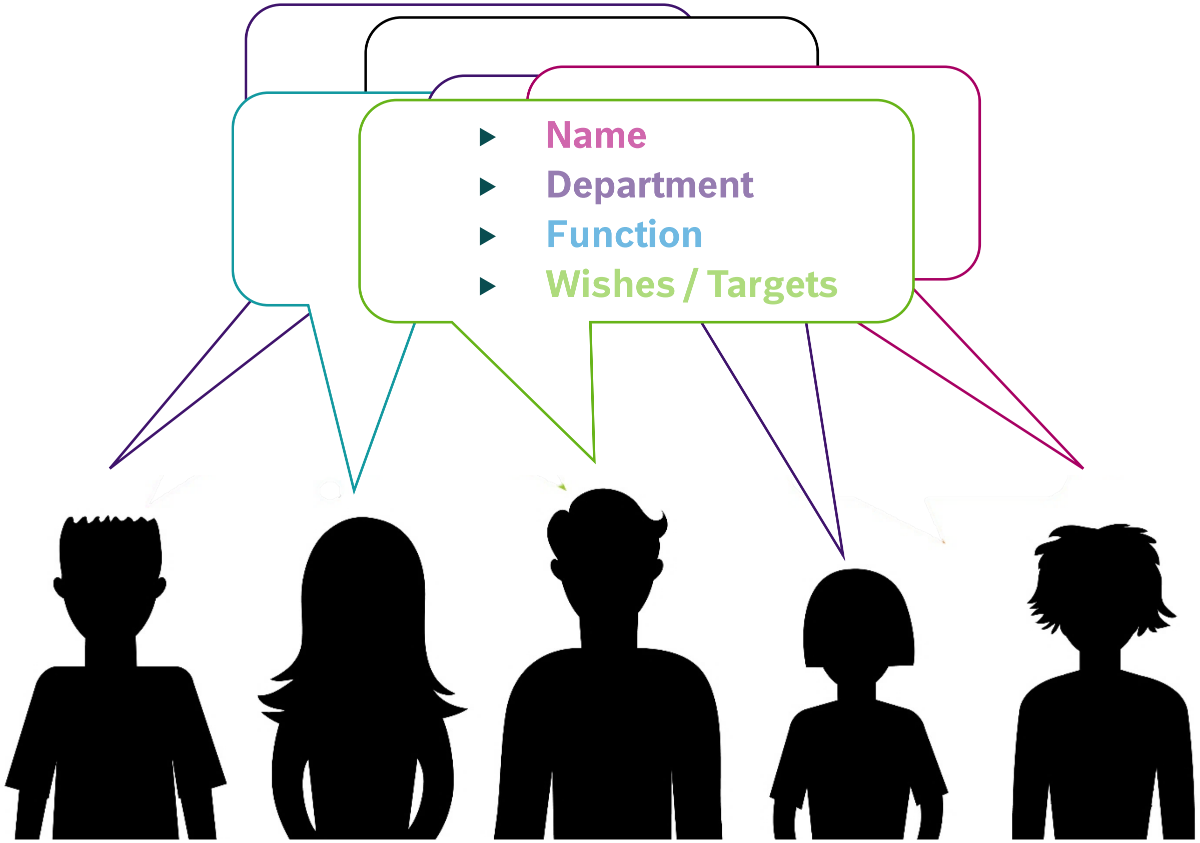
TIMESERIES FORECASTING (B)

ALEXANDROS PATELIS (CR/AEU2)

Timeseries Forecasting
Introduction of Participants and Trainer



BOSCH

Timeseries Forecasting
Introduction of Trainer: Patelis Alexandros (CR\AEU2)

- Timeseries Forecasting for Decision Support Systems
  - PhD, National Technical University of Athens, 1999

- ***Forecasting Working Fields***:
  - Stock Exchange
  - Banking
  - Insurance
  - Media & Advertisement
  - Telecommunications
  - Energy
  - Environment
  - Sales
  - Production
  - Customer Management
  - Travel & Tourism
  - Logistics
  - Real Estate
  - Medicine
  - Military

Timeseries Forecasting
Introduction of Participants

	Name	Department	Function	Wishes / Targets
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

Timeseries Forecasting: Beginners

Agenda

1 Introduction

2 Characteristics & Components

3 Statistical Analysis

4 Errors & Accuracy (Basic)

5 Forecasting Methods

6 Naive

7 Moving Average

8 Linear Regression

9 Exponential Smoothing

10 Exercise

11 Who can be a forecaster ?

12 Method & Data Selection

13 Special Topic Outliers

14 Special Topic Missing Values

15 Use Case Demand Forecasting

16 Use Case Energy Consumption

18 Use Case Minute Data

18 More... ?

Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

Timeseries Forecasting
Content

- What we will see today:
  - Some theory…
    - but not more than necessary!

  - Mathematical equations…
    - but only for completion!

$$S(\omega) = \frac{\alpha g^2}{\omega^5} e^{\left[-0.74\left\{\frac{\omega U_\omega 19.5}{g}\right\}^{-4}\right]}$$
$$= \frac{\alpha g^2}{\omega^5} \exp\left[-0.74\left\{\frac{\omega U_\omega 19.5}{g}\right\}^{-4}\right]$$

- But also:
  - Analytical examples 
  - Key messages & conclusions 
  - Best advices & rules of thumb 
  - Live examples with your participation 

Introduction
What about your data?

- Discussion (all)
  - What type of timeseries do you handle?
    - Yearly? Monthly?
  - What type of data?
    - Sales? Demand? Budget? Cost?
  - What type of problems are you facing in daily work?
    - Lack of data?
    - Abnormal values?
    - ...

BOSCH

1. INTRODUCTION

“A journey of a thousand miles begins with a single step."

Lao Tzu

Introduction
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? Biases?
- Can we trust a computer method?

Forecasting: Common sense + Probabilities + Calibration + Gut + Expertise + Computer algorithm

BOSCH

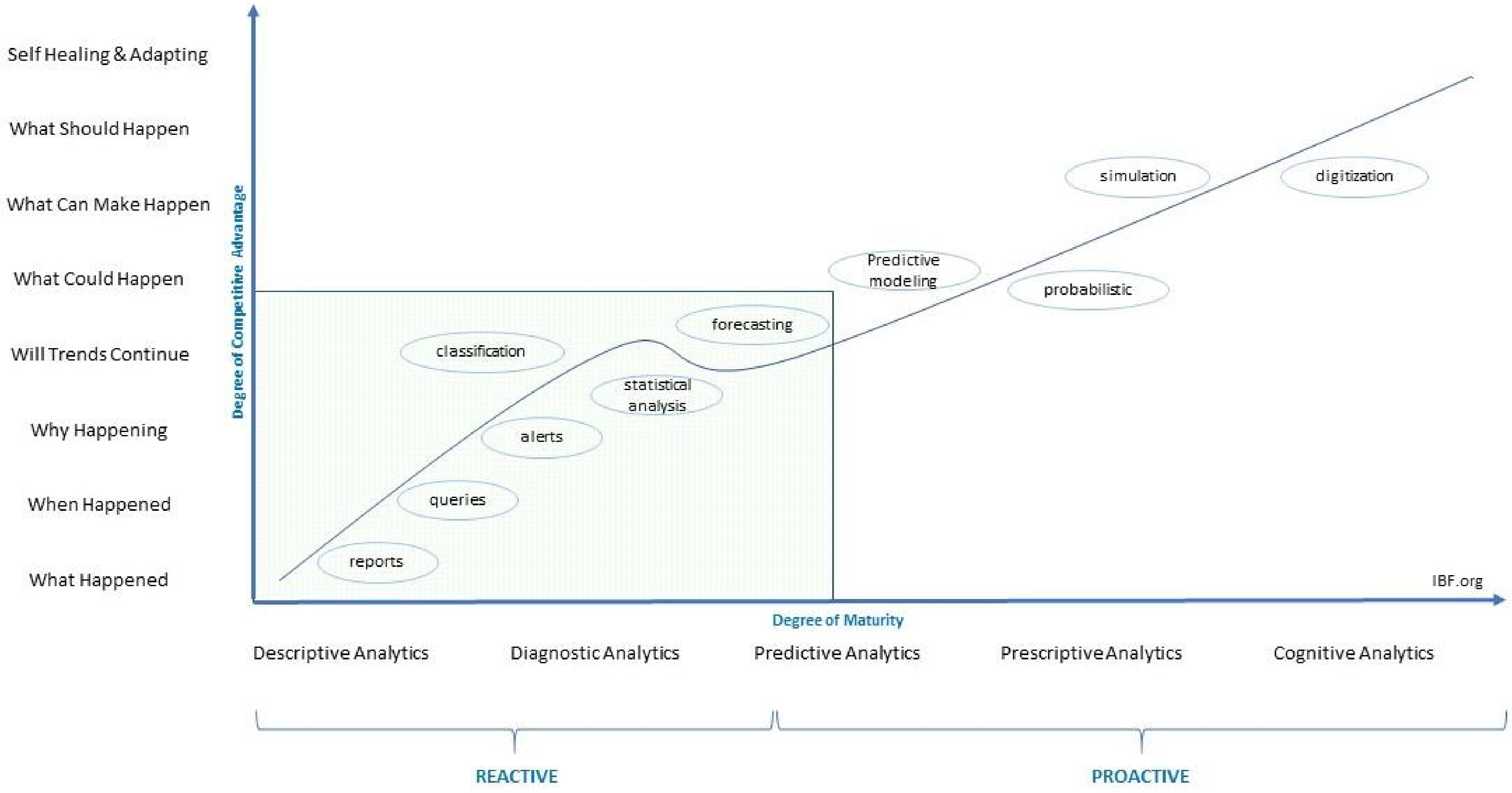
Source: Forecasting – Principles and Practice

Introduction
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 for many factors, such as:
    - future demand,
    - technological innovations,
    - cost,
    - prices,
    - competitors' plan,
    - labor,
    - legislation,
    - etc.

Introduction
Why forecasting?



Introduction
Forecasting problem

Data	
Year	**Value**
2009	1167
2010	1169
2011	1167
2012	1170
2013	1159
2014	1174
2015	1180
2016	1177
2017	1178
2018	1170

- What we have:
  - Some data
    - Table
    - Graph
- What we want:
  - Understand historical values
  - Estimate future values
    - next year value
- What is your estimation for 2019?

Can you find the black cat in the room?

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

Steve Davidson

Introduction
Some Basic Definitions

Forecasting:

- To predict how the sequence of observations will continue in the future.
  - About events whose actual outcomes have not yet been observed.
  - Based on past and present data.

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

Introduction
Forecasting: The real start

Journal of FORECASTING

Article

The accuracy of extrapolation (time series) methods: Results of a forecasting competition

S. Makridakis[1,2,3,4,5,6,7,†], A. Andersen[1,2], R. Carbone[1,‡], R. Fildes[1,3,§], M. Hibon[1,¶], R. Lewandowski[1,‖], J. Newton[1,††], E. Parzen[1,‡‡] and R. Winkler[1,§§]

Version of Record online: 20 SEP 2006

DOI: 10.1002/for.3980010202

Issue

Journal of Forecasting

Volume 1, Issue 2, pages 111–153, April/June 1982

Introduction
Some Basic Definitions

Timeserie: A sequence of data points, measured at **successive points in time** spaced in uniform time intervals.

- ***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.

- **Used in:** statistics, pattern recognition, econometrics, finance, weather, astronomy, ...
  - in **any domain** of applied science and engineering which involves **temporal measurements**.

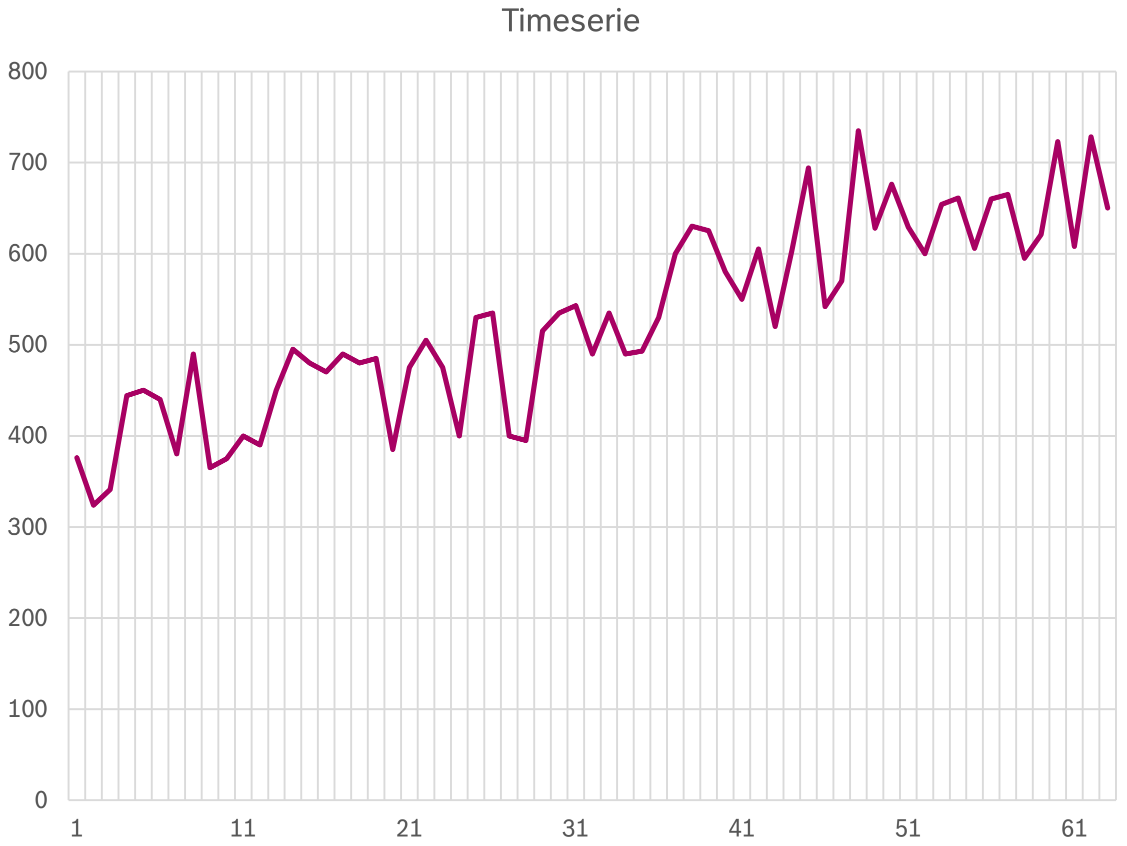
Introduction
Some Basic Definitions

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 Horizon:

- The length of time into the future for which forecasts are to be estimated.
  - **Short-term**: usually when the forecast horizon is less than or equal to 2 periods.
    - Up to 1 year (for annual data)
    - Cases: Inventory, Warehouse design.
  - **Mid-term**: usually when the forecast horizon is 1+ economical year (thus, 12-15 months if we refereed to a monthly timeserie).
    - Up to 5 years
    - Cases: Budget, Economical Planning.
  - **Long-term**: usually when the forecast horizon is more than 5 years.
    - Cases: Investment planning and development.

➢ Depending on the forecast horizon, the appropriate statistical forecasting method must be selected.

Introduction
Some Basic Definitions

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.

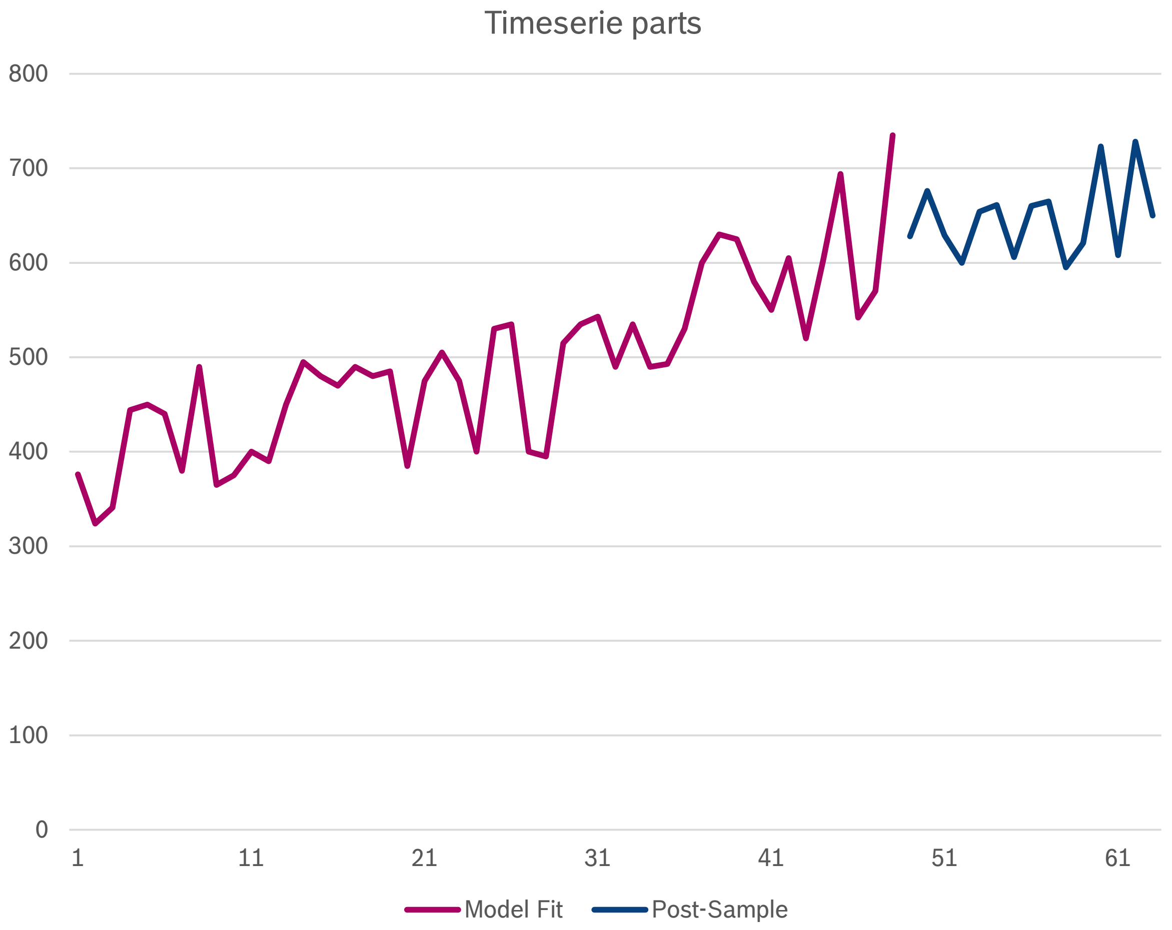
Introduction
Some Basic Definitions

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 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

Introduction
Some Basic Definitions

Relationship:

- The ***influence of one or more factors*** to the forecasted item,
  - 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
Where to use forecast?

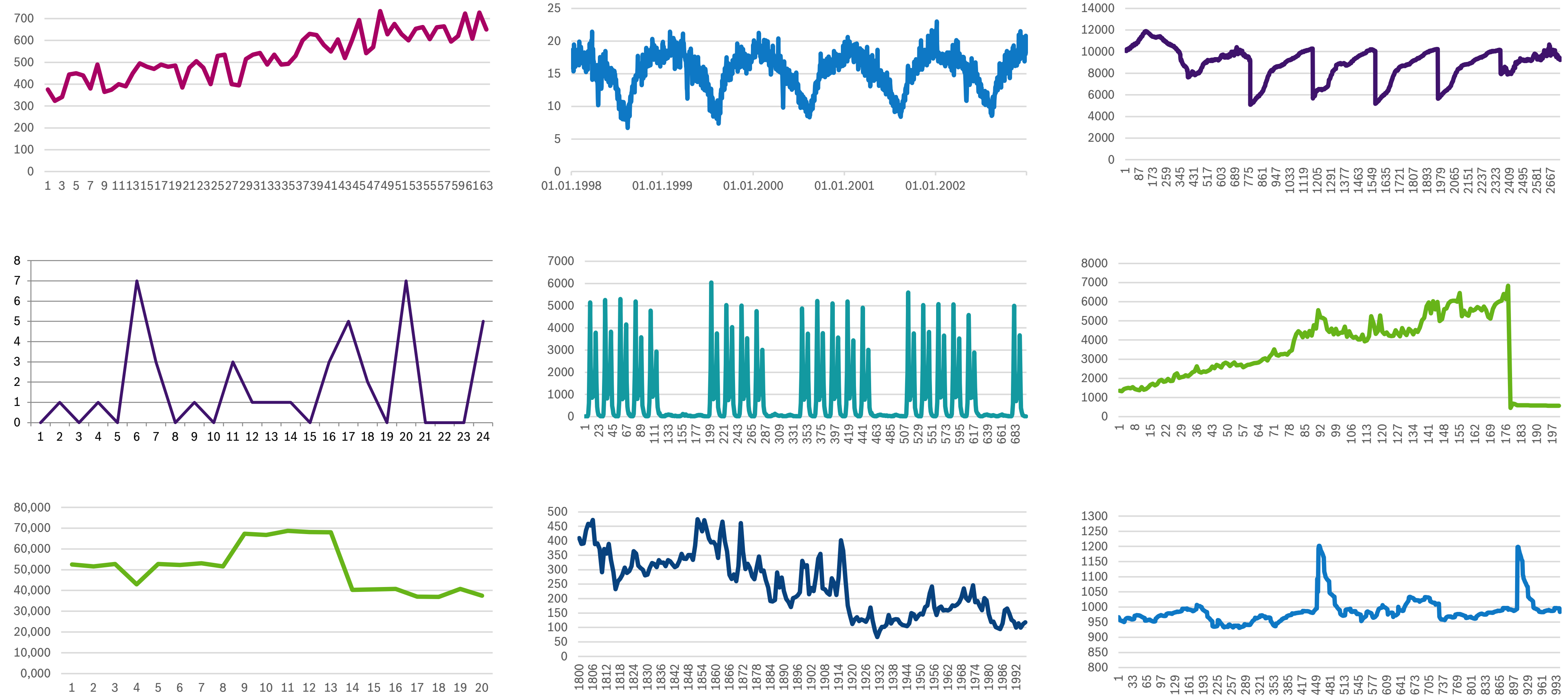
BOSCH

Introduction
Where to use forecast?

- 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.

Introduction

Time serie examples



2. TIMESERIE CHARACTERISTICS & COMPONENTS

"The most reliable way to forecast the future is to try to understand the present."

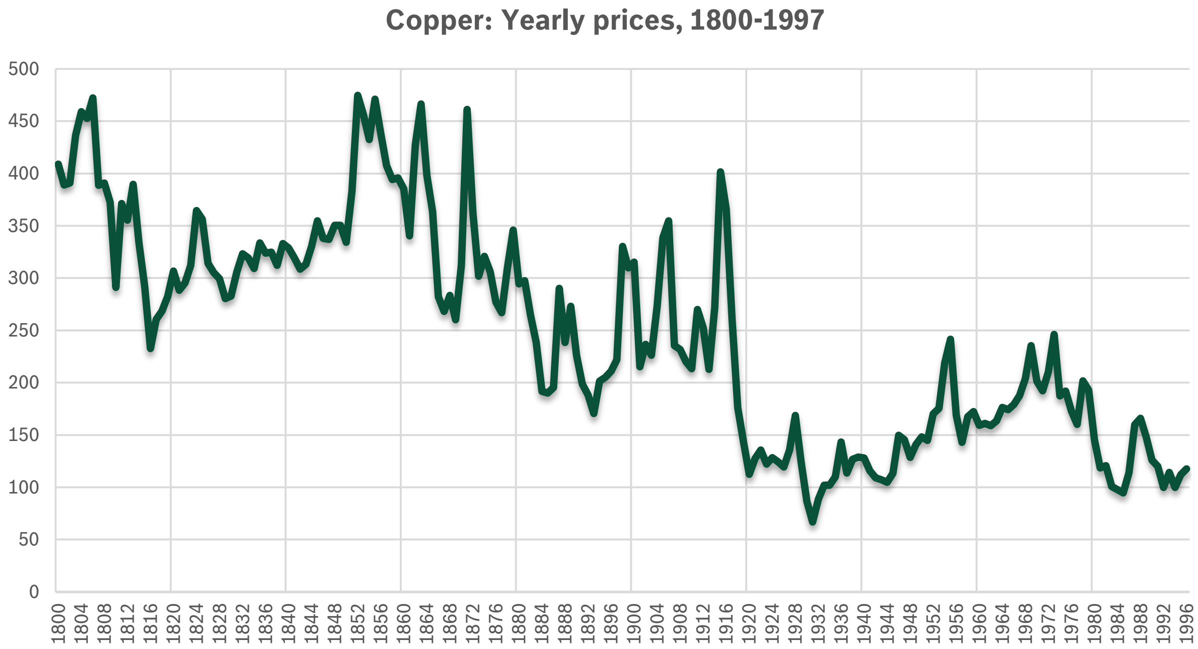
John Naisbitt (Author in area of future studies)

BOSCH

Timeserie Characteristics & Components
Graphical Representation

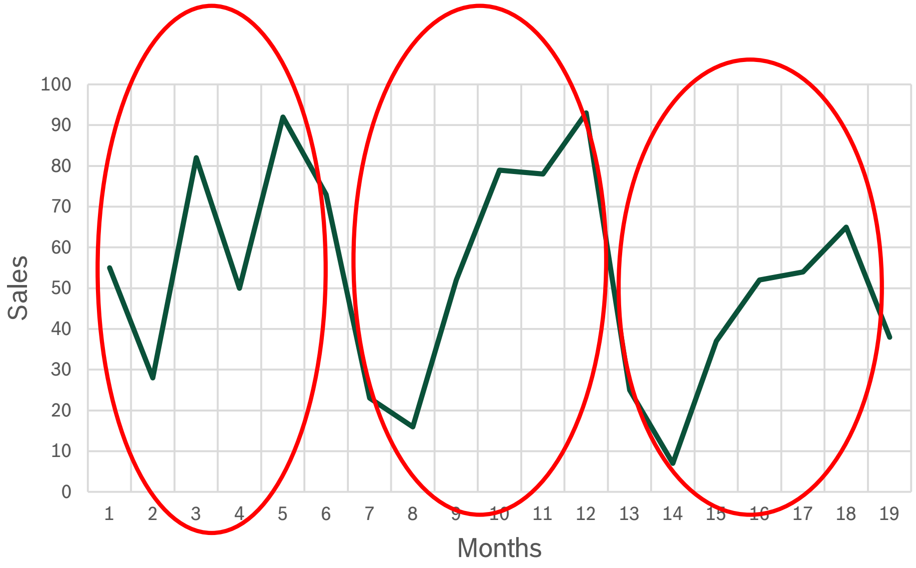
- An important tool for the analysis of the data and the forecasting procedure.
  - a 2-dimensional representation of the timeseries available data points (data values in y-axis, time in x-axis).

- Observe and understand the ***qualitative characteristics*** of the timeseries
- Helps to choose between alternative methodologies and tools.
- Reveal ***outliers*** and ***fault values***.



Timeserie Characteristics & Components
Studying the graph

- Can we derive the characteristics by **just studying the graph**?
  - We can see ***3 underlying cycles*** in the sales, each lasting about 6 months.
  - Within each of these cycles the sales 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.

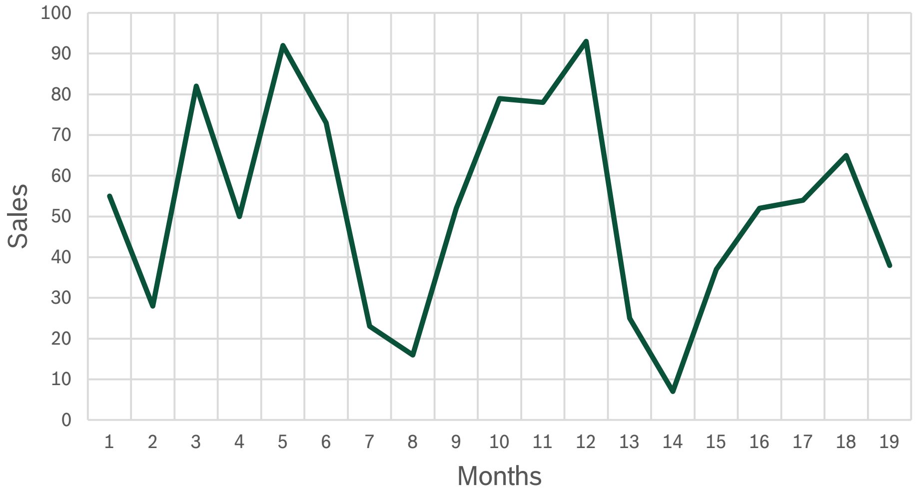


- Forecast:
  - Extend the pattern for 6 more months…

Timeserie Characteristics & Components
Studying the graph

- There is only a problem…
  - the apparent ‘sales figures’ on the graph are actually **random numbers** from 1 to 100.

- There is ***no systematic pattern*** underlying the numbers
- Forecasts based on the 6-month cycle are likely to be well off target.

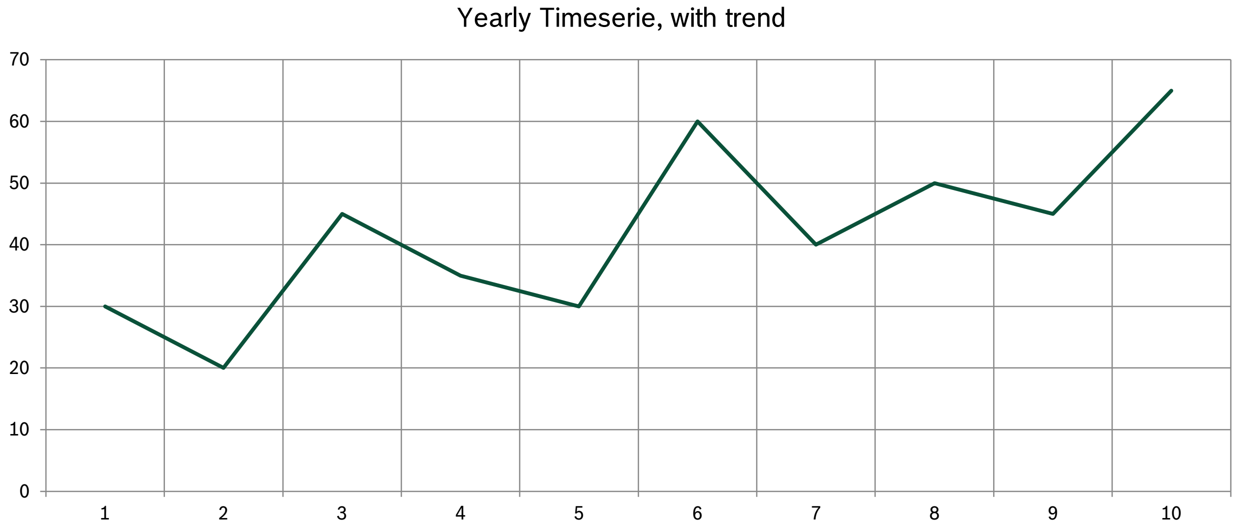


- Not only we are adept at finding patterns where there are none, but we are also ***brilliant at inventing theories*** to explain these patterns.

