TIMESERIES FORECASTING (B)

ALEXANDROS PATELIS (CR/AEU2)

Timeseries Forecasting – Data Analytics & Quantitative Methods
What is about?

Is about **forecasting**!

Not just the weather, not just politics, but forecasting sales, sports, the economy, almost anything.

Forecasting in everyday life!

We all have our own opinion (forecast) on things. But:

- When to trust our ***gut***?
- When to disrupt our ***biased nature***?
- What about ***human intuition***? Expertise? Human heuristics? Biases? Judgment?
- Can we trust a ***computer method***?

Forecasting: Is about common sense. It's about probabilities. It's about calibration. It's about knowing when to trust your gut, someone's expertise, or a computer algorithm.

Internal | CR/AEU2 - Patelis | 2018-03-14

BOSCH

Source: Forecasting – Principles and Practice

Timeseries Forecasting (B)
Agenda

1. INTRODUCTION

Timeseries Forecasting
Why Forecasting?

Why forecasting:

- we want to understand **the past** → to find a model that fits to the observations…
- we need to estimate **the future** → to find a model that can predict the future…
- we have so ***many data*** to use…
  - …We are drowning in information, but starved for knowledge….

- Because most of the decisions that we make involve forecasting!
  - Decision for building a factory requires forecasts about: future demand, technological innovations, cost, prices, competitors' plan, labor, legislation, etc.

Introduction
Some Basic Definitions

Timeserie:

- A sequence of data points, measured typically at **successive points in time** spaced in uniform time intervals.
- They can be ***yearly*** (gross annual product of a country), ***quarterly*** (sales of a product), ***monthly*** (average temperature of a city), ***weekly*** (total arrivals in an airport), ***daily*** (daily value of a stock), ***hourly*** (energy consumption of a building), etc.
- They are used in statistics, signal processing, pattern recognition, econometrics, mathematical finance, weather forecasting, earthquake prediction, astronomy, and largely in **any domain** of applied science and engineering which involves **temporal measurements**.

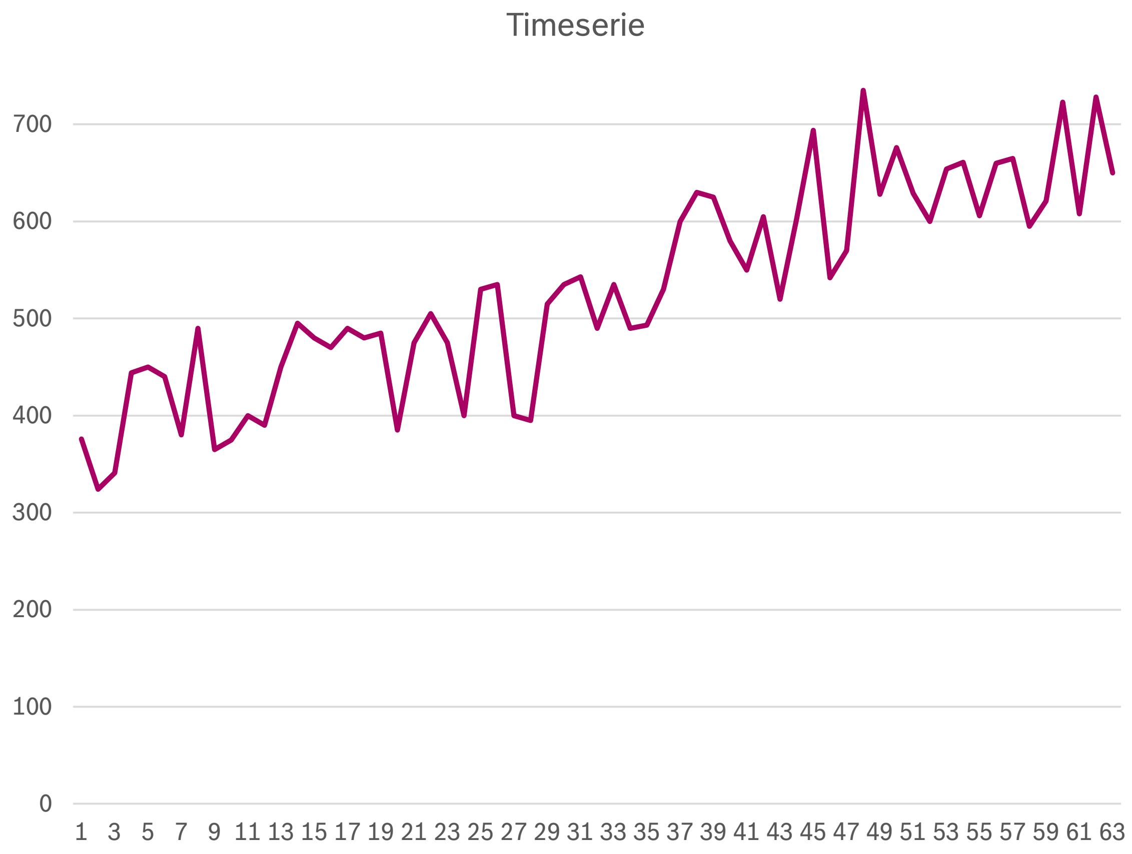
Introduction
Some Basic Definitions

Timeserie Data:

- The original ***historical data*** of the timeserie.

Timeserie Analysis:

- Comprises methods for analysing timeserie data, in order to ***extract meaningful statistics*** and other data characteristics.



Introduction
Some Basic Definitions

Forecasting:

- Definitions:
  - ***To predict how the sequence of observations will continue in the future.***
  - The process of ***making statements*** about events whose actual outcomes have not yet been observed.
  - The process of ***making predictions*** of the future based on past and present data.

- ***Aim***:
  - to ensure that the ***best possible accurate forecasts*** will be estimated, making the most from all available historical data and information.

Introduction
Some Basic Definitions

Forecasting Horizon:

- The length of time into the future for which forecasts are to be estimated.
- Three categories:
  - Short-term: Up to 1 year (for annual data)
  - Mid-term: up to 5 years
  - Long-term: more that 5 years

Forecasting Error:

- The difference between the actual value and the forecasted value of a timeserie.

Forecasting Accuracy:

- An estimation of how good the forecasting process is.

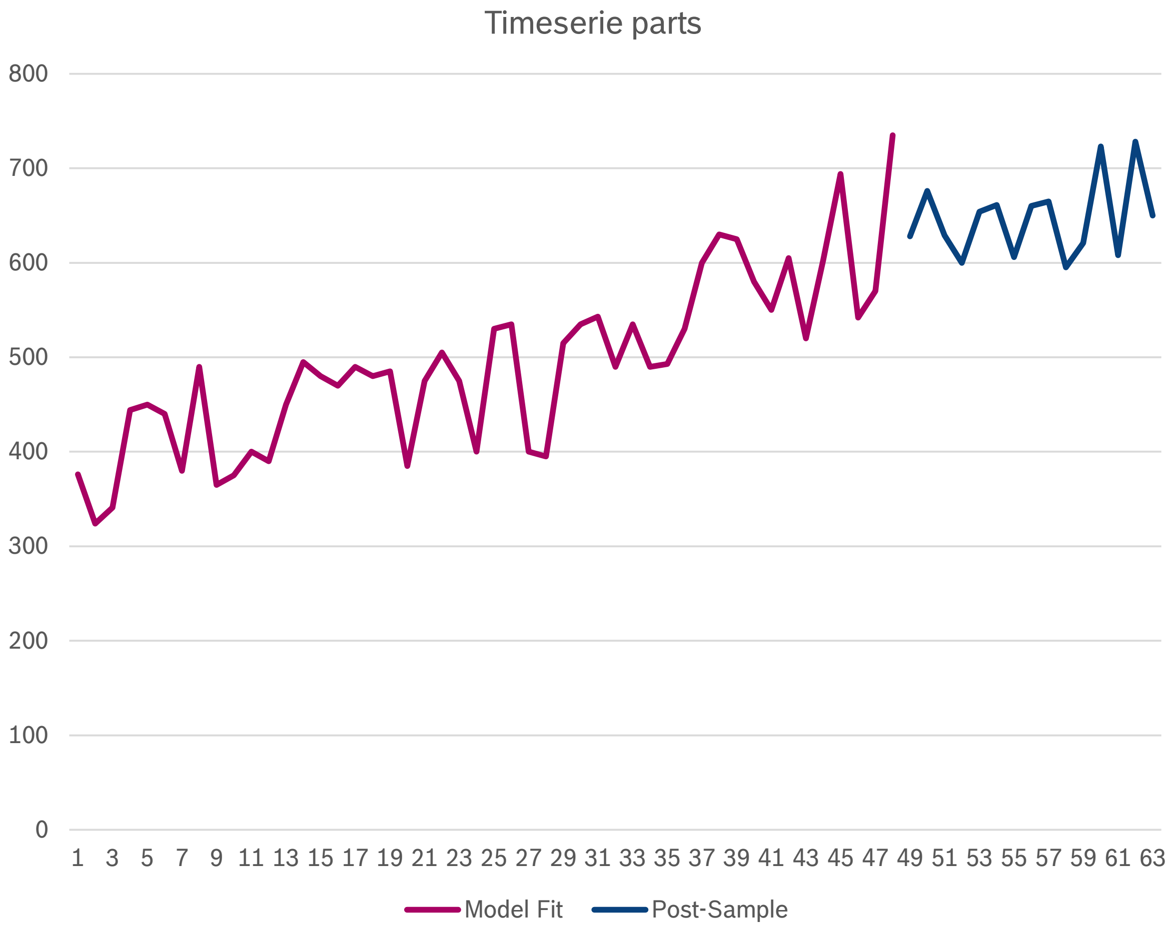
Introduction
Some Basic Definitions

Model-fit:

- A part of timeseries data that we use for ***fitting*** a model, by estimating model-fit forecasts.

Post-sample:

- A part (last part) of timeseries data that we use for ***evaluating*** our model, by estimating post-sample forecasts.



Introduction
Some Basic Definitions

Pattern:

- A ***repeated observation*** that can be ***gathered*** and ***analyzed*** (identified).
  - can be recorded directly in our memory and then used intuitively (implicitly) for prediction purposes, or
  - can be recorded in an external (explicit) location, such as figures recorded in a notebook, and then used in a forecasting method.
    - Example: estimate the driving time we need every day, from our home to our office.
- Patterns are timeseries' main components.
    - Seasonality: the day of the week
    - Trend: the gradual increase in traffic as the population in the suburbs increased
    - Cyclicality: reflected at a time of energy crisis (such as 1973)
    - Randomness: an accident slowing traffic on a day

BOSCH

Introduction
Some Basic Definitions

Relationship:

- The ***influence of one or more factors*** to the forecasted item, which can be discovered by analysing repeated observations
  - the heavier the snow fall, the slower the traffic.
  - a highway under repair increases traffic to alternative roads.
  - Some relationships are very noticeable, other more difficult to identify.
- Accurate forecasts requires the ***identification of key relationships***, such as:
  - those between advertising or promotions and sales volume,
  - the influence of price and revenues fluctuations, and
  - the way in which competitors' actions are reflected in the demand.

Introduction
Limits of predictability

- All techniques are ***extrapolative*** in nature.
  - They work well when the ***future is similar to the past***, or when changes happen to cancel out.
  - No simple and reliable way to predict when established patterns or relationships change.
- Although the complains about forecasting have increased over the years, the number of requests for forecasting have also increased.
  - Not a paradox!
  - In a turbulent environment with high uncertainty, ***the need for forecasts is great***.

➢ ***Forecasting is not crystal-balling***. Realistic expectations should be set.

Introduction
Factors that influence predictability (1)

- ***Number of Items***: The larger the number of items involved, the more accurate the forecasts.
  - Because of the statistical law of large numbers, the size of forecasting error, and therefore the accuracy, decreases as the number of items being forecast increases, and vice versa.
    - Thus, it is more accurate to predict the number of telephone calls arriving at a switching station during a five-minute interval that the number of personal computers sold on a certain day.
- ***Homogeneity of Data***: The more homogeneous the data, the more accurate the forecasts, and vice versa.
    - Thus, data referring to a single region can predict seasonality more accurate that data covering many regions with varying weather patterns.

Introduction
Factors that influence predictability (2)

- ***Elasticity of Demand***: The more inelastic, the more accurate the forecasts.
    - Thus, the demand for necessities can be forecast with a higher degree of accuracy than the demand for luxuries.
  - Related to the elasticity of demand is the influence of business cycles. Such cycles have the least impact on inelastic demand and the greatest impact on elastic demand.

- ***Competition***: The greater the competition, the more difficult it is to forecast.
  - Competitors can use the forecasts to change the course of future events and thus invalidate the forecasts.

WHERE TO USE FORECAST

Introduction
Where to use forecast? (1)

Introduction
Where to use forecast? (2)

- Different markets can benefit from analyzing their data and use forecasting techniques:
  - Telecommunications.
  - Online Sector.
  - Insurance.
  - Banking and Stock Market.
  - Energy sector.
  - Agriculture.
  - Management Consulting.
  - Transportation and Logistics.
  - Tourism Industry.
  - Digital Media and Press.
  - Advertisement.
  - Retail.
  - Food industry.
  - Manufacturing.
  - Gaming industry.
  - Politics.
  - Car rental companies.
  - Social Media Analytics.
  - Data Analytics.
  - etc.

Where to use forecast? (3)

- ***Marketing***:
  - ***Sales of each product type***, sales by geographic area, sales by customer, prices, inventory levels, sales promotions
  - ***Total sales***, product categories, major products, product groups, prices, advertising selection, sales promotions, new product introduction
  - ***Social attitudes***, Social trends, tastes, area of social concern, total sales, product categories, general economic categories, prices,
  - ***New product introduction***, saturation points of existing products, total sales, customer preferences and tastes
- ***Production***:
  - ***Demand*** of each product, plant loading,
  - Total demands, demand of product categories and product groups, scheduling, employment level, costs, new product introduction, product analysis, pricing
  - ***Costs, budget allocation,*** buying or ordering equipment and machinery, employment level, new product development, data driver product design
  - Costs, ***facility investments, expansion of plant and equipment***, ordering of heavy machinery and equipment, demand of production facilities, new technologies

Introduction
Where to use forecast? (4)

- ***Top Management***:
  - Total sales, sales breakdown, ***pricing***
  - Demand for sales, costs and other expenses, cash position, ***general economic conditions***, control objectives
  - Total sales, costs and other expenses, ***social and economic trends***, goals, ***objectives and strategies***, new products, pricing policies

- ***Purchasing***:
  - Production, ***cash availability***, purchasing of supplies and materials
  - Demand for products, demand for materials, lead time for purchasing
  - Demand for products, demand for raw and other materials, ***Real estate values***
  - ***Contracts for buying raw materials***, customer preferences and tastes

Introduction
Where to use forecast? (5)

- ***Finance & Accounting***:
  - ***Sales revenue***, production costs, inventory costs, leading indicators, cash inflows and outflows
  - Availability of money, ***interest rates***, total demand, inventory levels, cash flows analysis, short-term borrowing, prices
  - Budget allocation, cash flows, fiscal and monetary policies
  - ***Trends in taxation rating***, depreciation, Concept of free market, total sales, investment selections, capital expenditure, allocation of resources, capital programs, cash flows analysis
- ***Inventory***:
  - ***Demand for each products production***, demand for material, demand for semi-finished products, weather conditions
  - ***Possible strikes*** in suppliers or transportation facilities
  - Total sales, expansion of warehouses

Introduction
Where to use forecast? (6)

- ***Research & Development***:
  - ***New product introduction***, R&D selections, Available technologies, Capital investment
  - Areas of technological innovations, R&D Selections, available technological opportunities, Total sales, Technological, social,
  - ***political and economical conditions of future***, new product development
  - patient Diagnostics, patients states projections, patient disease monitoring, disease early detections,
  - **abnormally detections**
  - Identify Customer Shifting patterns, ***Fraud detection***, Risk analysis, customer profiles
  - ***Energy load forecasting***, reduction of expanding electrical grid, heating management
  - ***predictive error recognition***, predictive fail recognition, service and maintenance recommendations
  - ***social media data analysis***, targeting marketing, product suggestions based on users behaviour
  - prediction of traffic patterns and traffic jams

Introduction
Where to use forecast? (7)

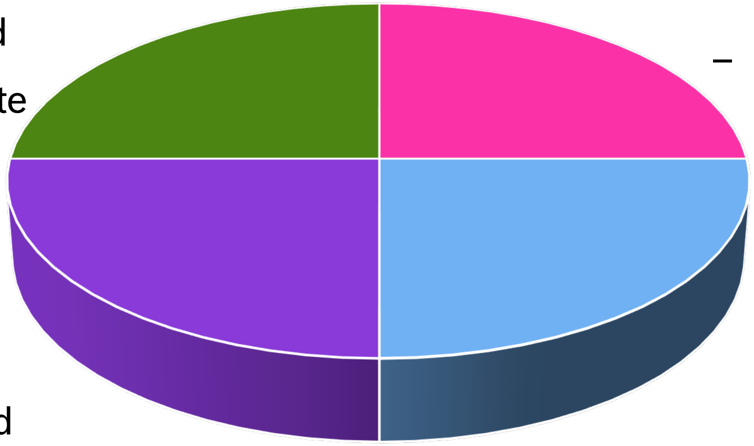
- ***Environment***:
  - ***Weather conditions***, Energy Consumption, Usage of water resources, ***Air Pollution levels***, Noise pollution, Water & Soil Pollution, Energy demand, Heating and Cooling demand,
  - ***Crops, Wind conditions, Rainfall***, Snowfall, Cloud Cover, Sunshine, Temperature, Humidity, Pressure
  - General environmental constrains (pollution level, availability of raw materials, etc.)

- ***Personnel***:
  - ***Human resource planning***
  - Training schedules
  - labor turnover, retirement age, absenteeism, tardiness

- ***Economic Unit***:
  - Level of economic activity
  - General economic conditions, turning point in economy, level of economic activity
  - State and type of economy, level of economic activity, sales of industry

Introduction
Who can be a forecaster?

- Making accurate forecasts does not depend on any special talent – ***it is a learned skill***.
- A forecaster should have:
  - ***Thinking*** Style:
    - Open-Minded, Intelligent and Curious, Reflective, Numerate
  - ***Philosophic*** Outlook:
    - Cautious, Humble, Nondeterministic



  - ***Forecasting*** Style:
    - Pragmatic, Analytical, Dragonfly-Eyed, Probabilistic, Thoughtful Updaters, Intuitive Psychologist
  - ***Work*** Ethic:
    - Growth Mindset, Effort and passion to achieve a goal.

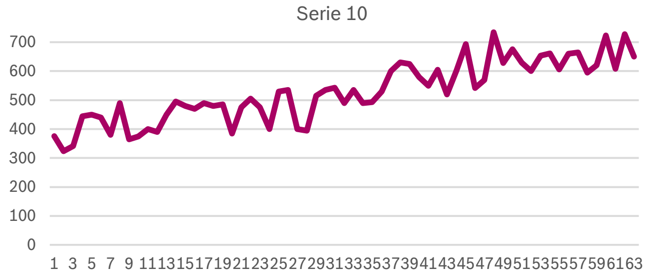
TIMESERIE EXAMPLES

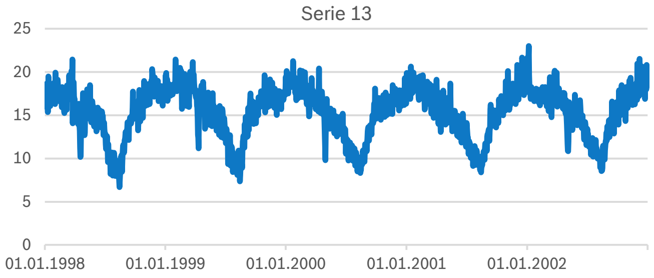
Introduction

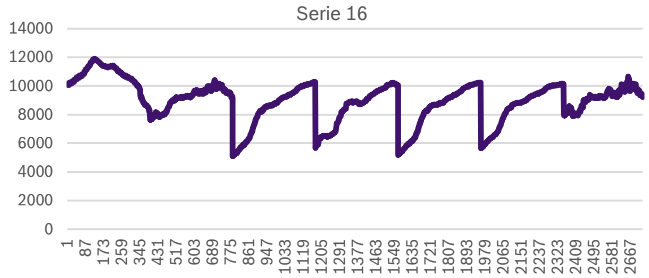
Time serie examples

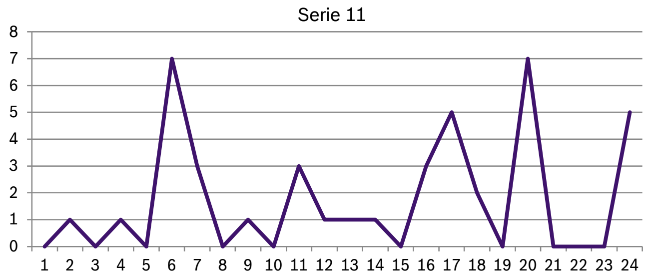
Introduction

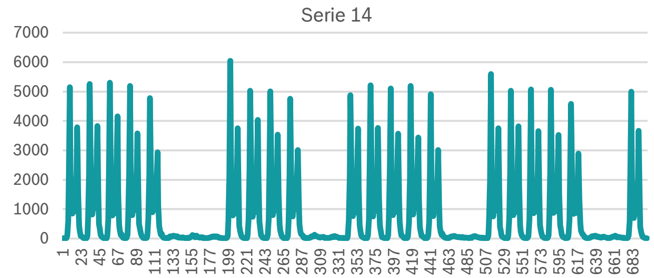
Time serie examples

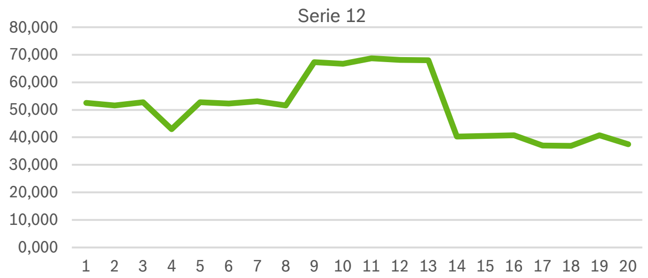


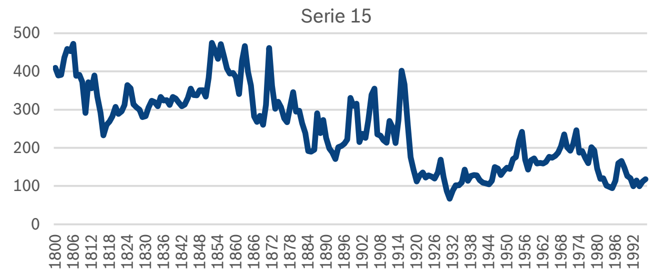












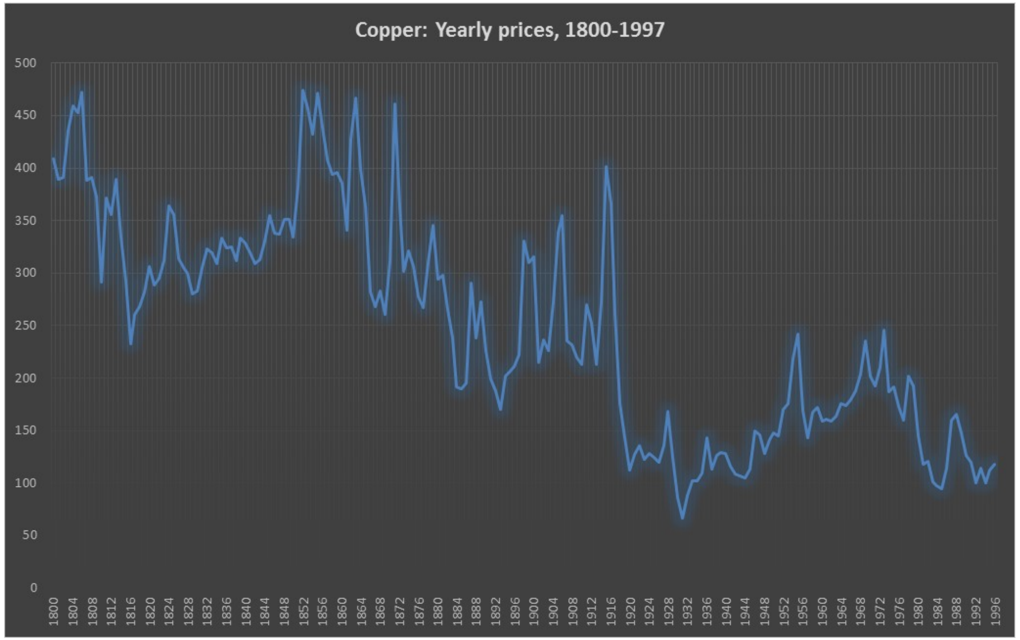
2. TIMESERIE CHARACTERISTICS & COMPONENTS

"The most reliable way to forecast the future is to try to understand the present."
John Naisbitt (Author in area of future studies)

BOSCH

Timeserie Characteristics & Components
Graphical Representation

- An important tool for the analysis of the data and the forecasting procedure.
- Is a 2-dimensional representation of the timeseries available data points (data values in y-axis, time in x-axis).
- We can observe and understand the ***qualitative characteristics*** of the timeseries, and helps us to choose between alternative methodologies and tools, and to locate the one which will have the best results and the smallest errors.
- In addition, the graphical representation can reveal ***outliers*** and ***fault values***.

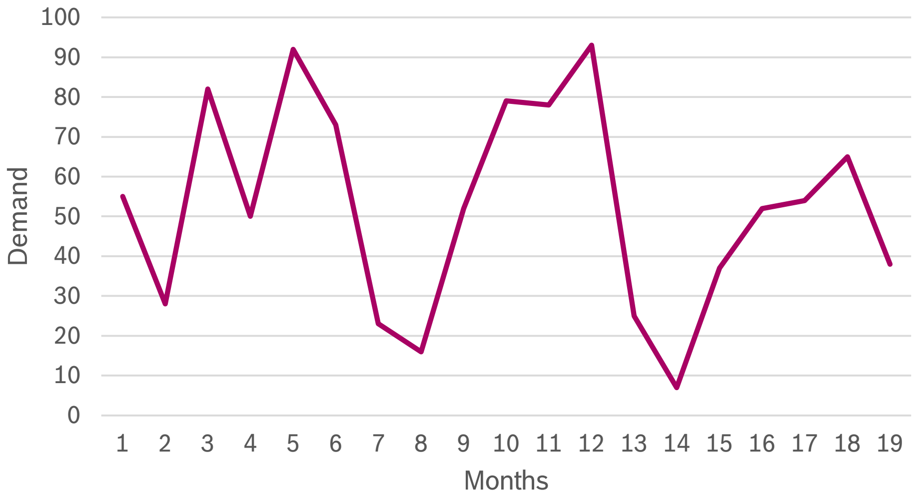


Internal | CR/AEU2 - Patelis | 2018-03-14

BOSCH

Timeserie Characteristics & Components
Studying the graph

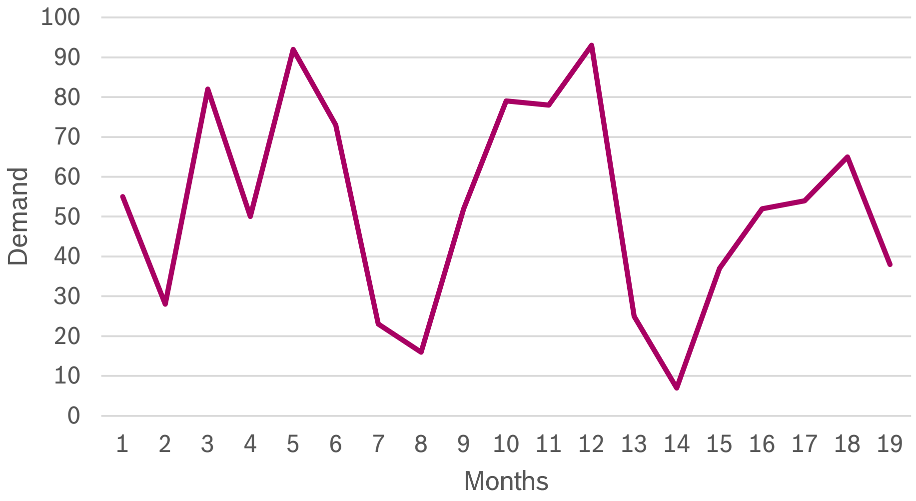
- Can I derive the characteristics by just studying the graph?
  - We can see ***3 underlying cycles*** in the demand, each lasting about 6 months.
  - Within each of these cycles the demand rises to a peak and then declines. But why?
    - Perhaps the product sells well in the summer and is sold in both Europe and Australia, so the 6-month peak coincide with summers in the northern and southern hemispheres.
    - Perhaps the product is an alcoholic drink that sells well in the summer and also around the festive season.
- Forecast:
  - Extend the pattern for 6 more months…



Timeserie Characteristics & Components

Studying the graph

- There is only a problem...
  - the apparent ‘demand figures’ on the graph are actually **random numbers** from 1 to 100, generated by a computer.
  - There is ***no systematic pattern*** underlying the numbers at all so forecasts based on the 6-month cycle are likely to be well off target.

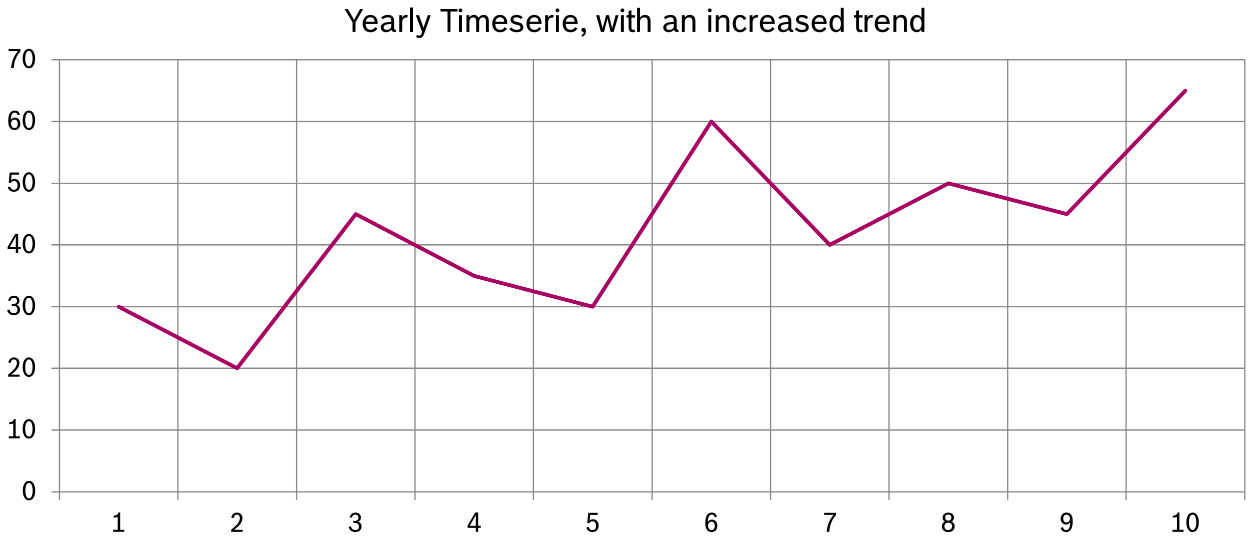


- Not only we are adept at finding patterns where there are none, but we are also ***brilliant at inventing theories*** to explain these patterns.

Timeserie Characteristics & Components
Graphical Representation

Trend/Stationary:

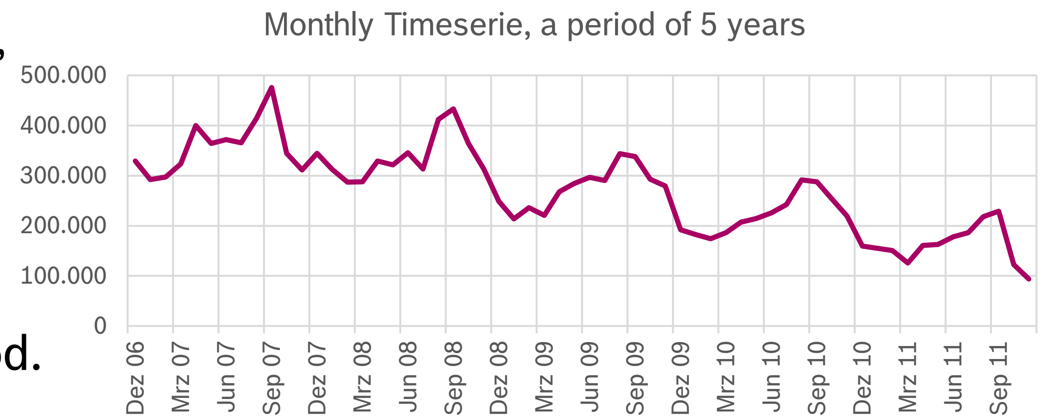
- The timeserie data tend to increase or decrease over time? Or they are stationary?
- A trend exists when there is a ***long-term increase or decrease*** in the data.
  - It does not have to be linear.
- Sometimes trend can be referred as "***changing direction***" when it might go from an increasing trend to a decreasing trend.



Timeserie Characteristics & Components
Graphical Representation

Seasonal:

- Is there a ***regularly repeating pattern*** of highs and lows related to calendar time, such as seasons, quarters, months, weekdays, and so on?
- A seasonal pattern exists when a series is influenced by:
  - ***calendar effects*** (e.g., the quarter of the year, the month, or day of the week),
  - ***institutional influences*** (e.g. stock shares, bonds),
  - ***Weather conditions***,
  - ***Expectations*** (e.g. prices before Christmas).

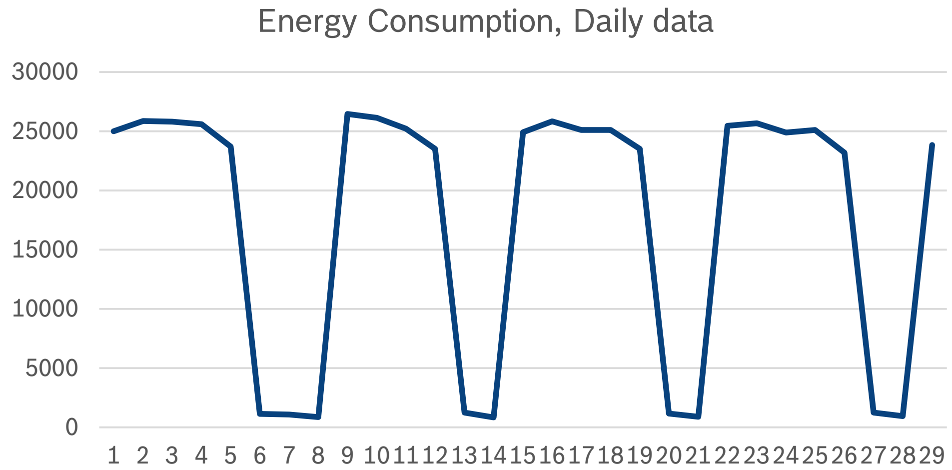


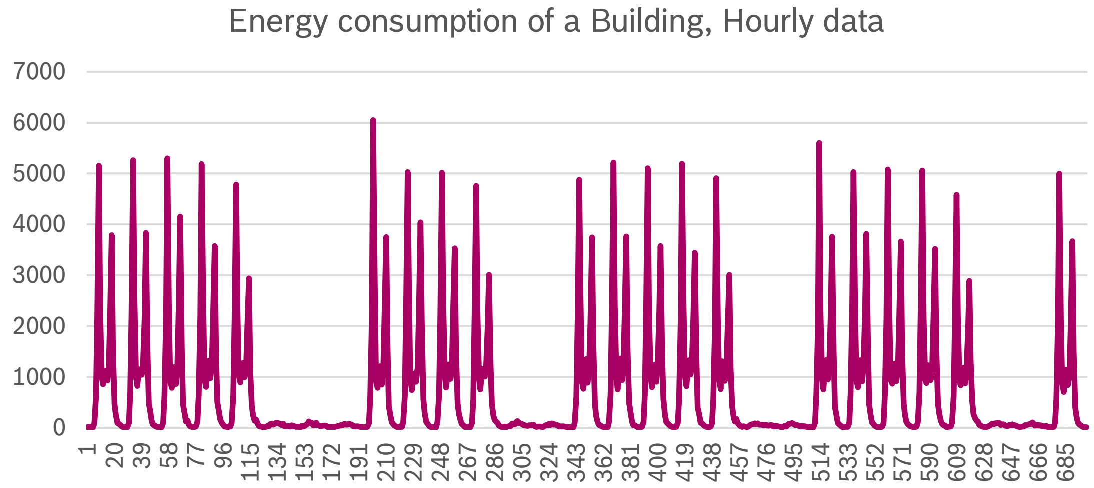
➢ Seasonality is always of a fixed and known period.

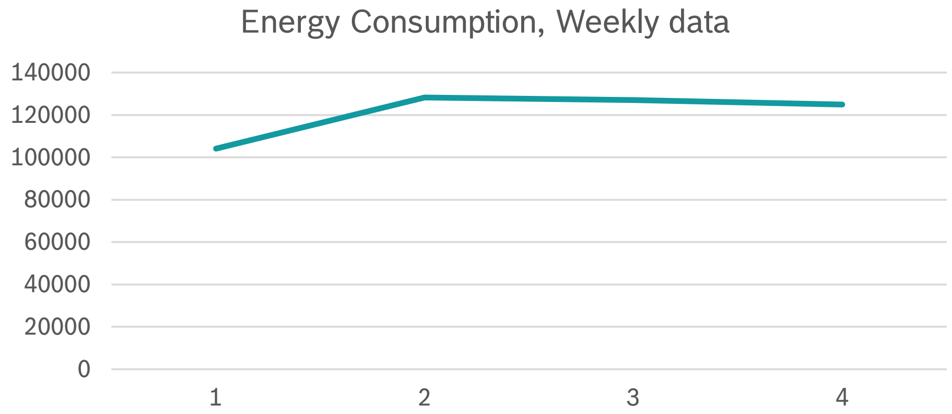
Timeserie Characteristics & Components
Graphical Representation

Seasonal:

- In some cases, a ***multiple seasonality*** may exist!
  - Example: Hourly data may have both:
    - 24H / day seasonality, and
    - 7D / week seasonality.





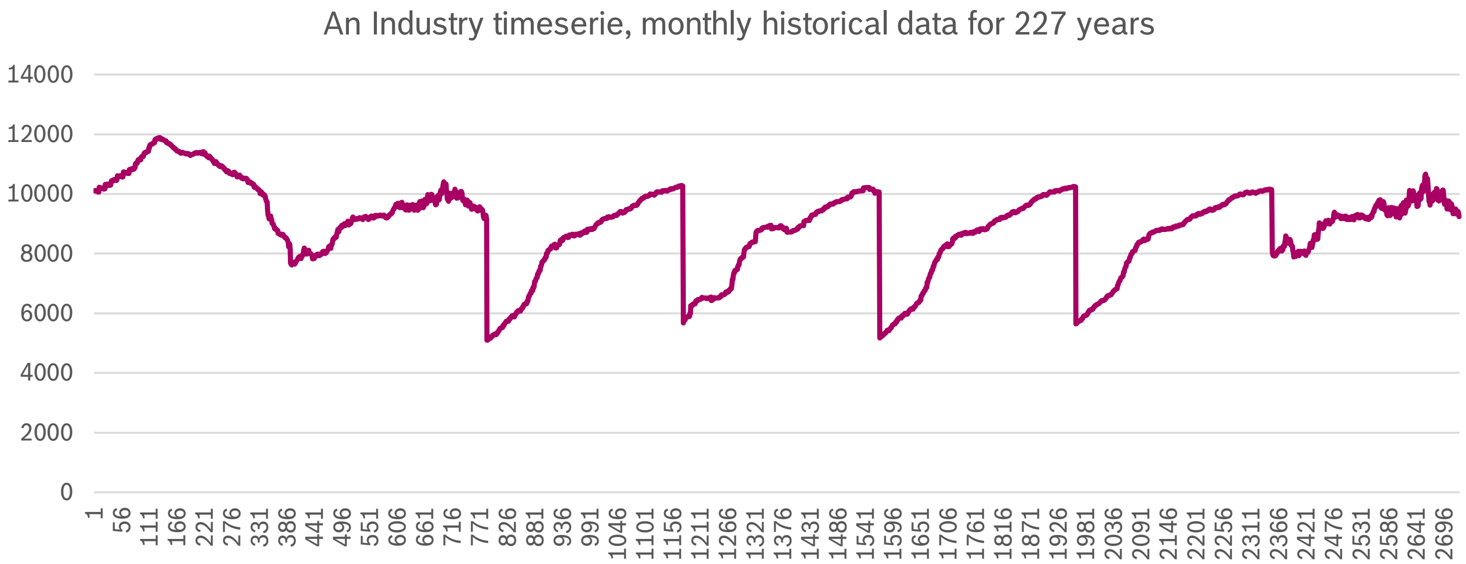


Timeserie Characteristics & Components
Graphical Representation

Cycle:

- Is there a ***long-run cycle*** or period ***unrelated to seasonality factors***?
- A cyclic pattern exists when data exhibit rises and falls that are not of fixed period.

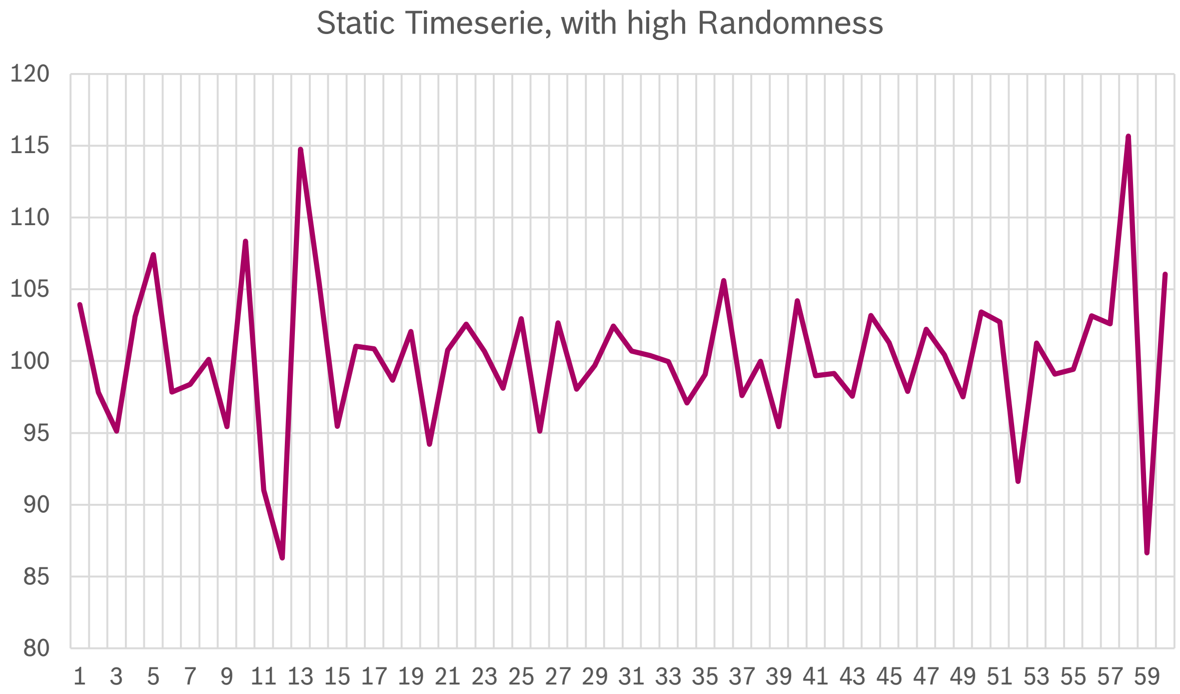
➢ The duration of these fluctuations is usually ***of at least 2 years***.



