Table of Contents

1 Introduction to Time Series Analysis

1.1 The Subject of Time Series Analysis

Time series is the set of the digitized indexed samples, in which each samples is connected with other in some relation.

As a rule there statistical relation have place to be.

- Usually Time series is considering as digitized

  sequential series of some values which are taken with some **time-step**.

  Often equidistant time-steps are used.

- Each series value can be called **sample**.

- **Time Series Analysis** is one or a sequential set of method/algorithm to obtain (or estimate/predict) some of the series parameters (or features).

Such features can be the further (previous) values of samples, some series parameters or its model at all.

- As a rule, the series is treated by statistically-based methods of analysis.

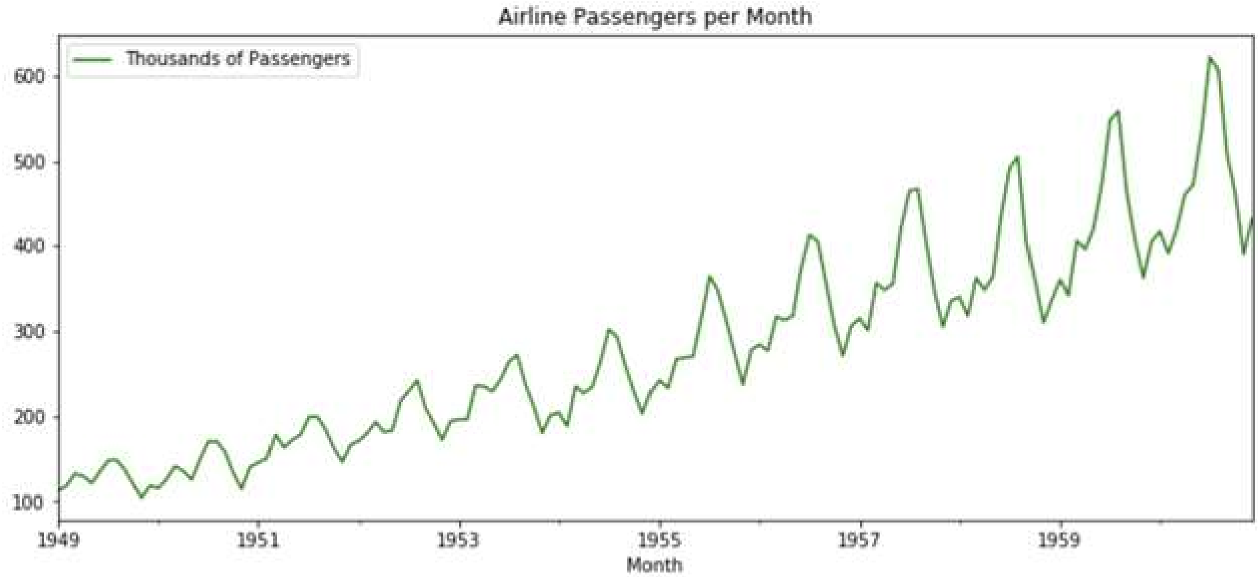
1.2 Examples of Time Series Applications

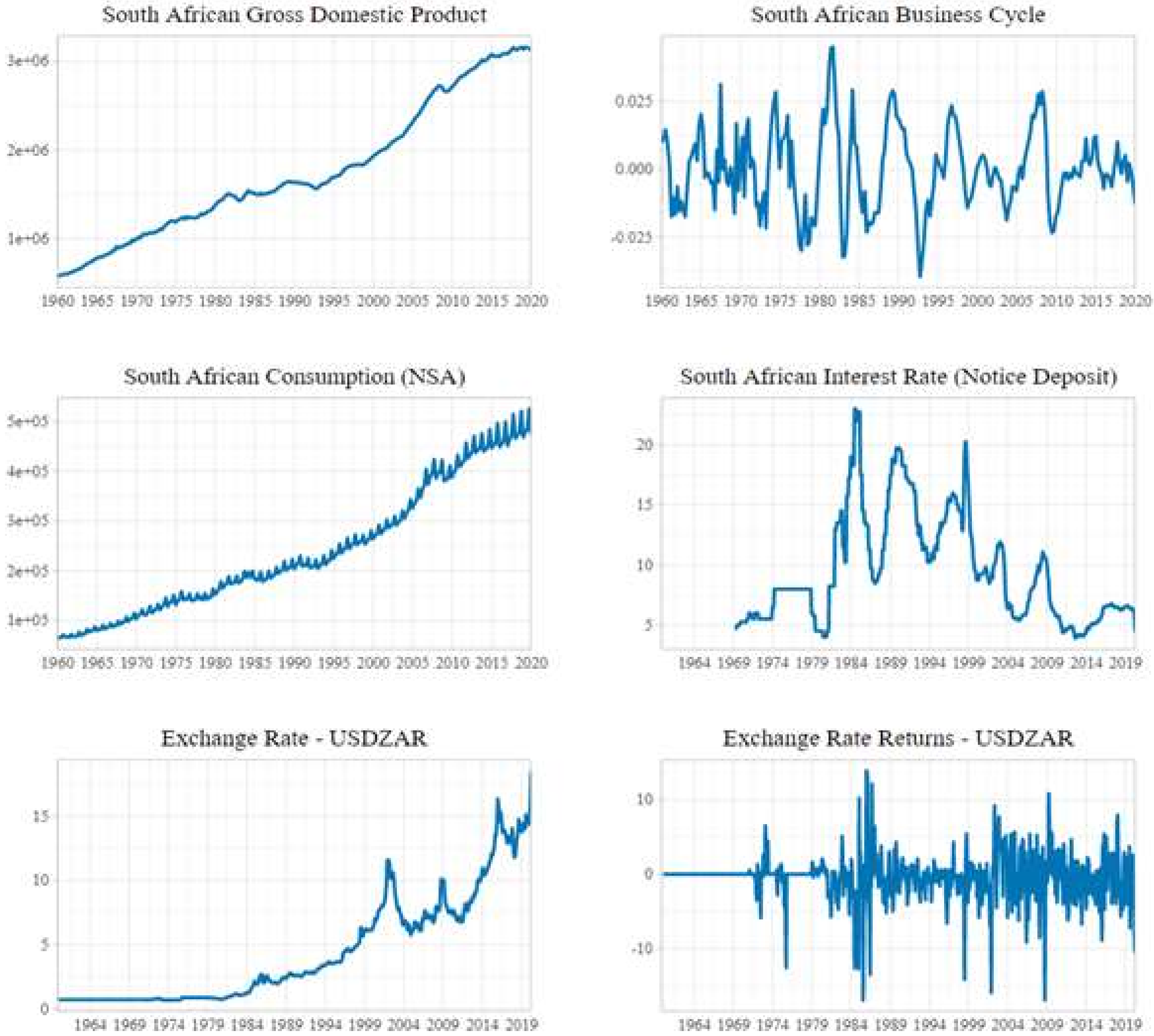
The examples of the **Time Series Analysis Applications** are:

- **Economic Forecasting** (Econometrics Analysis, Stock Market Analysis, resource consumption).
- **Industrial machine diagnostic** (for instance, machine fault diagnostic by vibration analysis).
- Physical and chemical **processes behavior prediction** (for instance, sun spot prediction, weather forecasting).
- **Medical diagnostic** (EEG, EMG and e.t.c.).
- **Signal processing** (for instance, measurement signals from some industrial sensors).
- **Industrial systems control** and diagnostic (for instance, forecast of products quality by sensors measurement results).
- **Time series database (TSDB)** is a database optimized for time-stamped or time series data. Time series data are simply measurements or events that are

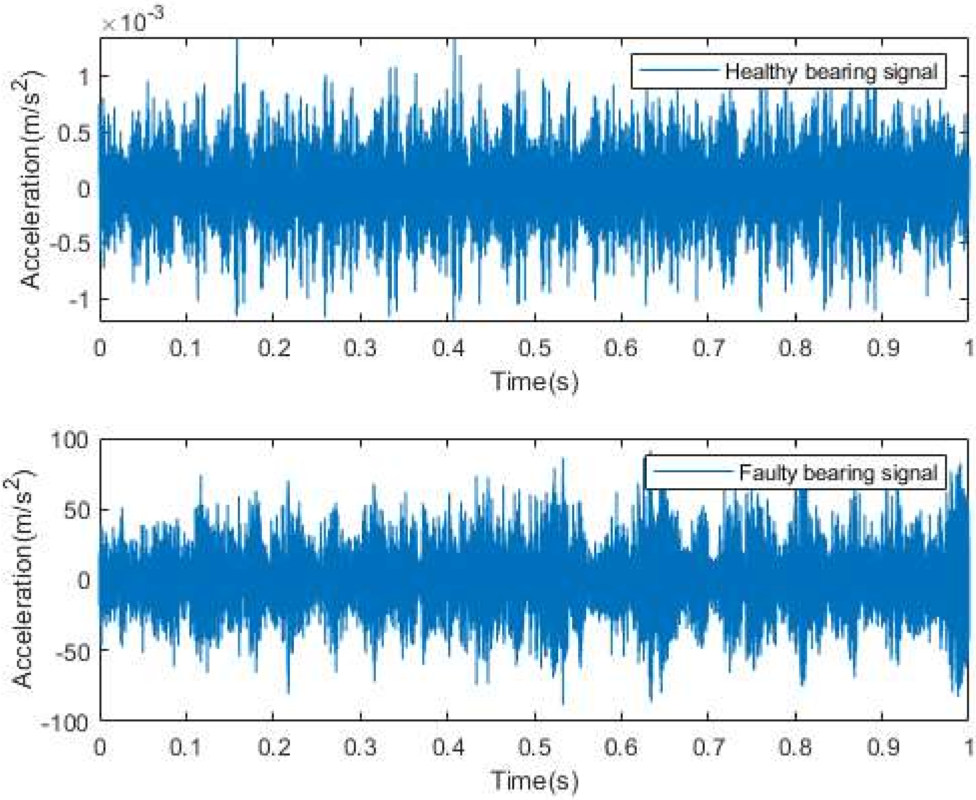
tracked, monitored, downsampled, and aggregated over time. A time series database is built specifically for handling metrics and events or measurements that are time-stamped.

Economic Forecasting.

