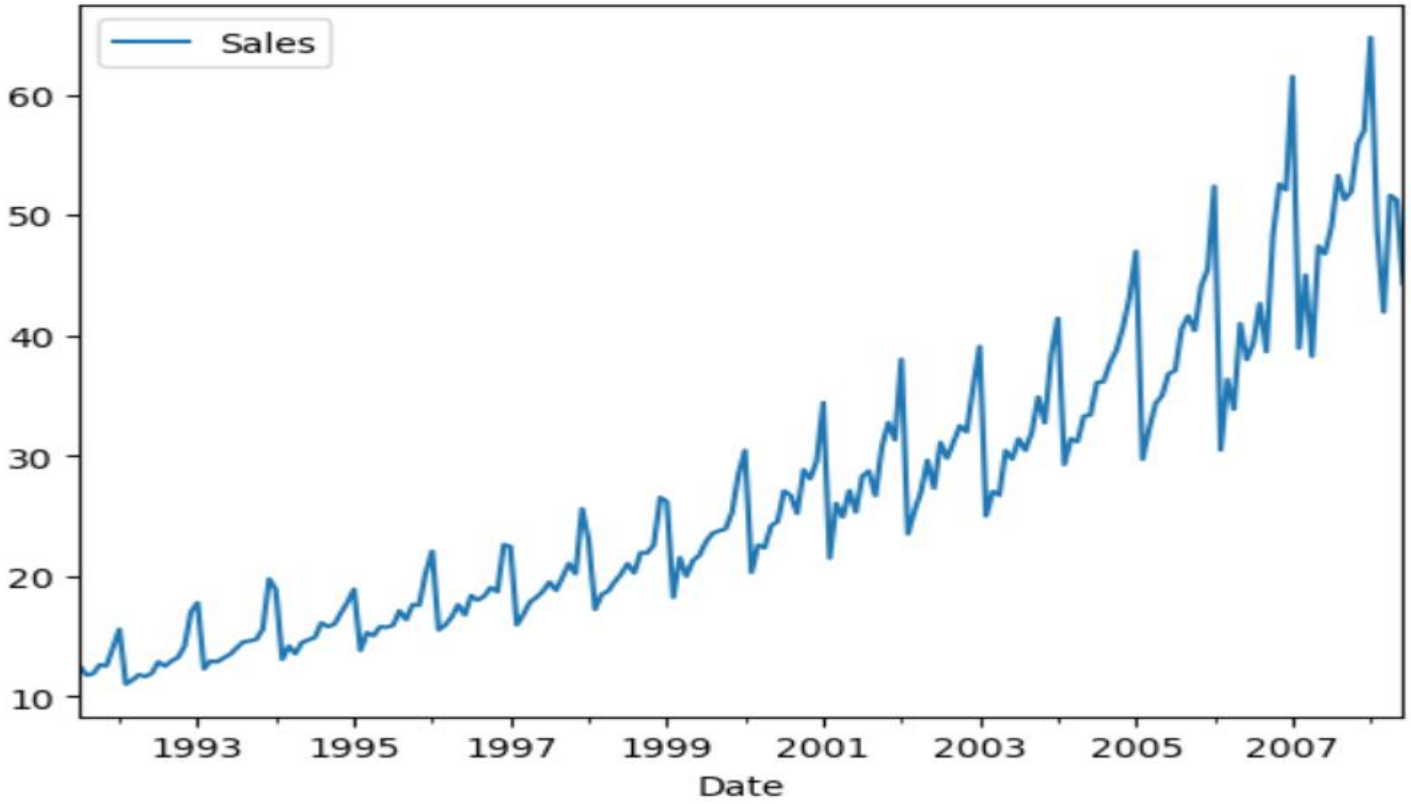


Time Series Data

WHAT IS A TIME SERIES?

It's a set of evenly spaced numerical data obtained by observing the response variable at regular time periods.



Period can be:

- Annual or longer
- Quarterly
- Monthly
- Daily
- ..
- Milisecond or less!

