NATIONAL RESEARCH
UNIVERSITY

School of Data Analysis and Artificial
Intelligence Department of Computer Science

DATA SCIENCE FOR BUSINESS

Lecture 7. Time series forecasting

Moscow, May 27th, 2022.

lzhukov@hse.ru

TIME SERIES FORECASTING ACROSS INDUSTRIES

Logistics & Transportation

- Forecasting of **shipped packages:** workforce planning

Retail grocery

- Forecasting of **sales during promotions:** optimizing warehouses

Insurance

- **Claims prediction:** determining insurance policies

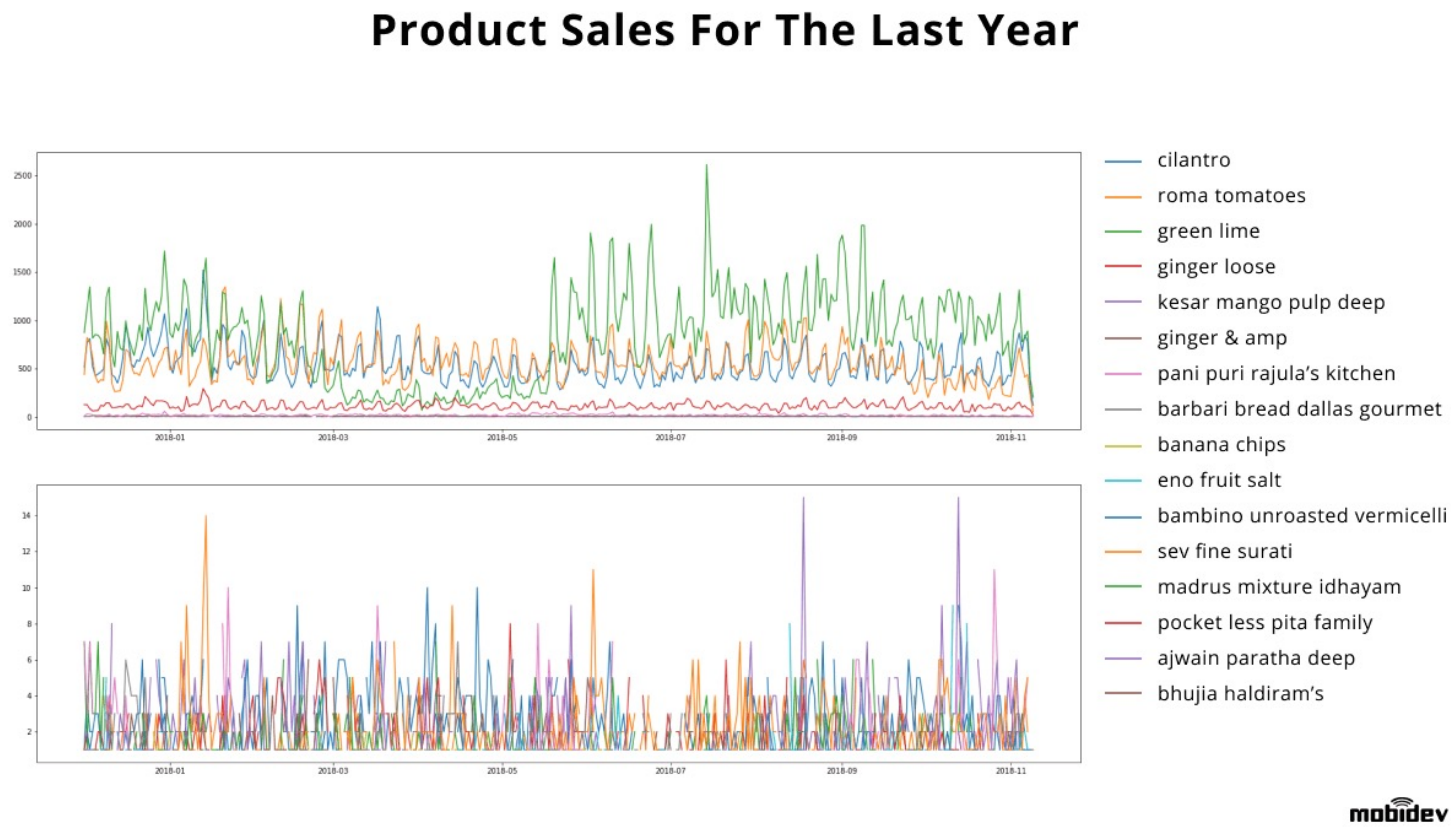
Manufacturing

- **Predictive Maintenance:** improving operational efficiency

Energy & Utilities

- **Energy load forecasting:** better planning and trading strategies

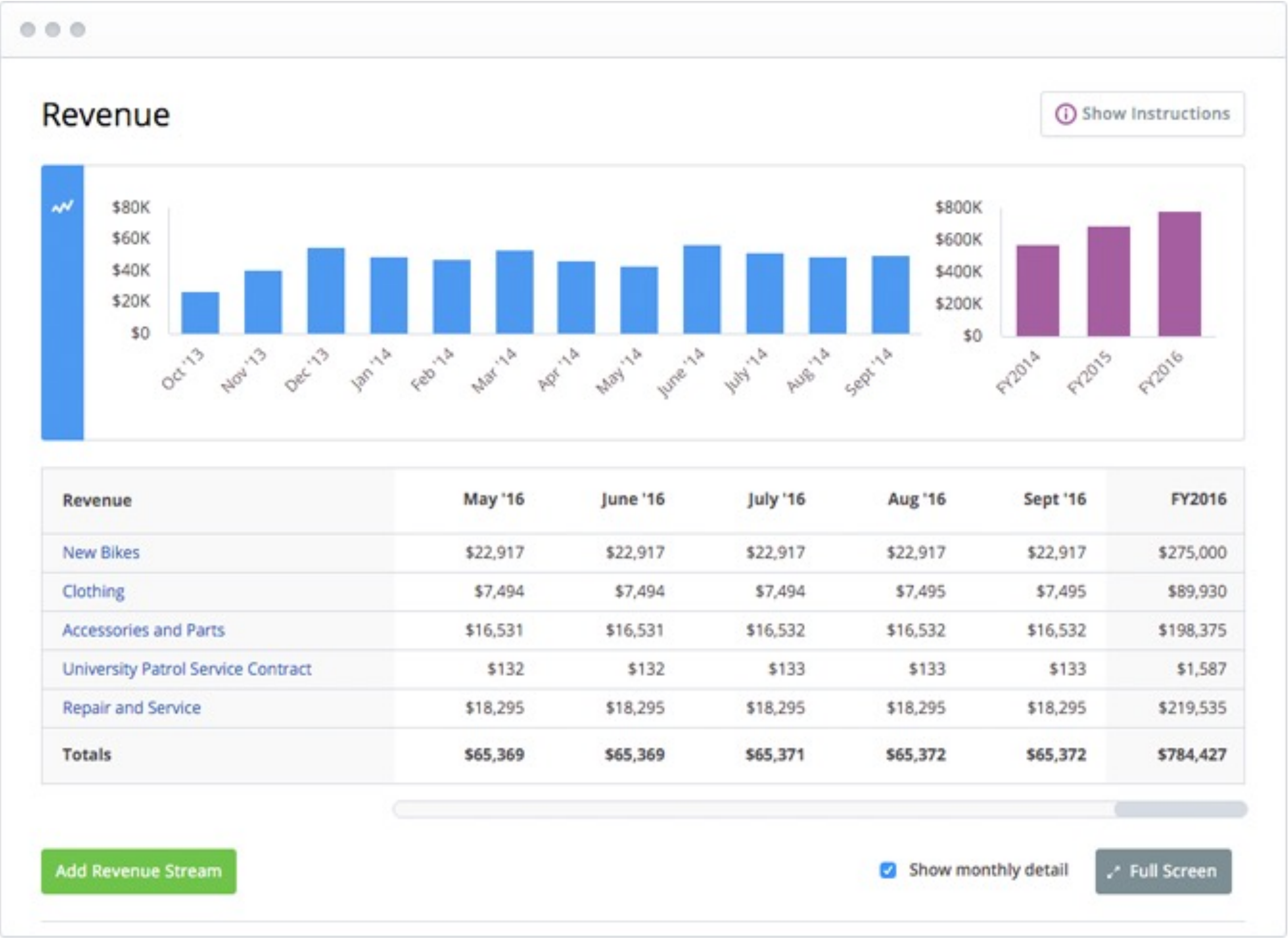
RETAIL SALES DATA



School of Data Analysis and Artificial Intelligence
Department of Computer Science

SALES FORECASTING

Estimating the future sales



Revenue	May '16	June '16	July '16	Aug '16	Sept '16	FY2016
New Bikes	$22,917	$22,917	$22,917	$22,917	$22,917	$275,000
Clothing	$7,494	$7,494	$7,494	$7,495	$7,495	$89,930
Accessories and Parts	$16,531	$16,531	$16,532	$16,532	$16,532	$198,375
University Patrol Service Contract	$132	$132	$133	$133	$133	$1,587
Repair and Service	$18,295	$18,295	$18,295	$18,295	$18,295	$219,535
Totals	$65,369	$65,369	$65,371	$65,372	$65,372	$784,427

From…

What we have

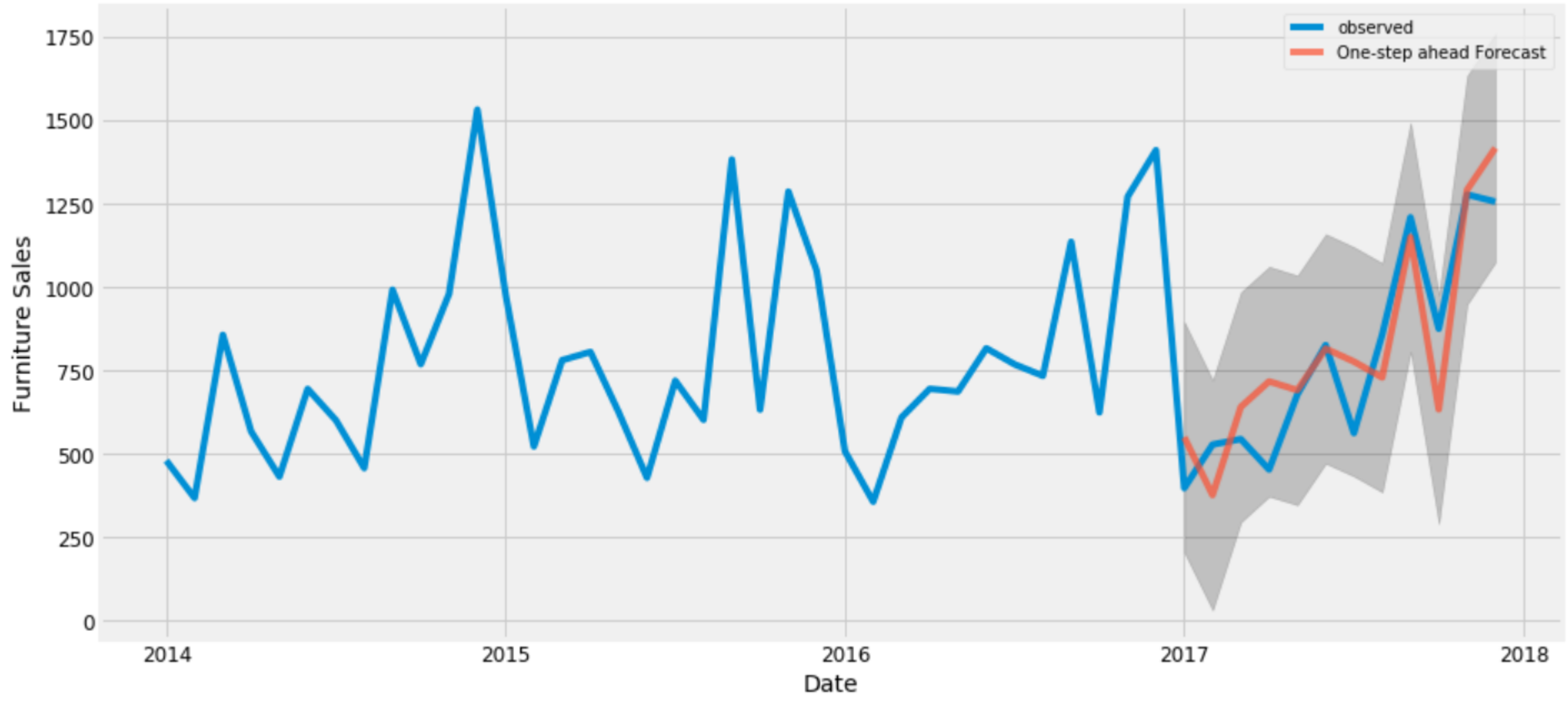
Past sales (transaction)