Timeserie Characteristics & Components
Graphical Representation

Trend/Stationary:

- Timeserie data tend over time: a) increase, b) decrease, c) 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.

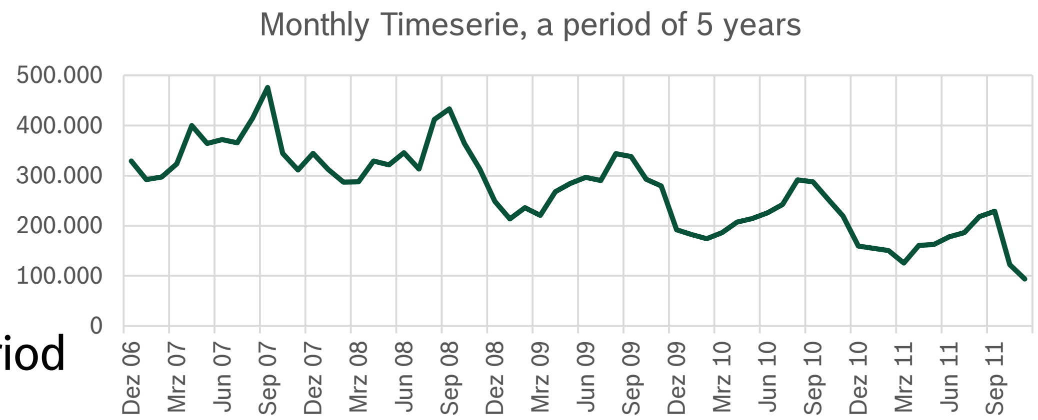


Timeserie Characteristics & Components
Graphical Representation

Seasonal:

- A ***regularly repeating pattern*** of highs and lows related to calendar time
  - Calendar time: seasons, quarters, months, weeks, weekdays, etc.
- 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)

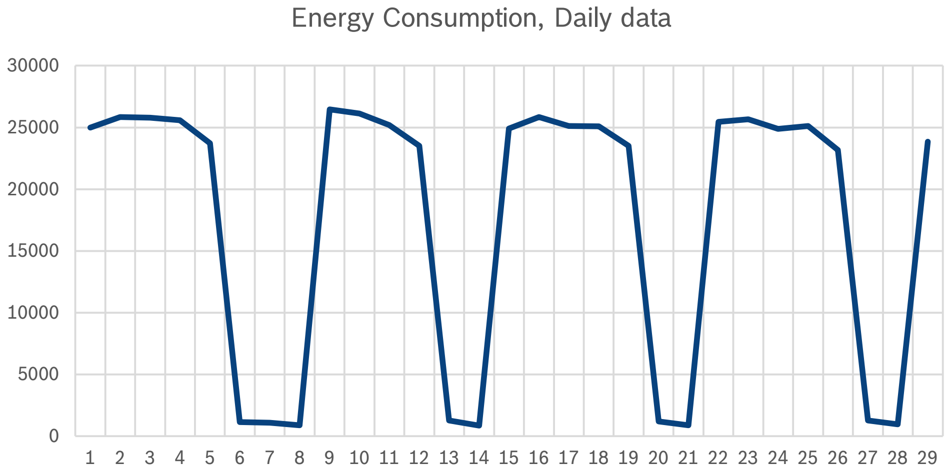
Seasonality is always of a fixed and known period

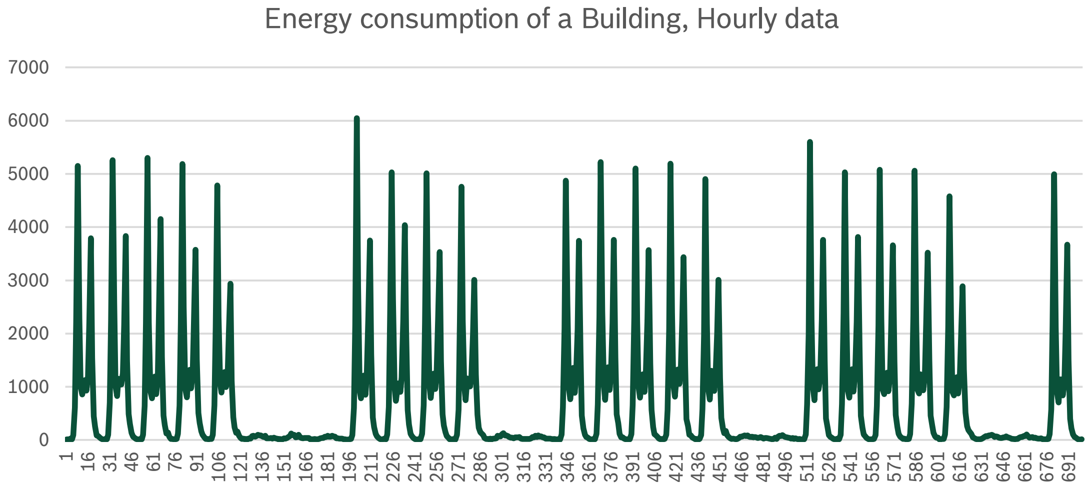


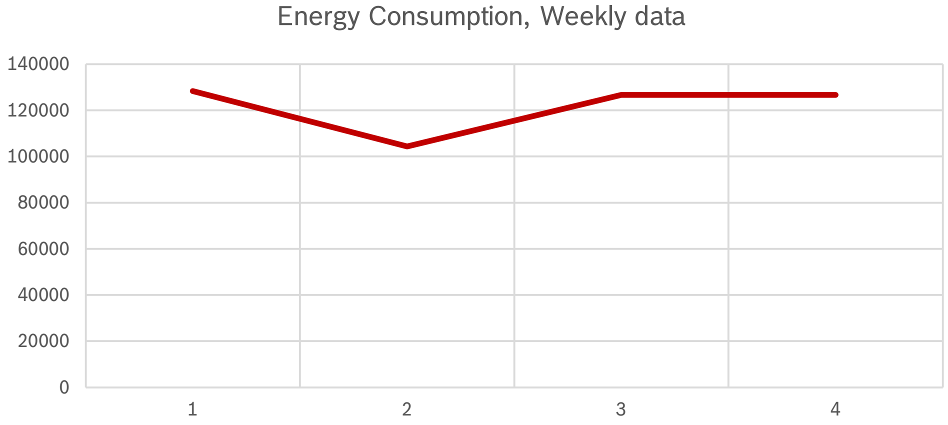
Timeserie Characteristics & Components
Graphical Representation

Seasonal:

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







BOSCH

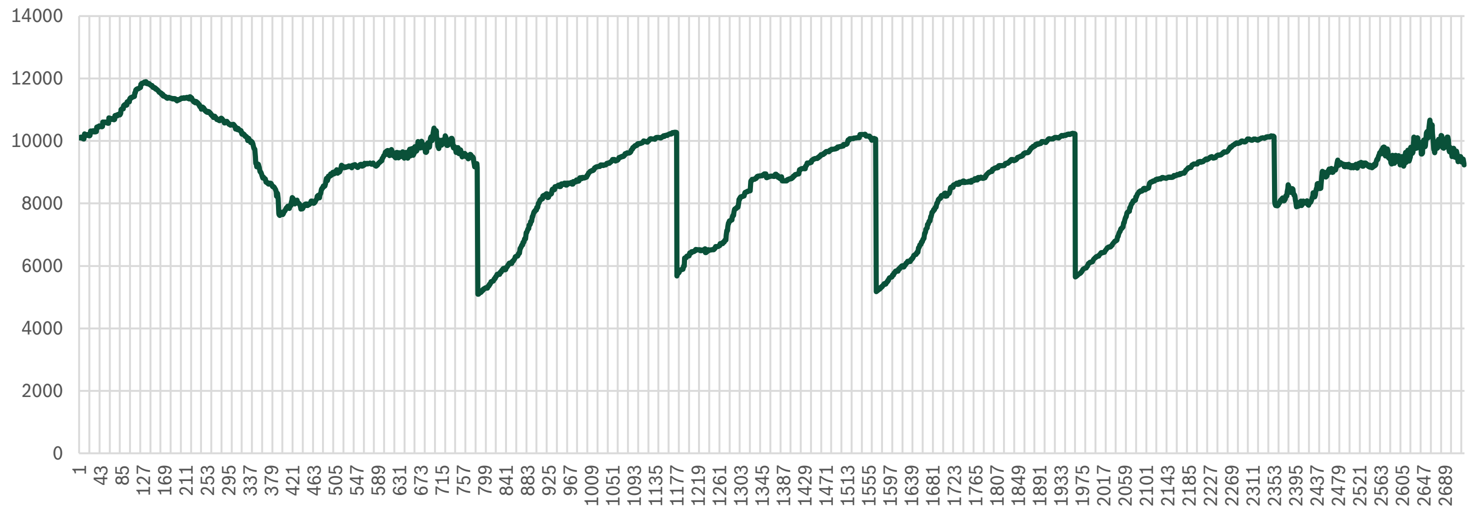
Timeserie Characteristics & Components
Graphical Representation

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***.

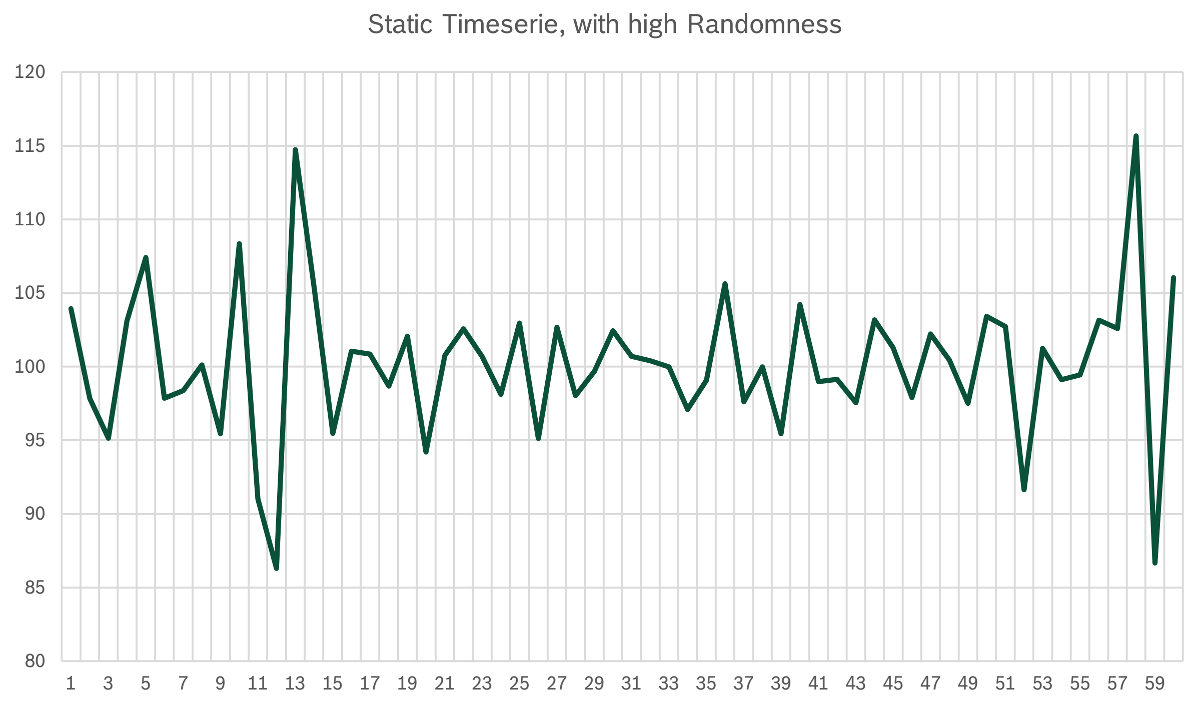
An Industry timeserie, monthly historical data for 227 years



Timeserie Characteristics & Components
Graphical Representation

Randomness:

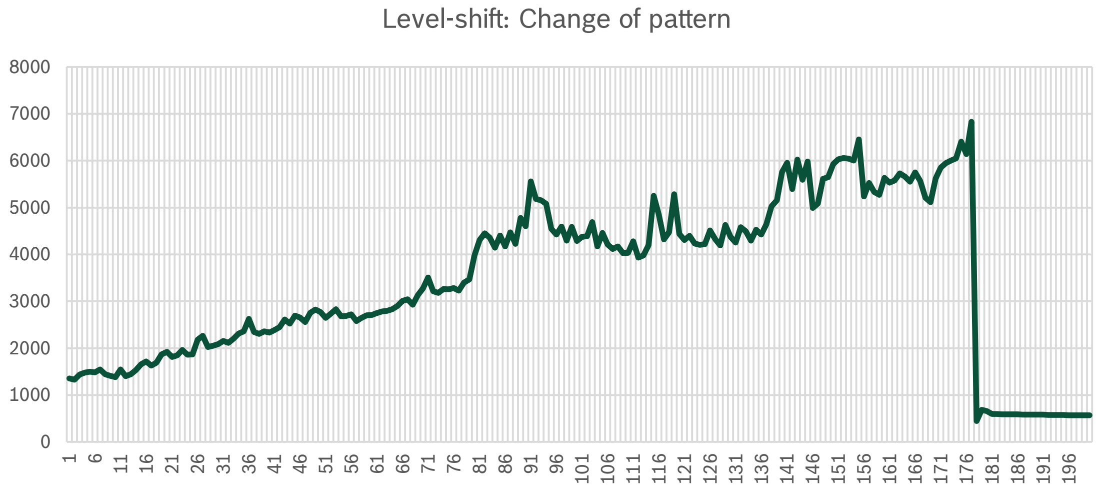
- How large randomness exists on the data?

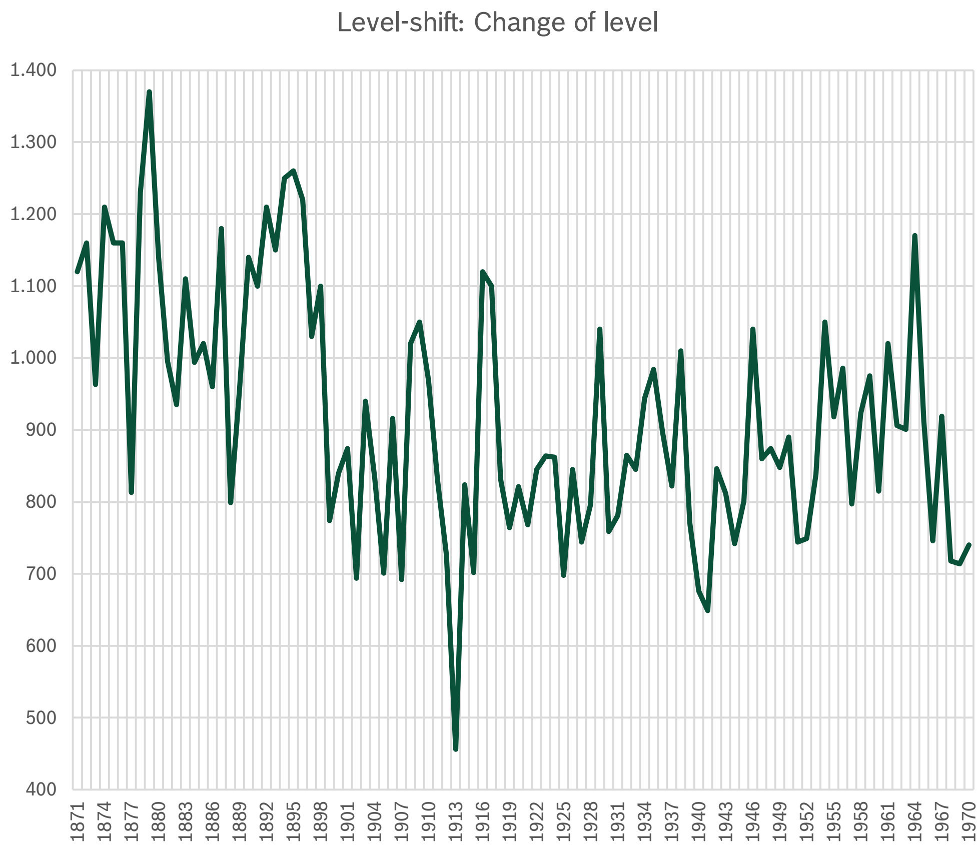


Timeserie Characteristics & Components
Graphical Representation

Level-shift:

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

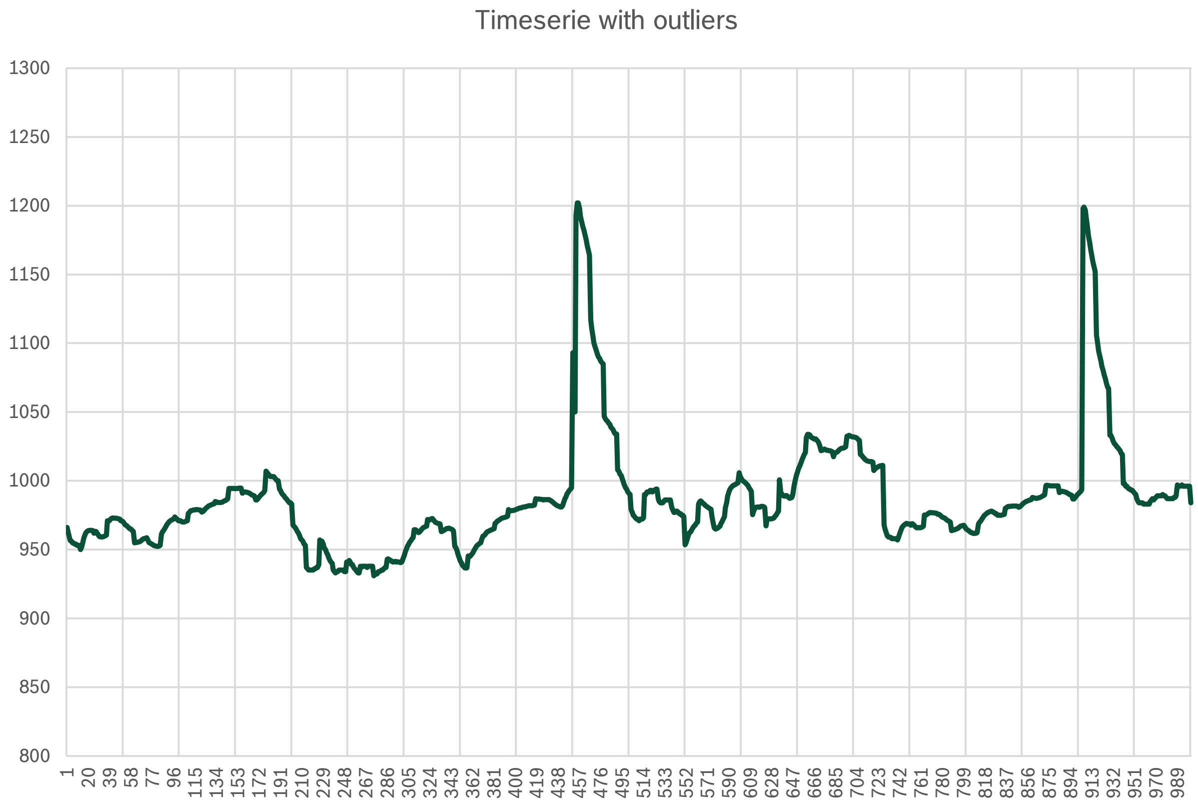




Timeserie Characteristics & Components
Graphical Representation

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?

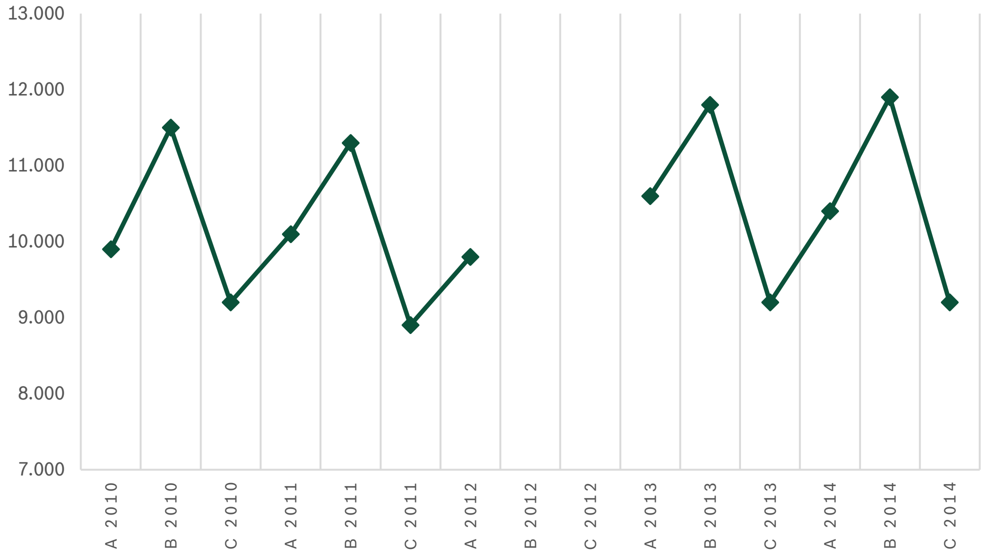


Timeserie Characteristics & Components
Graphical Representation

Fault or Missing values:

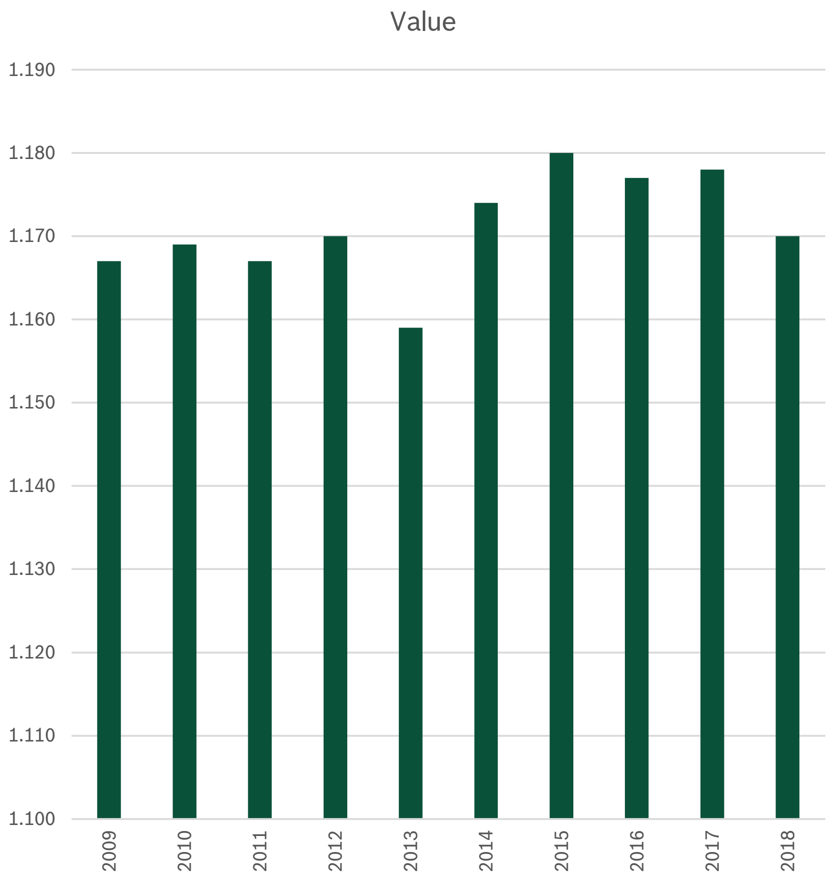
- Are some ***data point values missing*** from the timeseries data?
- Do we believe that some data point values are **fault**?

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



Timeserie Characteristics & Components
An example

Data	
Year	**Value**
2009	1.167
2010	1.169
2011	1.167
2012	1.170
2013	1.159
2014	1.174
2015	1.180
2016	1.177
2017	1.178
2018	1.170



- What you can observe?
  - 10 data points
  - Yearly data
  - No trend over the years (stationary timeserie)
  - No “fault” or “missing” values
  - No significant trend changes during years
  - Average (mean value) = 1.171

3. TIMESERIE STATISTICAL ANALYSIS

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

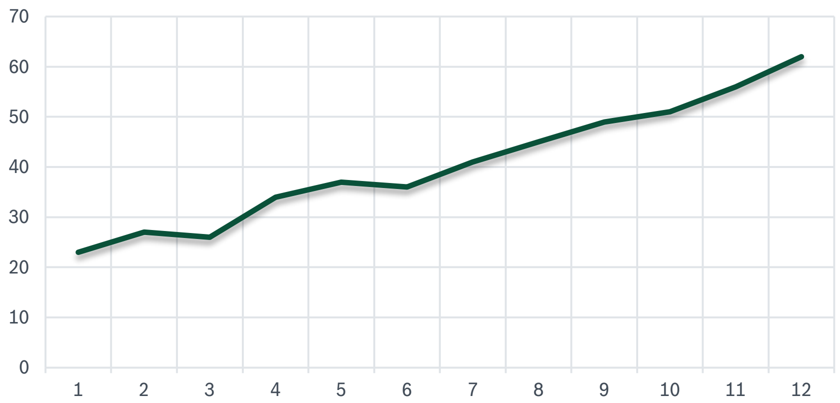
W. Allen Wallis (American Economist)

BOSCH

Timeserie Statistical Analysis

Basic Statistical Analysis

- ***Average***: $\bar{Y} = \frac{1}{n} \times \sum_{i=1}^{n} Y_i$



t	Y_t	$Y_t - \bar{Y}$	$(Y_t - \bar{Y})^2$
1	23	-17,58	309,17
2	27	-13,58	184,51
3	26	-14,58	212,67
4	34	-6,58	43,34
5	37	-3,58	12,84
6	36	-4,58	21,01
7	41	0,42	0,17
8	45	4,42	19,51
9	49	8,42	70,84
10	51	10,42	108,51
11	56	15,42	237,67
12	62	21,42	458,67

- ***Standard Deviation***: $\sigma = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1}}$

Y Mean Value	40,58
Standard Deviation	12,3543

- 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***
      - greater values of one variable mainly correspond with greater values of the other variable, and vice versa.
    - ***Negative*** covariance: the variables tend to show ***opposite behaviour***
      - 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.

t	Y_t	$X_t - \bar{X}$	$Y_t - \bar{Y}$	$(X_t - \bar{X}) * (Y_t - \bar{Y})$
1	23	-5,50	-17,58	96,71
2	27	-4,50	-13,58	61,13
3	26	-3,50	-14,58	51,04
4	34	-2,50	-6,58	16,46
5	37	-1,50	-3,58	5,38
6	36	-0,50	-4,58	2,29
7	41	0,50	0,42	0,21
8	45	1,50	4,42	6,63
9	49	2,50	8,42	21,04
10	51	3,50	10,42	36,46
11	56	4,50	15,42	69,38
12	62	5,50	21,42	117,79

Covariance = 40,375

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.

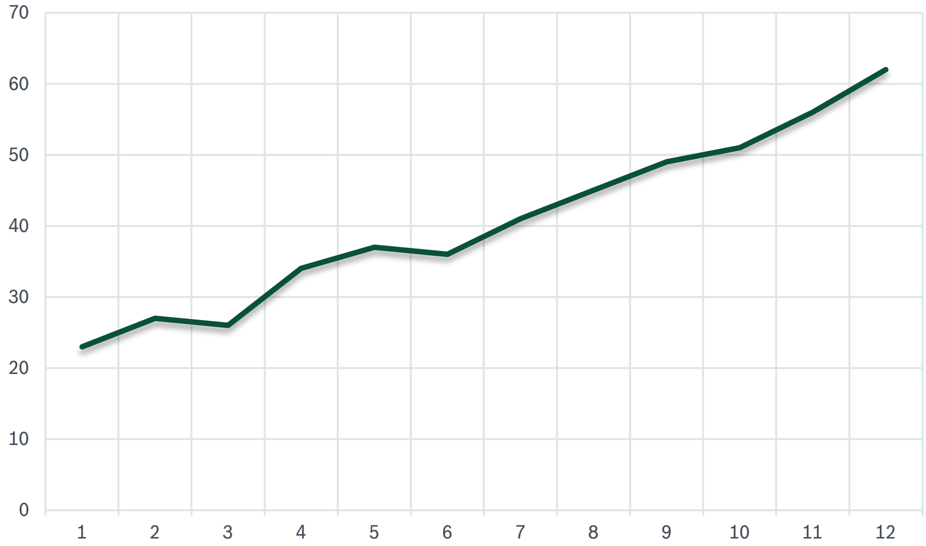
Timeserie Statistical Analysis
Basic Statistical Analysis

▶ Linear Correlation Coefficient:

$$r(X,Y) = \frac{\sum_{i=1}^{n}[(X-\bar{X})(Y-\bar{Y})]}{\sqrt{\sum_{i=1}^{n}(X-\bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y-\bar{Y})^2}}$$

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

Linear Correlation Coefficient = 0,989



t	Y_t	$X_t - \bar{X}$	$Y_t - \bar{Y}$	$(X_t - \bar{X}) * (Y_t - \bar{Y})$	$(X_t - \bar{X})^2$	$(Y_t - \bar{Y})^2$
1	23	-5,50	-17,58	96,71	30,25	309,17
2	27	-4,50	-13,58	61,13	20,25	184,51
3	26	-3,50	-14,58	51,04	12,25	212,67
4	34	-2,50	-6,58	16,46	6,25	43,34
5	37	-1,50	-3,58	5,38	2,25	12,84
6	36	-0,50	-4,58	2,29	0,25	21,01
7	41	0,50	0,42	0,21	0,25	0,17
8	45	1,50	4,42	6,63	2,25	19,51
9	49	2,50	8,42	21,04	6,25	70,84
10	51	3,50	10,42	36,46	12,25	108,51
11	56	4,50	15,42	69,38	20,25	237,67
12	62	5,50	21,42	117,79	30,25	458,67
			sum	**484,50**	**143,00**	**1678,92**

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

- Forecast: Votes
  - Variables: a) inflation 4%, b) unemployment 1.2 million, c) growth was 2%
  - Forecast: government will receive 45% of the votes.
  - Actual: government ***only received 32%*** of the votes → we are disappointed by our formula accuracy.
  - How to improve? Add a variable!
    - annual ***number of umbrellas*** left behind on the tube during election years.
  - New forecast: 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.