TIME SERIES VS CROSS-SECTIONAL DATA

Time in x-axis *X factor vs Y response*

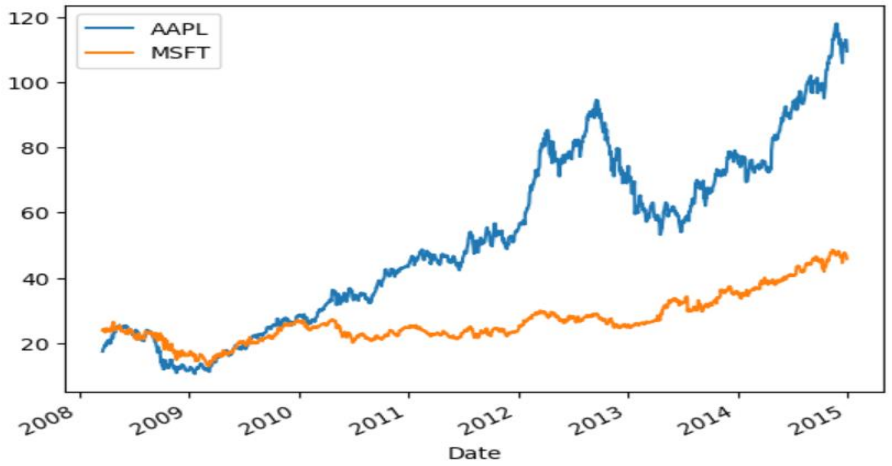
- **Time series data** is a temporal sequence of observations.
- It is desirable that the observation times are regular (same period).
- It allows to track time evolution of a variable of interest.

- **Cross-sectional data** is a collection of variables measured at the same time.
- There may by different measurements at different times, but no need to be regularly spaced (often the time measurement is not even recorded).
- It allows to see relationships between variables of interest.

- In cross-sectional data we typically **interpolate** (although we can extrapolate as well).
- In time series data we typically **extrapolate** beyond the last data point.

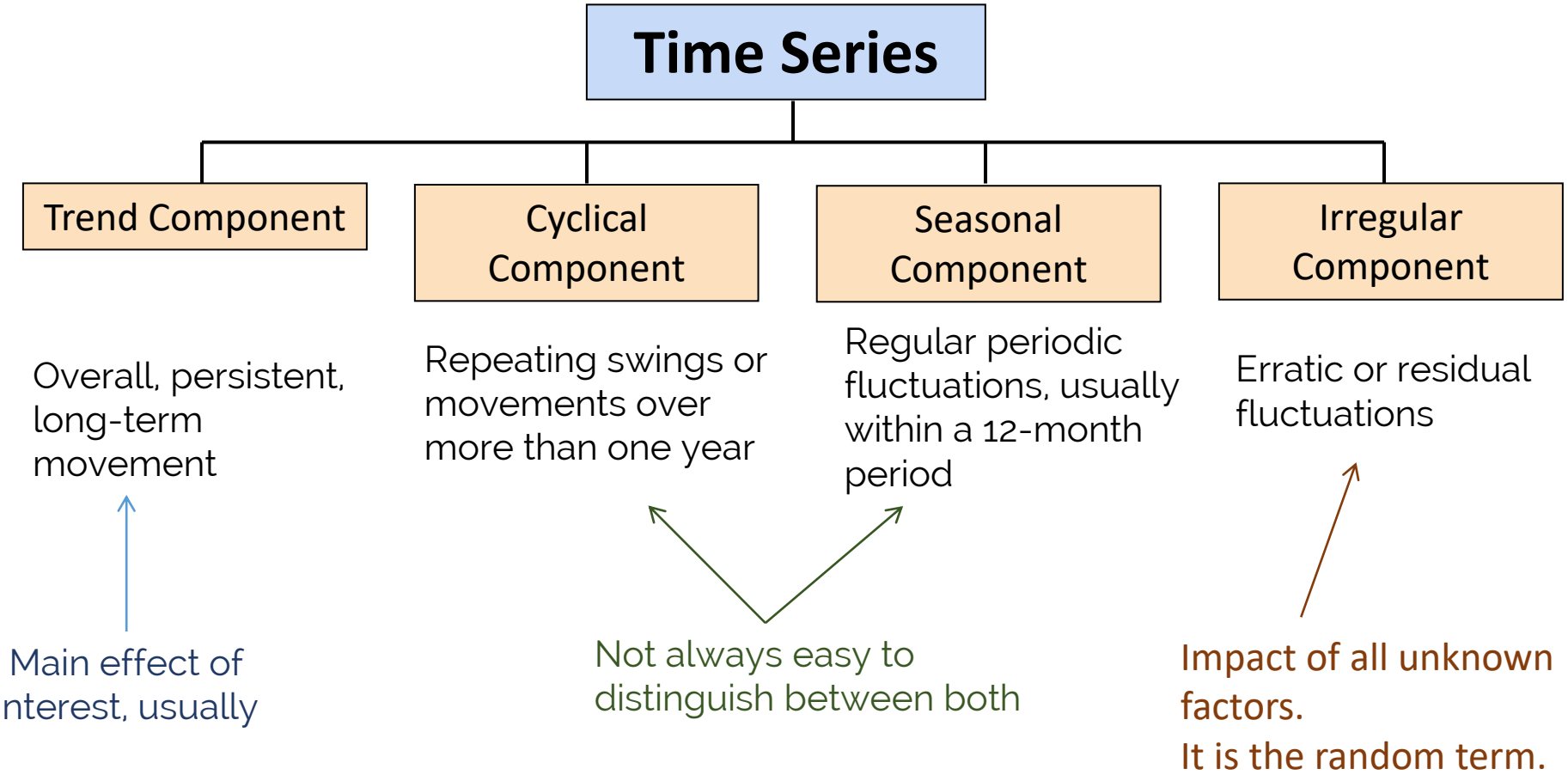
PANEL DATA

- Panel data combines time series and cross-sectional data.
- You have several time series, in which each variable has been measured at the same periods, so you can put them together to analyse the common temporal behaviour.
- It is very useful to investigate if there are external or common factors that affect more than one variable.