→

…You can predict

What we have

Future sales



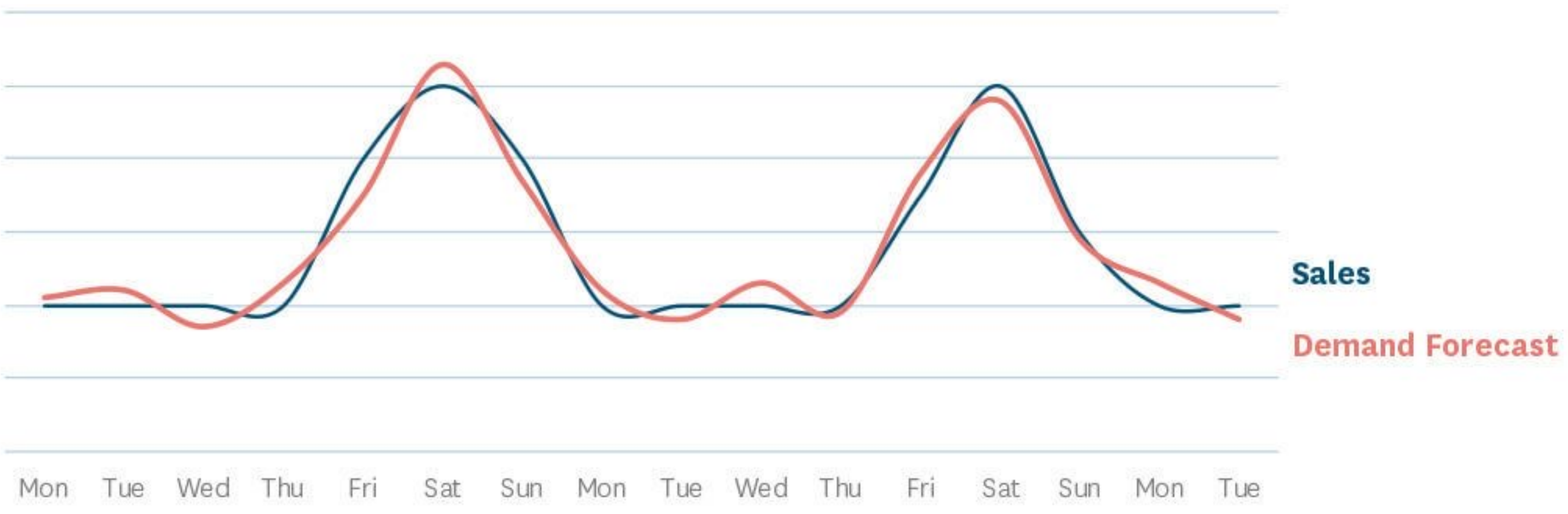
DEMAND FORECASTING

Demand forecasting is the process of predicting what the demand for certain products will be in the future.

- **Supplier relationship management** - calculate how many products to order
- **Customer relationship management -** predict which categories of products should be available the next period from a specific store location. This improves customer satisfaction and commitment to your brand.
- **Order fulfillment and logistics -** optimizing supply chains, the product will be more likely to be in stock for ordering, and unsold goods won’t occupy prime retail space.
- **Marketing campaigns -** adjust ads and marketing campaigns and influence the number of sales.
- **Manufacturing flow management.** Being part of the ERP, the time series-based demand forecasting predicts production needs

DEMAND FORECASTING

Demand forecasting is the process of predicting what the demand for certain products will be in the future.



Retailers need accurate day-level forecasts for effective replenishment of fresh products as well as for managing capacity in all parts of their supply chains.

DEMAND FORECASTING

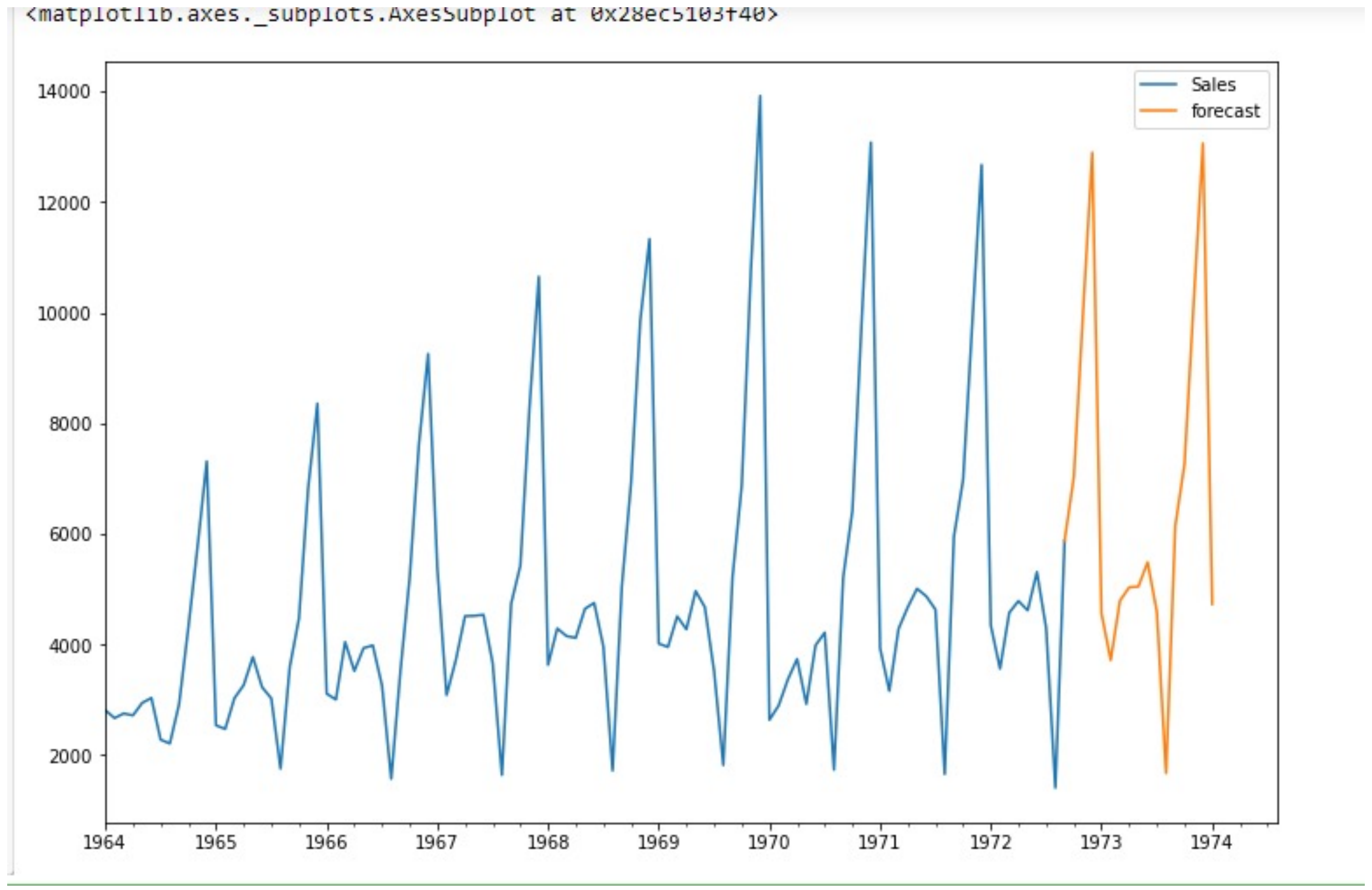
Price changes, promotions, and other business decisions impacting demand

WHAT IS TIME SERIES FORECASTING

Everything that is observed sequentially over time is a time series

Forecasting time series – how the sequence of observations will continue into the future

1. Numerical information about the past is available
2. Reasonable to assume that past patterns will continue into the future



TIME SERIES FORECASTING

Classical time series analysis

- Predicting future values by past observations of the same variable through autocorrelation
- Most often univariate
- Can use few external factors

y(t+1) = f (y(t), y(t-1), y(t-2), …)

- Decomposition models (trend, seasonality)
- Moving average
- Exponential smoothing
- ARIMA

ML methods, forecasting as supervised learning

- History of comparable signals
- Many explanatory factors
- Large datasets

y(t) = f (x1(t), x2(t), …))
y(t+h) = f(x1(t), x2(t),…)

ML algorithms (regression)

