An Introduction to Forecasting: Chapter 1

Dr. Esam Mahdi

Qatar University - Department of Mathematics, Statistics, and Physics

December 23, 2020

Learning objectives

- The difference between cross-sectional and time series data.
- The components of a time series.
- Introduce the two different kinds of forecasting methods- Qualitative and Quantitative.
- The difference between point forecasts and prediction interval forecasts.
- Explain how to measure the forecasting errors:
  - Absolute deviation (AD).
  - Mean absolute deviation (MAD).
  - Squared error (SE).
  - Mean squared error (MSE).
  - Absolute percentage errors (APE_t).
  - Mean absolute percentage errors (MAPE).
- Use statistical software R with some applications.

What is a time series?

Definition

A time series is a chronological sequence of observations on a particular variable.

- A collection of data $\{Y_t\}$ recorded over a period of time (daily, weekly, monthly, quarterly, yearly, etc.), analyzed to understand the past, in order to predict the future (forecast), helping managers and policy makers to make well-informed and sound decisions.
- An important feature of most time series is that observations close together in time tend to be correlated (serially dependent).
- Time could be discrete: $t = 1, 2, ...,$ or continuous: $t \in (a, b)$ where $a, b > 0$.

- Unit sales of product over time.
- Total dollars sales for a company over time.
- Number of employed over time.
- Unemployment rate over time.
- Air or water quality over time.
- Population of a city over time.
- Daily temperature over time.

Table: Time series data: Quarterly values of time deposits

Year	Quarter	Value of time deposits (in millions dollars)
2001	1	35.3
	2	37.6
	3	35.1
	4	38.5
2002	1	37.4
	2	36.8
	3	32.9
	4	41.1
⋮	⋮	⋮

If the values of the independent and dependent random variables in regression analysis are observed over time, the data are called time series data. On the other hand, if these values are observed at a point in time, the data are called cross-sectional data.

Cross-sectional data

Definition

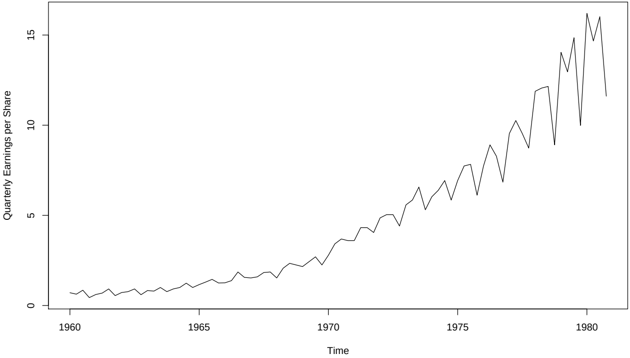
Cross-sectional data are values observed at one point in time.

- Starting salary and GPA (Grade Point Average) for graduate students last spring.
- Home upkeep cost in the past year and current value for home in an area.
- Labor hours, occupied bed days, and average length of stay for hospitals last month.

In these examples, values of the first variable would be predicted from the values of the remaining variables for a particular graduate, home, or hospital.

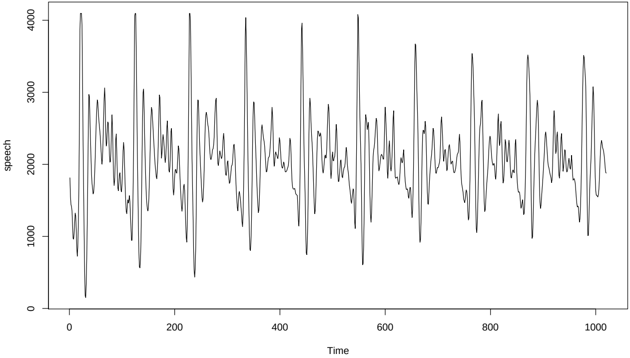
The nature of time series data - Example 1

Quarterly earnings per share for 1960Q1 to 1980Q4 of the U.S. company, Johnson & Johnson, Inc. (Upward trend).



