# Mining Time Series<br>
La Torge & Ta Rudolecky<br>
La Torge & Ta Rudolecky Mining Time Series<br>L. Torgo & T. Rudolecky<br>L. Torgo & T. Rudolecky

# Time Series Introduction<br>A Definition

## **Definition**

- A time series is a set of observations of a variable thatare Time Series Introduction<br>
Finition<br>
A time series is a set of observations of a variable that are<br>
ordered by time.<br>
E.g.,<br>  $x_1, x_2, \cdots, x_{t-1}, x_t, x_{t+1}, \cdots, x_n$ <br>
where x, is the observation of variable X at time  $t$
- 

 $x_1, x_2, \cdots, x_{t-1}, x_t, x_{t+1}, \cdots, x_n$ 

where  $x_t$  is the observation of variable X at time  $t$ .

A multivariate time series is a set of observations of a set of efinition<br>A time series is a set of observations of a variable that are<br>ordered by time.<br>E.g.,<br> $x_1, x_2, \dots, x_{t-1}, x_t, x_{t+1}, \dots, x_n$ <br>where  $x_t$  is the observation of variable X at time  $t$ .<br>A multivariate time series is a set ervations of a variable that are<br>  $\frac{K_{n}}{n}$ <br>
of variable X at time *t*.<br>
a set of observations of a set of<br>
od of time.<br>
Time Series<br>
2/45



Goals

# **Explanation**

Obtaining a Time Series Model help us to have <sub>Goals</sub><br>Deper Understanding of the Mechanism<br>The Deeper Understanding of the Mechanism<br>The Mechanism<br>The Series Data. Coals<br>Containing a Time Series Model help us to have<br>a Deeper Understanding of the Mechanism<br>that Generated the Observed Time Series Data. ries Model help us to have<br>anding of the Mechanism<br>bbserved Time Series Data.<br>Time Series<br>3/45





## Forecasting





Goals

<sub>Goals</sub><br>Time Series Data Mining<br>Main Time Series Data Mining Tasks

- o<sub>oals</sub><br>Main Time Series Data Mining<br>Main Time Series Data Mining Tasks<br>■ *Indexing (Query by Content)*<br>Given a query time series Q and a similarity measure *D*(Q, *X*) <sup>Goals</sup><br>
Indexing (Query by Content)<br>
Indexing (Query by Content)<br>
Given a query time series Q and a similarity measure *D*(Q, *X*)<br>
find the most similar time series in a database **D**<br>
Clustering <sup>Gives</sup><br>
Given a Given a Mining Tasks<br> *Indexing (Query by Content)*<br>
Given a query time series Q and a similarity measure  $D(Q, X)$ <br>
find the most similar time series in a database **D**<br> *Clustering* find the most similar time series in a database D some similarly conditions of the Series Data Mining<br>
1 Time Series Data Mining Tasks<br>
Indexing (Query by Content)<br>
Given a query time series Q and a similarity measure  $D(Q, X)$ <br>
find the most similar time series in a datab Exercise Data Mining<br>
Time Series Data Mining Tasks<br>
Indexing (Query by Content)<br>
Given a query time series Q and a similarity measure  $D(Q, X)$ <br>
find the most similar time series in a database **D**<br>
Clustering<br>
Find the nat Time Series Data Mining Tasks<br>
Indexing (Query by Content)<br>
Given a query time series Q and a similarity measure  $D(Q, X)$ <br>
find the most similar time series in a database **D**<br>
Clustering<br>
Find the natural goupings of a set A and a similarity measure  $D(Q, X)$ <br>
Fires in a database **D**<br>
Fa set of time series in a database **D**<br>
Inter  $D(Q, X)$ <br>
Tries Q, assign it a label C from a set of
	- Clustering

Find the natural goupings of a set of time series in a database D

**Classification** 





- Standard descriptive statistics (mean, standard deviation, etc.) do
- 
- So, for proving summaries of TS data we will be interested in concepts like trend, seasonality and correlation between n marries of Time Series Data<br>Standard descriptive statistics (mean, standard deviation, etc.) do<br>not allways work with time series (TS) data.<br>TS may contain trends, seasonality and some other systematic<br>components, making cs (mean, standard deviation, etc.) do<br>eries (TS) data.<br>sonality and some other systematic<br>stats misleading.<br>of TS data we will be interested in<br>ality and correlation between<br>ne TS.





# Exploratory Analysis Variation<br>Types of Variation<br>Secondary Analysis

Exploratory Analysis Variation<br>Types of Variation<br>Seasonal Variation<br>Some time series exhibit a variation that is annual in period<br>demand for ice cream. Some time series exhibit a variation that is annual in period, e.g. Exploratory Analysis Variation<br>
Types of Variation<br>
Seasonal Variation<br>
Seasonal Variation<br>
Seme time series exhibit a variation that is annual in period, e.g<br>
Other Cyclic Variation<br>
Seme time series beys periodic veriati

Exploratory Analysis Variation<br>
Types of Variation<br>
Seasonal Variation<br>
Seasonal Variation<br>
Some time series exhibit a variation that is annual in period, e<br>
Other Cyclic Variation<br>
Some time series have periodic variation Some time series have periodic variations that are not related to Exploration<br>Seasonal Variation<br>Some time series exhibit a variation that is annual in period, e.g.<br>demand for ice cream.<br>Other Cyclic Variation<br>Some time series have periodic variations that are not related to<br>seasons but Seasonal Variation<br>Some time series exhibit a variation that is annual in period, e.g.<br>demand for ice cream.<br>Other Cyclic Variation<br>Some time series have periodic variations that are not related to<br>seasons but to other fac tion that is annual in period, e.g.<br>variations that are not related to<br>some economic time series.<br>the mean level of the time series.<br> $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ <br>Time Series