Timeserie Characteristics & Components
Graphical Representation

Randomness:

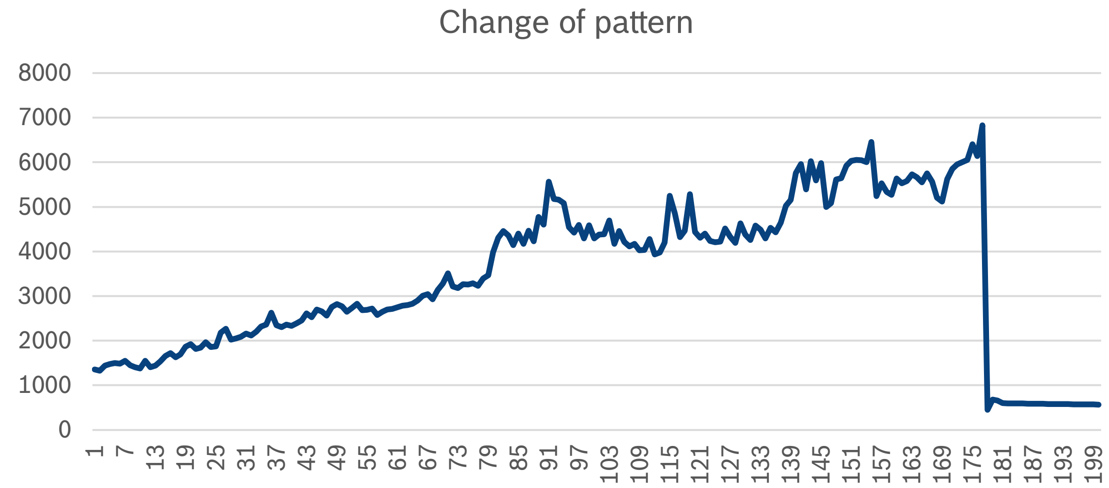
- How large randomness exists on the data?

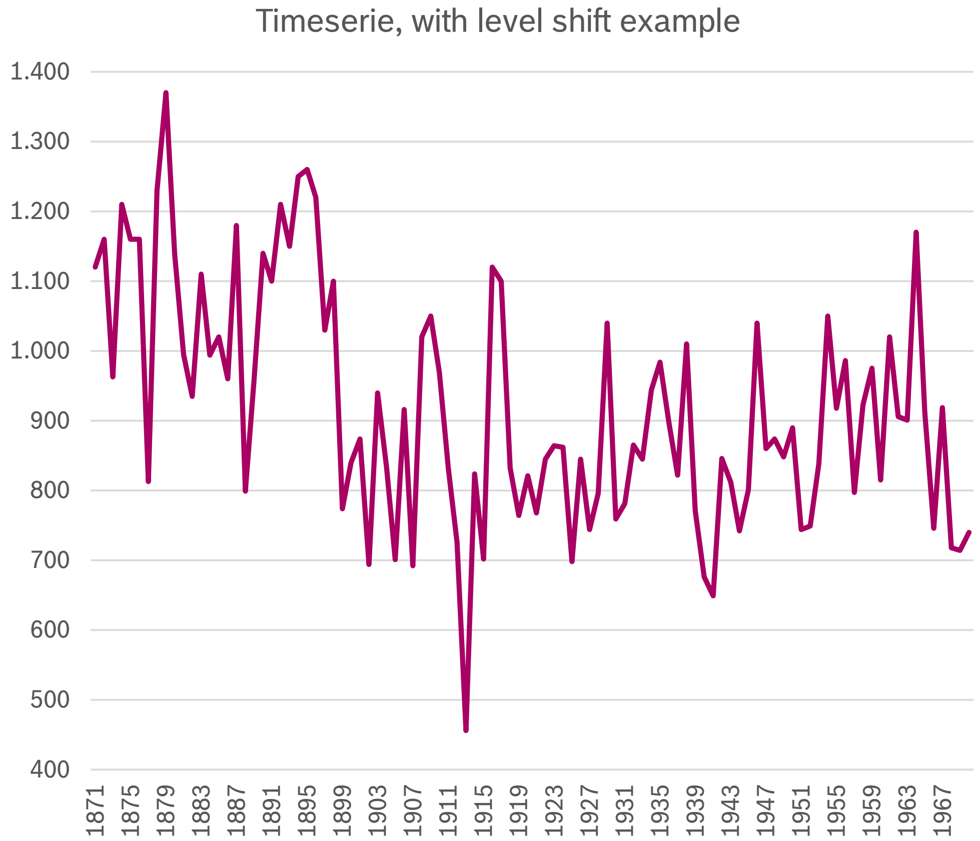


Timeserie Characteristics & Components
Graphical Representation

Level-shift:

- Are there any ***abrupt changes*** to the level of the timeserie?
- Can we identity the cause?



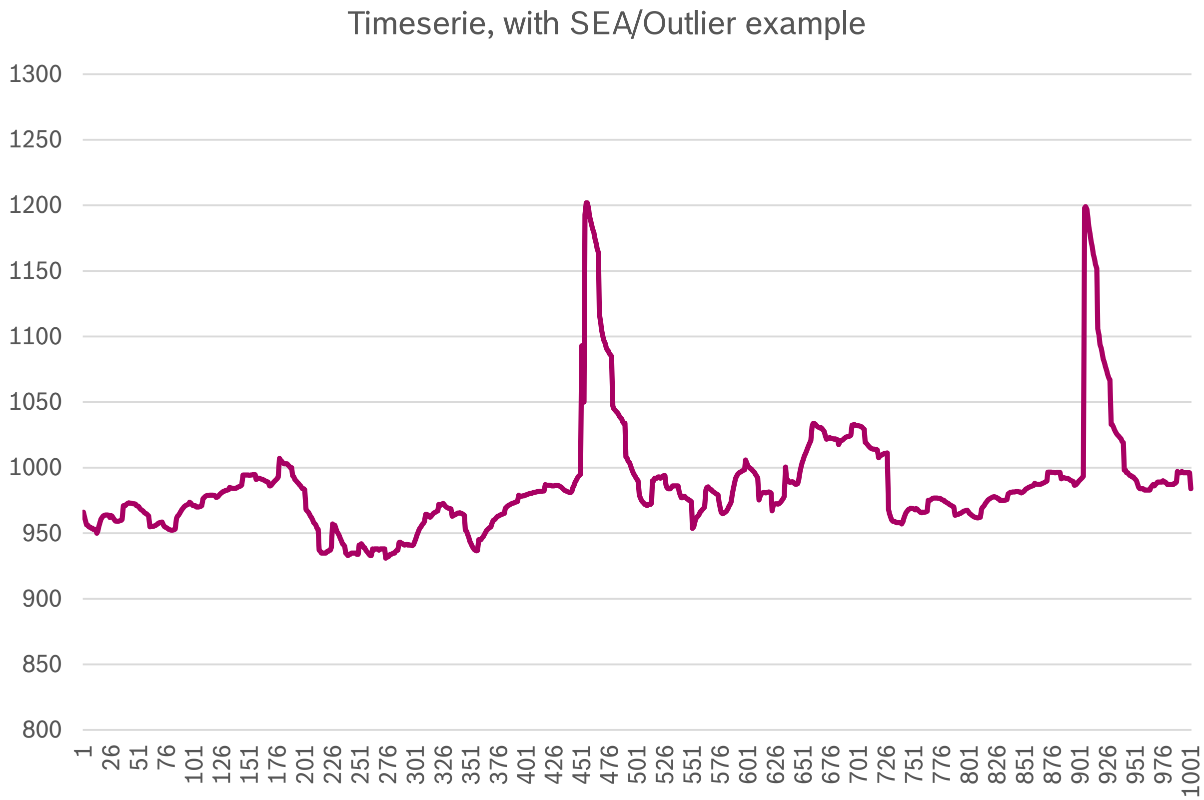


Internal | CR/AEU2 - Patelis | 2018-03-14

Timeserie Characteristics & Components
Graphical Representation

Outliers:

- Are there data points ***far away*** from the other data?
- ***Can we identify the cause***?
  - advertisement, promotion of product?
  - increase or decrease or price?
  - creation of a substitute product?
  - unusual weather conditions?
  - strikes? other?

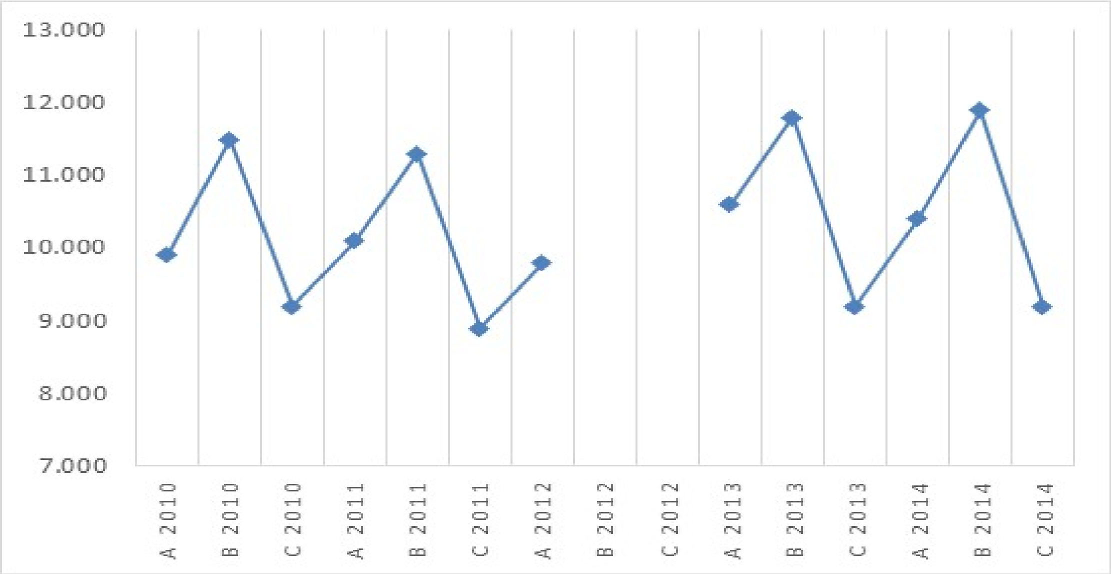


Timeserie Characteristics & Components
Graphical Representation

Fault or Missing values:

- Are some ***data point values missing*** from the timeseries data?
- Do we believe that some data point values are fault?

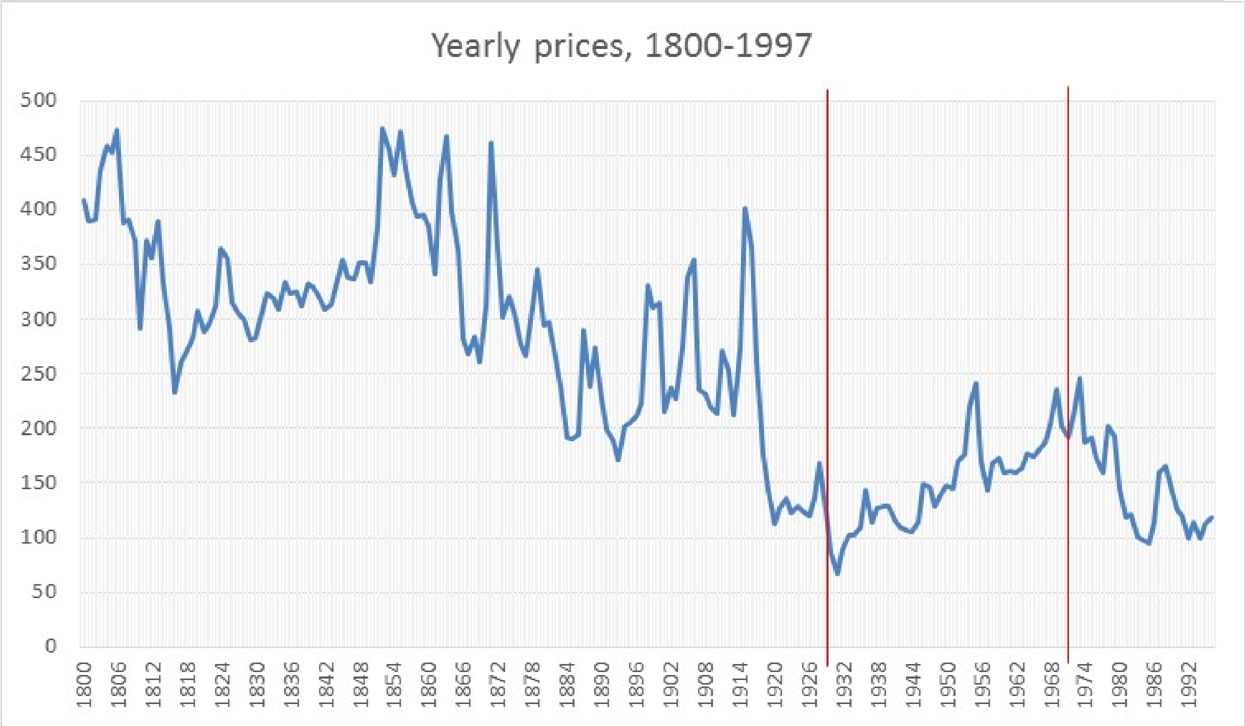
➢ These values ***must be corrected*** before applying forecast.



Timeserie Characteristics & Components
Graphical Representation

Volume of historical data:

- Choose carefully the volume of historical data (depending on our forecasting horizon)
  - Example: taken from the book "*Micro-Economics*", Makridakis, Wheelwright and Hyndman (1998).



3. TIMESERIE STATISTICAL ANALYSIS

"Statistics may be defined as 'a body of methods for making wise decisions in the face of uncertainty'."

W. Allen Wallis (American Economist)

Timeserie Statistical Analysis
Basic Statistical Analysis

- ***Average***: $\bar{Y} = \frac{1}{n} \times \sum_{i=1}^{n} Y_i$
- ***Standard Deviation***: $\sigma = \sqrt{\frac{\sum_{i=1}^{n}(Yi - \bar{Y})^2}{n-1}}$
  - shows how much variation or dispersion exists from the average (mean), or expected value.
  - ***Low*** standard deviation → indicates that the data points tend to be ***very close to the mean***.
  - ***High*** standard deviation → indicates that the data points are ***spread out over a large range*** of values.

Timeserie Statistical Analysis
Basic Statistical Analysis

- ***Covariance***: $cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}[(X-\bar{X})(Y-\bar{Y})]$
  - Is a measure of how much two variables change together.
  - shows ***the tendency in the linear relationship*** between the variables.
  - ***Sign*** of covariance:
    - ***Positive*** covariance: the variables tend to show ***similar behaviour***
      - The greater values of one variable mainly correspond with the greater values of the other variable, and the same holds for the smaller values.
    - ***Negative*** covariance: the variables tend to show ***opposite behaviour***
      - The greater values of one variable mainly correspond to the smaller values of the other.
  - ***Magnitude*** of covariance:
    - Not always easy to interpret.
      - The normalized version of the covariance (correlation coefficient) shows by its magnitude the strength of the linear relation.

Timeserie Statistical Analysis
Basic Statistical Analysis

- Linear Correlation Coefficient: $$r(X,Y) = \frac{\sum_{i=1}^{n}[(X-\bar{X})(Y-\bar{Y})]}{\sqrt{\sum_{i=1}^{n}(X-\bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y-\bar{Y})^2}}$$
  - ***Positive*** correlation:
    - If x and y have a strong positive linear correlation, r is close to +1.
    - An r value of exactly +1 indicates a ***perfect positive fit***.
    - Positive values indicate a relationship between x and y such that as values for x increase, values for y also increase.
  - ***Negative*** correlation:
    - If x and y have a strong negative linear correlation, r is close to -1.
    - An r value of exactly -1 indicates a ***perfect negative fit***.
    - Negative values indicate a relationship between x and y such that as values for x increase, values for y decrease.
  - ***No*** correlation:
    - If there is no linear correlation or a weak linear correlation, r is close to 0.
    - A value near zero means that there is a random, nonlinear relationship between the two variables.

- *+1, -1: Data on a straight line*
- *$r > 0.8$: Strong*
- *$0.8 > r > 0.5$: Average*
- *$r < 0.5$: Weak*

- Note that r is a ***dimensionless quantity***; that is, it does not depend on the units employed.

Timeserie Statistical Analysis
Basic Statistical Analysis

- ***Autocorrelation Coefficient***: $ACF_k = \frac{\sum_{i=1+k}^{n}[(Y_i-\bar{Y})(Y_{i-k}-\bar{Y})]}{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}$
  - Is the cross-correlation of a variable with itself.
    - Informally, it is the similarity between observations as a function of the time lag between them.
  - It is the mathematical tool for ***finding repeating patterns***.
- ***Coefficient of Determination***: $R^2 = \frac{\sum_i(\hat{y}_i-\bar{y})^2}{\sum_i(y_i-\bar{y})^2} = 1 - \frac{\sum_i(y_i-\hat{y}_i)^2}{\sum_i(y_i-\bar{y})^2}$
  - How close the data are to the fitted regression line.
  - It is the percentage of the response variable variation that is explained by a linear model.
  - R-squared is always between 0 and 100%:
    - 0% indicates that the model explains none of the variability of the response data around its mean.
    - 100% indicates that the model explains all the variability of the response data around its mean.
  - In general, the higher the R-squared, the better the model fits your data.

Timeserie Statistical Analysis
Basic Statistical Analysis: Correlation & Causation

- ***Correlation and Causality***: Caution when looking for correlation between variables!
  - Sometimes useful when making predictions:
    - A San Francisco company suggested that ***orange used*** cars are more ***reliable*** than used cars in other colors.
    - A US online lender found that people default on loans more often when they complete their loan application forms using only ***capital*** letters.
  - But, sometimes there is no causality:
    - Strong correlation between ***Brazil's population*** in each year since 1945 and the ***average cost of a train journey*** in Britain in these years. So, can we blame Brazilians when we get to the station to discover fares have risen again?
    - We can't ban ***ice cream*** because a spike in sales correlated with an increase in ***deaths by drowning***! Here the correlation arises because of a third factor: the weather.

➢ ***Just because 2 things are correlated it doesn't necessarily mean that one is causing the other.***

Timeserie Statistical Analysis
Basic Statistical Analysis: Correlation & Causation

- ***Open topic***: Do we need to worry about understanding why two things are correlated?
  - Some people say NO.
    - They argue that worrying about why these is a correlation is a waste of time. Petabytes allows us to say that correlation is enough.
    - "*If the computer finds things I'm unaware of, I don't care what they are just so long as they forecast. I'm not trying to explain*", (Richard Berk, professor of Criminology and Statistics, Warton School of the University of Pennsylvania).

- Predicting without understanding might work as long as nothing fundamental changes.
- But, if it does, we can be ***seriously mislead***.

Timeserie Statistical Analysis
Basic Statistical Analysis: Correlation & Causation

- In a past election year, when inflation was 4%, unemployment was 1.2 million and growth was 2%, a method estimated that government will receive 45% of the votes.
  - But, the government ***only received 32%*** of the votes, thus we are disappointed by our formula accuracy.
  - To try to improve on this, we expand the range of information to include the annual ***number of umbrellas*** left behind on the tube during election years. Now the method estimates a 31% of the votes – a massive improvement in accuracy!!
  - Did we discover some deep insight into voter behaviour?
  - No, we have just picked up on a ***coincidental short term association*** between voting behaviour and lost umbrellas which is unlikely to prevail in the future.

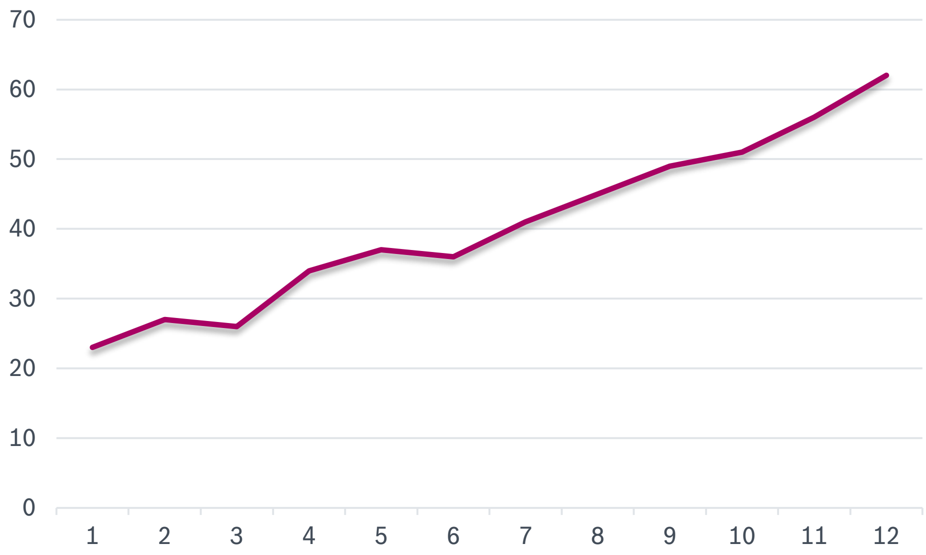
➢ ***Simplicity*** and ***common sense*** are the antidotes to overfitting.

Timeserie Statistical Analysis
Basic Statistical Analysis

$$GrowthRate = \frac{\frac{1}{ppy}\sum_{i=n-ppy+1}^{n}(Y_i) - \frac{1}{n-ppy}\sum_{i=1}^{n-ppy}(Y_i)}{\frac{1}{n-ppy}\sum_{i=1}^{n-ppy}(Y_i)} 100$$

- ***Growth Rate***:
  - Measures the ***increasing or decreasing trend*** of a timeserie for a given time duration.
  - Usually, it compares the values of the observations of the last year with the rest of the observations.

t	Y(t)
1	23
2	27
3	26
4	34
5	37
6	36
7	41
8	45
9	49
10	51
11	56
12	62



$$GR = \frac{(\frac{1}{4}\times\sum_{i=9}^{12} Y_i) - (\frac{1}{8}\times\sum_{i=1}^{8} Y_i)}{\frac{1}{8}\times\sum_{i=1}^{8} Y_i} \times 100\%$$

$$= \frac{\frac{49+51+56+62}{4} - \frac{23+27+\cdots+45}{8}}{\frac{23+27+\cdots+45}{8}} \times 100\%$$

$$= \frac{\frac{218}{4} - \frac{269}{8}}{\frac{269}{8}} \times 100\% = \frac{54{,}5 - 33{,}625}{33{,}625} \times 100\%$$

$= 62{,}08$ %

4. TIMESERIE PREPARATION & ANALYSIS

"By failing to prepare, you are preparing to fail"

Benjamin Franklin

Timeseries Preparation & Analysis

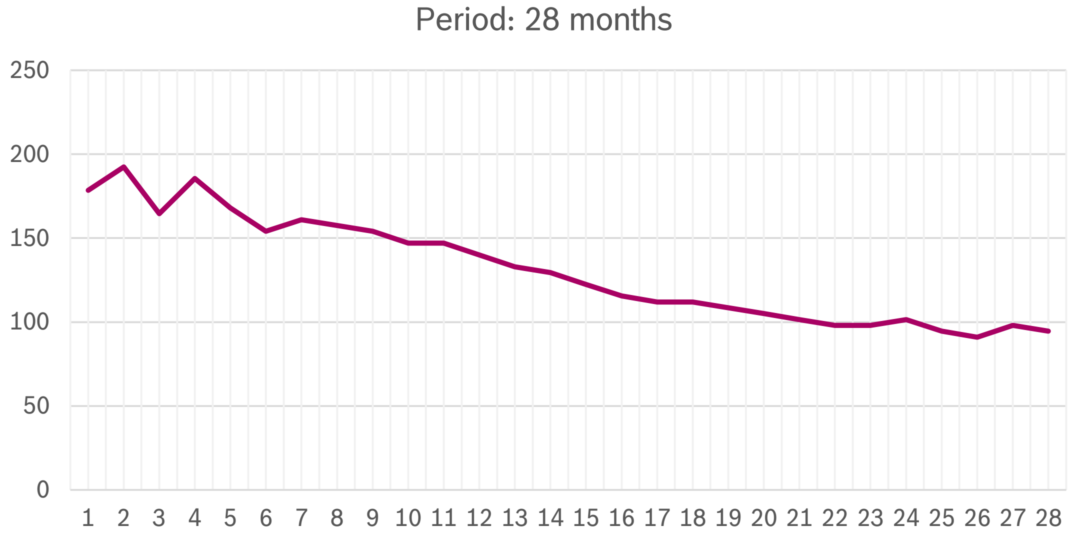
Topics

- Volume of historical data.
- Managing Fault / Missing values.
- Non-Working Days Adjustments.
- Special Events & Actions.
- Estimate Components.

VOLUME OF HISTORICAL DATA

Timeserie Preparation & Analysis
Volume of historical data (1)

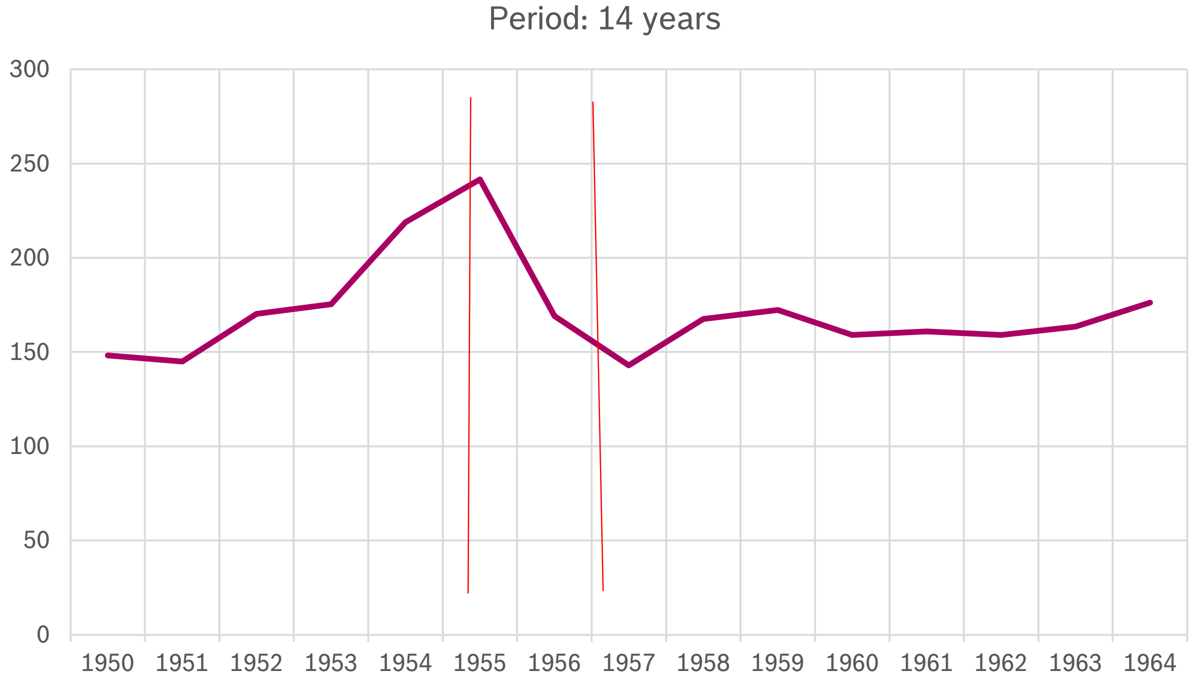
- In general, it is good to have ***as many historical data as possible***.
  - Not all historical data will be used in a method, but they will help to identify timeseries characteristics.
  - We have to ***act with caution***, when selecting historical data, since ***they will affect the results*** of our model selection and prediction results.
  - Example: Monthly prices, 28 months
    - ***Decreased trend***
    - R^2 = 0,974
    - Low error variance



Source: "Micro-Economics", Makridakis, Wheelwright & Hyndman, 1998).

Timeserie Preparation & Analysis
Volume of historical data (2)

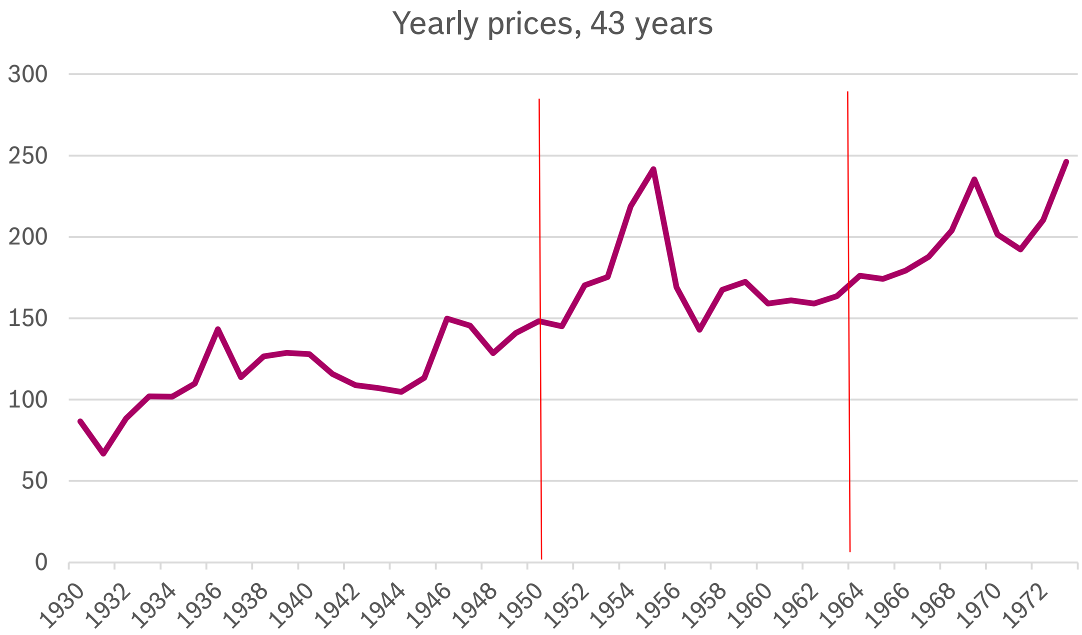
- Bigger picture: Yearly prices, 14 years
  - ***No decreased trend***, stable prices
  - R^2 = 0,007
  - Low error variance



Source: "Micro-Economics", Makridakis, Wheelwright & Hyndman, 1998).

Timeserie Preparation & Analysis
Volume of historical data (3)

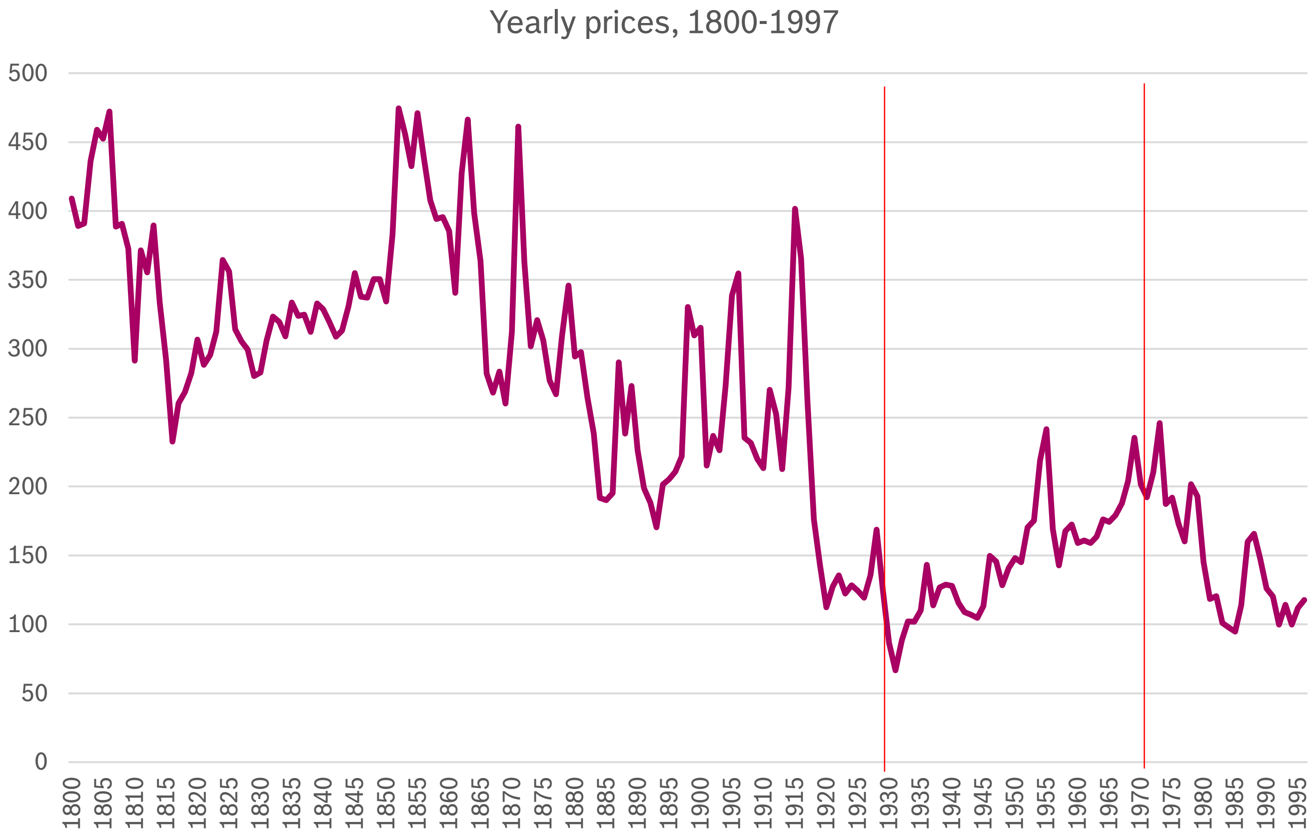
- Bigger picture: Yearly prices, 43 years
  - Now we observe ***increased trend***
  - $R^2 = 0,743$
  - Low error variance



Source: "Micro-Economics", Makridakis, Wheelwright & Hyndman, 1998).

Timeserie Preparation & Analysis
Volume of historical data (3)

- The biggest picture: Yearly prices, almost 200 years
  - ***Decreased trend***
  - Cyclical behavior, with different duration
  - R^2 = 0,618
  - Low error variance

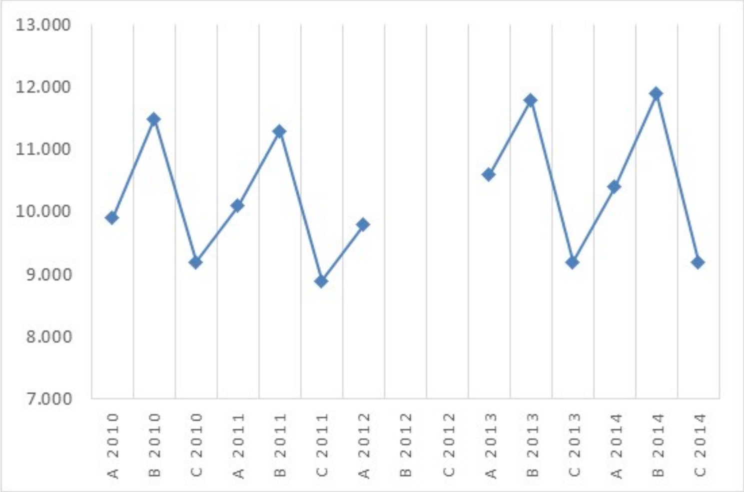


Source: "Micro-Economics", Makridakis, Wheelwright & Hyndman, 1998).

MANAGING FAULT / MISSING VALUES

Timeserie Preparation & Analysis
Managing Fault/Missing values (1)

- Manage the missing values before applying a forecast method:
  - Try to find the missing values from ***other possible sources***.
  - Direct define the missing values if a safe ***judgmental estimation*** exists.
  - ***Statistically estimate*** the missing values.



Timeserie Preparation & Analysis
Managing Fault/Missing values (2)

If a timeserie is characterized by ***stagnation*** and has ***no seasonal behavior***:

- the missing value can be defined as the ***average*** of the previous and the following values.

If a timeserie has a clear ***seasonal behavior*** but ***no trend***:

- the missing value can be defined as the average value of the respective periods.
  - For example, if the data consists of monthly observations and a missing value in on April of a given year, then the missing value can be defined as the ***average of all other April*** (average of all data points for April for all the years).

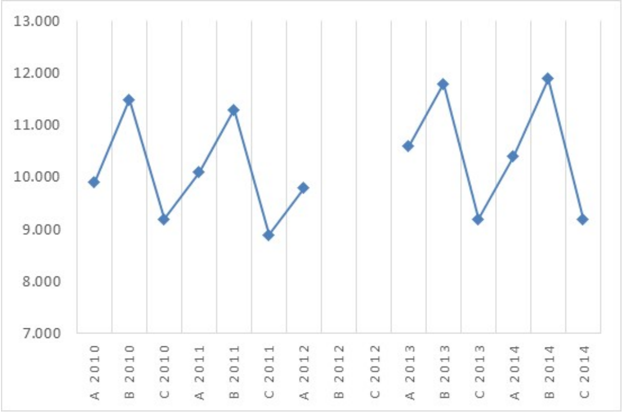
if a timeserie has a clear ***seasonal behavior*** and a ***trend***:

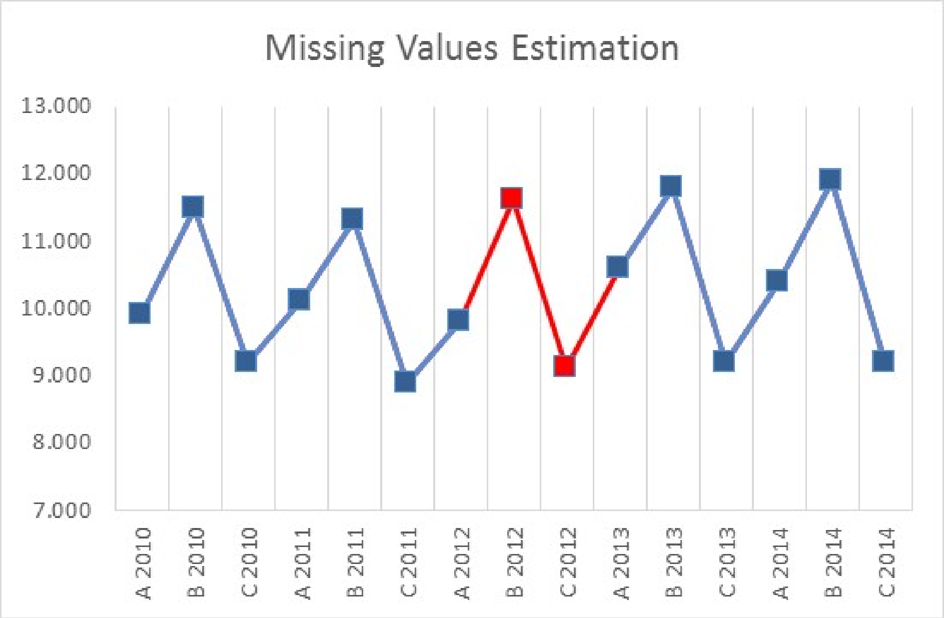
  - only the previous and following April values must be taken into calculation.

Timeserie Preparation & Analysis
Managing Fault/Missing values (3)

Example: A timeserie with 15 data points, and 2 missing values

	Period	Value
1	A 2010	9.900
2	B 2010	11.500
3	C 2010	9.200
4	A 2011	10.100
5	B 2011	11.300
6	C 2011	8.900
7	A 2012	9.800
8	B 2012	
9	C 2012	
10	A 2013	10.600
11	B 2013	11.800
12	C 2013	9.200
13	A 2014	10.400
14	B 2014	11.900
15	C 2014	9.200





The timeserie has no trend, but has seasonality. Thus:

- Y(8) = 1/4 * (Y(2) + Y(5) + Y(11) + Y(14)) = 11.625
- Y(9) = 1/4 * (Y(3) + Y(6) + Y(12) + Y(15)) = 9.125

BOSCH

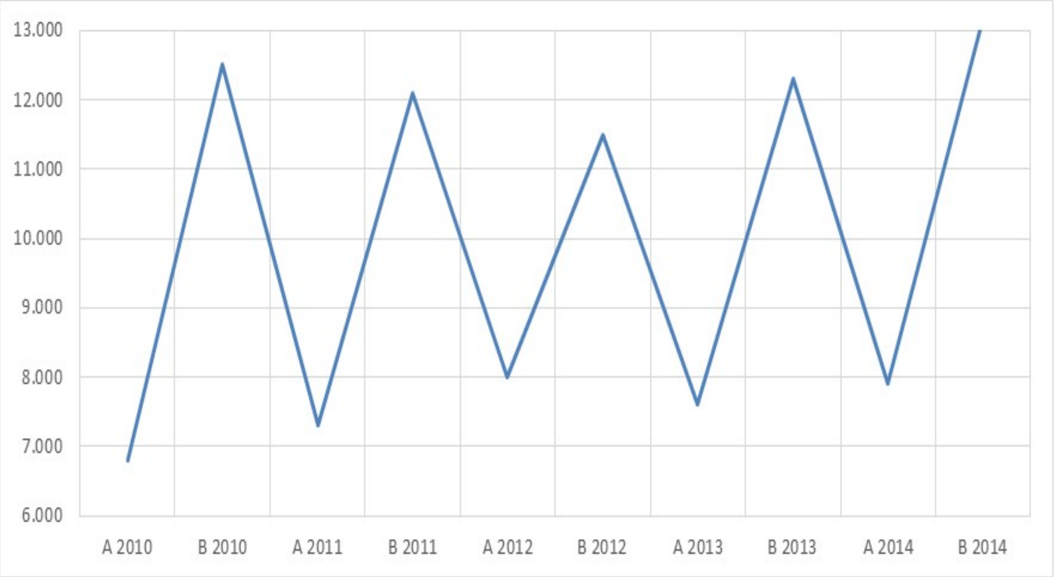
NON-WORKING DAYS ADJUSTMENTS

Timeserie Preparation & Analysis
Non-Working Days Adjustments (1)

An adjustment could be made, by taking into account ***the number of working days*** in each period.