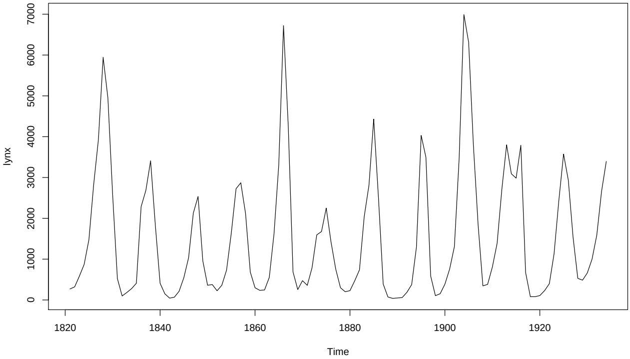
The nature of time series data - Example 2

A small .1 second (1000 points) sample of recorded speech for the phrase "aaa...hhh". (Regular repetition of small wavelets).



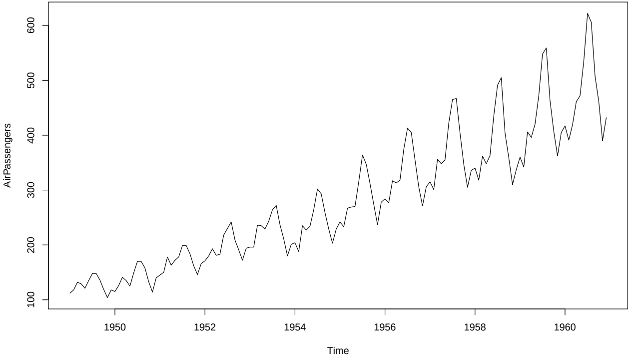
The nature of time series data - Example 3

Annual numbers of lynx trappings in McKenzie river in Northwest Territories of Canada over the years 1821-1934. (Aperiodic cycles of approximately 10 years).



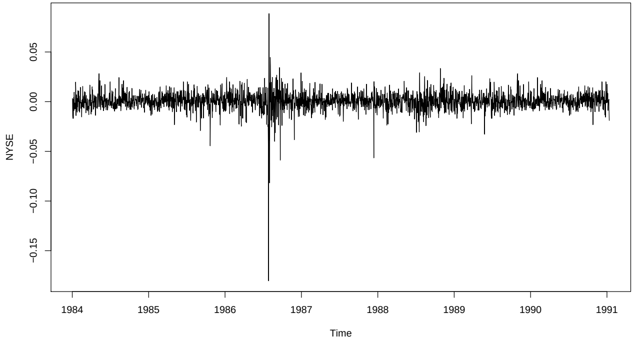
The nature of time series data - Example 4

Monthly Airline Passenger Numbers 1949-1960. (Seasonality appears to increase with the general trend).



The nature of time series data - Example 5

Returns of the New York Stock Exchange (NYSE) from February 2, 1984 to December 31, 1991. (Average return of approximately zero, however, volatility (or variability) of data changes over time).



Components of a time series

In general, the time series can be decompose into 4 components: **Secular Trend, Seasonal Variation, Cyclical Variation, and Irregular Variation** that can be modelled deterministically with mathematical functions of time (see the next slide).

- General Trend: The smooth long term direction (upward or downward) of a time series.
- Seasonal Variation: Patterns of change in a time series within a year which tend to repeat each year.
- Cyclical Variation: The rise and fall of a time series over periods longer than one year.
- Irregular fluctuations: Random and follow no regularity in the occurrence pattern.

Three types of time series decomposition (see chapter 7)

1. **The additive decomposition model:**

$$Y_t = T_t + S_t + C_t + I_t$$

   - Y_t = Original data at time t
   - T_t = Trend value at time t
   - S_t = Seasonal variation at time t,
   - C_t = Cyclical variation at time t,
   - I_t = Irregular fluctuation at time t.