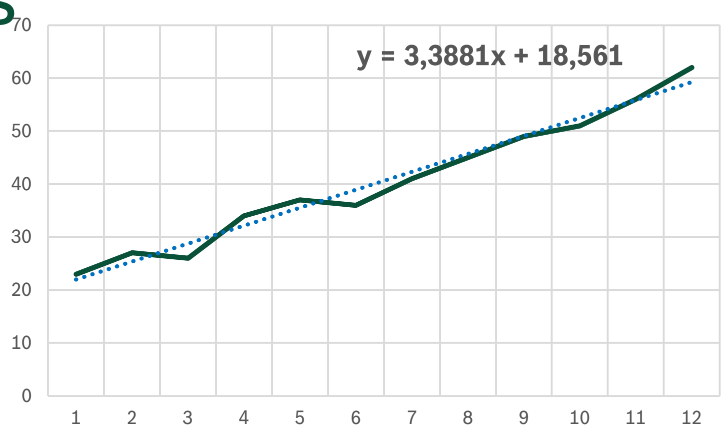
Simplicity and ***common sense*** are the antidotes to overfitting.

Timeserie Statistical Analysis

Basic Statistical Analysis

- ***Coefficient of Determination***:

$$R^2 = \frac{\sum_i (\hat{y}_i - \bar{y})^2}{\sum_i (y_i - \bar{y})^2} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$$



t	Y_t	$\widehat{Y}_t$	$(Y_t - \widehat{Y})^2$	$(Y_t - \overline{Y})^2$
1	23	21,95	1,10	309,17
2	27	25,34	2,76	184,51
3	26	28,73	7,43	212,67
4	34	32,11	3,56	43,34
5	37	35,50	2,25	12,84
6	36	38,89	8,35	21,01
7	41	42,28	1,63	0,17
8	45	45,67	0,44	19,51
9	49	49,05	0,00	70,84
10	51	52,44	2,08	108,51
11	56	55,83	0,03	237,67
12	62	59,22	7,74	458,67
		sum	**37,38**	**1678,92**

R^2 = 0,978

- 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

- ***Autocorrelation Coefficient*:**

$$ACF_k = \frac{\sum_{i=1+k}^{n}[(Y_i - \bar{Y})(Y_{i-k} - \bar{Y})]}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

t	Y_t	$Y_t - \bar{Y}$	$Y_{t-k} - \bar{Y}$	$(Y_t - \bar{Y}) * (Y_{t-k} - \bar{Y})$	$Y_t - \bar{Y}$	$(Y_t - \bar{Y})^2$
1	23				-17,58	309,17
2	27				-13,58	184,51
3	26				-14,58	212,67
4	34				-6,58	43,34
5	37	-3,58	-17,58	63,01	-3,58	12,84
6	36	-4,58	-13,58	62,26	-4,58	21,01
7	41	0,42	-14,58	-6,08	0,42	0,17
8	45	4,42	-6,58	-29,08	4,42	19,51
9	49	8,42	-3,58	-30,16	8,42	70,84
10	51	10,42	-4,58	-47,74	10,42	108,51
11	56	15,42	0,42	6,42	15,42	237,67
12	62	21,42	4,42	94,59	21,42	458,67
		sum		**113,22**		**929,22**

ACF(4) = 0,122

- 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***.

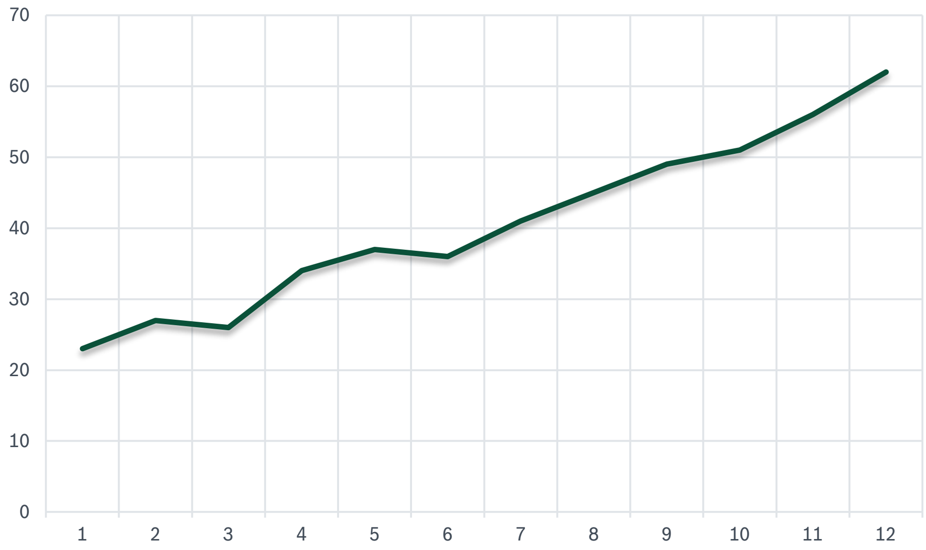
Timeserie Statistical Analysis
Basic Statistical Analysis

- ***Growth Rate***:

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

  - 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



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

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

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

$$= 62{,}08\ \%$$

4. FORECASTING ERRORS & ACCURACY (BASICS)

"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

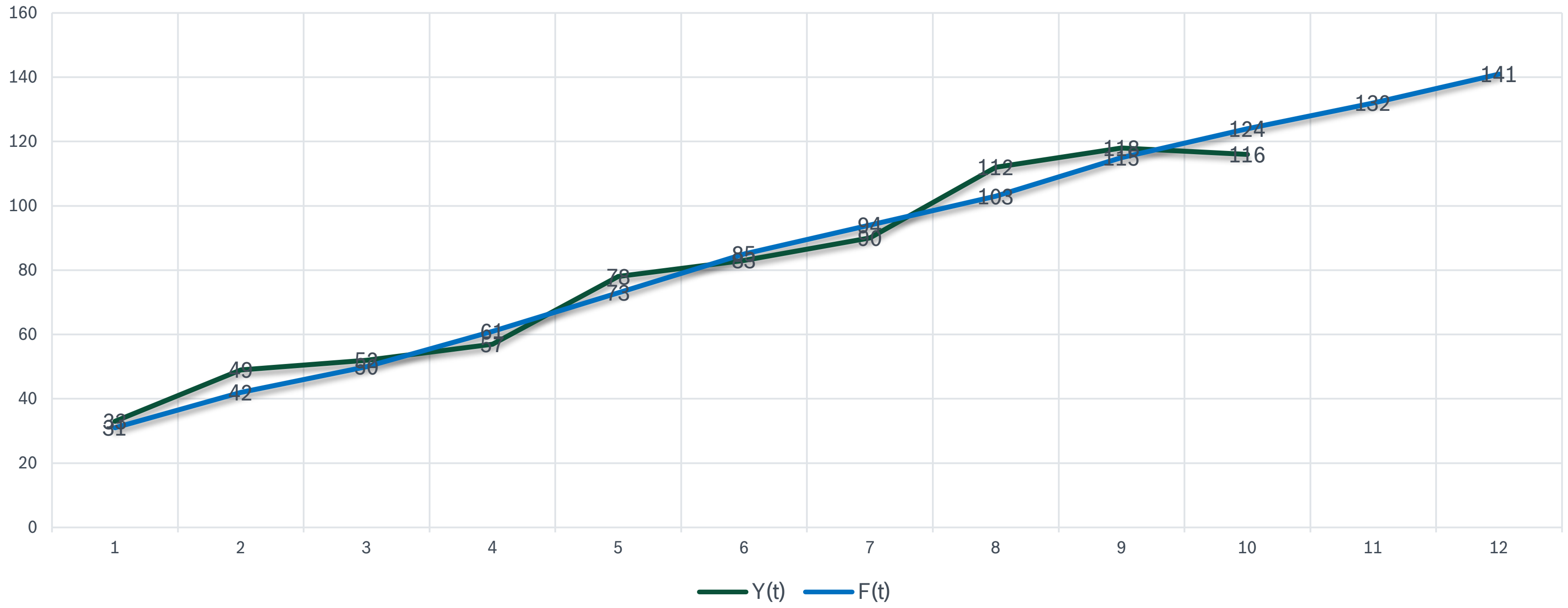
BOSCH

Forecasting Errors & Accuracy (Basics)
Definition

Forecasting Error:

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

t	Y_t	$\widehat{Y_t}$
1	33	31
2	49	42
3	52	50
4	57	61
5	78	73
6	83	85
7	90	94
8	112	103
9	118	115
10	116	124
11		132
12		141



Forecasting Errors & Accuracy (Basics)

Definition

- ***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.
  - $ME = \frac{\sum_{i=1}^{n}(Y_i - F_i)}{n}$

t	Y_t	$\widehat{Y_t}$	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	

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

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

Forecasting Errors & Accuracy (Basics)
Error Types

t	Y_t	$\widehat{Y_t}$	e_t	$\lvert e_t \rvert$	e_t^2
1	33	31	2	2	4
2	49	42	7	7	49
3	52	50	2	2	4
4	57	61	-4	4	16
5	78	73	5	5	25
6	83	85	-2	2	4
7	90	94	-4	4	16
8	112	103	9	9	81
9	118	115	3	3	9
10	116	124	-8	8	64
11		132			
12		141			
		sum	**10,0**	**46,0**	**272,0**
		Mean	**1,0**	**4,6**	**27,2**

- ***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
- ***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.

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

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

Forecasting Errors & Accuracy (Basics)
Relative Accuracy

t	Y_t	$\widehat{Y0}_t$	$\widehat{Y1}_t$	$\widehat{Y2}_t$	$e0_t$	$e1_t$	$e2_t$
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			

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

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

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

- ***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 $R\_MAE_0 = \frac{MAE_0}{MAE_0} = \frac{8,7}{8,7} = 1{,}000$

Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

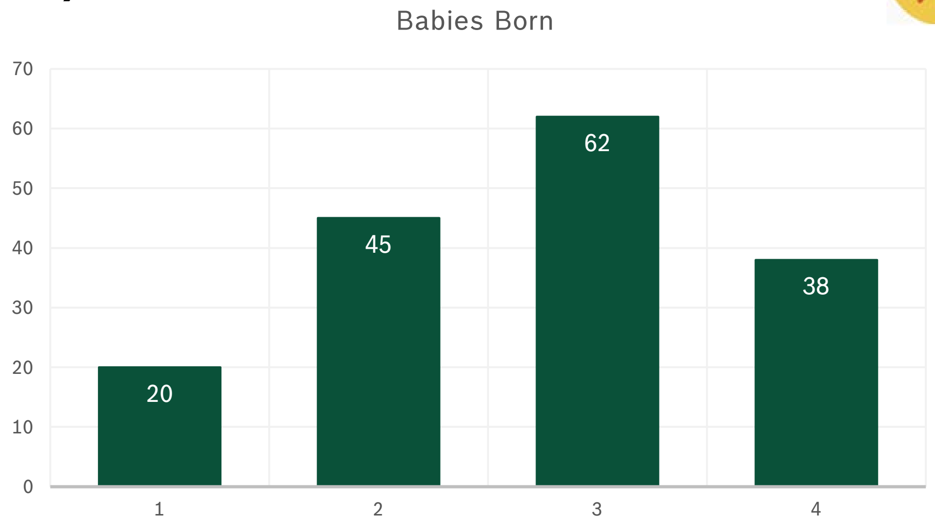
Forecasting Errors & Accuracy (Basics)
Minimizing the error

- General rule:
  - We select one (or more) error indexes
  - We try to ***minimize the selected error indexes*** over the know observations (data points)
  - Use a **benchmark** method (simple) and estimate **relative errors**.
- Be careful!
  - Don't use the same data set for:
    - Obtaining the model, and for
    - Estimating the model accuracy
  - This practice is almost always exaggerating the accuracy

Testing is needed against unseen observations.

Forecasting Errors & Accuracy (Basics)
Minimizing the error - Example

- Suppose the number of babies born over four weeks:
- Selected models:
  - Naive:
  - Mean:
  - Model A:



$$Babies(t) = Babies(t-1)$$

Week	Babies Born	Naïve Forecast	$\lvert e_t \rvert$
1	20		
2	45	20	25
3	62	45	17
4	38	62	24
		Average =	**22**

$$Babies(t) = \frac{1}{N} * \sum_1^N Babies(i)$$

Week	Babies Born	Mean Forecast	$\lvert e_t \rvert$
1	20		
2	45	20	25
3	62	33	30
4	38	42	4
		Average =	**20**

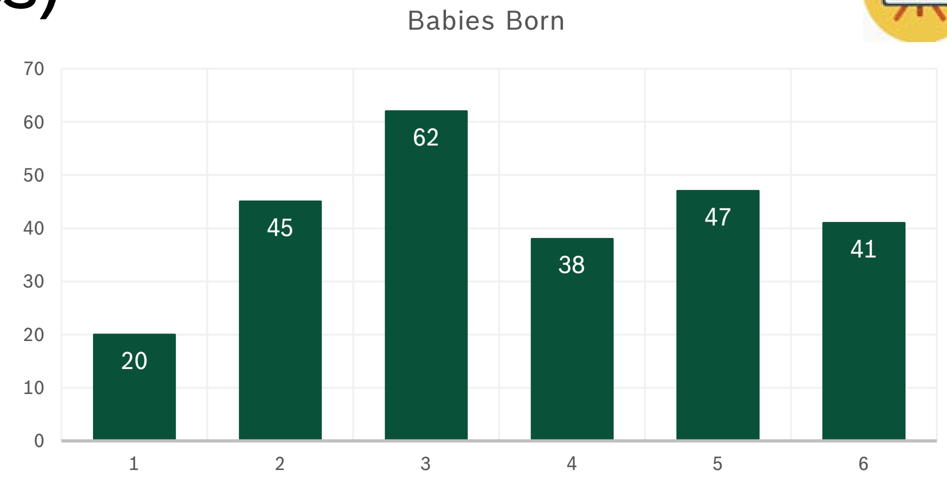
$$Babies(t) = -5.5 * t^3 + 29 * t^2 - 23.5 * t + 20$$

Week	Babies Born	Model Forecast	$\lvert e_t \rvert$
1	20	20	0
2	45	45	0
3	62	62	0
4	38	38	0
		Average =	**0**

- Model A has the lowest MAE = 0 .
- We have a perfect model!

Forecasting Errors & Accuracy (Basics)

Minimizing the error - Example



- ***No!*** We are misguided.
  - The model only represents the pattern of randomness.
  - This is **overfitting**!

$$Babies(t) = Babies(t-1)$$

Week	Babies Born	Naïve Forecast	$\lvert e_t \rvert$
1	20		
2	45	20	25
3	62	45	17
4	38	62	24
5	47	**38**	9
6	41	**47**	6
		Average =	**16**

$$Babies(t) = \frac{1}{N} * \sum_1^N Babies(i)$$

Week	Babies Born	Mean Forecast	$\lvert e_t \rvert$
1	20		
2	45	20	25
3	62	33	30
4	38	42	4
5	47	**41**	6
6	41	**42**	1
		Average =	**13**

$$Babies(t) = -5.5 * t^3 + 29 * t^2 - 23.5 * t + 20$$

Week	Babies Born	Model Forecast	$\lvert e_t \rvert$
1	20	20	0
2	45	45	0
3	62	62	0
4	38	38	0
5	47	**-60**	107
6	41	**-265**	306
		Average =	**69**

➢ ***Keep in mind: A very small MAE does not guarantee that our model is the best to use!***

5. FORECASTING METHODS

BOSCH

Forecasting Methods
Forecasting

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

- These methods are usually complementary.
  - Statistical methods are ***rigorous but consistent***
    - they can analyze large datasets of information very quickly.
  - People 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

- ***Advantages***:
  - Can be applied directly.
  - ***Relatively accurate***, taken into account the Confidence Intervals.
  - Does not require technical and statistical knowledge (when used as a "black box" from Managers).
  - Require ***significant little time*** and few computational resources.

- ***Disadvantages***:
  - Assumption: the behaviour of the given timeserie ***will continue*** in the future, which is not always true.
  - Does not take into account specific special events and actions that may take place (like athletic events, holidays, etc.)
  - Some methods require a large number of historical observations.

Forecasting Methods
Forecasting: Judgmental methods

- **Judgmental 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

- ***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. NAÏVE METHODS

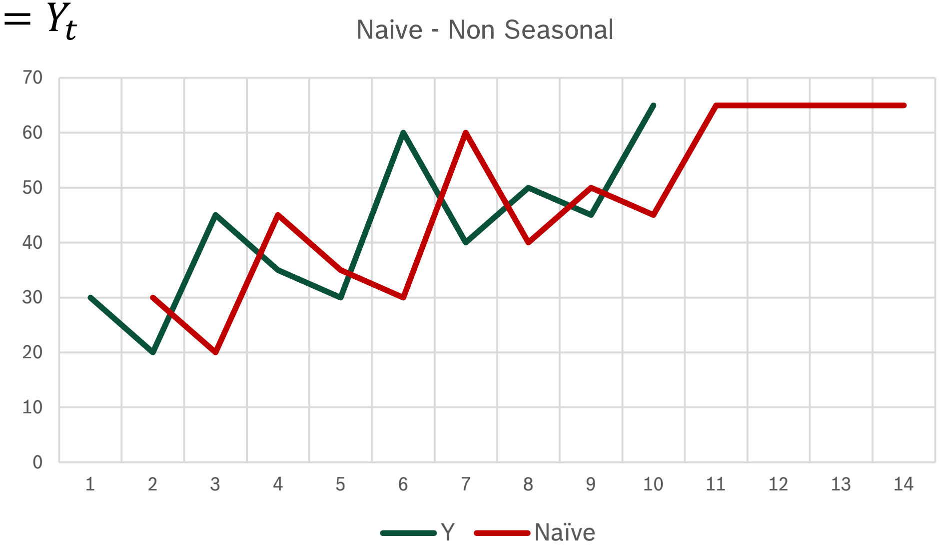
“A simple and surprisingly effective random walk towards success”

Naïve Methods
Naïve 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 (year)	Y (data)	Forecast
2010	30	
2011	20	30
2012	45	20
2013	35	45
2014	30	35
2015	60	30
2016	40	60
2017	50	40
2018	45	50
2019	65	45
2020		65
2021		65
2022		65
2023		65

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

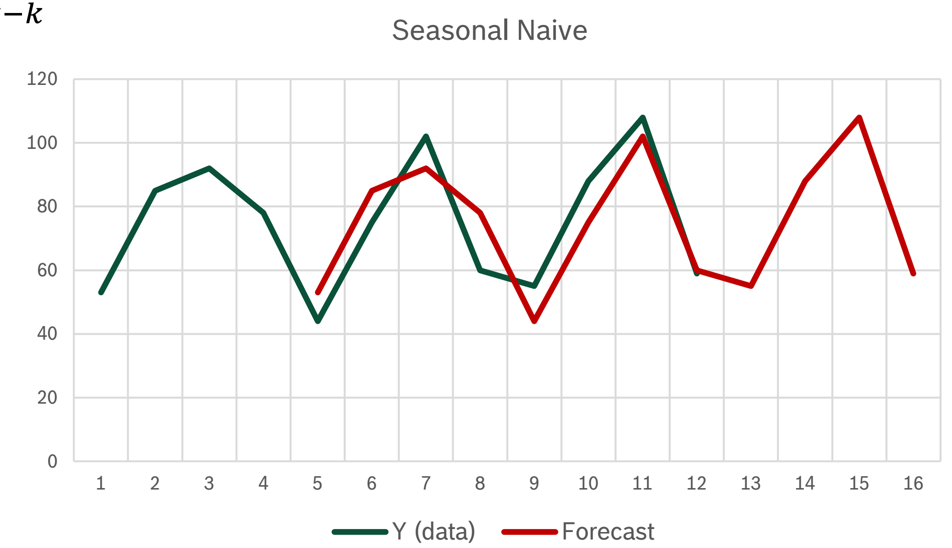


Naïve Methods
Seasonal Naïve

- ***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 (quarters)	Y (data)	Forecast
2016 A	53	
2016 B	85	
2016 C	92	
2016 D	78	
2017 A	44	53
2017 B	75	85
2017 C	102	92
2017 D	60	78
2018 A	55	44
2018 B	88	75
2018 C	108	102
2018 D	59	60
2019 A		55
2019 B		88
2019 C		108
2019 D		59

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



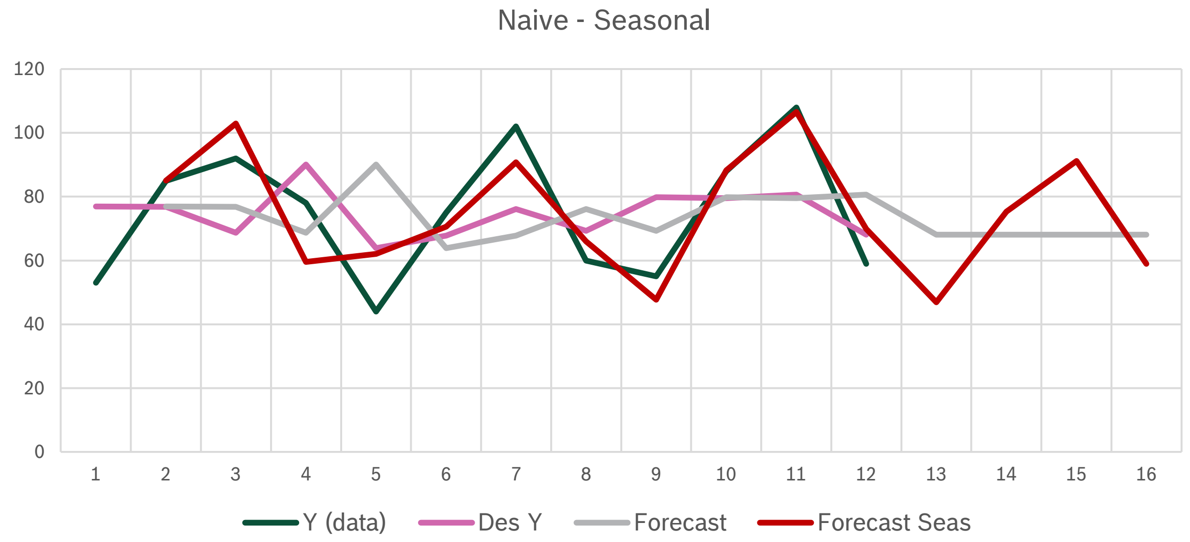
Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

BOSCH

Naïve Methods
Naïve 2

- ***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 (data)	DMA(4)	Des-Y	Forecast	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

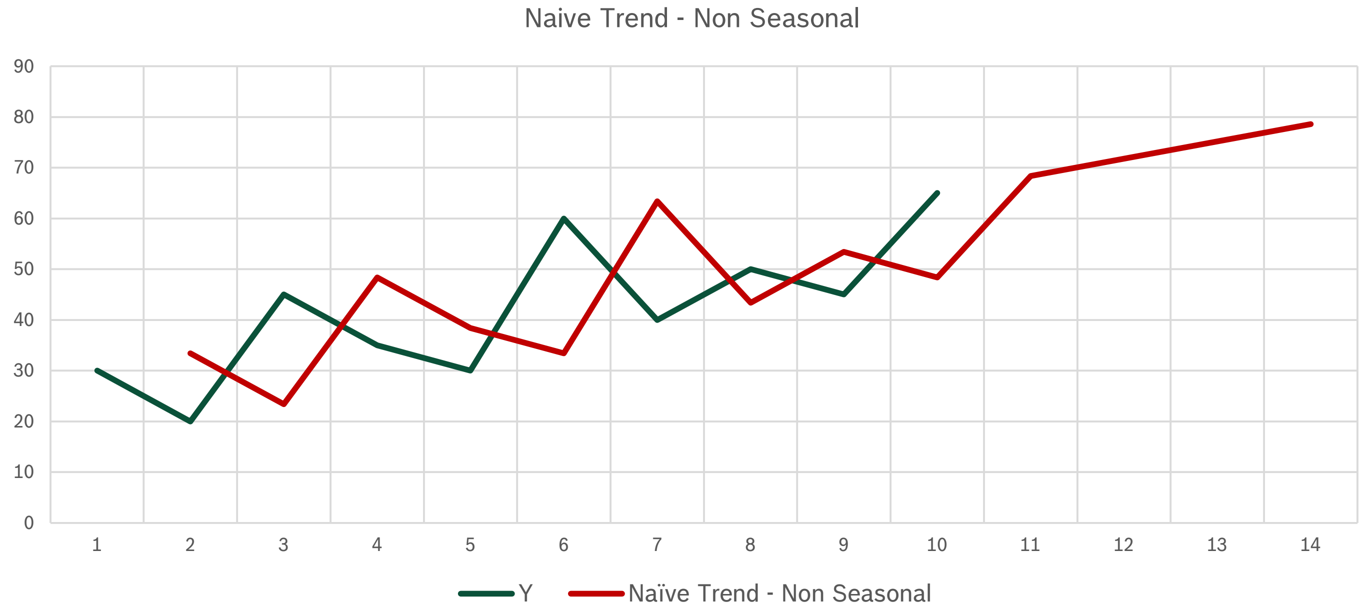


Naïve Methods
Naïve Trend

- ***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

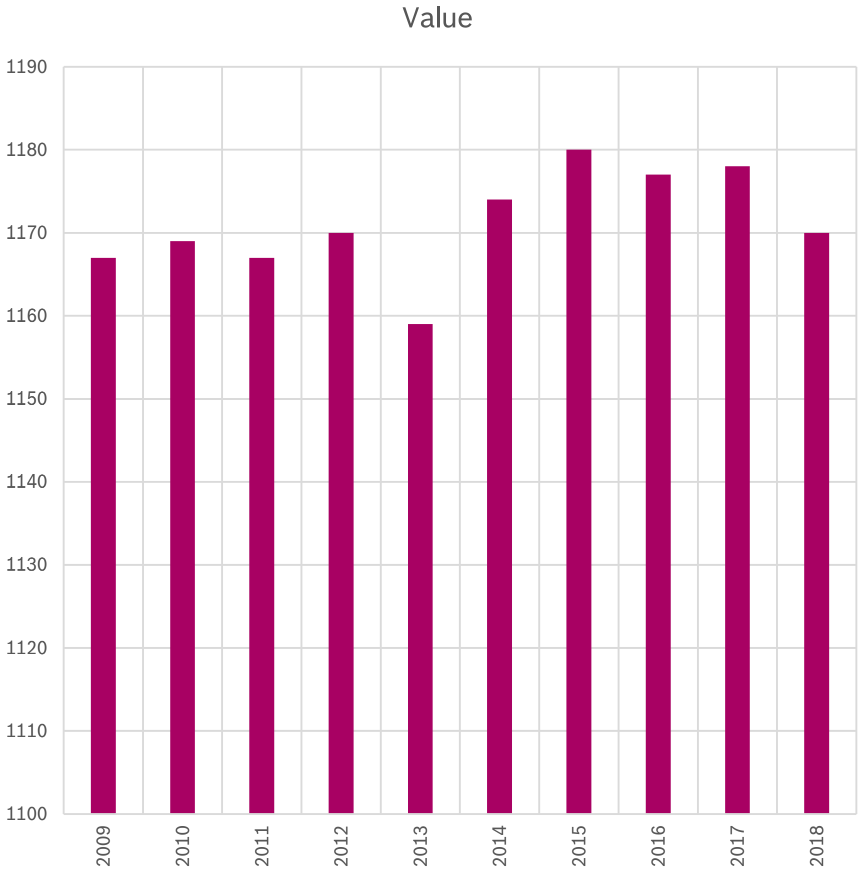
(slope = 3,3939)



Naïve Methods

Try to forecast with a method: Naive

Data	
Year	**Value**
2009	1167
2010	1169
2011	1167
2012	1170
2013	1159
2014	1174
2015	1180
2016	1177
2017	1178
2018	1170



- What you can observe?
  - 10 data points
  - Yearly data
  - No trend over the years (stationary timeserie)
  - No “strange” or “missing” values
  - No significant trend changes during years
  - Average (mean value) = 1171
- Let’s try to forecast:
  - Value for 2019
  - Horizon: 1 year (Short-term)

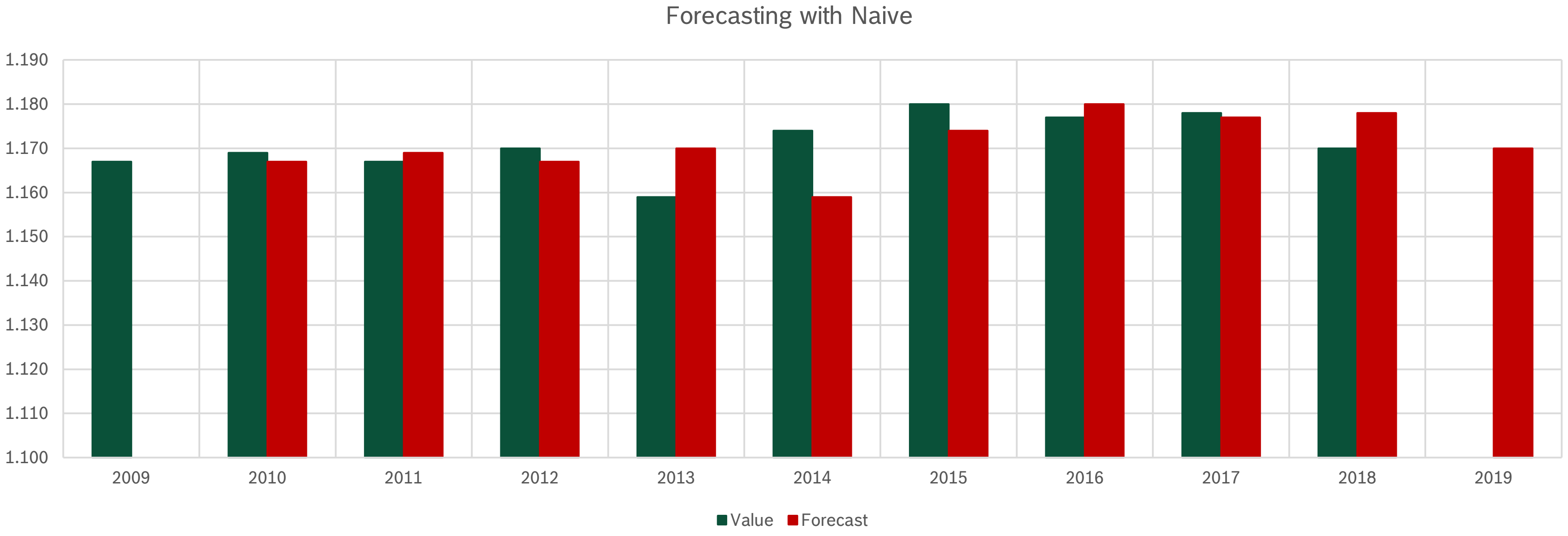
BOSCH

Naïve Methods
Try to forecast with a method: Naive

- ***Naïve:*** The simplest method for estimation
  - The forecast for a period (t+1) is equal with the previous (t) historical observation

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

Data		
Year	**Value**	**Forecast**
2009	1.167	
2010	1.169	1.167
2011	1.167	1.169
2012	1.170	1.167
2013	1.159	1.170
2014	1.174	1.159
2015	1.180	1.174
2016	1.177	1.180
2017	1.178	1.177
2018	**1.170**	1.178
2019		**1.170**

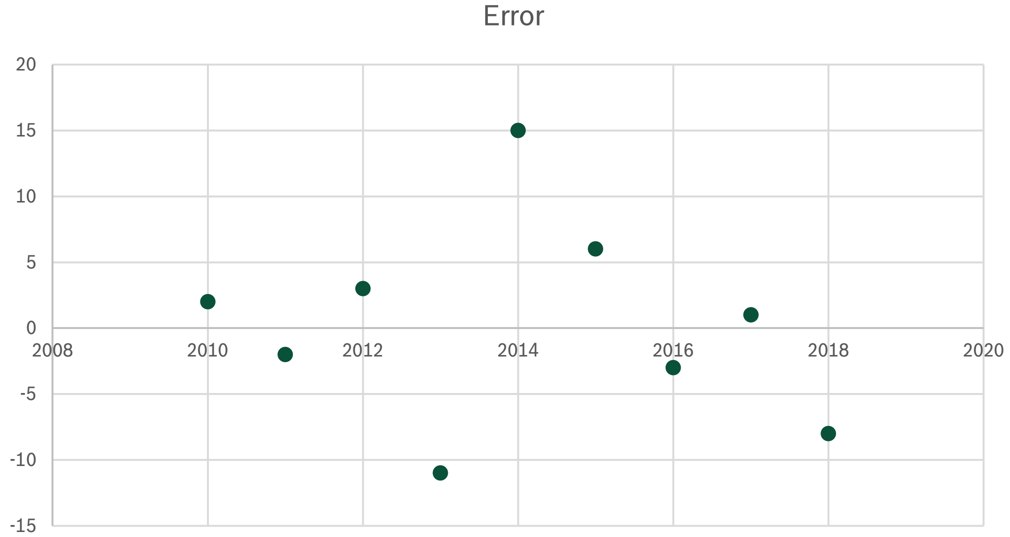


Naïve Methods
Try to forecast with a method: Naive

- ***Evaluate*:**
  - How good the forecast is (model evaluation)?
    - Let's calculate the error

Data			
Year	**Value**	**Forecast**	**Error**
2009	1.167		
2010	1.169	1.167	2
2011	1.167	1.169	-2
2012	1.170	1.167	3
2013	1.159	1.170	-11
2014	1.174	1.159	15
2015	1.180	1.174	6
2016	1.177	1.180	-3
2017	1.178	1.177	1
2018	1.170	1.178	-8
2019		1.170	

- ***Results*:**
  - Some positive & some negative errors.
  - Mean Error (ME): 0,3
    - No bias!