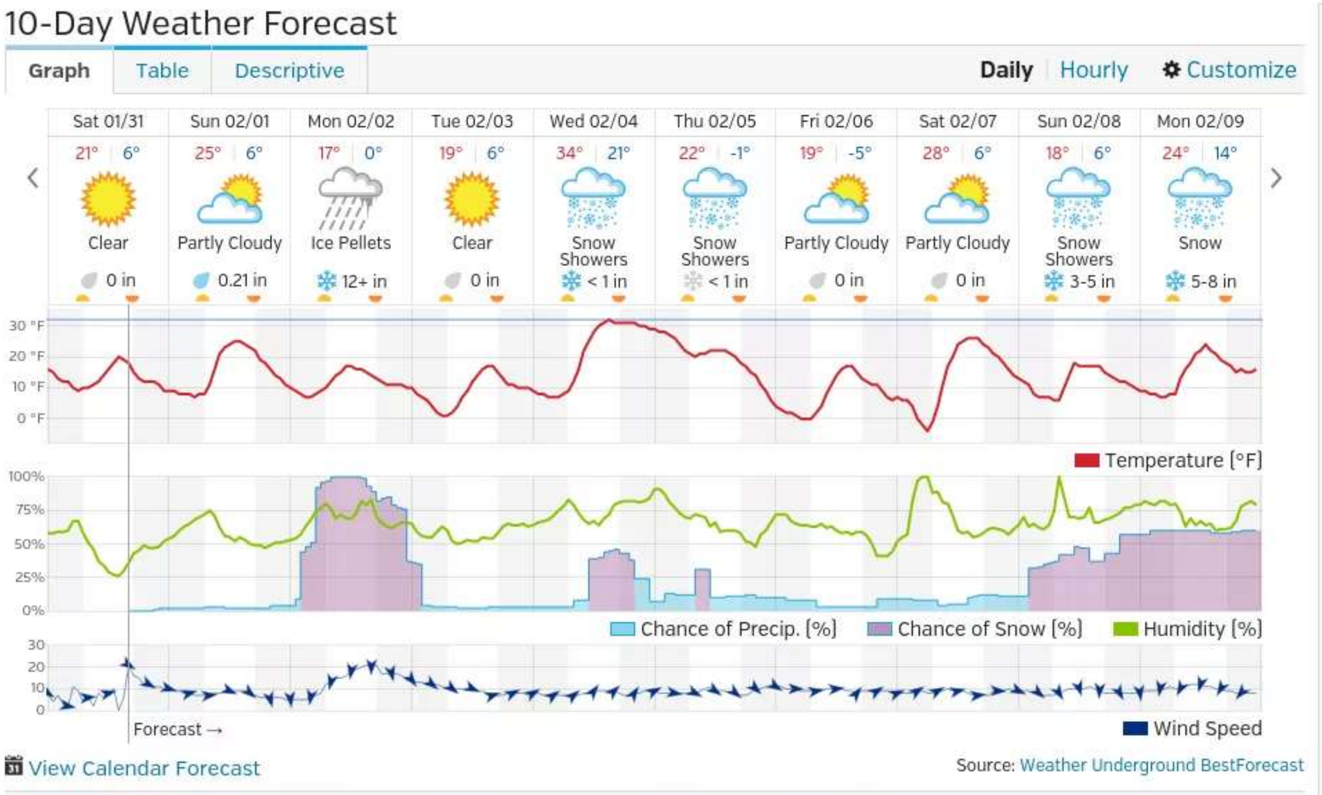




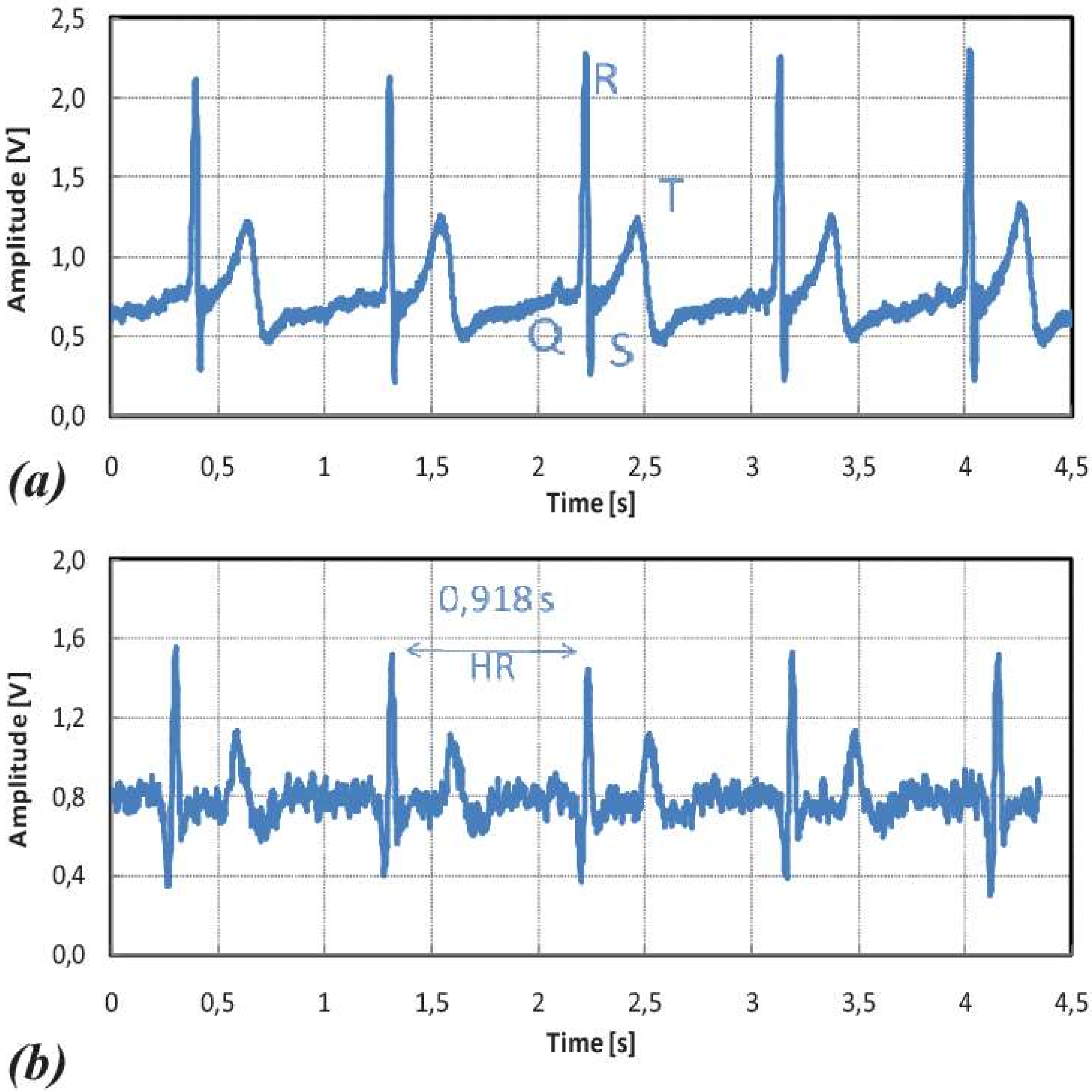
- Machine fault diagnostic



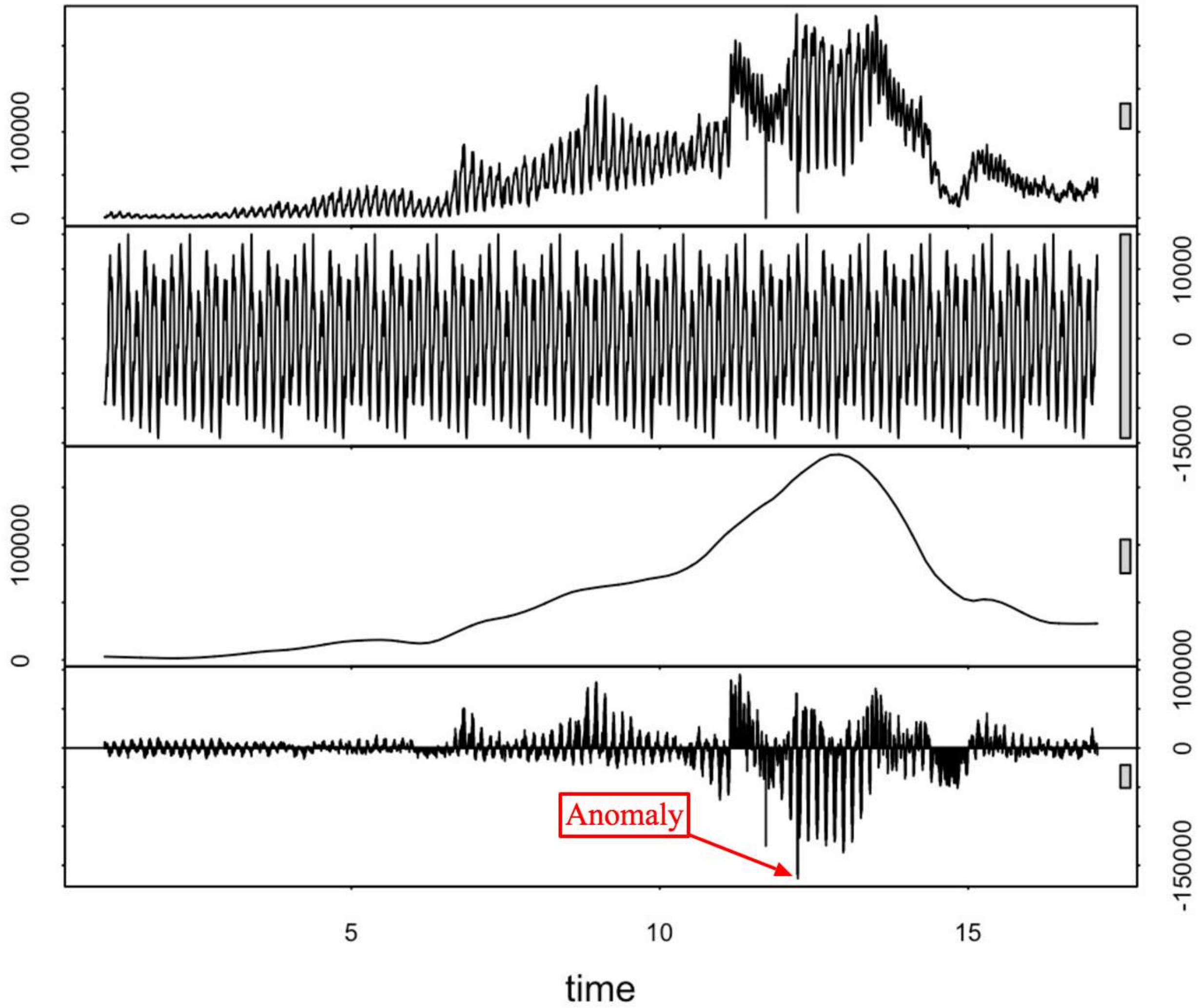
- Weather forecasting



- Medical diagnostic



- Signal processing



Example: Industrial systems control and diagnostic

Time series data set

Sensor ID	Time Stamp	Value 1
Sensor_1	01/01/2020	236
Sensor_1	01/02/2020	133
Sensor_1	01/03/2020	148
Sensor_1	01/04/2020	152
Sensor_1	01/05/2020	241

Time series as supervised learning problem

Sensor ID	Time Stamp	Value x	Value y
Sensor_1	01/01/2020	NaN	236
Sensor_1	01/01/2020	236	133
Sensor_1	01/02/2020	133	148
Sensor_1	01/03/2020	148	152
Sensor_1	01/04/2020	152	241
Sensor_1	01/05/2020	241	Value to be predicted

Machine Learning Prediction

Multivariate time series data set

Sensor ID	Time Stamp	Value 1	Value 2
Sensor_1	01/01/2020	236	23
Sensor_1	01/02/2020	133	34
Sensor_1	01/03/2020	148	32
Sensor_1	01/04/2020	152	31
Sensor_1	01/05/2020	241	22

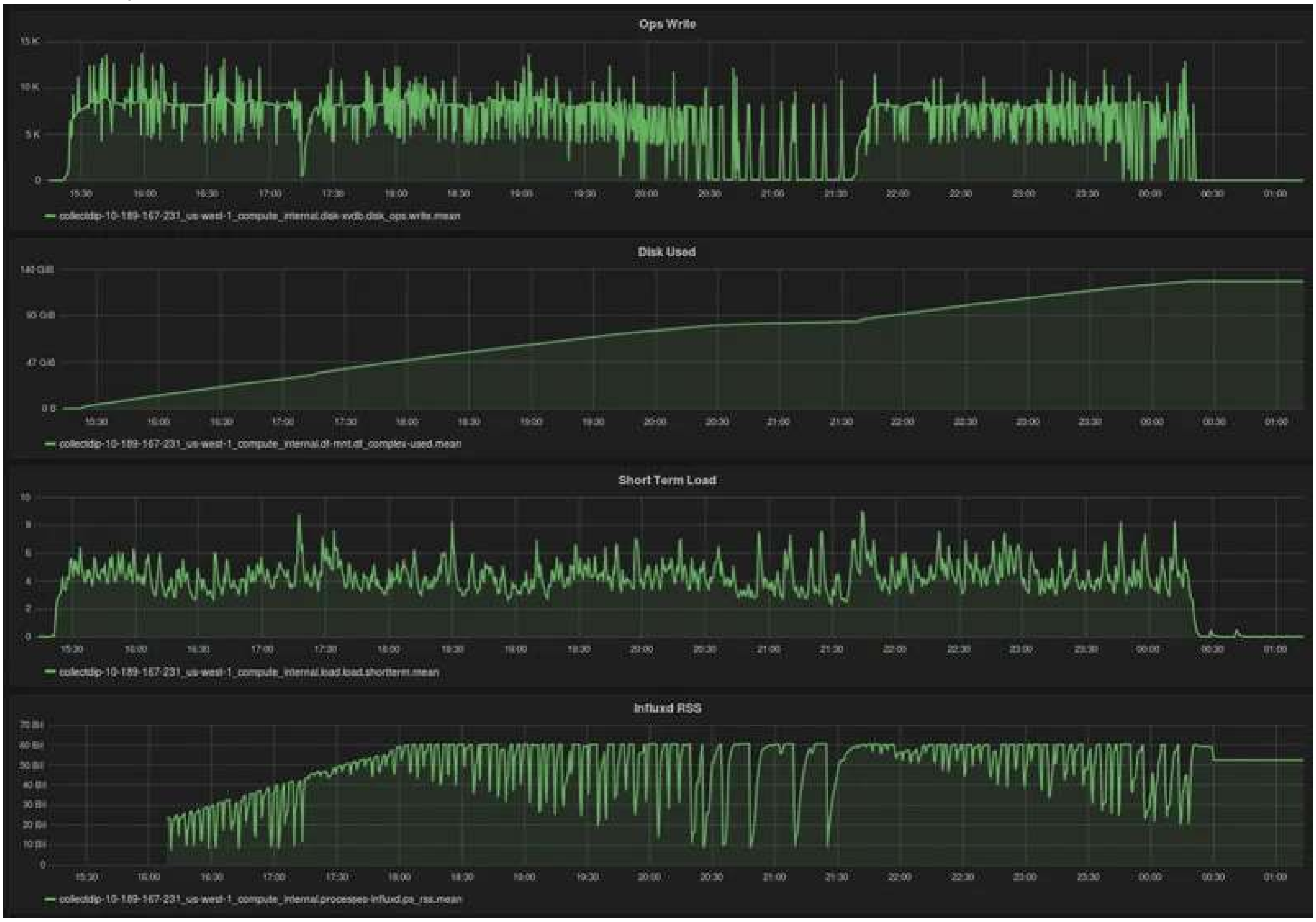
→

Multivariate time series as supervised learning problem

Sensor ID	Time Stamp	Value x	Value x2	Value x3	Value y
Sensor_1	01/01/2020	NaN	NaN	236	23
Sensor_1	01/01/2020	236	23	133	34
Sensor_1	01/02/2020	133	34	148	32
Sensor_1	01/03/2020	148	32	152	31
Sensor_1	01/04/2020	152	31	241	22
Sensor_1	01/05/2020	241	22	NaN	Value to be predicted

Machine Learning Prediction

Time-stamped or time series data base



1.3 Time Series Models

1.3.1 Model of time series.

A **Time Series** is a collection of observations (samples) indexed by time (in general other indexed variable can also be used).