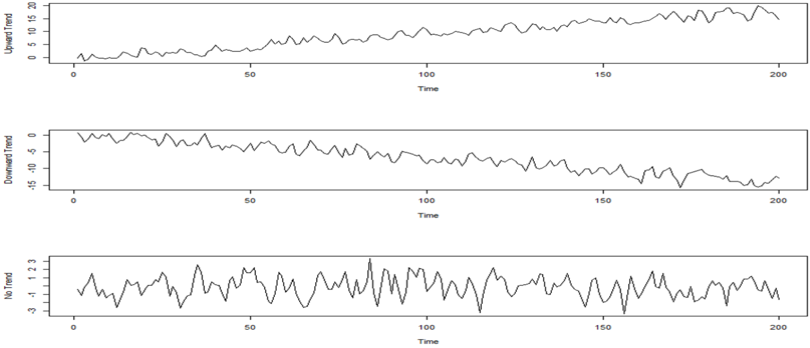
2. **The multiplicative decomposition model:**

$$Y_t = T_t \times S_t \times C_t \times I_t$$

3. **The Mixed decomposition model:** For example, if T_t and C_t are correlated to each others, but they are independent from S_t and I_t, the model will be $Y_t = T_t \times C_t + S_t + I_t$

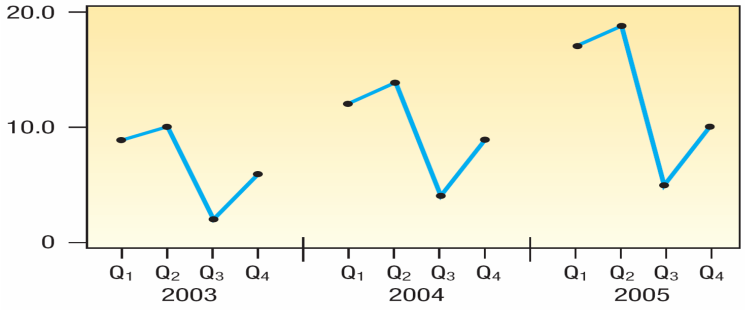
Secular (general) trend: sample chart

The increase or decrease in the movements of a time series that does not appear to be periodic is known as a trend.



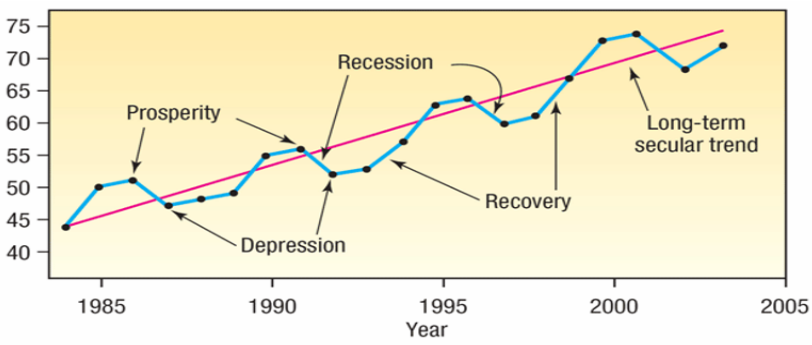
Seasonal variation: sample chart

Short-term fluctuation in a time series which occur periodically in a year. This continues to repeat year after year, although the term is applied more generally to repeating patterns within any fixed period, such as restaurant bookings on different days of the week.



Cyclical variation: sample chart

Recurrent upward or downward movements in a time series where the period of cycle is greater than a year. Also these variations are not regular as seasonal variation.



Time series decomposition in R (see chapter 7)

In R, the function `decompose()` estimates trend, seasonal, and irregular effects using the Moving Averages method (MA). In this case, the series is decompose into 3 components: **trend-cycle component, seasonal component, and irregular component**.

1. **The additive decomposition model:**

$$Y_t = T_t + S_t + I_t$$

2. **The multiplicative decomposition model:**

$$Y_t = T_t \times S_t \times I_t$$

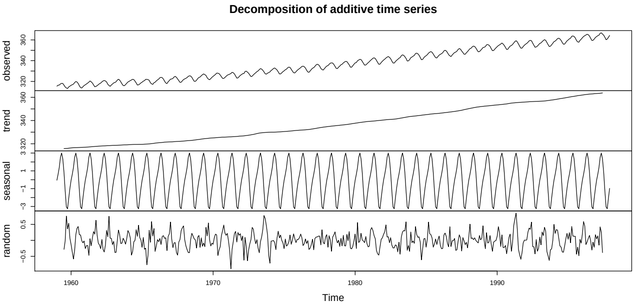
where T_t is the trend-cycle component (containing both trend and cycle components).