## **Trends**



## **Stationarity**

Exploratory Analysis<br>
Stationarity<br>
An Informal Definition<br>
A time series is said to be stationary if<br>
there is no systematic change in mean (no trend),<br>
if there is no systematic change in variance and Exploratory Analysis Stationarity<br>
Stationarity<br>
An Informal Definition<br>
A time series is said to be stationary if<br>
there is no systematic change in mean (no trend),<br>
if there is no systematic change in variance and<br>
if st Exploratory Analysis Stationarty<br>Therman is no systemation<br>there is no systematic change in mean (no trend),<br>there is no systematic change in variance and<br>if strictly periodic variations have been removed. Exploratory Analysis Stationarity<br>
informal Definition<br>
me series is said to be stationary if<br>
there is no systematic change in mean (no trend),<br>
if there is no systematic change in wariance and<br>
if strictly periodic varia Exploratory Analysis Stationarity<br>
Informal Definition<br>
me series is said to be stationary if<br>
there is no systematic change in mean (no trend),<br>
if there is no systematic change in variance and<br>
if strictly periodic varia Stationarity<br>An Informal Definition<br>A time series is said to be stationary if<br>there is no systematic change in mean (no trend),<br>if there is no systematic change in variance and<br>if strictly periodic variations have been rem nary if<br>
e in mean (no trend),<br>
ige in variance and<br>
ave been removed.<br>
s like mean, standard deviation,<br>
rmation!<br>
Time Series<br>
8/45

Note that in these cases statistics like mean, standard deviation,



# Exploratory Analysis Stationarity<br>
Stationarity<br>
The Contract of the Contract<br>
The Contract of the Contract of the Contract of the Contract of

# **Stationarity**



# Exploratory Analysis Time Plots



- **Ploting the time series** values against time is one of the most important tools for Ploting the time series<br>values against time is<br>one of the most<br>important tools for<br>analysing its behaviour.<br>Time plots show<br>important features like<br>trands assessity. Ploting the time series<br>values against time is<br>one of the most<br>important tools for<br>analysing its behaviour.<br>Time plots show<br>important features like<br>trends, seasonality,<br>outliers and<br>discontinuities.
- Time plots show trends, seasonality, outliers and From the most be most trime series<br>values against time is<br>one of the most<br>important tools for<br>analysing its behaviour.<br>Time plots show<br>important features like<br>trends, seasonality,<br>outliers and<br>discontinuities.



Exploratory Analysis Transformations - I<br>|
| ILION dest transformations : Exploratory Analysis Transformations - I<br>
Transformations - I<br>
Plotting the data may suggest transformations :<br>
To stabilize the variance

Exploratory Analysis Transformations - I<br>Plotting the data may suggest transformations :<br>To stabilize the variance<br>Symptoms: trend with the variance increasing with the mean. Exploratory Analysis Transformations - I<br>Transformations - |<br>Plotting the data may suggest transformations :<br>To stabilize the variance<br>Symptoms: trend with the variance increasing with the mean.<br>Solution: logarithmic trans Exploration Section Transformations - I<br>
Symptoms: To stabilize the variance<br>
Symptoms: trend with the variance increasing with the mean.<br>
Solution: logarithmic transformation.<br>
To make the seasonal effects additive Exploratory Analysis Transformations - I<br>Transformations - I<br>Plotting the data may suggest transformations :<br>To stabilize the variance<br>Symptoms: trend with the variance increasing with the mean.<br>Solution: logarithmic trans

Exploration S - |<br>
Fransformations - |<br>
Plotting the data may suggest transformations :<br>
To stabilize the variance<br>
Symptoms: trend with the variance increasing with the mean.<br>
Solution: logarithmic transformation.<br>
To mak Symptoms: there is a trend and the size of the seasonal effect Exploratory Analysis Transformations - I<br>
Plotting the data may suggest transformations :<br>
To stabilize the variance<br>
Symptoms: trend with the variance increasing with the mean.<br>
Solution: logarithmic transformation.<br>
To m Exploratory Analysis Transformations -1<br>
Fransformations - |<br>
Plotting the data may suggest transformations :<br>
To stabilize the variance<br>
Symptoms: trend with the variance increasing with the mean.<br>
Solution: logarithmic t **Transformations - I**<br>
Plotting the data may suggest transformations :<br>
To stabilize the variance<br>
Symptoms: trend with the variance increasing with the mea<br>
Solution: logarithmic transformation.<br>
To make the seasonal eff

*Examplems:* trend with the variance increasing with the *Ilution:* logarithmic transformation.<br> **Invertigary in the Seasonal effects additive**<br> *Imptoms:* there is a trend and the size of the seasonality.<br> *Ilution:* log Plotting the data may suggest transformations :<br>
To stabilize the variance<br>
Symptoms: trend with the variance increasing with the mean.<br>
Solution: logarithmic transformation.<br>
To make the seasonal effects additive<br>
Sympto Solution 1: first order differentiation ( $\mathbb{X}_t = X_t - X_{t-1}$ ). Solution 2: model the trend and subtractit from the original series  $(Y = X_t - r_t)$ . ce increasing with the mean.<br>
Station.<br>
additive<br>
the size of the seasonal effect<br>
cative seasonality).<br>
Extends the mean.<br>
Station ( $\mathbb{X}_t = X_t - X_{t-1}$ ).<br>
subtractit from the original series<br>
Time Series 11/45





<sup>omations</sup> -an example (1)<br>
<br>
An example time series with<br>
trend and a multiplicative<br>
seasonality effect. trend and a multiplicative Formations - an example (1)<br>
straight (1)<br>
An example time series with<br>
trend and a multiplicative<br>
seasonality effect. An example time series with<br>trend and a multiplicative<br>seasonality effect.