> **In common case** time variable is discrete with uniformly taken time steps in which case we have a **discrete time series**,
>
> however it could be continuous in this case we can say that we have continuous time series.

> A **Time Series Model** for a Time Series $\{y_n\}$ is a specification of the joint probability distribution of the model values y to time steps n.

Model of time series.

The Simplest case of Time series is the 1-dimensional (univariate) value-to-time dependence, given in the following form

$$y(t) = a_0 + trend(t) + seasonal(t) + cyclic(t) + noise(t),$$

where

- $y(t)$ is the series - the set of samples indexed by some variable t, usually t is the time-steps, if time-step discreet it can be also denoted as n (sample number), in this case real time-step value will corresponds to $t = n \cdot T_s$,

where T_s is the period of step n (sampling period with which samples are taken).

- a_0 is some start constant level (equal to bias).
- $trend$ is the presence of some trend, which is the slow change part of the dependence.
- $seasonal$ is the seasonality or some "relatively rapidly change" periodic-like components - is a relatively fast-changing part of the relation.
- $cyclic$ is the some "relatively slow change" periodic-like components with irregular period and relatively high intensity.
- $noise$ is the some undetermined (i.e. stochastic) distortion of the output values - irregular or random fluctuations (variations).

Notes

- Often the $cyclic$ and a_0 are included into $trend$, in this case the model can be given as

$$y(t) = trend(t) + seasonal(t) + noise(t).$$

- When some trend-seasonal decomposition carried out the $noise$ are considering as all **residual** part.
- All components of $y(t)$ expect of $noise$ have deterministic nature.

  However in some cases components can be considered as stochastic also (only formally).
- In some cases you may also meet multiplicative model of time series

$$y(t) = trend(t) \cdot seasonal(t) \cdot cyclic(t) \cdot noise(t)$$

  in opposite to multiplicative model the above one can be called additive model.
- In general any time series model can be simulated as combination of a several multiplicative model and additive models, for instance

$$y(t) = a_0 + (trend(t) \cdot cyclic(t) + seasonal_1(t)) \cdot seasonal_2(t$$

- In more complex case series can contain several trends or other components in different relations one to others.

  For instance, facebook prophet model assume to include rare periodic events in the model (like holidays), in this case the model can be given as

$$y(t) = trend(t) + \sum_{i=0}^{M} seasonal(t) + \sum_{i=0}^{H} holidays(t) + nois$$

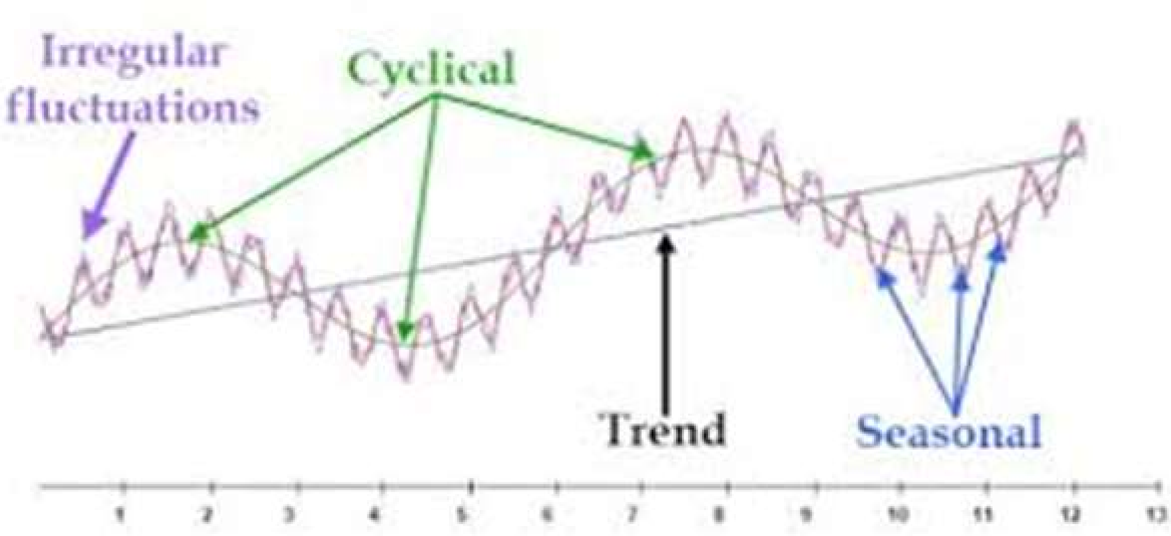
- Time series are typically assumed to be generated at regularly spaced interval of time (e.g. daily temperature), and so are called regular time series.

But the data in a time series doesn't obligate to come in regular time intervals. In that case it is called irregular time series. Account deposits or withdrawals from an ATM are examples of an irregular time series.

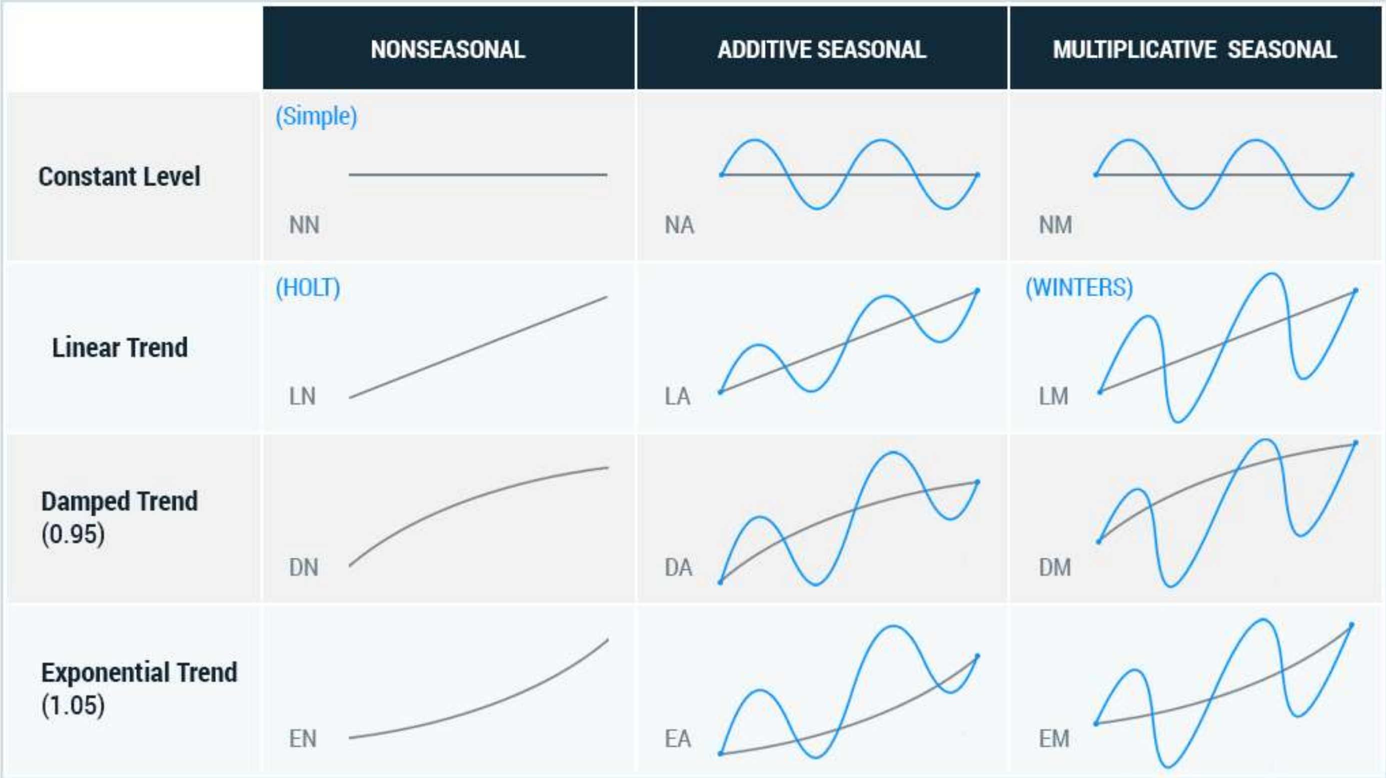
- Anomaly events can be also add to the time series model,

$$y(t) = trend(t) + \sum_{i=0}^{M} seasonal(t) + \sum_{i=0}^{H} holidays(t) + noise(t)$$