- GLM
- Random forest
- Gradient boosting

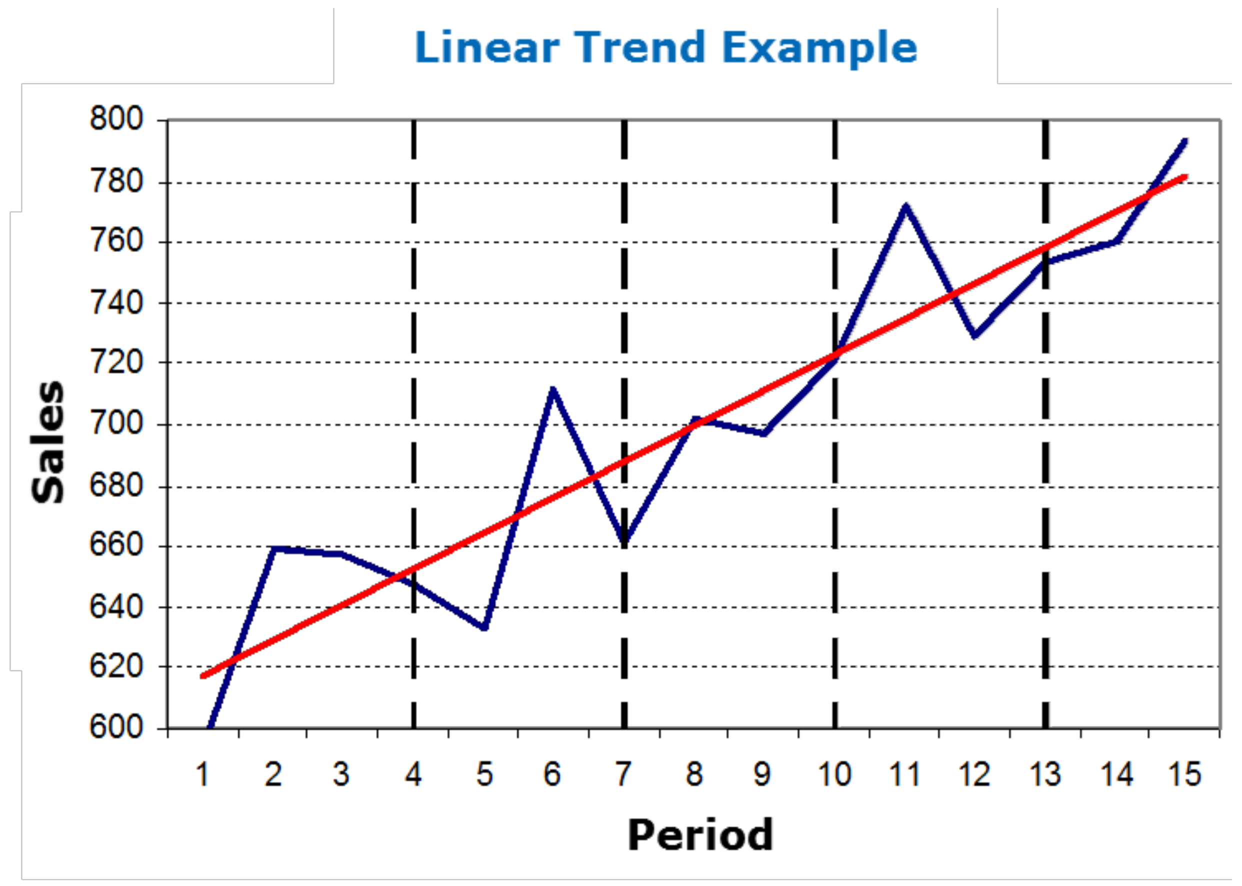
Time order dependence between observations!

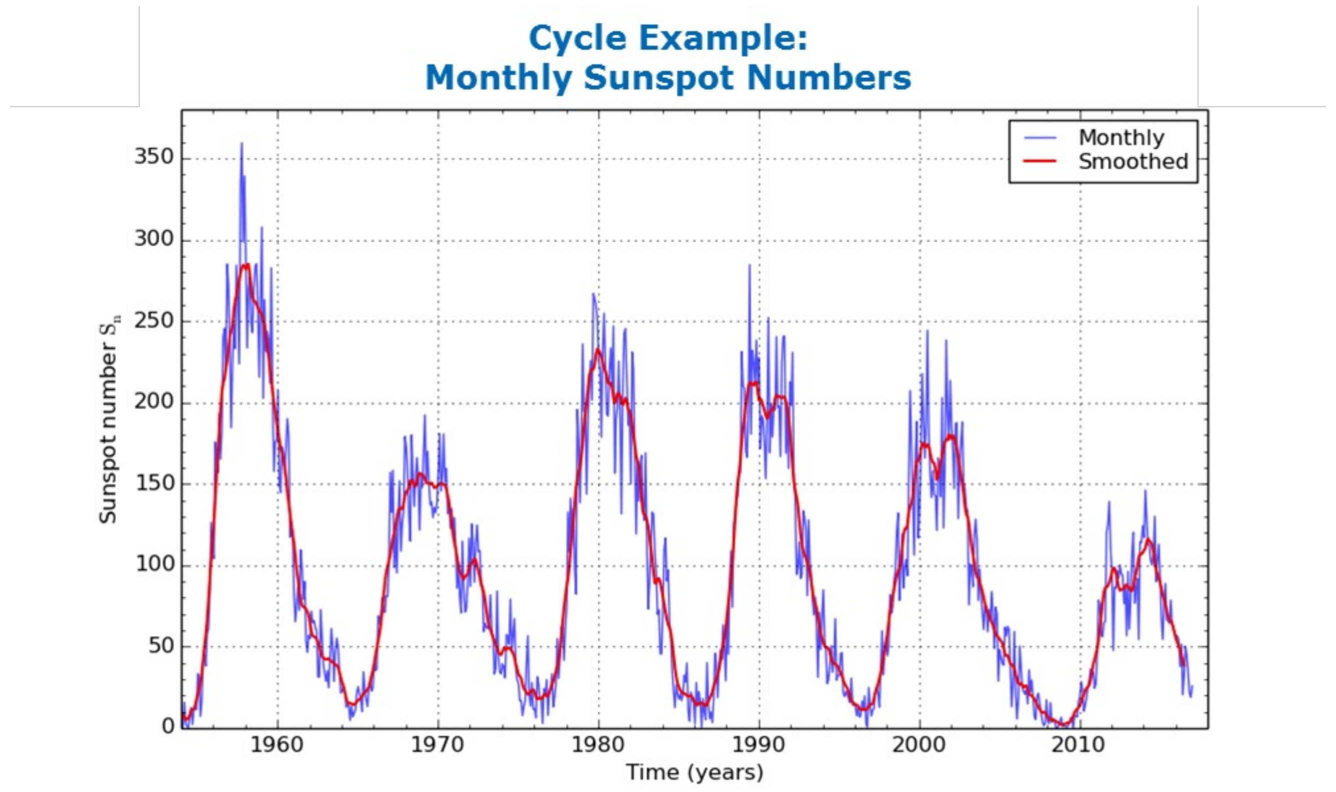
TIME SERIES DECOMPOSITION

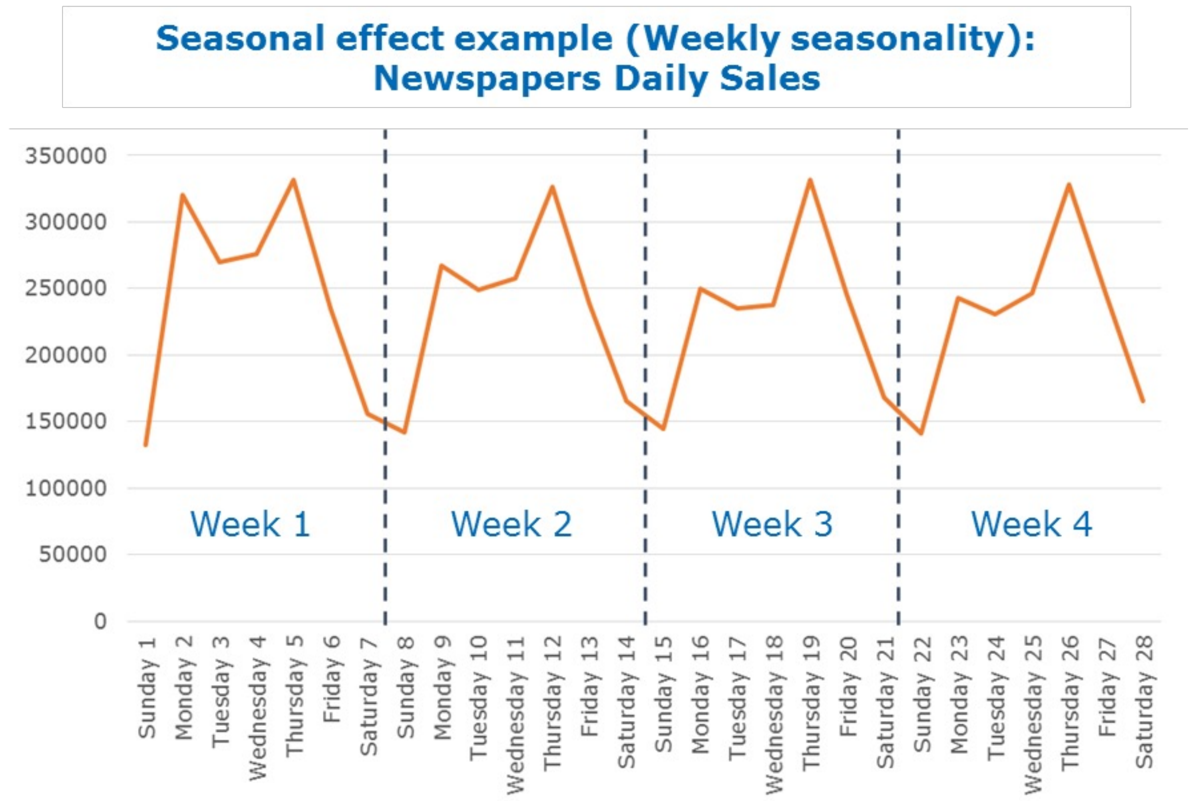
Trend, Seasonal, Cyclic, what seasonality is

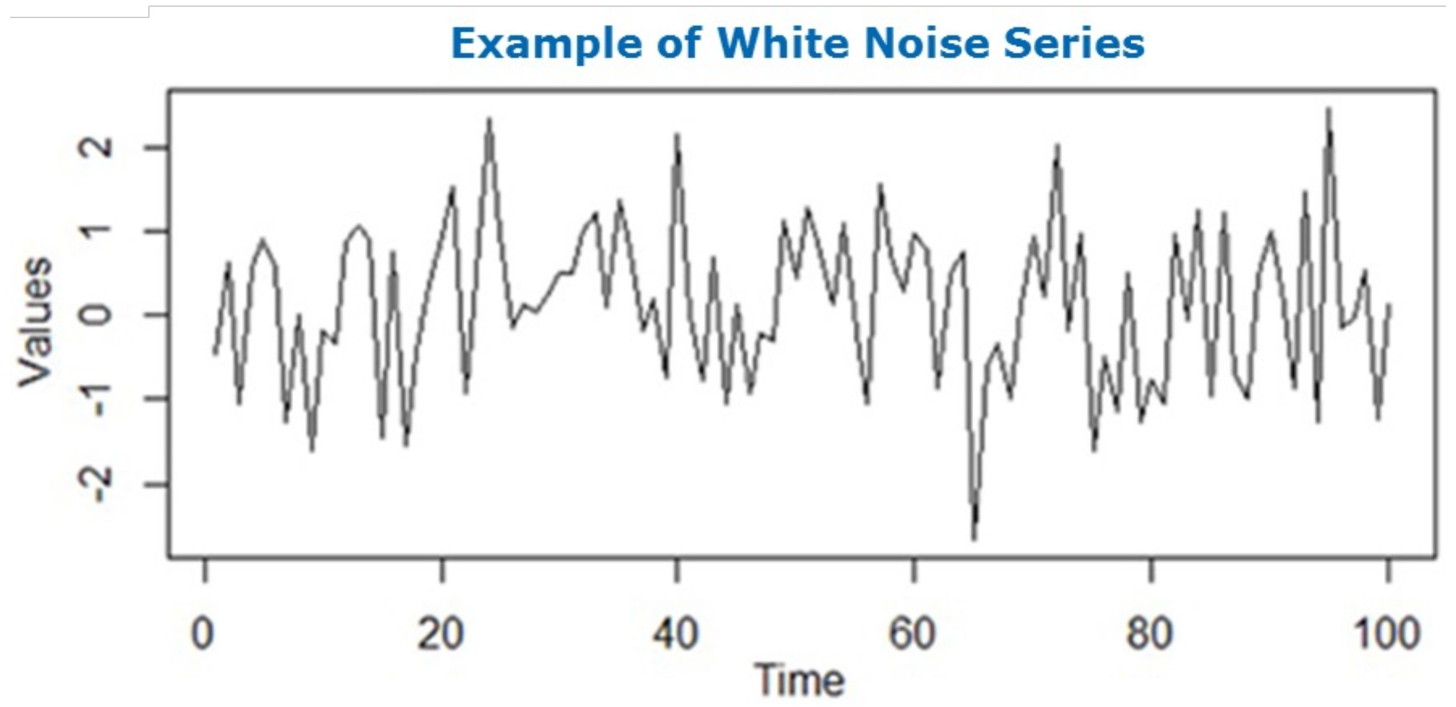
- **Trend:** Steady, long-term, moving gradually in one direction, long term increase or decrease
- **Seasons:** Regular short term variations often associated with months or quarters (fixed and known period)
- **Cycle:** A variation that occurs regularly, but may vary in length, oscillatory component not caused by seasonal factors that
- **Random Components/Residual:** Removal of trends and cyclical variations from time-series data uncovers fluctuations that is irregular, random factors often not relevant to prediction