- Define the number of working days in each period.
- Define the country and find the holidays.
- Adjust the data point values, by using:

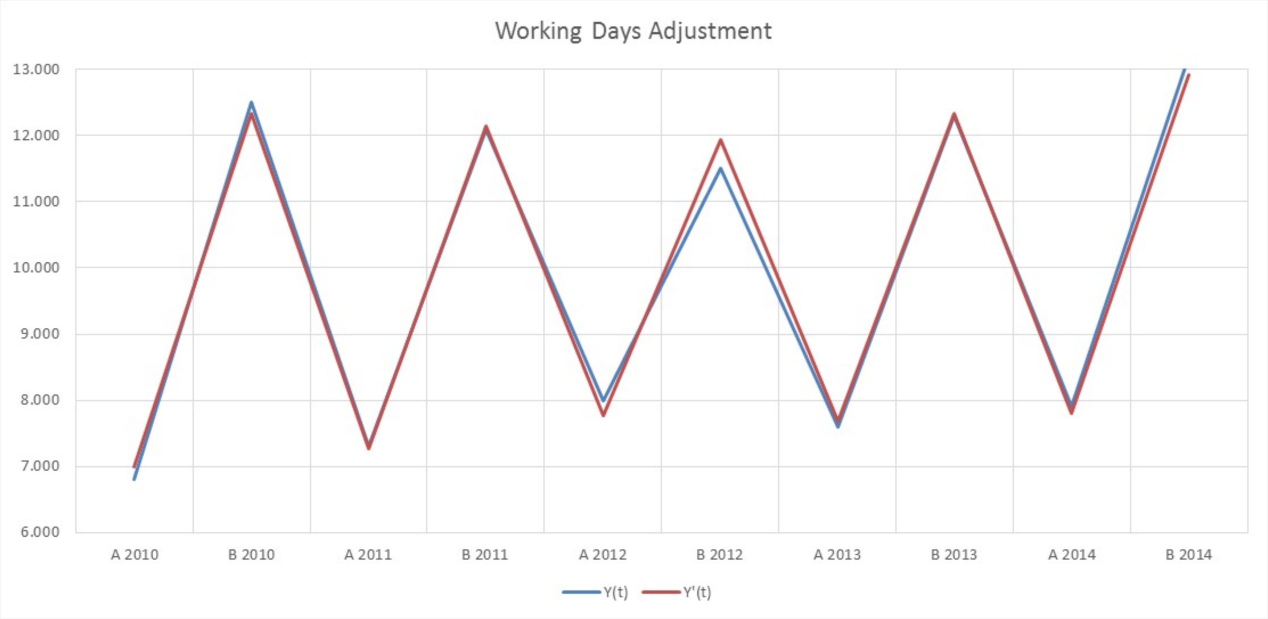
$$Y_t' = Y_t \times \frac{number\ of\ average\ trading\ days\ in\ all\ periods}{number\ of\ trading\ days\ in\ current\ period\ (t)}$$



Timeserie Preparation & Analysis
Non-Working Days Adjustments (2)

Example:

	Y(t)	Y'(t)	WD
A 2010	6.800	6.996	118
B 2010	12.500	12.337	123
A 2011	7.300	7.264	122
B 2011	12.100	12.140	121
A 2012	8.000	7.770	125
B 2012	11.500	11.932	117
A 2013	7.600	7.689	120
B 2013	12.300	12.341	121
A 2014	7.900	7.797	123
B 2014	13.200	12.923	124
Average Working Days			121,4



$$Y_1' = Y_1 \times \frac{Average}{WD1} = 6800 \times \frac{121{,}4}{118} = 6996$$

SPECIAL EVENTS & ACTIONS

Timeserie Preparation & Analysis
Special Events & Actions - SEA (1)

Level difference between current data of a timeserie and an initial level (or the level of the average value), could be the impact from a special event or action.

- Abnormal values (outliers)
- Level change (level shifts)

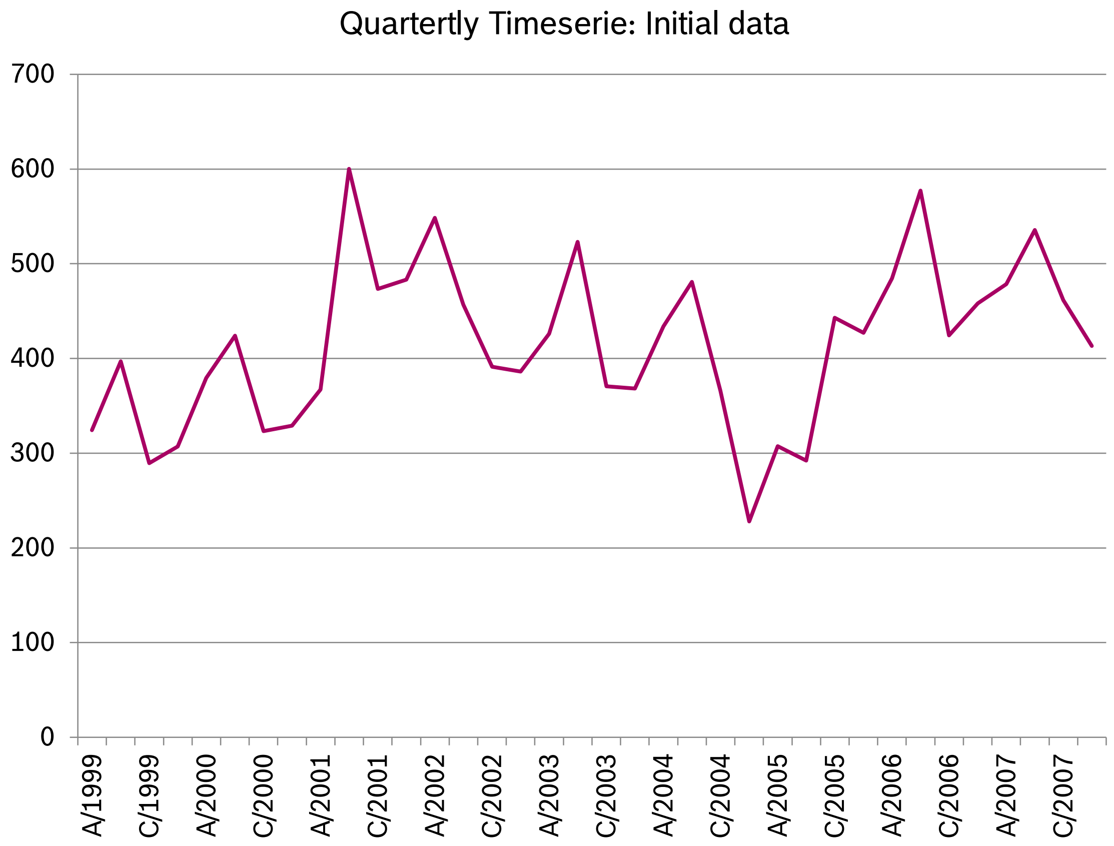
- The systematic recording and classification of impacts of various types of SEA could in turn lead to ***more accurate forecasts***, through appropriate ***critical interventions*** on statistical forecasts.
- These interventions will be based on ratios of past and upcoming events and actions.

BOSCH

Timeserie Preparation & Analysis
Special Events & Actions - SEA (2)

Example:

- A quarterly timeserie with 36 data points, 9 years.
- Do we see impact of SEA?
- How we can locate them?
- Can we use some detection methods?

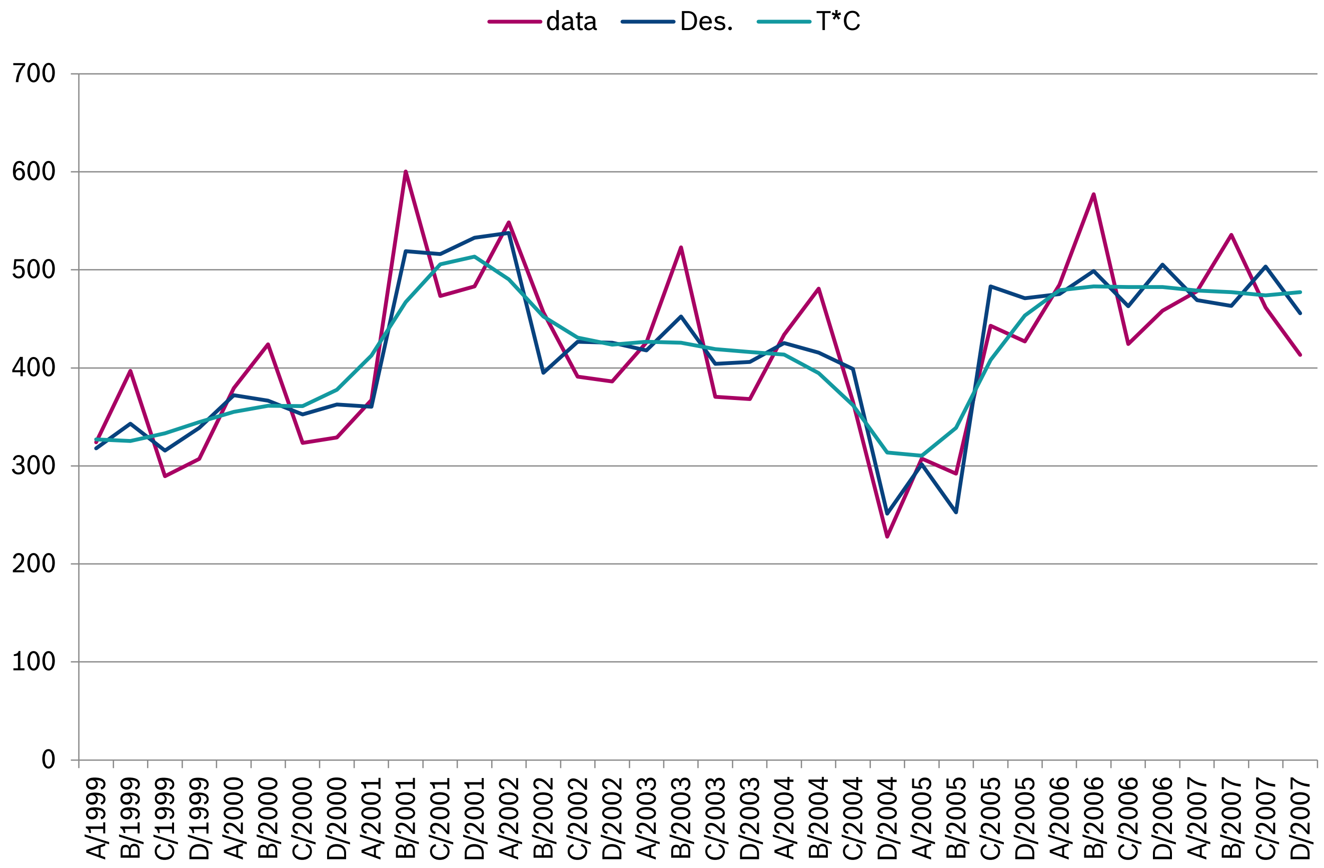


BOSCH

Timeserie Preparation & Analysis
Special Events & Actions - SEA (3)

Detection methods:

- 3 common methods.
- They must be applied into:
  - a **deseasonalized timeserie**, a timeserie with no seasonality component (TxCxR), or
  - a deseasonalized timeserie without Randomness (TxC).



Timeserie Preparation & Analysis
Special Events & Actions – Detection Method A (1)

- Estimate the ***Ratio1*** and ***Ratio2*** indexes for every period *t* of the timeserie.

$$Ratio1_t = \frac{D_t}{T * C_t} \qquad Ratio2_t = \frac{D_t}{F_t}$$

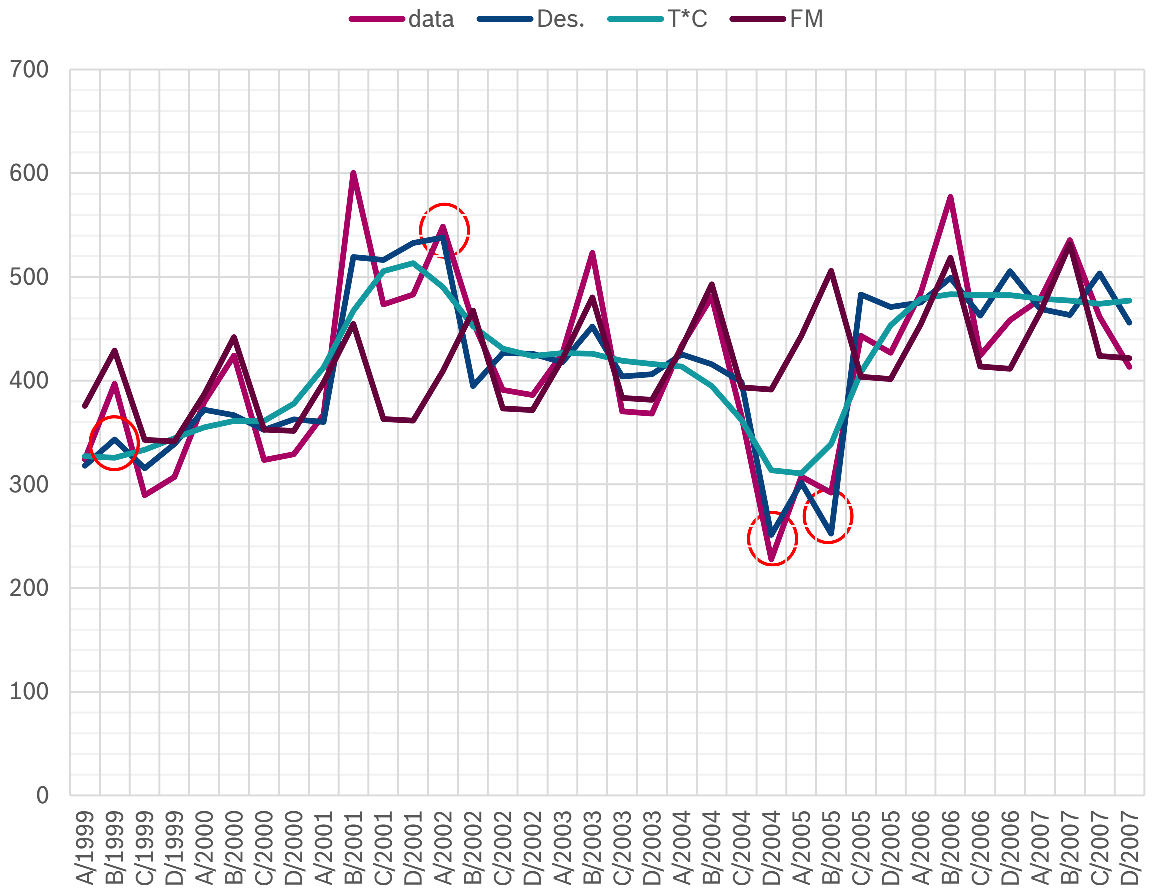
- Given that $t_a \leq 10$ and $t_\beta \leq 25$, the value of the period *t* of the timeserie is an unusual value when:

$$\left\{ \begin{matrix} Ratio1_t \geq 1.1 - \frac{t_a}{100} \\ or \\ Ratio1_t \leq 0.9 + \frac{t_a}{100} \end{matrix} \right\} and \left\{ \begin{matrix} Ratio2_t \geq 1.25 - \frac{t_\beta}{100} \\ or \\ Ratio2_t \leq 0.75 + \frac{t_\beta}{100} \end{matrix} \right\}$$

➢ ***Threshold A & Threshold B (t_a & t_b):*** Sensitive parameters for detecting special events and actions

Timeserie Preparation & Analysis
Special Events & Actions – Detection Method A (2)

month	data	Deseas.	T*C	FM	Ratio1		Ratio2		is SEA?
A/1999	324,25	317,96	327,07	375,60	0,972	NO	0,847	NO	NO
B/1999	396,82	343,12	325,56	429,15	1,054	YES	0,800	YES	YES
C/1999	289,42	315,61	333,40	342,81	0,947	YES	0,921	NO	NO
D/1999	307,17	338,78	344,60	341,46	0,983	NO	0,992	NO	NO
A/2000	379,36	372,00	355,02	386,87	1,048	NO	0,962	NO	NO
B/2000	424,10	366,71	361,21	441,93	1,015	NO	0,830	NO	NO
C/2000	323,36	352,62	361,01	352,94	0,977	NO	0,999	NO	NO
D/2000	328,91	362,76	377,75	351,48	0,960	NO	1,032	NO	NO
A/2001	367,38	360,25	412,58	398,14	0,873	YES	0,905	NO	NO
B/2001	600,29	519,05	467,31	454,71	1,111	YES	1,142	NO	NO
C/2001	473,40	516,24	505,63	363,08	1,021	NO	1,422	YES	NO
D/2001	483,13	532,85	513,41	361,49	1,038	NO	1,474	YES	NO
A/2002	548,50	537,86	490,21	409,40	1,097	YES	1,314	YES	YES
B/2002	456,72	394,91	452,48	467,48	0,873	YES	0,845	NO	NO
C/2002	391,18	426,58	430,76	373,21	0,990	NO	1,143	NO	NO
D/2002	386,09	425,82	423,71	371,51	1,005	NO	1,146	NO	NO
A/2003	426,02	417,75	426,69	420,67	0,979	NO	0,993	NO	NO
B/2003	523,13	452,33	425,84	480,26	1,062	YES	0,942	NO	NO
C/2003	370,50	404,03	419,14	383,34	0,964	NO	1,054	NO	NO
D/2003	368,30	406,20	416,16	381,53	0,976	NO	1,065	NO	NO
A/2004	433,78	425,36	413,64	431,94	1,028	NO	0,985	NO	NO
B/2004	480,79	415,72	394,78	493,04	1,053	YES	0,843	NO	NO
C/2004	365,72	398,82	361,93	393,47	1,102	YES	1,014	NO	NO
D/2004	227,87	251,32	313,65	391,55	0,801	YES	0,642	YES	YES
A/2005	307,47	301,50	310,47	443,20	0,971	NO	0,680	YES	NO
B/2005	292,07	252,54	338,81	505,82	0,745	YES	0,499	YES	YES
C/2005	443,09	483,19	408,16	403,60	1,184	YES	1,197	NO	NO
D/2005	427,04	470,99	453,49	401,56	1,039	NO	1,173	NO	NO
A/2006	484,69	475,28	479,09	454,47	0,992	NO	1,046	NO	NO
B/2006	577,09	498,99	483,29	518,60	1,032	NO	0,962	NO	NO
C/2006	424,36	462,76	482,41	413,73	0,959	NO	1,119	NO	NO
D/2006	458,35	505,52	482,49	411,58	1,048	NO	1,228	YES	NO
A/2007	478,35	469,07	478,96	465,74	0,979	NO	1,007	NO	NO
B/2007	535,68	463,19	477,28	531,37	0,970	NO	0,872	NO	NO
C/2007	461,52	503,29	474,08	423,87	1,062	YES	1,187	NO	NO
D/2007	413,23	455,76	477,30	421,60	0,955	NO	1,081	NO	NO



Internal | CR/AEU2 - Patelis | 2018-03-14

Timeserie Preparation & Analysis
Special Events & Actions – Detection Method B (1)

- Estimate the ***mean values*** of the deseasonalized timeserie (D) and the forecast model (F):

$$\bar{D} = \frac{\sum_{i=1}^{n} D_i}{n}, \bar{F} = \frac{\sum_{i=1}^{n} F_i}{n}$$

- Estimate the ***standard deviation*** of the forecast model values from the average value:

$$StD_F = \sqrt{\frac{\sum_{i=1}^{n} (F_i * \bar{F})^2}{n}}$$

- Given that $t_a \leq 3$, the value of the period *t* of the timeserie is an unusual value when:

$$\left\{ \begin{matrix} D_t \geq \bar{D} + (3 - t_a) * StD_F \\ or \\ D_t \leq \bar{D} - (3 - t_a) * StD_F \end{matrix} \right\}$$

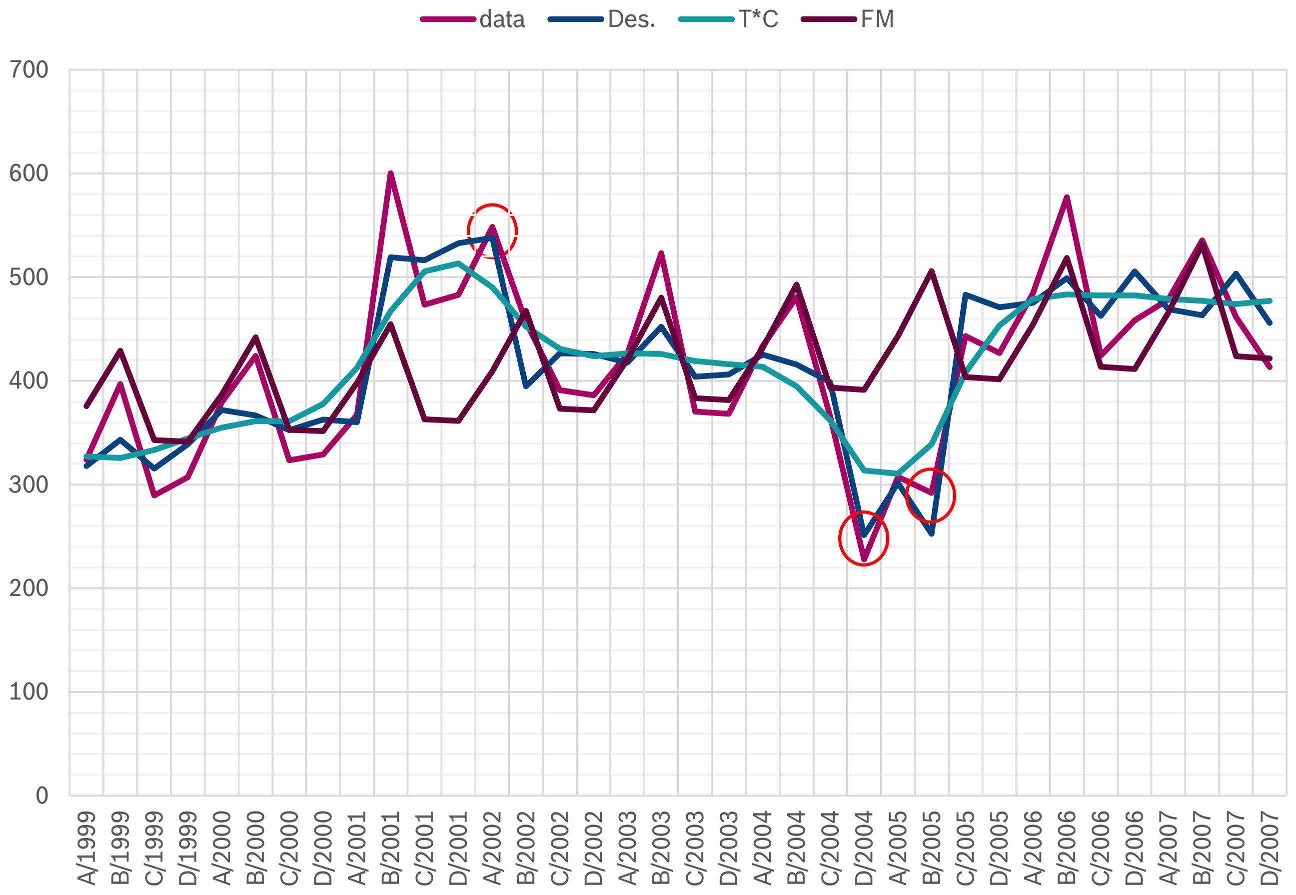
Internal | CR/AEU2 - Patelis | 2018-03-14

Timeserie Preparation & Analysis

Special Events & Actions – Detection Method B (2)

month	data	Deseas.	T*C	FM	Upper Limit	Lower Limit	is SEA?
A/1999	324,25	317,96	327,07	375,60	534,68	298,48	NO
B/1999	396,82	343,12	325,56	429,15	534,68	298,48	NO
C/1999	289,42	315,61	333,40	342,81	534,68	298,48	NO
D/1999	307,17	338,78	344,60	341,46	534,68	298,48	NO
A/2000	379,36	372,00	355,02	386,87	534,68	298,48	NO
B/2000	424,10	366,71	361,21	441,93	534,68	298,48	NO
C/2000	323,36	352,62	361,01	352,94	534,68	298,48	NO
D/2000	328,91	362,76	377,75	351,48	534,68	298,48	NO
A/2001	367,38	360,25	412,58	398,14	534,68	298,48	NO
B/2001	600,29	519,05	467,31	454,71	534,68	298,48	NO
C/2001	473,40	516,24	505,63	363,08	534,68	298,48	NO
D/2001	483,13	532,85	513,41	361,49	534,68	298,48	NO
A/2002	548,50	537,86	490,21	409,40	534,68	298,48	YES
B/2002	456,72	394,91	452,48	467,48	534,68	298,48	NO
C/2002	391,18	426,58	430,76	373,21	534,68	298,48	NO
D/2002	386,09	425,82	423,71	371,51	534,68	298,48	NO
A/2003	426,02	417,75	426,69	420,67	534,68	298,48	NO
B/2003	523,13	452,33	425,84	480,26	534,68	298,48	NO
C/2003	370,50	404,03	419,14	383,34	534,68	298,48	NO
D/2003	368,30	406,20	416,16	381,53	534,68	298,48	NO
A/2004	433,78	425,36	413,64	431,94	534,68	298,48	NO
B/2004	480,79	415,72	394,78	493,04	534,68	298,48	NO
C/2004	365,72	398,82	361,93	393,47	534,68	298,48	NO
D/2004	227,87	251,32	313,65	391,55	534,68	298,48	YES
A/2005	307,47	301,50	310,47	443,20	534,68	298,48	NO
B/2005	292,07	252,54	338,81	505,82	534,68	298,48	YES
C/2005	443,09	483,19	408,16	403,60	534,68	298,48	NO
D/2005	427,04	470,99	453,49	401,56	534,68	298,48	NO
A/2006	484,69	475,28	479,09	454,47	534,68	298,48	NO
B/2006	577,09	498,99	483,29	518,60	534,68	298,48	NO
C/2006	424,36	462,76	482,41	413,73	534,68	298,48	NO
D/2006	458,35	505,52	482,49	411,58	534,68	298,48	NO
A/2007	478,35	469,07	478,96	465,74	534,68	298,48	NO
B/2007	535,68	463,19	477,28	531,37	534,68	298,48	NO
C/2007	461,52	503,29	474,08	423,87	534,68	298,48	NO
D/2007	413,23	455,76	477,30	421,60	534,68	298,48	NO

$$StD_F = 49,93$$



Timeserie Preparation & Analysis
Special Events & Actions – Detection Method C (1)

- Estimate the ***Ratio*** index for every period *t* of the timeserie:

$$Ratio_t = \frac{CMA(7)_t^D}{CMA(5)_t^D}$$

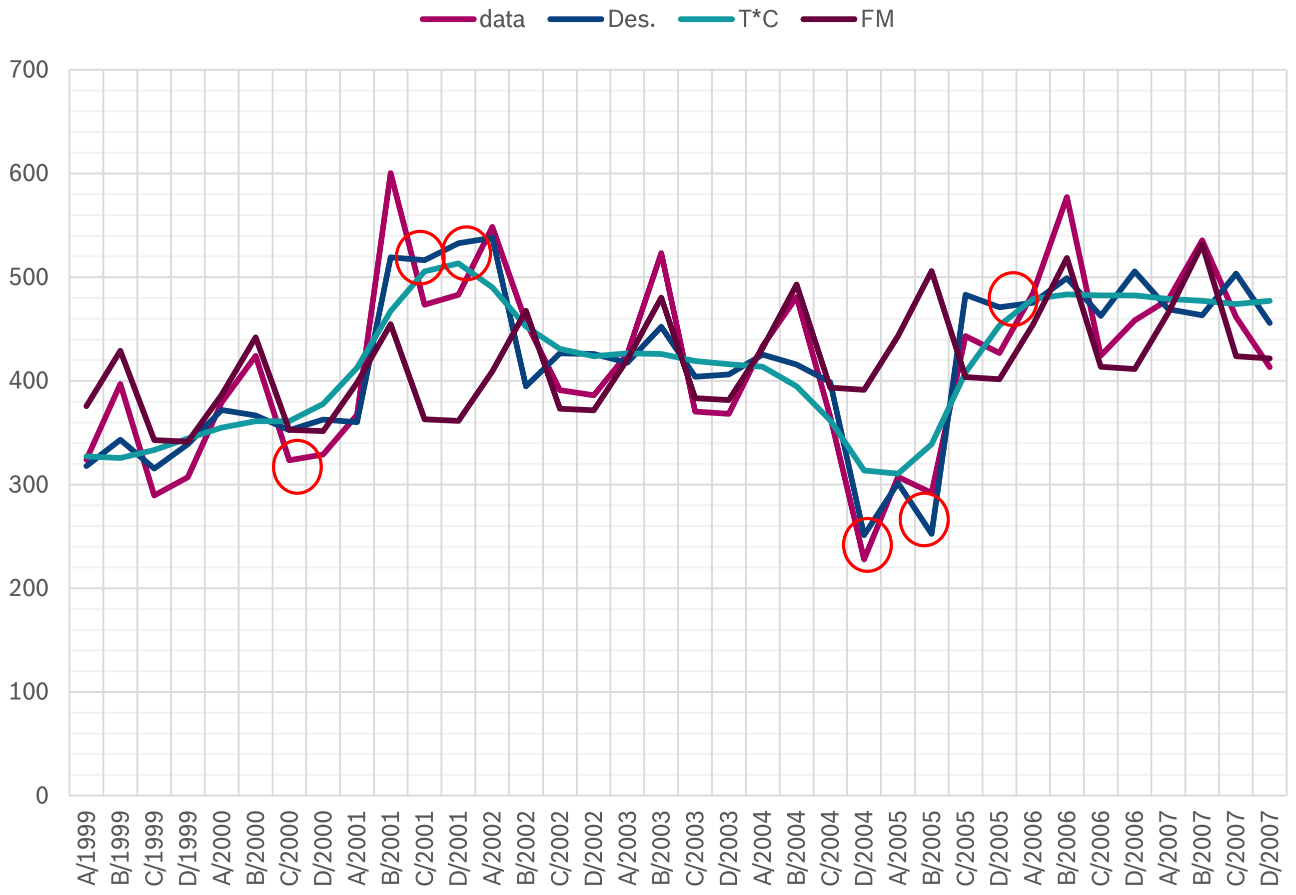
- Given that $t_a \leq 5$, the value of the period *t* of the timeserie is an unusual value when:

$$\left\{\begin{matrix} Ratio_t \ \geq 1.05 - \dfrac{t_a}{100} \\ or \\ Ratio_t \ \leq 0.95 + \dfrac{t_a}{100} \end{matrix}\right\}$$

Timeserie Preparation & Analysis

Special Events & Actions – Detection Method C (2)

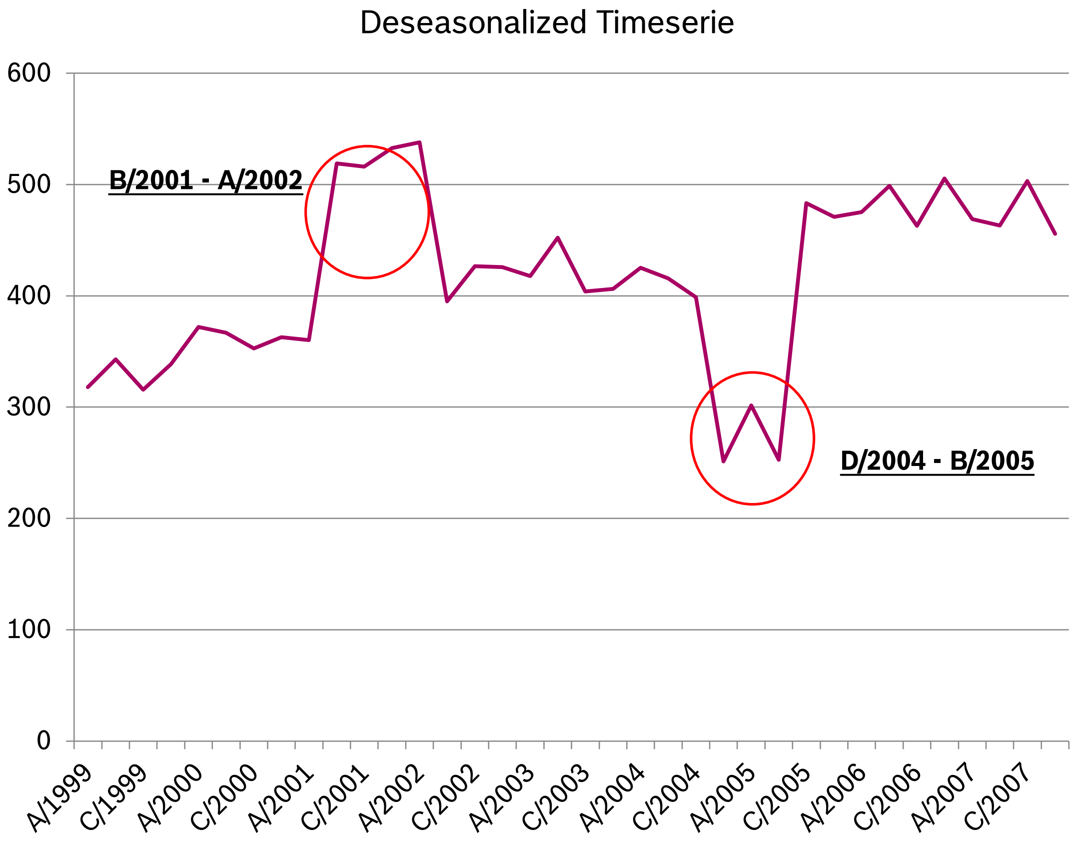
month	data	Deseas.	MA(5)	MA(7)	Ratio	is SEA?
A/1999	324,25	317,96				NO
B/1999	396,82	343,12				NO
C/1999	289,42	315,61	337,49			NO
D/1999	307,17	338,78	347,24	343,83	0,990	NO
A/2000	379,36	372,00	349,14	350,23	1,003	NO
B/2000	424,10	366,71	358,57	352,68	0,984	NO
C/2000	323,36	352,62	362,87	381,74	1,052	YES
D/2000	328,91	362,76	392,28	407,09	1,038	NO
A/2001	367,38	360,25	422,19	430,07	1,019	NO
B/2001	600,29	519,05	458,23	454,52	0,992	NO
C/2001	473,40	516,24	493,25	460,56	0,934	YES
D/2001	483,13	532,85	500,18	469,68	0,939	YES
A/2002	548,50	537,86	481,69	479,05	0,995	NO
B/2002	456,72	394,91	463,61	464,57	1,002	NO
C/2002	391,18	426,58	440,59	455,44	1,034	NO
D/2002	386,09	425,82	423,48	437,04	1,032	NO
A/2003	426,02	417,75	425,30	418,23	0,983	NO
B/2003	523,13	452,33	421,23	422,58	1,003	NO
C/2003	370,50	404,03	421,14	421,03	1,000	NO
D/2003	368,30	406,20	420,73	417,17	0,992	NO
A/2004	433,78	425,36	410,03	393,40	0,959	NO
B/2004	480,79	415,72	379,49	371,85	0,980	NO
C/2004	365,72	398,82	358,55	350,21	0,977	NO
D/2004	227,87	251,32	323,98	361,21	1,115	YES
A/2005	307,47	301,50	337,48	367,73	1,090	YES
B/2005	292,07	252,54	351,91	376,24	1,069	YES
C/2005	443,09	483,19	396,70	390,55	0,984	NO
D/2005	427,04	470,99	436,20	420,75	0,965	NO
A/2006	484,69	475,28	478,24	449,90	0,941	YES
B/2006	577,09	498,99	482,71	480,83	0,996	NO
C/2006	424,36	462,76	482,33	477,97	0,991	NO
D/2006	458,35	505,52	479,91	482,59	1,006	NO
A/2007	478,35	469,07	480,77	479,80	0,998	NO
B/2007	535,68	463,19	479,36			NO
C/2007	461,52	503,29				NO
D/2007	413,23	455,76				NO



Internal | CR/AEU2 - Patelis | 2018-03-14

Timeserie Preparation & Analysis
Special Events & Actions – Identify (1)

- Each of the 3 detection methods:
  - examine every period value independently and separately, answering the question if the given period value is an unusual value or not.
- Potential SEA:
  - Every unusual value recognised by a detection method, ***is not necessarily a SEA***.
  - We need to verify the findings by combining:
    - a ***graphical representation*** of the data, and
    - possible ***available information*** for internal or external company sources.



Timeserie Preparation & Analysis
Special Events & Actions – Estimate smoothed TS & Impact (1)

- Two methods for estimating a smoothed timeserie (thus, with no impact of SEA):
  - For constant level data having ***no trend***, the smoothed value of a period t (which has an unusual value due to a SEA) is equal to the ***value of the previous period*** (just before the SEA).

$$D'_t = D_{t_o}$$

  - For data with ***strong trend***, the smoothed value of a period t (which has an unusual value due to a SEA) is calculated by a ***linear interpolation*** based on ***the preceding and the following period*** of the SEA.

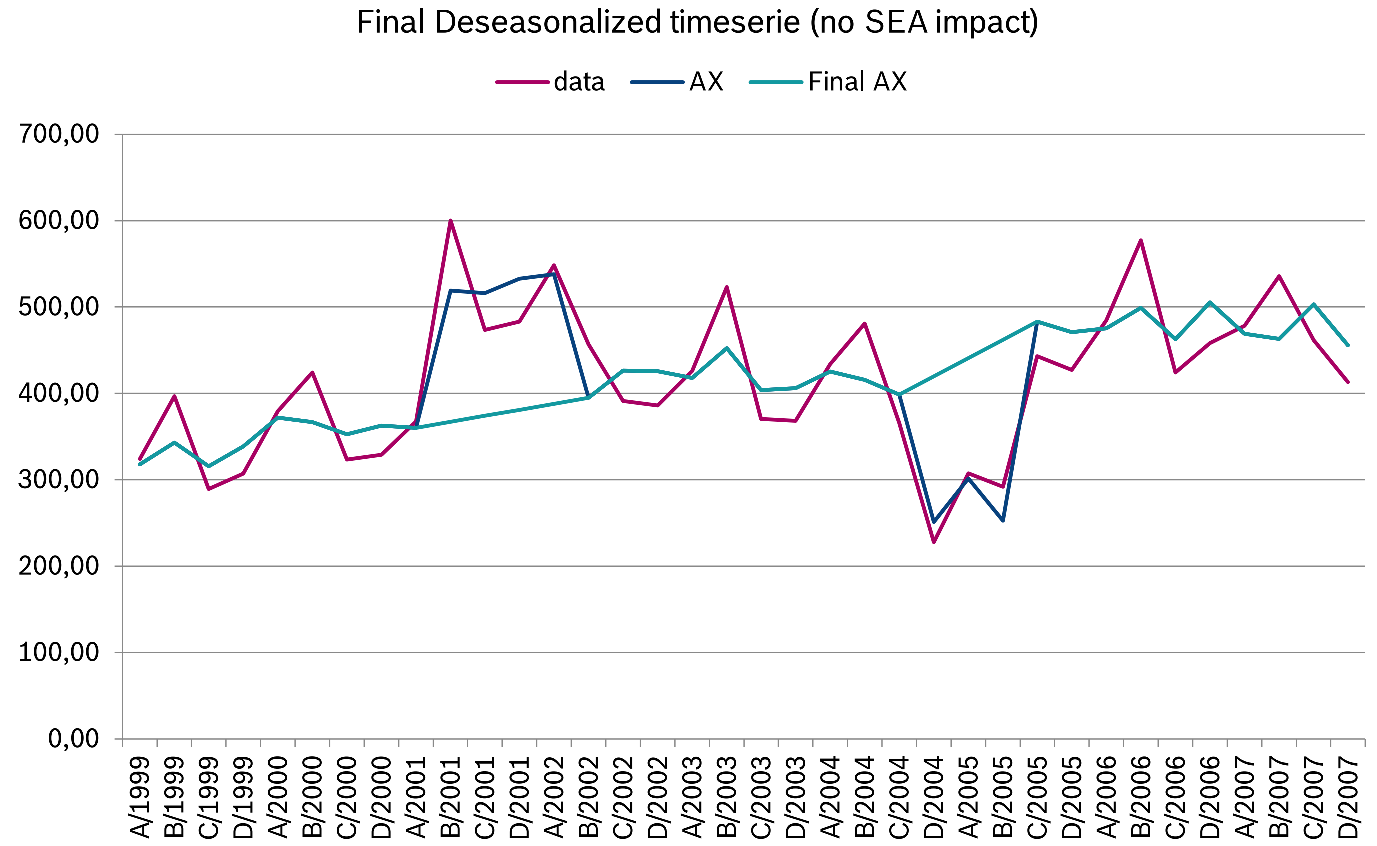
$$D'_t = (t - t_o) * \frac{D_{t_o + n+1} - D_{t_o}}{n+1} + D_{t_o}$$

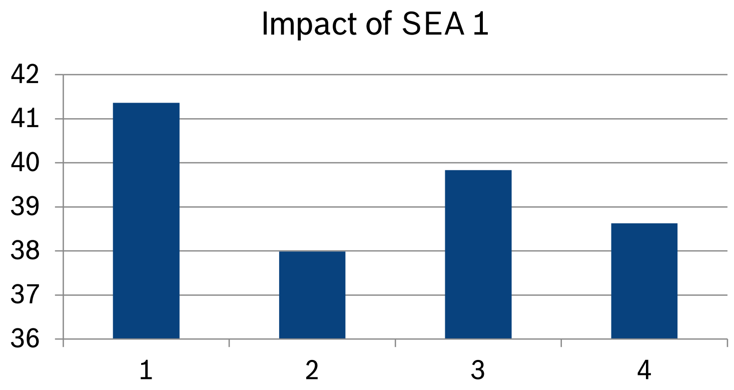
- The impact of an SEA for each period is calculated as the ***% ratio*** between the difference of the initial and the smoothed value and the smoothed value

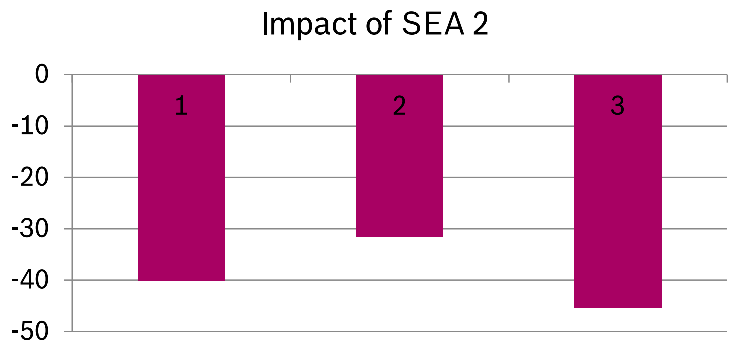
$$Impact_t = \frac{D_t - D'_t}{D'_t} * 100\ (\%)$$

Timeserie Preparation & Analysis

Special Events & Actions – Estimate smoothed TS & Impact (2)



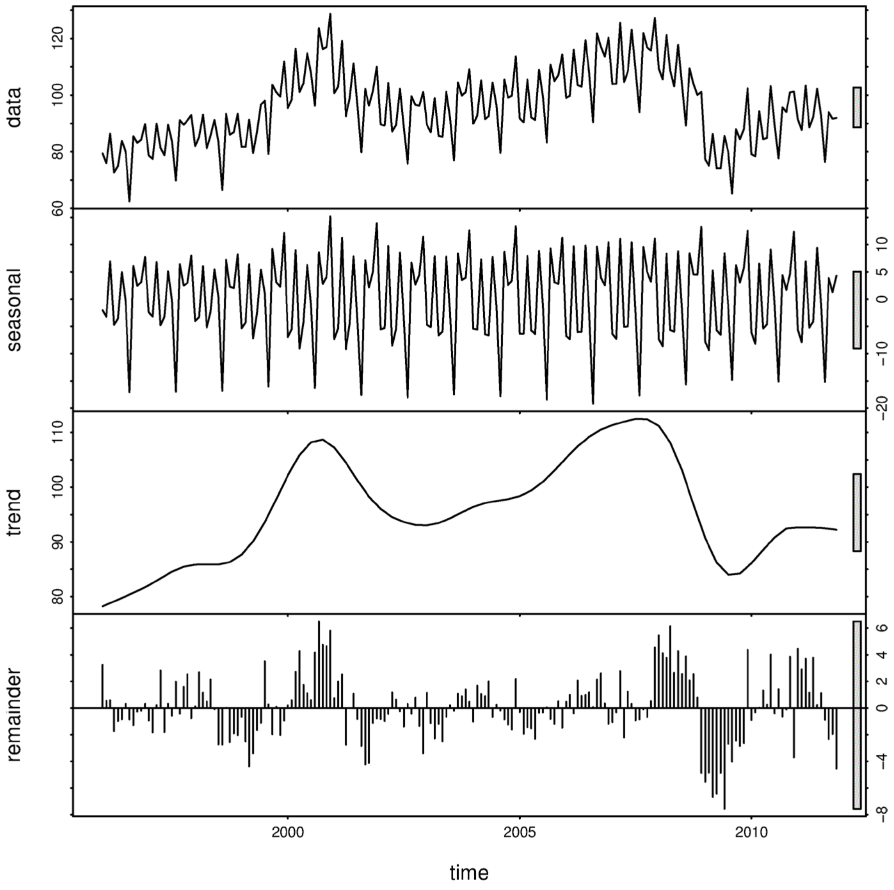




ESTIMATE COMPONENTS

Timeserie Preparation & Analysis
Estimate Components (1)

- Deconstructs a timeserie (Y) into the basic four parts (components):
  - ***Trend*** (T) component.
  - ***Cyclical*** (C) component.
  - ***Seasonal*** (S) component.
  - ***Random*** (R) or Irregular component.