anomaly(t) -rare events

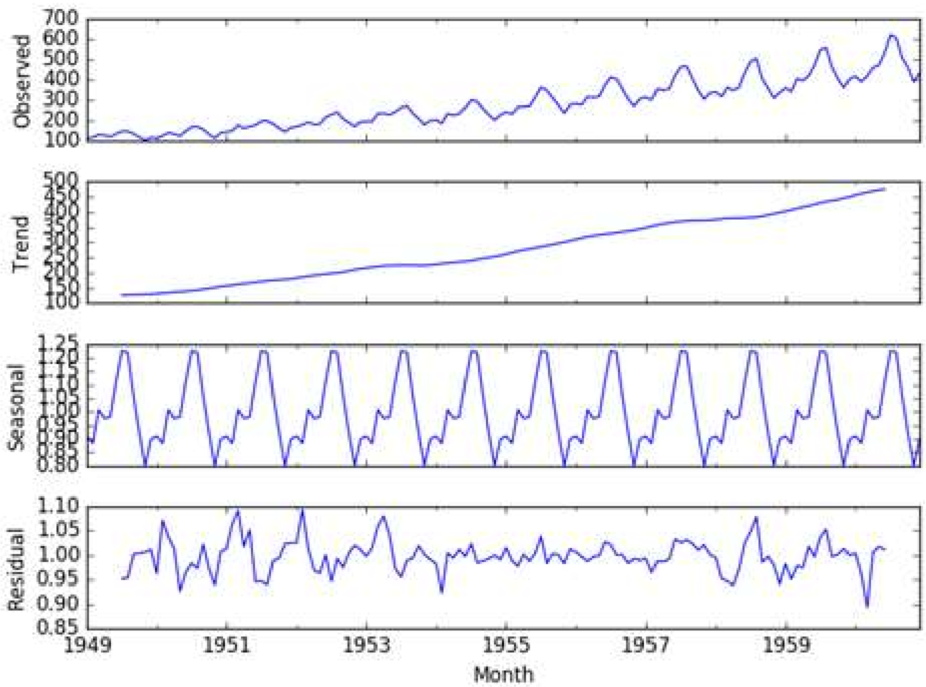


Examples of additive and multiplicative seasonality on trend

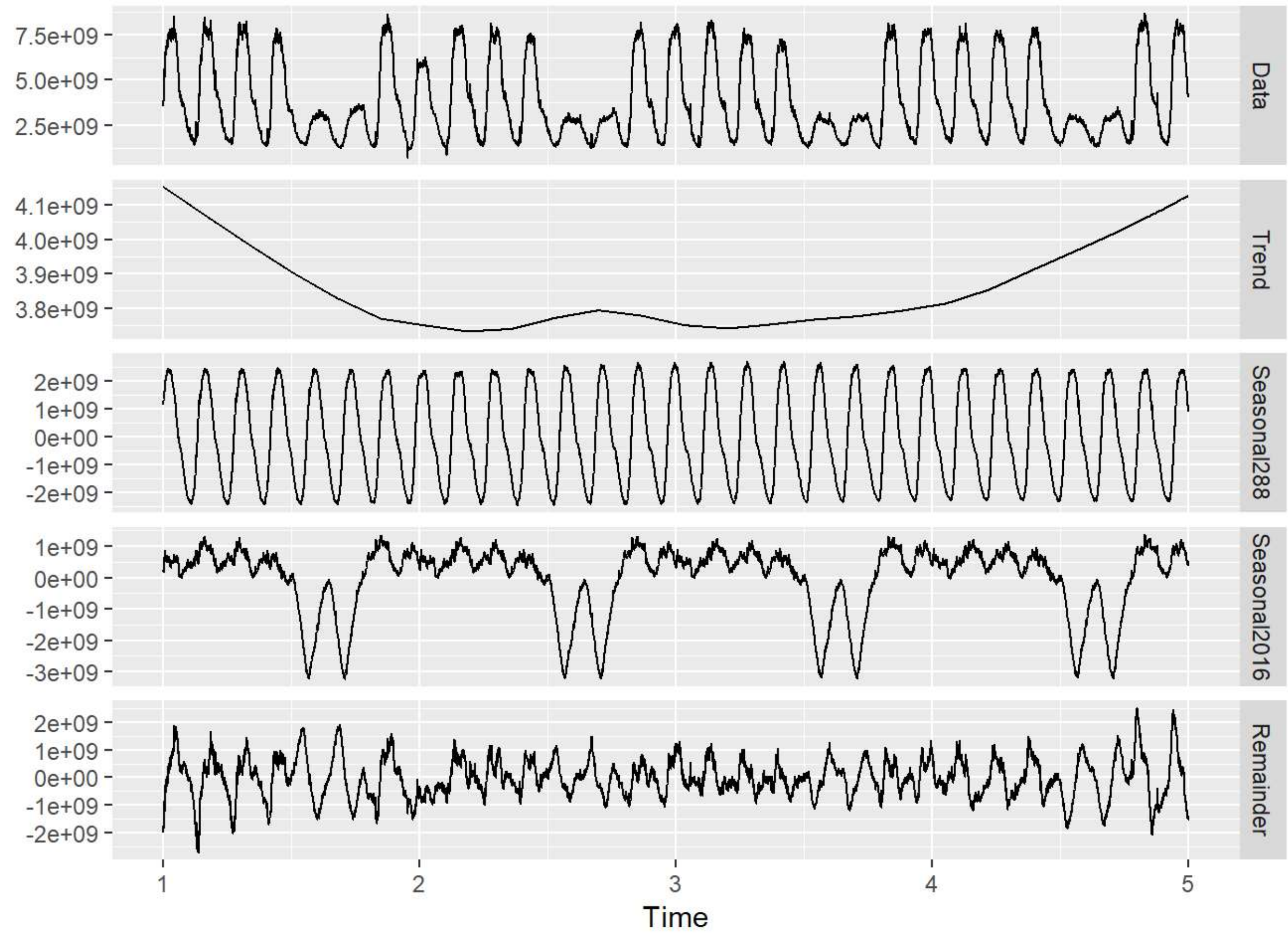


As a rule, the time series can presented (decomposed) into its main components by methods of the time series analysis.

The study of methods of time series analysis is the subject of our course.



Example of multi-seasonality



The interpretation of our data features as time series components depends on the our goal.

Before make decision about time series model we need to make and check a several hypothesis.

> For instance, lets considering the price of Bitcoin.
>
> If we wont to predict a day-change of the price we will considering the several hour-changing as seasonality, a week change as trend, a day-change as cyclic components and a minute change as noise.

But if we wan to estimate a month price, than the day change will be considered as a noise part, and also we can we can identify several trends (or trend to time dependence!).

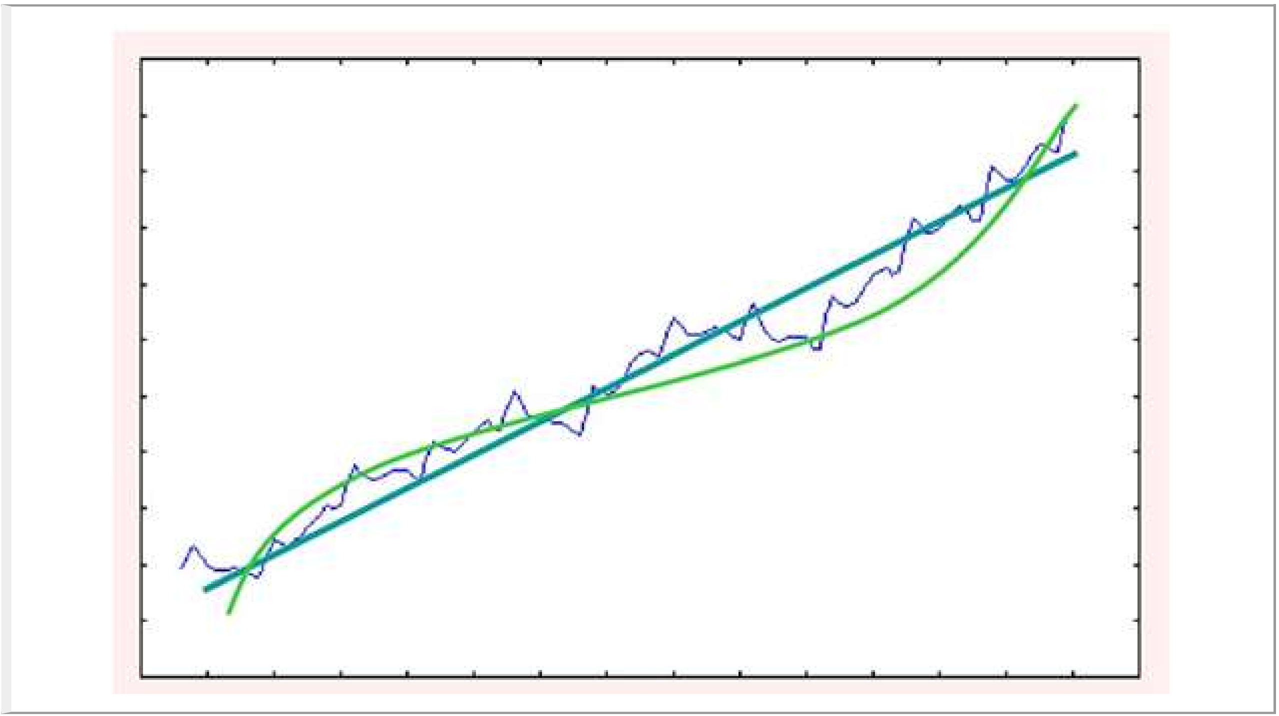
Thus the first task in the time series analysis (at least in its classical form) is to introduce the model of the series. Ideally we need to have the full formally described model of the series, but at least we need to have the trend model (and, or including cyclic part).

1.3.2 About the Time Series Trend

There are a several simplest **types of the trends** can be introduced within time series:

- lack of trend.
- random trend (or stochastic trend, like random walk).
- line trend $y(t) = a \cdot t + b$
- parabolic trend $y(t) = a \cdot t^2 + b \cdot x + c$
- polynomial trend $y(t) = a \cdot t^b + c$
- hyperbolic trend $y(t) = \frac{a}{t^b + c} + d$
- exponential trend $y(t) = \exp{(a \cdot t + b)}$
- saturating(logistic) trend $y(t) = \frac{c}{1 + exp(-k(t-m))}$
- logarithmic trend $y(t) = \log_b{(a \cdot t)}$
- many other functions that are, as a rule, smoothed, very slow changed or even monotonic.

> Example when there are a several hypothesis about trend are possible.



The important about the trend.

1. *Frequently we do not know how to approximate trend the best expect of linear trend.*
2. <u>The error of the trend approximation have main influence on the error of the time series prediction or its parameters estimation.</u>
3. **For most of the real-world series, the trend is not the monotonic growing (or falling) function**.

    The trend can have inflexion points and/or saturation parts which are necessary to be predicted with the most accuracy (for instance, see currency price prediction graphic above).

The piecewise-monotonic behavior is natural for many real-world processes, thus it is most important to identify the the inflection point

4. *If trend has a complex form it could be hard to distinguish it in series.*
5. In many cases the cyclical part of the series can be considered either as its complex trend or its seasonality.
6. **The deterministic trend can be affected by several random effects, such as a random walk and drift.**

   These effects, in fact, can be considered as noise but mostly influencing only on the trend part.

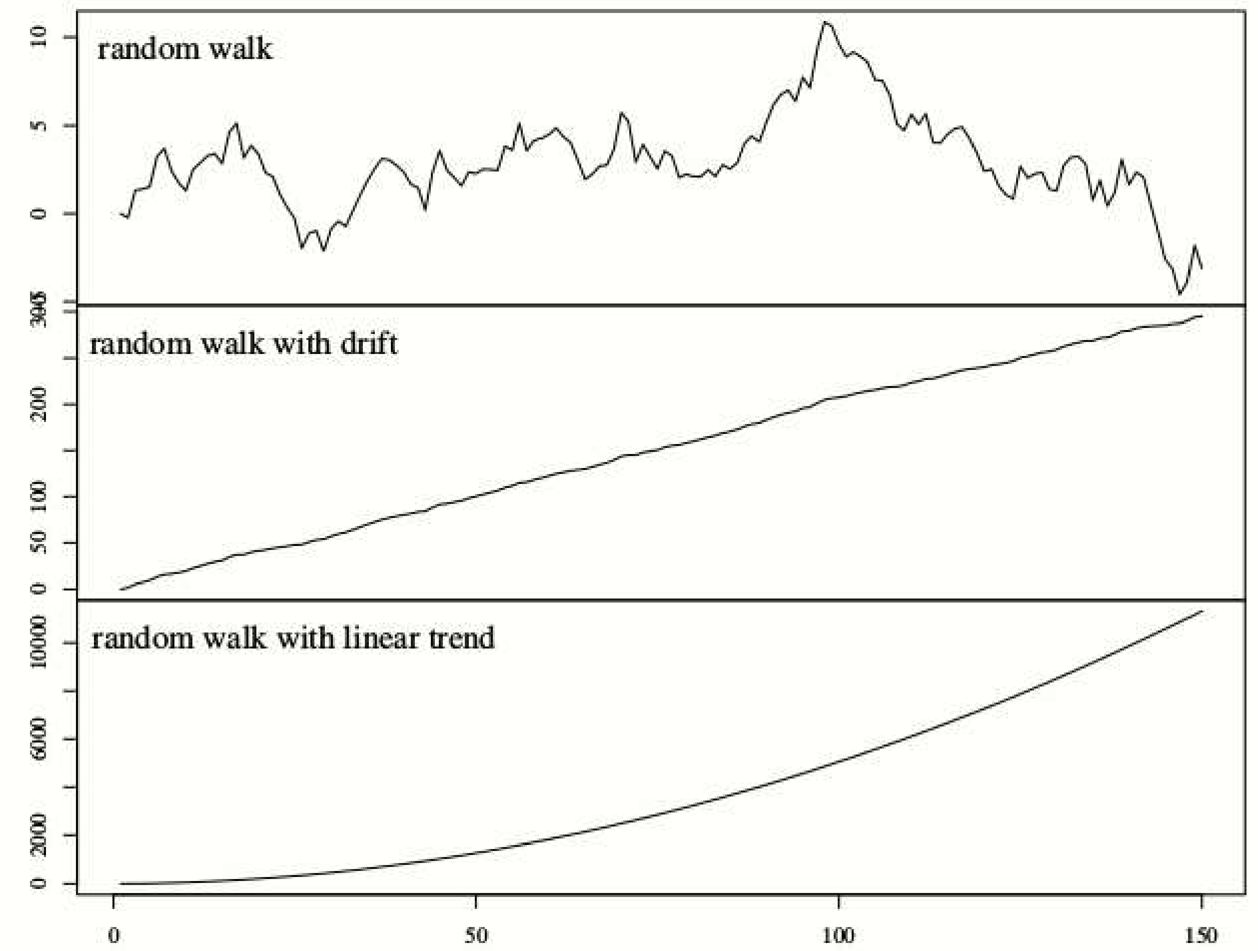
Random behavior simulation of time series trend