TIME SERIES DECOMPOSITION

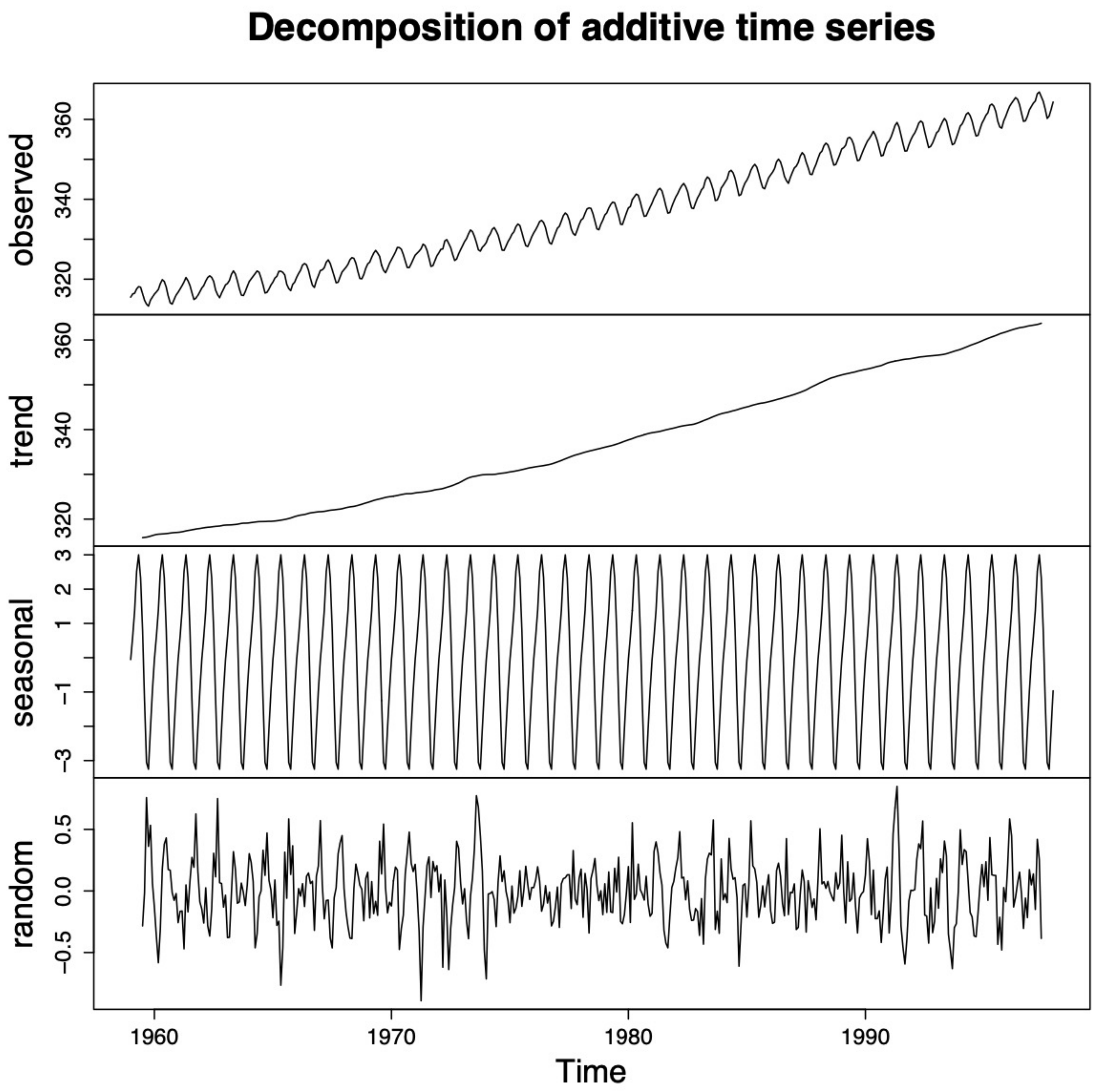








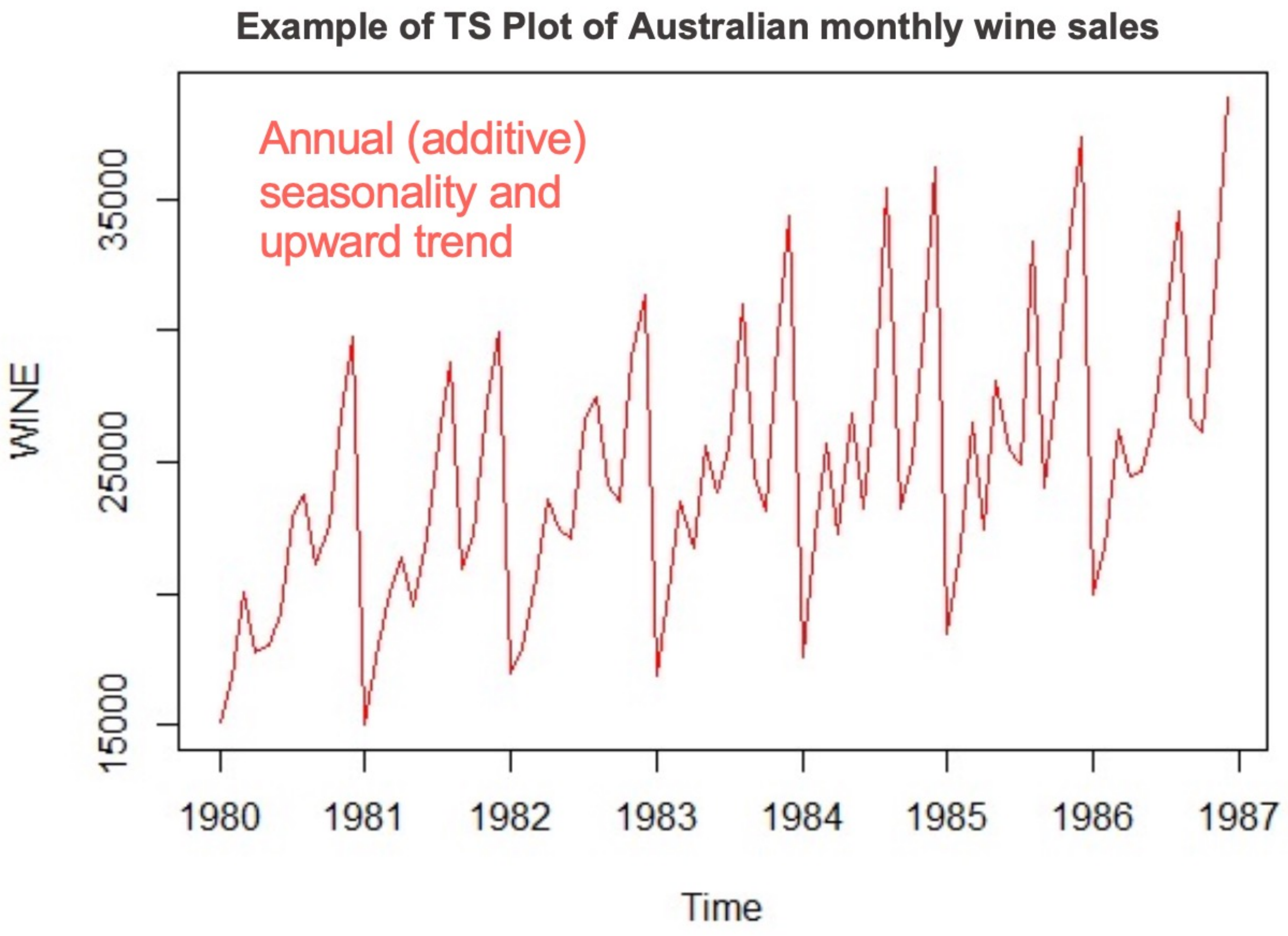
TIME SERIES DECOMPOSITION

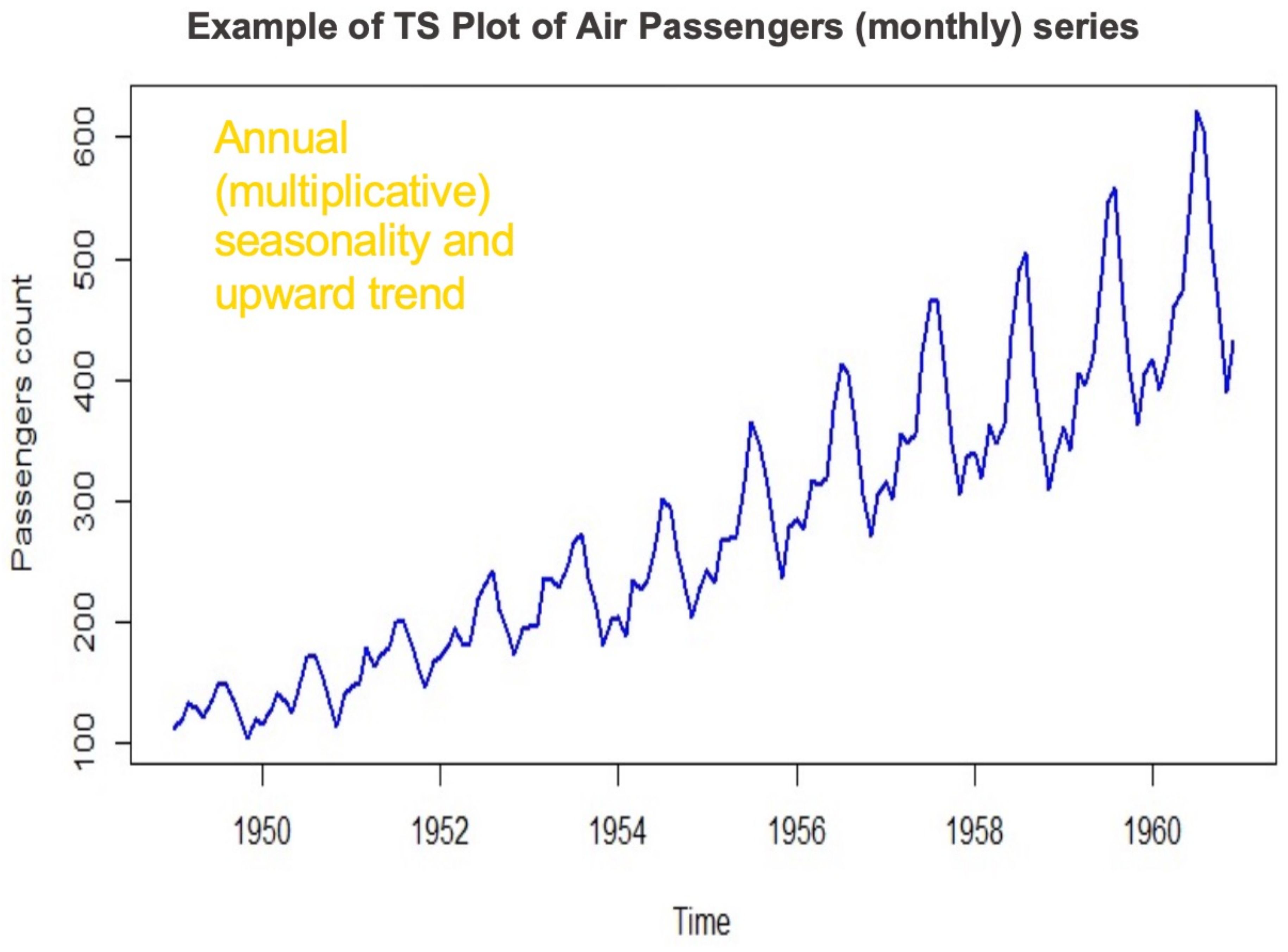


Classical decomposition:

- Estimating trend T(t) through a rolling mean
- Computing S(t) as the average detrended series Y(t)-T(t) for each season (e.g. for each month)
- Computing the remainder series as R(t)=Y(t)-T(t)-S(t)

TIME SERIES DECOMPOSITION



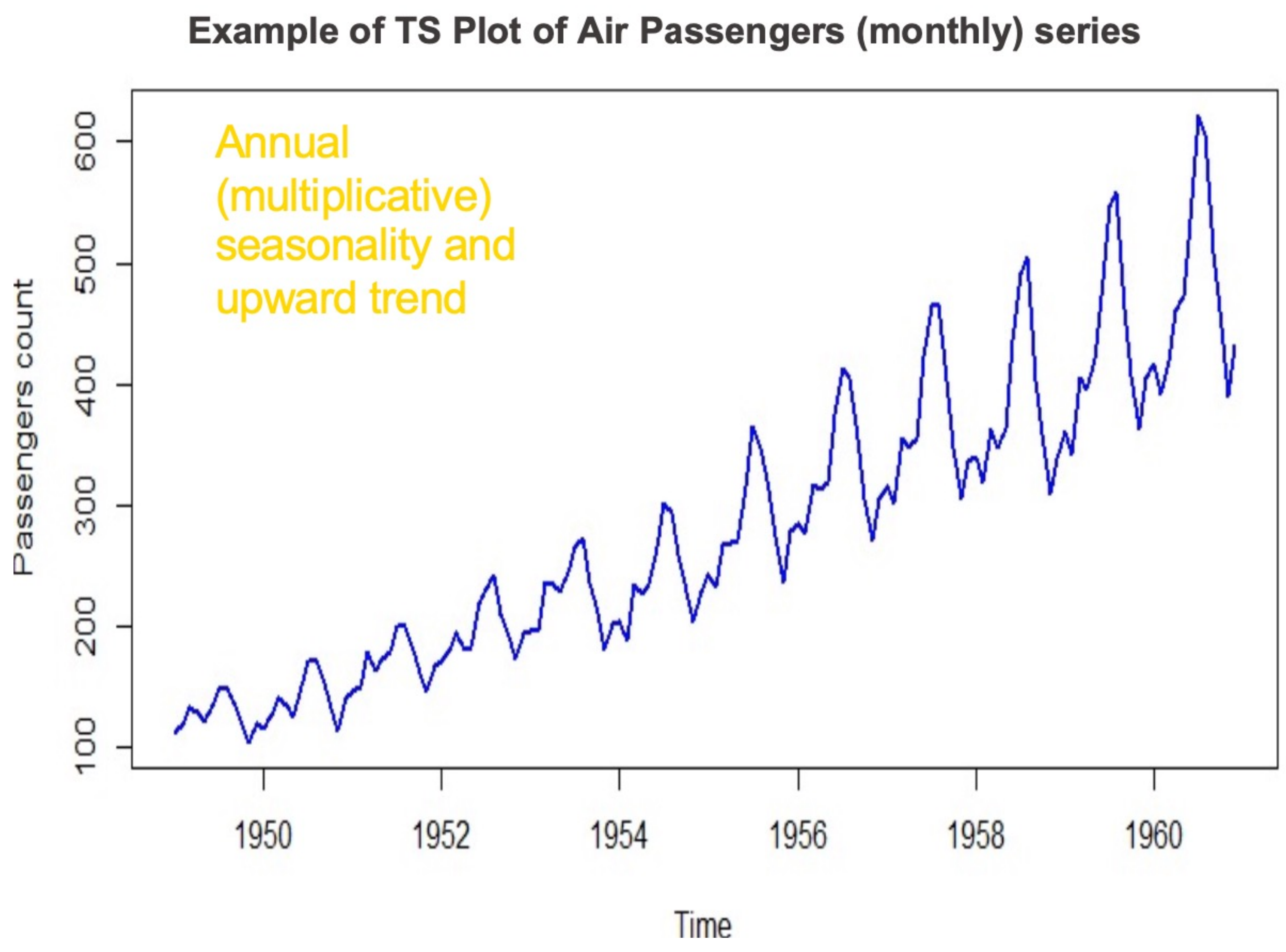


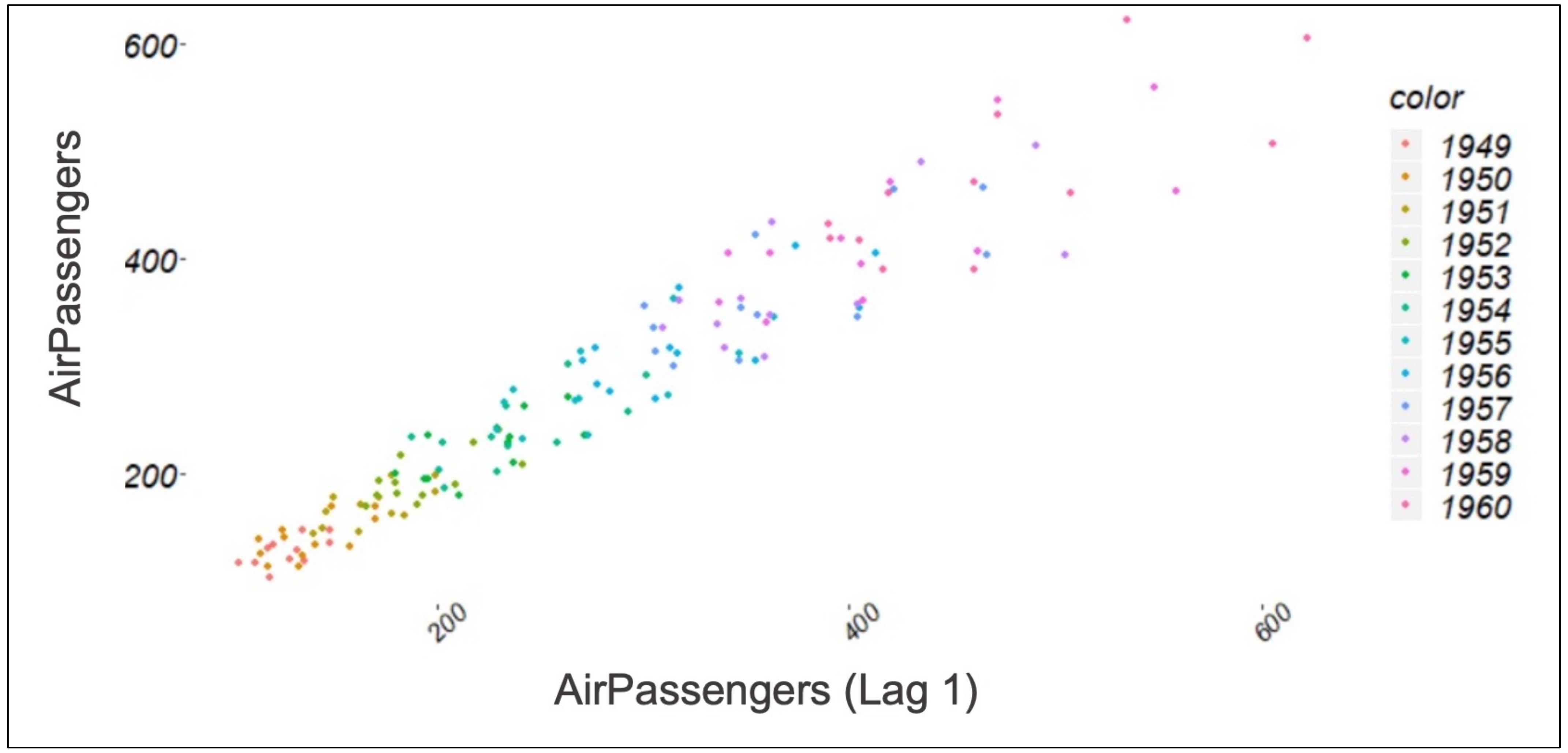
$$y_t = T_t + C_t + S_t + I_t,$$

$$y_t = T_t \times C_t \times S_t \times I_t.$$

CLASSICAL TIME SERIES ANALYSIS

Autocorrelation – correlation (linear relationship) between the lagged values of time series y(t) and y(t-k)





CLASSICAL TIME SERIES ANALYSIS

- **Classical Decomposition:** considers the time series as the overlap of several elementary components (i.e. trend, cycle, seasonality, error)

$$\hat{y}_{T+h|T} = y_{T+h-m(k+1)},$$

- **Exponential Smoothing:** method based on the weighting of past observations, taking into account the overlap of some key time series components (trend and seasonality)

$$\hat{y}_{t+1|t} = \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \ \dots$$