BOSCH

Source: Forecasting – Principles and Practice

Timeserie Preparation & Analysis
Estimate Components (2)

- We must:
  - isolate the seasonal component from a timeserie, and
  - create the deseasonalized version of the timeserie

- We must always apply forecasting methods on the deseasonalized timeseries.
  - This ***increases the reliability*** of the forecasts.

- Commonly used:
  - ***Classical Decomposition method***, based on Moving Averages
  - Example will follow…

5. FORECASTING TYPES

Forecasting Methods
Forecasting

- The forecasting processes can be divided into the following different types:
  - ***quantitative*** (statistical) methods
  - ***judgemental*** methods

- We have to keep in mind, that these methods are usually complementary.
  - Statistical methods are ***rigorous but consistent***, and they can analyze large datasets of information very quickly.
  - People (judgmental methods) adapt more easily and can ***take into consideration events outside the timeserie pattern***. But they are inconsistent and show increased bias.

Forecasting Methods

Forecasting: Statistical methods (1)

- **Statistical (Quantitative) methods**:
  - ***Time-series methods***: seek to identify historical patterns (using time as a reference) and then forecast using a time-based extrapolation of these patterns.
    - Naive, Smoothing, Decomposition, Auto-regressive moving average, Theta, ADIDA , etc.
  - ***Explanatory methods***: seek to identify the relationships that led to (caused) observed outcomes in the past and then forecast by applying those relationships in the future.
    - Vector Auto-regressive, Regression, Econometrics, etc.
  - ***Monitoring methods***: not yet in widespread use, seek to identify changes in patterns and relationships.
    - They are used primarily to indicate when extrapolation of past patterns or relationships in not appropriate.

Forecasting Methods
Forecasting: Statistical methods (2)

- ***Advantages***:
  - Can be applied directly.
  - ***Relatively accurate***, taken into account the Confidence Intervals.
  - Does not require technical and statistical knowledge in order to estimate the forecasts (when used as a "black box" from Managers).
  - They require ***significant little time*** and few computational resources.
- ***Disadvantages***:
  - They assume that the behaviour of the given timeserie ***will continue*** in the future, which is not always true.
  - They does not take into account specific special events and actions that may take place (like athletic events, holidays, etc.)
  - Some statistical methods require a large number of historical observations, in order to estimate forecasts.

Forecasting Methods
Forecasting: Judgmental methods (1)

- **Judgmental (Qualitative) methods**:
  - Most commonly used in business and government organizations.
  - The data for these methods are product of ***intuition***, ***judgment*** and ***accumulated knowledge***.
  - Made as ***individual judgment*** or by ***committee agreements*** or ***decisions***.
    - Individual judgment, Decision rules, Sales Force Estimates, Delphi, Role playing, Juries of executive opinion.

Forecasting Methods
Forecasting: Judgmental methods (2)

- ***Advantages***:
  - Need less data.
  - May take into account SEA.
  - Ability to compensate inadequacies and gaps in historical data.
  - Appropriate when ethical issues are raised which outweigh the financial costs or technology factors.
  - Allow the processing of forecasting, in cases when the director of a company wishes to have control over the product whose demand will be forecasted.
  - Can produce ***more acceptable forecasts*** (Notice: *more acceptable, not necessarily more accurate*).

- ***Disadvantages***:
  - The biggest disadvantage is ***bias*** (prejudice), thus the innate tendency of people to be optimistic or pessimistic.
  - Other disadvantages:
    - Inconsistency, Conservatism, Persistence on recent events, Availability, Incorrect correlations, False estimation of regression, Yield of success and failure, Devaluation of uncertainty, Selective perception.

6. FORECASTING ERRORS & ACCURACY

"Some of the biggest, most sophisticated organizations in the world – which spend millions on forecasting – do not scientifically test the accuracy of the forecasting…"

Dan Gardner

Forecasting Errors & Accuracy
Errors & Accuracy

Forecasting Error:

- The difference between the actual value and the forecasted value of a timeserie.
- In order to evaluate forecasts for accuracy, we have to be able to understand exactly what the forecast says. ***This is more difficult that it seems.***
- What is probability to happen?
  - A forecaster says that ***it is likely*** that something will happen.
  - People were asked for their thoughts on the exact likelihood it would happen:
    - Answers ranged from 20% to 80% probability, illustrating that sometimes people have very different ideas about what something means.

Forecasting Errors & Accuracy
Errors & Accuracy

- ***Forecast: 70% chance of rain***
  - If it doesn't rain, some people may think the forecast was wrong. True?
    - The only real way to judge the accuracy of the forecast would be rerun the weather 100 times and see how often it rains.
    - But since we can't do that, all we can really say is that the forecast was not disproved.
  - We can't rerun history so we can't judge an isolated forecast.
  - What we can do, however, is ***look at a large number of forecasts together*** - look at the track record of a meteorologist.
    - Of all the times he said 70% chance of rain, did it rain 70% of the time?
    - This is calibration, and by calibrating the meteorologist's forecasts (plotting percent correct by number of forecasts), we can identify whether he is under confident or overconfident.
  - But in order to make these assessments, we need to have lots of data. It doesn't work very well with rare events.

Forecasting Errors & Accuracy
Error Types (1)

- ***Error*** $(\boldsymbol{e})$: The difference between the actual value and the forecasted value, for a given time.
  - $e_i = Y_i - F_i$
- ***Mean Error (ME)***: The average error between the actual values and the forecasted values, for a given period.
  - $ME = \frac{\sum_{i=1}^{n}(Y_i - F_i)}{n}$
- ***Mean Absolute Error (MAE)***: The average absolute error for a given period
  - $MAE = \frac{\sum_{i=1}^{n}|Y_i - F_i|}{n}$
  - Commonly used

Forecasting Errors & Accuracy
Error Types (2)

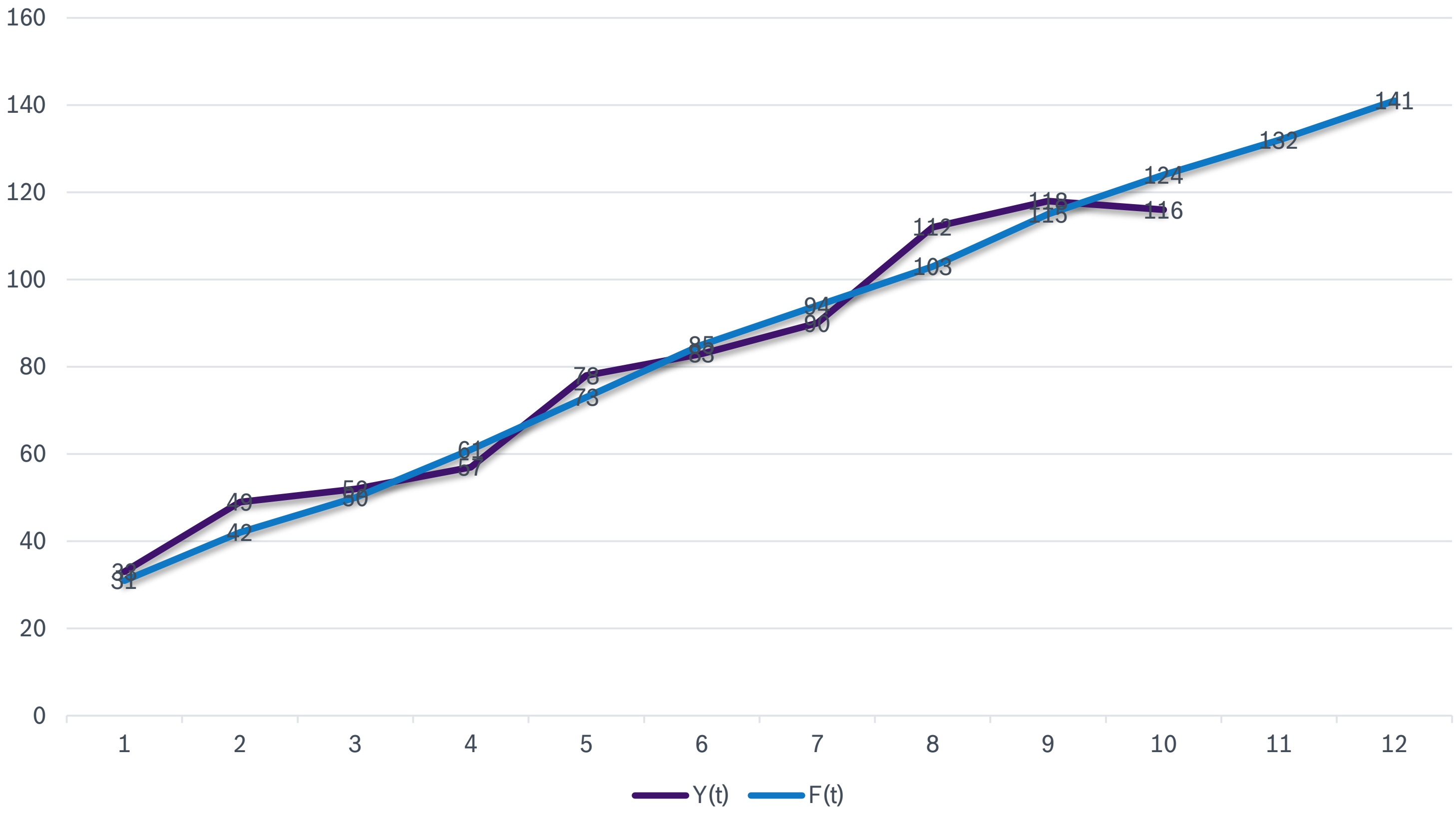
- ***Mean Squared Error (MSE)***: Average of all square errors
  - $MSE = \frac{\sum_{i=1}^{n}(Y_i - F_i)^2}{n}$
  - Compared to MAE, it amplifies large errors.
- ***Root Mean Squared Error (RMSE)***: Root of the average of all square errors
  - $RMSE = \sqrt{MSE} = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - F_i)^2}{n}}$
  - The use of RMSE is very common and it makes an excellent general purpose error metric for numerical forecasts.

Forecasting Errors & Accuracy
Error Types (3)

- ***Mean Absolute Percentage Error (MAPE)***: the average absolute % error for each time period divided by actuals
  - $MAPE = \frac{1}{n}\sum_{i=1}^{n}|\frac{Y_i - F_i}{Y_i}|100$
  - Most commonly used, but with zeros or near-zeros, it can give a distorted picture of error.
- ***Symmetric Mean Absolute Percentage Error (sMAPE)***: Alternative to MAPE
  - $sMAPE = \frac{1}{n}\sum_{i=1}^{n}|\frac{(Y_i - F_i)}{\frac{(Y_i + F_i)}{2}}|100 = \frac{1}{n}\sum_{i=1}^{n}|\frac{2(Y_i - F_i)}{(Y_i + F_i)}|100$
- ***Mean Absolute Scaled Error (MASE)***:
  - $MAsE = \frac{\frac{1}{n}\sum_{i=1}^{n}|Y_i - F_i|}{\frac{1}{n-1}\sum_{i=2}^{n}|Y_i - Y_{i-1}|}$
  - Tends to be the a standard when comparing forecasting accuracies.

Forecasting Errors & Accuracy
Error Types – Examples (1)

t	Y(t)	F(t)	e(t)
1	33	31	2
2	49	42	7
3	52	50	2
4	57	61	-4
5	78	73	5
6	83	85	-2
7	90	94	-4
8	112	103	9
9	118	115	3
10	116	124	-8
11		132	
12		141	



Forecasting Errors & Accuracy

Error Types – Examples (2)

t	Y(t)	F(t)	e(t)
1	33	31	2
2	49	42	7
3	52	50	2
4	57	61	-4
5	78	73	5
6	83	85	-2
7	90	94	-4
8	112	103	9
9	118	115	3
10	116	124	-8
11		132	
12		141	

$$e_1 = Y_1 - F_1 = 33 - 31 = 2$$

$$ME = \frac{1}{n} \times \sum_{i=1}^{n} e_1 = \frac{2 + 7 + \cdots - 8}{10} = \frac{10}{10} = 1$$

$$MAE = \frac{1}{n} \times \sum_{i=1}^{n} |e_1| = \frac{|2| + |7| + \cdots + |-8|}{10} = \frac{46}{10} = 4{,}6$$

$$MSE = \frac{1}{n} \times \sum_{i=1}^{n} (e_1)^2 = \frac{2^2 + 7^2 + \cdots + (-8)^2}{10} = \frac{272}{10} = 27{,}2$$

$$RMSE = \sqrt{MSE} = 5{,}22$$

BOSCH

Forecasting Errors & Accuracy
Error Types – Examples (3)

t	Y(t)	F(t)	e(t)
1	33	31	2
2	49	42	7
3	52	50	2
4	57	61	-4
5	78	73	5
6	83	85	-2
7	90	94	-4
8	112	103	9
9	118	115	3
10	116	124	-8
11		132	
12		141	

$$MAPE = \frac{1}{n} \times \sum_{i=1}^{n} \frac{|e_i|}{Y_i} 100\% = \frac{\frac{|2|}{33} + \frac{|7|}{49} + \cdots + \frac{|-8|}{116}}{10} 100\% = \frac{0{,}619}{10} 100\% = 6{,}19\%$$

$$sMAPE = \frac{1}{n} \times \sum_{i=1}^{n} \frac{2 \times |e_i|}{Y_i} 100\% = 2 \frac{\frac{|2|}{33+31} + \frac{|7|}{49+42} + \cdots + \frac{|-8|}{116+124}}{10} 100\% = 2 \frac{0{,}317}{10} 100\% = 6{,}33\%$$

$$MASE = \frac{\frac{1}{n}\sum_{i=1}^{n}|e_i|}{\frac{1}{n-1}\sum_{i=2}^{n}|Y_i - Y_{i-1}|} = \frac{\frac{|2| + |7| + \cdots + |-8|}{10}}{\frac{|49-33| + |52-49| + \cdots + |116-118|}{9}} = \frac{\frac{46}{10}}{\frac{87}{9}} = 0{,}476$$

Internal | CR/AEU2 - Patelis | 2018-03-14

Forecasting Errors & Accuracy

Error Types – Examples (4) – Relative Accuracy

t	Y(t)	F0(t)	F1(t)	F2(t)	e0(t)	e1(t)	e2(t)
1	33	33	31	32	0	2	1
2	49	33	42	43	16	7	6
3	52	49	50	51	3	2	1
4	57	52	61	58	5	-4	-1
5	78	57	73	75	21	5	3
6	83	78	85	87	5	-2	-4
7	90	83	94	95	7	-4	-5
8	112	90	103	105	22	9	7
9	118	112	115	114	6	3	4
10	116	118	124	126	-2	-8	-10
11		116	132	134			
12		116	141	143			

$$MAE_0 = \frac{1}{n} \times \sum_{i=1}^{n} |e_1| = \frac{|0| + |16| + \cdots + |-2|}{10} = \frac{87}{10} = 8{,}7$$

$$MAE_1 = \frac{1}{n} \times \sum_{i=1}^{n} |e_1| = \frac{|2| + |7| + \cdots + |-8|}{10} = \frac{46}{10} = 4{,}6$$

$$MAE_2 = \frac{1}{n} \times \sum_{i=1}^{n} |e_1| = \frac{|1| + |6| + \cdots + |-10|}{10} = \frac{42}{10} = 4{,}2$$

- ***Relative MAE (R_MAE)***:
  - Compare MAE with a benchmark MAE
  - If F0(t) is the benchmark method (Naïve), then:
    - o $R\_MAE_{_1} = \frac{MAE_1}{MAE_0} = \frac{4{,}6}{8{,}7} = 0{,}529$
    - o $R\_MAE_{_2} = \frac{MAE_2}{MAE_0} = \frac{4{,}2}{8{,}7} = 0{,}483$
  - And of course:
    - o $R\_MAE_{_0} = \frac{MAE_0}{MAE_0} = \frac{8{,}7}{8{,}7} = 1{,}000$

Forecasting Errors & Accuracy
Sources of errors & future uncertainty (1)

- ***Erroneous Identification of Patterns and Relationships***:
  - An illusory pattern or relationship might be identified when none really exists.
    - People often glimpse illusory correlation, while statistical models based on a small number of observations can "identify" a ***pattern that is not maintained*** over a longer period.
    - Similar, a relationship between two variables might be spurious, existing only because a ***third factor*** causes both variables to move in the same direction.
    - Alternatively, patterns or relationships that exist might be ***incorrectly identified or ignored*** because insufficient information is available, or because reality is too complex to be understood or modelled with a limited number of variables.

  - Illusory or inappropriate identification ***can cause serious and non-random forecasting errors***, since the future could turn out to be very different from what was postulated by an erroneous pattern or relationship.

BOSCH

Forecasting Errors & Accuracy
Sources of errors & future uncertainty (2)

- ***Inexact Patterns or Imprecise Relationships***:
  - Although an average pattern or relationship can be identified, fluctuations around such an average exist in almost all cases.
    - The purpose of statistical modelling is to identify patterns or relationships in such a way as to make ***past fluctuations around the average as small and random*** as possible.
    - Whether or not this is a good strategy is questionable, but even if it is appropriate, it does not guarantee that future errors will be random or symmetric or that they will not exceed a certain magnitude.
- ***Changing Patterns or Relationships***:
  - Patterns or relationships may constantly ***changing over time*** in a way that is not predictable in the great majority of cases.
    - Changes in patterns or relationships can cause large persistent errors whose magnitude cannot be known in advance.
    - The size of such errors depends on the magnitude and duration of the change.

7. STATISTICAL FORECASTING METHODS – BASIC

"Forecasting future events is often like searching for a black cat in an unlit room, that may not even be there."

Steve Davidson

Forecasting Methods
Forecasting: Horizon

- Rarely we need to predict only the following period (t+1) value of a timeserie.
- In practice, we are asked to estimate forecasts for several periods in the future.
- ***Forecasting Horizon***:
  - ***Short***-Term forecasting: usually when the forecast horizon is less than or equal to 2 periods.
    - Cases for selecting short-term: Inventory, Warehouse design.
  - ***Mid***-Term forecasting: usually when the forecast horizon is 1+ economical year (thus, 12-15 months if we refereed to a monthly timeserie).
    - Cases for selecting mid-term: Budget, Economical Planning.
  - ***Long***-Term forecasting: usually when the forecast horizon is more than 3 years.
    - Cases for selecting Long-Term: Investment planning and development.

- Not all forecasting methods are suitable for short or long term forecasting.
- Depending on the forecast horizon, the appropriate statistical forecasting method must be selected.

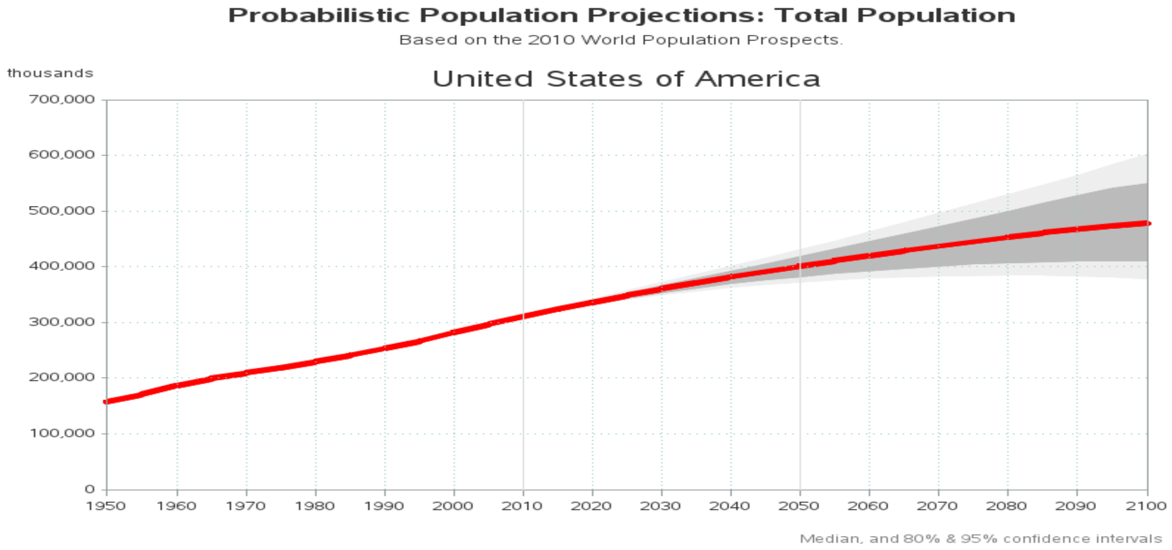
Forecasting Methods
Forecasting: Intervals

- ***Probability*** (confidence interval): is a type of interval estimate of a population parameter
  - It is used to indicate the reliability of an estimate.
  - The probability of the value of the measure to be included within the confidence intervals, is determined by the confidence level. By increasing the desired level, the confidence interval "widens“.

$$F_i = F_i \pm (t * RMSE * \sqrt{i-n})$$

where:

- F: is the forecast values
- t: is the confidence parameter, depending of the confidence level
  - o when 99% --> t = 2.580
  - o when 98% --> t = 2.330
  - o when 95% --> t = 1.960
  - o when 90% --> t = 1.645
  - o when 80% --> t = 1.280
- n: is the number of available observations
- RMSE: is the square root of the Mean Square Error (MSE)



 Internal | CR/AEU2 - Patelis | 2018-03-14

BOSCH

Source: SAS blog

Forecasting Methods
Forecasting

- There is ***no such thing as the best approach or method***.
- We are not looking for “winners” and “losers”.
- We need to understand:
  - how various forecasting methods differ from each other, and
  - how information can be provided so that forecasting users can be able to make rational choices for their situations.

Internal | CR/AEU2 - Patelis | 2018-03-14

Forecasting Methods
Forecasting

- Can we estimate forecasts by using our brain?
  - For a timeserie such as 2, 4, 6, 8, 10, 12, 14 it is easy.
  - When numbers grow in a non-linear way, such as 2, 8, 32, 128, our predictions are usually much too low.
- Question:
  - An A4 sheet of paper is 0.193mm thick.
  - Let's assume that you can fold it 40 times.
  - How ***thick*** do you think the folded sheet would be?
    - More than 2cm?
    - Less?

Forecasting Methods
Forecasting

- Answer:
  - Most people give answers such as 1cm, or possible 2cm.
  - If fact, if we do the calculations we'll get 0,193mm * 2^{40} = ***212.206 km***!
- That's more than half the distance to the moon.

- Surely, it is impossible to fold the sheet 40 times, it would become too thick after about six folds.
- But, would also be impossible without a ladder into outer space…

BOSCH

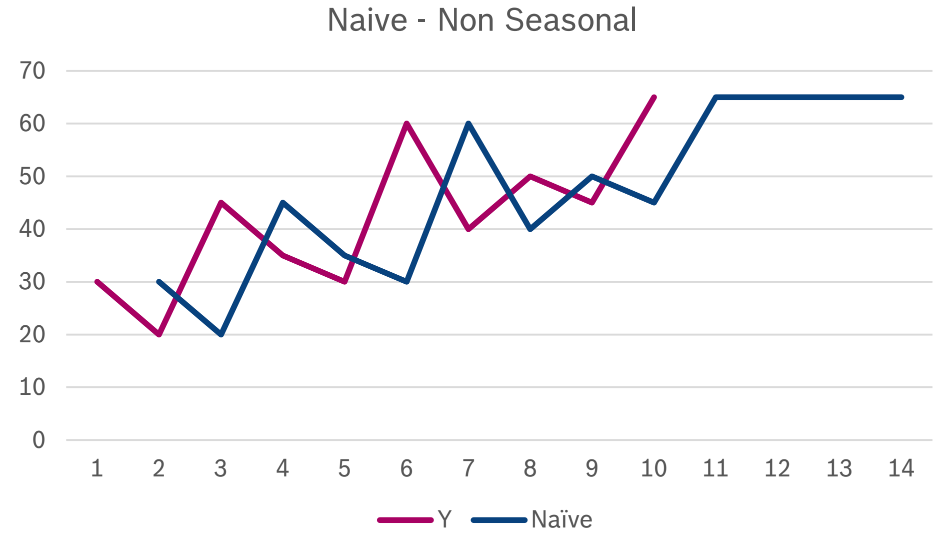
SIMPLE METHODS

Forecasting Methods
Simple methods (1)

- ***Naïve – Non Seasonal (Naïve 1)*:** The simplest method for estimation
  - the most cost-effective and efficient objective forecasting model
  - provide a ***benchmark*** against which other more sophisticated models can be compared.
  - The forecast for the next period (t+1) is equal with the last (t) historical observation.

X	Y	Naïve
1	30	
2	20	30
3	45	20
4	35	45
5	30	35
6	60	30
7	40	60
8	50	40
9	45	50
10	65	45
11		65
12		65
13		65
14		65

$$Y_{t+1} = Y_t$$

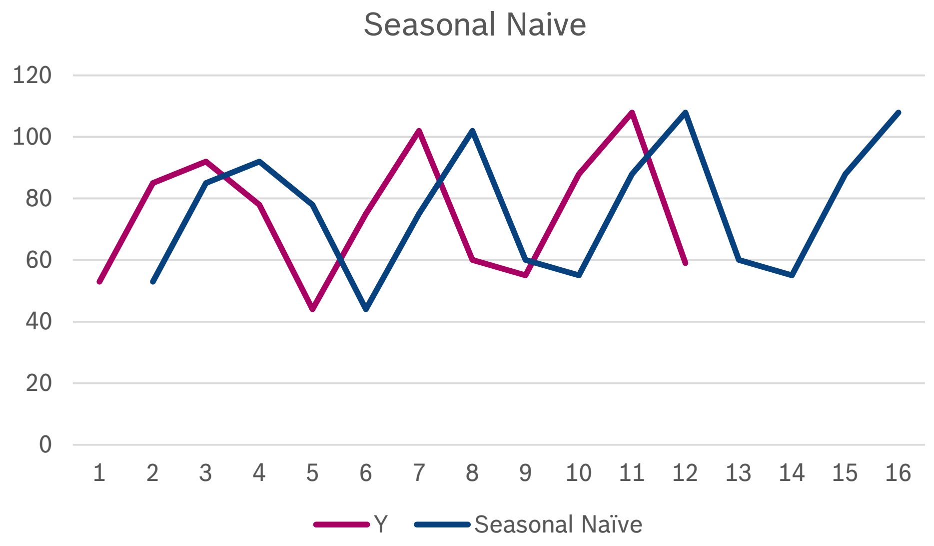


Forecasting Methods
Simple methods (2)

- ***Seasonal Naïve***:
  - Same as Naïve 1, but taking into account last-observed seasonality.
  - provide also a ***benchmark*** against which other more sophisticated models can be compared.
  - The forecast for the next period (t+1) is equal with the last (t-k) historical observation of the same period.

X	Y	Seasonal Naïve
1	53	
2	85	53
3	92	85
4	78	92
5	44	78
6	75	44
7	102	75
8	60	102
9	55	60
10	88	55
11	108	88
12	59	108
13		60
14		55
15		88
16		108

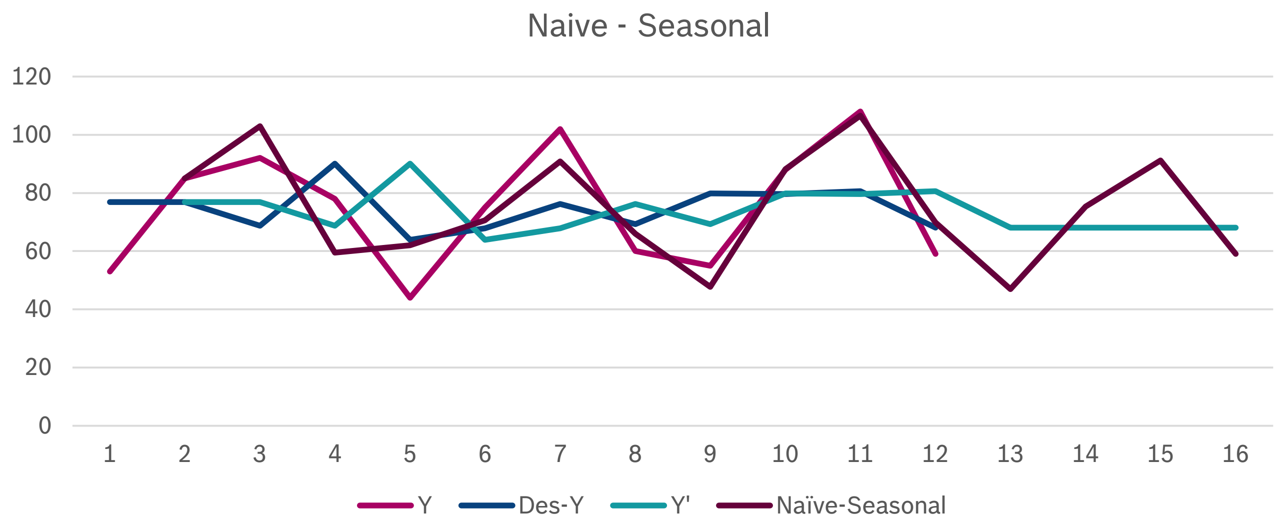
$$Y_{t+1} = Y_{t-k}$$



Forecasting Methods
Simple methods (3)

- ***Naïve – Seasonal (Naïve 2)***:
  - Same as Naïve 1, but the data is seasonal adjusted, by applying Decomposition for removing Seasonality.
  - provide also a ***benchmark*** against which other more sophisticated models can be compared.
  - The forecast for the next period (t+1) is: $Y_{t+1} = Y_t \times SI_k$

X	Y	DMA(4)	Des-Y	Y'	Naïve-Seasonal
1	53	73	76,9		
2	85	77	76,9	76,9	85,0
3	92	76	68,7	76,9	102,9
4	78	74	90,0	68,7	59,5
5	44	74	63,9	90,0	62,0
6	75	73	67,8	63,9	70,6
7	102	72	76,2	67,8	90,8
8	60	75	69,3	76,2	66,0
9	55	77	79,8	69,3	47,7
10	88	78	79,6	79,8	88,3
11	108	81	80,7	79,6	106,6
12	59	84	68,1	80,7	69,9
13				68,1	46,9
14				68,1	75,3
15				68,1	91,2
16				68,1	59,0

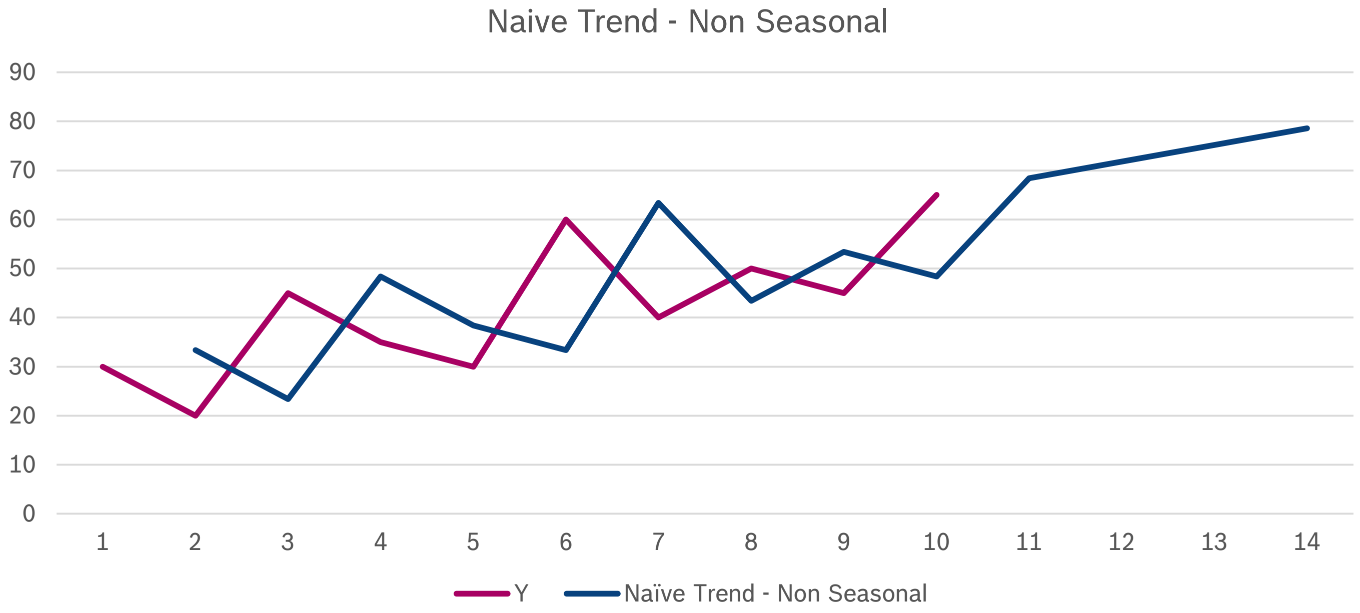


Forecasting Methods
Simple methods (4)

- ***Naïve Trend – Non Seasonal***:
  - Same as Naïve 1, but we take into account the data slope, by applying Linear Regression.
  - The forecast for the next period (t+1) is: $Y_{t+1} = Y_t + Slope$

X	Y	Naïve Trend - Non Seasonal
1	30	
2	20	33,4
3	45	23,4
4	35	48,4
5	30	38,4
6	60	33,4
7	40	63,4
8	50	43,4
9	45	53,4
10	65	48,4
11		68,4
12		71,8
13		75,2
14		78,6

(slope = 3,3939)

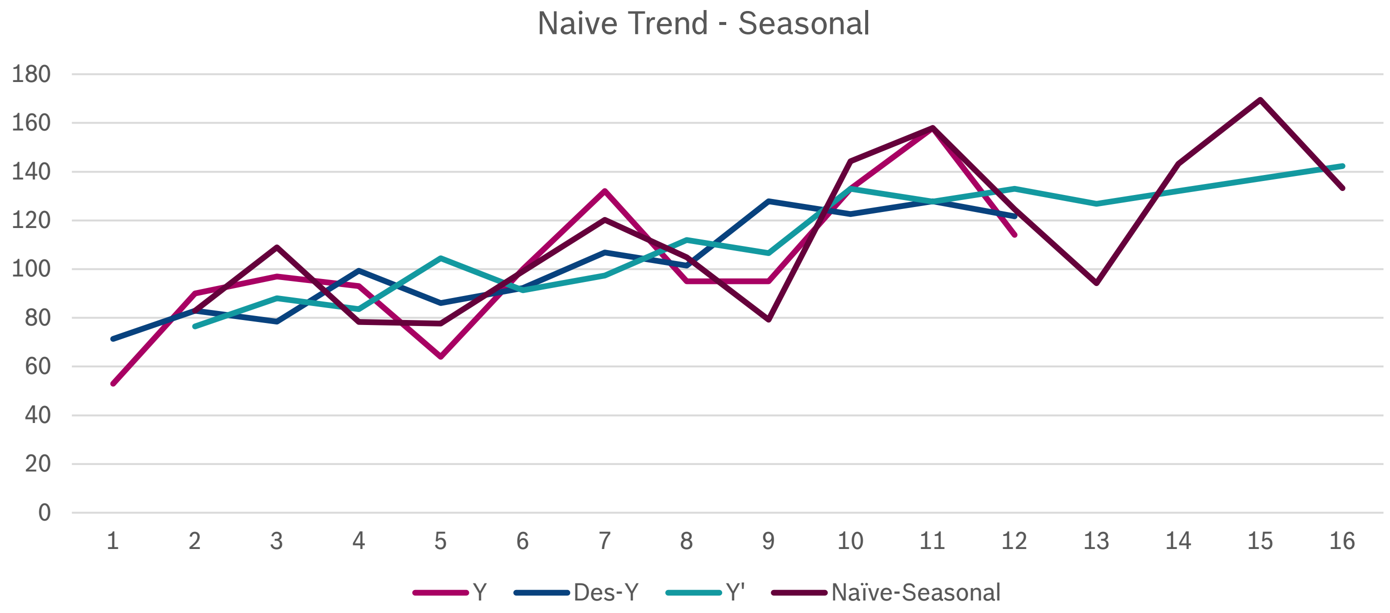


Forecasting Methods
Simple methods (5)

- ***Naïve Trend – Seasonal***:
  - Same as Naïve Trend, but the data is seasonal adjusted, by applying Decomposition for removing Seasonality.
  - The forecast for the next period (t+1) is: $Y_{t+1} = (Y_t + Slope) \times SI_k$

X	Y	Des-Y	Y'	Naïve-Seasonal
1	53	71,3		
2	90	83,0	76,5	83,0
3	97	78,5	88,1	108,9
4	93	99,3	83,6	78,3
5	64	86,1	104,4	77,6
6	100	92,2	91,3	99,0
7	132	106,8	97,3	120,3
8	95	101,4	111,9	104,8
9	95	127,8	106,6	79,2
10	133	122,6	133,0	144,3
11	158	127,8	127,7	157,9
12	114	121,7	133,0	124,5
13			126,9	94,3
14			132,0	143,2
15			137,1	169,5
16			142,3	133,3

(slope = 5,1408)



MOVING AVERAGES

Forecasting Methods
Moving Averages

- Forecasting methods for estimating ***mean values and minimizing random components***.
- They can be used as:
  - A **deseasonalization method**, in order to smooth the data, thus we estimate the Trend-Cycle component, by using the mean value of the nearby observation points.
  - A **forecasting method**, thus, we estimate the value of the (t+1) observation, by using the mean value of the latest observations.

- Four types of Moving Averages:
  - ***Simple*** Moving Average (SMA)
  - ***Weighted*** Moving Average (WMA)
  - ***Double*** Moving Average (DMA)
  - ***Centered*** Moving Average (CMA)

Forecasting Methods
Simple Moving Average – SMA (1)

- SMA: The ***unweighted mean*** of the previous n data points.
  - However, in science and engineering the mean is normally taken from an **equal number of data** on either side of a central value.
- The scope is to eliminate part of the random component, and estimate the Trend-Cycle component.
- ***Hypothesis***: "*Observations that are nearby in time as also likely to be close in value, and the average eliminates some of the data randomness, leaving a smooth trend-cycle component*".
- The order (m) of the moving average determines the smoothness of the trend-cycle estimate.
- In general, a ***larger order means a smoother curve***.

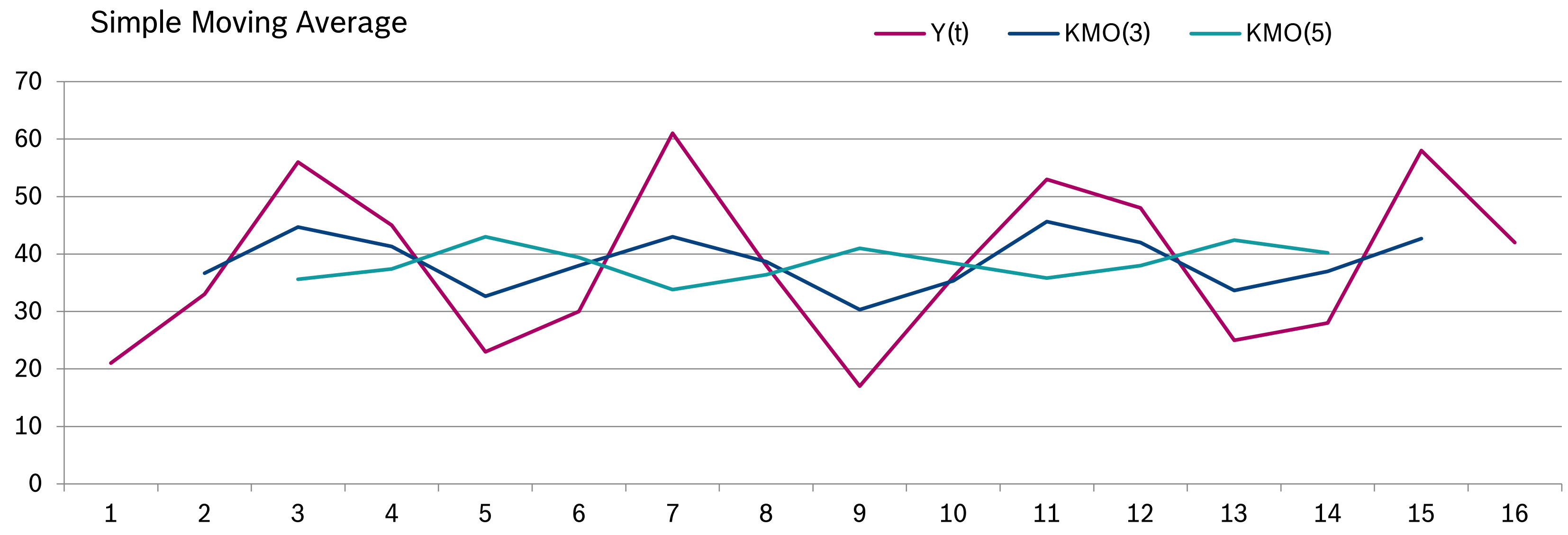
$$SMA(m)_t = \frac{1}{m} \times \sum_{i=-k}^{k} Y_{t+i}$$,where m = 2k+1.