Continue to next slide ...

The function `decompose()` in R - Example

Atmospheric concentrations of CO2 (monthly from 1959 to 1997).

```
R> D <- decompose(co2)
R> season.term <- D$figure; trend.term <- D$trend
R> random.term <- D$random; plot(D)
```



Forecasting

Definition

Prediction of future events and conditions are called forecasts, and the act of making such predictions is called forecasting.

- Government of a country must be able to forecast such things as air quality, water quality, unemployment rate, inflation rate, and welfare payment in order to formulate its policies.
- The university might wish to forecast the daily mean temperature so that it can plan its fuel purchases for the coming month.
- The local school board must be able to forecast the number of children of elementary school age who will be living in the school district years in the future in order to decide whether a new school should be built.
- Any organization must be able to make forecasts in order to make intelligent decisions.

Qualitative and quantitative forecasting methods

- Qualitative forecasting methods: Generally use the opinions of experts to predict future events subjectively (judgmental forecasting methods).
- Such methods are often required when historical data either are not available at all or are scare.
- The common qualitative forecasting techniques are:
  - Subjective curve fitting.
  - Delphi method.
  - Technological comparisons.
- Quantitative forecasting methods: Involve the analysis of historical data in an attempt to predict future values of a variable of interest.
- Quantitative forecasting models can be grouped into two-kinds- univariate models and casual models.

Quantitative forecasting methods: univariate and casual models

- A univariate forecasting models predicts the future values of a time series solely on the basis of the past values of the time series.
  - Useful in forecasting when conditions based on historical data are expected to remain the same.
  - Not useful in forecasting when the historical conditions keep changing; e.g., not good in predicting the changes in sales that might result from a price increase, increased advertising expenditure, or a new advertising campaign.
- A casual forecasting models involves the identification of other variables that are related to the variable to be predicted; e.g., the sales of a product (dependent variable) might be related to the price of the product, advertising expenditure to promote the product, competitors' prices for similar products, and so on (independent variables).

Types of forecasts: point and prediction interval forecasts

- Point forecasts is a single number that represents the best prediction (or guess) of the actual value of the variable being forecasted.
- Prediction interval forecasts is an interval (or range) of numbers that is calculated so that we are very confident-(for instance, 95% confident) that the actual value will be contained in the interval.

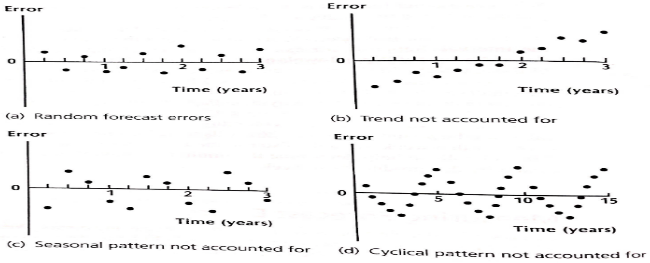
Measuring forecast errors

Consider the value of the variable of interest in time t as y_t and the predicted value as $\hat{y}_t$. Then

Definition

The forecast error for a particular forecast $\hat{y}_t$ is

$$e_t = y_t - \hat{y}_t$$



Measuring forecast errors

Definition

The absolute deviation (AD) of the forecasting errors is

$$\text{Absolute deviation (AD)} = |e_t| = |y_t - \hat{y}_t|$$

Definition

The mean absolute deviation (MAD) of the forecasting is

$$\text{Mean absolute deviation (MAD)} = \frac{\sum_{t=1}^{n} |e_t|}{n} = \frac{\sum_{t=1}^{n} |y_t - \hat{y}_t|}{n}$$

Definition

The squared error (SE) of the forecasting is

$$\text{Squared error (SE)} = (e_t)^2 = (y_t - \hat{y}_t)^2$$

Example

Compute the mean absolute deviation (MAD)

Actual value y	Predicted value $\hat{y}_t$	Error $e_t = y_t - \hat{y}_t$	Absolute deviation $\lvert e_t \rvert = \lvert y_t - \hat{y}_t \rvert$
25	22	3	3
28	30	-2	2
29	30	-1	1

$$\sum_{t=1}^{3} |e_t| = 6$$

$$\text{Mean absolute deviation (MAD)} = \frac{\sum_{t=1}^{3} |e_t|}{3} = \frac{6}{3} = 2$$