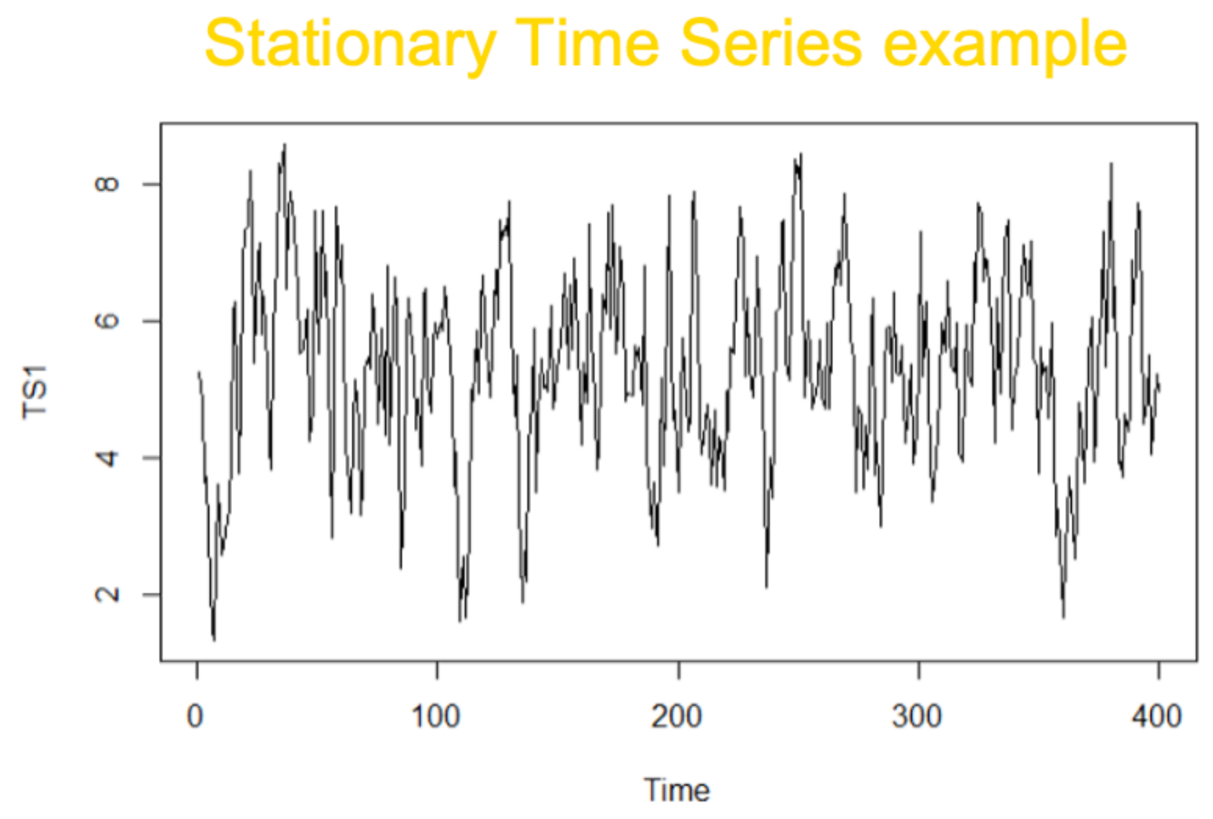
- **ARIMA** (*AutoRegressive Integrated Moving Average*): class of statistical models that aim to treat the correlation between values of the series at different points in time using a regression-like approach and controlling for seasonality

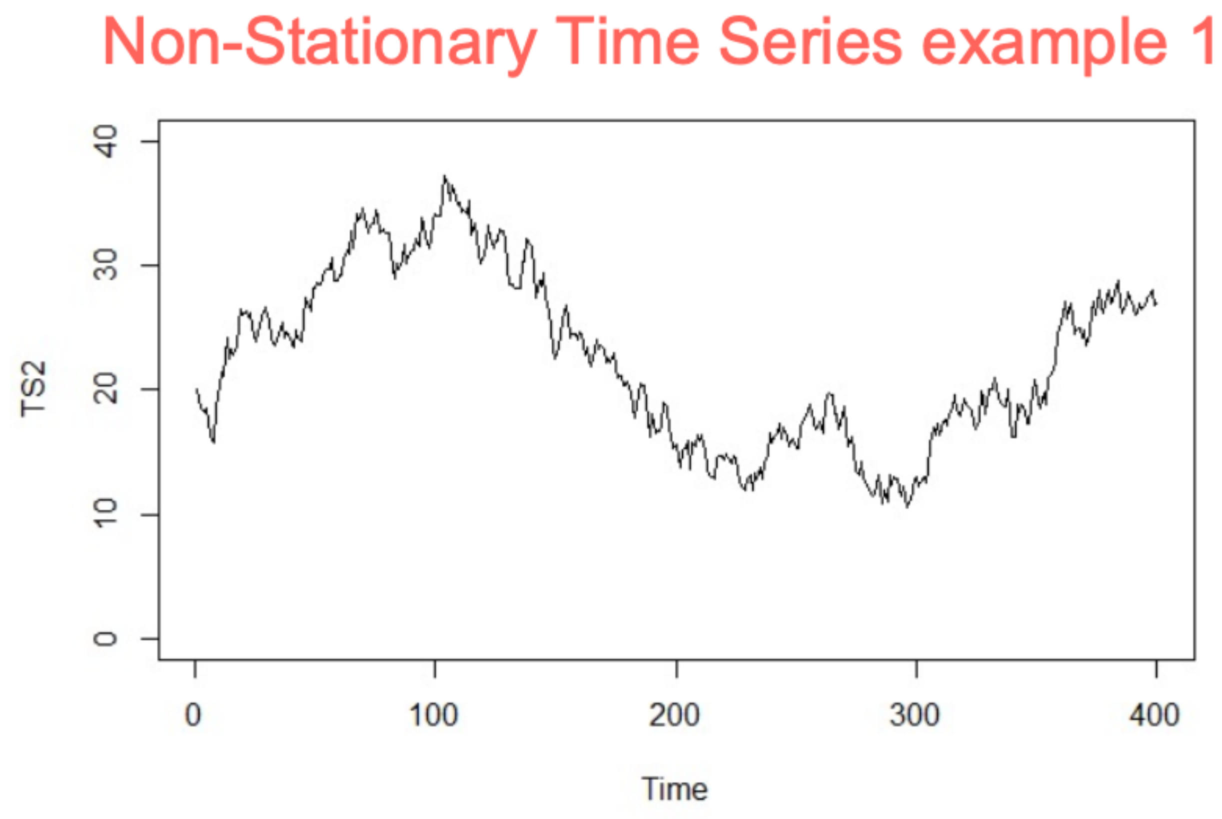
$$y_t = c + \underbrace{\phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p}}_{\text{Autoregressive component of order p}} + \underbrace{\theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}}_{\text{Moving Average component of order q}} + \varepsilon_t$$

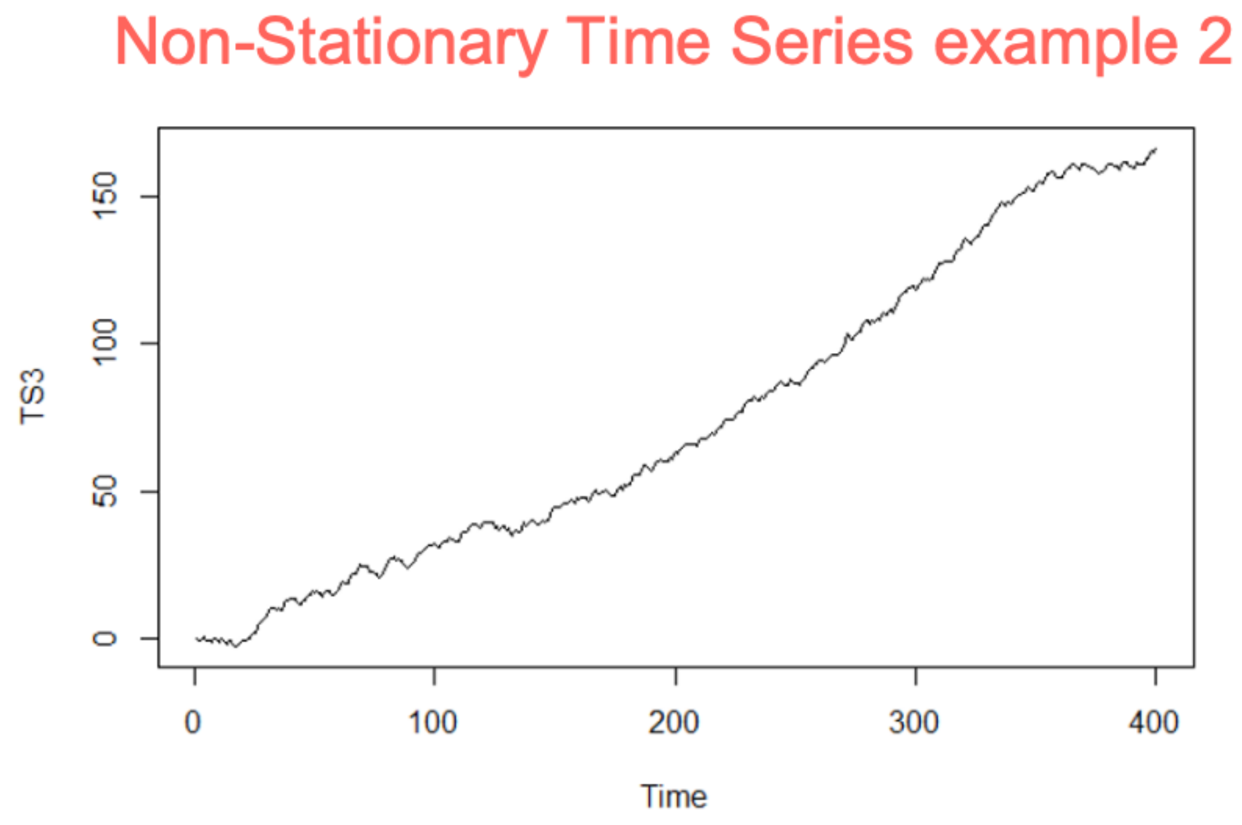
CLASSICAL TIME SERIES ANALYSIS

A time series can be defined as “stationary” when *its properties does not depend on the time at which the series is observed*, so that:

- the values oscillate frequently around the mean, independently from time
- the variance of the fluctuations remains constant across time
- the autocorrelation structure is constant over time and no periodic fluctuations exist