# Exploratory Analysis Transformations - an example (2)<br>a simple example (2) Exploratory Analysis Transformations - an example (2)<br>Transformations - a simple example  $(2)$ <br>y<sub>i-x-(7.708+42.521×0)</sub>













# Exploratory Analysis Randomness<br>Tests of Randomness<br>Erequently we want to test the bypothesis that the obs

Exploratory Analysis<br>Frequently we want to test the hypothesis that the observed time<br>series is random. series is random. Exploratory Analysis Randomness<br>
A possible way is to inspect the correlogram.<br>
A possible way is to inspect the correlogram.<br>
A possible way is to inspect the correlogram.<br>
Ernative, which is frequently used, is the runs Exis of Randomness<br>
Frequently we want to test the hypothesis that the observed time<br>
series is random.<br>
A possible way is to inspect the correlogram.<br>
An alternative, which is frequently used, is the runs test.<br>
This test

of  $x_t$  is above (below) the median value of the series, or the number of times there is a sucession of monotonically increasing (decreasing) ant to test the hypothesis that the observed time<br>series is random.<br>So series is random.<br>So series the correlogram.<br>So, which is frequently used, is the runs test.<br>Abecks for things like the number of times the value<br>of th S is random.<br>
D inspect the correlogram.<br>
Equently used, is the runs test.<br>
Ings like the number of times the value<br>
In value of the series, or the number of<br>
nonotonically increasing (decreasing)<br>
Series and so on.



Exploratory Analysis Check List<br>Data<br>Analysis Check List

Exploratory Analysis Check List<br>Handling Real World Data<br>A Check List of Common Sense Things to Do Exploratory Analysis Check List<br>
A Check List of Common Sense Things to Do<br>
(taken from Chatfield, 2004)<br>
■ Do you understand the context? Have the right variables been

- Exploratory Analysis<br>
Handling Real World Data<br>
A Check List of Common Sense Things to Do<br>
(taken from Chatfield, 2004)<br>
Do you understand the context? Have the right variables been<br>
measured? Do you understand the context? Have the right variables been measured? Exploratory Analysis<br>
Exploratory Analysis<br>
Heck List of Common Sense Things to Do<br>
Pho you understand the context? Have the right variables been<br>
measured?<br>
Have all the time series been plotted?<br>
Are there any outliers? Espleratory Analysis Check List<br>neck List of Common Sense Things to Do<br>nen from Chatfield, 2004)<br>Do you understand the context? Have the right variables been<br>measured?<br>Have all the time series been plotted?<br>Are there missi reck List of Common Sense Things to Do<br>
Do you understand the context? Have the right variables been<br>
measured?<br>
Have all the time series been plotted?<br>
Are there missing values? If so, what should be done about them?<br>
Are neck List of Common Sense Things to Do<br>
an from Chatfield, 2004)<br>
Do you understand the context? Have the right variables been<br>
measured?<br>
Have all the time series been plotted?<br>
Are there missing values? If so, what shoul **ISTANG INTERT IS SEASURE IN SEAST AND MONOR INTER IS SEVERTHERT AND NO YOU understand the context? Have the right variables been measured?**<br>Have all the time series been plotted?<br>Are there missing values? If so, what shou ext? Have the right variables been<br>
n plotted?<br>
So, what should be done about them?<br>
The series 19 of what do they mean?<br>
So what do they mean?<br>
So what be done about it?<br>
Should be done about it?<br>
Time Series<br>
Time Series
	-
	- Are there missing values? If so, what should be done about them?
	-
	- Are there any discontinuities? If so, what do they mean?
	-
	-
	-

# w<sub>hy?</sub><br>Measuring Similarity<br>Why?

## Why?

Most time series data mining tasks require the similarity between <sup>why?</sup><br>Measuring Similarity<br>Most time series data mining tasks require the similarity between<br>series to be asserted (e.g. indexing, clustering, classification, etc.).<br>Types of matching why?<br>
Measuring Similarity<br>
Why?<br>
Most time series data mining tasks require the similarity b<br>
series to be asserted (e.g. indexing, clustering, classificat<br>
Types of matching<br>
There are essentially two variants of similar why?<br>
Why?<br>
Most time series data mining tasks require the similarity between<br>
series to be asserted (e.g. indexing, clustering, classification, etc.).<br>
Types of matching<br>
There are essentially two variants of similarity m <sup>Why?</sup><br>Susuring Similarity<br>Where time series data mining tasks require the similarity be<br>as to be asserted (e.g. indexing, clustering, classification<br>whole matching<br>Whole matching<br>Where the query time series is matched (as Fraction Similarity<br>
The series data mining tasks require the similarity between<br>
time series data mining tasks require the similarity between<br>
Subsequence are essentially two variants of similarity matching:<br>
Whole matchi  $\frac{1}{2}$ <br>  $\frac{1}{2}$ <br> <sup>2</sup><br>subsection matching the similarity between<br>subsection matching<br>example are essentially two variants of similarity matching:<br>whole matching<br>where the query time series is matched (as a whole) against all<br>time series in

Why?

where the query time series is matched (as a whole) against all ing, clustering, classification, etc.).<br>
Its of similarity matching:<br>
is matched (as a whole) against all<br>
.<br>
data base are searched for a<br>
ne query subsequence.<br>
Time Series<br>
20/45

where all time series in the data base are searched for a

Distance Measures<br>
e Function<br>
" Distance Measures<br>
Defining a Distance Function<br>
Distance (or dissimilarity) functions<br>
Civen any two time series suand se their distance (or dissimilarity) is