The random behavior of time series can be described as the model of its random walk and drift models. Random walk process, in particular for trend, can be imagined from the case where each value of the time series at time t is the value of the series at time $t-1$, plus a some random movement noted by ε_t

$$y_n = y_{n-1} + \varepsilon_n.$$

If y_n also has some constant increment (or decrement) μ, that part will be called drift.

$$y_n = y_{n-1} + \mu + \varepsilon_n.$$

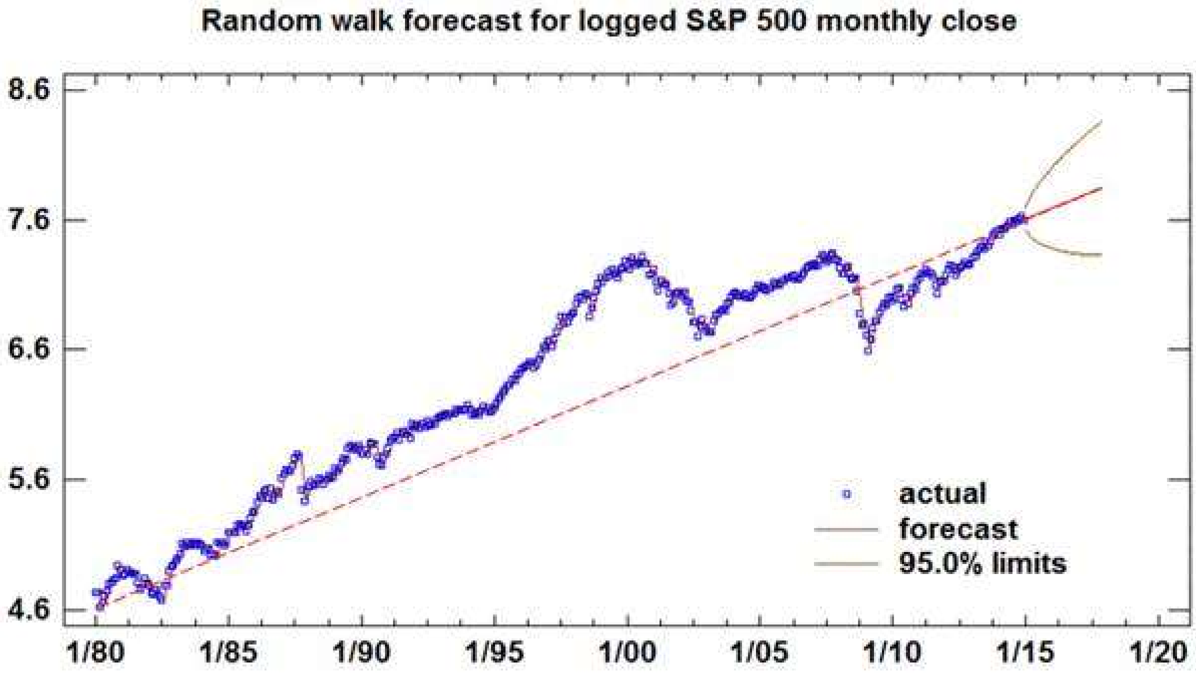


By repeating the random-walk we can re-write that with taken into account dumping of the previous values influence on the current value as

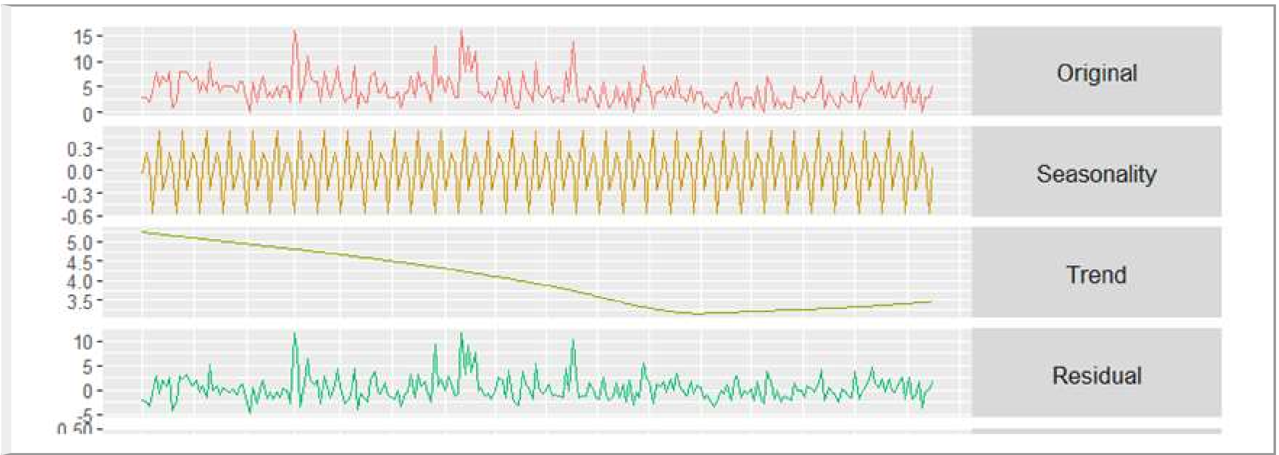
$$y_n = \sum_{i=0}^{n} \beta_n \varepsilon_i$$

where β_n is the coefficient of the previous random values influence on the current value.

Example of random walk and drift of trend



Example, when it is hard to extract trend correctly due to its relatively low slop and high noise influence.



The main goal of time series analysis.

In the time series analysis our main goal is to estimate and possibly predict the following values of the trend and deviation from the trend or to estimate the trend and cyclical parts.

In the last case the goal must be to estimate the trend duration (For instance, if we want to predict the trend change time).

1.3.3 Types of Time Series Representation

It important to introduce several definitions concerning time series analysis:

- **Univariate analysis** is the simplest form of time series analysis where the data being analyzed contains only one variable.
- **Bivariate analysis** is slightly more complex case than Univariate analysis. When the data set contains two variables both indexed by time and researchers aim to undertake comparisons between the two data or predict both of them.
- **Multivariate analysis** is a more complex form of time series analysis technique and used when there are more than two variables in the data set.
- **Analysis with exogenous factors**(covariate) is form of time series analysis when we suppose that the analyzed data depends on the some other data and time indexes.

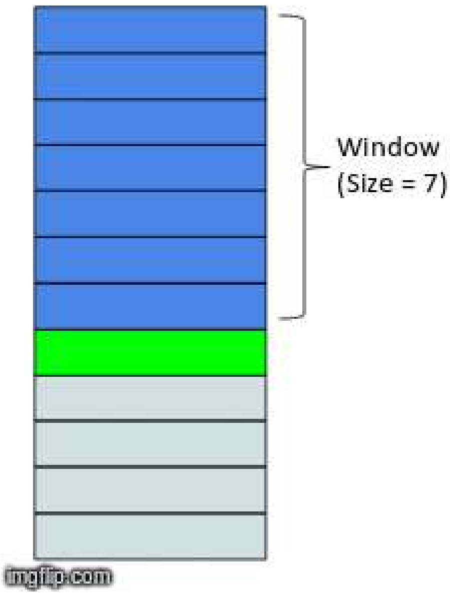
  The covariate factors can be considered as a priory information for analyzed series.
- **Date Time Features**: these are components of the time step itself for each observation.

- **Lag Features**: these are values at prior time steps.

    For instance, you are predicting the stock price for a company. So, the previous day's stock price is important to make a prediction, right? In other words, the value at time t is greatly affected by the value at time t-1. The past values are known as lags, so t-1 is lag 1, t-2 is lag 2, and so on.

- **Window Features**: these are a summary of values over a fixed window of prior time steps.

Example of the windowed feature.



1.3.4 Time Series Types

Time series types

1. **Stationary series** in strong meaning, when the series parts are the same with no matter in which moment the part is taken.

   In weak definition Stationary series it is the series in which you have constant mean and variance.

2. **Non-stationary series** if different parts of the series differ, for instance, when the spread of sample values change.
3. **Linear series model** when components of the series have linear or reduced to linear relations.

   - As a rule, for linear model we can say about certain simple relations, frequently with equidistant time stamps and without empty values and so on.
   - For linear model we can introduce some known analytical or heuristically-analytical model of time series.
   - For linear model we can distinguish and formally describe some patterns in behavior.

4. **Non-linear series model** when components of the series have complex or non-linear relations with each other.

   As a rule, for Non-linear model we can not say about certain patterns.

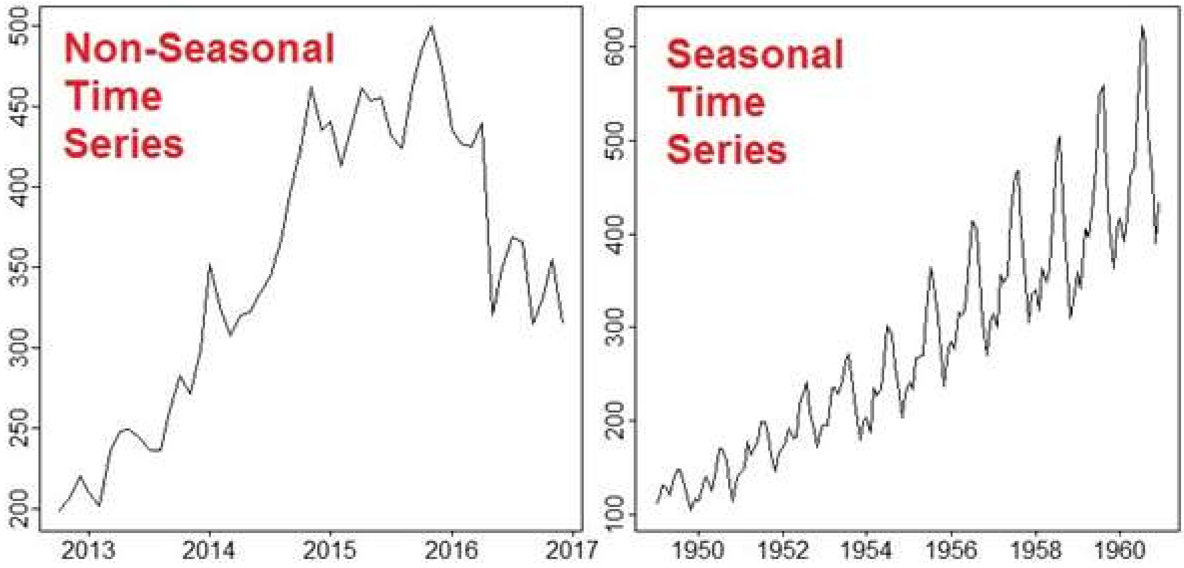
5. **Seasonal series** the data have some periodic patterns, in general fixed period is not necessary, for instance dependence of values per week).
6. **Non-seasonal series** when data have a no-periodic pattern or not-known behavior.
7. **Univariate series** when the forecast is based on only one variable that varying over time.
8. **Multivariate series** when the forecast is based on multiple variables that varying over time.
9. **Unstructured series** - series with no obvious systematic time-dependent pattern in a time series variable.

    As particular cases: **random-trend (or stochastic trend)** and **random-seasonality**.

