Time Series Analysis

The set of data collected on the basis of time is called time series.

Time series
use-cases
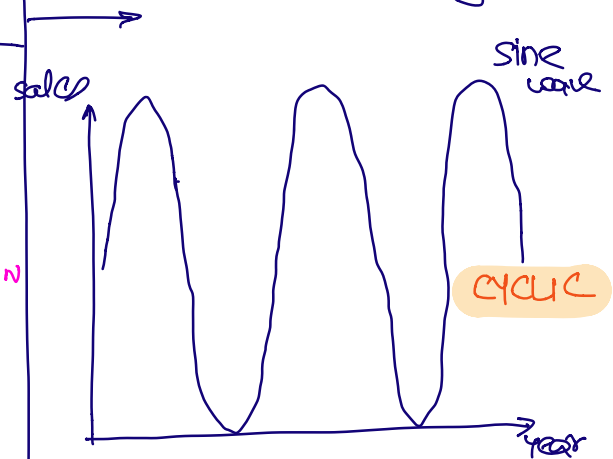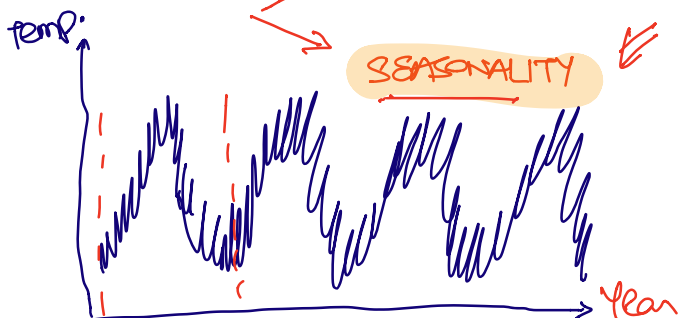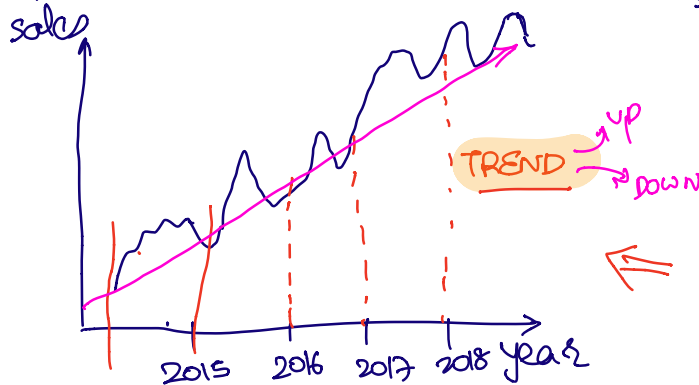
Univariate Time Series

time	sales

Multivariate time series

time	f1	f2	...	fn

Patterns that may encounter while working on time series data:

sales

2015 2016 2017 2018 year

TREND → UP
       → DOWN

sales

sine wave

CYCLIC

year

temp.

SEASONALITY

Year

RANDOM

Steps to perform when dealing with Time series forecasting problem: (Python)

① Ensure your dataset must hold a column that represent time (year, month, days, hours, min, sec)

② Ensure your time column is represented as datetime datatype

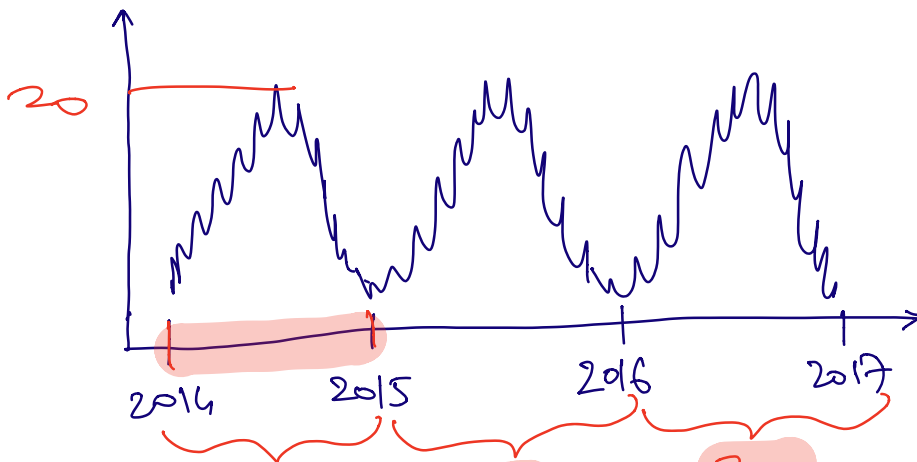③ Ensure your time column is a row index in pandas dataframe.

	Sales
2016-01	266
2016-02	146

datetime (datatype)

④ Check whether the data is stationary or not.

Stationary Data in Time Series:

It's a term which defines, the

mean
Variance   } Same over a given period
Co-variance

when can I say the data is stationary?

Lets assume period = 1 year

3 periods

mean → 20 ~ 20 ~ 20
variance → 0.2 ~ 0.2 ~ 0.2
covariance → 0.1 ~ 0.1 ~ 0.1

if its stationary then go for step 5.
Else report data as inappropriate for
time-series modelling

Note:- Incase your data is a trend data
use Integral with 1 step to convert the
same into Seasonality (Stationary Data)

5. Create train test split to understand the
accuracy & generalization of model

6. Create the model.

→ (a) Auto regressive model (AR model)

small time frame data → ⓑ Moving Average mode (MA model)

with moredata →ⓒ ARIMA mode

7. Predict the values of test data and check the error score. Try to ensure your model has less error score.

8. If satisfied with model, deploy the same.

Time series modelling techniques

① A R modelling (Auto - Regressive model)

② MA modelling (Moving - Average model)

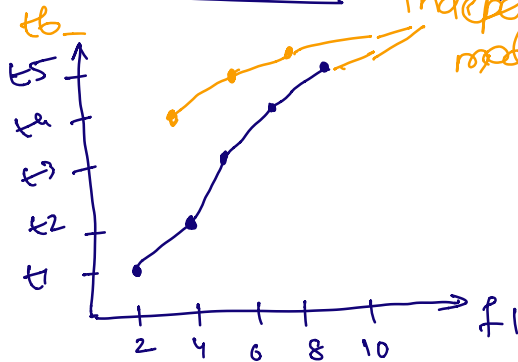③ ARIMA modelling (Auto Regressive Integrated moving Average)

AR modelling:

$$y_t = f(\underbrace{y_{(t-1)}, y_{(t-2)}, \cdots})_{\text{past data}}$$

$$= b_0 + b_1 y_{(t-1)} + b_2 y_{(t-2)} + \cdots$$

Moving average model = not an independent model   period (q) = 3

3 SmA
Simple
moving
Average

t	f1
t	2
t2	4
t3	5
t4	7
t5	9

$$\frac{2+4+5}{3} = 3.\sim$$

$$\frac{4+5+7}{3} =$$

baseline for ARIMA

Its a lagging indicator for ARIMA correction.

ARIMA

3 parameters

Period $\longrightarrow$ (P) $\longrightarrow$ 1   AR

Integral order $\longrightarrow$ (d) $\rightarrow$ 1   I

moving average $\rightarrow$ (q) $\rightarrow$ 1   MA