Distance Measures<br>Defining a Distance Function<br>Distance (or dissimilarity) functions<br>Given any two time series  $s_1$  and  $s_2$  their distance (or dissimilarity) is<br>denoted by  $D(s_1, s_2)$ . Given any two time series  $s_1$  and  $s_2$  their distance (or dissimilarity) is Defining a Distance Function<br>
Distance (or dissimilarity) functions<br>
Given any two time series  $s_1$  and  $s_2$  their distance (or dissimilarity) is<br>
denoted by  $D(s_1, s_2)$ .<br>
Desirable Properties of a Distance Function<br>
Sy Distance Measures<br>
Distance Function<br>
ance (or dissimilarity) functions<br>
n any two time series  $s_1$  and  $s_2$  their distance (or dissimilated by  $D(s_1, s_2)$ .<br>
Firable Properties of a Distance Function<br>
Symmetry<br>  $D(X, Y) = D$ Distance Measures<br>
ining a Distance Function<br>
ance (or dissimilarity) functions<br>
in any two time series  $s_1$  and  $s_2$  their distance (or dissimilarity)<br>
ited by  $D(s_1, s_2)$ .<br>
irable Properties of a Distance Function<br>
Sy Fining a Distance Tunction<br>
ance (or dissimilarity) functions<br>
in any two time series  $s_1$  and  $s_2$  their distance (or dissimilated by  $D(s_1, s_2)$ .<br>
irable Properties of a Distance Function<br>
Symmetry<br>  $D(X, Y) = D(Y, X)$ <br>
Con In any two time series  $s_1$  and  $s_2$  their distance (or dissimilarit<br>
of the by  $D(s_1, s_2)$ .<br>
Firable Properties of a Distance Function<br>
Symmetry<br>  $D(X, Y) = D(Y, X)$ <br>
Constancy of Self-Similarity<br>  $D(X, X) = 0$ <br>
Positivity<br>  $D(X$ ance (or dissimilarity) functions<br>
in any two time series  $s_1$  and  $s_2$  their distance (or dissimilated by  $D(s_1, s_2)$ .<br>
irable Properties of a Distance Function<br>
Symmetry<br>  $D(X, Y) = D(Y, X)$ <br>
Constancy of Self-Similarity<br> Time Series<br>21/45<br>21/45<br>21/45

■ Symmetry

- 
- **Positivity**
- Firable Properties of a Distance Function<br>
Symmetry<br>  $D(X, Y) = D(Y, X)$ <br>
Constancy of Self-Similarity<br>  $D(X, X) = 0$ <br>
Positivity<br>  $D(X, Y) = 0$  iff  $X = Y$ <br>
Triangular Inequality<br>  $D(X, Y) \ge D(X, Z) + D(Y, Z)$ <br>
Time Series<br>
Time Series

Distance Measures<br>Functions<br>Functions Distance Measures<br>Types of Distance Functions

Distance Functions<br>
Metric - satisfy all properties<br>
e.g. Euclidean, correlation, etc.<br>
Non-metric - do not satisfy any of the properties<br>
Non-metric - do not satisfy any of the properties Distance Measures<br>
e.g. Of Distance Functions<br>
Metric - satisfy all properties<br>
e.g. Euclidean, correlation, etc.<br>
Non-metric - do not satisfy any of the properties<br>
e.g. time warping, LCSS, etc. Distance Measures<br>
Distance Functions<br>
Metric - satisfy all properties<br>
e.g. Euclidean, correlation, etc.<br>
Non-metric - do not satisfy any of the properties<br>
e.g. time warping, LCSS, etc. Distance Measures<br>
E.g. of Distance Functions<br>
E.g. Euclidean, correlation, etc.<br>
Mon-metric - do not satisfy any of the properties<br>
e.g. time warping, LCSS, etc. s<br>etc.<br>C.<br>Time Series 22/45



Distance Measures Minkowski Metrics<br>
Trics<br>
United States Minkowski Metrics<br>
United States

# Distance Measures Minkowski Metrics<br>The Minkowski Metrics<br>Allowski Metrics

$$
D(X, Y) = \left(\sum_{i=1}^k (x_i - y_i)^p\right)^{\frac{1}{p}}
$$

City Block  $(p = 1)$ Euclidean ( $p = 2$ )

$$
D(X, Y) = \left(\sum_{i=1}^{k} (x_i - y_i)^p\right)^p
$$
  

$$
D(X, Y) = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2}
$$
  

$$
Time Series
$$
  

$$
23/45
$$

Distance Measures Correlation<br>
Distance Measures Correlation<br>
Distance Measures Correlation **Distance Measures Correlation**<br>Correlation between two time series

$$
\rho_{x,y} = \frac{\sum_{t=1}^{N} (x_t - \bar{x})(y_t - \bar{y})}{N \sigma_x \sigma_y}
$$

$$
=\frac{\sum_{t=1}^{N}(x_t - x)(y_t - y)}{N\sigma_x \sigma_y}
$$
  
where  

$$
\rho_{x,y} = \frac{1}{N} \sum_{t=1}^{N} x_t y_t
$$

# Distance Measures Dynamic Time Warping

Distance Measures Dynamic Time Warping<br>Dynamic Time Warping - introduction<br>Dynamic Time Warping (DTW) is a non-metric distance function.<br>Main Ideas of DTW Dynamic Time Warping - introduction<br>Dynamic Time Warping (DTW) is a non-metric distance function.<br>Main Ideas of DTW<br>- Allow for local deformations (stretch and shrink) along the time

- Distance Measures Dynamic Time Warping<br>
Dynamic Time Warping introduction<br>
Dynamic Time Warping (DTW) is a non-metric distance function<br>
Main Ideas of DTW<br>
 Allow for local deformations (stretch and shrink) along the t<br> **Allow for local deformations (stretch and shrink) along the time** axis. Distance Measures Dynamic Time Warping<br>
Able Time Warping - introduction<br>
The Warping (DTW) is a non-metric distance function.<br>
Allow for local deformations (stretch and shrink) along the time<br>
axis.<br>
Able to handle series (stretch and shrink) along the time<br>
serent lengths<br>  $T_{\text{time Series}}$ <br>  $\begin{bmatrix}\n\frac{2}{3} \\
25/45\n\end{bmatrix}$ 
	-