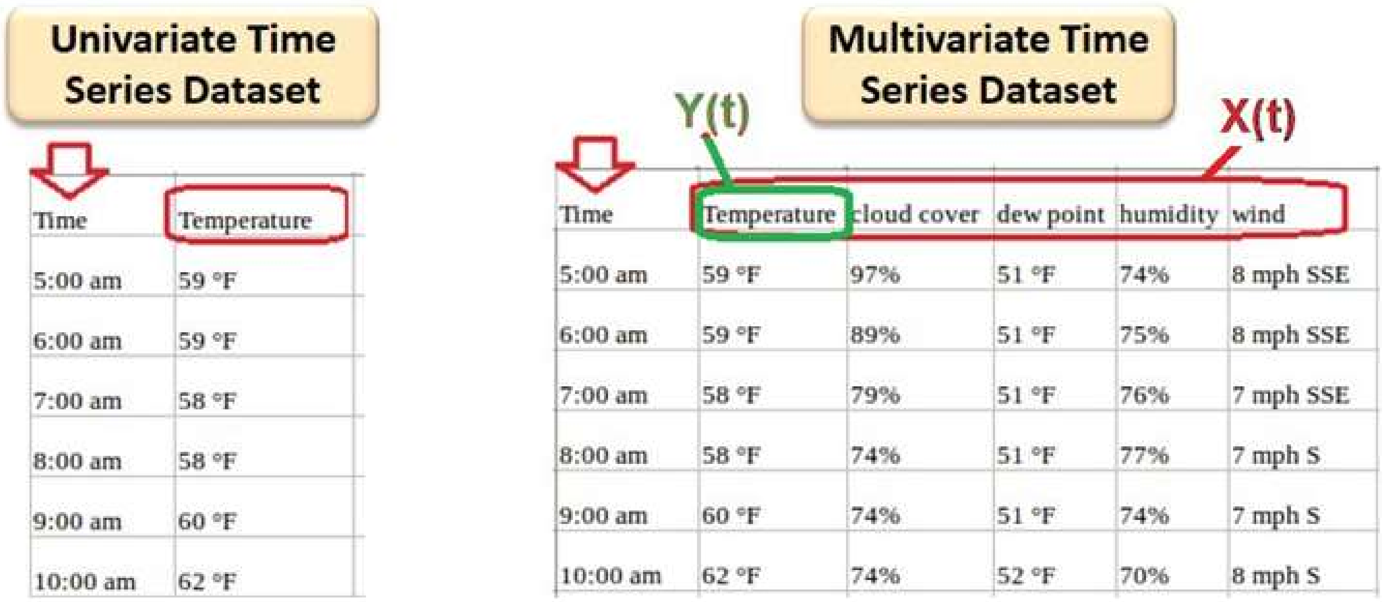
10. **Structured series** - series with systematic time-dependent patterns in a time series variable (e.g. trend and/or seasonality).

    As particular cases: **deterministic trend** and **deterministic seasonality**.

Examples of time series with cyclic trend and seasonality with linear trend

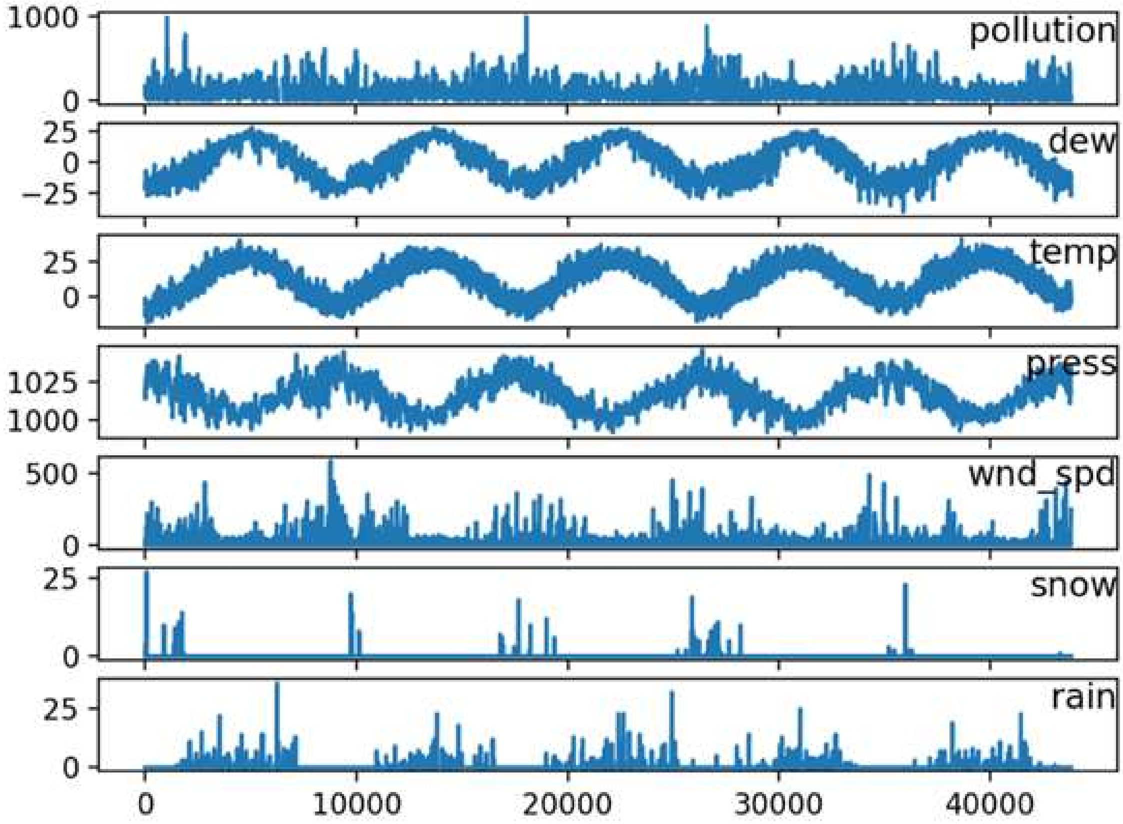


Examples of univariate and multivariate time series

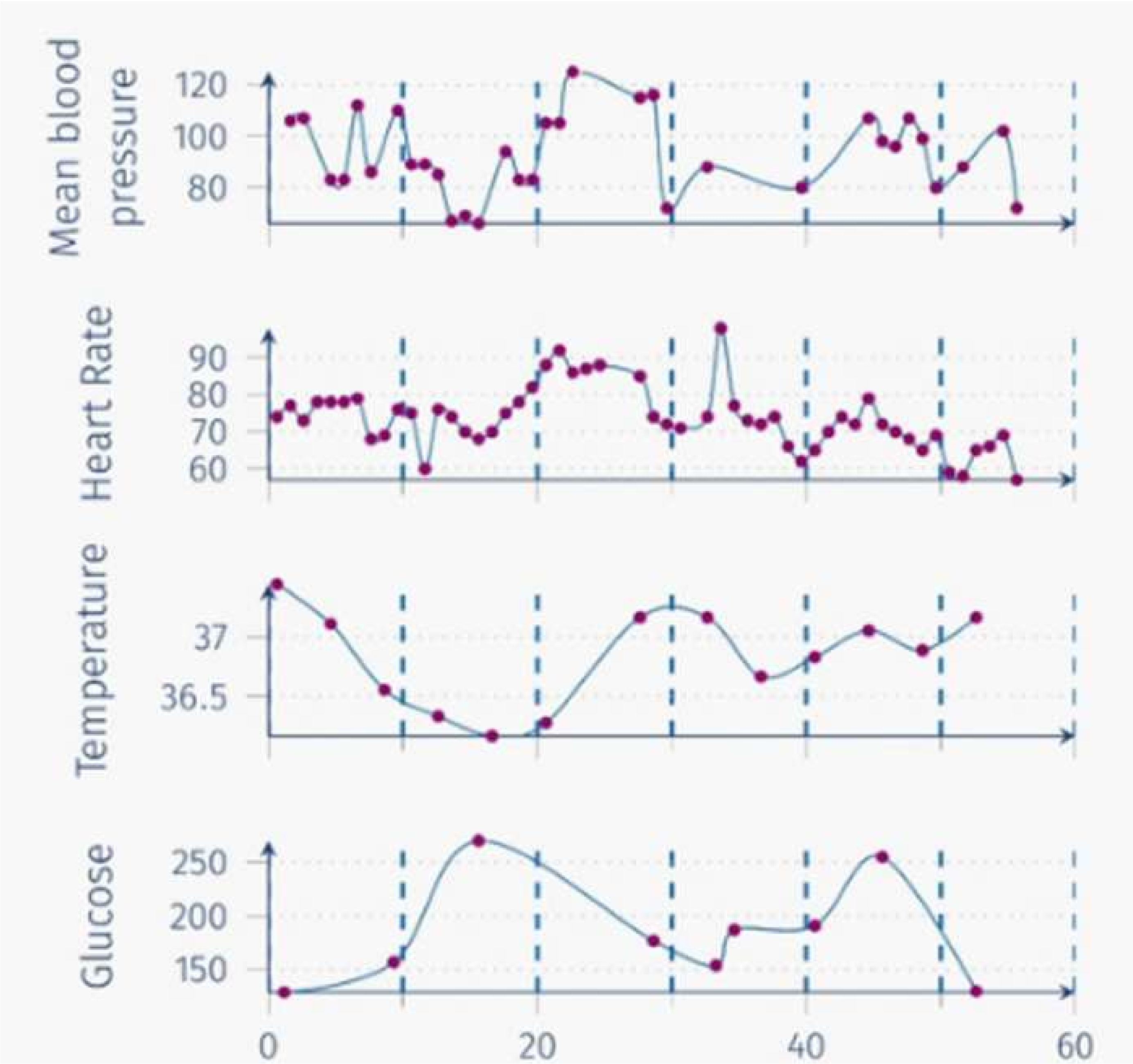


Time	Temperature
5:00 am	59 °F
6:00 am	59 °F
7:00 am	58 °F
8:00 am	58 °F
9:00 am	60 °F
10:00 am	62 °F

Time	Temperature	cloud cover	dew point	humidity	wind
5:00 am	59 °F	97%	51 °F	74%	8 mph SSE
6:00 am	59 °F	89%	51 °F	75%	8 mph SSE
7:00 am	58 °F	79%	51 °F	76%	7 mph SSE
8:00 am	58 °F	74%	51 °F	77%	7 mph S
9:00 am	60 °F	74%	51 °F	74%	7 mph S
10:00 am	62 °F	74%	52 °F	70%	8 mph S



Example of multivariate time series with measurements that are taken at a different frequency and irregular times, probably with missing values.



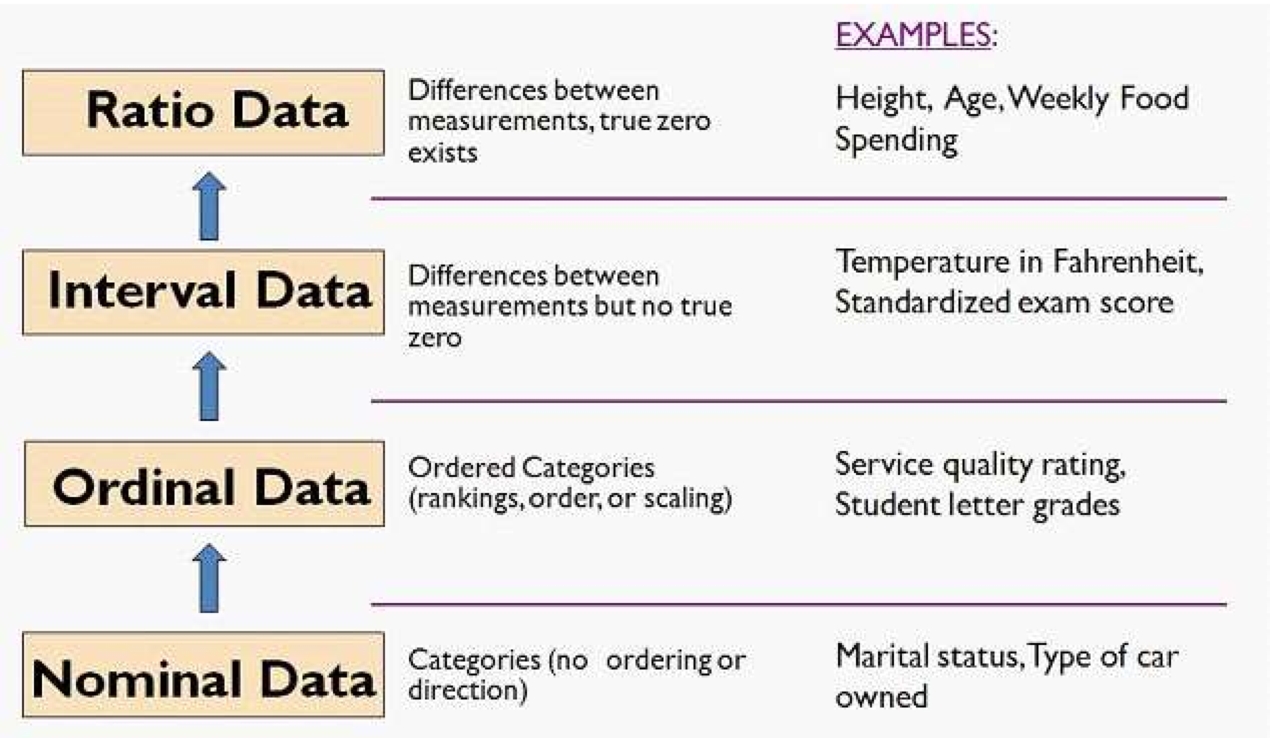
In statistics and Machine Learning, data types are also called levels of measurement.