 Internal | CR/AEU2 - Patelis | 2018-03-14

BOSCH

Forecasting Methods
Simple Moving Average – SMA Example (2)

	Y(t)	SMA(3)	SMA(5)
1	21,0		
2	33,0	36,67	
3	56,0	44,67	35,60
4	45,0	41,33	37,40
5	23,0	32,67	43,00
6	30,0	38,00	39,40
7	61,0	43,00	33,80
8	38,0	38,67	36,40
9	17,0	30,33	41,00
10	36,0	35,33	38,40
11	53,0	45,67	35,80
12	48,0	42,00	38,00
13	25,0	33,67	42,40
14	28,0	37,00	40,20
15	58,0	42,67	
16	42,0		



$$SMA(5)_3 = \frac{1}{5} \times \sum_{i=-2}^{2} Y_{3+i} = \frac{1}{5} \times (21 + 33 + 56 + 45 + 23) = 35{,}6$$

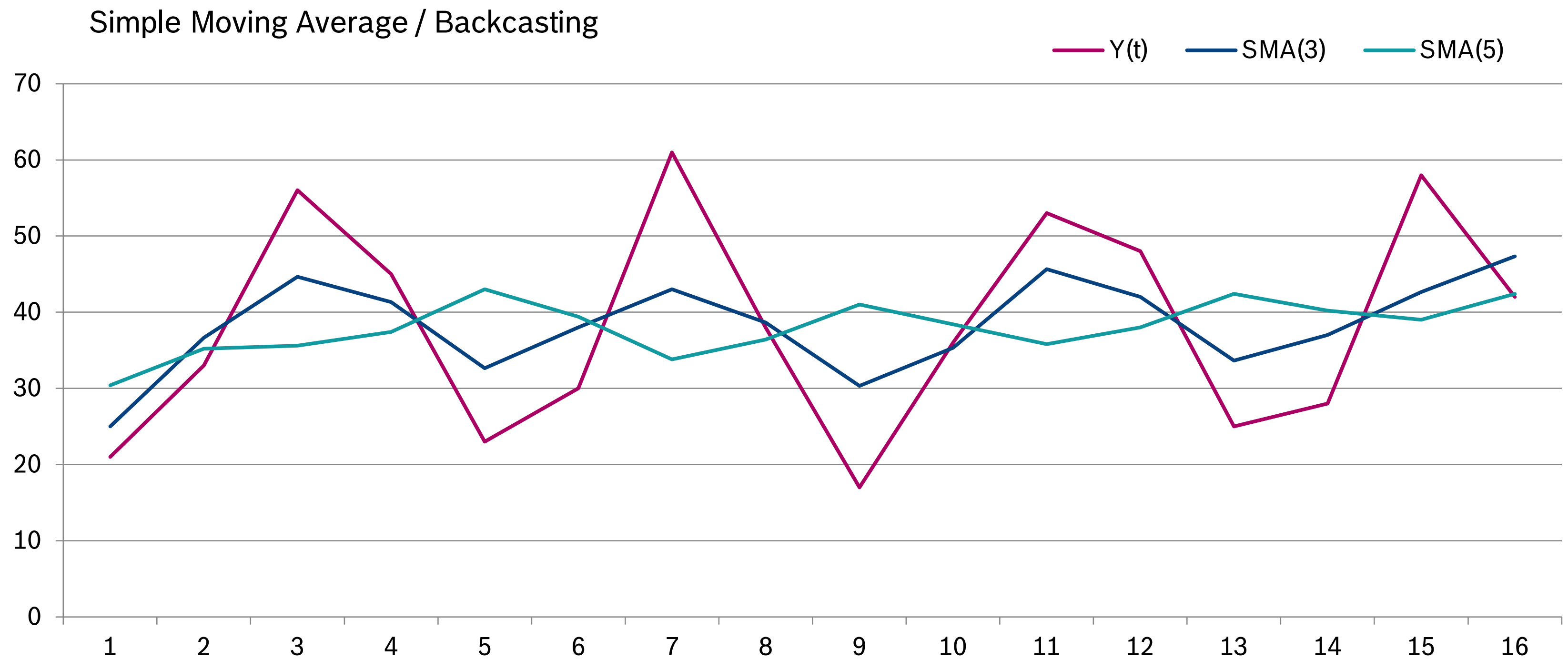
- OK! But, what about the missing points? → ***Backcasting technique***

BOSCH

Forecasting Methods

Simple Moving Average – SMA Example with Backcasting (3)

	Y(t)	SMA(3)	SMA(5)
-1	**21**		
0	**21**		
1	21	25,00	30,40
2	33	36,67	35,20
3	56	44,67	35,60
4	45	41,33	37,40
5	23	32,67	43,00
6	30	38,00	39,40
7	61	43,00	33,80
8	38	38,67	36,40
9	17	30,33	41,00
10	36	35,33	38,40
11	53	45,67	35,80
12	48	42,00	38,00
13	25	33,67	42,40
14	28	37,00	40,20
15	58	42,67	39,00
16	42	47,33	42,40
17	**42**		
18	**42**		



- Extend the timeserie, by using the first and the last values (Naïve method).

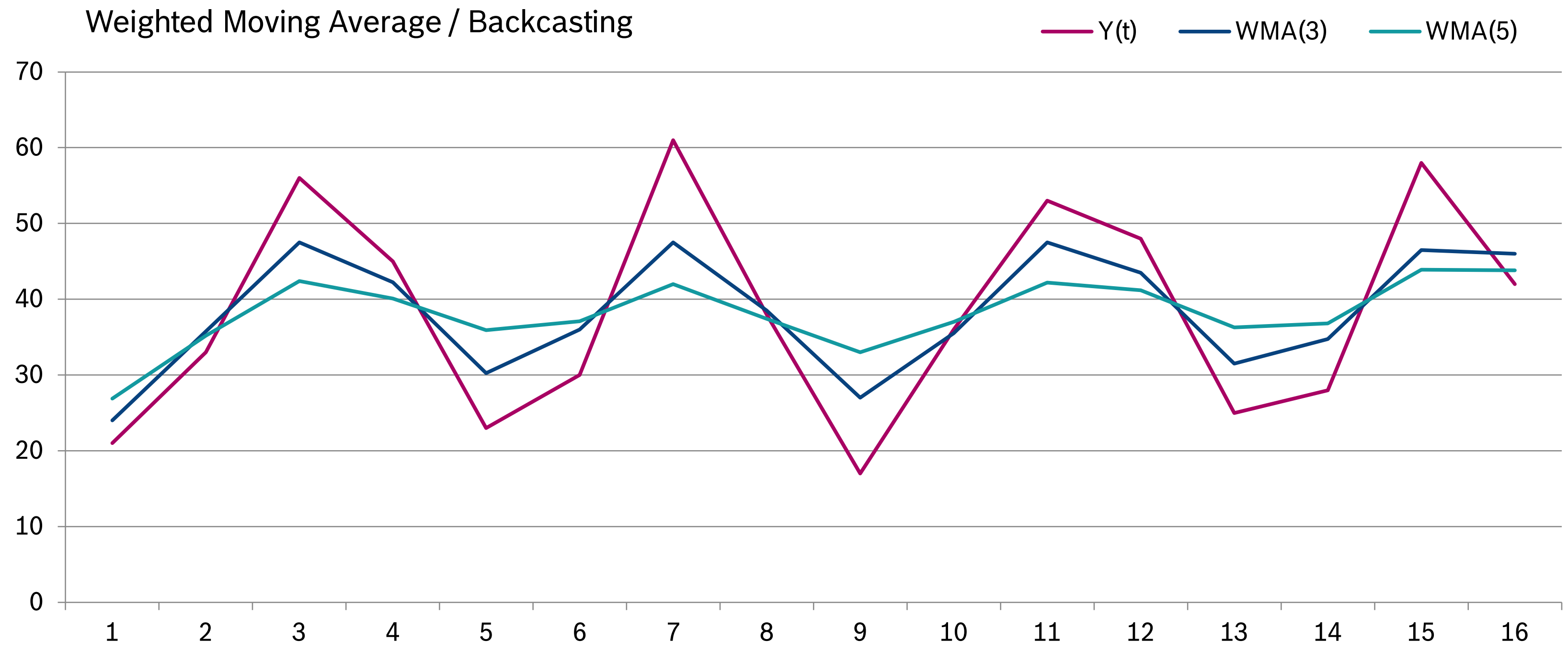
Forecasting Methods
Weighted Moving Average – WMA (1)

- WMA: The ***weighted mean of equal number of data*** on either side of a central value
- Weights:
  - Decreasing when we are moving away from the central value.
    - the nearby data have large weight, and the further away data have smaller weight.
  - The selection of the weights is in a symmetric way.
  - ***The sum of weights must be equal to 1.***
- Result: A more smoothed curve.

$$WMA(m)_t = \sum_{i=-k}^{k} a_i \times Y_{t+i}$$

Forecasting Methods
Weighted Moving Average – WMA Example (2)

	Y(t)	WMA(3)	WMA(5)
-1	**21**		
0	**21**		
1	21	24,00	26,90
2	33	35,75	35,20
3	56	47,50	42,40
4	45	42,25	40,10
5	23	30,25	35,90
6	30	36,00	37,10
7	61	47,50	42,00
8	38	38,50	37,40
9	17	27,00	33,00
10	36	35,50	37,00
11	53	47,50	42,20
12	48	43,50	41,20
13	25	31,50	36,30
14	28	34,75	36,80
15	58	46,50	43,90
16	42	46,00	43,80
17	42		
18	42		



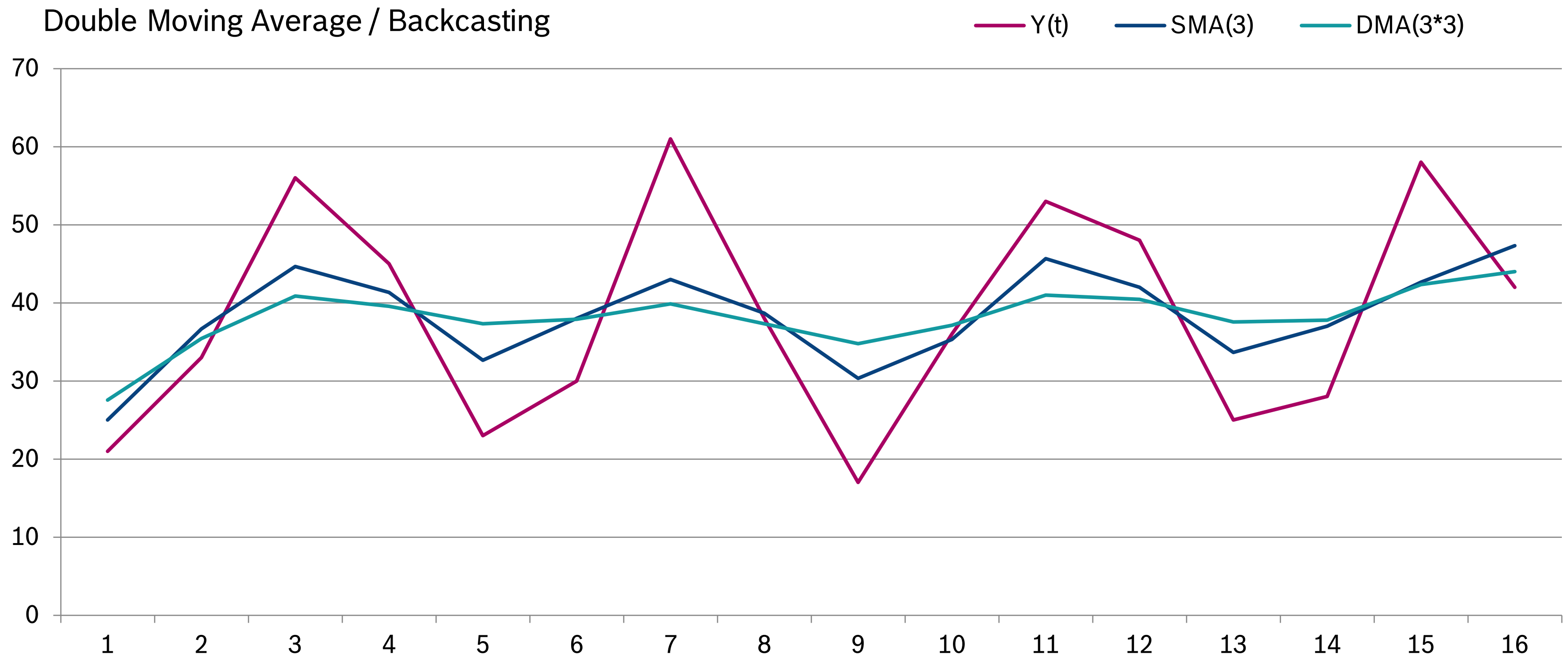
$$WMA(5)_3 = \sum_{i=-2}^{2} a_i \times Y_{3+i} = (0{,}1 \times 21) + (0{,}2 \times 33) + (0{,}4 \times 56) + (0{,}2 \times 45) + (0{,}1 \times 23) = 42{,}40$$

Forecasting Methods
Double Moving Average – DMA (1)

- DMA: the result of normalizing of a SMA, ***with a use of a second SMA*** (apply a moving average to a moving average).
- Example:
  - DMA(5x5): is a SMA(5) on an SMA(5)

Forecasting Methods
Double Moving Average – DMA Example (2)

	Y(t)	SMA(3)	DMA(3*3)
-1	21		
0	21	21,00	
1	21	25,00	27,56
2	33	36,67	35,44
3	56	44,67	40,89
4	45	41,33	39,56
5	23	32,67	37,33
6	30	38,00	37,89
7	61	43,00	39,89
8	38	38,67	37,33
9	17	30,33	34,78
10	36	35,33	37,11
11	53	45,67	41,00
12	48	42,00	40,44
13	25	33,67	37,56
14	28	37,00	37,78
15	58	42,67	42,33
16	42	47,33	44,00
17	42	42,00	
18	42		



$$SMA(3)_3 = \frac{1}{3} \times (33 + 56 + 45) = 44{,}67 \qquad SMA(3)_4 = \frac{1}{3} \times (56 + 45 + 23) = 41{,}33 \qquad SMA(3)_5 = \frac{1}{3} \times (45 + 23 + 30) = 32{,}67$$

$$DMA(3x3)_4 = \frac{1}{3} \times (SMA(3)_3 + SMA(3)_4 + SMA(3)_5)) = \frac{1}{3} \times (44{,}67 + 41{,}33 + 32{,}67) = 39{,}56$$

Forecasting Methods
Centered Moving Average – CMA (1)

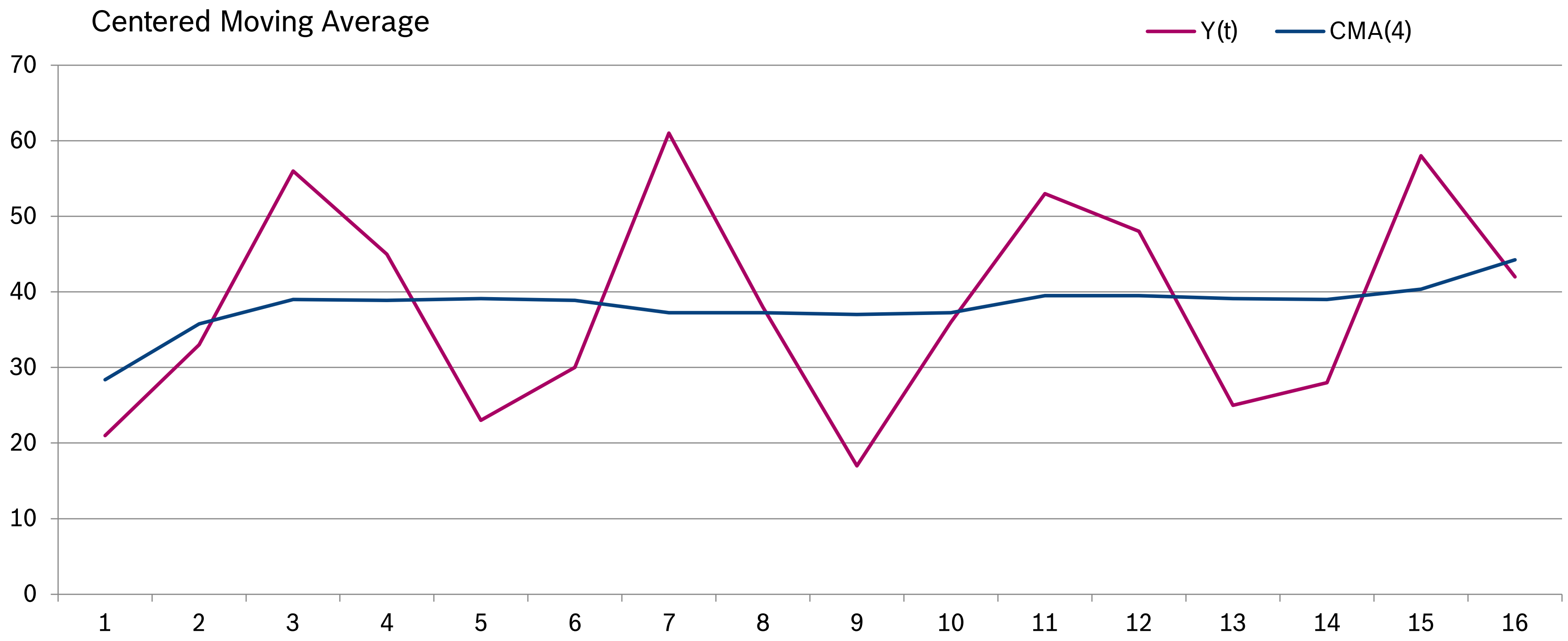
- CMA: Another version of WMA

$$CMA(m)_t = \frac{1}{2} \times \left(\frac{1}{m} \times \sum_{i=-(m\,mod\,2)}^{(m\,mod\,2)-1} Y_{t+i} \right) + \frac{1}{2} \times \left(\frac{1}{m} \times \sum_{i=-(m\,mod\,2)+1}^{(m\,mod\,2)} Y_{t+i} \right)$$

BOSCH

Forecasting Methods
Centered Moving Average – CMA Example (2)

	Y(t)	CMA(4)
-1	**21**	
0	**21**	
1	21	28,38
2	33	35,75
3	56	39,00
4	45	38,88
5	23	39,13
6	30	38,88
7	61	37,25
8	38	37,25
9	17	37,00
10	36	37,25
11	53	39,50
12	48	39,50
13	25	39,13
14	28	39,00
15	58	40,38
16	42	44,25
17	**42**	
18	**42**	

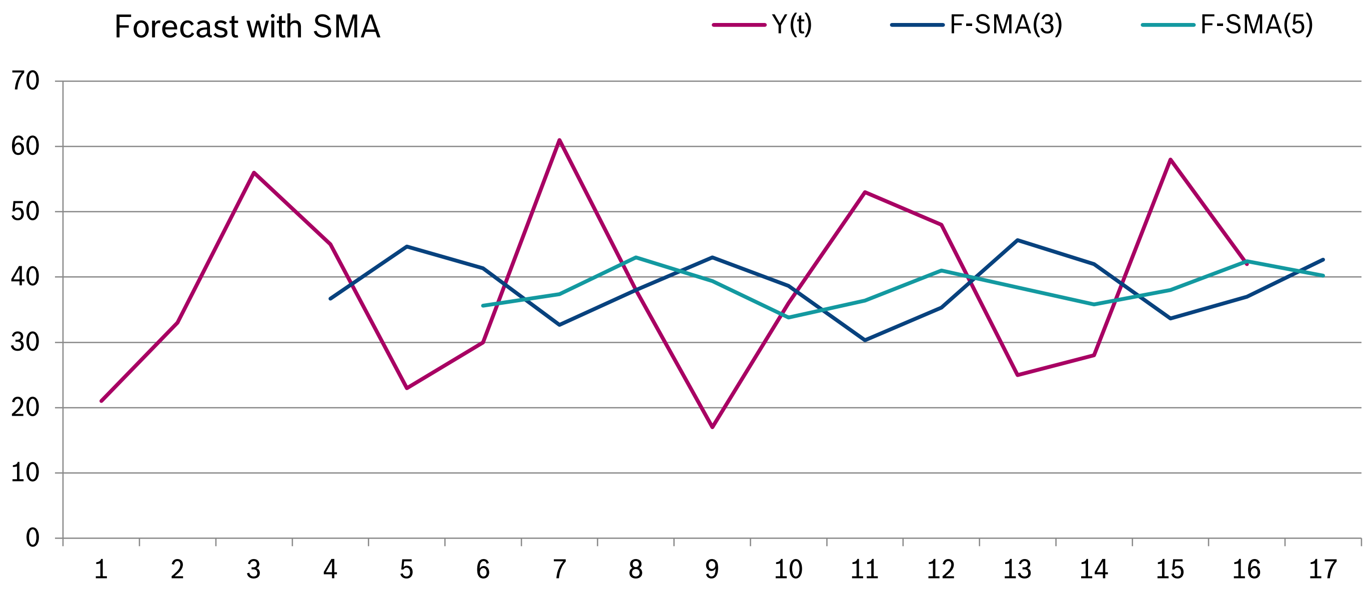


$$CMA(4)_3 = \frac{1}{2} \times \left(\frac{1}{4} \times \sum\nolimits_{i=-2}^{3} Y_{3+i} \right) + \frac{1}{2} \times \left(\frac{1}{4} \times \sum\nolimits_{i=-1}^{2} Y_{3+i} \right) = \left(\frac{1}{8} \times (Y_1 + Y_2 + Y_3 + Y_4) \right) + \left(\frac{1}{8} \times (Y_2 + Y_3 + Y_4 + Y_5) \right) = \frac{\mathbf{1}}{\mathbf{8}} Y_1 + \frac{\mathbf{1}}{\mathbf{4}} Y_2 + \frac{\mathbf{1}}{\mathbf{4}} Y_3 + \frac{\mathbf{1}}{\mathbf{4}} Y_4 + \frac{\mathbf{1}}{\mathbf{8}} Y_5 = 39{,}00$$

Forecasting Methods
Moving Average – How to forecast?

- All types of Moving Averages can be used as a forecasting method
  - How to use SMA: $F_t = SMA_{n_t} = \frac{1}{n} * \sum_{i=-n}^{-1} Y_{t+i}$
  - Example: $F_4 = SMA_{3_4} = \frac{1}{3} * \sum_{i=-3}^{-1} Y_{4+i}$

	Y(t)	SMA(3)	SMA(5)	F-SMA(3)	F-SMA(5)
1	21				
2	33	36,67			
3	56	44,67	35,60		
4	45	41,33	37,40	36,67	
5	23	32,67	43,00	44,67	
6	30	38,00	39,40	41,33	35,60
7	61	43,00	33,80	32,67	37,40
8	38	38,67	36,40	38,00	43,00
9	17	30,33	41,00	43,00	39,40
10	36	35,33	38,40	38,67	33,80
11	53	45,67	35,80	30,33	36,40
12	48	42,00	38,00	35,33	41,00
13	25	33,67	42,40	45,67	38,40
14	28	37,00	40,20	42,00	35,80
15	58	42,67		33,67	38,00
16	42			37,00	42,40
17				42,67	40,20



BOSCH

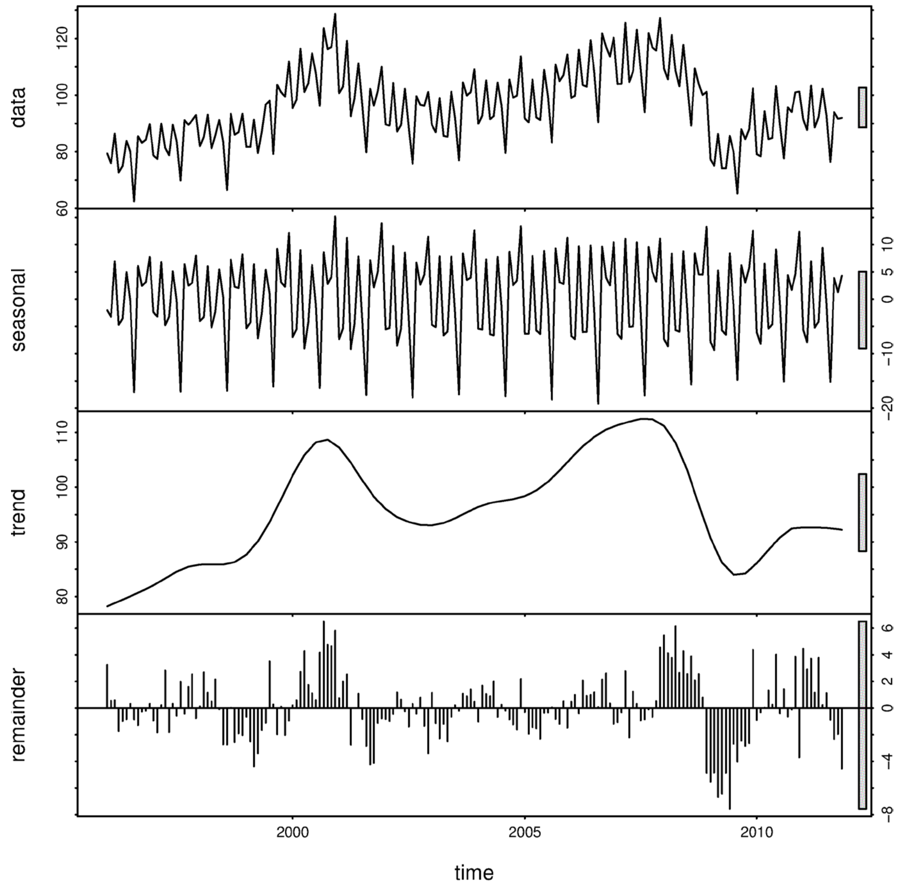
DECOMPOSITION

Forecasting Methods
Decomposition (1)

Decomposition:

- A statistical method that deconstructs a timeserie (Y) into the four basic parts (components):
  - ***Trend*** (T) component.
  - ***Cyclical*** (C) component.
  - ***Seasonal*** (S) component.
  - ***Random*** (R) or Irregular component.

$$Y_t = f(S_t, T_t, C_t, R_t)$$



- Two forms of decomposition:
  - ***Additive*** $Y_t = S_t + T_t + C_t + R_t$
  - ***Multiplicative*** $Y_t = S_t \times T_t \times C_t \times R_t$

 Internal | CR/AEU2 - Patelis | 2018-03-14

BOSCH

Source: Forecasting – Principles and Practice

Forecasting Methods
Decomposition (2)

Deseasonalization:

- A method for:
  - ***isolating the seasonal component*** from a timeserie, and
  - creating the deseasonalized version of the timeserie

- We must always apply forecasting methods on the deseasonalized timeseries.
  - This increases the reliability of the forecasts.

- Commonly used:
  - ***Classical Decomposition method***, based on Moving Averages

 Internal | CR/AEU2 - Patelis | 2018-03-14

Forecasting Methods
Decomposition (3)

Step 1: *Calculate Moving Average*

- MA(n): based on the length **n** of the seasonality:
  - example: MA(7) for daily timeserie, MA(12) for monthly timeserie.
  - Does not include seasonality, and
  - Contains very little or no randomness, since randomness is represented by random fluctuations that range around the average value of observations.
  - Gives a good estimate of the behaviour of timeseries Trend & Cyclical components.

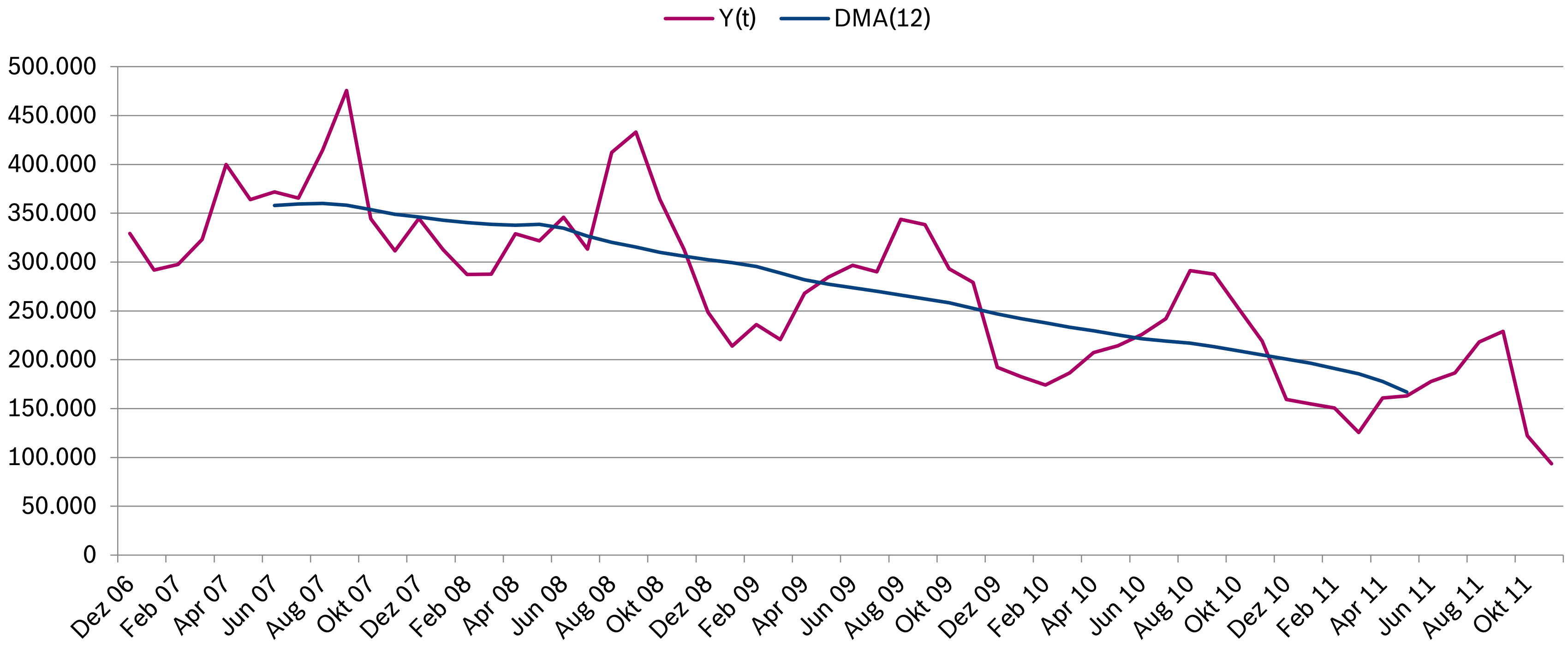
$$MA_n = T \times C$$

- ***In most cases a SMA is used***.
- But, in the case when the seasonality length n is even (like in monthly data with n = 12), then the usage of CMA is preferred.

Forecasting Methods
Decomposition (4)

Step 1: Calculate Moving Average – Example

	Y(t)	DMA(12)
Dez 06	329.225	
Jan 07	291.927	
Feb 07	297.449	
Mrz 07	323.086	
Apr 07	399.828	
Mai 07	363.939	
Jun 07	371.906	358.051,75
Jul 07	365.572	359.559,25
Aug 07	414.576	360.000,58
Sep 07	475.699	358.096,08
Okt 07	344.335	353.663,63
Nov 07	311.346	348.949,38
Dez 07	344.691	346.093,50
Jan 08	312.641	342.828,17
Feb 08	287.327	340.553,04
Mrz 08	287.500	338.678,29
Apr 08	329.035	337.731,54
Mai 08	321.590	338.628,25
Jun 08	345.714	334.691,46
Jul 08	313.396	326.568,42



Internal | CR/AEU2 - Patelis | 2018-03-14

Forecasting Methods
Decomposition (5)

Step 2: Initial Estimation of ***Seasonality***

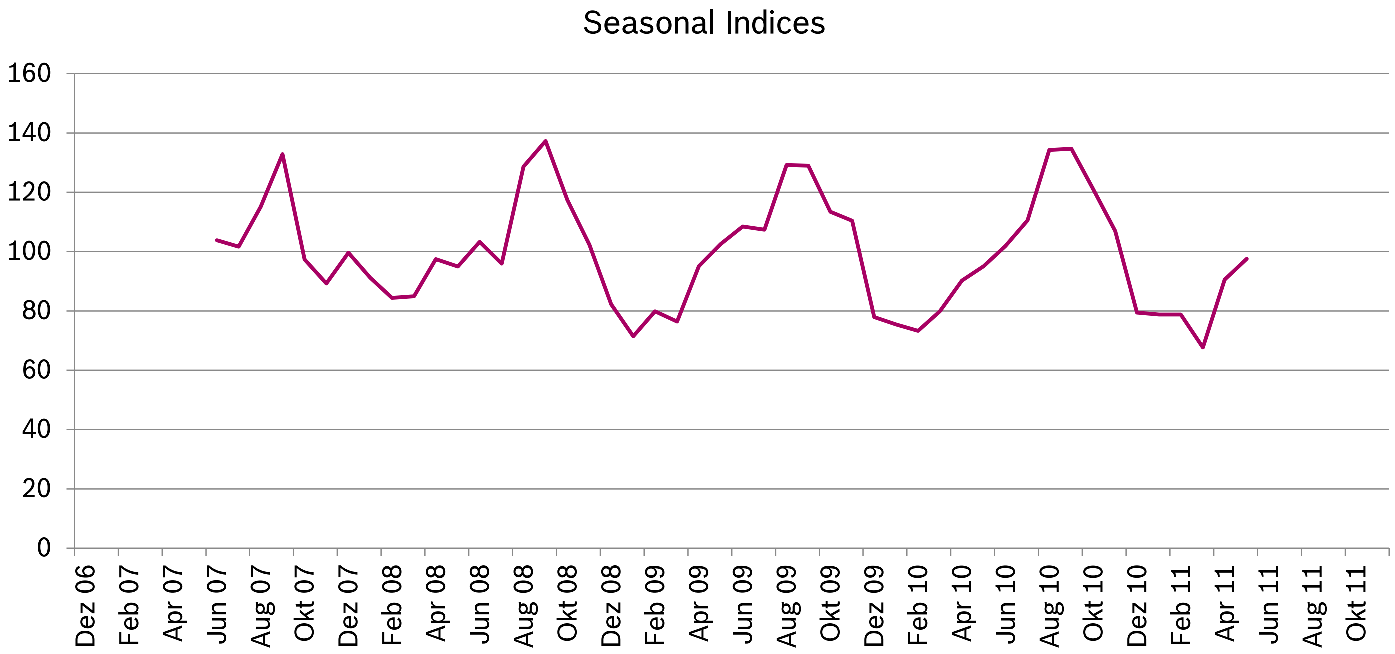
- Divide the actual observations Y_t with the T*C values (that we estimate on the 1st step).
- With this way, we can have a first estimation of Seasonality, which has also randomness.

$$\frac{Y}{MA_n} = \frac{T \times C \times S \times R}{T \times C} = S \times R$$

Forecasting Methods
Decomposition (6)

Step 2: Initial Estimation of Seasonality – Example

	Y(t)	DMA(12)	Seas.Ind.
Dez 06	329.225		
Jan 07	291.927		
Feb 07	297.449		
Mrz 07	323.086		
Apr 07	399.828		
Mai 07	363.939		
Jun 07	371.906	358.051,75	103,87
Jul 07	365.572	359.559,25	101,67
Aug 07	414.576	360.000,58	115,16
Sep 07	475.699	358.096,08	132,84
Okt 07	344.335	353.663,63	97,36
Nov 07	311.346	348.949,38	89,22
Dez 07	344.691	346.093,50	99,59
Jan 08	312.641	342.828,17	91,19
Feb 08	287.327	340.553,04	84,37
Mrz 08	287.500	338.678,29	84,89
Apr 08	329.035	337.731,54	97,43
Mai 08	321.590	338.628,25	94,97
Jun 08	345.714	334.691,46	103,29
Jul 08	313.396	326.568,42	95,97



Forecasting Methods
Decomposition (7)

Step 3: ***Remove Randomness*** from Seasonality

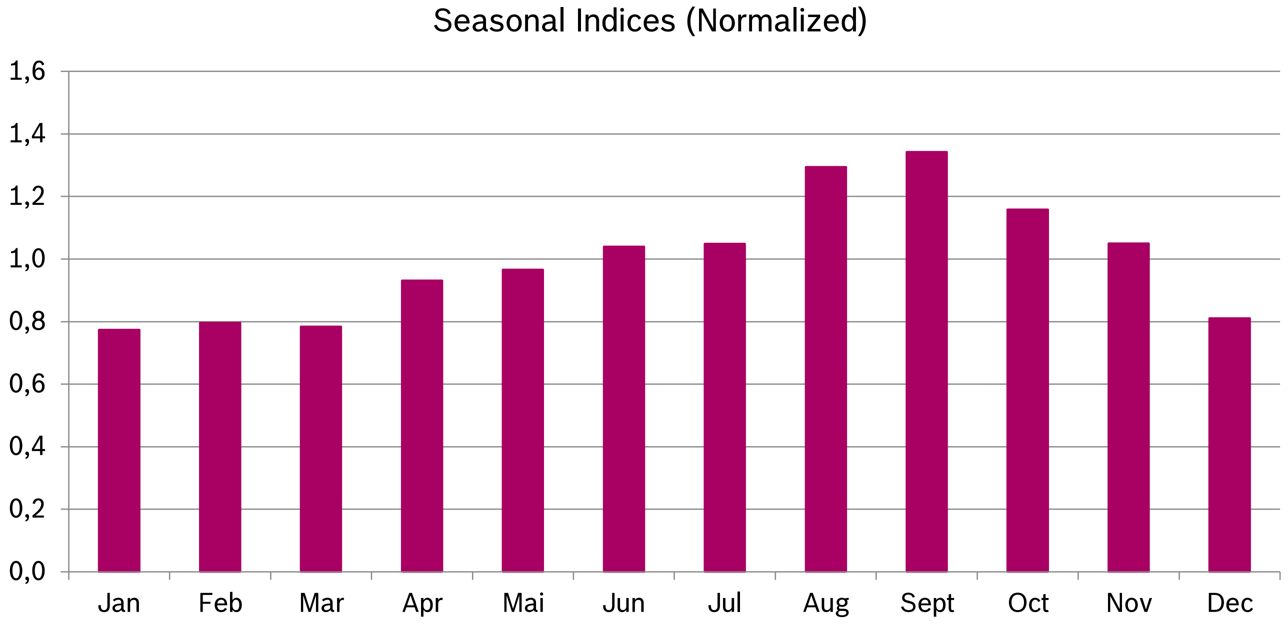
- This is achieved by finding the ***average of the corresponding seasonality indices***.
  - For a monthly timeserie, we calculate the seasonal indice for January, by averaging all seasonality indices corresponding to January.
- Do we need normalization?
  - In some cases, the seasonality indices need to be normalized, so that their ***sum is equal to the seasonality*** length n.
  - If the timeserie contains ***significant randomness or outliers***, it is suggested to use the Intermediate Averages, for the calculation of seasonality indices. The difference lies in the ***non-use of maximum and minimum*** seasonality indice in the calculation, in order to stabilize them.

Forecasting Methods
Decomposition (8)

Step 3: Remove Randomness from Seasonality – Example

	Seasonal Indices								
	2007	2008	2009	2010	2011	min	max	Average	Normalized
Jan		0,91	0,71	0,75	0,79	0,71	0,91	0,771	0,774
Feb		0,84	0,80	0,73	0,79	0,73	0,84	0,793	0,796
Mar		0,85	0,76	0,80	0,68	0,68	0,85	0,782	0,785
Apr		0,97	0,95	0,90	0,91	0,90	0,97	0,928	0,932
Mai		0,95	1,03	0,95	0,98	0,95	1,03	0,963	0,967
Jun	1,04	1,03	1,08	1,02		1,02	1,08	1,036	1,040
Jul	1,02	0,96	1,07	1,10		0,96	1,10	1,045	1,049
Aug	1,15	1,29	1,29	1,34		1,15	1,34	1,289	1,294
Sept	1,33	1,37	1,29	1,35		1,29	1,37	1,338	1,343
Oct	0,97	1,17	1,13	1,21		0,97	1,21	1,155	1,159
Nov	0,89	1,02	1,10	1,07		0,89	1,10	1,046	1,050
Dec	1,00	0,82	0,78	0,79		0,78	1,00	0,808	0,811
							sum	11,954	12,000

$$NF = \frac{\sum_{i=1}^{n} I_i}{n} = \frac{11{,}954}{12} = 0{,}9995 \qquad NI_i = \frac{I_i}{NF}$$



Forecasting Methods
Decomposition (9)

Step 4: Estimate ***deseasonalized timeserie***

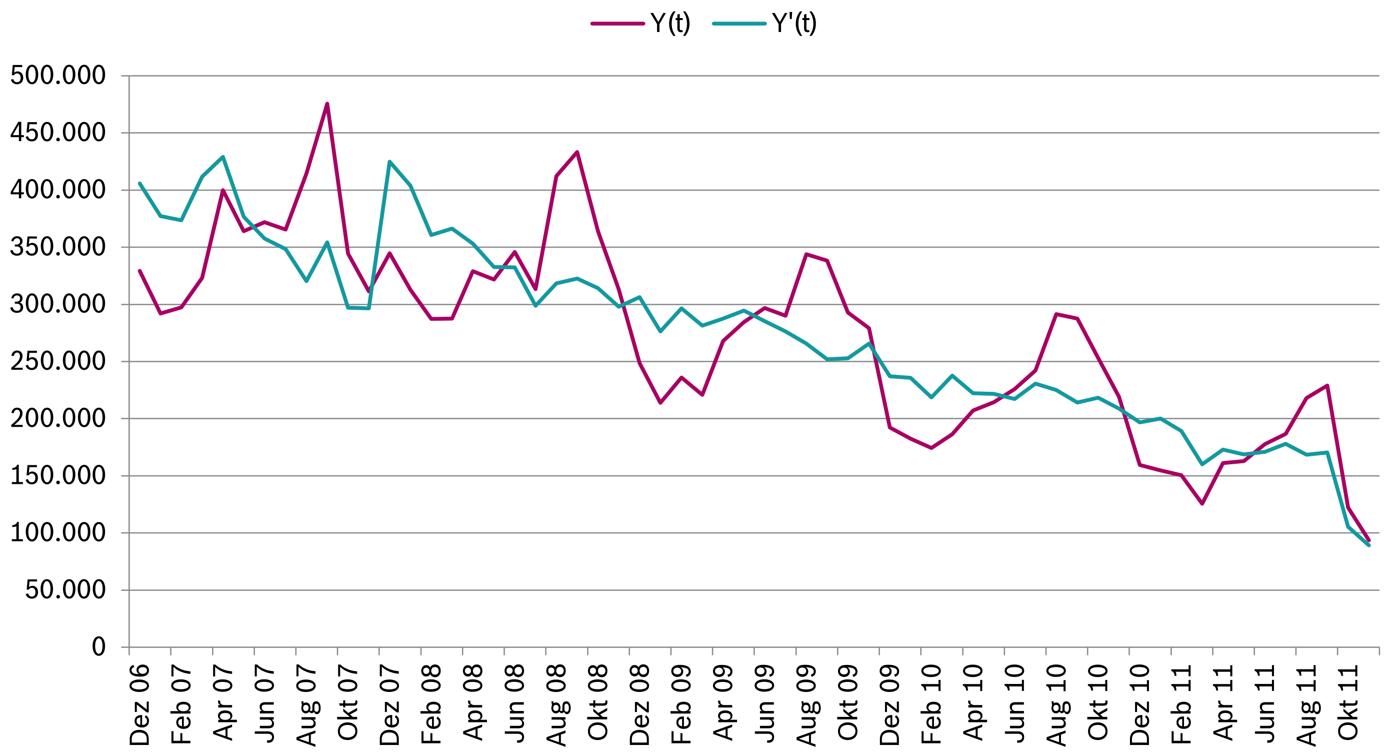
- Divide the actual observations with the seasonality indices, in order to estimate the deseasonalized timeserie.
- This timeserie now includes only Trend, Cycle and Random component.

$$\frac{Y}{S} = \frac{T \times C \times S \times R}{S} = T \times C \times R$$

Forecasting Methods
Decomposition (10)

Step 4: Estimate deseasonalized Timeserie – Example

	Y(t)	Seas.Ind.	Y'(t)
Dez 06	329.225	81,11	405.922,7
Jan 07	291.927	77,41	377.136,9
Feb 07	297.449	79,64	373.473,9
Mrz 07	323.086	78,46	411.762,6
Apr 07	399.828	93,19	429.048,4
Mai 07	363.939	96,66	376.521,8
Jun 07	371.906	103,98	357.682,2
Jul 07	365.572	104,93	348.410,0
Aug 07	414.576	129,41	320.361,3
Sep 07	475.699	134,30	354.207,6
Okt 07	344.335	115,90	297.106,3
Nov 07	311.346	105,03	296.446,9
Dez 07	344.691	81,11	424.991,7
Jan 08	312.641	77,41	403.897,0
Feb 08	287.327	79,64	360.764,9
Mrz 08	287.500	78,46	366.409,4
Apr 08	329.035	93,19	353.081,7
Mai 08	321.590	96,66	332.708,7
Jun 08	345.714	103,98	332.491,9
Jul 08	313.396	104,93	298.683,5



Forecasting Methods
Decomposition (11)

Step 5: ***Remove Randomness*** from deseasonalized Timeserie

- Calculate a ***Moving Average*** MA(3) or MA(6) of the deseasonalized timeserie.
- This new MA values are fairly smooth, they are an estimation of the Trend-Cycle component.
- For achieving optimal smoothing and elimination of randomness, it is recommended to use ***Double Moving Average DMA(3x3)***.