Naïve Methods

Try to forecast with a method: Naive

- ***Evaluate*:**
  - How good the forecast is (model evaluation)?
    - Let's calculate the ***absolute*** error

Data			
Year	**Value**	**Forecast**	**Abs. Error**
2009	1.167		
2010	1.169	1.167	2
2011	1.167	1.169	2
2012	1.170	1.167	3
2013	1.159	1.170	11
2014	1.174	1.159	15
2015	1.180	1.174	6
2016	1.177	1.180	3
2017	1.178	1.177	1
2018	1.170	1.178	8
2019		1.170	

- ***Results*:**
  - Mean Absolute Error (MAE): 5,7
    - **This is our benchmark!**
  - Pretty good, compared to mean value (1.171)

- Is this the best method?
  - **We need to try more methods**

7. MOVING AVERAGES

"A technique for calculating the overall trend in a data set"

Moving Averages
Definition

- 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)

Moving Averages
Simple Moving Average – SMA

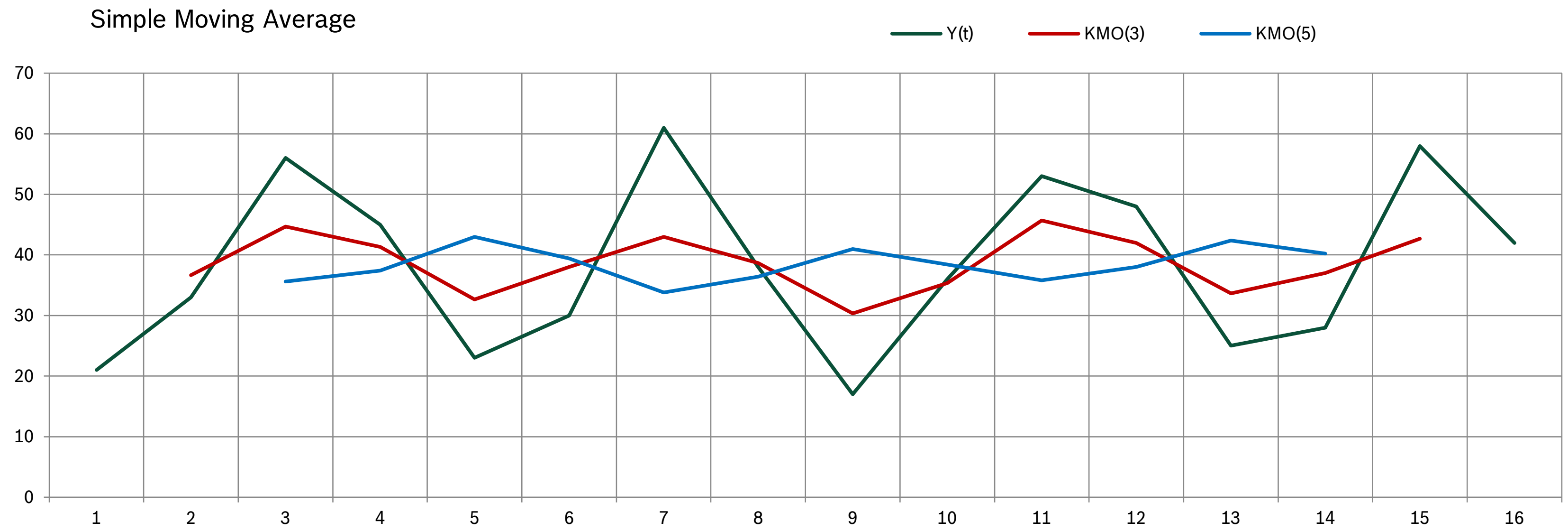
- SMA: The ***unweighted mean*** of the previous n data points.
  - However, in science and engineering the mean is normally taken from an **equal number of data** on either side of a central value.
- The scope is to eliminate part of the random component, and estimate the Trend-Cycle component.
- ***Hypothesis***: "*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.

Moving Averages
Simple Moving Average – SMA

	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		

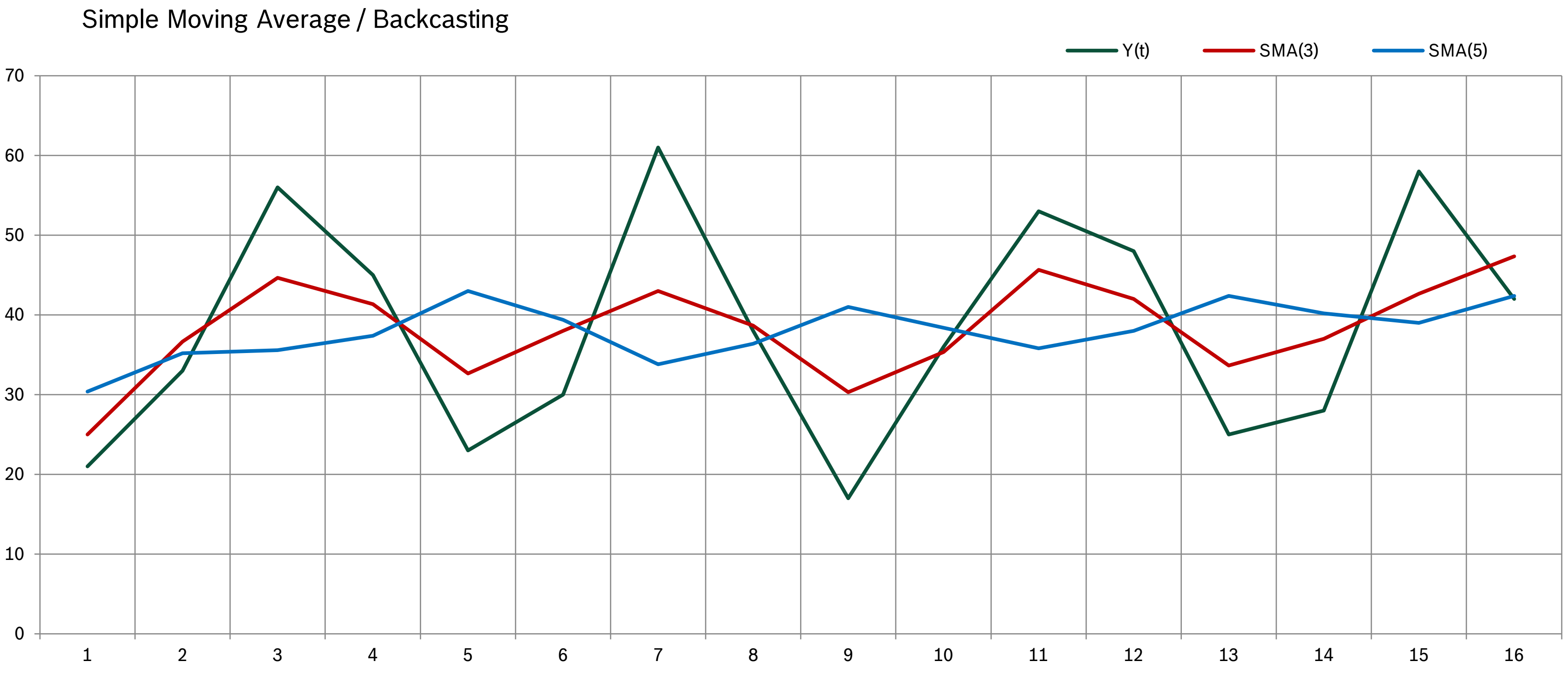


$$SMA(5)_3 = \frac{1}{5} \times \sum_{i=-2}^{2} Y_{3+i} = \frac{1}{5} \times (21 + 33 + 56 + 45 + 23) = 35{,}6$$

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

Moving Averages
Simple Moving Average – SMA with Backcasting

	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).

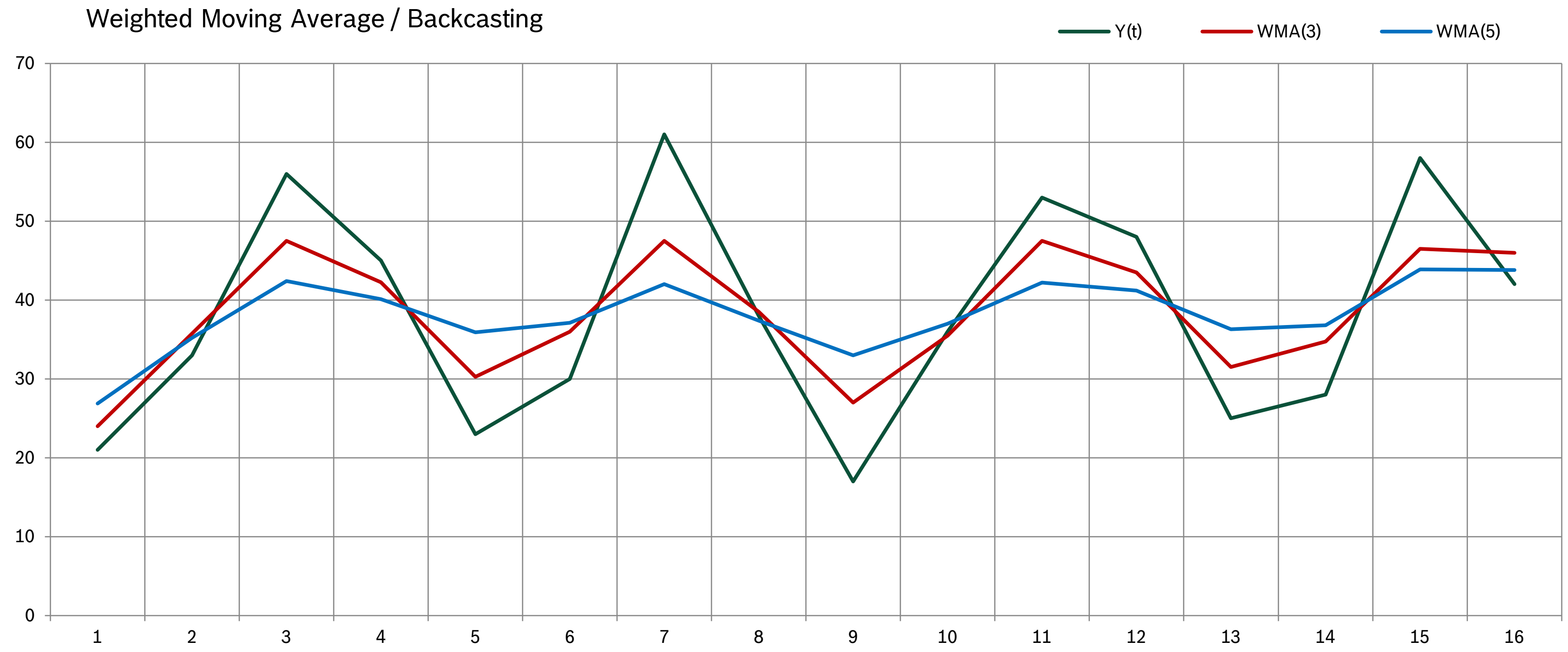
Moving Averages
Weighted Moving Average – WMA

- 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.

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

Moving Averages
Weighted Moving Average – WMA Example

	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**		



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

BOSCH

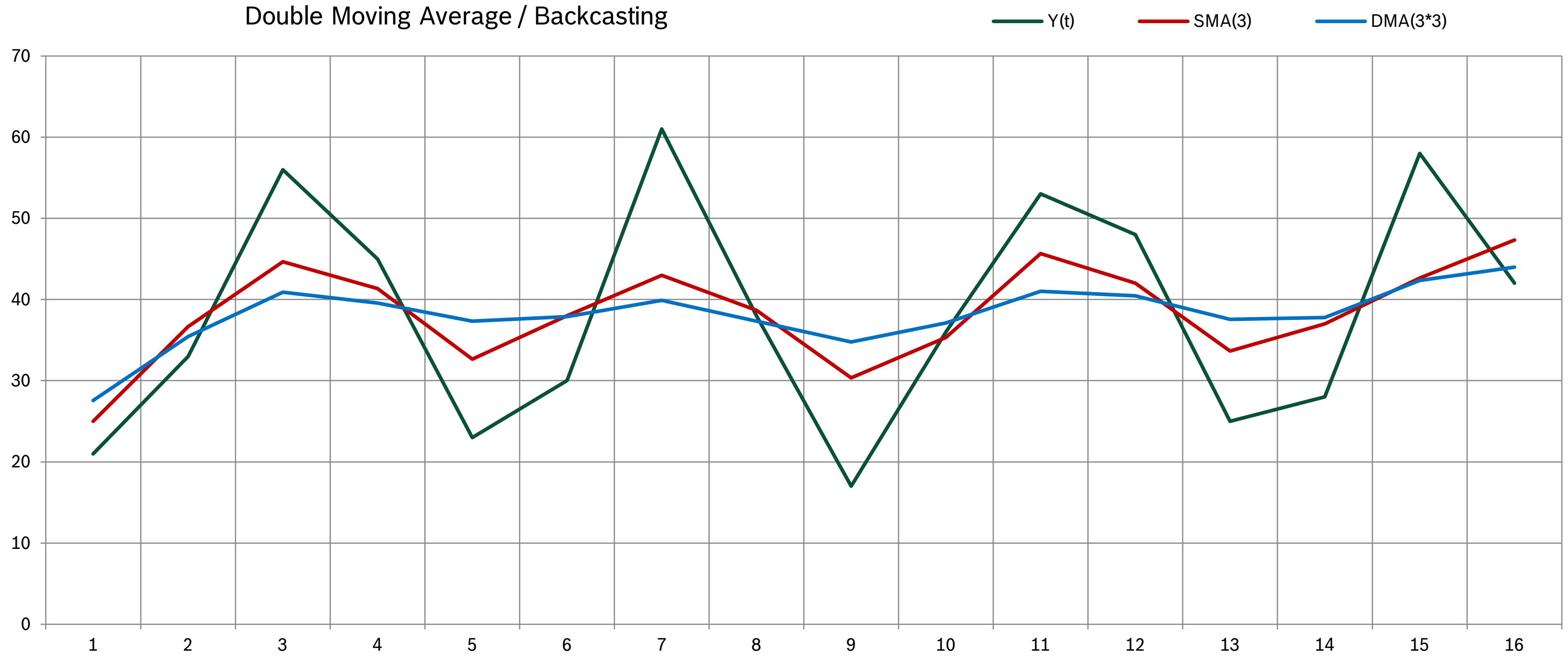
Moving Averages
Double Moving Average – DMA

- 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(3x3): is a SMA(3) on an SMA(3)

Moving Averages

Double Moving Average – DMA Example

	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		



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

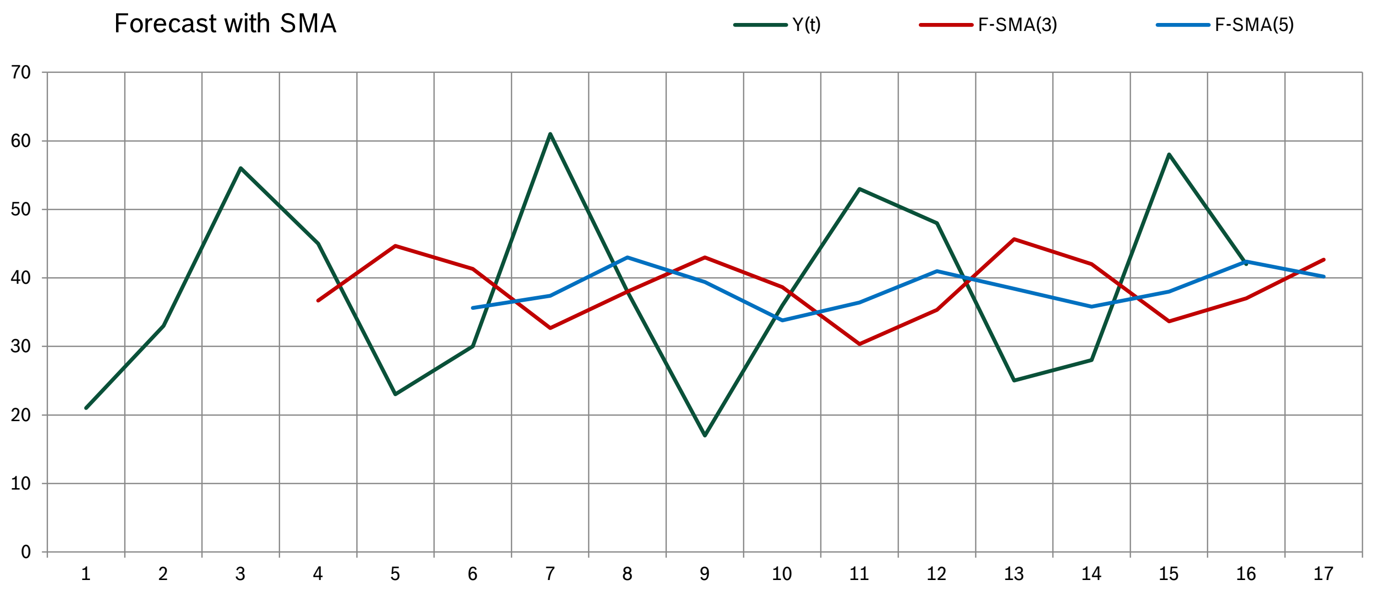
$$DMA(3x3)_4 = \frac{1}{3} \times (SMA(3)_3 + SMA(3)_4 + SMA(3)_5)) = \frac{1}{3} \times (44{,}67 + 41{,}33 + 32{,}67) = 39{,}56$$

Moving Averages

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**

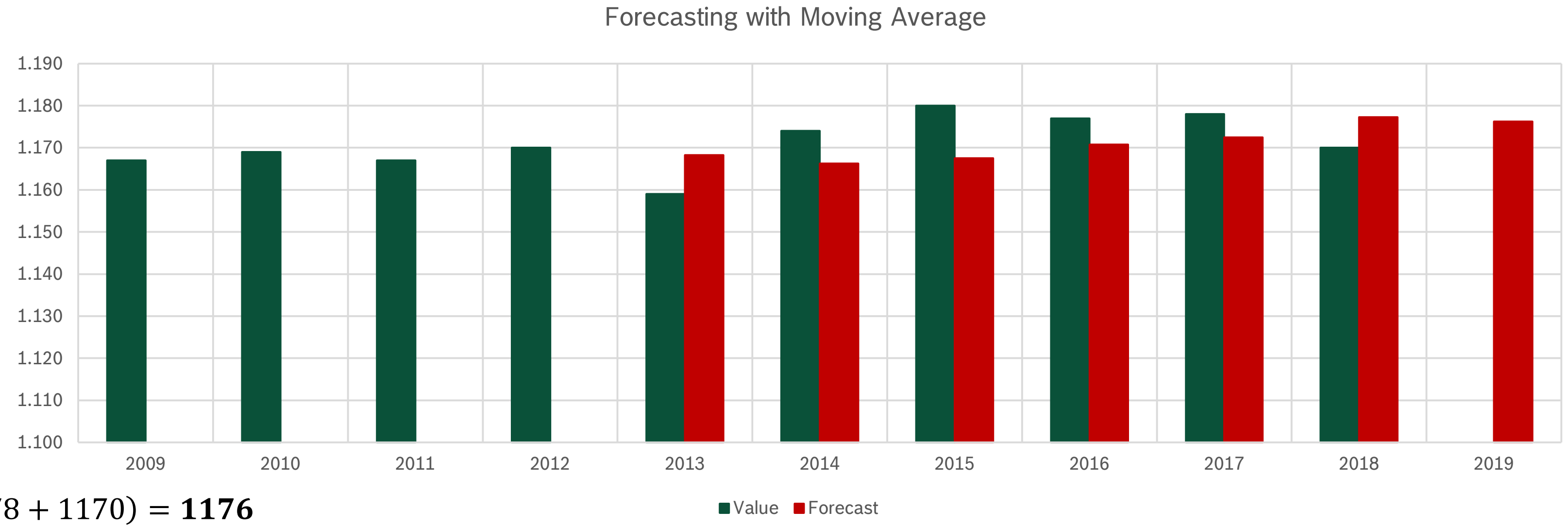


Moving Averages
Try to forecast with a method: SMA

- ***Moving Average:*** The unweighted mean of previous data points
  - The forecast for a period (t+1) is equal with the average of the previous **m** historical observation

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

Data		
Year	**Value**	**Forecast**
2009	1.167	
2010	1.169	
2011	1.167	
2012	1.170	
2013	1.159	1.168
2014	1.174	1.166
2015	**1.180**	1.168
2016	**1.177**	1.171
2017	**1.178**	1.173
2018	**1.170**	1.177
2019		**1.176**



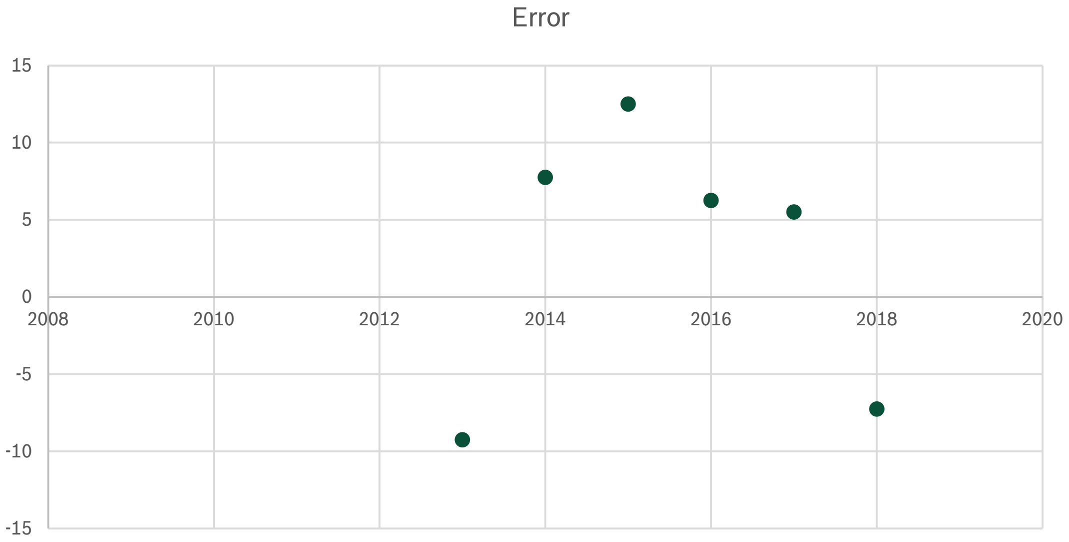
$$SMA(4)_{2019} = \frac{1}{4} * (1180 + 1177 + 1178 + 1170) = \mathbf{1176}$$

Moving Averages
Try to forecast with a method: SMA

- ***Evaluate*:**
  - How good the forecast is (model evaluation)?
    - Let's calculate the error

Data			
Year	**Value**	**Forecast**	**Error**
2009	1.167		
2010	1.169		
2011	1.167		
2012	1.170		
2013	1.159	1.168	-9
2014	1.174	1.166	8
2015	1.180	1.168	13
2016	1.177	1.171	6
2017	1.178	1.173	6
2018	1.170	1.177	-7
2019		1.176	

- ***Results*:**
  - More positive & less negative errors.
  - Mean Error (ME): 2,6
    - Some positive bias



Moving Averages
Try to forecast with a method: SMA

- ***Evaluate*:**
  - How good the forecast is (model evaluation)?
    - Let's calculate the ***absolute*** error

Data			
Year	**Value**	**Forecast**	**Abs. Error**
2009	1.167		
2010	1.169		
2011	1.167		
2012	1.170		
2013	1.159	1.168	9
2014	1.174	1.166	8
2015	1.180	1.168	13
2016	1.177	1.171	6
2017	1.178	1.173	6
2018	1.170	1.177	7
2019		1.176	

- ***Results*:**
  - Mean Absolute Error (MAE): **8,1**
  - Naive (MAE): **5,7**
  - Relative error: 8,1/5,7 = **1,421** ☹

- Best method so far:
  - Naïve was better (MAE = 5.7)

Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

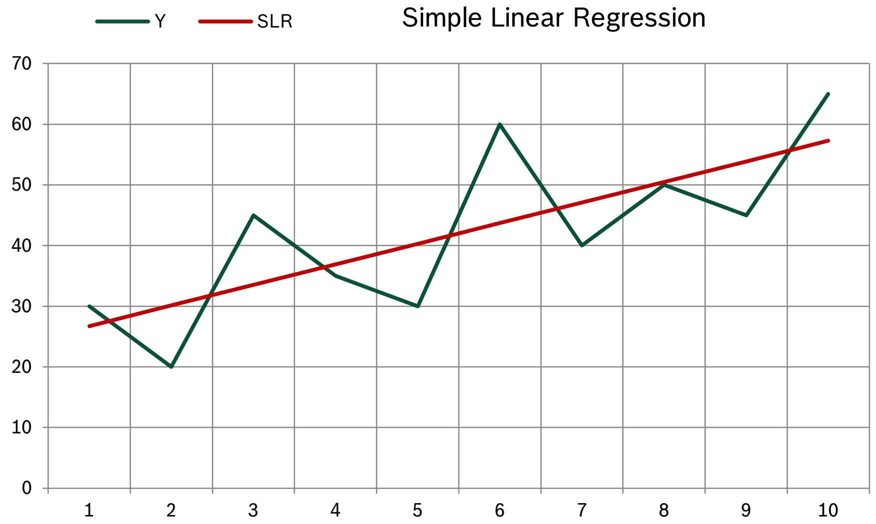
8. LINEAR REGRESSION

Linear Regression
Definition

- 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.