Distance Measures Dynamic Time Warping<br>Thing - how to calculate? Distance Measures Dynamic Time Warping<br>Dynamic Time Warping - how to calculate?<br>A Service Structure of the Service Structure of the Service Structure of the Service Structure of the Service









# Distance Measures Dynamic Time Warping LCSS - Longest common subsequence

$$
A = ((a_{x_1,1},\ldots,a_{x_p,1}),\ldots,(a_{x_1,n},\ldots,a_{x_p,n})),
$$
  

$$
B = ((b_{x_1,1},\ldots,b_{x_p,1}),\ldots,(b_{x_1,m},\ldots,b_{x_p,m})).
$$

$$
Head(A) = ((a_{x_1,1},\ldots,a_{x_p,1}),\ldots,(a_{x_1,n-1},\ldots,a_{x_p,n-1})).
$$

$$
B = ((b_{x_1,1},\ldots,b_{x_p,1}),\ldots,(b_{x_1,m},\ldots,b_{x_p,m})).
$$
  
rajectory *A*, let  $Head(A)$  be the sequence:  

$$
Head(A) = ((a_{x_1,1},\ldots,a_{x_p,1}),\ldots,(a_{x_1,n-1},\ldots,a_{x_p,n-1})).
$$
an integer  $\delta$  and a real number  $0 < \epsilon < 1$ , the similarity function  $LCSS_{\delta,\epsilon}(A,B)$  is  
using the recurrent algorithm (4) [32]. *N* and *M* are the size of the sequences *A* respectively at the first step of the recurrent algorithm.  
  
respectively at the first step of the recurrent algorithm.  
  

$$
LCSS_{\delta,\epsilon}(A,B) = \begin{cases} 0 & \text{if } A \text{ or } B \text{ is empty,} \\ 0 & \text{if } A(a_{x_k,n},b_{x_k,m}) < \epsilon, \forall 1 \leq k \leq p, \\ 1 + LCSS_{\delta,\epsilon}(Head(A),Head(B)), \\ \text{and } |n-m| \leq \delta \text{ and } |N-n-M+m| \leq \delta, \\ 0 & \text{and } |n-m| \leq \delta \text{ and } |N-n-M+m| \leq \delta, \\ \text{max } (LCSS_{\delta,\epsilon}(Head(A),B), LCSS_{\delta,\epsilon}(A,Head(B))) \end{cases}
$$
and  
times  $LESS_{\delta,\epsilon}(Head(A),B), LCSS_{\delta,\epsilon}(A,Head(B)))$   
times  $LESS_{\delta,\epsilon}(A,B)$ 



# Distance Measures Dynamic Time Warping LCSS - Longest common subsequence







and (C) distance based on the longest common subsequence (LCSS). we<sup>3</sup>



Goals

# <sub>Goals</sub><br>Goals of an Evaluation Method<br>**The golden rule:**

<sup>Goals</sup><br>als of an Evaluation Method<br>The golden rule:<br>*The data used for evaluating (or comparing) any models c*<br>*be seen during model development.* The data used for evaluating (or comparing) any models cannot be seen during model development.

- 
- Soals<br>
Soals<br>
Soals<br>
Soals<br>
Soals<br>
The golden rule:<br>
The data used for evaluating (or comparing) any models cannot<br>
be seen during model development.<br>
The goal of any evaluation procedure:<br>
Nigh probability of achieving th <sup>Goats</sup><br>
of an Evaluation Method<br>
golden rule:<br> *e data used for evaluating (or comparing) any models cannot<br>
be seen during model development.<br>
Solatin a reliable estimate of some evaluation measure.<br>
High probability of* <sup>Goals</sup><br>
Solden rule:<br>
Solden rule:<br> *e data used for evaluating (or comparing) any models cannot<br>
be seen during model development.<br>
And a reliable estimate of some evaluation measure.<br>
High probability of achieving the s* soals<br>
osals<br>
osals<br>
of an Evaluation Method<br>
golden rule:<br>
e *data used for evaluating (or comparing) any models callel*<br> *be seen during model development.*<br>
Obtain a reliable estimate of some evaluation measure.<br>
High p of an Evaluation Method<br>golden rule:<br>e *data used for evaluating (or comparing) any models can*<br>be seen during model development.<br>goal of any evaluation procedure:<br>Obtain a reliable estimate of some evaluation measure.<br>*Hi* golden rule:<br>
Exploden rule:<br>
A data used for evaluating (or comparing) any models<br>
be seen during model development.<br>
Solatin a reliable estimate of some evaluation measure.<br>
High probability of achieving the same score o golden rule:<br>
e data used for evaluating (or comparing) any models cannot<br>
the seen during model development.<br>
goal of any evaluation procedure:<br>
Contain a reliable estimate of some evaluation measure.<br>
High probability of may any models cannot<br>ting model development.<br>procedure:<br>of some evaluation measure.<br>ing the same score on other samples of<br>...<br>...<br>...<br>Time Series<br>31/45
- **Evaluation Measures** 
	-
	-
	-