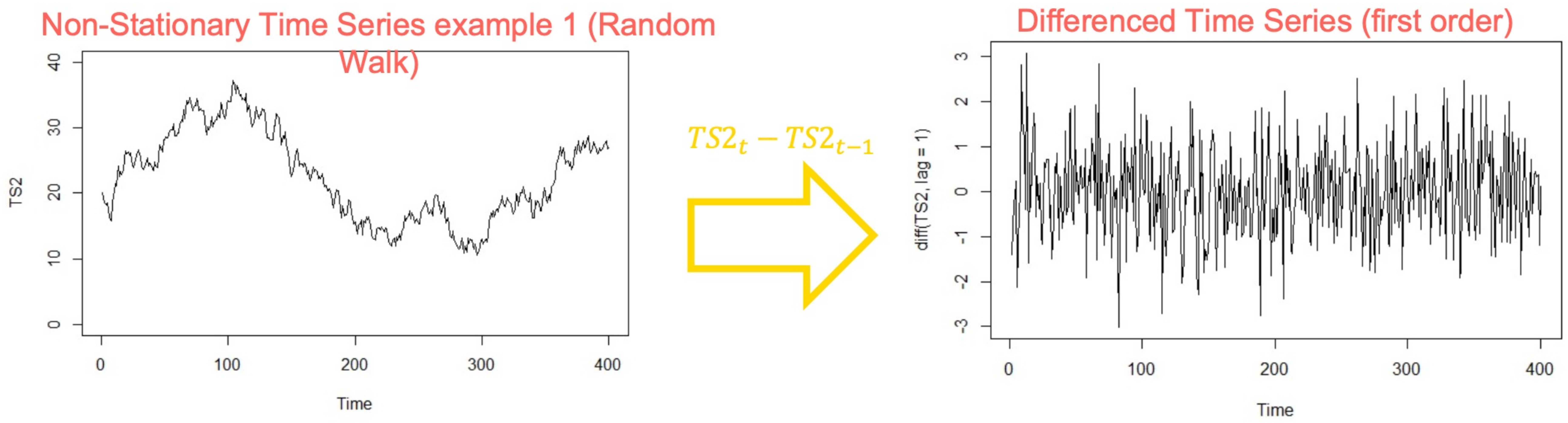






CLASSICAL TIME SERIES ANALYSIS

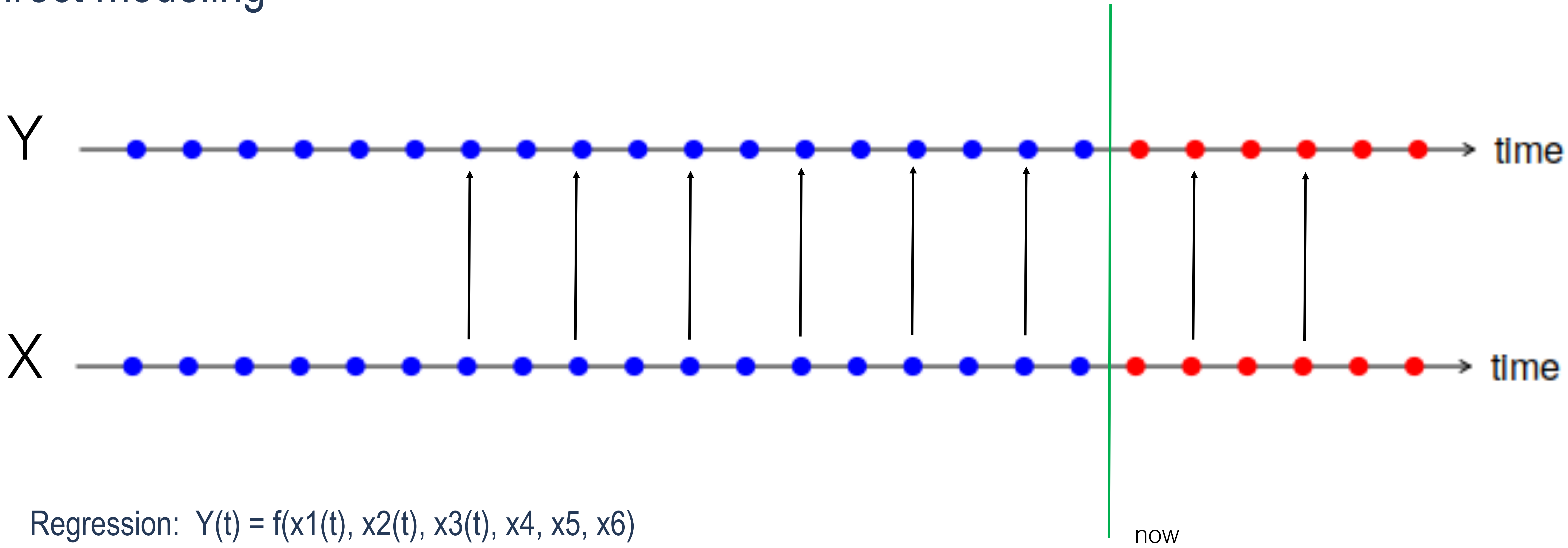
Differencing – computing the difference between consecutive observations. $y_t' = y_t - y_{t-1}$



Can make non-stationary series stationary. Widely used as a preprocessing for ML methods

FORECASTING AS SUPERVISED LEARNING PROBLEM - REGRESSION

Direct modeling

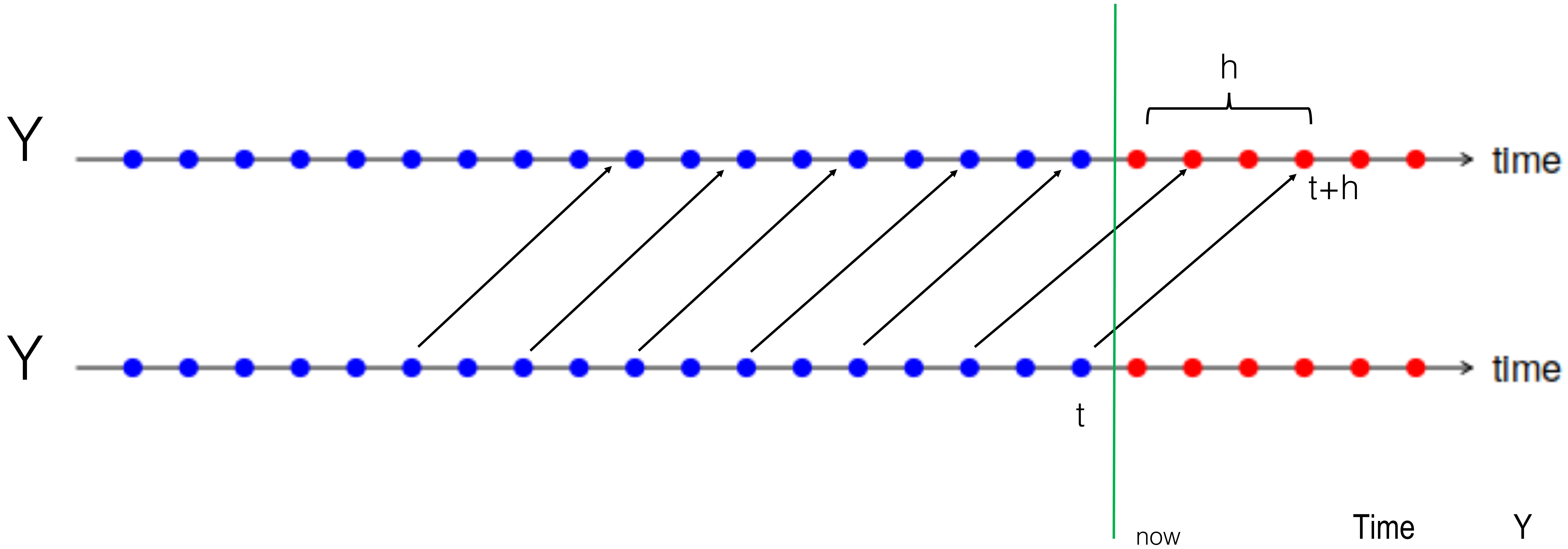


Regression: Y(t) = f(x1(t), x2(t), x3(t), x4, x5, x6)

100 110 108 115 120

FORECASTING WITH REGRESSION

Modeling with time lag



Time	Y
1,	100
2,	110
3,	108
4,	115
5,	120

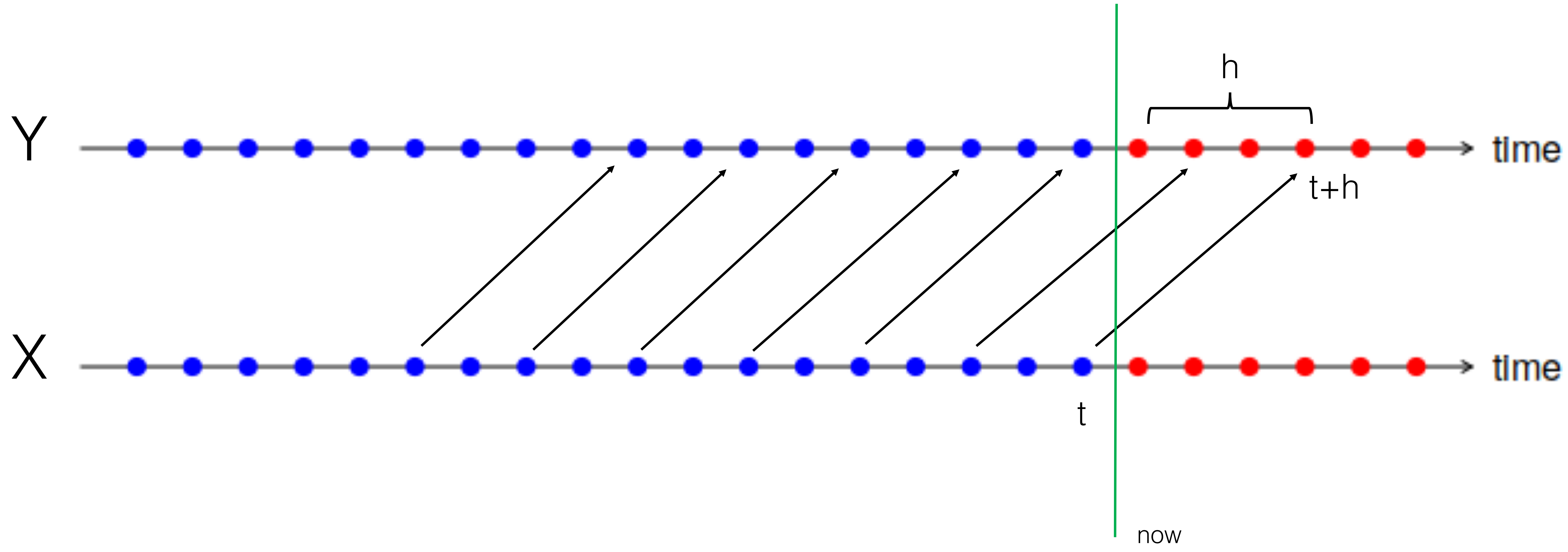
X	Y
?,	100
100,	110
110,	108
108,	115
115,	120
120,	?

Regression: Y(t+h) = f(Y(t))

using previous time steps as input variables and use the next time step as the output variable.