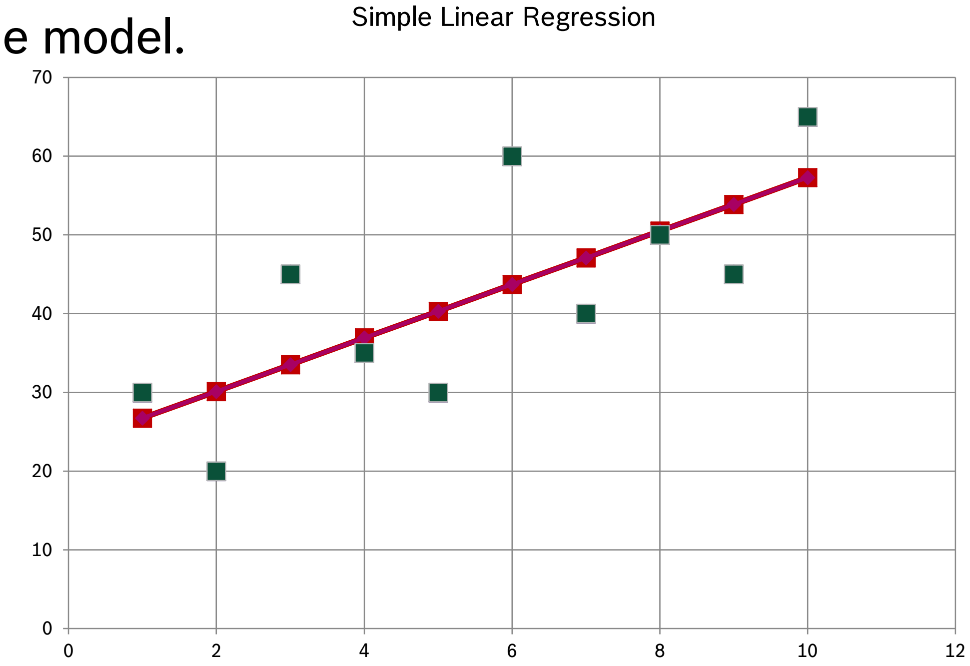
Linear Regression
Definition

- ***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
  - 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 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.



Linear Regression
Definition

- 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.



Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

BOSCH

Linear Regression
Equations

- 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***.

Linear Regression
Correlation Coefficient

- ***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.

Linear Regression
Correlation Coefficient

- 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:

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

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

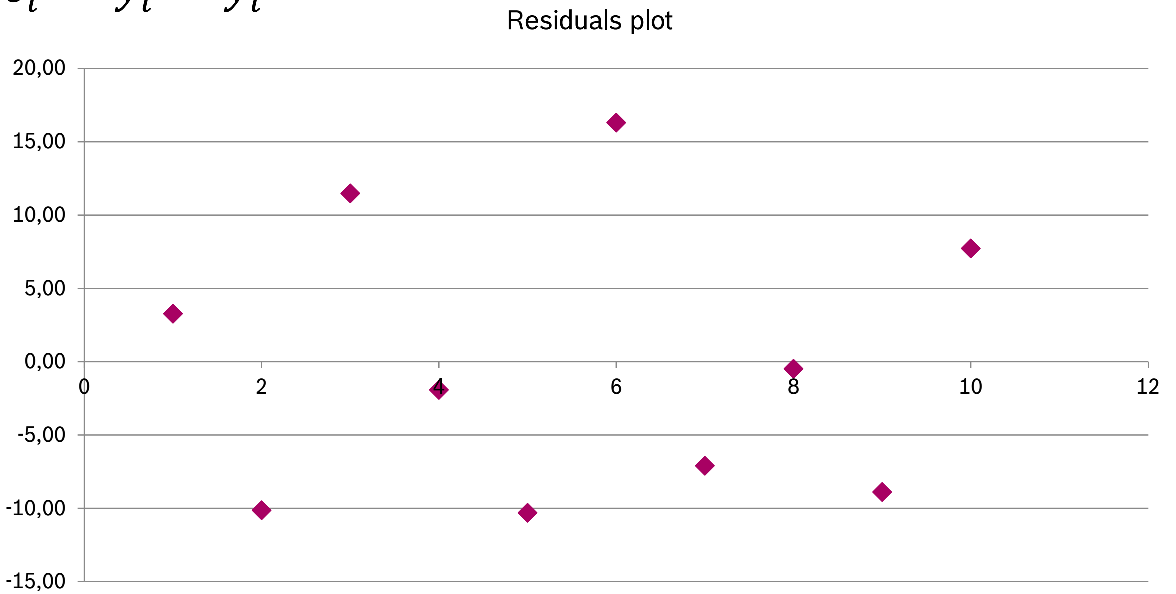
- It is obvious that ***correlation and regression are strongly linked***.
  - The advantage of a regression model over correlation is that is asserts a predictive relationship between the two variables and quantifies this in a way that is useful for forecasting.

Linear Regression
Residual

- ***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.

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

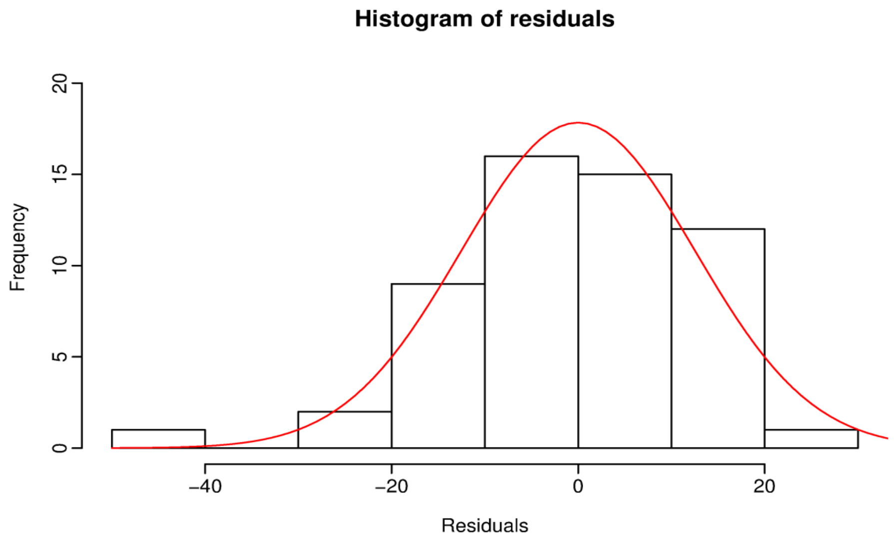
- A non-random pattern may indicate that:
  - a non-linear relationship may be required, or
  - some heteroscedasticity is present (residuals have non constant variance), or
  - there is some unexplained serial correlation.



Linear Regression
Residual

- ***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.



Linear Regression
Evaluation

- ***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.

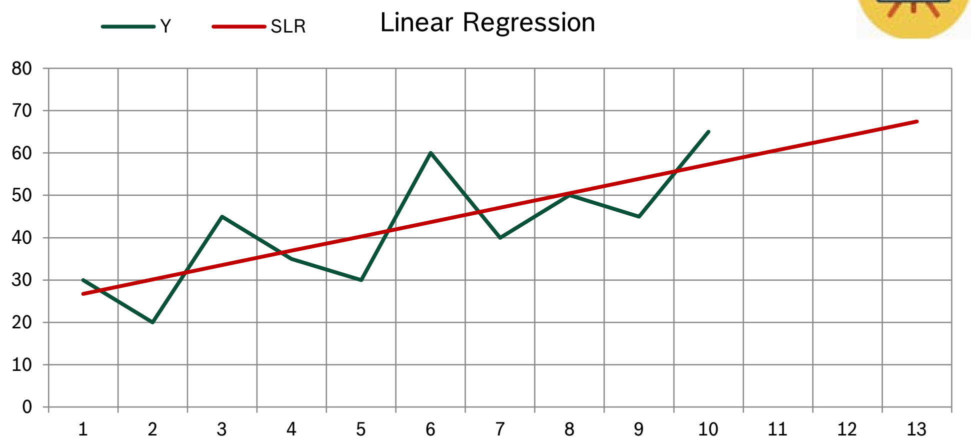
Linear Regression
How to use

- 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.

Linear Regression
Example

Data		Numerator			Denominator				
X	Y	A $X_t-\bar{X}$	B $Y_t-\bar{Y}$	A * B	$(X_t-\bar{X})^2$	$(Y_t-\bar{Y})^2$	$(\hat{Y}_t-\bar{Y})^2$	$(Y_t-\hat{Y})^2$	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								



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

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

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

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

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

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

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

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

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

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

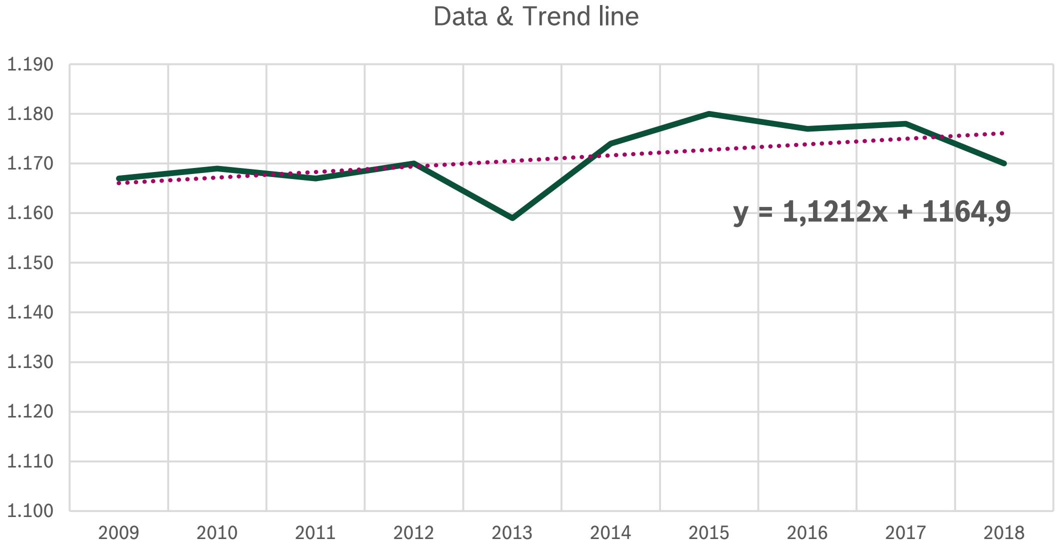
$$r_{XY} = \frac{COV_{XY}}{\sqrt{COV_{YY} * COV_{XX}}} = \frac{COV_{XY}}{S_Y * S_X} = 0{,}735$$

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

Linear Regression
Try to forecast with a method: LR

- ***Linear Regression:*** A fitted line though all historical data points
  - fits a straight line in such a way that makes the sum of squared differences (the distances between the points of the data sets and the fitted line) as small as possible.

Data	
Year	**Value**
2009	1.167
2010	1.169
2011	1.167
2012	1.170
2013	1.159
2014	1.174
2015	1.180
2016	1.177
2017	1.178
2018	1.170



Data				
Year	Value	Line	Difference	Square Diff
2009	1.167	1.166	1	0,89
2010	1.169	1.167	2	3,33
2011	1.167	1.168	-1	1,68
2012	1.170	1.169	1	0,34
2013	1.159	1.171	-12	133,16
2014	1.174	1.172	2	5,47
2015	1.180	1.173	7	52,10
2016	1.177	1.174	3	9,59
2017	1.178	1.175	3	8,86
2018	1.170	1.176	-6	37,77
		SUM	**0,0**	**253,2**
		Average	**0,0**	**25,3**

Slope	1,121
Intercept	1.164,9

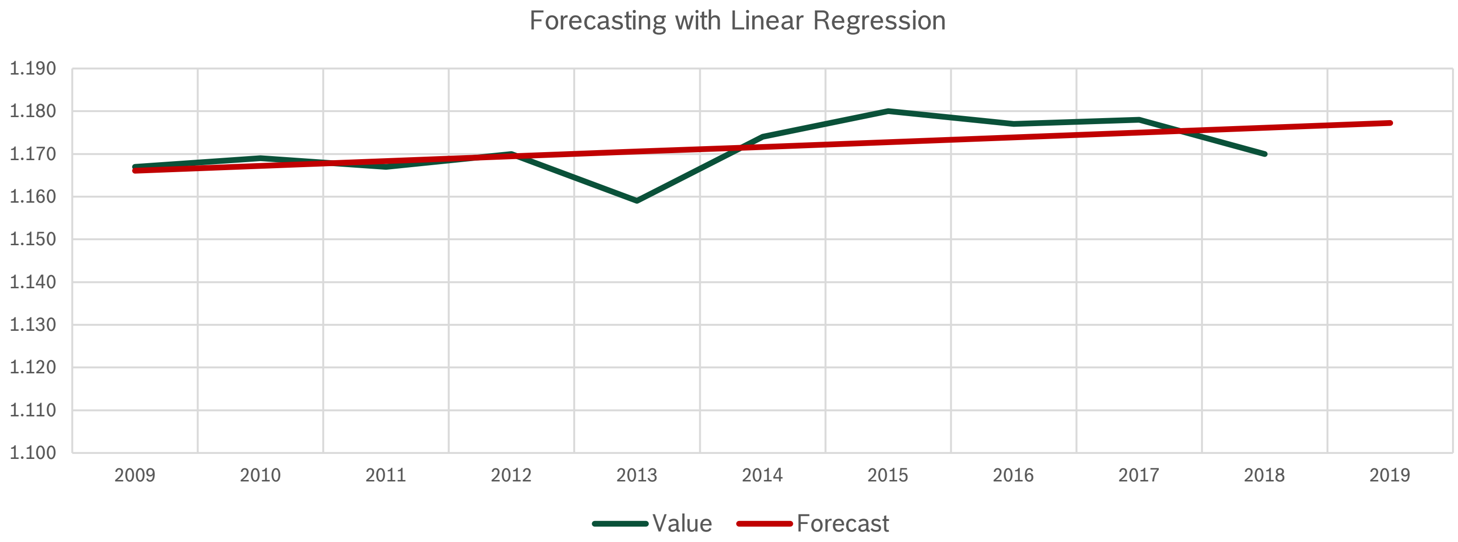
Linear Regression

Try to forecast with a method: LR

▶ ***Estimate Forecasts***

	Data		
	Year	**Value**	**Forecast**
1	**2009**	1.167	1.166
2	**2010**	1.169	1.167
3	**2011**	1.167	1.168
4	**2012**	1.170	1.169
5	**2013**	1.159	1.171
6	**2014**	1.174	1.172
7	**2015**	1.180	1.173
8	**2016**	1.177	1.174
9	**2017**	1.178	1.175
10	**2018**	1.170	1.176
11	**2019**		**1.177**

Slope	1,121
Intercept	1.164,9



$$F_{2019} = Intercept + (slope * t) = 1164{,}9 + (1{,}121 * 11) = \mathbf{1177}$$

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

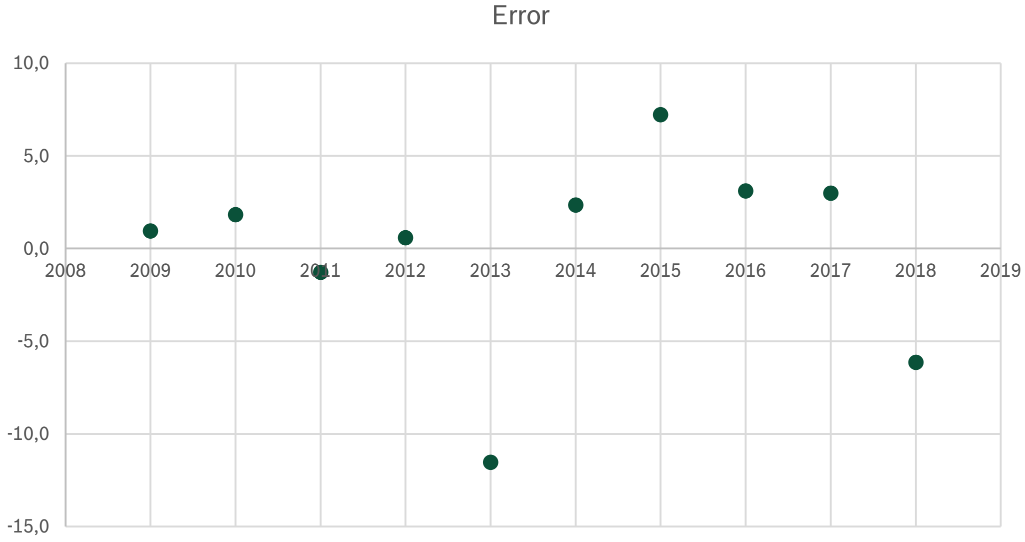
BOSCH

Linear Regression
Try to forecast with a method: LR

- ***Evaluate*:**
  - How good the forecast is (model evaluation)?
    - Let's calculate the error

Data			
Year	**Value**	**Forecast**	**Error**
2009	1.167	1.166	1
2010	1.169	1.167	2
2011	1.167	1.168	-1
2012	1.170	1.169	1
2013	1.159	1.171	-12
2014	1.174	1.172	2
2015	1.180	1.173	7
2016	1.177	1.174	3
2017	1.178	1.175	3
2018	1.170	1.176	-6
2019		1.177	

- ***Results*:**
  - More positive & less negative errors.
  - Mean Error (ME): 0,0



Linear Regression

Try to forecast with a method: LR

- ***Evaluate*:**
  - How good the forecast is (model evaluation)?
    - Let's calculate the ***absolute*** error

Data			
Year	**Value**	**Forecast**	**Abs. Error**
2009	1.167	1.166	1
2010	1.169	1.167	2
2011	1.167	1.168	1
2012	1.170	1.169	1
2013	1.159	1.171	12
2014	1.174	1.172	2
2015	1.180	1.173	7
2016	1.177	1.174	3
2017	1.178	1.175	3
2018	1.170	1.176	6
2019		1.177	

- ***Results*:**
  - Mean Absolute Error (MAE): **3,8**
  - Naive (MAE): **5,7**
  - Relative error: 3,8/5,7 = **0,667** ☺

- Best method so far:
  - Naïve was worst (MAE = 5.7)
  - Moving Average was worst (MAE = 8.1)

BOSCH

9. EXPONENTIAL SMOOTHING

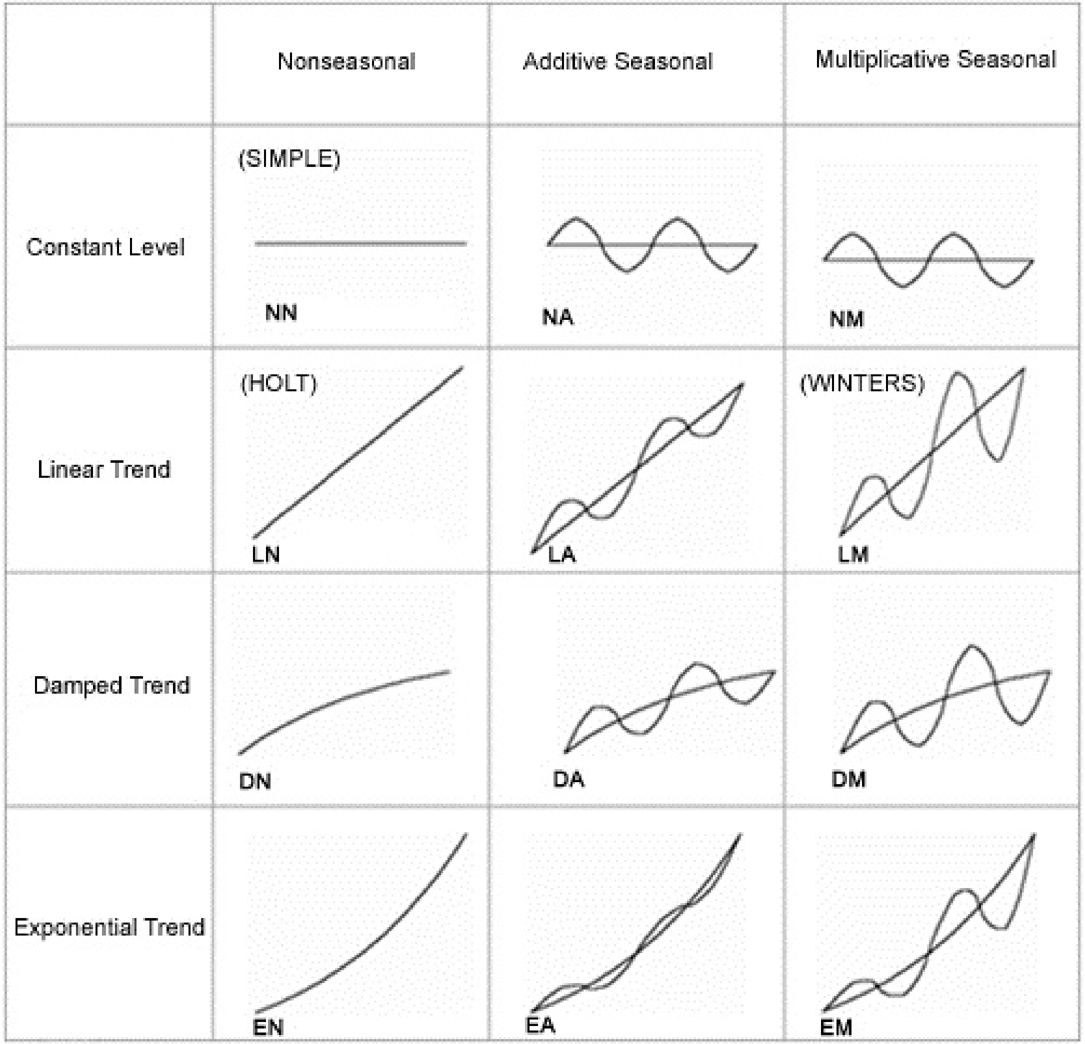
Exponential Smoothing
Definition

- 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.
- Generates ***reliable forecasts quickly*** and for a wide spectrum of timeseries
  - They perform better on data that are ***stagnating*** or ***having small growth or reduction*** during time.
    - a great advantage and of major importance to applications in industry.
- One of the most popular methods of forecasting among businessmen, mainly because of:
  - their convenience (easy to use)
  - the minimum requirement in computational time, and
  - the need of relatively few observations.

Exponential Smoothing
Definition

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



Source: Forecast Pro

Exponential Smoothing
Constant Level

Simple Exponential Smoothing (SES):

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

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

- Thus:

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

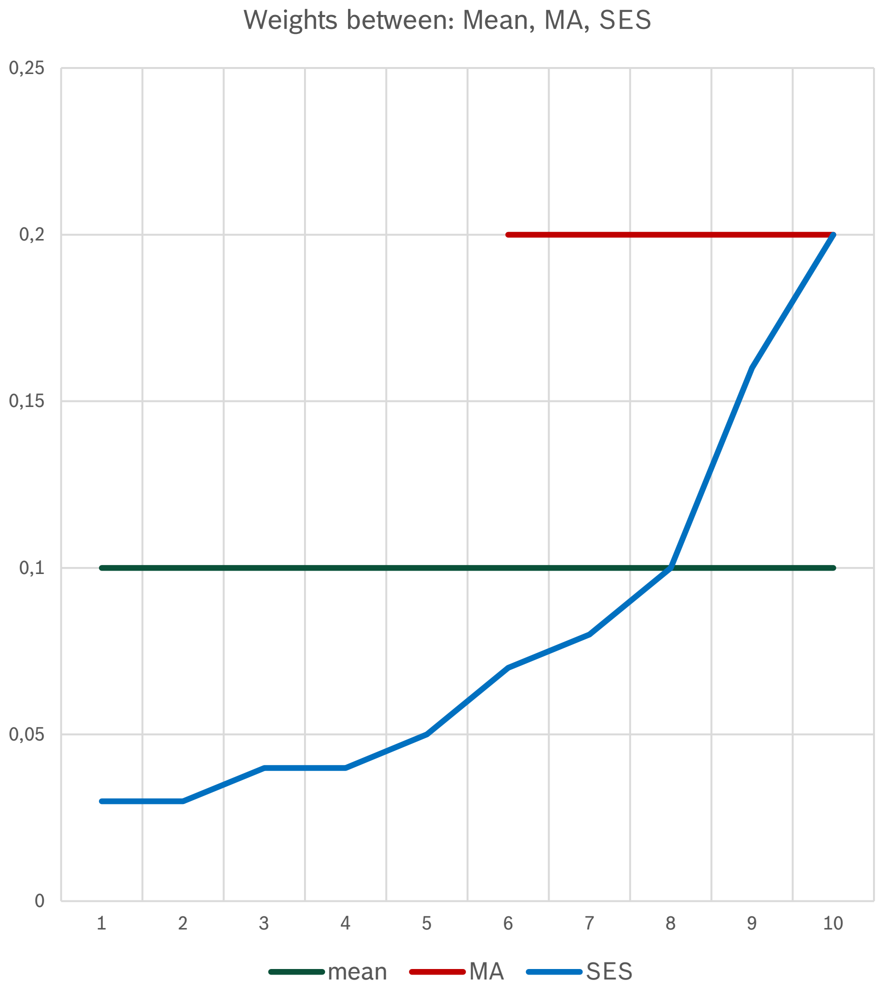
$$F_{t+1} = (a \times Y_t) + (a \times (1-a) \times Y_{t-1}) + (a \times (1-a)^2 \times Y_{t-2}) + (a \times (1-a)^3 \times Y_{t-3}) + \cdots + \big(a \times (1-a)^{(t-1)} \times Y_1\big) + \big((1-a)^t \times F_1\big)$$

BOSCH

Exponential Smoothing
Constant Level

Simple Exponential Smoothing (SES):

- The weighting factors are decreasing exponentially
- The **first** forecast F_1 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



Exponential Smoothing
Constant Level

Initialization:

- As first forecast 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.

Exponential Smoothing
Constant Level

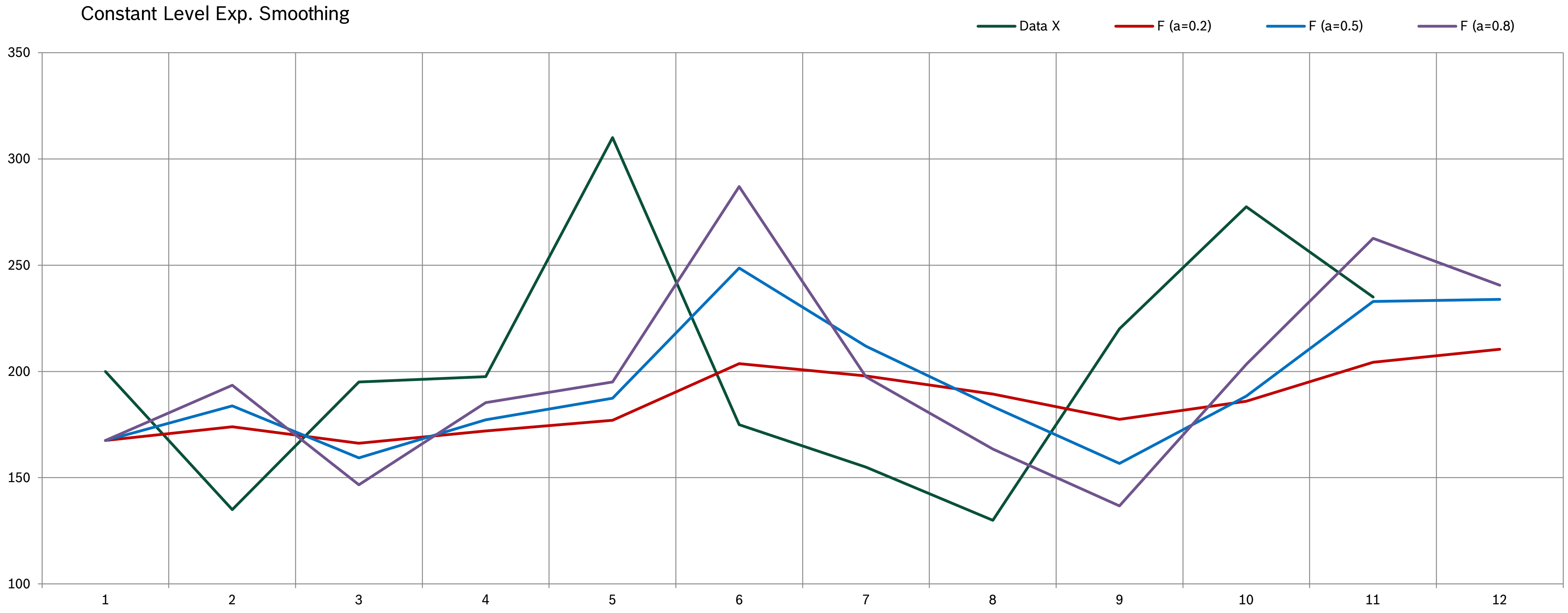
- ***Smoothing parameter: Value between 0 and 1***.
  - if the value is 1, then all forecasts are equal with the first period forecast.
  - a large value (such as 0.8) is creating a small smoothing effect.
  - a small value (such as 0.2) is creating a large smoothing effect.
  - If the value is 0, then 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.

Exponential Smoothing
Constant Level – Example

- Starting level (first forecast):
  - Average of 1st and 2nd data point values.
- Smoothing parameter:
  - $\alpha = 0.2$
  - $\alpha = 0.5$
  - $\alpha = 0.8$
- 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**	**Forecast**	**Error**	**Level S**	**Forecast**	**Error**	**Level S**	**Forecast**	**Error**	**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%**		

Exponential Smoothing
Constant Level – Example



BOSCH

Exponential Smoothing
Linear Trend

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).