# Reliable Estimates<br>Estimates<br>Estimates

- Reliable Estimates<br>
Obtaining Reliable Estimates<br>
The usual techniques for model evaluation revolve around **The usual techniques for model evaluation revolve around** resampling. Reliable Estimates<br>
Ing Reliable Estimates<br>
Usual techniques for model evaluation revolve around<br>
mpling.<br>
Simulating the reality.<br>
Dobtain an evaluation estimate for unseen data.<br>
mples of Resampling-based Methods<br>
Holdou Reliable Estimates<br>
Reliable Estimates<br>
Reliable Estimates<br>
Il techniques for model evaluation revolve around<br>
Il techniques<br>
Dotain an evaluation estimate for unseen data.<br>
Sof Resampling-based Methods<br>
S-validation. Reliable Estimates<br>
The usual techniques for model evaluation revolve around<br>
resampling.<br> **Examples of Researce Methods**<br> **Examples of Resampling-based Methods**<br>
Holdout.<br>
Resampling-based Methods<br>
Resampling-based Method <table>\n<tbody>\n<tr>\n<th>Obtaining Reliable Estimates</th>\n</tr>\n<tr>\n<td>■ The usual techniques for model evaluation revolve around<br/>resampling.</td>\n</tr>\n<tr>\n<td>\_\_ Simulating the reality.</td>\n</tr>\n<tr>\n<td>\_\_ Obtain an evaluation estimate for unseen data.</td>\n</tr>\n<tr>\n<td>\_\_ Examples of Resampling-based Methods</td>\n</tr>\n<tr>\n<td>\_\_ Holdout.</td>\n</tr>\n<tr>\n<td>\_\_ Cross-validation.</td>\n</tr>\n<tr>\n<td>\_\_ Bootstra.</td>\n</tr>\n</tbody>\n</table>\n<p>Time Series Data Are Special!</p>\n<p>Any form of resampling changes the natural order of the data!</p>\n<p>\_\_3</p> The usual techniques for model evaluation revolve around<br>
resampling.<br>
Simulating the reality.<br> **Any form of resampling-based Methods**<br>
Methods<br> **Any form of resampling changes the natural order of the data!**<br>
Any form of
	- -
	- - **Holdout.**
		- Cross-validation.
		- Bootstrap.

stimate for unseen data.<br>
ased Methods<br>
I<br>
the natural order of the data!<br>
Time Series<br>
32/45<br>
32/45



# Evaluation Methodology Correct Evalution of Time Series Models Evaluation Methodology<br>
Figure 1997 - The Series Models<br>
General Guidelines<br>
Do not "forget" the time tags of the observations.<br>
Do not evaluate a model on past data. Evaluation Methodology<br>Discrimed Services Models<br>Do not "forget" the time tags of the observations.<br>Do not evaluate a model on past data. Evaluation Methodology<br>
Do not "forget" the time tags of the observations.<br>
Do not "forget" the time tags of the observations.<br>
Do not evaluate a model on past data.<br>
Sisible method Evaluation Methodology<br>
For Evaluation of Time Series Models<br>
General Guidelines<br>
Do not "forget" the time tags of the observations.<br>
Do not evaluate a model on past data.<br>
A possible method<br>
Divide the existing data in tw Evalution Mehodology<br>
Evalution of Time Series Models<br>
Po not "forget" the time tags of the observations.<br>
Do not evaluate a model on past data.<br>
Sible method<br>
Divide the existing data in two time windows<br> **Container a** Pa Evalution of Time Series Models<br>
aral Guidelines<br>
Do not "forget" the time tags of the observations.<br>
Do not evaluate a model on past data.<br>
Sible method<br>
Divide the existing data in two time windows<br> **Exacted alternatives** Suidelines<br>
Suidelines<br>
Suidelines<br>
ot "forget" the time tags of the observations.<br>
ot evaluate a model on past data.<br>
e method<br>
e the existing data in two time windows<br>
Past data (observations till a time t).<br>
"Future" da

- 
- Guidelines<br>ot "forget" the time tags of the observations.<br>ot evaluate a model on past data.<br>e method<br>past data (observations till a time *t*).<br>Past data (observations after *t*).<br>Then of these three learn-test alternatives Suidelines<br>
ot "forget" the time tags of the observations.<br>
ot evaluate a model on past data.<br>
e method<br>
e the existing data in two time windows<br>
Past data (observations till a time *t*).<br>
"Future" data (observations after
	-

- - **Past data (observations till a time t).**
	- **F** "Future" data (observations after  $t$ ).
- on past data.<br>
where windows<br>
still a time t).<br>
still a time t).<br>
arn-test alternatives<br>  $\begin{bmatrix}\n\cdot & \cdot & \cdot & \cdot \\
\cdot & \$ 
	-
	-
	-







A single model is obtained with the available "training" data, and Evaluation Methodology<br>
Learn-Test Strategies<br>
A single model is obtained with the available "training" data, and<br>
applied to all test period.<br>
Growing Window<br>
Even: w test cases a new model is obtained wing all data avail

Every  $w_v$  test cases a new model is obtained using all data available Fixed Window<br>
A single model is obtained with the available "training" data, and<br>
applied to all test period.<br>
Growing Window<br>
Every  $w_v$  test cases a new model is obtained using all data available<br>
iil then.<br>
Sliding Win Eixed Window<br>
A single model is obtained with the available "training" data, and<br>
applied to all test period.<br>
Growing Window<br>
Every w<sub>y</sub> test cases a new model is obtained using all data available<br>
fill then.<br>
Sliding Win Time Series 34 / 45<br>Time Series 34 / 45<br>Time Series 34 / 45<br>Time Series 34 / 45

# Evaluation Measures<br>Evaluating Predictive Performance Evaluation Measures<br>Some Metrics for Evaluating Predictive Performance<br>Absolute Measures<br>Relative Measures

$$
MSE = \frac{1}{n}\sum_{i=1}^n (\hat{x}_i - x_i)^2
$$

Mean Absolute Deviation (MAD)

$$
MAD = \frac{1}{n}\sum_{i=1}^{n} |\hat{x}_i - x_i|
$$

Theil Coefficient  
\n
$$
U = \frac{\sqrt{\sum_{i=1}^{n} (\hat{x}_i - x_i)^2}}{\sqrt{\sum_{i=1}^{n} (x_i - x_{i-1})^2}}
$$
\nMean Absolute Percentage  
\nError (MAPE)  
\n
$$
MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{(\hat{x}_i - x_i)}{x_i} \right|
$$
\nTime Series

$$
MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{(\hat{x}_i - x_i)}{x_i} \right|
$$