FORECASTING WITH REGRESSION

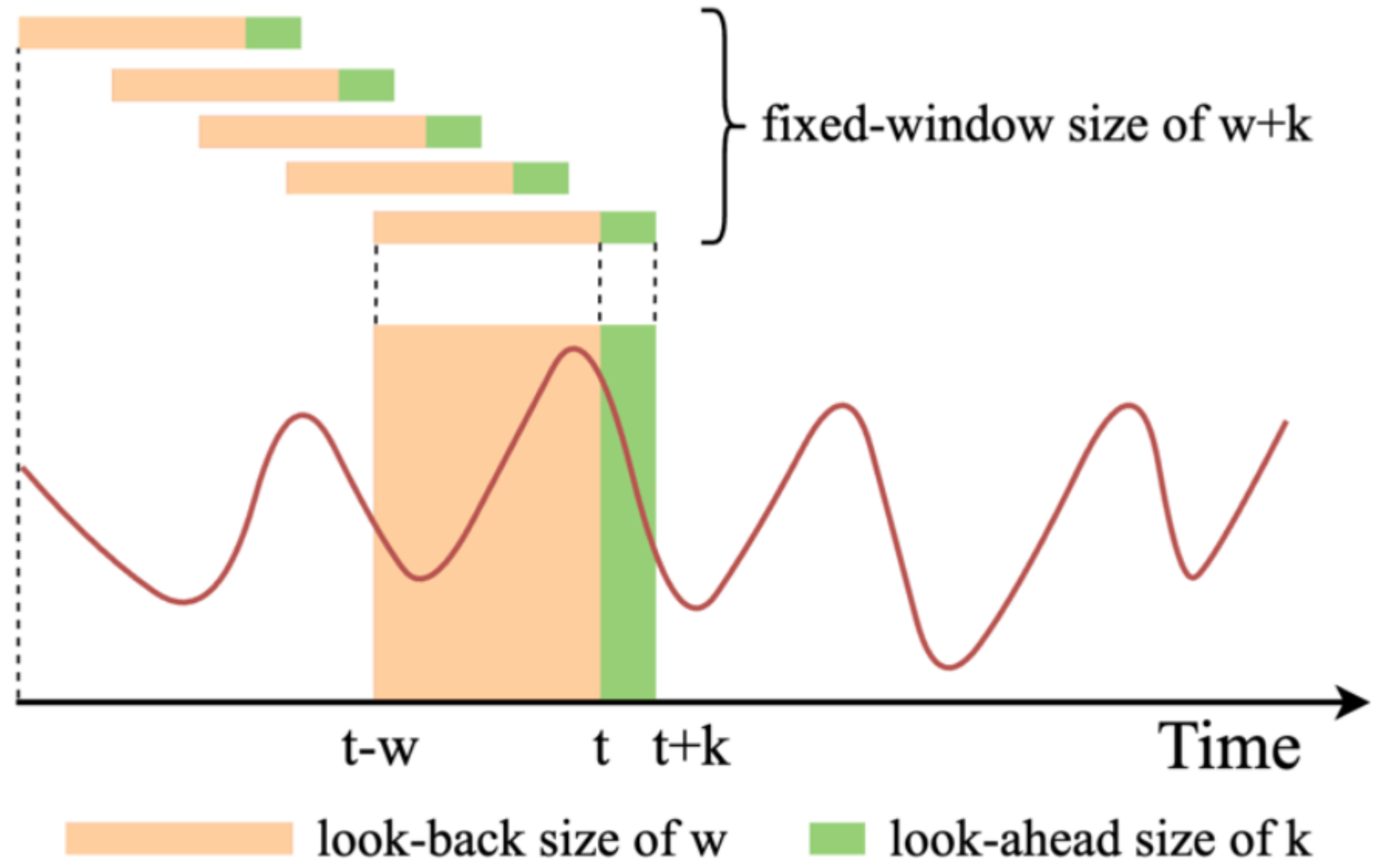
Modeling with time lag



Regression: Y(t+h) = f(Y(t), x1(t), x2(t), x3(t), x4, x5, x6)

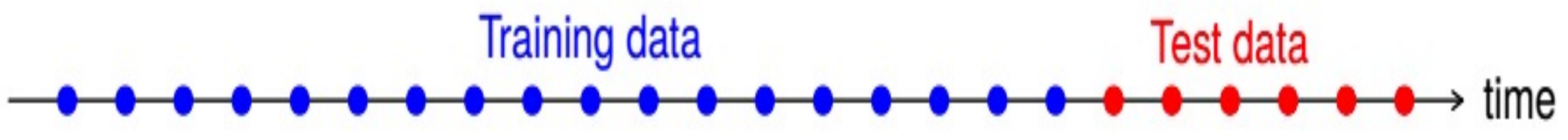
FORECASTING WITH REGRESSION

Sliding window



$$\hat{y}_{t+k} = f_k(\boldsymbol{x}_{t-w}, \dots, \boldsymbol{x}_{t-1}, \boldsymbol{y}_{t-w}, \dots, \boldsymbol{y}_{t-1})$$

TRAINING & EVALUATION



Always train on the earlier timestamps, test and evaluate on the later timestamps

Avoid “signal leakage” !!

Standard quality metrics

Mean absolute error: $MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}|$

Mean squared error: $MSE = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y})^2$

Root mean squared error: $RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y})^2}$

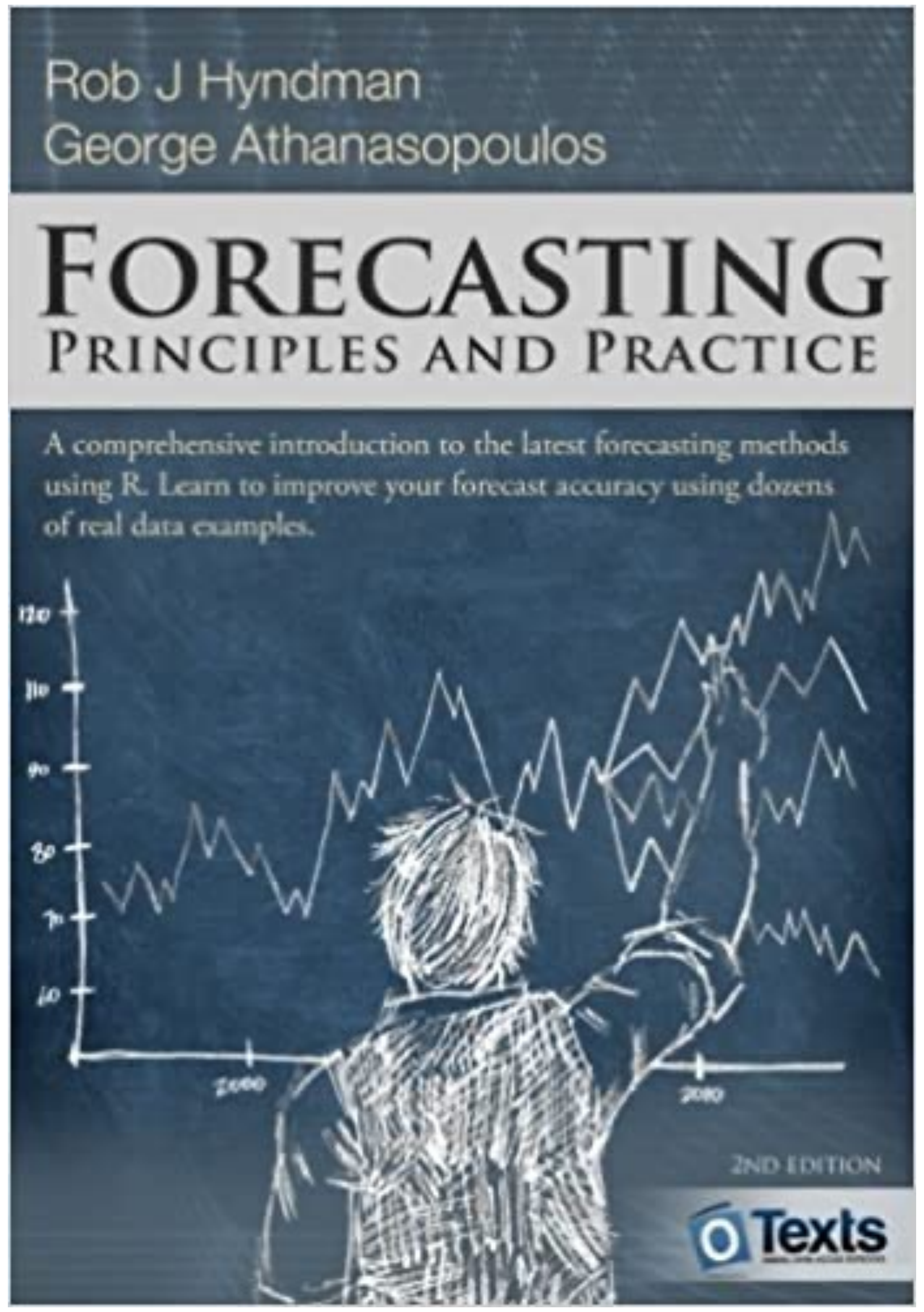
R-squared: $R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2}$

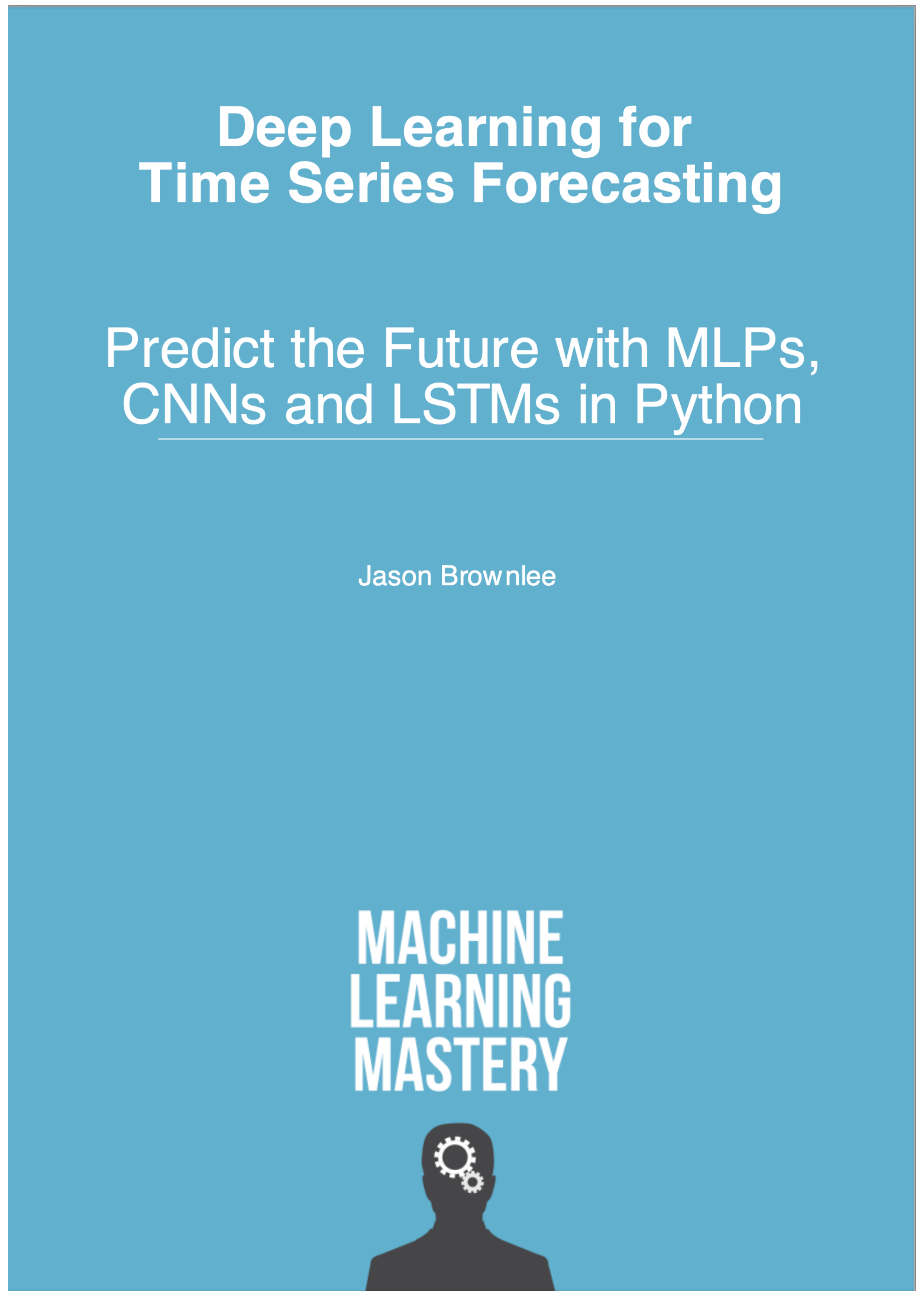
Where,

$\hat{y}$ $-$ *predicted value of y*

$\bar{y}$ $-$ *mean value of y*

MORE BOOKS


Rob J Hyndman
George Athanasopoulos
FORECASTING
PRINCIPLES AND PRACTICE
A comprehensive introduction to the latest forecasting methods using R. Learn to improve your forecast accuracy using dozens of real data examples.
2ND EDITION
OTexts


Deep Learning for
Time Series Forecasting
Predict the Future with MLPs,
CNNs and LSTMs in Python
Jason Brownlee
MACHINE
LEARNING
MASTERY

HIGHER SCHOOL OF ECONOMICS

NATIONAL RESEARCH
UNIVERSITY