Exponential Smoothing
Linear Trend

***Initialization*:**

- Linear Regression must be performed, in order to estimate the linear regression equation.
- the initial level $\boldsymbol{S_0}$ is equal to the **constant parameter** of the regression.
- the initial trend $\boldsymbol{T_0}$ 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)

Exponential Smoothing
Linear Trend – Example

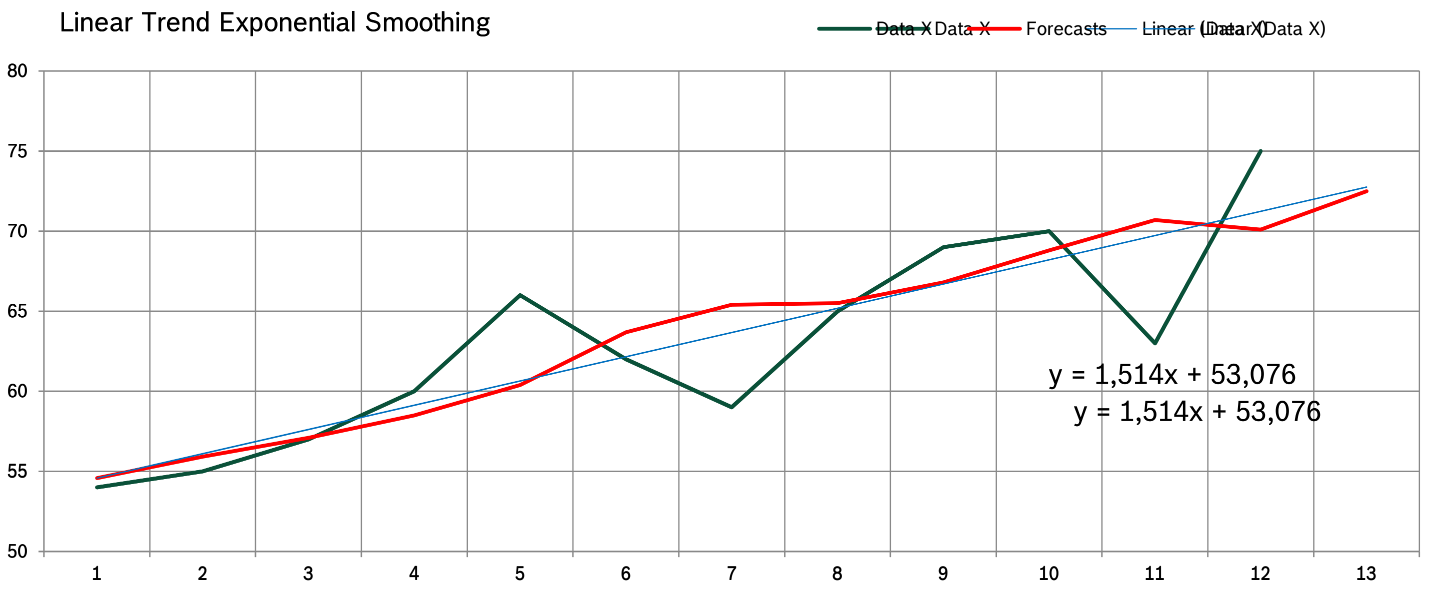
$$S_1 = S_0 + T_0 + (a * e_1)$$

$$T_1 = T_0 + (\beta * e_1)$$

$$F_1 = S_0 + T_0$$

- Starting level: Constant of regression, starting trend: Slope of regression.
- Smoothing parameter: a = 0.2, β = 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**	**Forecasts**	**Error e**	**Level S**	**Trend T**
1	54	54,59	-0,59	54,47	1,46
2	55	55,93	-0,90	55,70	1,40
3	57	57,10	-0,10	57,10	1,40
4	60	58,50	1,50	58,80	1,60
5	66	60,40	5,60	61,50	2,20
6	62	63,70	-1,70	63,40	2,00
7	59	65,40	-6,40	64,10	1,40
8	65	65,50	-0,50	65,40	1,40
9	69	66,80	2,20	67,20	1,60
10	70	68,80	1,20	69,00	1,70
11	63	70,70	-7,70	69,20	0,90
12	75	70,10	4,90	71,10	1,40
13		72,50		72,50	1,40
		73,90		73,90	1,40
		75,30		75,30	1,40
		MAPE:	4,33%		



Exponential Smoothing
Damped Trend

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 *α, β, φ* must be estimated by minimizing (usually) the mean squared error (MSE).

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

Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

Exponential Smoothing
Damped Trend

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)

Exponential Smoothing
Damped Trend

Initialization:

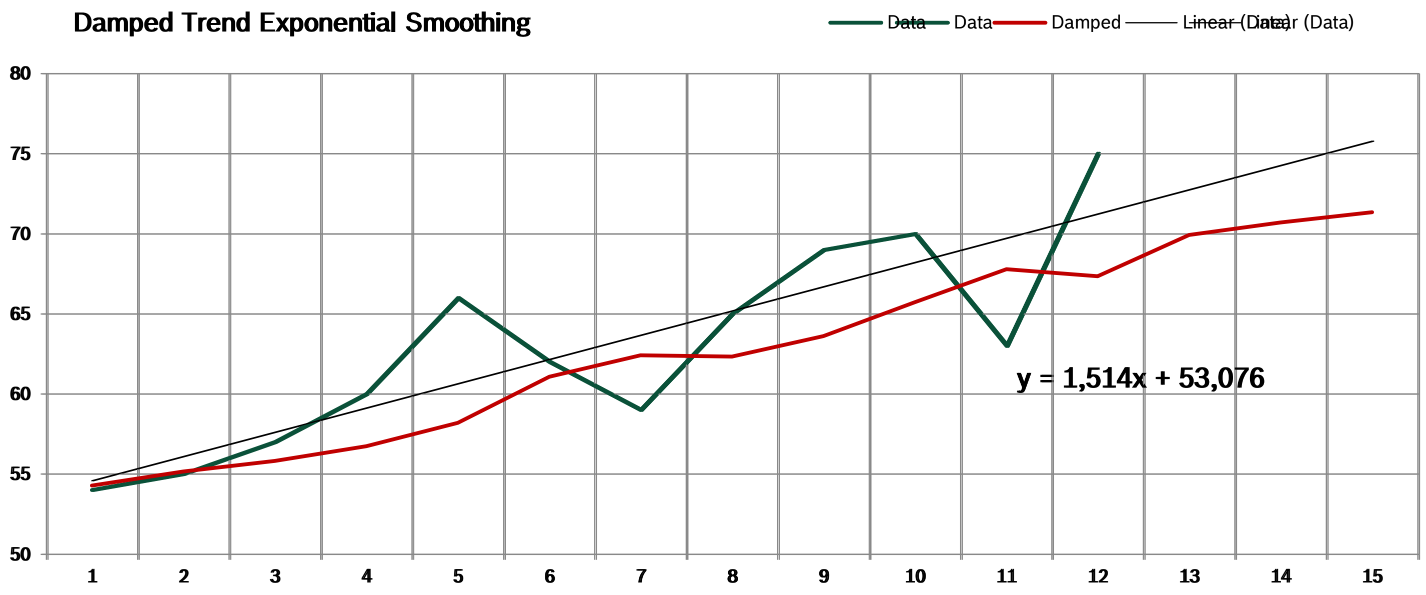
- Damping factor φ:
  - $\varphi = 0$, for the Constant Level
  - $\varphi < 1$, for the Damped Trend
  - $\varphi = 1$, for the Linear Trend
  - $\varphi > 1$, for the Exponential Trend

Exponential Smoothing
Damped – Example

$$S_1 = S_0 + (\varphi * T_0) + (a * e_1)$$
$$T_1 = (\varphi * T_0) + (\beta * e_1)$$
$$F_1 = S_0 + (\varphi * T_0)$$

- Starting level: Constant of regression, starting trend: Slope of regression.
- Smoothing parameter: a = 0.2, β = 0.1, φ = 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	Forecast	Error	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%		



Exponential Smoothing
Model Selection

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

Exponential Smoothing
Model Selection

***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

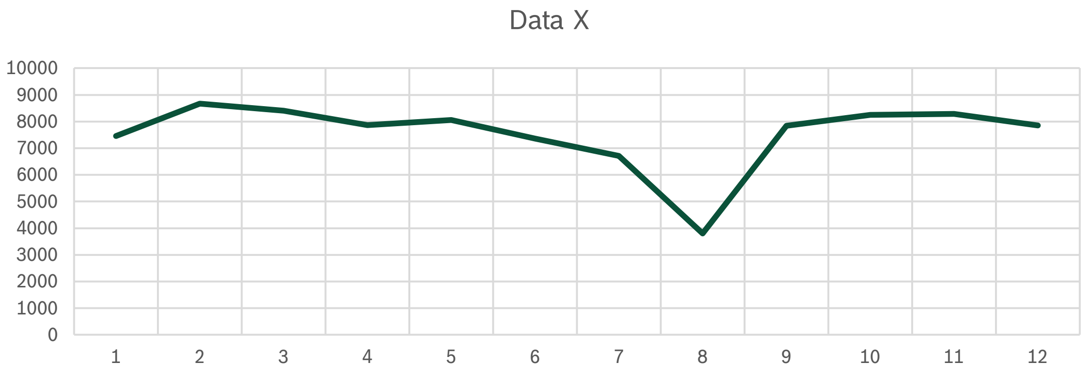
$$Variance = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n}$$

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

Exponential Smoothing
Model Selection

	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**

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

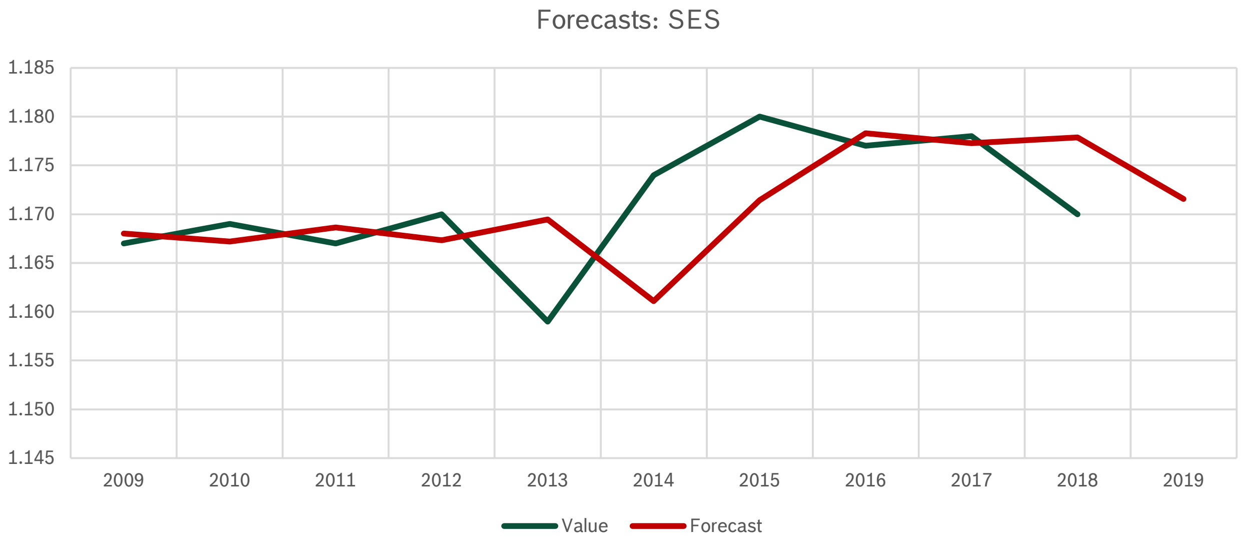


➢ Model to use: SES

Exponential Smoothing
Try to forecast with a method: SES

- **SES**: Simple Exponential smoothing
  - Initialization: S_1 is the mean value of ***the first two*** historical observations
  - Smoothing factor: 0.8 $F_{t+1} = S_t$, where $S_t = S_{t-1} + (a \times e_t)$, $e_t = Y_t - F_t$

	S(0) =	1.168	α		0,8
Time (t)	Year	Data X	Forecast	Error e	Level S
1	2009	1.167	1.168,0	-1,0	1.167,2
2	2010	1.169	1.167,2	1,8	1.168,6
3	2011	1.167	1.168,6	-1,6	1.167,3
4	2012	1.170	1.167,3	2,7	1.169,5
5	2013	1.159	1.169,5	-10,5	1.161,1
6	2014	1.174	1.161,1	12,9	1.171,4
7	2015	1.180	1.171,4	8,6	1.178,3
8	2016	1.177	1.178,3	-1,3	1.177,3
9	2017	1.178	1.177,3	0,7	1.177,9
10	2018	1.170	1.177,9	-7,9	1.171,6
11	2019		**1.171,6**		



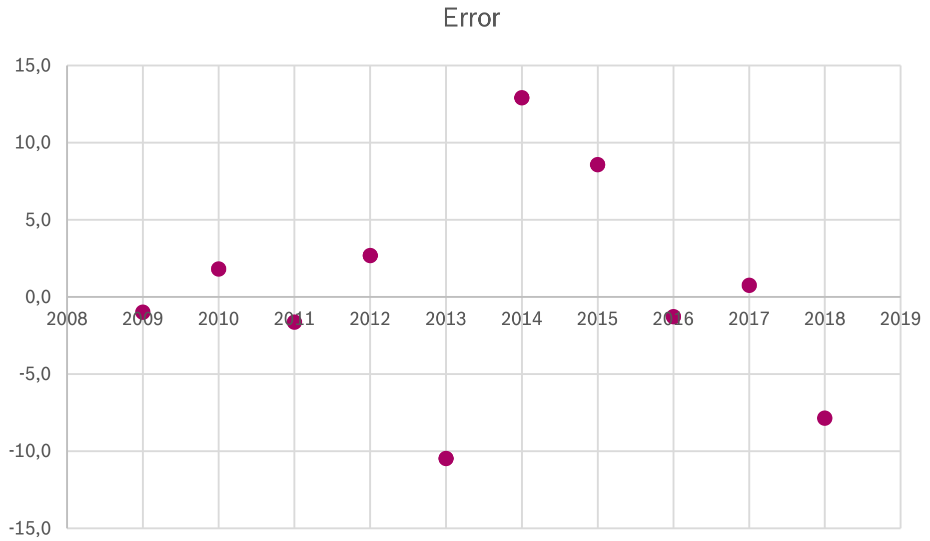
BOSCH

Exponential Smoothing
Try to forecast with a method: SES

- ***Evaluate*:**
  - How good the forecast is (model evaluation)?
    - Let's calculate the error

Data			
Year	**Value**	**Forecast**	**Error**
2009	1.167	1.168,0	-1,0
2010	1.169	1.167,2	1,8
2011	1.167	1.168,6	-1,6
2012	1.170	1.167,3	2,7
2013	1.159	1.169,5	-10,5
2014	1.174	1.161,1	12,9
2015	1.180	1.171,4	8,6
2016	1.177	1.178,3	-1,3
2017	1.178	1.177,3	0,7
2018	1.170	1.177,9	-7,9

- ***Results*:**
  - Equal positive & negative errors
  - Mean Error (ME): 0,45



 Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

Exponential Smoothing
Try to forecast with a method: SES

- ***Evaluate*:**
  - How good the forecast is (model evaluation)?
    - Let's calculate the ***absolute*** error

Data			
Year	**Value**	**Forecast**	**Abs. Error**
2009	1.167	1.168,0	-1,0
2010	1.169	1.167,2	1,8
2011	1.167	1.168,6	-1,6
2012	1.170	1.167,3	2,7
2013	1.159	1.169,5	-10,5
2014	1.174	1.161,1	12,9
2015	1.180	1.171,4	8,6
2016	1.177	1.178,3	-1,3
2017	1.178	1.177,3	0,7
2018	1.170	1.168,0	-1,0
2019		1.177	

- ***Results*:**
  - Mean Absolute Error (MAE): **4,9**
  - Naive (MAE): **5,7**
  - Relative error: 4,9/5,7 = **0,860** ☺

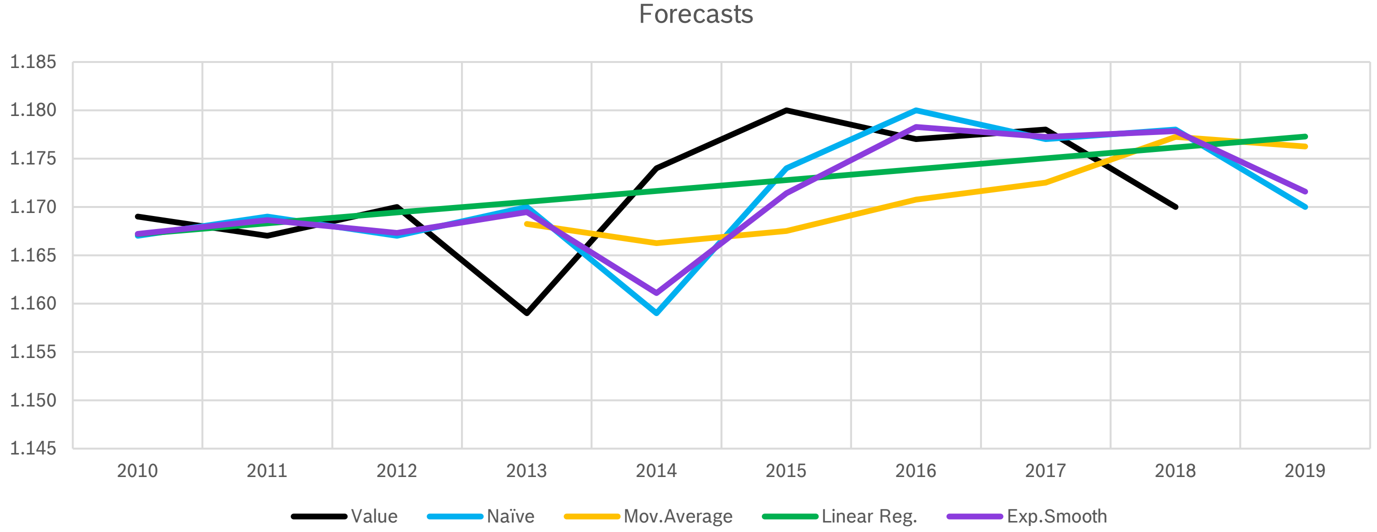
- Best method so far:
  - Naïve was worst (MAE = 5.7)
  - Moving Average was worst (MAE = 8.1)
  - Linear Regression was better (MAE = 3.8)

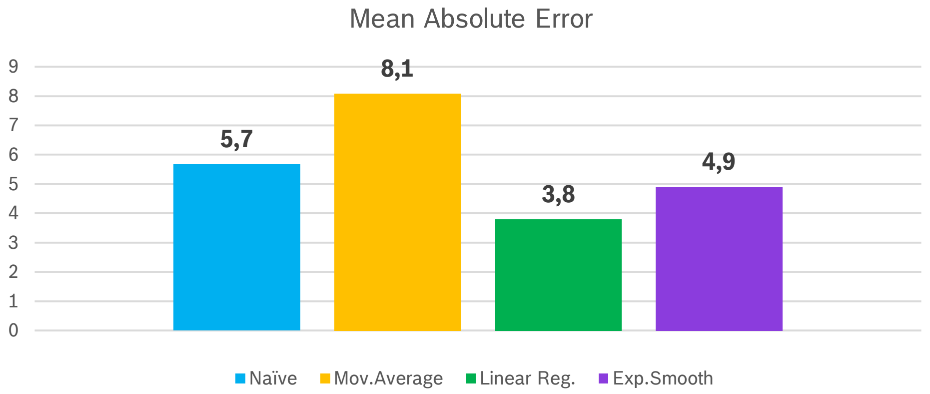
BOSCH

Exponential Smoothing
Compare methods & results

Data		Forecasts			
Year	**Value**	**Naïve**	**Moving Average**	**Linear Reg.**	**Exp. Smoothing**
2009	1.167			1.166,1	1.168,0
2010	1.169	1.167		1.167,2	1.167,2
2011	1.167	1.169		1.168,3	1.168,6
2012	1.170	1.167		1.169,4	1.167,3
2013	1.159	1.170	1.168,3	1.170,5	1.169,5
2014	1.174	1.159	1.166,3	1.171,7	1.161,1
2015	1.180	1.174	1.167,5	1.172,8	1.171,4
2016	1.177	1.180	1.170,8	1.173,9	1.178,3
2017	1.178	1.177	1.172,5	1.175,0	1.177,3
2018	1.170	1.178	1.177,3	1.176,1	1.177,9
2019		1.170	1.176,3	1.177,3	1.171,6

Mean Absolute Error			
Naïve	**Moving Average**	**Linear Reg.**	**Exp. Smoothing**
		0,9	1,0
2,0		1,8	1,8
2,0		1,3	1,6
3,0		0,6	2,7
11,0	9,3	11,5	10,5
15,0	7,8	2,3	12,9
6,0	12,5	7,2	8,6
3,0	6,3	3,1	1,3
1,0	5,5	3,0	0,7
8,0	7,3	6,1	7,9
5,7	**8,1**	**3,8**	**4,9**





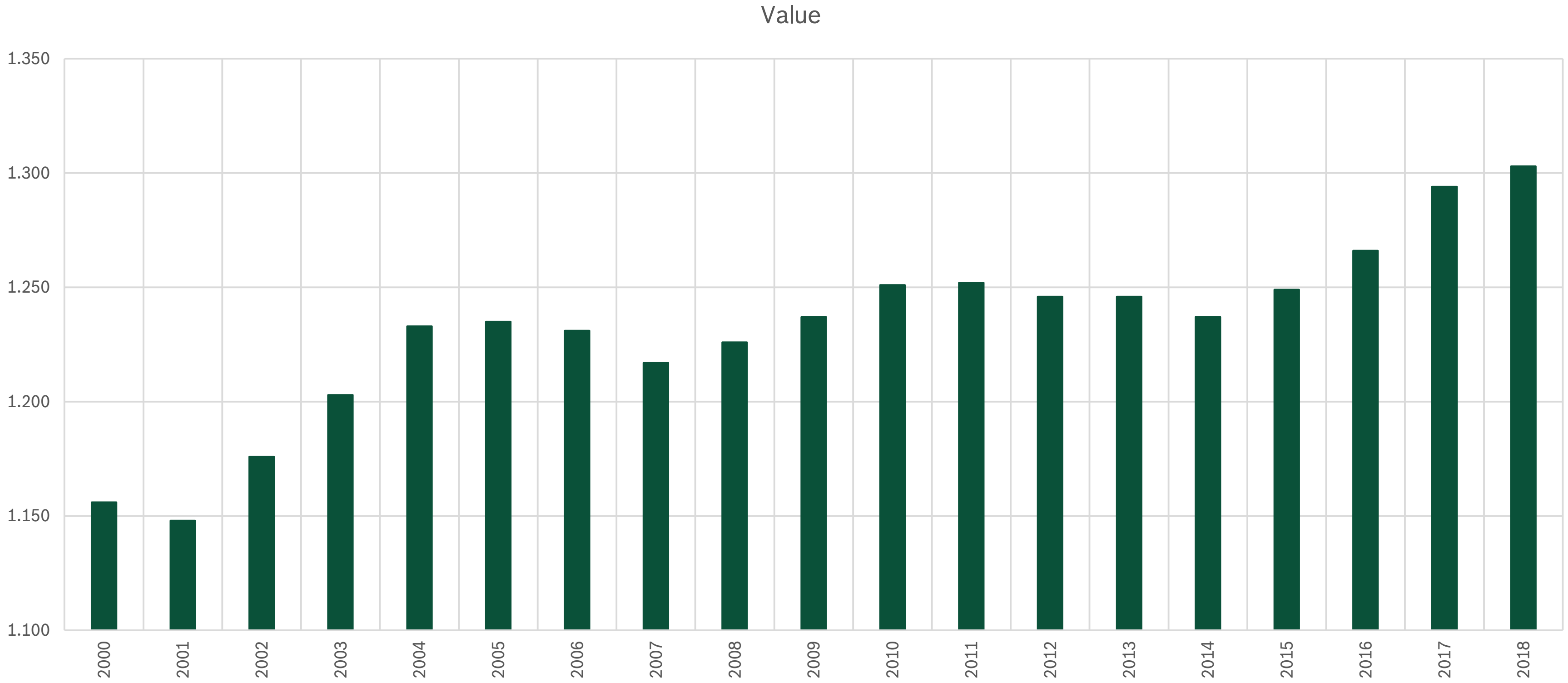
TIMESERIES FORECASTING (B)

BREAK TIME

10. FORECAST: EXERCISE

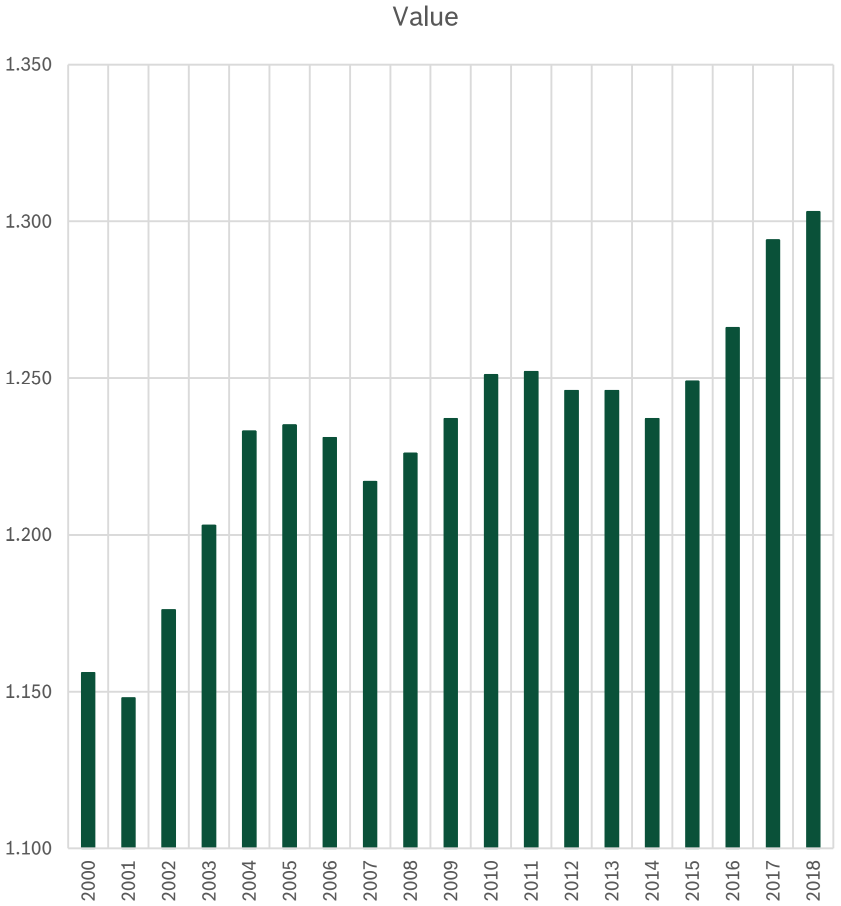
Forecast: Exercise
Timeserie

Data	
Year	**Value**
2000	1.156
2001	1.148
2002	1.176
2003	1.203
2004	1.233
2005	1.235
2006	1.231
2007	1.217
2008	1.226
2009	1.237
2010	1.251
2011	1.252
2012	1.246
2013	1.246
2014	1.237
2015	1.249
2016	1.266
2017	1.294
2018	1.303
2019	
2020	



Forecast: Exercise
Timeserie Characteristics

Data	
Year	**Value**
2000	1.156
2001	1.148
2002	1.176
2003	1.203
2004	1.233
2005	1.235
2006	1.231
2007	1.217
2008	1.226
2009	1.237
2010	1.251
2011	1.252
2012	1.246
2013	1.246
2014	1.237
2015	1.249
2016	1.266
2017	1.294
2018	1.303
2019	
2020	



- What you can observe?
  - 19 data points
  - Yearly data
  - Increasing trend over the years
  - No “strange” or “missing” values
- Let’s try to forecast:
  - Values: for 2019 and 2020
  - Horizon: 2 years (Mid-term)
  - Methods: Naïve, Moving Average, Linear Regression, Exponential Smoothing.

 Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

BOSCH

Forecast: Exercise
Timeserie Characteristics

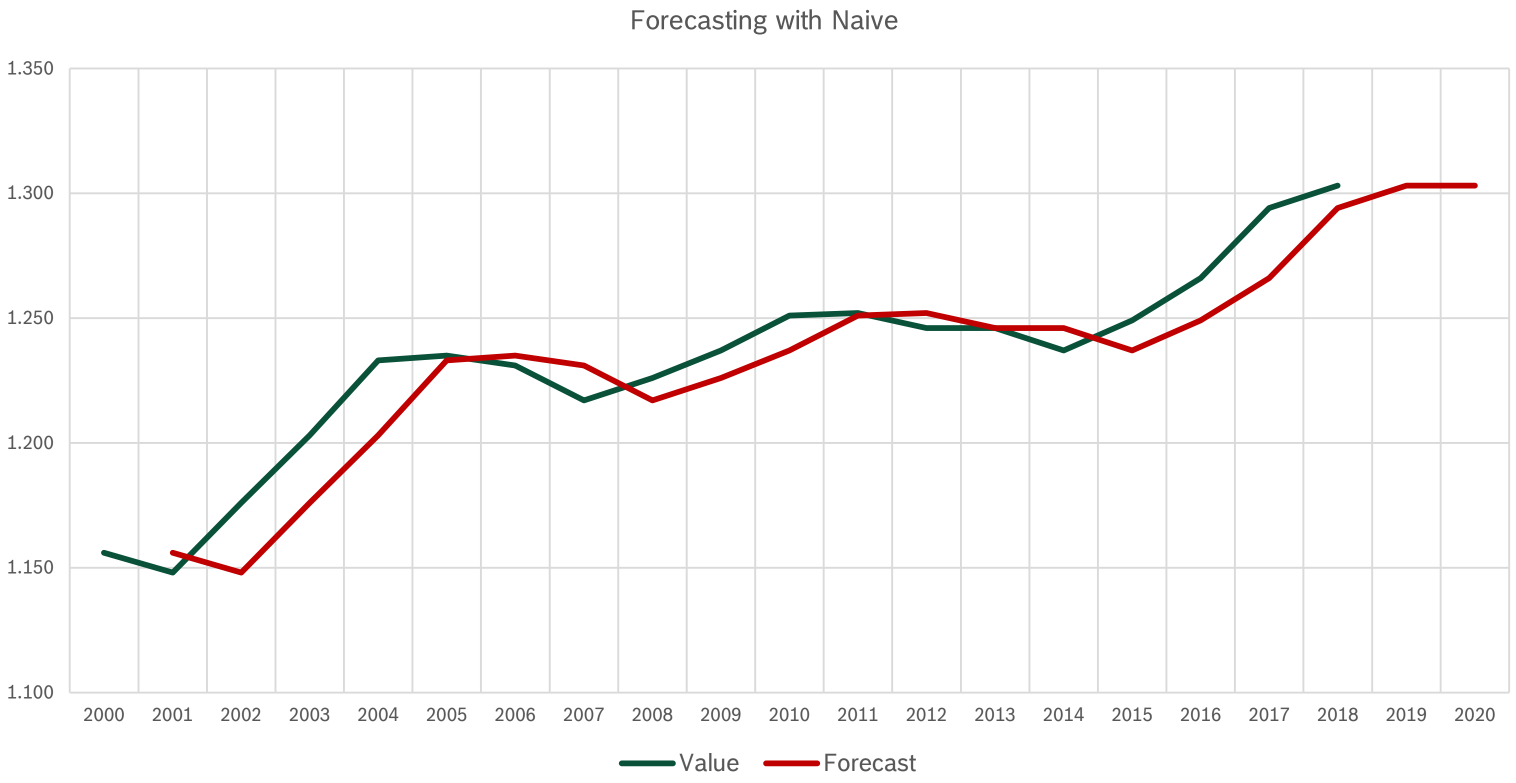
- Exercise steps: http://bos.ch/Cfd
  - Open file “**Forecasting Example - 20c - Exercise for Students (Empty).xlsx**”
  - Try to estimate forecasts:
    - Naïve
    - Moving Average
    - Linear Regression
    - Exponential Smoothing
  - Estimate errors
  - Compare the methods accuracy
- Duration:
  - **10 minutes**
- Your Goal:
  - Find best method and smallest MAE
- If you need help:
  - Open file “**Forecasting Example - 20d - Exercise for Students (Solution).xlsx**”
  - Just ask

Forecast: Exercise
Discussion

- Each participant:
  - Did you complete the exercise?
  - Did you manage to forecast?
    - Naïve
    - Moving Average
    - Linear Regression
    - Exponential Smoothing
  - Did you manage to estimate errors?
  - Did you find the best method?
- What was the most difficult?
  - The estimation of forecasts?
  - How to use the excel?
  - The forecast for the 2nd horizon 2020?
  - Something else?