## Assumptions

## Assumptions of "Classical" Linear Approaches

**Linearity** 

**Stationarity** 

Assumptions<br>
Umptions of "Classical" Linear Approaches<br>
Unearity<br>
The model of the time series behaviour is linear on its inputs.<br>
Stationarity<br>
The underlying equations governing the behaviour of the system<br>
do not change The underlying equations governing the behaviour of the system Assumptions<br>
Assumptions<br>
Linear Approaches<br>
Linearity<br>
The model of the time series behaviour is linear on its inputs.<br>
Stationarity<br>
The underlying equations governing the behaviour of the system<br>
do not change with time

Most "classical" approaches assume stationary time series, thus one usually needs to transform non-stationary time series into stationary equations governing the behaviour of the system<br>with time.<br>or descriptions are trained with time.<br>or description and the series, thus one<br>ansform non-stationary time series into stationary<br>ones before using these tools. Substitution is interaction is in puts.<br>
Soverning the behaviour of the system<br>
sume stationary time series, thus one<br>
-stationary time series into stationary<br>
using these tools.



Moving Averages Moving Averages<br>Moving Average Models<br>Definition<br>A moving average of order  $\alpha$ , MA( $\alpha$ ) is a time series given by

## **Definition**

Moving Average Models<br>
Definition<br>
A moving average of order q, MA(q),is a time series given by<br>  $Y_t = \sum_{j=1}^{q} \beta_j X_{t-j}$ 

$$
Y_t = \sum_{i=0}^q \beta_i X_{t-i}
$$



Exponential MAs<br>| Average Models<br>| Exponential Moving Average Models<br>
Definition<br>
An exponential moving average is a series given by

## **Definition**

Exponential Mas

\nExponential May

\nDefinition

\nAn exponential moving average is a series given by

\n
$$
Y_t = a \cdot X_t + (1 - a) \cdot \text{EMA}_\alpha(X_{t-1})
$$
\n
$$
Y_1 = X_1
$$
\nwhere  $a \in [0..1]$  is a smoothing parameter.

\n
$$
X_t = \frac{a}{1 - \frac{1}{2} \cdot 1} \cdot \frac{1}{2} \cdot \frac{1}{
$$



Autoregressive AR

Autoregressive AR<br>Autoregressive (AR) Models<br>Definition<br>An autoregressive model of order n is a series given by

## **Definition**

Autoregressive (AR) Models<br>
Definition<br>
An autoregressive model of order  $p$  is a series given by<br>  $Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i}$ 

$$
Y_t = \sum_{i=0}^p \alpha_i Y_{t-i}
$$





**Autoregressive ARMA**<br>The and Moving Average Models Autoregressive ARMA<br>Mixed Autoregressive and Moving Average Models<br>Definition<br>A mixed ARMA model of order a later series given by Andregressive ARMA<br>
Andregressive and Moving Average Models<br>
Definition<br>
A mixed ARMA model of order p, q is a series given by<br>  $Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{i=1}^{q} \beta_i X_{t-i}$ 

## **Definition**

$$
Y_t = \sum_{i=0}^p \alpha_i Y_{t-i} + \sum_{i=0}^q \beta_i X_{t-i}
$$







# Autoregressive ARIMA<br>- ARIMA) Models<br>-Autoregressive ARIMA<br>Integrated ARMA (or ARIMA) Models<br>Definition<br>An integrated ARMA (or ARIMA) model of order p. d. q is a series

## **Definition**

Antorgressive ARMA<br>
An integrated ARMA (or ARIMA) Models<br>
An integrated ARMA (or ARIMA) model of order p, d, q is a series<br>
given by<br>  $y_t' = c + \phi_1 y_{t-1}' + \cdots + \phi_p y_{t-p}' + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t$ Autoregressive ARIMA<br>
Integrated ARMA (or ARIMA) Models<br>
Definition<br>
An integrated ARMA (or ARIMA) model of order p, d, if<br>
given by<br>  $y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t$ 

 $\begin{align*} \mathcal{P}_{t-1} + \cdots + \theta_q e_{t-q} + e_t \end{align*}$ <br>ay have been differenced more than and side include both lagged values of  $y_t$ <br> $\begin{align*} \mathbf{A}(p,d,q) \text{ model, where} \end{align*}$ <br>Fine Series  $\begin{align*} \mathbf{I}(\mathbf{f}(t)) = \mathbf{I}(\mathbf{f}(t)) \mathbf{I}(\mathbf{f}(t)) \end{align*}$ 



# Case Dependecies

# **Clustering**



[1]



## **Clustering**

Case Dependecies Whole time-series clustering is considered as clustering of a set of Case Dependecies<br>
Clustering<br>
Whole time-series clustering is considered as clustering of a set of<br>
individual time-series with respect to their similarity. Here, clustering<br>
means applying conventional (usually) clusterin Case Dependecies<br>
Clustering<br> **Whole time-series clustering** is considered as clustering of a set of<br>
individual time-series with respect to their similarity. Here, clustering<br>
means applying conventional (usually) cluster Case Dependecies<br>
Clustering<br>
Whole time-series clustering is considered as clustering of<br>
individual time-series with respect to their similarity. Here,<br>
means applying conventional (usually) clustering on discret<br>
object Curring<br>
Subsequence considered as clustering of a set of<br>
individual time-series with respect to their similarity. Here, clustering<br>
means applying conventional (usually) clustering on discrete objects, where<br>
objects are Case Dependedes<br>
Clustering<br>
Whole time-series clustering is considered as clustering of a set of<br>
individual time-series with respect to their similarity. Here, clustering<br>
means applying conventional (usually) clustering Case Dependecies<br>
Clustering<br>
Whole time-series clustering is considered as clustering of a set of<br>
individual time-series with respect to their similarity. Here, clusterin<br>
means applying conventional (usually) clustering