Date	AAPL	MSFT
2008-03-19	17.400475	23.840880
2008-03-20	17.883556	24.307368
2008-03-24	18.723591	24.299038
2008-03-25	18.918164	24.274047
2008-03-26	19.465665	23.790898
2008-03-27	18.820201	23.366061
2008-03-28	19.190571	23.249439
2008-03-31	19.256324	23.640956
2008-04-01	20.065493	24.573933
2008-04-02	19.791744	24.290708

TIME SERIES COMPONENTS (I)



TIME SERIES COMPONENTS (II)

*T: **Trend** component*, also called "Tendency"
This component expresses its **structural characteristic** of evolution. It indicates whether the trend of a series is increasing or decreasing. It can be linear or nonlinear.

*C: **Cyclical** component*
This component includes **cyclical macroeconomic phenomena**.

*S: **Seasonal** component*
This **periodic** component, more or less regular, is linked to **seasonal influences**.

*R: **Residual** or "random" component*

This component includes everything that has **not been taken into account by the components T, C and S**, that is to say: background noise, accidental variations and new information.

How do all components interact?

These components are assumed to follow either an **additive** or a **multiplicative** model.

If the model is **additive**:

- All the components addition directly in the observed variable.
- The data can be expressed as follows:

$$X=T+C+S+R$$

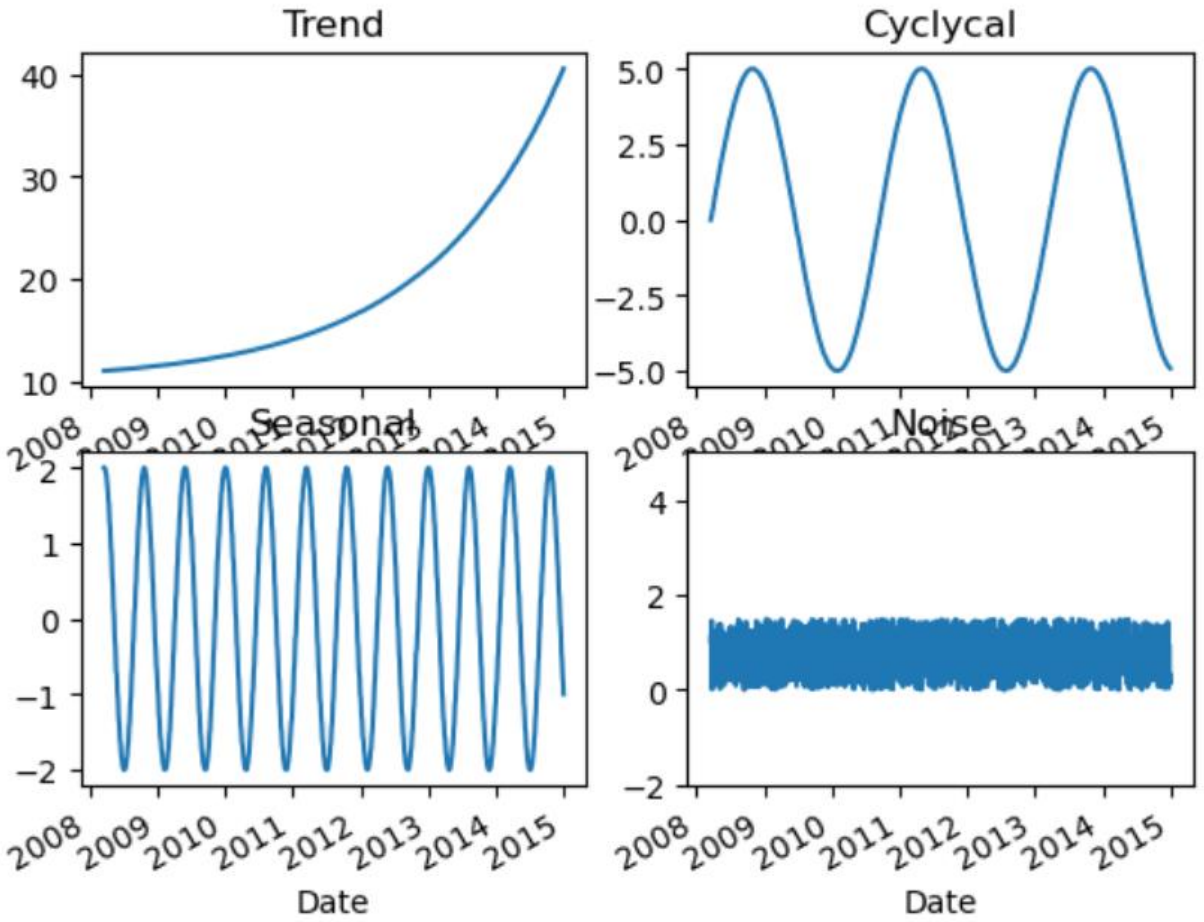
If the model is **multiplicative:**

- The components may addition to the rate of growth of the observed variable.
- The data can be expressed as follows:

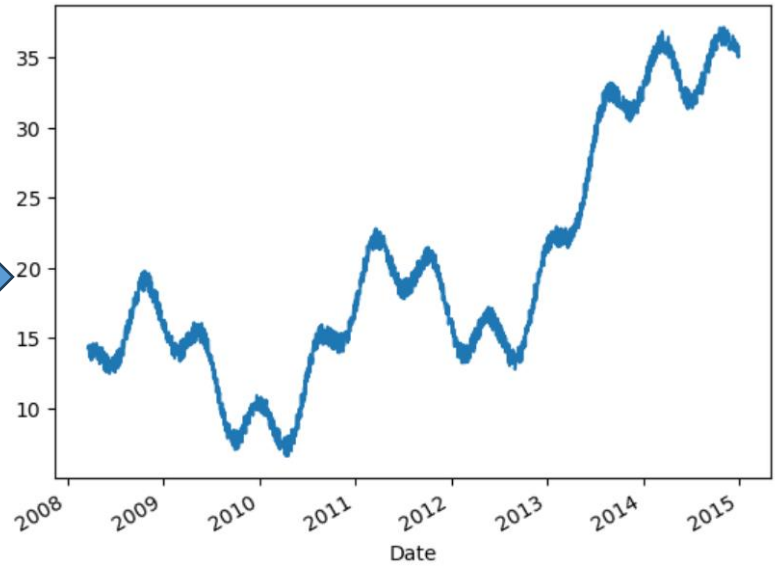
$$Y = T \times C \times S + R \text{ (partial multiplicative)}$$

$$Y = T \times C \times S \times R \text{ (full multiplicative)}$$

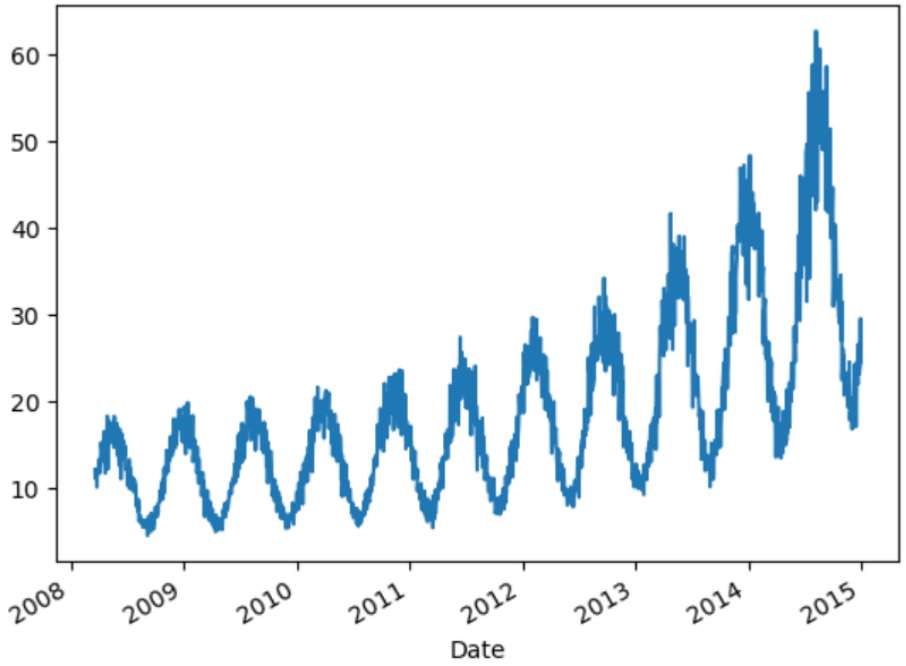
EXAMPLE 1: ADDITIVE MODEL



Additive model



EXAMPLE 2: MULTIPLICATIVE MODEL



Can you see why this could be a multiplicative model?

In the next section we'll use Python to separate components of a time series.