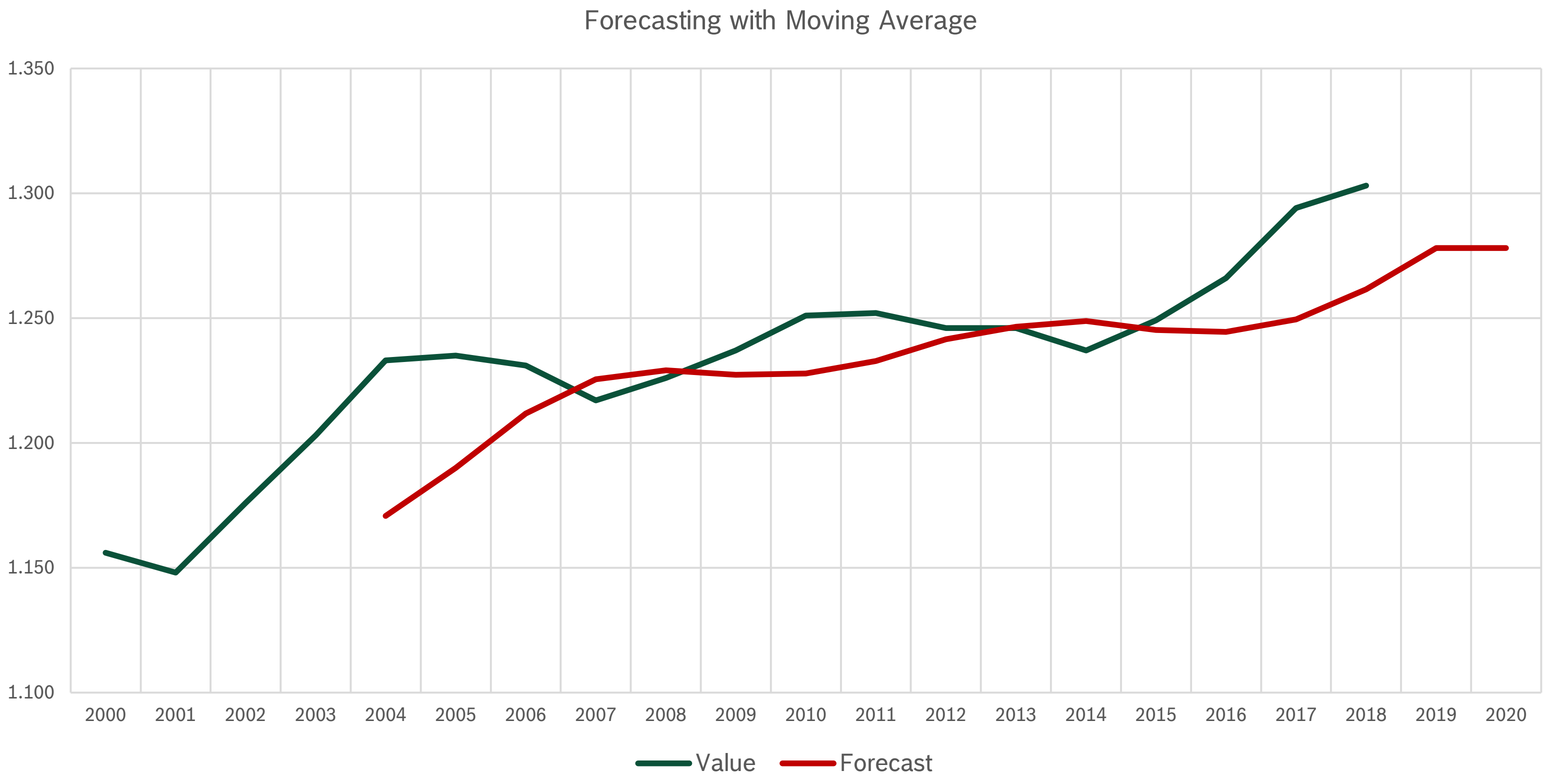
Forecast: Exercise
Solution: Naive

Data			
Year	**Value**	**Forecast**	**Abs. Error**
2000	1.156		
2001	1.148	1.156	8
2002	1.176	1.148	28
2003	1.203	1.176	27
2004	1.233	1.203	30
2005	1.235	1.233	2
2006	1.231	1.235	4
2007	1.217	1.231	14
2008	1.226	1.217	9
2009	1.237	1.226	11
2010	1.251	1.237	14
2011	1.252	1.251	1
2012	1.246	1.252	6
2013	1.246	1.246	0
2014	1.237	1.246	9
2015	1.249	1.237	12
2016	1.266	1.249	17
2017	1.294	1.266	28
2018	1.303	1.294	9
2019		**1.303**	
2020		**1.303**	
		Mean Error	**12,7**



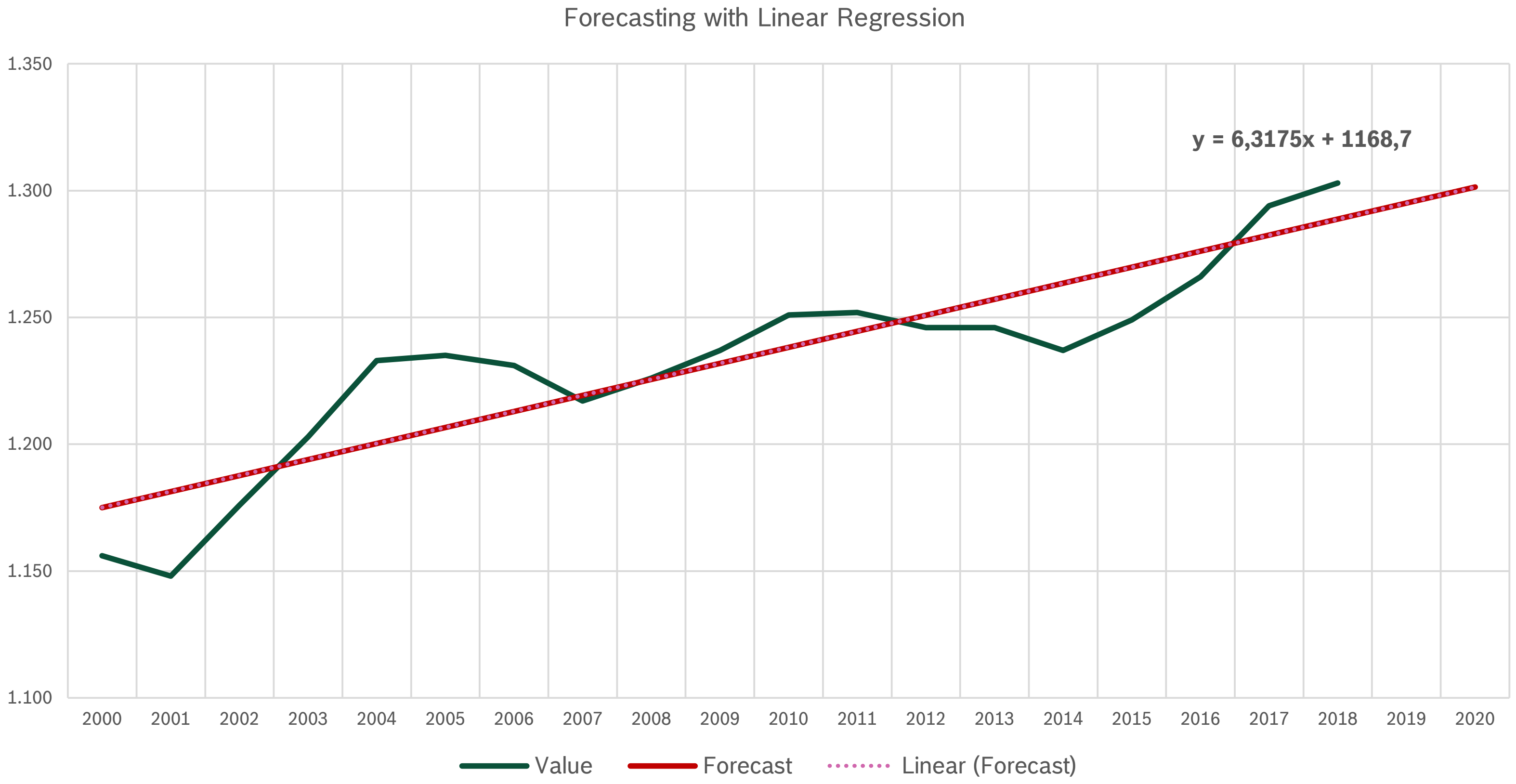
Forecast: Exercise
Solution: Simple Moving Average

Data			
Year	**Value**	**Forecast**	**Abs. Error**
2000	1.156		
2001	1.148		
2002	1.176		
2003	1.203		
2004	1.233	1.170,8	62,3
2005	1.235	1.190,0	45,0
2006	1.231	1.211,8	19,3
2007	1.217	1.225,5	8,5
2008	1.226	1.229,0	3,0
2009	1.237	1.227,3	9,8
2010	1.251	1.227,8	23,3
2011	1.252	1.232,8	19,3
2012	1.246	1.241,5	4,5
2013	1.246	1.246,5	0,5
2014	1.237	1.248,8	11,8
2015	**1.249**	1.245,3	3,8
2016	**1.266**	1.244,5	21,5
2017	**1.294**	1.249,5	44,5
2018	**1.303**	1.261,5	41,5
2019		**1.278,0**	
2020		**1.278,0**	
		Mean Error	**21,2**



Forecast: Exercise
Solution: Linear Regression

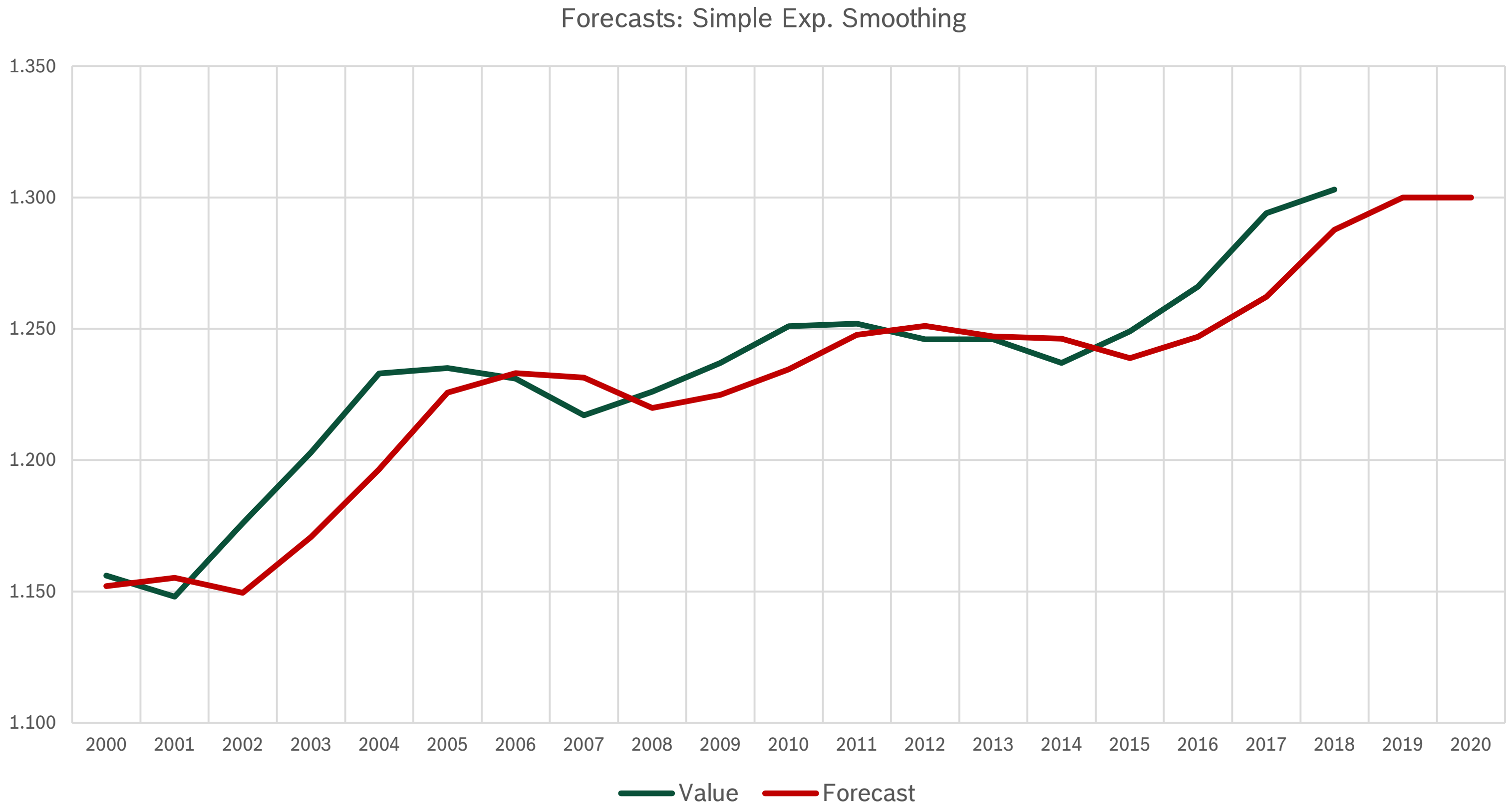
Data			
Year	Value	Forecast	Abs. Error
2000	1.156	1.175,0	19,0
2001	1.148	1.181,4	33,4
2002	1.176	1.187,7	11,7
2003	1.203	1.194,0	9,0
2004	1.233	1.200,3	32,7
2005	1.235	1.206,6	28,4
2006	1.231	1.212,9	18,1
2007	1.217	1.219,3	2,3
2008	1.226	1.225,6	0,4
2009	1.237	1.231,9	5,1
2010	1.251	1.238,2	12,8
2011	1.252	1.244,5	7,5
2012	1.246	1.250,8	4,8
2013	1.246	1.257,2	11,2
2014	1.237	1.263,5	26,5
2015	1.249	1.269,8	20,8
2016	1.266	1.276,1	10,1
2017	1.294	1.282,4	11,6
2018	1.303	1.288,8	14,2
2019		**1.295,1**	
2020		**1.301,4**	
		Mean Error	**14,7**



Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

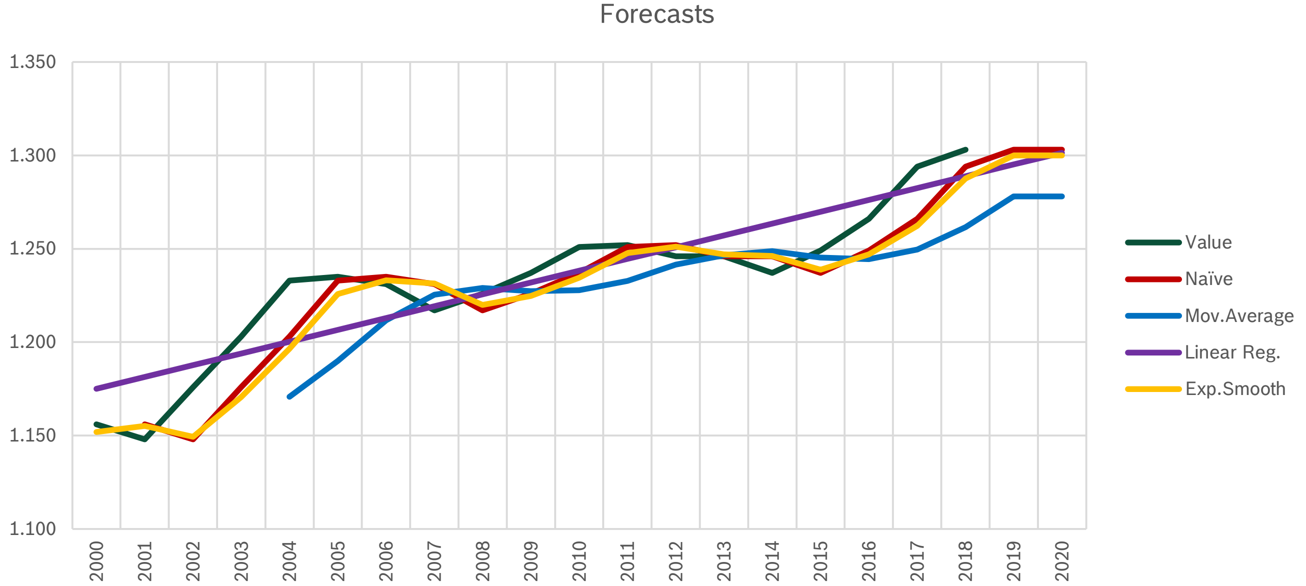
Forecast: Exercise
Solution: Exponential Smoothing

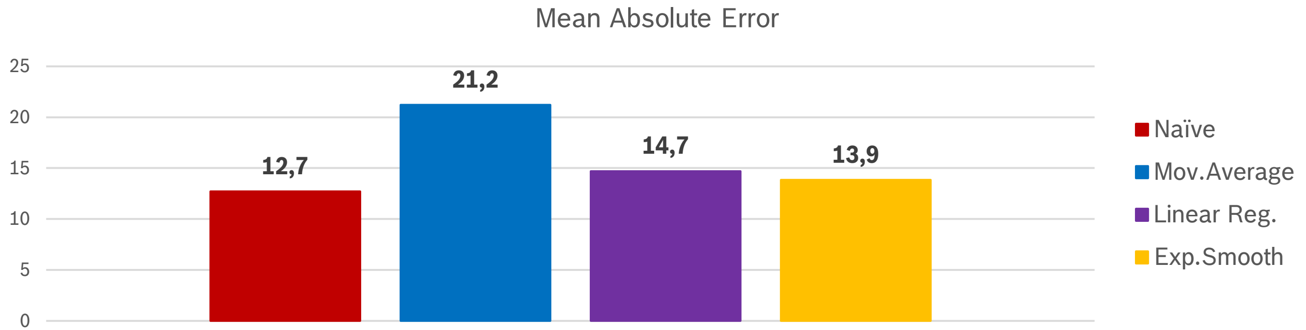
Data			
Year	Value	Forecast	Abs. Error
2000	1.156	1.152,0	4,0
2001	1.148	1.155,2	7,2
2002	1.176	1.149,4	26,6
2003	1.203	1.170,7	32,3
2004	1.233	1.196,5	36,5
2005	1.235	1.225,7	9,3
2006	1.231	1.233,1	2,1
2007	1.217	1.231,4	14,4
2008	1.226	1.219,9	6,1
2009	1.237	1.224,8	12,2
2010	1.251	1.234,6	16,4
2011	1.252	1.247,7	4,3
2012	1.246	1.251,1	5,1
2013	1.246	1.247,0	1,0
2014	1.237	1.246,2	9,2
2015	1.249	1.238,8	10,2
2016	1.266	1.247,0	19,0
2017	1.294	1.262,2	31,8
2018	1.303	1.287,6	15,4
2019		**1.299,9**	
2020		**1.299,9**	
		Mean Error	**13,9**



Forecast: Exercise
Solution: Comparison

Data		Forecasts			
Year	Value	Naïve	Mov. Average	Linear Regression	Exponential Smoothing
2000	1.156			1.175,0	1.152,0
2001	1.148	1.156		1.181,4	1.155,2
2002	1.176	1.148		1.187,7	1.149,4
2003	1.203	1.176		1.194,0	1.170,7
2004	1.233	1.203	1.170,8	1.200,3	1.196,5
2005	1.235	1.233	1.190,0	1.206,6	1.225,7
2006	1.231	1.235	1.211,8	1.212,9	1.233,1
2007	1.217	1.231	1.225,5	1.219,3	1.231,4
2008	1.226	1.217	1.229,0	1.225,6	1.219,9
2009	1.237	1.226	1.227,3	1.231,9	1.224,8
2010	1.251	1.237	1.227,8	1.238,2	1.234,6
2011	1.252	1.251	1.232,8	1.244,5	1.247,7
2012	1.246	1.252	1.241,5	1.250,8	1.251,1
2013	1.246	1.246	1.246,5	1.257,2	1.247,0
2014	1.237	1.246	1.248,8	1.263,5	1.246,2
2015	1.249	1.237	1.245,3	1.269,8	1.238,8
2016	1.266	1.249	1.244,5	1.276,1	1.247,0
2017	1.294	1.266	1.249,5	1.282,4	1.262,2
2018	1.303	1.294	1.261,5	1.288,8	1.287,6
2019		1.303	1.278,0	1.295,1	1.299,9
2020		1.303	1.278,0	1.301,4	1.299,9





11. WHO CAN BE A FORECASTER?

"Fortunate who was able to know the cause of things"

Publius Vergilius

Who can be a forecaster?
What I need?

- Making accurate forecasts does not depend on any special talent – ***it is a learned skill***.
- To make predictions:
  - Keep an open mind
  - Understand what you don't know
  - Think twice, or more
  - Admit and learn from your fails
  - We learn new things by doing. We improve those skills by doing more
  - Don't take your forecasting success for granted
  - Remain humble

Who can be a forecaster?
Beware the “Illusion of control”

- **Flipping a coin 100 times**: Can you predict the outcomes?
  - The majority of forecasters: Accurately predicted the coin toss about 50% of the time.
  - Some will mostly be wrong.
  - Some others mostly right! Do they have a coin-flipping forecasting ability?
    - No! There is no skill involved. It’s only luck!
  - Extremely unlikely for a person to forecast correct >70% of 100 coin tosses.
    - But, if you have 3000 persons, the unlikely becomes quiet likely.

➢ This is the “illusion of control”, or “illusion of prediction”.

Who can be a forecaster?
Beware your success

- You beat the forecast six or seven times in a row?
  - Don’t take your forecasting success for granted
  - Remain humble
  - Reality is infinite complex

- This cause a paradox
  - Success can lead to acclaim, that can undermine the habits of mind that produced the success
  - Such hubris often afflicts highly accomplished individuals
  - In business cycles, it is called “***CEO disease***”

BOSCH

Who can be a forecaster?
How to think: Fermitize the problem

- **How many piano tuners are in Chicago?**
  - Enrico Fermi: An Italian-American physicist
  - Question posted to his students

- How to answer such a question?
  - "*Well, maybe…200*"
    - But the number came out of a black box
    - No idea how it was generated
- **Forecaster approach**:
  - Break the question down into smaller questions for facts, that can probably be answered!

 Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

BOSCH

Source: Nobel prize org

Who can be a forecaster?
How to think: Fermitize the problem

- **Four questions for four facts:**

1. Number of pianos in Chicago?
2. How often pianos are tuned each year?
3. How long it takes to tune a piano?
4. How many hours a year the average piano tuner works?

- By breaking down the question, we can better separate the knowable and the unknowable
  - Now let's go deeper with each question.

Source: Nobel prize org

Who can be a forecaster?
How to think: Fermitize the problem

1. Number of pianos in Chicago?

- How many people are there in Chicago? I don't know..
  - Chicago is the 3rd largest city after New York & Los Angeles. LA has 4 M people or so.
  - Chicago has more than, let's say, 1.5 M. And I'm pretty sure it has fewer that 3.5 M.
  - Where is the correct answer within this range? Let's take the midpoint and guess that Chicago has 2.5 M people.
- What percentage of people own a piano?
  - Pianos are too expensive for most families – and most who can afford one don't really want one. So I'll put in at 1%
  - That's mostly a black box guess but it's the best I can do.
- How many institutions (schools, concert halls, bars) own pianos?
  - Some institutes would own pianos. I'll again make a black box guess.
  - It's enough to double the per person number of pianos to roughly 2 %.
- Estimation: **there are 50.000 pianos in Chicago.**

Source: Nobel prize org

Who can be a forecaster?
How to think: Fermitize the problem

2. How often are pianos tuned?

- Maybe once a year. That strikes as reasonable. Why? I don’t know (black box guess).

3. How long does it take to tune a piano?

- I’ll say two hours (black box guess).

4. How many hours a year does the average piano tuner work?

- The standard American workweek is 40 hours, minus two weeks of vacation. Piano tuners also.
- Thus: Multiply 40 hours by 50 weeks to come up with 2000 hours per year.
- What about travelling between pianos?
  - How much time do they spend between jobs? I’ll guess 20% of their work hours.
- Estimation: the average piano tuner works 1600 hours per year.

BOSCH

Source: Nobel prize org

Who can be a forecaster?
How to think: Fermitize the problem (5)

- **Let's summarize:**
  - 50.000 pianos, need tuning 1/year, and takes 2 hours to tune.
  - A total of 100.000 hours.
  - Divide 100.000 / 1.600 working hours = 62,5 piano tuners.
  - Final forecast: **63** (actual value: 83 listings on yellow pages, but including duplicates…). Good forecast!

➢ *With little or no information at disposal, we could do some calculations like this to come up with a number that is quite accurate!*

- Can you estimate the number of cm^2 of pizzas consumed in your city during a year? ☺

BOSCH

Source: Nobel prize org

Who can be a forecaster?
How to think: Outside first

- **Use-case**:
  - A family lives in a small house. Massimo is 44 and works as an accountant. Mary is 35 and work part-time as nurse. Tommy is 5. Grandmother Camila lives with them.
  - **How likely is that they have a pet?**
- Thoughts:
  - Are they Italians? If so, Massimo grew up with lots of brothers and sisters. But he has only 1 child.
  - Maybe they want a bigger family, but they can’t afford it. So, they may have a pet.
- Or…
  - People get pets for their kids, and Tommy isn’t old enough to take care a pet…

BOSCH

Source: Nobel prize org

Who can be a forecaster?
How to think: Outside first

- All these thoughts are valid.
- But, this is the “**inside view**”.
- A forecaster should:
  - Avoid the psychological concept of “**anchoring**”
  - Start with the “**outside view**”:
    - What is the % of American Households own a pet? Around 62%
  - Then turn to the inside view, all those details about the family:
    - Use them to adjust that initial 62% up or down.

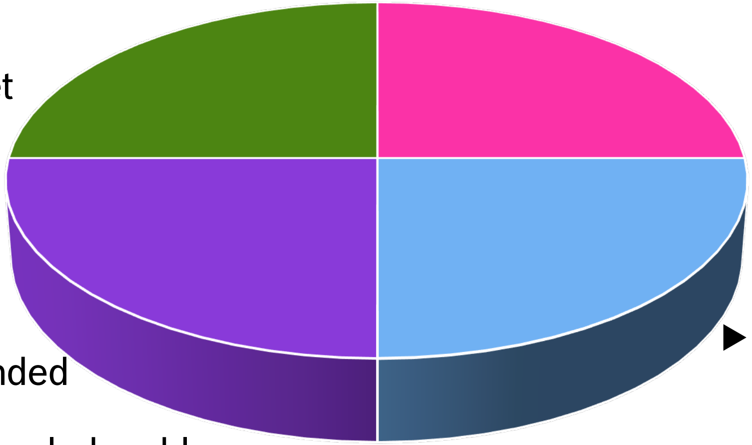
- *Coming up with an outside view, an inside view, and a synthesis of the two isn’t the end. It’s a good beginning.*

BOSCH

Source: Nobel prize org

Who can be a forecaster?
What a forecaster should have

- A forecaster should have:



- ***Work*** Ethic
  - A growth mindset
  - Grit
- ***Philosophic*** Outlook
  - Cautious
  - Humble
  - Non deterministic
- ***Thinking*** Style
  - Actively open-minded
  - Intelligent and Knowledgeable
  - Reflective
  - Numerate
- ***Forecasting*** Style
  - Pragmatic
  - Analytical
  - Dragonfly-eyed
  - Probabilistic
  - Thoughtful updaters
  - Good intuitive Psychologists

➢ *Not every attribute is equally important!*

➢ *And not every good forecaster has every attribute!*

12. METHOD & DATA SELECTION

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

Control Magazine, Institute of Operation Management

BOSCH

Method & Data Selection
Before selecting the method

- 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.