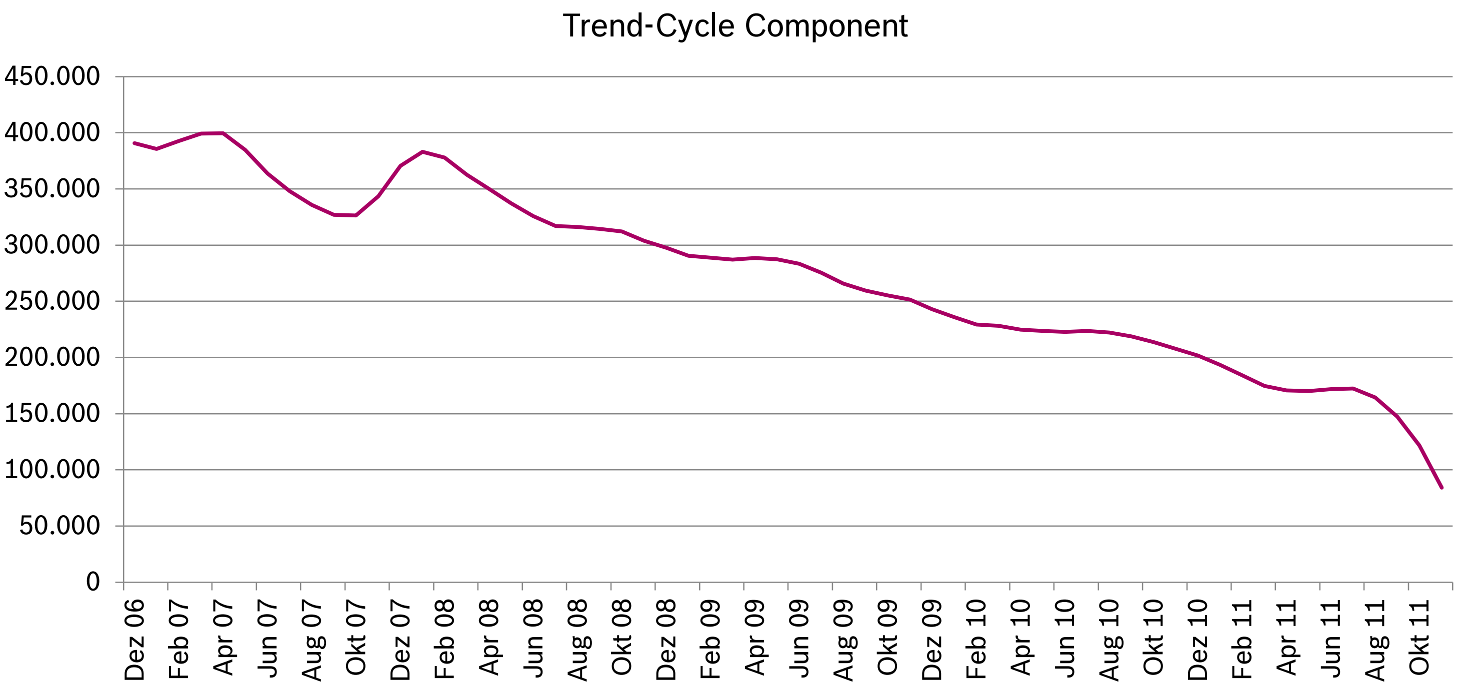
$$CMA(3x3) = T \times C$$

$$\frac{T \times C \times R}{CMA(3x3)} = \frac{T \times C \times R}{T \times C} = R$$

Timeserie Preparation & Analysis
Decomposition (12)

Step 5: Remove Randomness from deseasonalized Timeserie – Example

	Y(t)	Seas.Ind.	Y'(t)	CMA(3)	CMA(3x3)
Dez 06	329.225	81,11	405.922,7		390.556,5
Jan 07	291.927	77,41	377.136,9	385.511,2	385.511,2
Feb 07	297.449	79,64	373.473,9	387.457,8	392.576,9
Mrz 07	323.086	78,46	411.762,6	404.761,7	399.332,4
Apr 07	399.828	93,19	429.048,4	405.777,6	399.430,0
Mai 07	363.939	96,66	376.521,8	387.750,8	384.799,9
Jun 07	371.906	103,98	357.682,2	360.871,4	363.591,1
Jul 07	365.572	104,93	348.410,0	342.151,2	348.005,2
Aug 07	414.576	129,41	320.361,3	340.993,0	335.678,6
Sep 07	475.699	134,30	354.207,6	323.891,7	326.935,0
Okt 07	344.335	115,90	297.106,3	315.920,3	326.442,3
Nov 07	311.346	105,03	296.446,9	339.515,0	343.515,7
Dez 07	344.691	81,11	424.991,7	375.111,9	370.392,7
Jan 08	312.641	77,41	403.897,0	396.551,2	382.895,6
Feb 08	287.327	79,64	360.764,9	377.023,8	377.886,8
Mrz 08	287.500	78,46	366.409,4	360.085,3	362.614,1
Apr 08	329.035	93,19	353.081,7	350.733,3	350.082,0
Mai 08	321.590	96,66	332.708,7	339.427,4	337.151,8
Jun 08	345.714	103,98	332.491,9	321.294,7	325.758,6
Jul 08	313.396	104,93	298.683,5	316.553,8	317.025,2



Forecasting Methods
Decomposition (13)

Step 6: Calculate the ***Trend***

- The result of the previous step is a timeserie that includes only Trend and Cycle components.
- In case, that we need to separate these 2 components, we must find the Trend model which describes best the timeserie.
- If we assume ***Linear Trend***, then the calculation of the Trend component can be achieved by using simple linear regression (SLR) with a least square straight line.

$$T = a + \beta * t \qquad \beta = \frac{\frac{\sum_{i=1}^{n}(t_i * TC_i)}{n} - (\bar{t} - \overline{TC})}{\frac{\sum_{i=1}^{n} t_i^2}{n} - \overline{t^2}}, \qquad \alpha = \overline{TC} - (\beta * \bar{t})$$

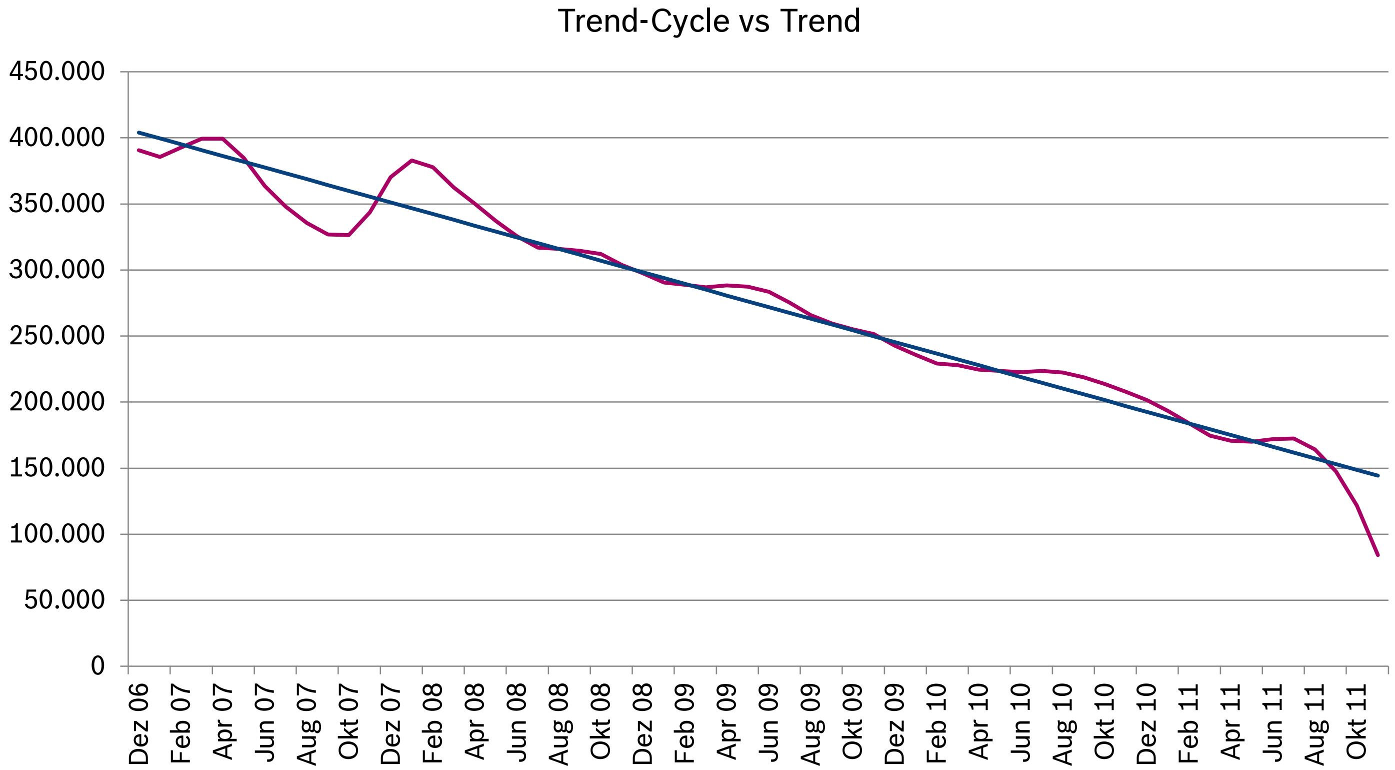
Forecasting Methods
Decomposition (14)

Step 6: Calculate the Trend – Example

$$T = a + \beta \times t$$

$$a = 408.373{,}9$$

$$\beta = -4.402{,}47$$



Forecasting Methods
Decomposition (15)

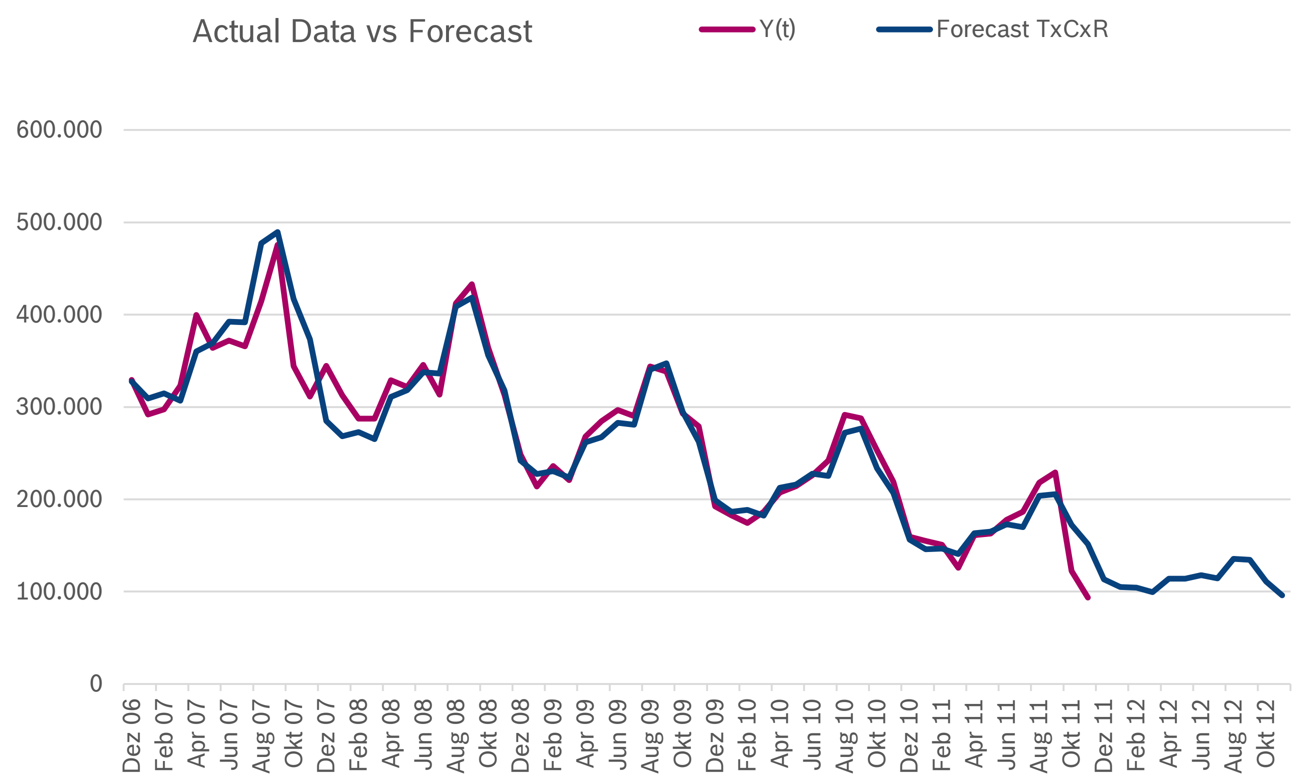
Step 7: ***Forecasting***

- Use the timeserie components for forecasting future values.
- Thus:
  - ***Extent the TxC into the future***, depending on the trend type
  - ***Multiply*** (or add) the ***seasonal*** component.

Forecasting Methods
Decomposition (16)

Step 7: Forecasting – Example

X		Y(t)	Forecast TxC	S	Forecast TxCxS
1	Dez 06	329.225	403.971,4	0,811	327.642,4
2	Jan 07	291.927	399.569,0	0,774	309.290,8
3	Feb 07	297.449	395.166,5	0,796	314.725,8
4	Mrz 07	323.086	390.764,0	0,785	306.609,6
5	Apr 07	399.828	386.361,5	0,932	360.048,4
...	...	...	...	...	...
58	Sep 11	228.940	153.030,6	1,343	205.519,4
59	Okt 11	122.245	148.628,2	1,159	172.254,5
60	Nov 11	93.667	144.225,7	1,050	151.474,3
61	Dez 11		139.823,2	0,811	113.404,1
62	Jan 12		135.420,8	0,774	104.823,9
63	Feb 12		131.018,3	0,796	104.348,0
64	Mrz 12		126.615,8	0,785	99.348,0
65	Apr 12		122.213,4	0,932	113.890,0
66	Mai 12		117.810,9	0,967	113.873,8
67	Jun 12		113.408,4	1,040	117.918,3
68	Jul 12		109.005,9	1,049	114.375,4
69	Aug 12		104.603,5	1,294	135.366,2
70	Sep 12		100.201,0	1,343	134.569,4
71	Okt 12		95.798,5	1,159	111.026,9
72	Nov 12		91.396,1	1,050	95.989,5



Forecasting Methods
Decomposition (17)

Important Notice:

- In cases that the timeserie shows no significant seasonal behaviour, the ***steps 1-4 can be skipped***.
- We can evaluate this, by comparing:
  - The Autocorrelation of the data, with a delay period (k) equal to the number of the seasonality length (pos).
  - The autocorrelations of the data, with delay periods (k) till 1 period smaller than the seasonality length (pos).
- A timeserie has significant seasonal behaviour (with 90% confidence), only when:

$$\left|ACF_{pos}\right| > Limit$$

$$ACF_k = \frac{\sum_{i=1+k}^{n}\left[(Y_i - \bar{Y}) * (Y_{i-k} - \bar{Y})\right]}{\sum_{i=1}^{n}(Y_i - \bar{Y})}$$

$$Limit = 1{,}645 \times \sqrt{\frac{1 + 2 * (ACF_1 + \sum_{i=2}^{pos-1} ACF_i^2)}{n}}$$

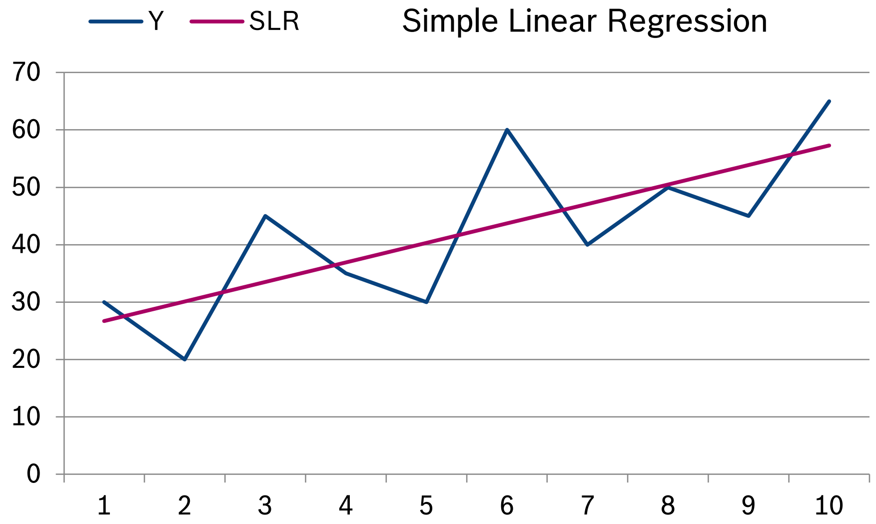
REGRESSION LINEAR

Forecasting Methods
Regression

- In statistics, regression analysis examines:
  - the ***relationship*** between a ***y*** dependent variable
    - regressand / dependent / explained variable / variable reaction / response
  - with certain independent variables ***x***
    - regressors / independents / explanatory variables.

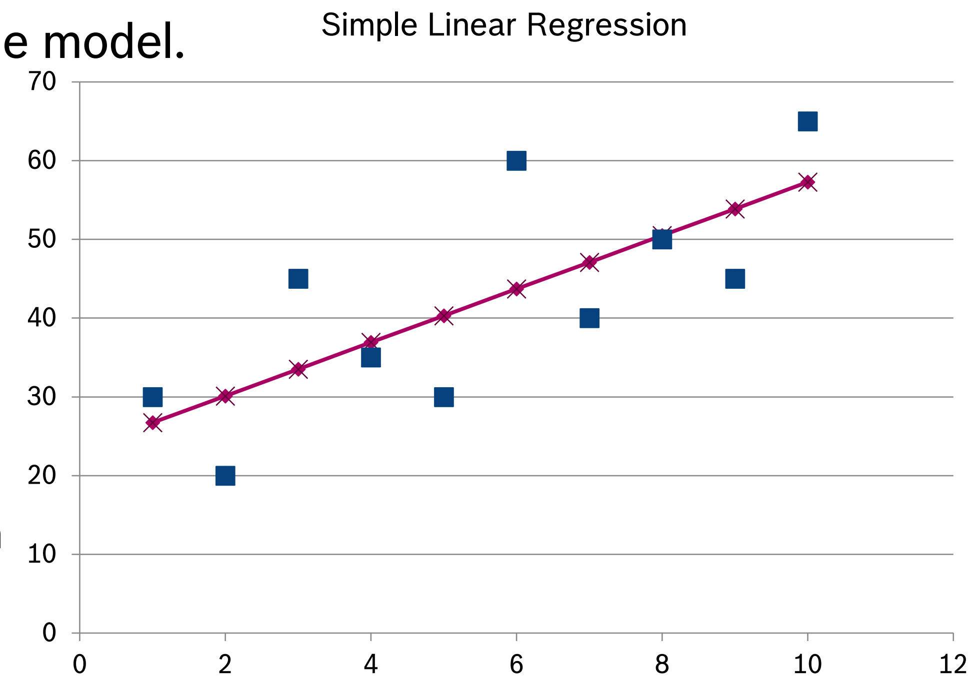
Forecasting Methods
Simple Linear Regression (1)

- ***Simple Linear Regression (SLR)***: is the least squares estimator of a linear regression model with a single explanatory variable.
  - SLR fits a straight line through the set of *N* points in such a way that makes the sum of squared residuals of the model (the vertical distances between the points of the data sets and the fitted line) as small as possible.
- The "Simple" refers to the fact that this method is one of the simplest.
  - The slope of the fitted line is equal to the ***correlation*** between *x* and *y,* corrected by the ratio of standard deviations of these variables.
  - The intercept of the fitted line is such that it passes through the ***center of mass*** of the data points.



Forecasting Methods
Simple Linear Regression (2)

- The data points do not lie on the straight line, but are ***scattered around it***.
- The distance of each data point from the line is the ***residual error***. The error term does not imply a mistake, but a deviation from the underlying straight line model.
- Error assumptions:
  - They have ***mean zero***; otherwise the forecasts will be systematically biased.
  - They are ***not auto-correlated***; otherwise the forecasts will be inefficient as there is more information to be exploited from the data.
  - They are ***unrelated to the predictor value***; otherwise there would be more information that should be included in the systematic part of the model.



BOSCH

Forecasting Methods
Simple Linear Regression (3)

- The SLR model equations:

$$Y_t = \alpha + \beta \times X_t$$

$$\beta = \frac{\sum_{i=1}^{N}\left((Y_i - \bar{Y}) \times (X_i - \bar{X})\right)}{\sum_{i=1}^{N}(X_i - \bar{X})^2}$$

$$\alpha = \bar{Y} - \beta \times \bar{\mathrm{X}}$$

- For each value of *X*, we can forecast a corresponding value of *Y*.
- The forecasted values are called ***fitted values***.
- The difference between the actual *Y* values and the fitted values are called ***residuals***.

Forecasting Methods
Simple Linear Regression: Correlation Coefficient (1)

- ***Prerequisite*** for using simple linear regression: The value of a variable depends on the value (or in the change of value) of another variable.
  - Often, however, two variables may be related but without being able to identify how the values of the first variable affect the values of the other variable.
- Correlation coefficient ***r***: A measure of the degree of correlation between two variables.
  - It measures the strength and the direction (positive or negative) of the linear relationship between two variables. It can be interpreted in two ways:
    - As an indicator of the ***direction*** of the relationship between the two variables, thus
      - if their values are increased or decreased at the same time, or
      - if the increase of one value implies the decrease of the other, or
      - if they are independent / uncorrelated.
    - As an indicator of the ***degree*** of correlation, since the larger the value of ***r*** (in an absolute value) the stronger is the correlation between the two variables.

Forecasting Methods
Simple Linear Regression: Correlation Coefficient (2)

- How to estimate the ***r***:

$$r_{XY} = \frac{COV_{XY}}{\sqrt{COV_{YY} * COV_{XX}}} = \frac{COV_{XY}}{S_Y * S_X} \qquad |r_{XY}| \leq 1$$

where:

$$COV_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X}) * (X_i - \bar{X})}{n}$$

$$COV_{XX} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n} = Var_X = S_X^2 \ , \qquad COV_{YY} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n} = Var_Y = S_Y^2$$

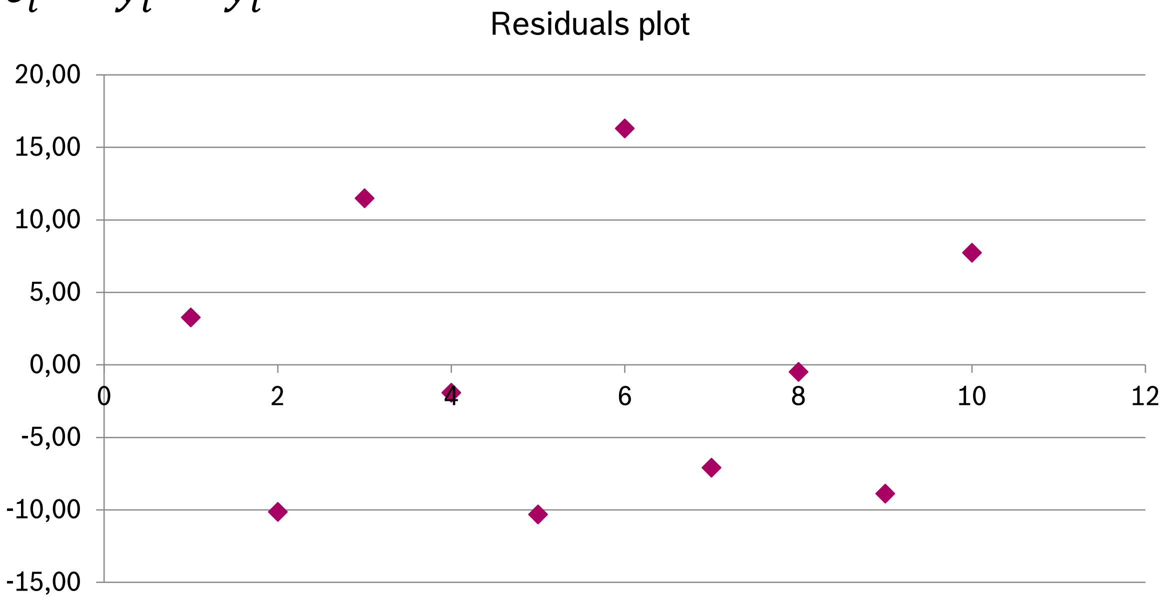
- It is obvious that ***correlation and regression are strongly linked***. The advantage of a regression model over correlation is that is asserts a predictive relationship between the two variables and quantifies this in a way that is useful for forecasting.

Forecasting Methods
Simple Linear Regression: Residual (1)

- ***Residual plot diagram***: A simple way for evaluating a simple linear regression model.
- Normally, the residuals errors must be randomly scattered without showing any systematic pattern.

$$e_i = y_i - \widehat{y_i}$$

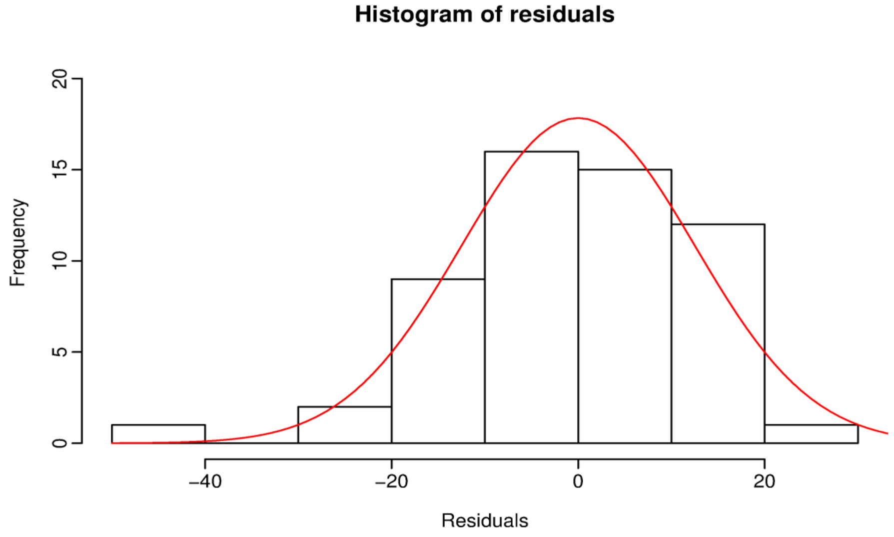
- A non-random pattern may indicate that:
  - a non-linear relationship may be required, or
  - some heteroscedasticity is present (residuals have non constant variance), or
  - there is some unexplained serial correlation.



BOSCH

Forecasting Methods
Simple Linear Regression: Residual (2)

- ***Residual Histogram***:
  - Another approach is to create a residuals histogram, in order to check if they are normally distributed.
  - This is not essential for forecasting, but it does make the calculation of prediction intervals much easier.
  - Example:
    - the residuals seem to be ***slightly negative skewed***.
    - This may be an outcome of an outlier.



Forecasting Methods
Simple Linear Regression: Evaluation(1)

- ***Coefficient of determination R^2***: A common way to evaluate how well a linear regression model fits the timeserie data.
  - The correlation of values resulting from the regression line equation (forecasts) and the actual values is denoted as ***R***.
  - In practice, this correlation is used in the quadratic form and it is always a ***positive number between 0 and 1***.
    - If the predictions are close to the actual values, R^2 is expected to be close to 1.
    - If the predictions are unrelated to the actual values, then $R^2 = 0$.

$$R^2 = \frac{\sum_{i=1}^{n}\left(\widehat{Y}_i - \bar{Y}\right)^2}{(Y_i - \bar{Y})^2} = r_{XY}{}^2$$

- A high R^2 does not always indicates a good forecasting model! There are no set rules of what a good R^2 value is and typical values of R^2 depend on the type of data used.
- Validating a model's out-of-sample forecasting performance is much better than measuring the in-sample R^2 value.

Forecasting Methods
Simple Linear Regression: Evaluation(2)

- ***Standard error of regression***: is the standard deviation of the residuals.

$$s_e = \sqrt{\frac{1}{N-2} * \sum_{i=1}^{N} e_i^2}$$

- The equation is slightly different from the usual standard deviation, where normally it is divided by ***N-1***. Here, an ***N-2*** is used, since two parameters (intercept and slope) are used for computing residuals. Also, the standard error can be highly subjective as it is scale dependent.
- The standard error is related to the size of the average error that the model produces. It can be compared with the sample mean of ***y*** or with the standard deviation of ***y,*** in order to gain some perspective on the accuracy of the model.

Forecasting Methods
Simple Linear Regression: Statistical Indexes (1)

- Considering the regression equation as a statistical model, some statistical indexes can be calculated that allows the assessment of:
  - the ***likelihood*** of future values of the dependent variable to differ from those specified (given quantity),
  - the ***reliability*** of the calculation of the regression line, and
  - the ***accuracy*** of the parameters α and β.
- Two indexes can be used:

1. Index ***F***: estimates the ***significance of the regression equation***, since it answers the question of whether there is a significant relationship between the variables X and Y.
2. Index ***t***: estimates the ***significance of the regression parameters***, especially if they are significantly different for hypothetical values.

Forecasting Methods
Simple Linear Regression: Statistical Indexes (2)

- Index **F**: Equations:

$$F = \frac{\frac{\Sigma\left(\hat{Y}_t - \bar{Y}\right)^2}{k-1}}{\frac{\Sigma\left(Y_t - \hat{Y}_t\right)^2}{n-k}} \qquad F = \frac{\frac{R^2}{k-1}}{\frac{1-R^2}{n-k}}$$

- Indexes **t**: Equations:

$$\widehat{\sigma_e} = \sqrt{\frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-k}} \qquad SE_a = \widehat{\sigma_e} \times \sqrt{\frac{1}{n} + \frac{\overline{X^2}}{\sum_{i=1}^{n}(\mathrm{X}_i - \bar{\mathrm{X}})^2}} \qquad SE_\beta = \widehat{\sigma_e} \times \sqrt{\frac{1}{\sum_{i=1}^{n}(\mathrm{X}_i - \bar{\mathrm{X}})^2}}$$

$$t_a = \frac{a - \alpha'}{\mathrm{SE}(a)} \qquad t_\beta = \frac{\beta - \beta'}{\mathrm{SE}(\beta)}$$

Forecasting Methods
Simple Linear Regression: How to use

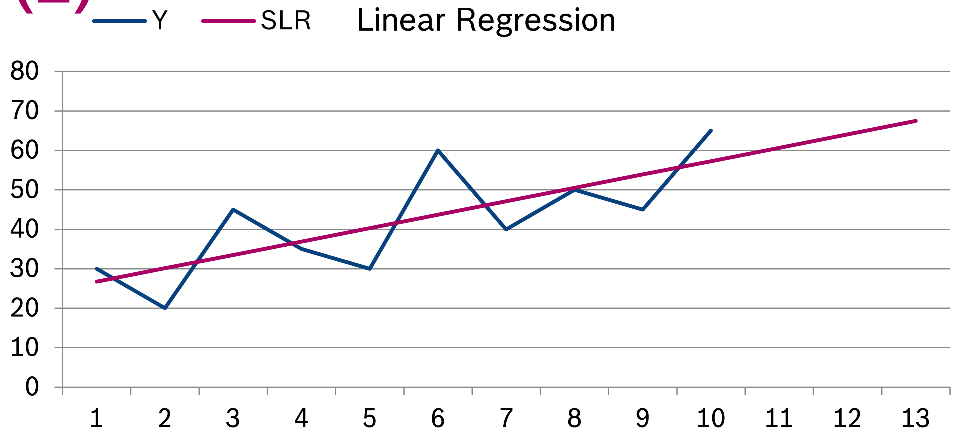
- Steps for performing an SLR:

1. Formulation of the Problem.
2. Selection of Financial and Other Related Indicators.
3. Initial testing use of regression.
4. Monitor the Simple Correlation Matrix.
5. Selecting the Regression Equation.
6. Estimation of R2 index.
7. Checking assumptions validity for Regression.
8. Preparing the model for estimation / forecasting.

Forecasting Methods

Simple Linear Regression: Example (1)

Data		Numerator			Denominator				
		A	B						
X	Y	x-mean(x)	y-mean(y)	A * B	(x-mean(x))^2	(Y-mean(Y))^2	(Yf-mean(Y))^2	(Y-Yf)^2	SLR
1	30	-4,5	-12	54	20,25	144	233,256	10,711	26,73
2	20	-3,5	-22	77	12,25	484	141,106	102,439	30,12
3	45	-2,5	3	-7,5	6,25	9	71,993	131,902	33,52
4	35	-1,5	-7	10,5	2,25	49	25,917	3,645	36,91
5	30	-0,5	-12	6	0,25	144	2,880	106,152	40,30
6	60	0,5	18	9	0,25	324	2,880	265,789	43,70
7	40	1,5	-2	-3	2,25	4	25,917	50,281	47,09
8	50	2,5	8	20	6,25	64	71,993	0,235	50,48
9	45	3,5	3	10,5	12,25	9	141,106	78,833	53,88
10	65	4,5	23	103,5	20,25	529	233,256	59,711	57,27
11									60,67
12									64,06
13									67,45
Average			SUM	280	82,5	1760	950,3030303	809,697	
X	5,5		Average						
Y	42								



$$R^2 = \frac{\sum_{i=1}^{n}\left(\hat{Y}_i - \bar{Y}\right)^2}{(Y_i - \bar{Y})^2} = r_{XY}{}^2 = 0{,}540$$

$$\widehat{\sigma_e} = \sqrt{\frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-k}} = 10{,}06$$

$$F = \frac{\dfrac{\Sigma\left(\hat{Y}_t - \bar{Y}\right)^2}{k-1}}{\dfrac{\Sigma\left(Y_t - \hat{Y}_t\right)^2}{n-k}} = 9{,}389$$

$$SE_a = \widehat{\sigma_e} \times \sqrt{\frac{1}{n} + \frac{\overline{X^2}}{\sum_{i=1}^{n}(\mathrm{X}_i - \bar{\mathrm{X}})^2}} = 6{,}873$$

$$SE_\beta = \widehat{\sigma_e} \times \sqrt{\frac{1}{\sum_{i=1}^{n}(\mathrm{X}_i - \bar{\mathrm{X}})^2}} = 1{,}108$$

$$\beta = \frac{\sum_{i=1}^{N}\left((Y_i - \bar{Y}) \times (X_i - \bar{X})\right)}{\sum_{i=1}^{N}(X_i - \bar{X})^2} = 3{,}39$$

$$COV_{XX} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n} = Var_X = S_X^2 = 8{,}25$$

$$COV_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X}) * (X_i - \bar{X})}{n} = 28$$

$$\alpha = \bar{Y} - \beta \times \bar{\mathrm{X}} = 23{,}33$$

$$COV_{YY} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n} = Var_Y = S_Y^2 = 176$$

$$r_{XY} = \frac{COV_{XY}}{\sqrt{COV_{YY} * COV_{XX}}} = \frac{COV_{XY}}{S_Y * S_X} = 0{,}735$$

$$t_a = \frac{a - \alpha'}{\mathrm{SE}(a)} = 3{,}395 \qquad t_\beta = \frac{\beta - \beta'}{\mathrm{SE}(\beta)} = 3{,}064$$

EXPONENTIAL SMOOTHING

Forecasting Methods
Exponential Smoothing (1)

- Forecasts are ***weighted averages of past observations***.
  - the weights decaying exponentially as the observations get older.
  - the more recent the observation the higher the associated weight is.
- This framework generates ***reliable forecasts quickly*** and for a wide spectrum of timeseries which is a great advantage and of major importance to applications in industry.
- They became one of the most popular methods of forecasting among businessmen, mainly because:
  - of their convenience,
  - the minimum requirement in computational time, and
  - the ***need of relatively few observations*** in order to estimate forecasts.

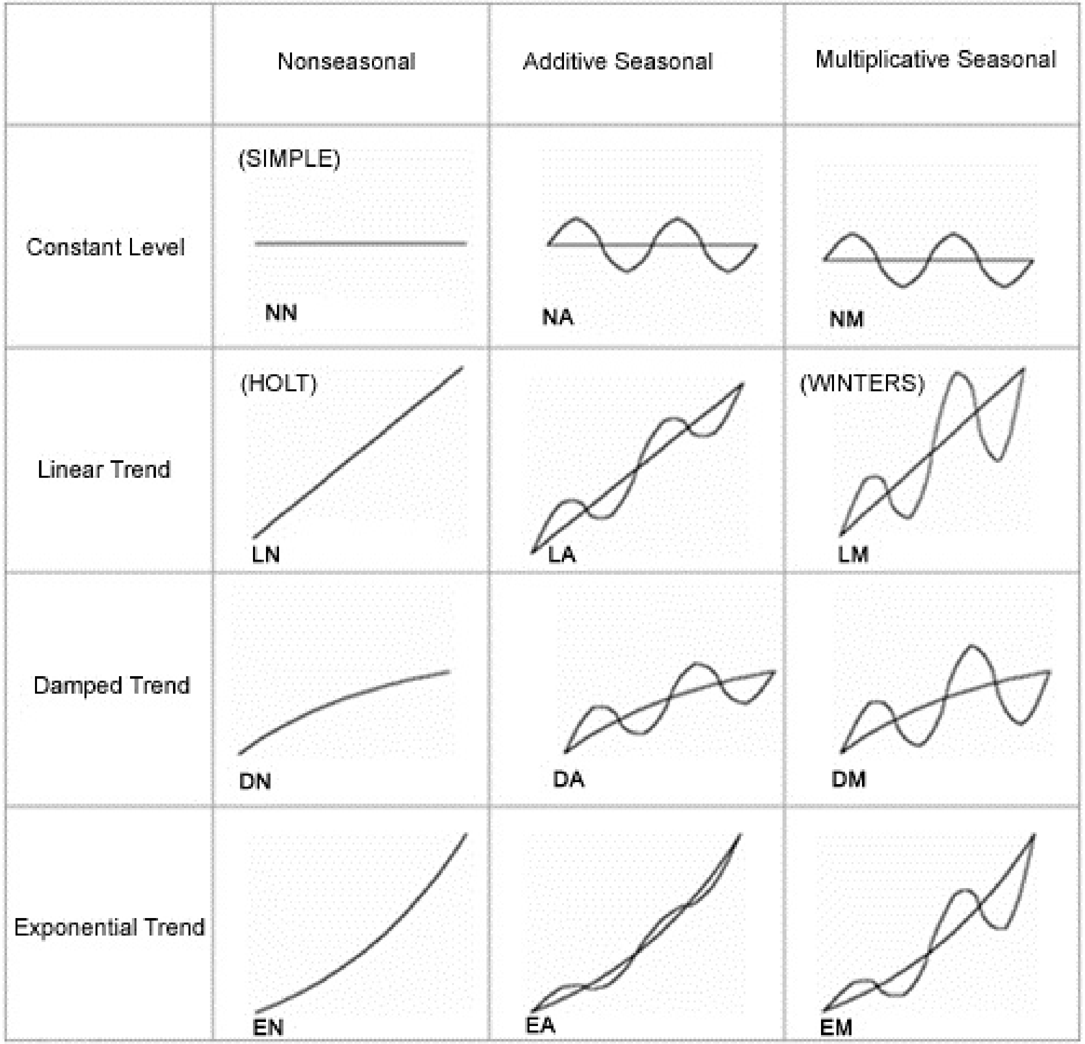
Forecasting Methods
Exponential Smoothing (2)

- Scope:
  - to ***isolate timeserie pattern*** from randomness.
- Usage:
  - Easy to use.
  - Require ***little historical data*** and computational time.
  - Widely used for ***short-term*** and ***medium-term*** forecasting.
- Performance:
  - They perform better on data that are ***stagnating*** or ***having small growth or reduction*** during time.
  - They are satisfactory more accurate than more sophisticated methods.

Forecasting Methods
Exponential Smoothing (3)

Smoothing types: according to trend pattern and seasonality pattern.