Definition

The mean squared error (MSE) for all forecasts is

$$\text{Mean squared error (MSE)} = \frac{\sum_{t=1}^{n}(e_t)^2}{n} = \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n}$$

Example

Actual value y	Predicted value $\hat{y}_t$	Error $e_t = y_t - \hat{y}_t$	Squared error $(e_t)^2 = (y_t - \hat{y}_t)^2$
25	22	3	9
28	30	-2	4
29	30	-1	1

$$\sum_{t=1}^{3}(e_t)^2 = 14$$

$$\text{Mean squared error (MSE)} = \frac{\sum_{t=1}^{3}(e_t)^2}{3} = \frac{14}{3} = 4.67$$

Mean absolute deviation versus mean squared error

Compare the errors produced by two different forecasting methods using MAD and MSE measurements

	Actual value y	Predicted value $\hat{y}_t$	Error $e_t = y_t - \hat{y}_t$	Absolute deviation $\lvert e_t \rvert = \lvert y_t - \hat{y}_t \rvert$	Squared error $(e_t)^2 = (y_t - \hat{y}_t)^2$
Forecasting	60	57	3	3	9
method	64	61	3	3	9
A	67	70	-3	3	9
				9	27

Mean absolute deviation (MAD) $= 9/3 = 3$

Mean squared error (MSE) $= 27/3 = 9$

	Actual value y	Predicted value $\hat{y}_t$	Error $e_t = y_t - \hat{y}_t$	Absolute deviation $\lvert e_t \rvert = \lvert y_t - \hat{y}_t \rvert$	Squared error $(e_t)^2 = (y_t - \hat{y}_t)^2$
Forecasting	60	59	1	1	1
method	64	65	-1	1	1
B	67	73	-6	6	36
				8	38

Mean absolute deviation (MAD) $= 8/3 = 2.67$

Mean squared error (MSE) $= 38/3 = 12.67$

Absolute average absolute percentage errors

Definition

The absolute percentage error is a measure of the forecasting error as a percentage of actual value. It allows comparison across different time series with values of different magnitude by dividing the absolute deviations by the actual value y_t and then multiply by 100.

$$\text{Absolute percentage error}(APE_t) = \frac{|e_t|}{y_t} \times 100 = \frac{|y_t - \hat{y}_t|}{y_t} \times 100$$

Definition

The average absolute percentage error (MAPE) is

$$\text{Average absolute percentage error (MAPE)} = \frac{\sum_{t=1}^{n} APE_t}{n}$$

Example

Compute the mean absolute percentage error

Actual value y	Predicted $\hat{y}_t$	Error $\lvert e_t \rvert = \lvert y_t - \hat{y}_t \rvert$	Absolute % error $APE_t = (\lvert e_t \rvert / y_t)(100)$
25	22	3	12.0
28	30	2	7.1
29	30	1	3.5

$$\sum_{t=1}^{3} APE_t = 22.6$$

$$\text{Mean absolute percentage errors (MAPE)} = \frac{\sum_{t=1}^{3} APE_t}{3} = \frac{22.6}{3} = 7.5$$

Homework 1

- Install R on your own PC (laptop).
- Install the following R packages: **t**series, **f**orecast,**T**SA,**z**oo **a**stsa, **t**imeSeries, **p**ortes, **s**arima, **M**ASS, **l**attice, **n**lme, **M**TS, **v**ars, **m**vtnorm, **f**racdiff, **a**kima, **f**Garch, **F**itAR.
- Download (as txt files) all data sets needed for this course, create a "ts" object, plot and explain each one:
  `http://www.stat.pitt.edu/stoffer/tsa4/`,
  `http://staff.elena.aut.ac.nz/Paul-Cowpertwait/ts/`, and
  `http://astro.temple.edu/~wwei/data.html`
- Read the summarized time series analysis on CRAN task view: `https://cran.r-project.org/web/views/TimeSeries.html`.
- Solve question 1.6 in the textbook page 25.