Four common ones are used:

- **Interval**: Interval scales provide information about order, and also ability to compare ranges; *eg. temperature measured either on a Fahrenheit or Celsius scale: measured in Fahrenheit units, the difference between a temperature of 46 and 42 is the same as the difference between 72 and 68.*

- **Nominal**: This is qualitative, not quantitative; *eg. Religious Preference: 1 = Buddhist, 2 = Muslim, 3 = Christian, 4 = Jewish, 5 = Other.*
- **Ordinal**: An ordinal scale that indicates ordering or direction in addition to providing nominal information; *eg. Low/Medium/High or Faster/Slower* are examples of ordinal levels of measurement.

  Ranking an experience as a "nine" on 1-10 scale tells us that it was higher than an experience ranked as a "six".
- **Ratio**: In addition to possessing the qualities of nominal, ordinal, and interval scales, a ratio scale has an absolute zero, a point where none of the quality being measured exists; *eg. income, years of work experience, number of children.*



General about time series model

- Try to make some initial hypothesis about way of analysis based on the task.
- Try to visualize time series.
- Make an assumption about exogenous factors, univariate or multivariate consideration of time series.
- Make a one or several assumption about linearity of the model.
- Make a one or several assumption about description of the model (trend, seasonal parts, dumped trend, saturation, inflection points and e.t.c.).
- Make an assumption about linearity of the model component relation.
- Make an assumption about stationarity of the model.
- Check your hypothesis to find the best suit one.

1.4 Time Series Analysis Tasks

Main time series tasks:

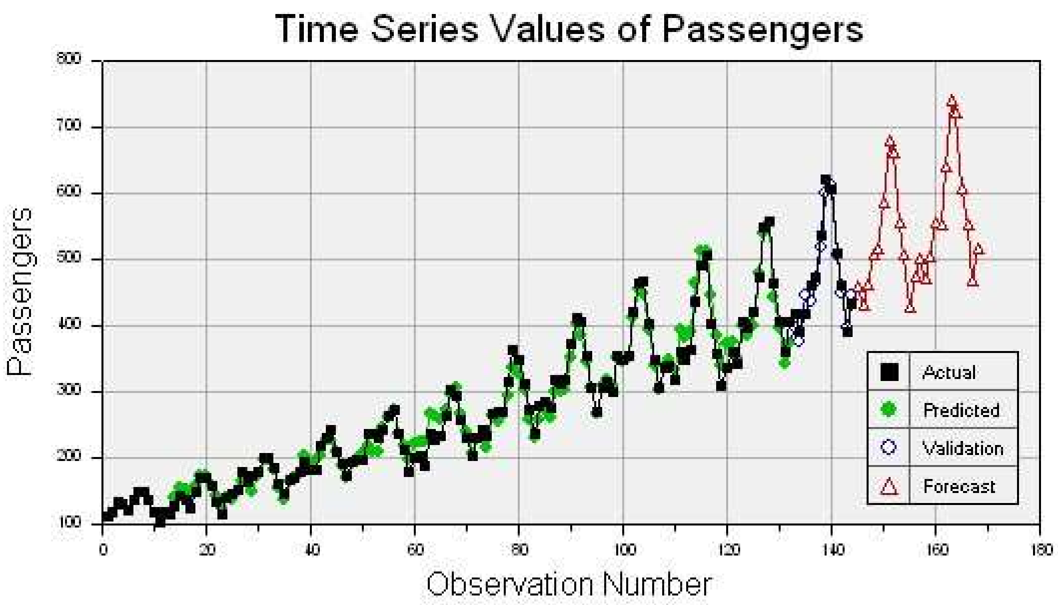
- forecast future values in the series.
- series parameter estimation (for instance, the period of seasonality (or cyclicity), the slope of the trend).
- anomaly detection.
- patterns in series recognition and classification.
- time-series segmentation

- incomplete series reconstruction (if some data are missing).
- time series pattern clustering (unsupervised task).

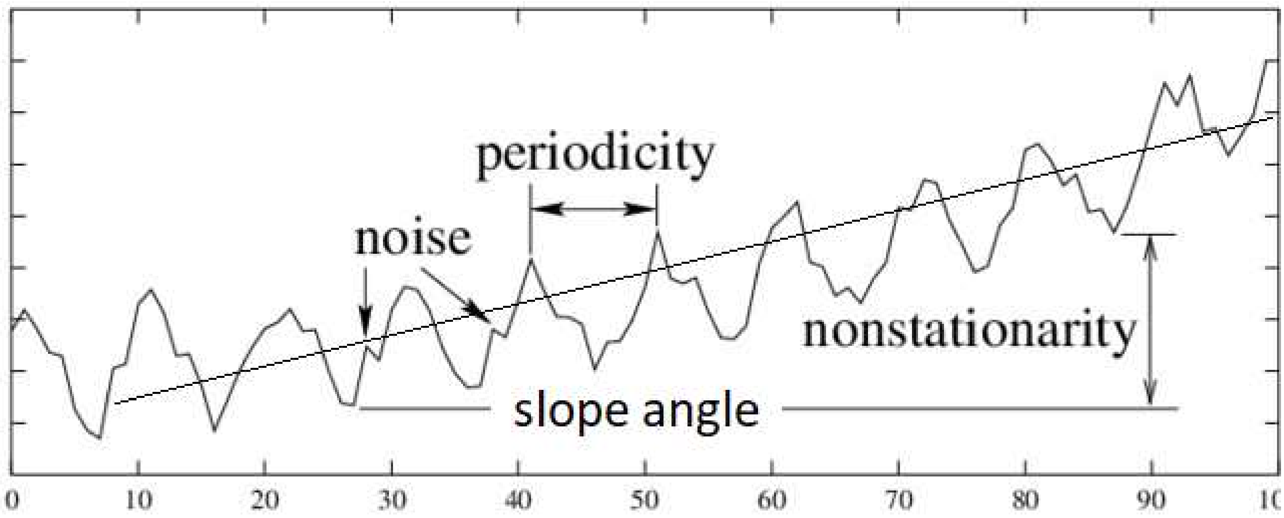
Auxiliary tasks:

- series denoising.
- series transformation (non-linear transformation for better representation of time series, the simplest are derivative, logarithm transforms).
- series decomposition (some components extraction, for instance complex trend reconstruction).
- series re-discretization and re-quantization (for instance, if we have series with ununiform steps or for dimension reduction).

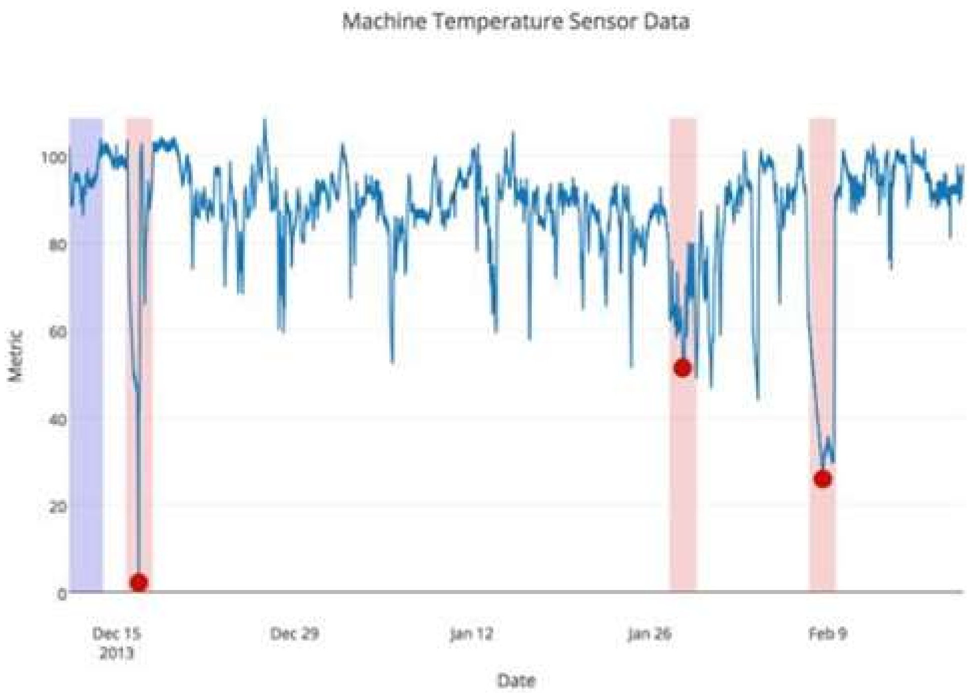
Forecast (predict) future values in the series



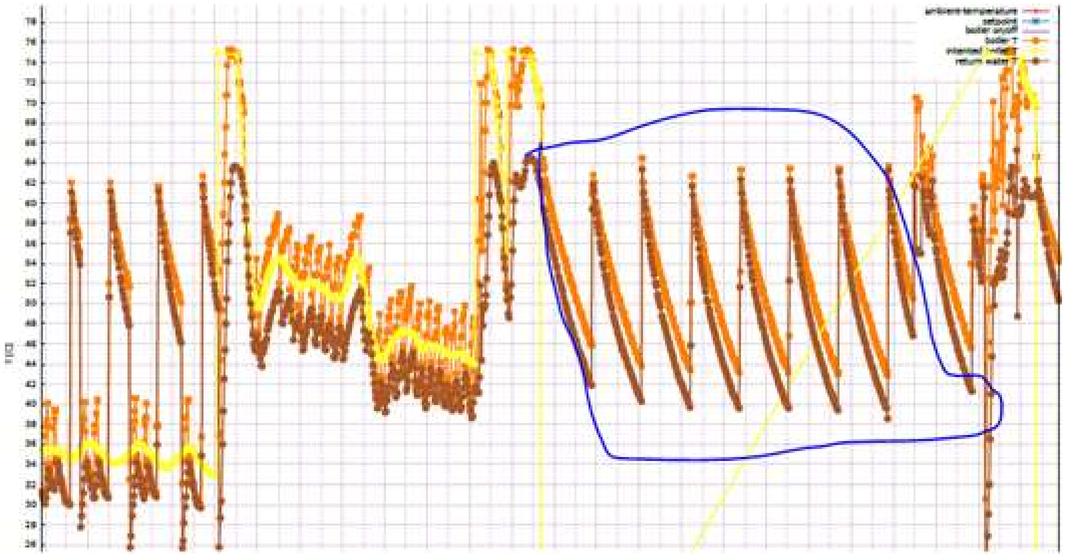
Series parameter estimation.



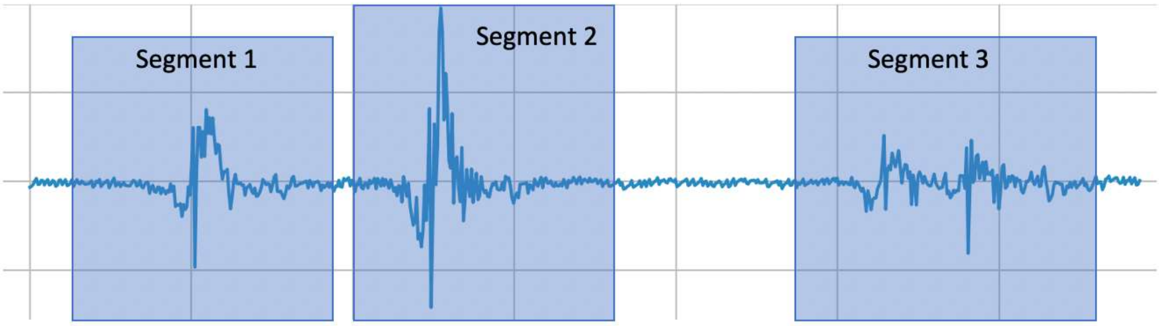
Anomaly detection



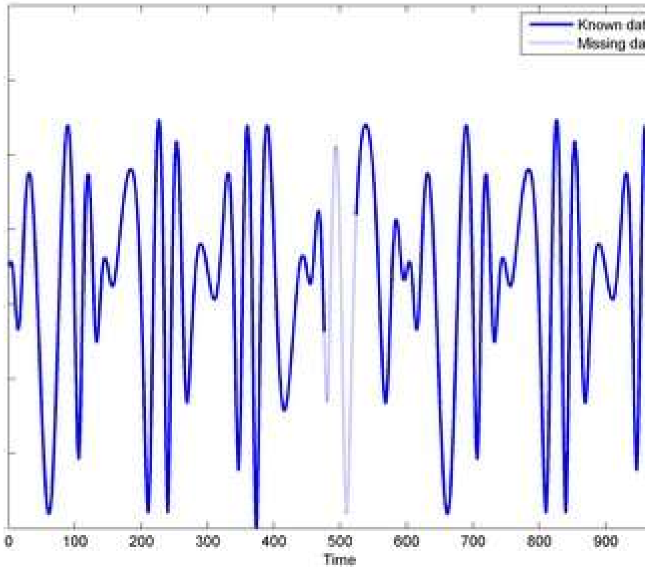
pattern recognition and classification in data