- ***Four trend patterns***:
  - Constant level
  - Linear trend
  - Damped trend
  - Exponential trend
- ***Three seasonality patterns***:
  - No seasonality
  - Additive seasonality
  - Multiplicative seasonality



BOSCH

Source: Forecast Pro

Forecasting Methods
Exponential Smoothing: Constant Level (1)

Simple Exponential Smoothing (SES):

- The simplest of the exponential smoothing methods.
- Forecast: $$F_{t+1} = S_t$$

where $S_t = S_{t-1} + (a \times e_t)$, $e_t = Y_t - F_t$

- Thus:

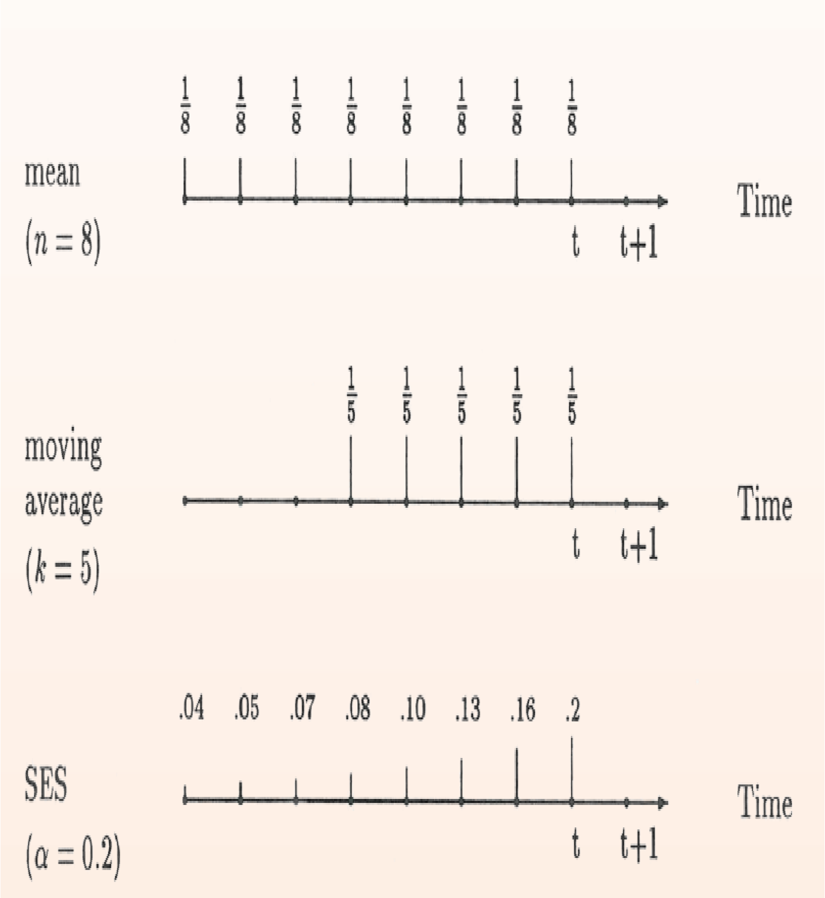
$$F_t = F_{t-1} + \big(a \times (Y_{t-1} - F_{t-1})\big) = (a \times Y_{t-1}) + \big((1-a) \times F_{t-1}\big)$$

$$F_{t+1} = (a \times Y_t) + (a \times (1-a) \times Y_{t-1}) + (a \times (1-a)^2 \times Y_{t-2}) + (a \times (1-a)^3 \times Y_{t-3}) + \cdots + \big(a \times (1-a)^{(t-1)} \times Y_1\big) + \big((1-a)^t \times F_1\big)$$

Forecasting Methods
Exponential Smoothing: Constant Level (2)

Simple Exponential Smoothing (SES):

- The weighting factors are decreasing exponentially
- The F_1 forecast is playing a significant role:
  - The smaller the value of α, the greater is the impact of F_1 to the other forecasts. Example, if $t=11$ then:
    - when a = 0.1, then the first period forecast is multiplied by 0.3138
    - when a = 0.5, then the first period forecast is multiplied by 0.0004
    - when a = 0.9, then the first period forecast is multiplied by 0.0000
  - The larger the number of periods t, the smaller is the impact of F_1 to the other forecasts. Example, if $a = 0.1$:
    - when t = 12, then the first period forecast is multiplied by 0.2824
    - when t = 24, then the first period forecast is multiplied by 0.0798



Internal | CR/AEU2 - Patelis | 2018-03-14

Exponential Smoothing: Constant Level (3)

Initialization:

- As F_1 we can use:
  - the mean value of ***all*** historical observations,
  - the mean value of ***the first four*** or five historical observations,
  - the ***first*** historical observation, or
  - the ***level*** for the model of linear regression.

Forecasting Methods
Exponential Smoothing: Constant Level (4)

- ***Smoothing parameter: Value between 0 and 1***.
  - if the value is 1, then all forecasts are equal with the first period forecast.
  - a large value (such as 0.8) is creating a small smoothing effect.
  - a small value (such as 0.2) is creating a large smoothing effect.
  - If the value is 0, then the model is the Naive model.
- The optimal value for the smoothing parameter (α) is usually defined from the ***minimization*** of the forecasting error (MSE, MAPE, other).
  - The parameter could be different if we aiming on minimization of MSE or MAPE.
  - a method for the optimization of α, is the estimation of MSE for a range of possible values (for example: 0.1, 0.2, ... 0.9) and the selection of the value that gives the smaller MSE,
  - another method is the usage of an non-linear optimization algorithm.

BOSCH

Forecasting Methods

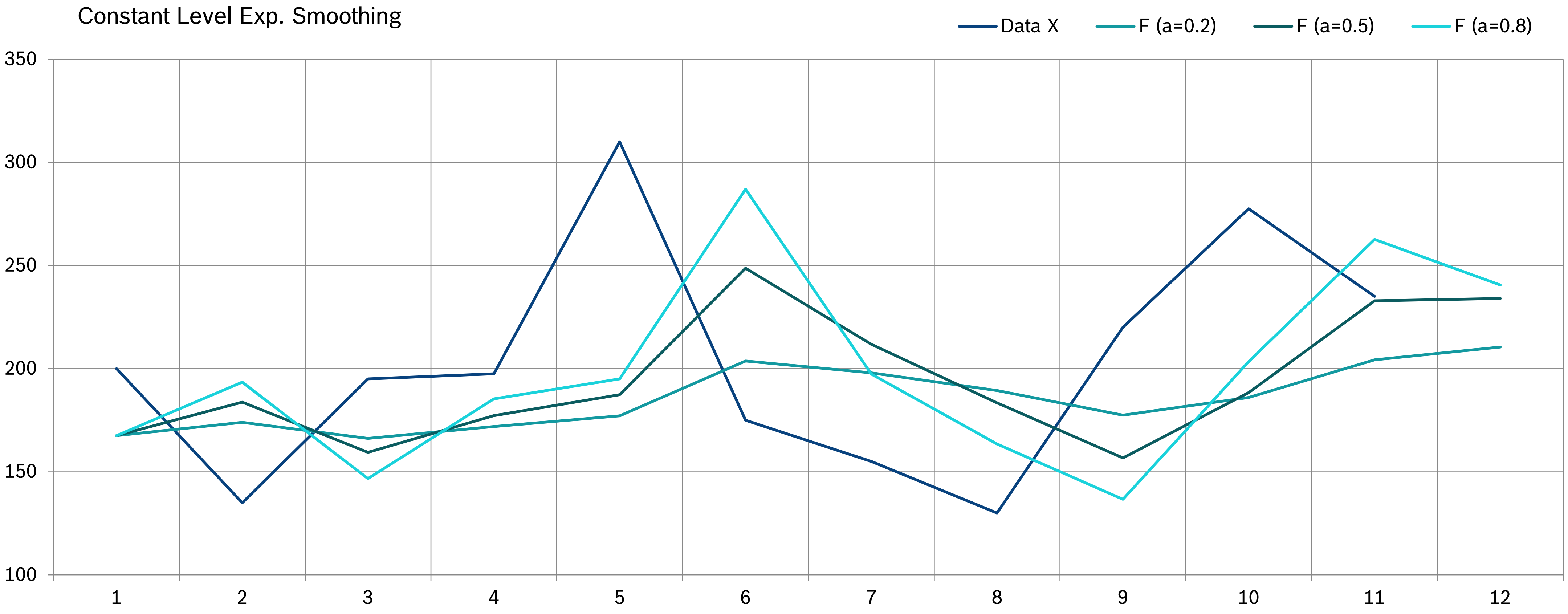
Exponential Smoothing: Constant Level – Example (5)

- Starting level:
  - Average of 1st and 2nd data point values.
- Smoothing parameter:
  - α = 0.2
  - α = 0.5
  - α = 0.8
- Evaluation with:
  - Mean error (ME).
  - Mean Absolute error (MAE).
  - Mean Absolute percentage error (MAPE).

S(0) =	167,5	α=0.2			α=0.5			α=0.8		
Time (t)	Data X	F (a=0.2)	Error e	Level S	F (a=0.5)	Error e	Level S	F (a=0.8)	Error e	Level S
1	200	167,5	32,5	174,0	167,5	32,5	183,8	167,5	32,5	193,5
2	135	174,0	-39,0	166,2	183,8	-48,8	159,4	193,5	-58,5	146,7
3	195	166,2	28,8	172,0	159,4	35,6	177,2	146,7	48,3	185,3
4	197,5	172,0	25,5	177,1	177,2	20,3	187,3	185,3	12,2	195,1
5	310	177,1	132,9	203,7	187,3	122,7	248,7	195,1	114,9	287,0
6	175	203,7	-28,7	197,9	248,7	-73,7	211,8	287,0	-112,0	197,4
7	155	197,9	-42,9	189,3	211,8	-56,8	183,4	197,4	-42,4	163,5
8	130	189,3	-59,3	177,5	183,4	-53,4	156,7	163,5	-33,5	136,7
9	220	177,5	42,5	186,0	156,7	63,3	188,4	136,7	83,3	203,3
10	277,5	186,0	91,5	204,3	188,4	89,1	232,9	203,3	74,2	262,7
11	235	204,3	30,7	210,4	232,9	2,1	234,0	262,7	-27,7	240,5
12		210,4			234,0			240,5		
Mean Error		19,51			12,08			8,30		
MAE		50,41			54,39			58,13		
MAPE		24,62%			27,46%			29,19%		

BOSCH

Forecasting Methods
Exponential Smoothing: Constant Level – Example (6)



BOSCH

Forecasting Methods
Exponential Smoothing: Linear Trend (1)

Linear Trend Exponential Smoothing (Holt):

- Forecast: $$F_{t+m} = S_t + m \times T_t$$

  where $S_t = S_{t-1} + T_{t-1} + (a \times e_t)$, $T_t = T_{t-1} + (\beta \times e_t)$, $e_t = Y_t - F_t$

- Parameters estimation:
  - The parameters a, β, must be estimated by minimizing (usually) the mean squared error (MSE).

BOSCH

Forecasting Methods
Exponential Smoothing: Linear Trend (2)

Initialization:

- Linear Regression must be performed, in order to estimate the linear regression equation.
- the initial level S_0 is equal to the **constant parameter** of the regression.
- the initial trend T_0 is equal to the **slope parameter** of the regression.

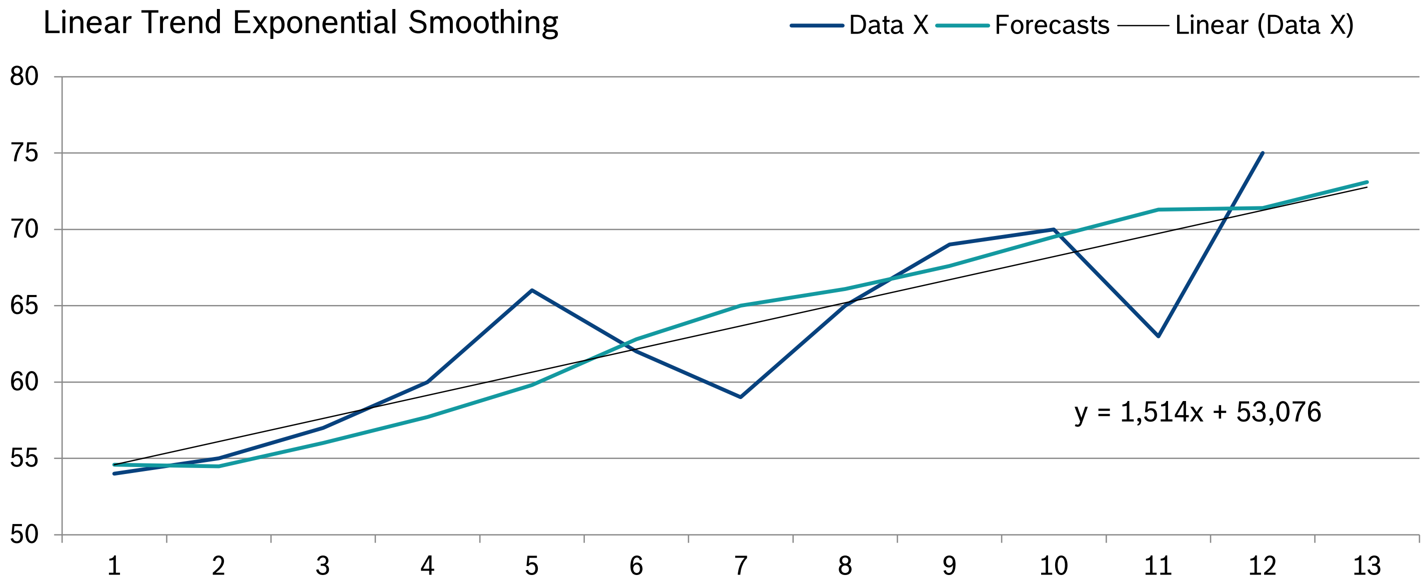
Alternative initialization:

- the initial level S_0:
  - can be equal to the first historical observation, or
  - can be equal to the mean value of the first N historical observations.
- the initial trend T_0:
  - can be equal to the difference between the second and first historical observations, or
  - can be equal to the difference between the Nth and the first historical observations, divided by (N-1)

Forecasting Methods
Exponential Smoothing: Linear Trend – Example (3)

- Starting level: Constant of regression, starting trend: Slope of regression.
- Smoothing parameter: $\alpha = 0.2$, $\beta = 0.1$
- Evaluation with: Mean Absolute percentage error.

S(0) =	53,076	T(0) =	1,514	α = 0.2, β = 0.1	
Time (t)	Data X	Forecasts	Error e	Level S	Trend T
1	54	54,6	-0,6	54,5	1,5
2	55	54,5	0,5	56,0	1,5
3	57	56,0	1,0	57,7	1,6
4	60	57,7	2,3	59,8	1,8
5	66	59,8	6,2	62,8	2,4
6	62	62,8	-0,8	65,0	2,3
7	59	65,0	-6,0	66,1	1,7
8	65	66,1	-1,1	67,6	1,6
9	69	67,6	1,4	69,5	1,7
10	70	69,5	0,5	71,3	1,8
11	63	71,3	-8,3	71,4	1,0
12	75	71,4	3,6	73,1	1,4
13		73,1		74,5	1,4
		74,5		75,9	1,4
		75,9		77,3	1,4
		MAPE:	4,24%		



Forecasting Methods

Exponential Smoothing: Damped Trend (1)

Damped Trend Exponential Smoothing:

- Forecast: $$F_{t+m} = S_t + \sum_{i=1}^{m} \varphi^i \times T_t$$

where $S_t = S_{t-1} + (\varphi \times T_{t-1}) + (a \times e_t)$, $T_t = (\varphi \times T_{t-1}) + (\beta \times e_t)$, $e_t = Y_t - F_t$

- Parameters estimation:
  - The parameters a, β, φ must be estimated by minimizing (usually) the mean squared error (MSE).

$$0 < \alpha < 1 \quad and \quad 0 < \beta < \alpha$$

 Internal | CR/AEU2 - Patelis | 2018-03-14

Forecasting Methods
Exponential Smoothing: Damped Trend (2)

Initialization:

- Linear Regression must be performed, in order to estimate the linear regression equation.
- the initial level S_0 is equal to the ***constant parameter*** of the regression.
- the initial trend T_0 is equal to the ***slope parameter*** of the regression.

Alternative initialization:

- the initial level S_0:
  - can be equal to the first historical observation, or
  - can be equal to the mean value of the first N historical observations.
- the initial trend T_0:
  - can be equal to the difference between the second and first historical observations, or
  - can be equal to the difference between the Nth and the first historical observations, divided by (N-1)

Forecasting Methods
Exponential Smoothing: Damped Trend (3)

***Initialization*:**

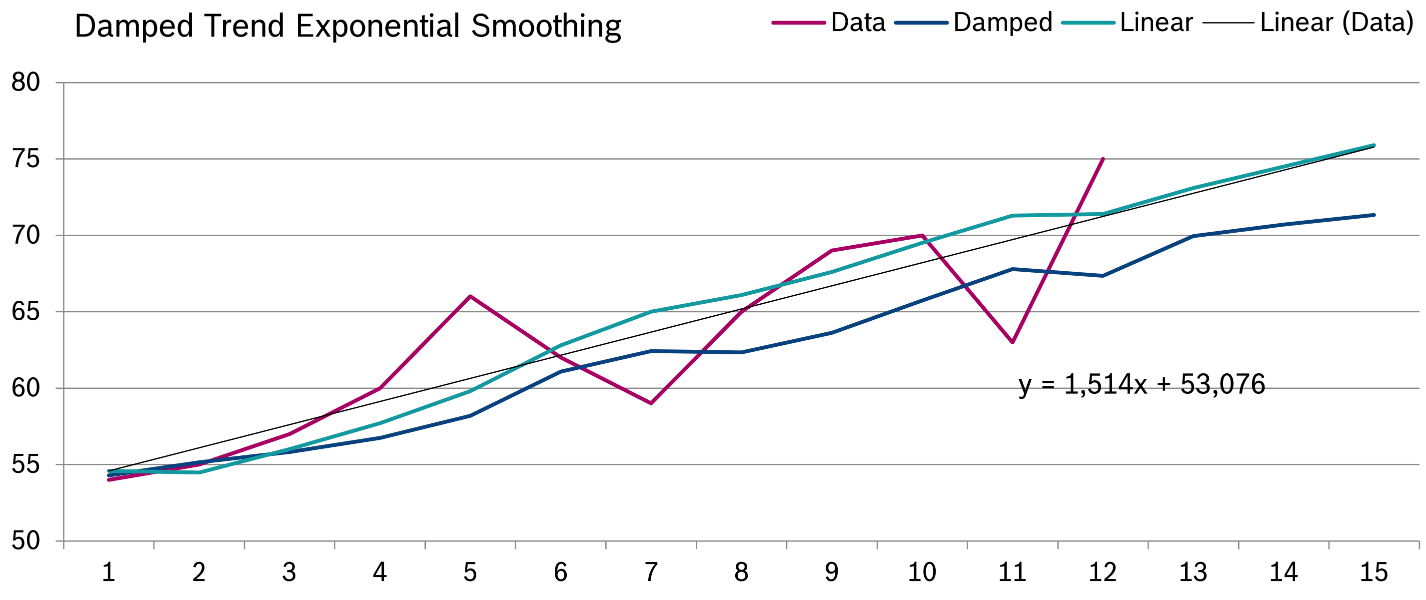
- Damping factor φ:
  - $\varphi = 0$, for the Constant Level
  - $\varphi < 1$, for the Damped Trend
  - $\varphi = 1$, for the Linear Trend
  - $\varphi > 1$, for the Exponential Trend

Forecasting Methods

Exponential Smoothing: Damped – Example (4)

- Starting level: Constant of regression, starting trend: Slope of regression.
- Smoothing parameter: $\alpha = 0.2$, $\beta = 0.1$, $\varphi = 0.8$
- Evaluation with: Mean absolute percentage error (MAPE).

S(0) =	53,076	T(0) =	1,514	α = 0.2, β = 0.1, φ = 0.8	
Time (t)	Data X	Forecasts	Error e	Level S	Trend T
1	54	54,3	-0,3	54,2	1,2
2	55	55,2	-0,2	55,1	0,9
3	57	55,8	1,2	56,1	0,8
4	60	56,7	3,3	57,4	1,0
5	66	58,2	7,8	59,8	1,6
6	62	61,1	0,9	61,3	1,4
7	59	62,4	-3,4	61,7	0,8
8	65	62,3	2,7	62,9	0,9
9	69	63,6	5,4	64,7	1,3
10	70	65,7	4,3	66,6	1,5
11	63	67,8	-4,8	66,8	0,7
12	75	67,4	7,6	68,9	1,3
13		69,9		69,9	1,0
		70,7		70,7	0,8
		71,3		71,3	0,6
		MAPE:	5,29%		



Forecasting Methods
Exponential Smoothing: What about seasonality? (1)

If a timeserie has seasonality pattern, then a ***Seasonal Index must be added*** into the non-seasonal models, for every period of the year.

- Forecast: $$F_{t+m} = S_t \times I_{t-p+m}$$

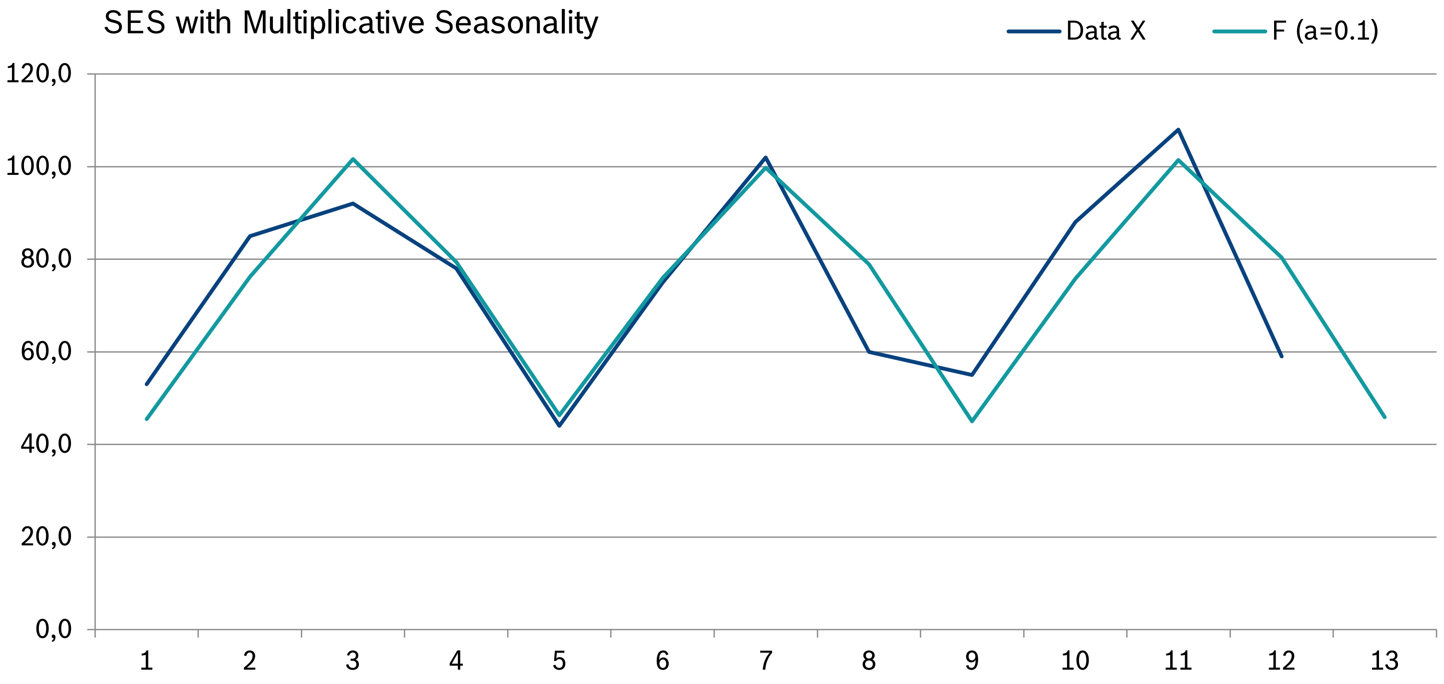
where $S_t = S_{t-1} + (\frac{(a \times e_t)}{I_{t-p}})$, $e_t = Y_t - F_t$, $I_t = I_{t-p} + (\frac{(\gamma \times e_t)}{S_t})$

For ***initialization*** of the model:

- Estimate the initial seasonal indexes, by using the decomposition model
- Estimate S_0 and T_0
- Estimate smoothing parameters, by using the linear method for best smoothing parameters

Forecasting Methods
Exponential Smoothing: What about seasonality? (2)

S(0) = 74,3		α=0.1 , γ = 0.01			
Time (t)	Data X	F (a=0.1)	Error e	Level S	Indexes I
-3					0,6122
-2					1,0086
-1					1,3303
0					1,0489
1	53,0	45,5	7,5	75,5	0,6132
2	85,0	76,2	8,8	76,4	1,0098
3	92,0	101,6	-9,6	75,7	1,3290
4	78,0	79,4	-1,4	75,5	1,0487
5	44,0	46,3	-2,3	75,2	0,6129
6	75,0	75,9	-0,9	75,1	1,0096
7	102,0	99,8	2,2	75,2	1,3293
8	60,0	78,9	-18,9	73,4	1,0461
9	55,0	45,0	10,0	75,1	0,6142
10	88,0	75,8	12,2	76,3	1,0112
11	108,0	101,4	6,6	76,8	1,3302
12	59,0	80,3	-21,3	74,7	1,0433
13		45,9			



Internal | CR/AEU2 - Patelis | 2018-03-14

Forecasting Methods
Exponential Smoothing – Model Selection (1)

Not all exponential smoothing methods can be used in all data cases. In general:

- Constant Level:
  - for 1 period forecasting, and
  - for timeseries with high noise or randomness.
- Linear Trend:
  - for stable increase in the future.
- Exponential Trend:
  - for exponential increase in the future (for example, in the first days of a product life-cycle).
  - they are over-optimistic for long-term forecasting.
- Damped Trend:
  - For mid-term forecasting

Forecasting Methods
Exponential Smoothing – Model Selection (2)

Model selection rule from Gardner & McKenzie (1998):

- Estimate Variance for:

A. Initial data

B. 1st level differences of A

C. 2nd level differences of A

D. Seasonal 1st level differences

E. 1st level differences of D

F. 2nd level differences of D

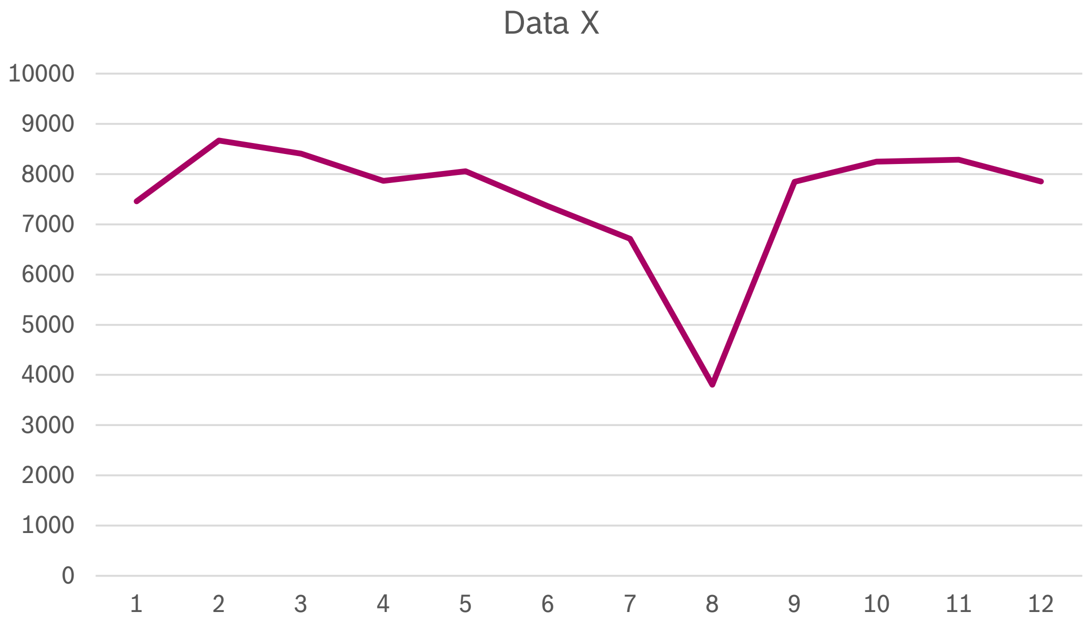
$$Variance = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n}$$

Case	Timeserie	Suggested Smoothing Model
A	Initial Data	SES
B	1st difference	Damped
C	2nd difference	Holt
D	Seasonal 1st difference	Seasonal SES
E	1st difference of D	Seasonal Damped
F	2nd difference of D	Seasonal Holt

Forecasting Methods
Exponential Smoothing – Model Selection (3)

	A	B	C	D	E	F
Time (t)	Data X	1st Differences	2nd Differences	Seasonal Differences	1st Differences of D	2nd Differences of D
1	7460					
2	8670	1210				
3	8410	-260	-1470			
4	7865	-545	-285			
5	8055	190	735	595		
6	7360	-695	-885	-1310	-1905	
7	6715	-645	50	-1695	-385	1520
8	3805	-2910	-2265	-4060	-2365	-1980
9	7845	4040	6950	-210	3850	6215
10	8250	405	-3635	890	1100	-2750
11	8285	35	-370	1570	680	-420
12	7855	-430	-465	4050	2480	1800
Variance	**1.527.077**	2.324.128	5.858.183	3.427.228	2.525.874	4.385.935

➢ Model to use: SES



BOSCH

8. METHOD SELECTION

"It is often said there are two types of forecasts… lucky or wrong!"

Control Magazine, Institute of Operation Management

BEFORE SELECTING THE METHOD

Method Selection
Before Selecting the method (1)

- ***Problem definition***: Sometimes the most difficult part.
  - Break the question down into ***smaller components***.
  - Identify the ***known*** and the unknown.
  - Look closely at all of your ***assumptions***.
  - Consider the ***outside view***, frame the problem not as a unique thing but as a variant in a wider class of phenomena.
  - Then, look at what it is that ***makes it unique***; look at how your opinions on it are the same or different from other people's viewpoints.
  - Taking in all this information with your dragonfly eyes, construct a unified vision of it; describe your judgement about it as clearly and concisely as you can, being as granular as you can be.

Method Selection
Before Selecting the method (2)

- ***Data gathering***:
  - We need to collect:
    - Statistical (mathematical) data
    - Judgmental data, experience from experts in the market.
  - Questions to be addressed:
    - How the data will be collected?
    - How they will be maintained?

Method Selection
Before Selecting the method (3)

- ***Data preparation***:
  - Scope: to have a first feeling
    - Do patterns exist? What information can we gain from the raw historical data?
    - Does the data have trend?
    - Does the data have seasonality (Calendar effects, Institutional influences, Weather, Expectations, etc.)
    - Do outliers exist?
  - This analysis will guide us:
    - in the proper decomposition methods and then
    - to the proper estimation method for sufficient results.

Method Selection
Before Selecting the method (4)

- ***Data analysis***:
  - Adjustments:
    - Missing values
    - Zero values
    - Working and trading days
    - Trend-Cycle graph with usage of moving averages
  - Data Analysis (statistics, mean, deviation, min, max, seasonal indices, growth rates, etc.)
  - Special Events & Actions (SEA)
    - Outliers
    - Level Shift
    - Estimation and correction of identified SEA impact.

FINDING THE PROPER METHOD

Method Selection
Find the proper method

- Selection of the proper method depends on many factors, such as:
  - The ***data type*** and the data values.
  - The timeserie ***characteristics***.
  - ***Data volume***.
  - The forecasting ***horizon***.
  - The selection of the ***explanatory variables***.
  - Our ***scope***:
    - Do we want to forecast future values?
    - Do we want to understand the past values?
    - Control vs Planning?
  - The method ***accuracy***.
  - ***Cost***.

Method Selection
Find the proper method: Data type

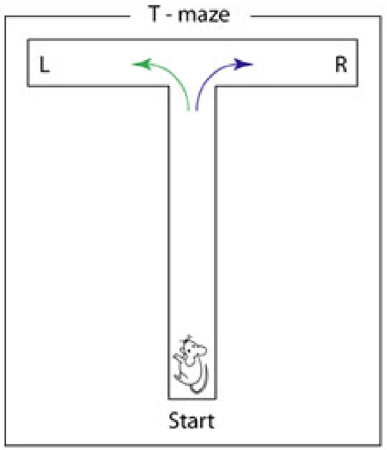
- Data type: Yearly, Quarterly, Monthly, Weekly, Daily, ...
  - ***Yearly data***: They include less randomness.
    - Simple methods can be used.
  - ***Quarterly / Monthly data***: Randomness is limited, trend and cycle exist.
    - More advanced methods can be used
  - ***Daily data***: Randomness is dominating, trend is not significant or can not be observed.
    - The use of smoothing methods is proposed.

➢ In general, the greater the level of detail (and frequency) that is required, the greater the need for an automated forecasting procedure.

Method Selection
Find the proper method: Timeserie Characteristics

- ***Pattern of Data*** (Seasonality, Trend, Cycle, Randomness…)
  - Seasonality:
    - Decomposition method is the most simple to use (but: almost all method can estimate seasonality).
  - Randomness / TrendCycle comparison:
    - If ***Randomness > TrendCycle*** → Short-term forecasting, by using exponential smoothing
    - If ***Randomness < TrendCycle***:
      - Little randomness → smoothing models, or ARIMA models
      - Almost no randomness → Holt method.
    - In general, ***in cases of great randomness we selected simple methods***.

Method Selection
An Experiment



- A rat was placed in a T-maze. At intervals, food was placed at L or R.
  - The sequence between left and right was random, but on 60% of occasions food appeared at the left side.
- The rat soon learned that the left side was more likely to deliver food and always went to that side so its prediction was ***correct 60% of the time***.
- When the sequence of left and right placements was showed to Yale University students, they started searching for patterns.
  - Perhaps after 2 right placements a left placement was more likely, or
  - perhaps after 3 lefts a right would nearly always follow.
  - Of course, these perceived patterns were false. As a result the students only predicted the ***correct side 52% of the time***.
- ➢ *The rat was the more accurate forecaster.*

Method Selection
Find the proper method: Data Volume

- Since different methods are based on ***historical information***, we must consider:
  - the quantity of data at hand,
  - the appropriateness of the data, and
  - what it would cost to gather additional data.

- ***Best strategy***:
  - often is most effective to ***start with a simple forecasting method*** that does not require many data, and then,
  - as we build experience and gather more data, increasingly ***sophisticated methods*** can be adopted.

Method Selection
Find the proper method: Horizon

- Method selection depends on the number of forecasts (horizon)

- Generally, the bigger the seasonality → greater number of forecasts
  - Daily timeseries request more forests that the monthly or yearly forecasts

- In practice:
  - When the number of forecasts increases (and the seasonality is shorter), it is better to use ***simple and automated forecasting methods.***

Method Selection
Find the proper method: Scope

- ***Example***: Forecasting sales
  - Variables: Advertisement, product price
  - An exponential smoothing shows a decreased trend of sales by 10%
  - A regression method shows a 13% increase of sales if we increase advertisement by 5%.

➢ Depending on our scope, we select a smoothing or a regression method.

BOSCH

Method Selection
Find the proper method: Control vs Planning

- ***Control***:
  - management by exception is the general procedure.
  - What is needed is some way to determine, as early as possible, when a ***process is out of control*** (shift of basic pattern).
  - A forecasting method in such situations should be ***able to recognise changes*** in patterns or relationships at an ***early stage***.
- ***Planning***:
  - It is generally assumed that ***existing patterns will continue*** in the future.
  - the major emphasis is on ***identifying*** those patterns and ***extrapolating*** them into the future.

Method Selection
Find the proper method: Accuracy

- Some factors are increasing the size of errors:
  - ***Forecast wrong variable***:
    - We need to forecast a product demand, but we lack such data. Therefore, we are estimating forecasts for other variables, such as orders, production, load, billing, etc.
  - ***Change of pattern*** or ***non-static relations***:
    - Statistical methods assume that the pattern and the relations are static, something not frequent in the real world.
  - ***Selection of wrong method***:
    - A statistical method may minimizes the error for forecasting 1 period ahead. This does not imply that this method is the best for forecasting with larger horizons.
- Closely related to the level of detailed required in a forecast is the ***needed accuracy***.
  - For some decision situations, a +-10% may be sufficient.
  - in others, a variation of +-5% could spell disaster.

Method Selection
Find the proper method: Cost

- ***Cost***:
  - Three elements of cost are involved in the application of a forecasting procedure:
    - development,
    - data preparation, and
    - actual operation.
  - There are also opportunity costs in terms of other techniques that might have been applied.

- The variation in costs obviously affects the attractiveness of different methods for different situations.
- Keep in mind the relative change ***between cost and accuracy***!
  - We can select a more straightforward and less expensive forecasting method if it achieves the required level of accuracy.

Method Selection
Find the proper method

- What if a selection decision as to the "best" method is unclear?
  - it has been shown to be beneficial to hedge by using ***more that one*** forecasting method or forecaster, and
  - then ***combine the predictions***.

- This has proved to be an extremely effective way of:
  - increasing forecasting accuracy, and
  - decreasing the variance in errors.

➢ Thus, when in doubt we should ***combine*** multiple forecasts that come from a ***variety of independent*** sources.

EXPERTS ADJUSTMENTS TO STATISTICAL FORECASTS

Method Selection
Experts Adjustments: (1)

- Characteristics of experts:
  - They rely on ***automated mental processes*** that allow them quickly to recognise patterns.
  - They have a ***wide-ranging knowledge*** of their field, including the latest development.
  - They can ***discriminate*** between information that is ***relevant*** and that which is irrelevant.
  - They can get to the heart of the matter ***more quickly*** that others.
  - They can ***make sense*** out of chaos.

- "Expert: one who knows more and more about less and less" (*Nicholas Murray Butler, Columbia University president*).

Method Selection
Experts Adjustments (2)

- ***General Rule***:
  - IF:
    - Computer methods and experts adjustments are ***working in tandem***.
    - AND
    - They are ***sticking to roles*** where each performs best.
  - THEN:
    - Can often produce ***more reliable forecasts*** than either working alone.

Method Selection
Experts Adjustments: (3)

- ***Lessons learned*** from the literature:
  - Model-based forecasts for Stock keeping Unit (SKU)-level data are frequently adjusted by experts (*Franses and Legerstee, 2009*).
  - Experts adjust model-based ***forecasts upwards more*** than downwards (Franses and Legerstee, 2009).
  - ***Positive*** adjustments were ***far less effective*** than negative ones (Fildes et al., 2009).
  - The optimism bias leads to positive adjustments having ***larger errors*** than negative ones (Trapero et al., 2013).
  - Expert forecasts often significantly differ from model forecasts (Franses and Legerstee, 2010).
  - However small adjustments are also the case: ownership (Fildes et al., 2009).

Method Selection
Experts Adjustments (4)

- ***Lessons learned*** from the literature:
  - Experts ***can reduce forecasting error*** when adjustments size is not too large (Trapero et al., 2013).
  - Where the forecasters' principal motivation is towards improved accuracy, they can ***add substantially*** to forecast accuracy (Fildes et al., 2009).
  - ***Combination leads to improvements*** (Fildes et al., 2009; Franses and Legerstee, 2011).
  - Small improvements to statistical forecasts can translate to significant gains in terms of utility (Synetos et al., 2009).
  - Big losses in judgmental adjustments are most probable to be followed by another big loss as a result of a large adjustment (Petropoulos et al., 2016).

UPDATE FORECASTS

“If you have to forecast, forecast often”

Edgar R. Fiedler

BOSCH

Method Selection
Update Forecasts (1)

- Predictions should be updated any time there is ***additional information***.
  - These updated forecasts tend to be more accurate, because the forecaster who is updating more often is likely to be better informed.
- It is tricky to update a forecast - one can underreact and one can overreact.
  - Often, when we are confronted with new information, we want to ***stick to our beliefs*** regardless of the new evidence.
    - People's opinions about things can actually be more about their own self-identity than any other thing.
    - Also, the more people who have an emotional investment in something the harder it is to admit one was wrong.
  - Another challenge: once people publicly take a stance on something, it's hard to get them to change their opinion. But you need to be able to ***change your opinion when the facts change***.

Method Selection
Update Forecasts (2)

- It is also tricky to ***distinguish*** important from irrelevant ***information***.
  - sometimes people think something is important but it's not, and irrelevant information can confuse and trigger biases.
  - when one doesn't feel committed to the results, they can overreact.
  - when they are really attached, they can underreact.

- The trick is to update a forecast frequently, but, in most cases, make only ***small adjustments***.
- Sometimes, of course, we need to make a ***dramatic change***. If we are really far off target, incremental change won't cut it.

Method Selection
Update Forecasts (3)

- ***Do not forget*:**
  - Most people will only recall the rare forecasts that are regarded as spectacular failures.
  - Get a thousand forecasts right but one very wrong and you'll probably be judged by that single forecast.

➢ *In forecasting, reputations are hard won, but very easily lost.*

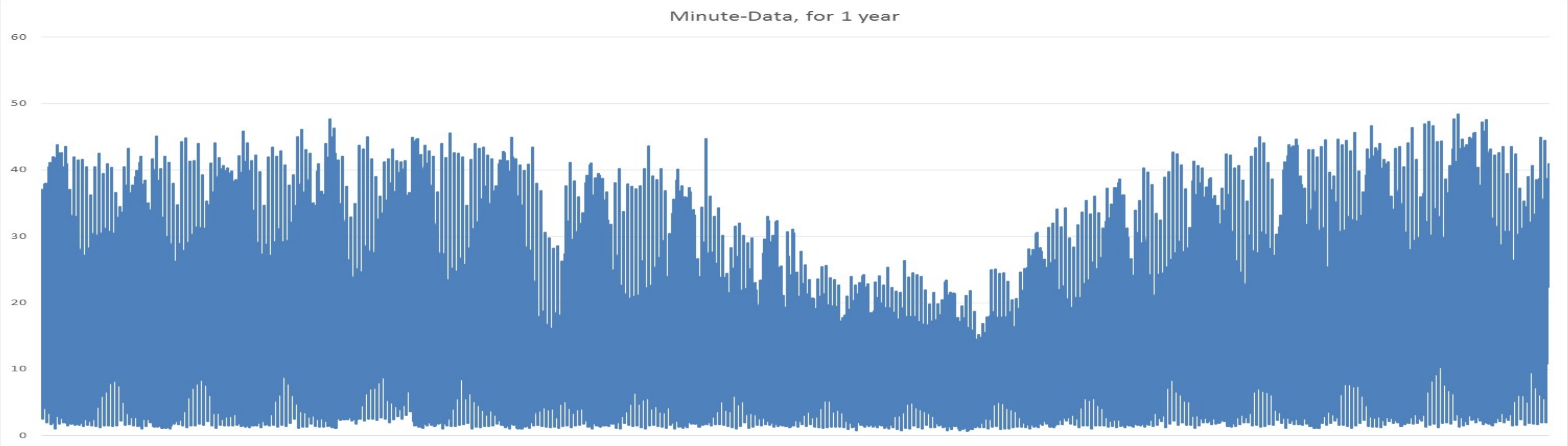
9. INTERESTING USE-CASES

Weather forecast tonight: Dark

FORECASTING WITH MINUTE DATA

Interesting Use-Cases
Forecasting with minute data (1)

- Data: a timeserie with min-by-min data, for 5 years, a total of 2.6 million data points.
- Scope: to forecast values for the next week.
- Problem: How we work with this data?