stering on a set of subsequences of a<br>a sliding window, that is, clustering of<br>series.<br>ategory of clustering which is seen in<br>e points based on a combination of their<br>and the similarity of the<br>ch is ismilar to time-series Caustering<br>
Caustering<br>
Whole time-series clustering is considered as clustering of a set of<br>
Individual time-series with respect to their similarity. Here, clustering<br>
means applying conventional (usually) clustering on d **Solution Exercise Solution**<br> **Solution Constrainers**<br> **Solution Constrainers**<br> **Solution Constrainers**<br> **Subsequence clustering** means clustering on a set of subsequences of a<br> **Subsequence clustering** means clustering on **VIOSTETTITY**<br> **Whole time-series clustering** is considered as clustering of a set of<br>
individual time-series with respect to their similarity. Here, clustering<br>
means applying conventional (usually) clustering on discrete Whole time-series clustering is considered as clustering of a set of<br>individual time-series with respect to their similarity. Here, clustering<br>means applying conventional (usually) clustering on discrete objects, where<br>obj Whole time-series customing is considered as clustering of a set of microscores with respect to their similarity. Here, clustering means applying conventional (usually) clustering on discrete objects, where objects are tim manyioual time-series with respect to their similarity. Here, custering<br>means applying conventional (usually) clustering on discrete objects, where<br>objects are time-series.<br>**Subsequence clustering** means clustering on a se

[5]

# Case Dependecies<br>Series subsequences

Clustering of time series subsequences<br>Clustering of time series subsequences<br>Subsequence Clustering: Given a single time series, sometimes in the<br>Sorm of streaming time series, individual time series (subsequences) are Subsequence Clustering: Given a single time series, sometimes in the form of streaming time series, individual time series (subsequences) are extracted with a sliding window. Clustering is then performed on the extracted time series subsequences.

**Abstract**<br> **Abstract**<br> **Abstract**<br>
In the agorithms, clustering of time series<br>
this are reduced much attention. In this work we make a<br>
puences is meaningless] More concretely, clusters extracted<br>
puences is meaningless



[6]

# Discretisation: SAX, PAA, TVA



# Case Dependecies

## **Resources**

Case Dependecies<br>
[1] Selina Chu, Eamonn Keogh, David Hart and Michael Pazzani. *Iterative*<br> *Deepening Dynamic Time Warping for Time Series*<br>
[2] Bilal Esmael, Arghad Arnaout, Rudolf K. Fruhwirth and Gerhard<br>
Thonhauser.

Case Dependecies<br>
21 Selina Chu, Eamonn Keogh, David Hart and Michael Pazzani. *Iterative<br>
Deepening Dynamic Time Warping for Time Series*<br>
21 Bilal Esmael, Arghad Arnaout, Rudolf K. Fruhwirth and Gerhard<br>
Thonhauser. Mult Case Dependecies<br>
[1] Selina Chu, Eamonn Keogh, David Hart and Michael Pazzani. *Iterative*<br>
Depending Dynamic Time Warping for Time Series<br>
[2] Bilal Esmael, Arghad Arnaout, Rudolf K. Fruhwirth and Gerhard<br>
Thonhauser. Mu Thonhauser. Multivariate Time Series Classification by Combining Trend-Based and Value-Based Approximations

Time Series

doolf K. Frunwirth and Gernard<br>s Classification by Combining Trend-<br>ions<br>Multiresolution Motif Discovery in<br>probay and Vincent Rialle. Mining<br>eries for Learning Meaningful Patterns:<br>e<br>irkhorshidi and Teh Ying Wah. Time-<br>Cl case Dependecies<br> **Resources**<br> **[1] Selina Chu, Eamonn Keogh, David Hart and Michael Pazzani.** *Iterative***<br>
Deepening Dynamic Time Warping for Time Series<br>
<b>[2] Bilal Esmael, Arghad Arnaout, Rudolf K. Fruhwirth and Gerhard** Case Dependecies<br>
[1] Selina Chu, Eamonn Keogh, David Hart and Michael Pazzani. *Iterative*<br> *Deepening Dynamic Time Warping for Time Series*<br>
[2] Bilal Esmael, Arghad Arnaout, Rudolf K. Fruhwirth and Gerhard<br>
Thonhauser. Heterogeneous Multivariate Time-Series for Learning Meaningful Patterns: Application to Home Health Telecare **[4] Selina Chu, Eamonn Keogh, David Hart and Michael Pazzani.** *Iterative***<br>Deepening Dynamic Time Warping for Time Series<br><b>[2] Bilal Esmael, Arghad Arnaout, Rudolf K. Fruhwirth and Gerhard**<br>**[2] Bilal Esmael, Arghad Arnaou** [1] Selina Chu, Eamonn Keogh, David Hart and Michael Pazzani. *Itera:*<br>Deepening Dynamic Time Warping for Time Series<br>[2] Bilal Esmael, Arghad Arnaout, Rudolf K. Fruhwirth and Gerhard<br>Thonhauser. Multivariate Time Series C [1] Selina Chu, Eamonn Keogh, David Hart and Michael Pazzani. *Iterative*<br>
Deepening Dynamic Time Warping for Time Series<br>
[2] Bilal Esmael, Arghad Arnaout, Rudolf K. Fruhwirth and Gerhard<br>
Thonhauser. Multivariate Time Se

is Meaningless: Implications for Previous and Future Researc

![](_page_41_Picture_10.jpeg)