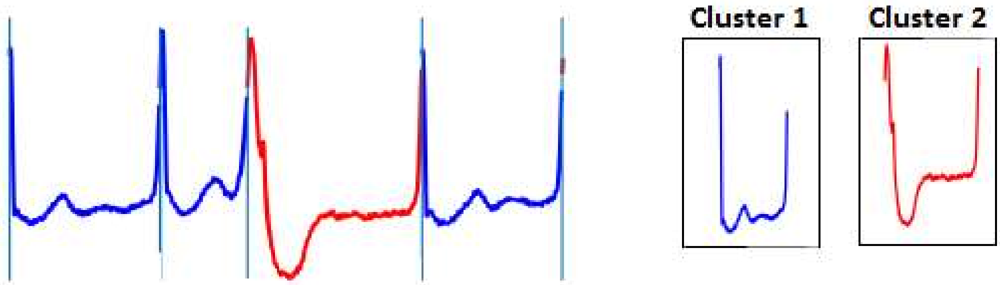
Time series segmentation



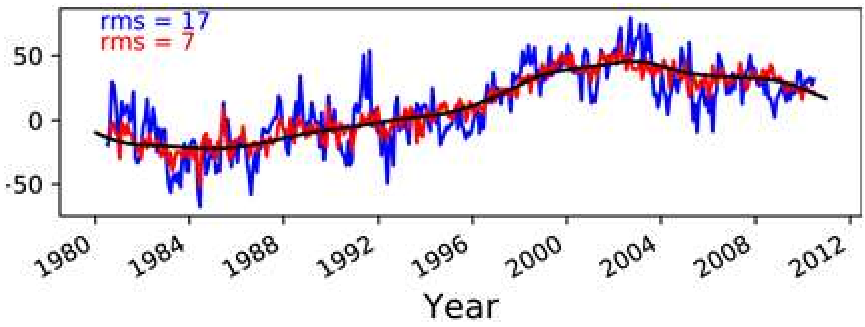
Incomplete time series reconstruction



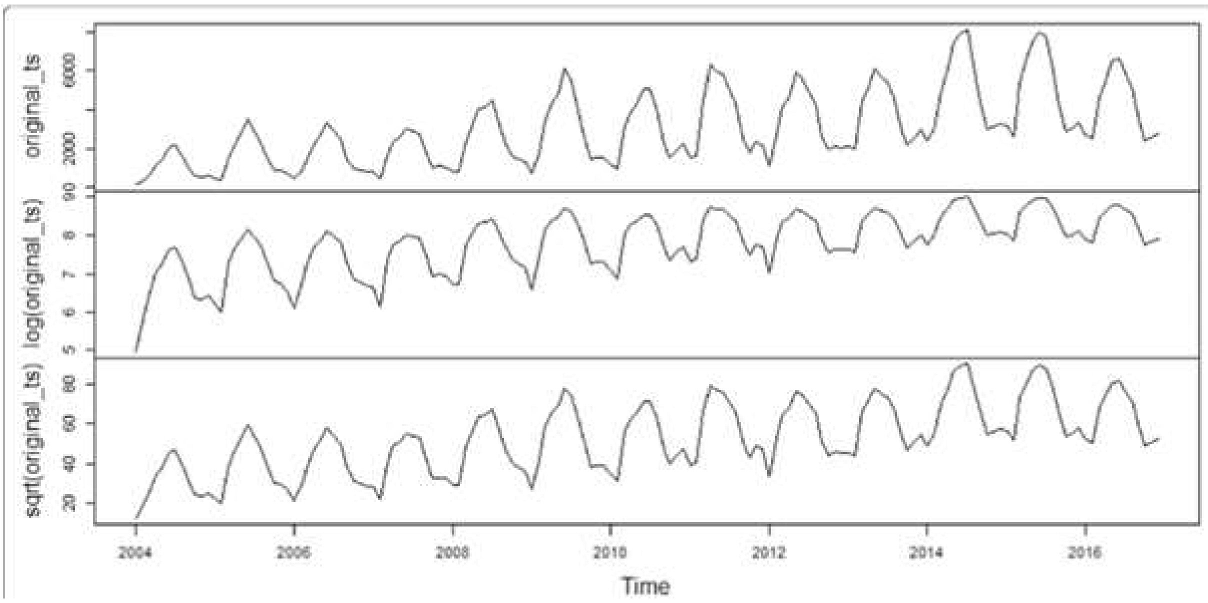
Time series pattern clustering



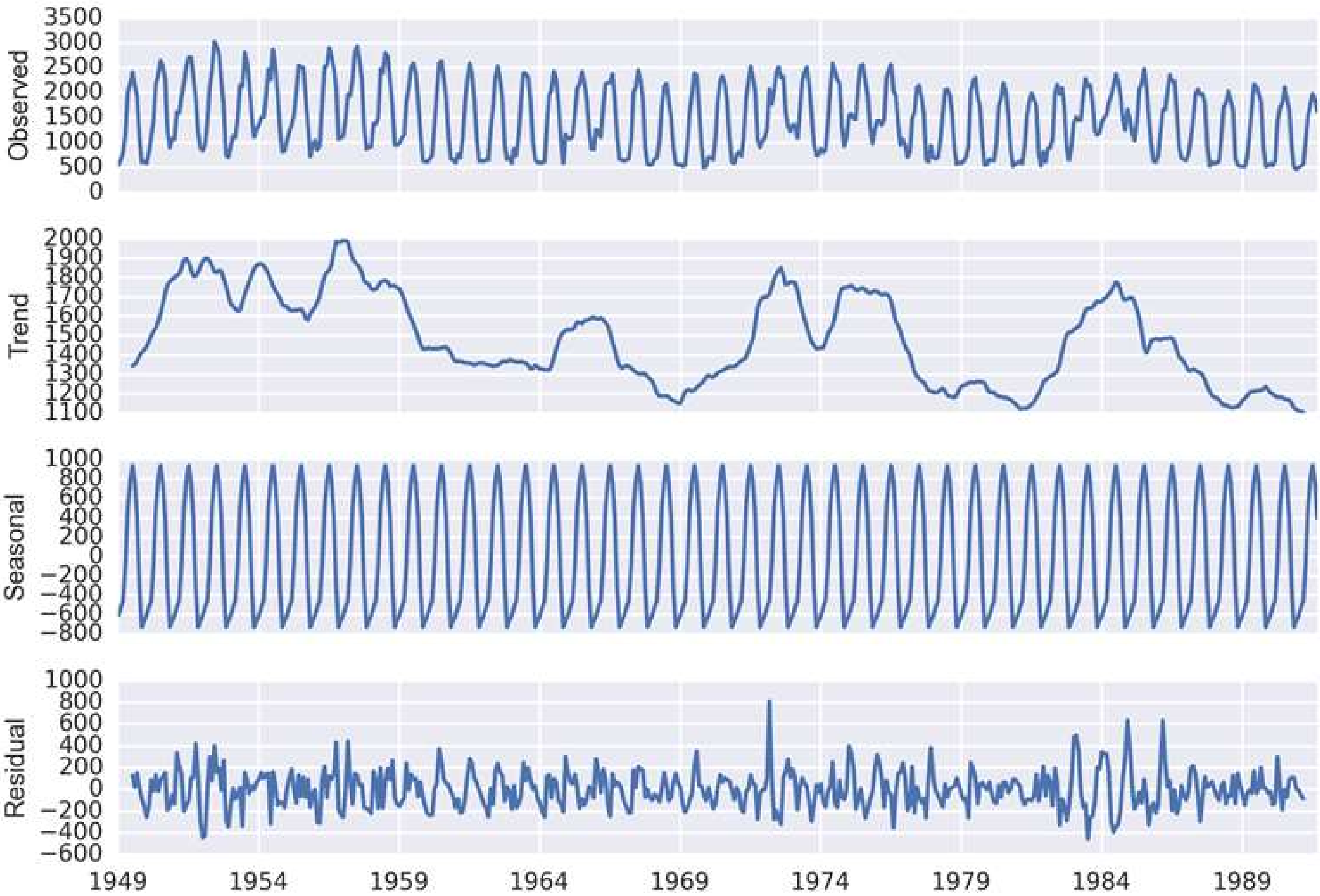
Time series denoiseing



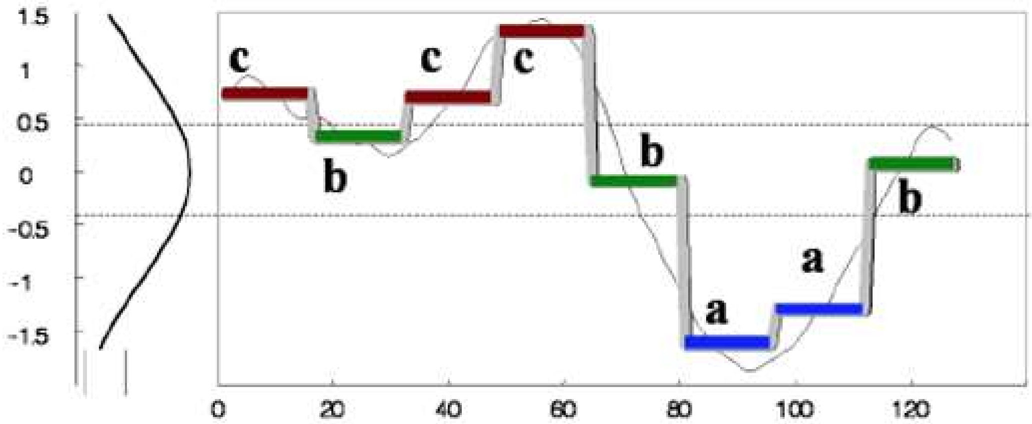
Time series transformation



Series Decomposition (some components extraction, for instance complex trend reconstruction)



Time series dimension reduction



1.5 Classification of Methods of the Time Series Analysis .

1. **Linear models**:

   A. Linear regression (or regression task reduced to linear, for instance polynomial regression).

   B. **AutoRegression-Moving Average (ARMA and its variants, including ARIMA, SARIMAX, ARIFMA, AR, MA, VAR and e.t.c.).**

   C. ANalysis Of VAriance (ANOVA and its variants, including F-Test, Multivariate ANOVA, and e.t.c.).

   D. **Filtration models, including Winer filter, Kalman fliter, linear adaptive filters, moving average models and e.t.c.**

   E. **Decomposition models, including Prony decomposition, Principle component analysis (PCA), Matrix penciling, Dynamic mode decomposition (DMD), singular signal analysis (SSA) and e.t.c.**

F. Time-frequency models (short-time Fourier transform, fractional Fourier transform and e.t.c.).

2. **Non-Linear parameter-based models:**

   A. Generalized AutoRegressive Conditional Heteroskedasticity (GARCH and its variants, including ARCH, and e.t.c.).

   B. Non-linear AutoRegression models (NAR).

   C. **Intrinsic mode decomposition models (empirical mode decomposition (EMD), variation mode decomposition (VMD), hilbert vibration transform (HVT), emperical wavelet transfrom (EWT) and e.t.c.)**.

   D. Time-pseudo-frequency models (wavelet transform, e.t.c.).

3. **Machine learning models (Data-driven models):**

   A. Classical supervised methods (knn, trees, boosting, svm, and e.t.c.).

   B. Deep neural networks (recurent networks, 1d-and 2d -convolution networks, and e.t.c.).

   C. **Non-linear unsupervised methods (autoencoders, gaussian-mixture models, nonlinear pca, t-sne, and e.t.c.).**

Note

that in other classification we can classify some of the methods above into **harmonic analysis** class:

- Decomposition models, including Prony decomposition, Principle component analysis (PCA), Matrix penciling, Dynamic mode decomposition (DMD), singular signal analysis (SSA) and e.t.c.
- Time-frequency models (short-time Fourier transform, fractional Fourier transform and e.t.c.).
- Intrinsic mode decomposition models (empirical mode decomposition (EMD), variation mode decomposition (VMD), hilbert vibration transform (HVT), emperical wavelet transfrom (EWT) and e.t.c.).
- Time-pseudo-frequency models (wavelet transform, e.t.c.).

In [ ]:

```

```