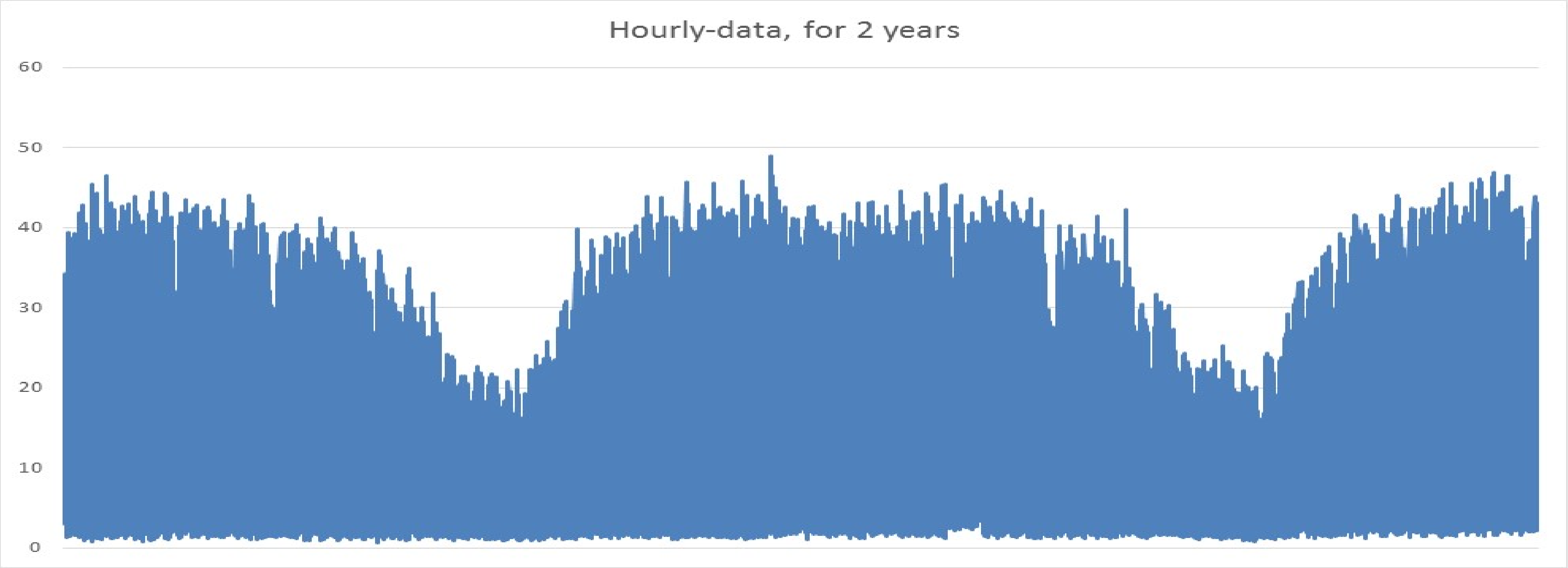


 Internal | CR/AEU2 - Patelis | 2018-03-14

BOSCH

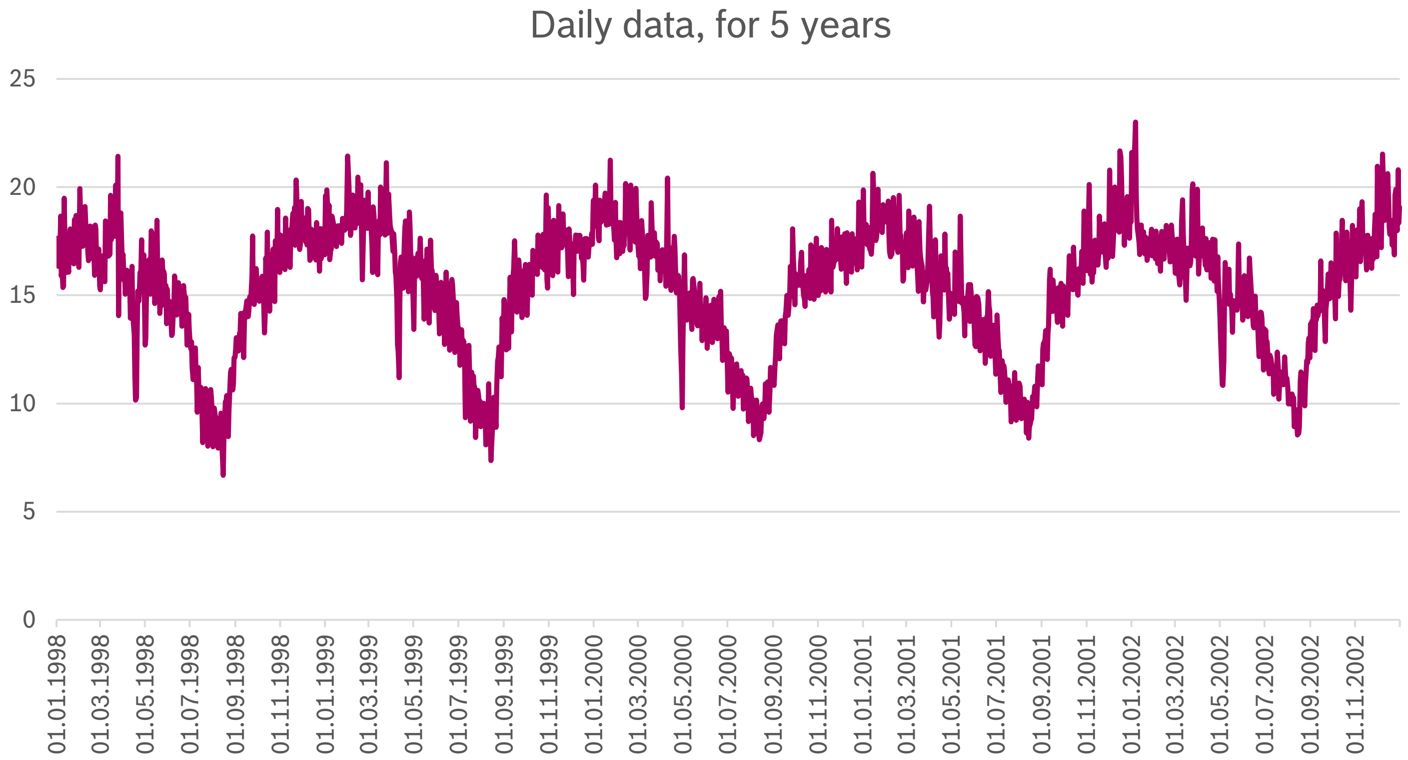
Interesting Use-Cases
Forecasting with minute data (2)

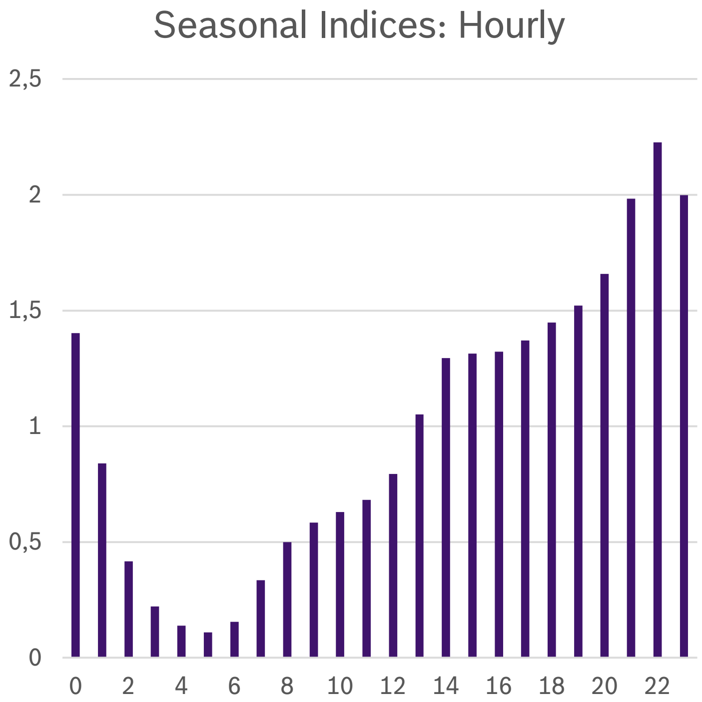
- ***Step 1***: Aggregate minute-data to Hourly-data.



Interesting Use-Cases
Forecasting with minute data (3)

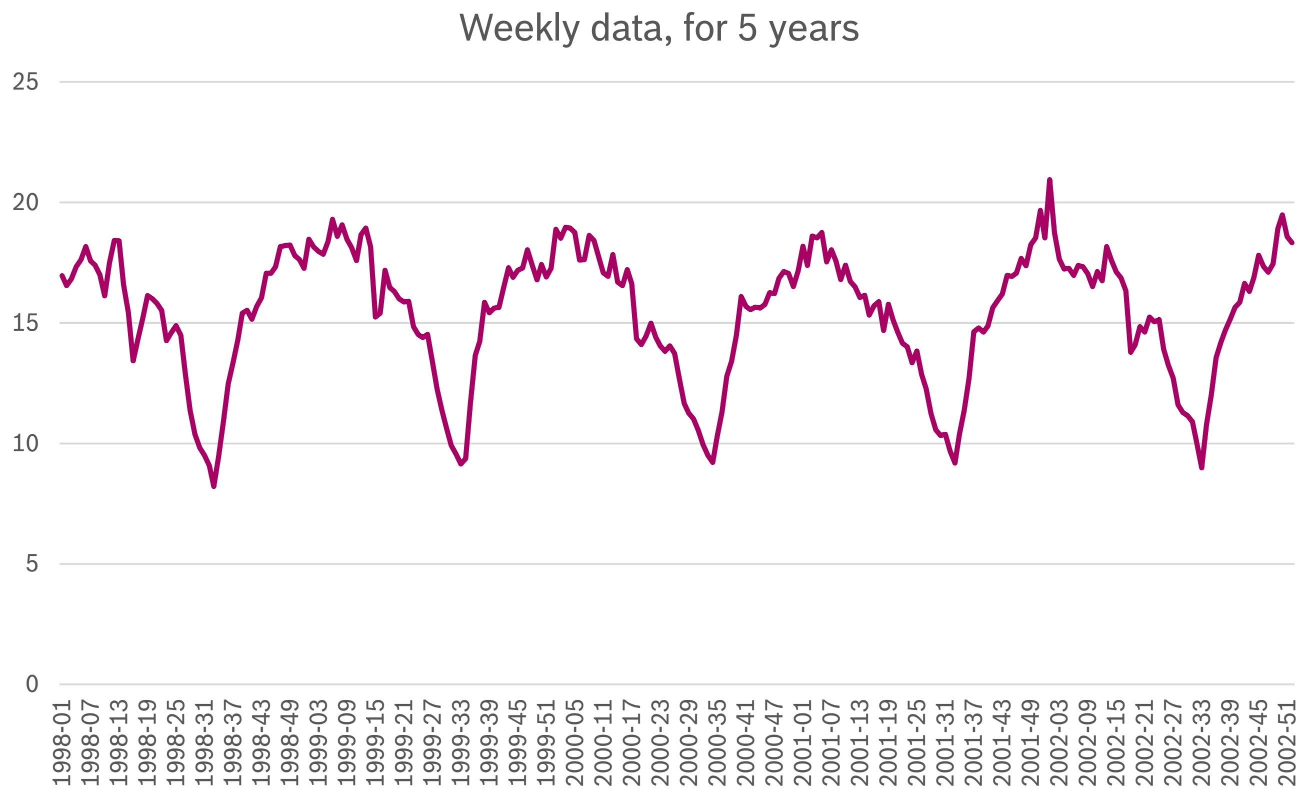
▶ ***Step 2***: Aggregate hourly-data to daily-data, and estimate Hourly seasonal indices.

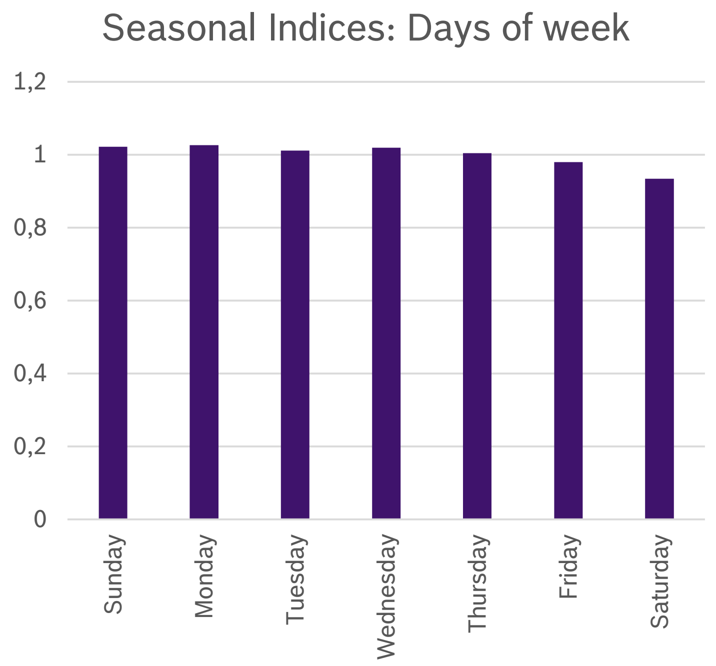




Interesting Use-Cases
Forecasting with minute data (4)

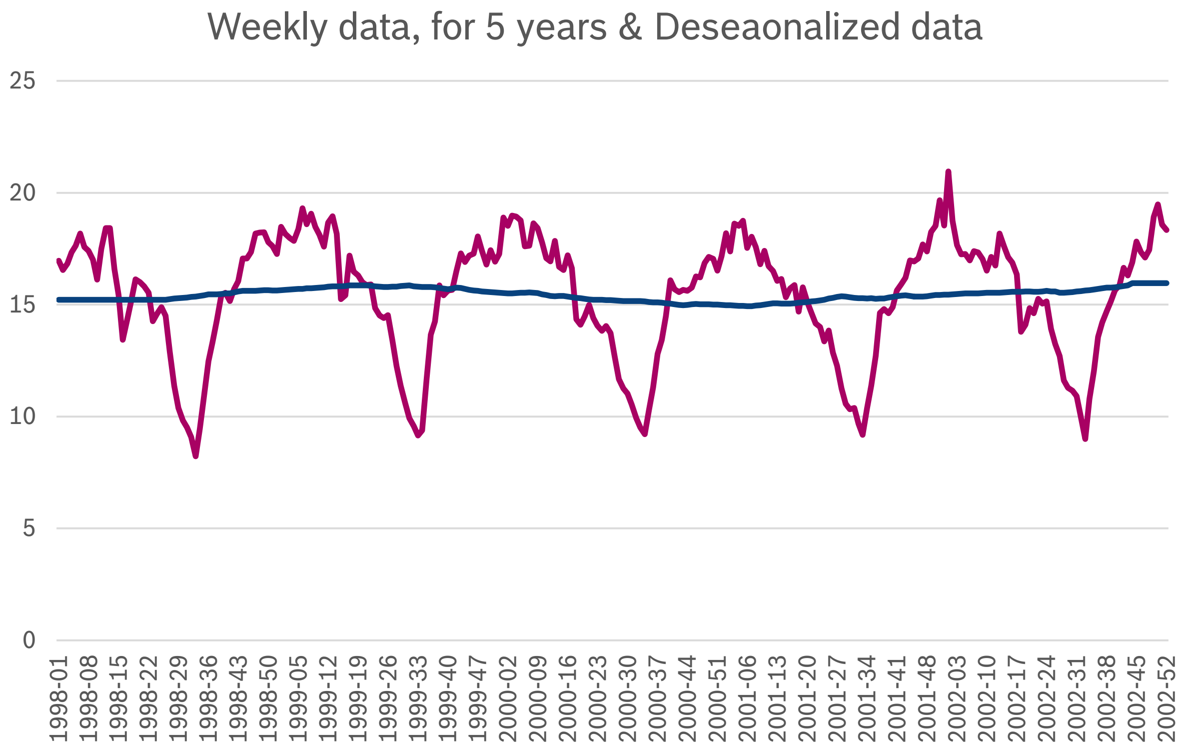
- ***Step 3***: Aggregate daily-data to weekly-data, and estimate Daily seasonal indices.

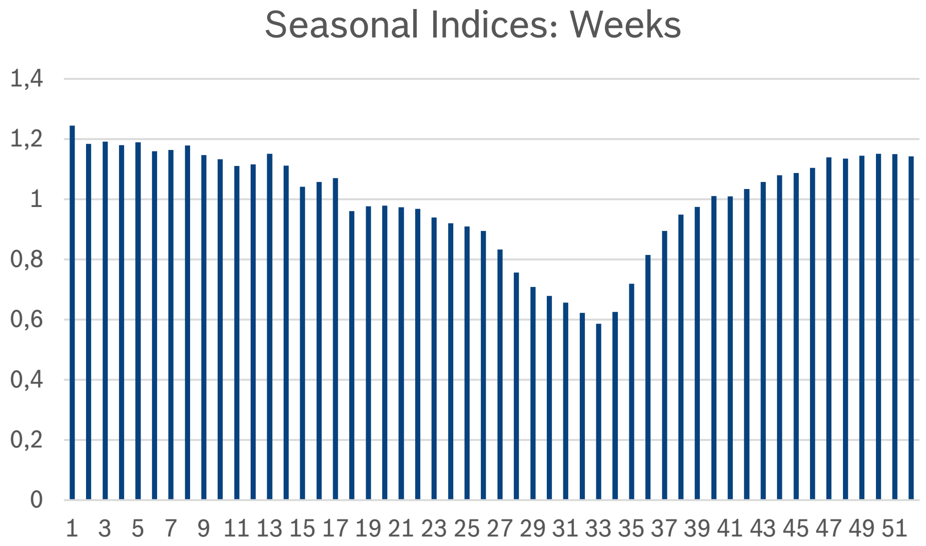




Interesting Use-Cases
Forecasting with minute data (5)

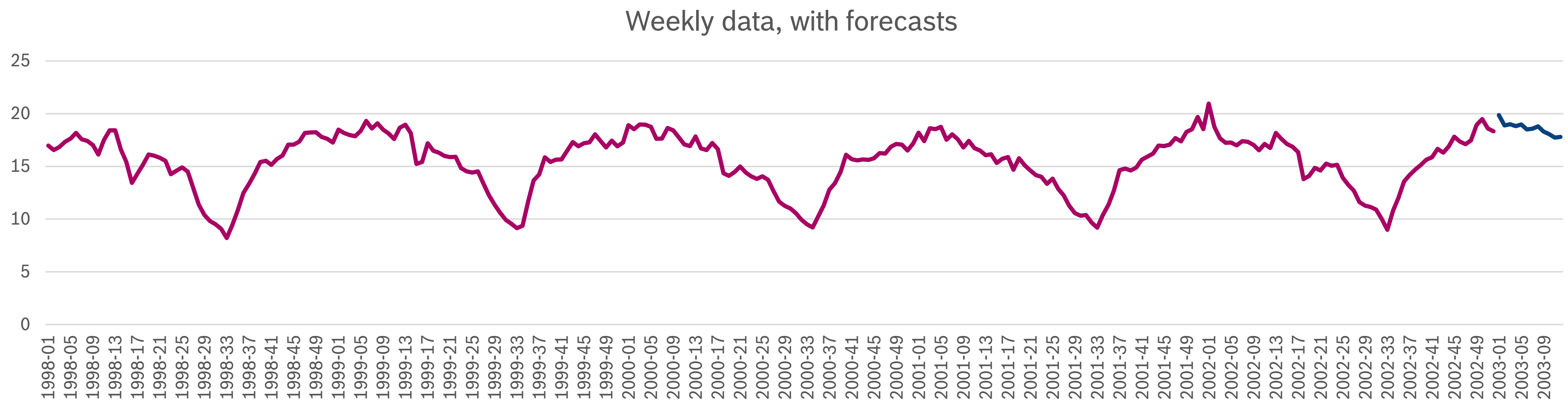
- ***Step 4***: Remove also weekly seasonality, and estimate Weekly seasonal indices.





Interesting Use-Cases
Forecasting with minute data (6)

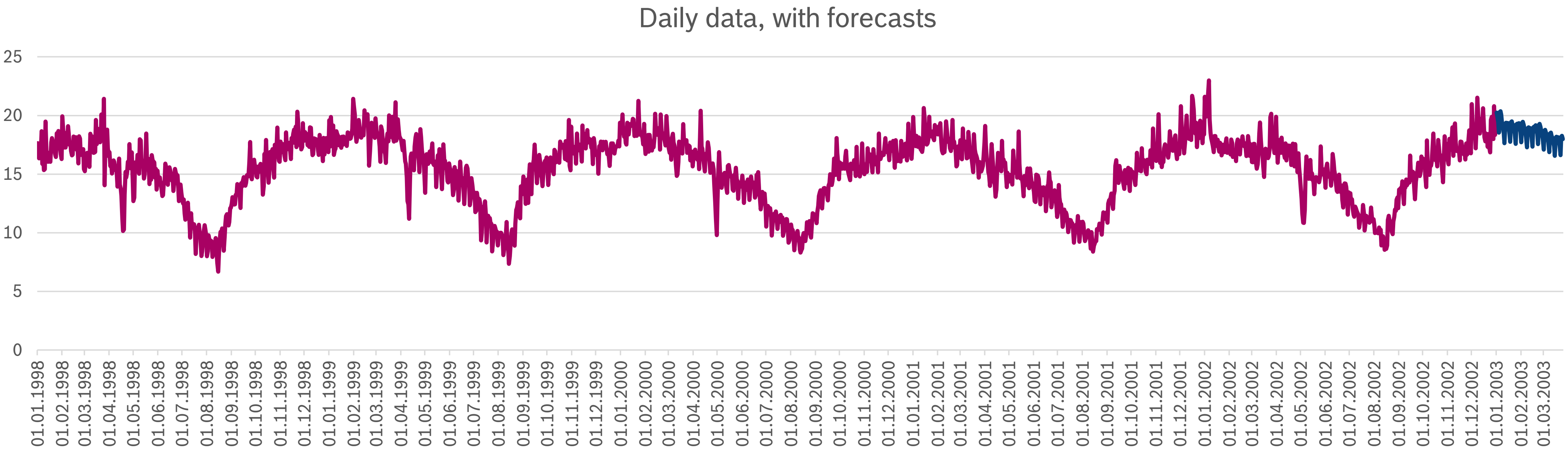
- ***Step 5***:
  - Estimate forecast(s) for next week(s).
  - Re-seasonalize the forecasts, by using the seasonal indexes for weeks.



BOSCH

Interesting Use-Cases
Forecasting with minute data (7)

- ***Step 6***:
  - Disaggregate forecasts, from weekly to daily, by using the seasonal indexes for days of week.
  - Disaggregate forecasts, from daily to hourly, by using the seasonal indexes for hours of day.



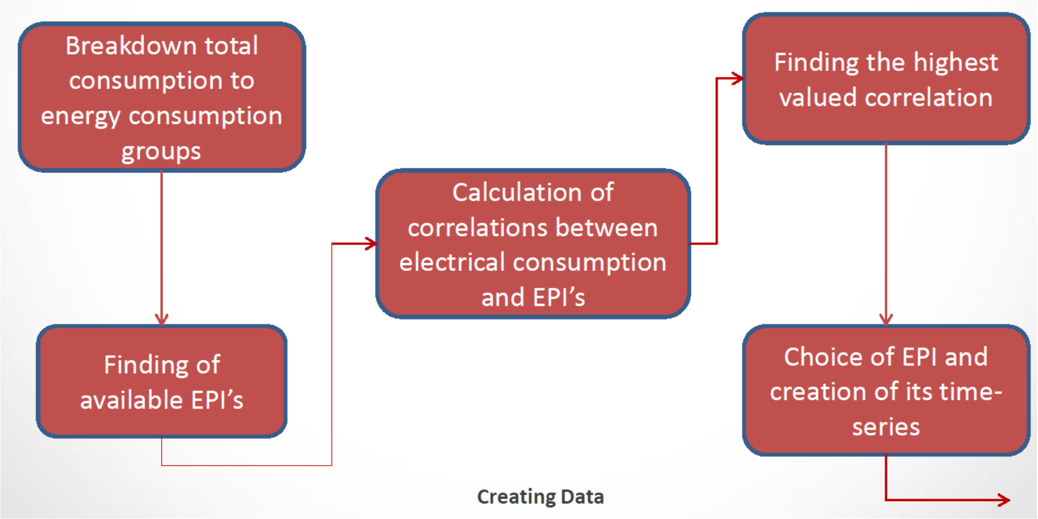
FORECASTING ENERGY CONSUMPTION OF BUILDINGS

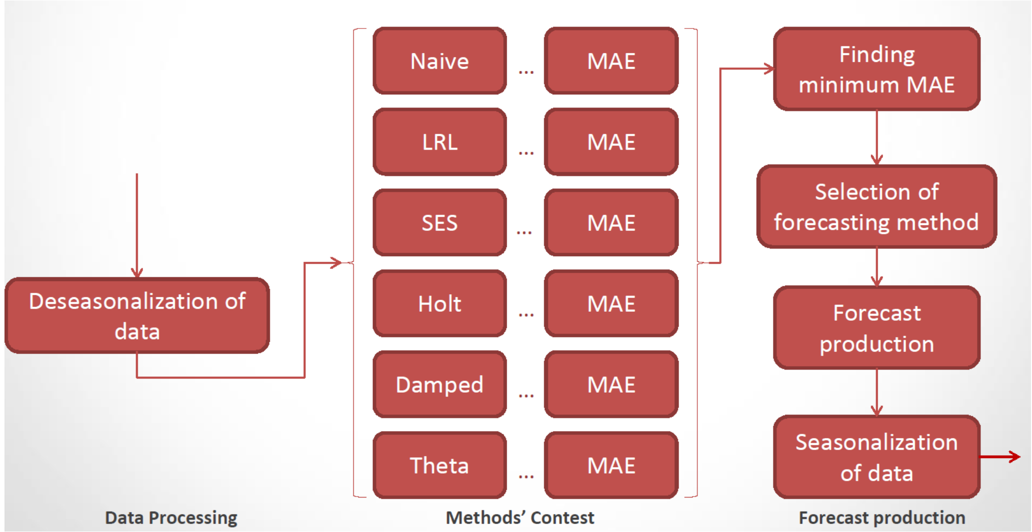
Interesting Use-Cases
Forecasting Energy Consumption of Buildings (1)

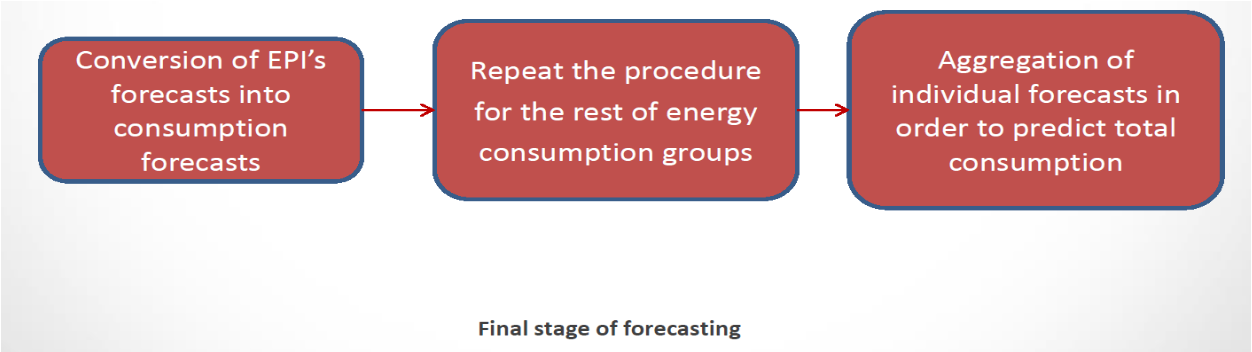
- ***Scope***: Use performance indicators for forecasting consumption in energy-intensive buildings.
  - Three types of indicators to investigate:
    - ***Physical Indicators***: Relation between energy consumption and a physical measured object.
    - ***Value Based Indicators***: Relation between energy consumption and the economic return that this implies.
    - ***Usual Energy Performance Index EPI*** (kWh/m2): Universality and convenience.

Source: *"Forecasting energy consumption of buildings using performance indicators", Spiliotis et. al., Forecasting & Strategy Unit (FSU), National Technical University of Athens (NTUA)*

Interesting Use-Cases
Forecasting Energy Consumption of Buildings (2)







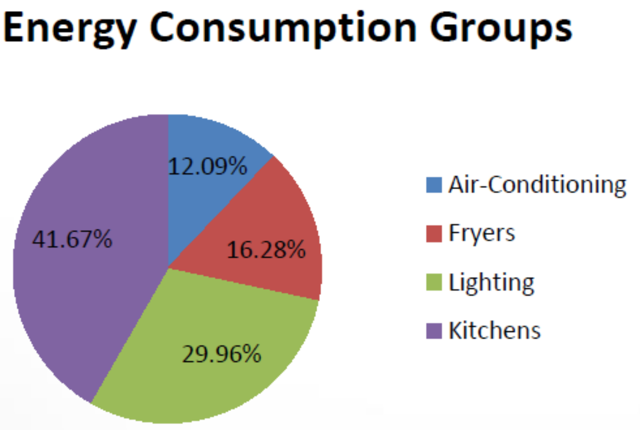
BOSCH

Interesting Use-Cases
Forecasting Energy Consumption of Buildings (3)

- ***Example***: Energy consumption in a restaurant

Energy Consumption Groups:

- ✓ Kitchen (two lines)
- ✓ Lighting
- ✓ Air-Conditioning
- ✓ Fryers



Available Data for EPI construction:

- ✓ Sales(euro)
- ✓ Number of Receipts
- ✓ Operational Hours
- ✓ Temperature(inside, outside & its difference)
- ✓ Percentage of electrical consumption in each energy consumption group
- ✓ Number of Employees
- ✓ Floor Area

Interesting Use-Cases
Forecasting Energy Consumption of Buildings (4)

- Find correlations in the energy consumption groups

Results for kitchen consumption group(line 1)

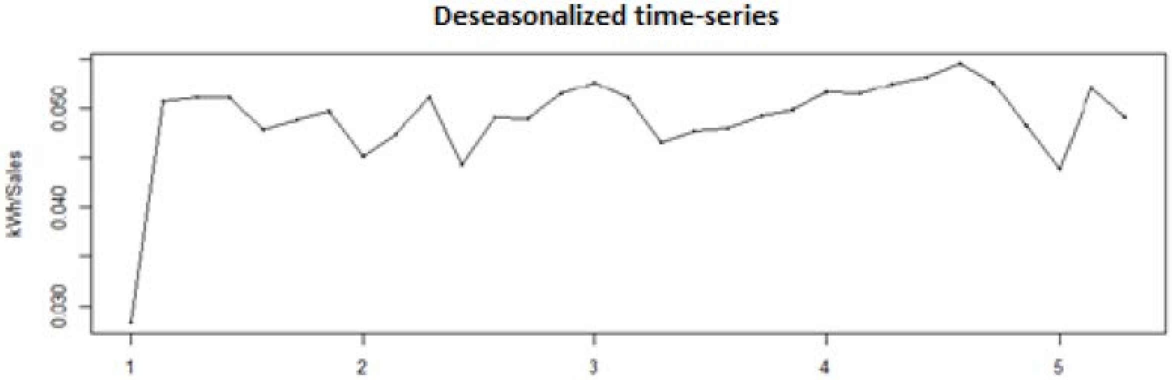
EPI's denominator	Correlation Value
O.H.	-0.194071
% Consumption	-0.239891
Inside Temperature	-0.275588
Outside Temperature	0.530509
Differential Temperature	-0.310188
Sales	**0.590446**
Receipts	0.340027

Summarized total results

Energy Consumption Group	*Optimal EPI*	*Correlation Value*
Kitchen 1	Sales	0.590446
Kitchen 2	Sales	0.784185
Fryers	Sales	0.967809
Air-Conditioning	% Consumption	0.983528
Lighting	% Consumption	0.480626

Interesting Use-Cases
Forecasting Energy Consumption of Buildings (5)

- Starting Forecast Production (KWh/Sales)



	ME	RMSE	MAE	MPE	MAPE
Naive	0.000672215	0.005021201	**2.90E-03**	1.132450058	5.914250798
Ses	2.56E-06	4.55E-03	**2.77E-03**	-1.28E+00	6.59E+00
Holt	1.39E-05	4.81E-03	**3.08E-03**	-6.59E-01	6.71E+00
Damped	6.79E-06	4.22E-03	**2.87E-03**	-1.08E+00	6.68E+00
Theta	7.16E-07	4.33E-03	***2.75E-03***	-1.20E+00	6.51E+00
LRL	-2.84E-17	4.26E-03	**2.84E-03**	-1.12E+00	6.65E+00

Internal | CR/AEU2 - Patelis | 2018-03-14

Interesting Use-Cases
Forecasting Energy Consumption of Buildings (6)

- Forecasting

Forecasts for the kitchen1 consumption group

Day	Real Total Consumption	Forecast kWh/Sales	Seasonality Index	seasonalized Data kWh/Sales	Final Forecast
1/11/2012	459.2939588	0.05045368	1.1973455	0.060410	579.322869
2/11/2012	461.1831554	0.05054342	0.7593826	0.038382	368.072735
3/11/2012	476.1334234	0.05063317	0.6755555	0.034206	328.023147
4/11/2012	460.3787153	0.05072291	0.8022659	0.040693	390.239185
5/11/2012	453.8522061	0.05081266	1.2816189	0.065122	624.509677
6/11/2012	456.8338446	0.0509024	1.2280855	0.062512	599.480753
7/11/2012	432.5857712	0.05099215	1.0557461	0.053835	516.263175

Total forecasts for the restaurant

Day	Real Total Con/tion	forecast kitchen1	forecast kitchen2	forecasta/c	forecast fryers	forecast lights	Final Forecast
1/11/2012	2252.739	579.322	437.465	396.2419	361.2875	646.0089	2420.326
2/11/2012	2272.101	368.072	302.0829	399.1489	278.8617	602.5187	1950.685
3/11/2012	2550.088	328.023	291.358	405.886	282.3684	589.1806	1896.816
4/11/2012	2523.604	390.2391	339.3794	359.3128	312.9879	640.8097	2042.729
5/11/2012	2163.014	624.509	516.9215	368.931	387.1452	639.2548	2536.762
6/11/2012	2244.768	599.480	472.8116	298.8419	371.8767	678.3153	2421.326
7/11/2012	2110.074	516.263	432.9509	353.9452	339.0651	619.9037	2262.128

Total accuracy of method

Day	ME%	MAE%
1/11/2012	-7.43926	7.439264
2/11/2012	14.14622	14.14622
3/11/2012	25.61762	25.61762
4/11/2012	19.05508	19.05508
5/11/2012	-17.279	17.27902
6/11/2012	-7.86533	7.86533
7/11/2012	-7.20608	7.206084
Average	2.71846	14.08695

- Quite good accuracy -14.09%.
- Very good behavior of the method in terms of bias –2.71% pessimism.

BOSCH

MORE?

ADVANCED TIMESERIES FORECASTING…

Friday, September 28th, 08:30 – 12:30

BOSCH

Timeseries Forecasting (A)
Agenda

10. Statistical Methods - Advanced

11. Judgmental forecasting methods

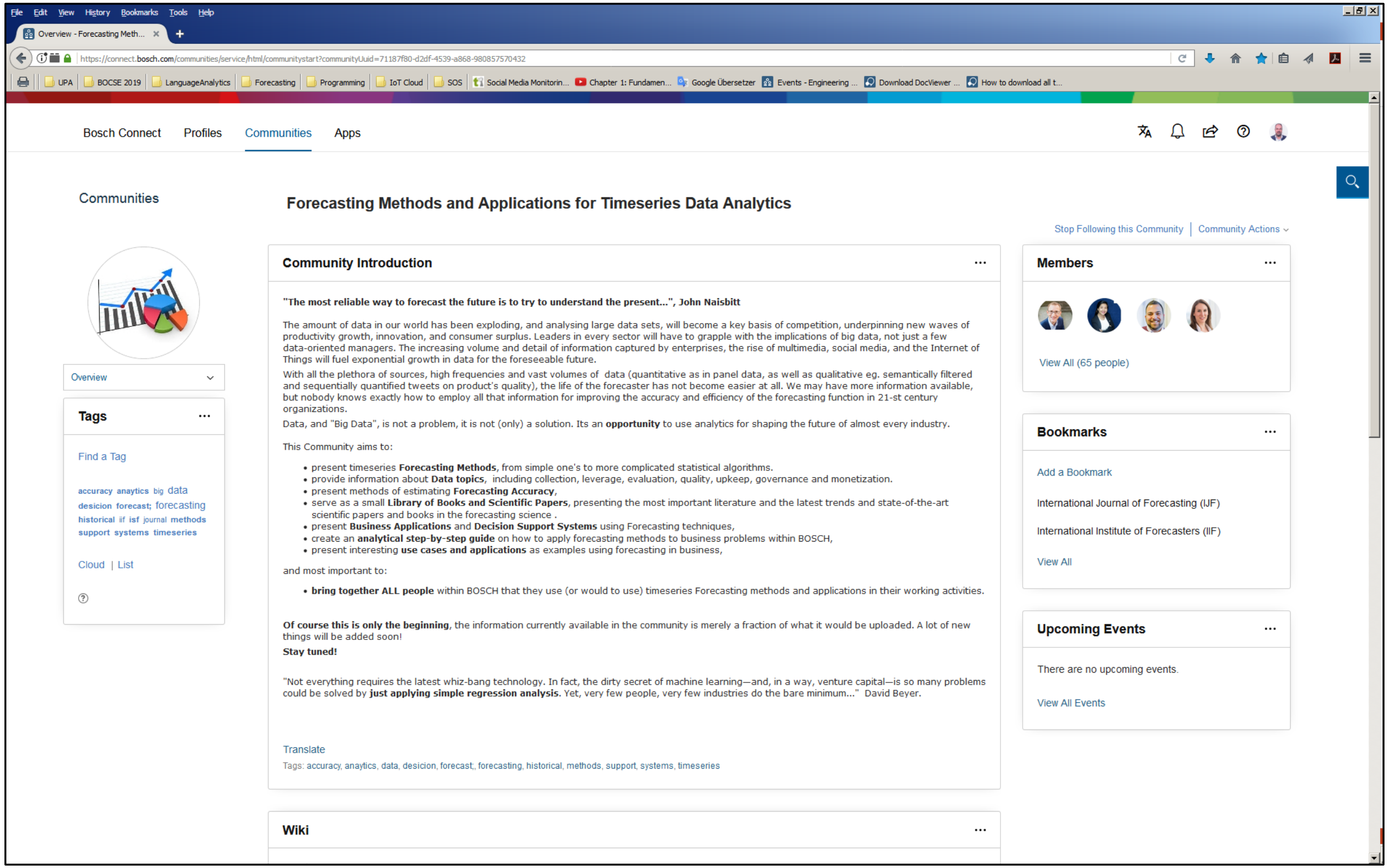
12. Special topics - Intermitted demand

13. Special topics - Structural changes detection

14. Special topics - Other

15. Interesting use cases

Timeseries Forecasting – Data Analytics & Quantitative Methods
Bosch Community



PROPOSED
LINKS & LITERATURE

Timeseries Forecasting – Data Analytics & Quantitative Methods
Proposed Links & Literature

- R. Hyndman, G. Athanasopoulos (2013): "*Forecasting: Principles and Practice*"
- J. Ord, R. Fildes (2012): "*Principles of Business Forecasting*"
- T. Ruud, S. Babangida (2009): *"On the bias of Croston's forecasting method*", European Journal of Operational Research, Vol.194, pp.177-183.
- J. Hanke, A. Reitsch (2008): "*Business Forecasting*"
- K. Nikolopoulos, P. Goodwin, A. Patelis, V. Assimakopoulos (2007): *"Forecasting with cue information: A comparison of multiple regression with alternative forecasting approaches"*, European Journal of Operational Research, Vol.180, pp.354-368.
- G. Rowe (2007): *"A guide to Delphi",* International Journal of Applied Forecasting 8, 11-16.
- P. Goodwin, R. Fildes, M. Lawrence, K. Nikolopoulos (2007): *"The process of using a forecasting support system"*, International Journal of Forecasting (IJF), Vol.23, pp.391-404.
- K. C. Green, J. S. Armstrong (2007): *"Structured analogies for forecasting"*, International Journal of Forecasting 23(3), 365–376.
- R. Hyndman, A. Koehler (2006): *"Another look at measures of forecast accuracy"*, IJF, Vol.22, pp.679-688.
- E. Gardner (2006): *"Exponential smoothing: The state of the art – Part II"*, IJF, Vol.22, pp.637-666.

Timeseries Forecasting – Data Analytics & Quantitative Methods
Proposed Links & Literature

- J. De Gooijer, R. Hyndman (2006): *"25 years of time series forecasting"*, IJF, Vol.22, pp.443-473.
- M. Lawrence, P. Goodwin, M. O'Connor, D. Önkal (2006): "*Judgmental forecasting: A review of progress over the last 25 years",* International Journal of Forecasting 22(3), 493–518.
- Synetos, J. Boylan (2005): *"The accuracy of intermittent demand estimates"*, IJF, Vol.21, pp.303-314.
- Synetos, J. Boylan, J. Croston (2005): *"On the categorization of demand patterns"*, Journal of Operational Research Society, Vol.56, pp.495-503.
- R. Buehler, D. Messervey, D. Griffin (2005), *"Collaborative planning and prediction: Does group discussion affect optimistic biases in time estimation?"*, Organizational Behavior and Human Decision Processes 97(1), 47–63.
- R.G. Brown (2004), *"Smoothing, Forecasting and Prediction of Discrete Time Series"*, Courier Corporation ,Technology & Engineering.
- Hyndman & Billah (2003): *"Theta: SES with drift?"*
- R. Hyndman, A. Koehler, R. Snyder, S. Grose (2002): *"A state space framework for automatic forecasting using exponential smoothing methods"*, IJF, Vol.18, pp.439-454.
- V. Assimakopoulos, K. Nikolopoulos (2000): *"The Theta model: A decomposition approach to forecasting"*, IJF, Vol.16, pp.521-530.
- S. Makridakis, M. Hibon (2000): *"The M3-competition: Results, conclusions and implications"*, IJF, Vol.16 (special issue), pp.451-476.

Timeseries Forecasting – Data Analytics & Quantitative Methods
Proposed Links & Literature

- G. Rowe. G. Wright (1999): *“The Delphi technique as a forecasting tool: issues and analysis”*, IJF, Vol.15, pp.353-375.
- S. Makridakis, S. Wheelwright, R. Hyndman (1998): “*Forecasting: Methods and Applications*”
- Assimakopoulos V. (1995): *“A successive filtering technique for identifying long-term trends”*, IJF Vol14, 35-43.
- J. D. Hamilton (1994): “*Timeseries Analysis*”
- J. Armstrong, F. Collopy (1992): *“Error Measures for generalizing about forecasting methods: Empirical comparisons”*, IJF, Vol.8, pp.69-80.
- S. Makridakis (1990): *“Sliding Simulation: A new approach to time series forecasting*”, Management Science, Vol.36, pp.505-512.
- S. Wheelwright, S. Makridakis (1985): “*Forecasting methods for Management*”
- S. Makridakis, A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, R. Winkler (1982): *“The accuracy of extrapolation (time series) methods: Results of a forecasting competition”*, IJF, Vol.1, pp 111-153.
- H. Kahn, A.J. Wiener (1967): “*The use of scenarios*”, Hudson Institute
- http://www.forecasters.org
- http://www.sciencedirect.com
- http://www.elsevier.com

THANK YOU