Method & Data Selection
Before selecting the method

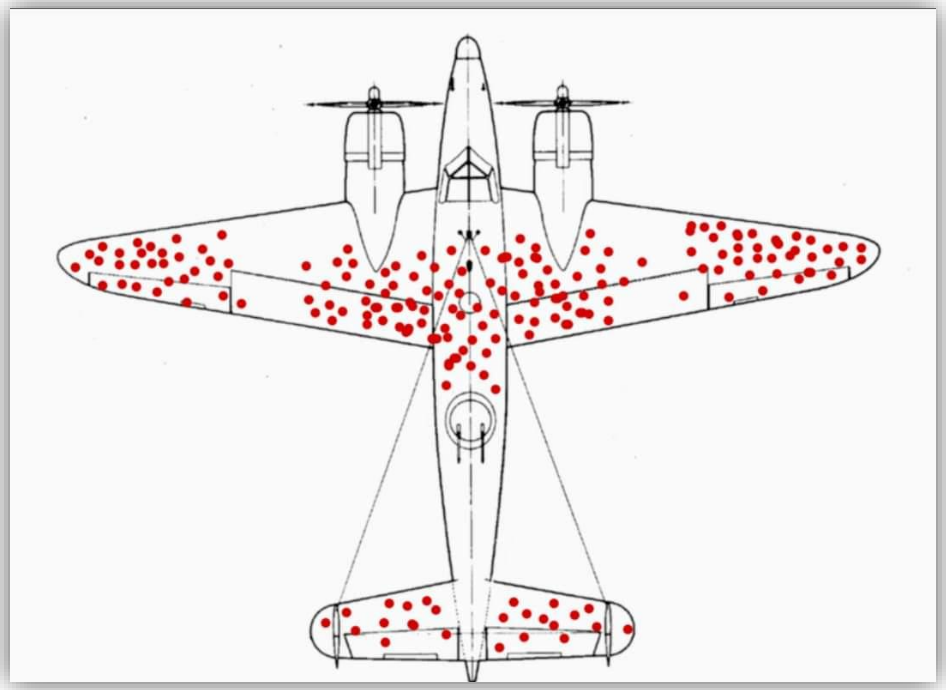
- ***Problem definition***: Sometimes the most difficult part.
  - Break the question down into ***smaller components***.
  - Identify the ***known*** and the unknown.
  - Look closely at all of your ***assumptions***.
  - Consider the ***outside view***, frame the problem not as a unique thing but as 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 & Data Selection
Before selecting the method

- ***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 & Data Selection
Before selecting the method

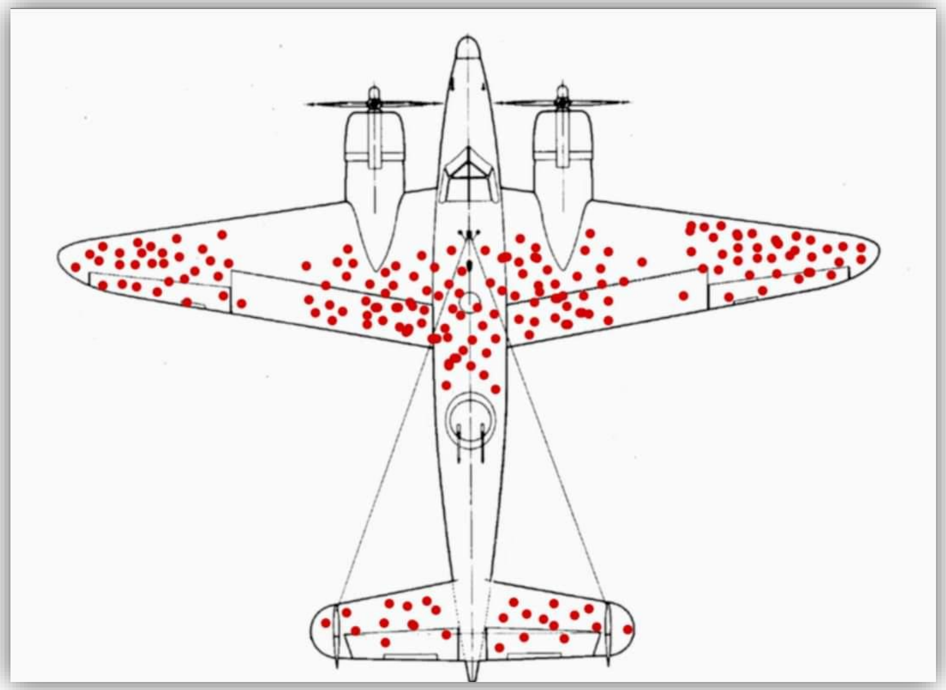
- During WWII, the Navy tried to determine where they needed to armor their aircraft to ensure they came back home.
  - They ran an analysis of where planes had been shot up
  - Obviously the places that needed to be up-armored are:
    - wingtips
    - central body
    - elevators



- But ***Abraham Wald***, a statistician, disagreed.

Method & Data Selection
Before selecting the method

- He thought they should better armor the nose area, engines, and mid-body. Why?
  - He realized what the others didn't!
  - The planes were getting shot there too, but they weren't making it home.
- The Navy:
  - analyzed where aircraft could suffer the most damage without catastrophic failure.
  - Those planes had been shot there and crashed.
  - **They weren't looking at the whole sample set**, only the survivors.



Are you sure that you are using the correct data?

BOSCH

Method & Data Selection
Before selecting the method

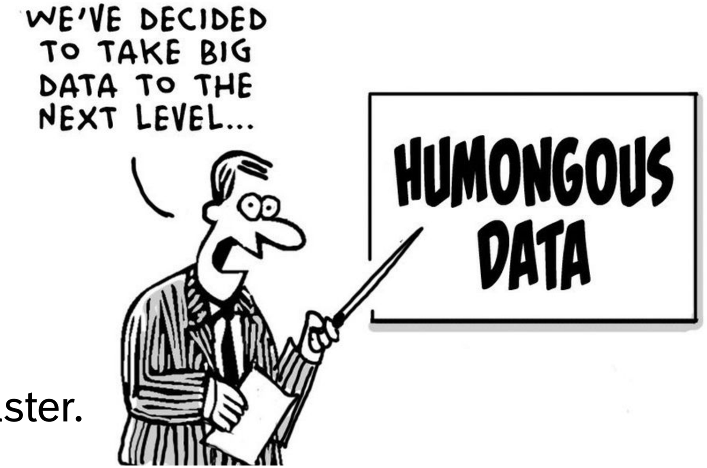
- ***Data preparation***:
  - Scope: to have a first feeling
    - Do patterns exist? What information can we gain from the raw historical data?
    - Does the data have trend?
    - Does the data have seasonality (Calendar effects, Institutional influences, Weather, Expectations, etc.)
    - Do outliers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

Method & Data Selection
Before selecting the method

- ***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.

BOSCH

Method & Data Selection
More data the better?



- Add more data and variables?
  - Can help identify more correlations, but
  - The proportion that are spurious and dangerously misleading rises even faster.

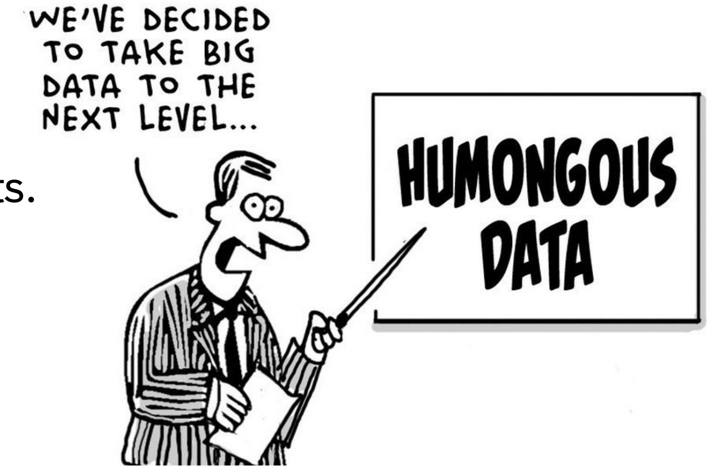
- Keep in mind the distinction between ***wisdom*** and ***information***!
  - Failing to put data into context, check its quality, or understand the questions that one is trying to answer will result in noise, not signal!
  - Forecasts will not be improved unless ***we are identifying why things are happening***.
  - More data in an uncertain world may create the impression that we are dealing with calculable risks when we actually are ***dealing with unknowns***.

BOSCH

Source: Recode

Method & Data Selection
More data the better?

- The most meaningful forecasts are the ones that provide:
  - ***useful explanations*** and
  - ***cogent framework*** for understanding current conditions and future events.
- What matters is:
  - not so much the amount of information a forecaster marshals, but
  - ***the depth of the wisdom underlying the analysis***.



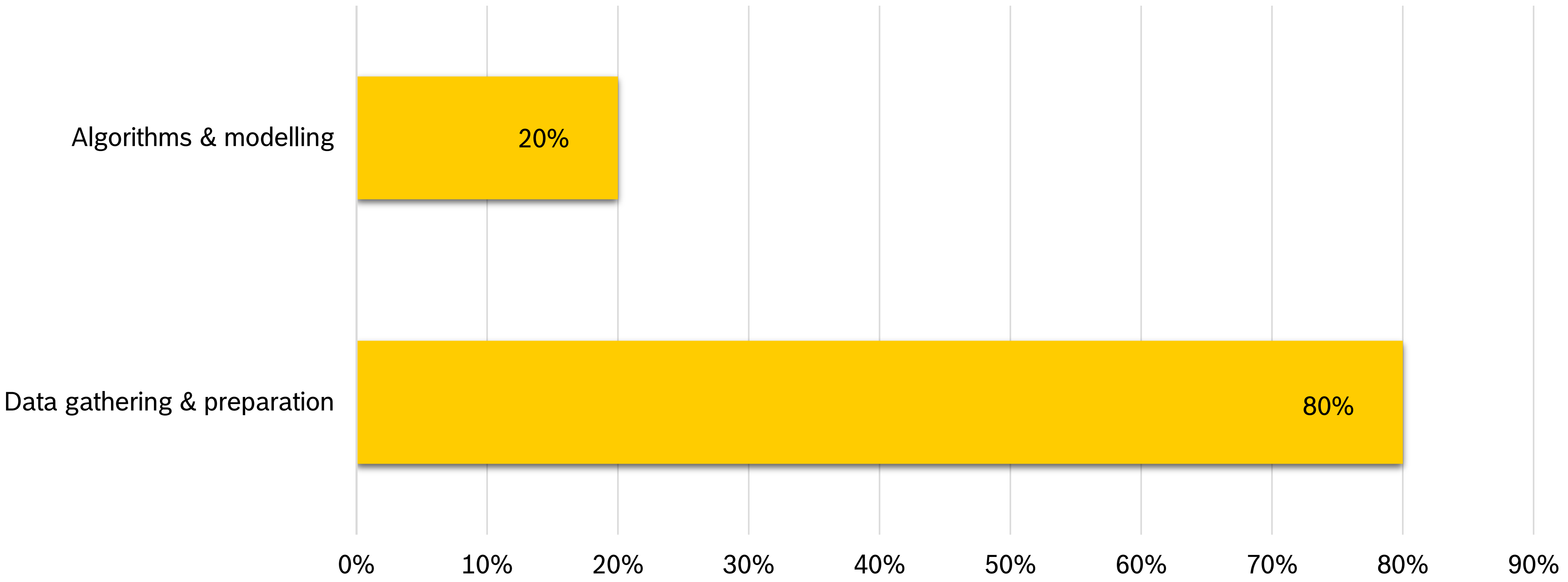
- Those who focus solely on information are liable to equate correlation with causation.

Wisdom comes from a deeper understanding of why things happen.

Source: Recode

Method & Data Selection
Before selecting the method

The 1st 80/20 Law



BOSCH

Method & Data Selection
Finding 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 & Data Selection
Finding 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.

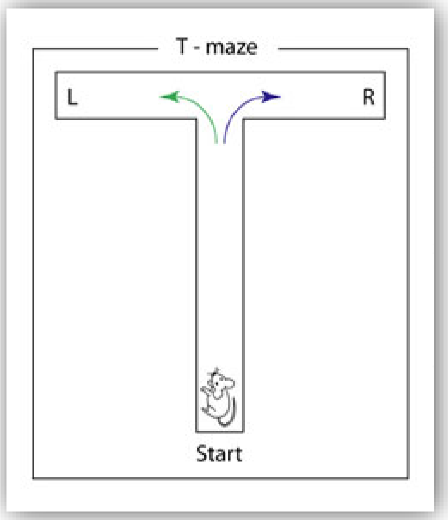
In general, the greater the level of detail (and frequency) that is required, the greater the need for an automated forecasting procedure.

Method & Data Selection
Finding 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 & Data 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.*

BOSCH

Method & Data Selection
Finding 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 & Data Selection
Finding 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.***

BOSCH

Method & Data Selection
Finding 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 & Data Selection
Finding 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.

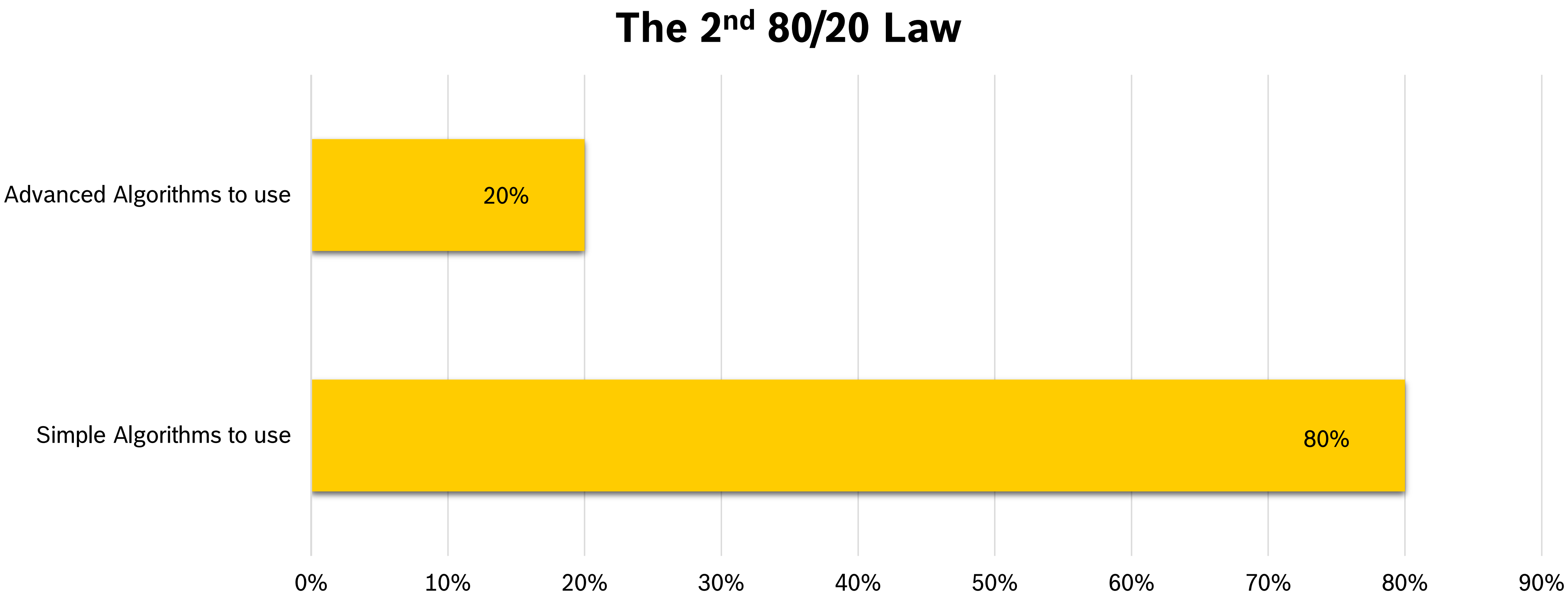
Method & Data Selection
Finding the proper method: Best ?

- What if a selection decision as to the "best" method is unclear?
  - it has been shown to be beneficial to hedge by using ***more than 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.

When in doubt we should ***combine*** multiple forecasts that come from a ***variety of independent*** sources.

Method & Data Selection
Finding the proper method



Method & Data Selection
Finding the proper method: Keep in mind

- Getting forecasts to fit past data patterns is relatively easy.
  - There is a temptation to think that a **good fit** is a sign of accurate forecasts to come.
  - But, if there is no rationale for these forecasts, their performance is likely to be disappointing.

- We need to start placing a **greater emphasis** on judging forecasts by **the quality of the process** that was used to produce them.
  - The past accuracy of a method will often be an ***imperfect basis for trust***.
    - Basing a decision on what a forecaster/model tells you just because they've previously made a few lucky forecasts to be accurate is almost ***a sure-fire way to be caught out***.

➢ ***And of course: Get serious about keeping score!***

Method & Data Selection
Finding the proper method: A true example

- Setup:
  - A set of 10 timeseries
  - A set of 5 forecasting methods to be tested
  - Historical data from 2010 to 2017

- Goal:
  - Find the proper method
  - Forecast for the next year 2018

Method & Data Selection
Finding the proper method: A true example

- Find the best method
  - A set of 5 methods will be tested over historical data
  - Try to minimize MAE index

	Find Best Methods: Historical Data from 2000 up to 2017					
	Mean Absolute Error (MAE)					Best Method
Timeserie	**Method A**	**Method B**	**Method C**	**Method D**	**Method E**	
1	12,7	12,8	11,1	12,5	**10,1**	E
2	46,7	**37,3**	41,1	37,4	39,8	B
3	**4,2**	4,3	4,4	4,4	5,1	A
4	124,5	110,9	**108,8**	109,5	111,1	C
5	1.450,5	1.233,4	1.169,7	1.087,1	**1.085,2**	E
6	610,3	522,6	519,6	**517,1**	531,6	D
7	6,2	5,9	4,4	**2,4**	3,4	D
8	31.758,4	27.077,1	26.800,0	25.720,0	**25.192,2**	E
9	214,8	**188,7**	190,5	189,1	194,0	B
10	**25,7**	26,1	26,4	25,8	27,9	A

	Find Best Methods: Historical Data from 2000 up to 2017					
	RELATIVE - Mean Absolute Error					Best Method
Timeserie	**Method A**	**Method B**	**Method C**	**Method D**	**Method E**	
1	1,000	1,008	0,874	0,985	**0,794**	E
2	1,000	**0,798**	0,880	0,800	0,852	B
3	**1,000**	1,024	1,048	1,048	1,214	A
4	1,000	0,890	**0,874**	0,880	0,892	C
5	1,000	0,850	0,806	0,749	**0,748**	E
6	1,000	0,856	0,851	**0,847**	0,871	D
7	1,000	0,944	0,711	**0,392**	0,552	D
8	1,000	0,853	0,844	0,810	**0,793**	E
9	1,000	**0,879**	0,887	0,880	0,903	B
10	**1,000**	1,016	1,027	1,004	1,086	A
Average	**1,000**	**0,912**	**0,880**	**0,839**	**0,870**	**0,812**

Method & Data Selection
Finding the proper method: A true example

- Use the best method to forecast for 2018
  - Wait for the actual values to arrive, and **measure** the accuracy

	Measure success for 2018: Forecasts vs Actual values					
	Mean Absolute Error (MAE)					Best Method
Timeserie	**Method A**	**Method B**	**Method C**	**Method D**	**Method E**	
1	16,2	14,7	11,9	11,2	10,8	E (10,1)
2	59,5	62,8	52,3	51,7	52,2	B (37,3) → D
3	5,3	5,9	5,8	5,5	5,7	A (4,2)
4	110,2	103,7	95,8	96,1	97,0	C (108,8)
5	1.390,4	1.227,9	1.150,0	1.164,5	1.201,5	E (1085,2) → C
6	738,9	531,1	539,9	508,0	540,1	D (517,1)
7	6,6	5,1	6,3	5,7	6,2	D (2,4) → B
8	29.354,8	29.554,9	30.007,9	29.505,7	29.600,4	E (25192,2) → A
9	246,6	202,9	211,0	202,7	204,9	B (188,7) → D
10	32,8	23,7	22,5	22,8	24,6	A (25,7) → C

	Measure success for 2018: Forecasts vs Actual values					
	RELATIVE - Mean Absolute Error					Best Method
Timeserie	**Method A**	**Method B**	**Method C**	**Method D**	**Method E**	
1	1,000	0,907	0,735	0,691	0,667	E
2	1,000	1,056	0,880	0,869	0,878	B → D
3	1,000	1,109	1,091	1,034	1,072	A
4	1,000	0,941	0,869	0,872	0,880	C
5	1,000	0,883	0,827	0,838	0,864	E → C
6	1,000	0,719	0,731	0,688	0,731	D
7	1,000	0,774	0,957	0,866	0,942	D → B
8	1,000	1,007	1,022	1,005	1,008	E → A
9	1,000	0,823	0,856	0,822	0,831	B → D
10	1,000	0,723	0,686	0,695	0,750	A → C
Average	**1,000**	**0,894**	**0,865**	**0,838**	**0,862**	**0,884**

- New "best" method for 6 timeseries!
  - "Best" methods for historical data, are not always suitable for the future! (overfitting)
  - A same stable method for all timeseries, can provide more accurate future forecasts!

BOSCH

Method & Data Selection
Experts Adjustments

- 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 & Data Selection
Experts Adjustments

- ***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 & Data Selection
Update Forecasts

- 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 & Data Selection
Update Forecasts

- In the light of new information:
  - If you don't adjust your view, you won't capture the value of that information.
  - If you are impressed by the new information and you base your forecast entirely on it, you will lose the value of the old information that underpinned your prior forecast.

- ***What to do?***
  - Carefully balance old and new information, in order to capture the value in both.
  - The best way is to update your forecast ***often*** but ***bit by bit***.
    - "*If you have to forecast, forecast often*", Edgar R. Fiedler

Method & Data Selection
Update Forecasts

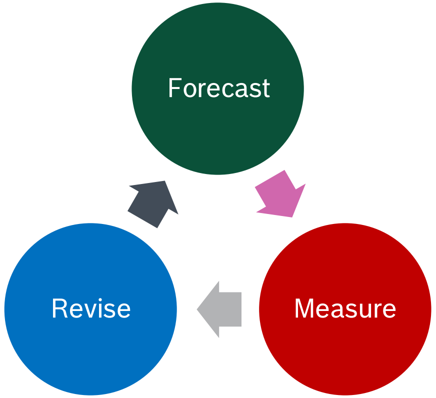
- It is also tricky to ***distinguish*** important from irrelevant ***information***.
  - sometimes people think something is important but it's not, and irrelevant information can confuse and trigger biases.
  - when one doesn't feel committed to the results, they can overreact.
  - when they are really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

- ***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.



Forecast – Measure – Revise – Repeat. It's a never ending process

13. SPECIAL TOPIC

OUTLIERS / SPECIAL EVENTS & ACTIONS

Outliers / Special Events & Actions
Definition

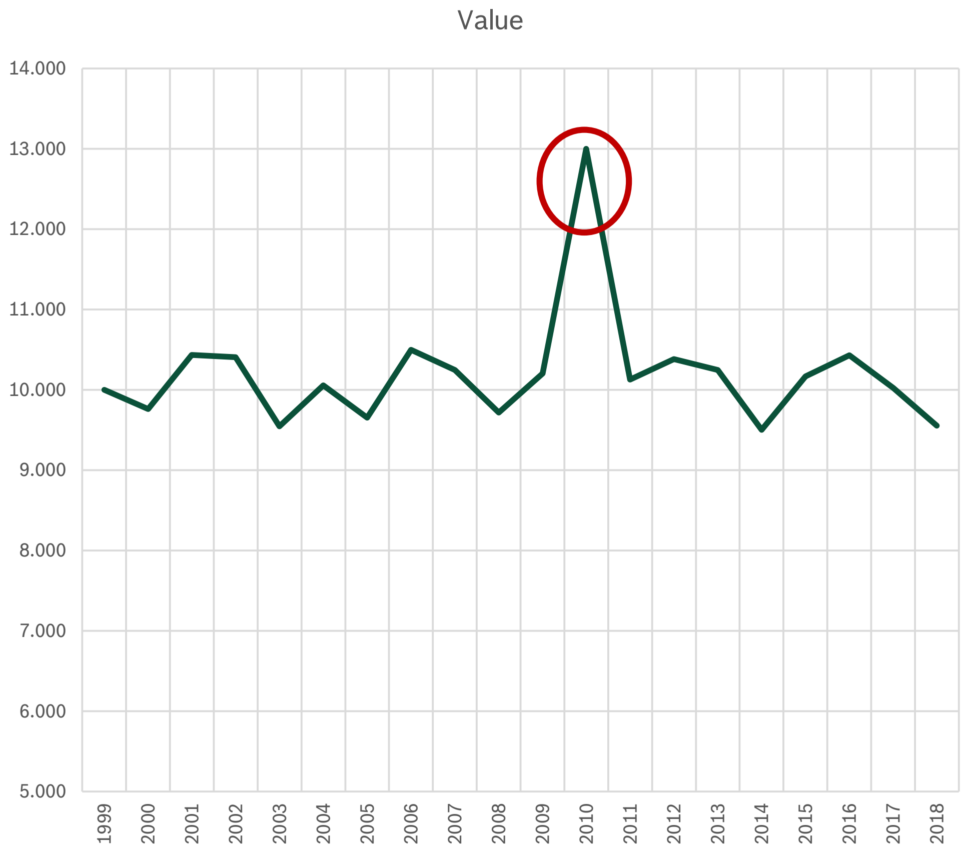
Level difference between current data of a timeserie and an initial level (or the level of the average value), could be the impact from a special event or action.

- 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.

Outliers / Special Events & Actions
Observe Yearly Outlier effect

Data	
Year	**Value**
1999	10.000
2000	9.760
2001	10.433
2002	10.408
2003	9.545
2004	10.054
2005	9.654
2006	10.499
2007	10.248
2008	9.719
2009	10.203
2010	13.000
2011	10.127
2012	10.381
2013	10.248
2014	9.504
2015	10.167
2016	10.431
2017	10.025
2018	9.552

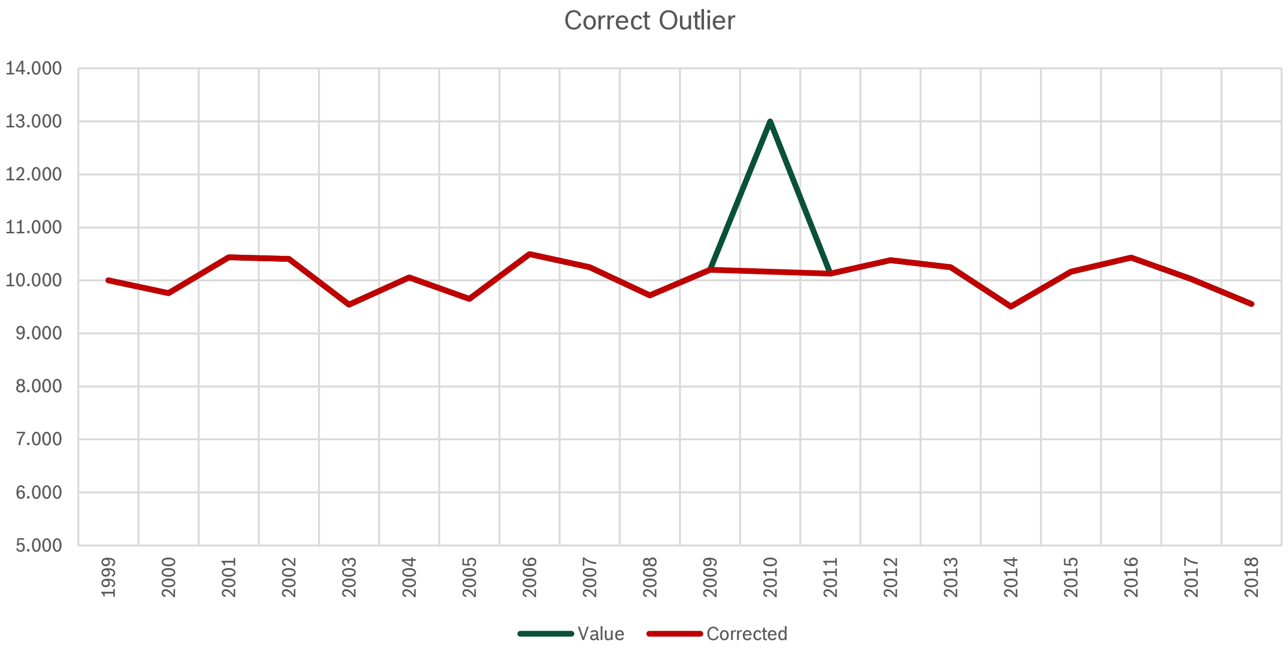


- What you can observe?
  - 20 data points
  - Yearly data
  - No trend
  - Year 2010 → too high
    - Outlier due to a SEA.

BOSCH

Outliers / Special Events & Actions
Removing Yearly Outlier

Data	
Year	**Value**
1999	10.000
2000	9.760
2001	10.433
2002	10.408
2003	9.545
2004	10.054
2005	9.654
2006	10.499
2007	10.248
2008	9.719
2009	**10.203**
2010	**13.000**
2011	**10.127**
2012	10.381
2013	10.248
2014	9.504
2015	10.167
2016	10.431
2017	10.025
2018	9.552



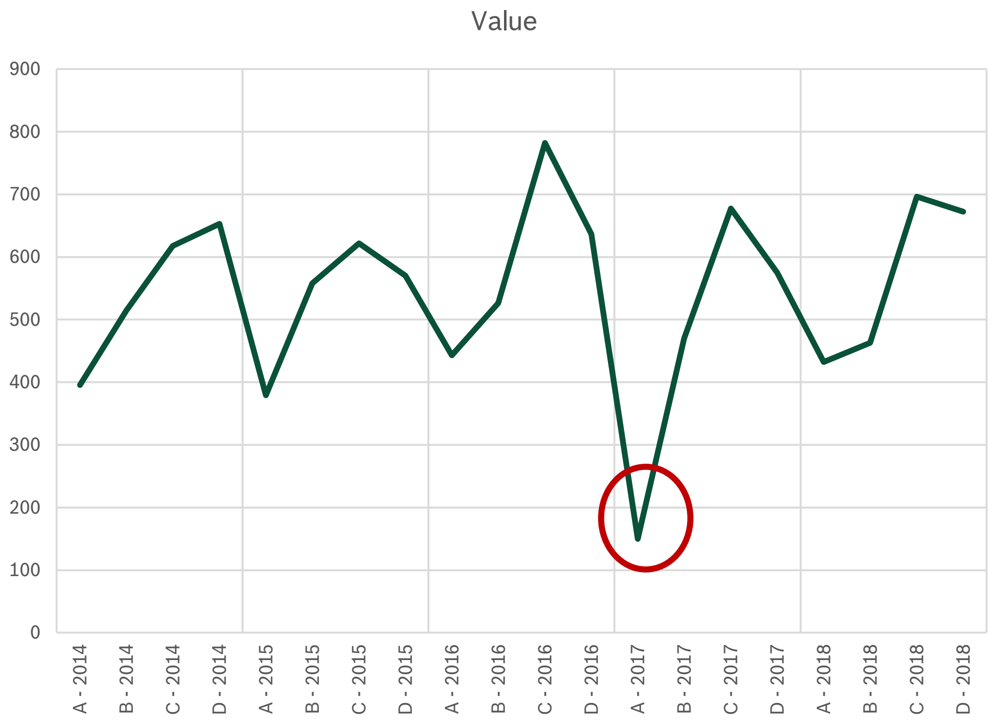
Data	
Year	**Corrected**
1999	10.000
2000	9.760
2001	10.433
2002	10.408
2003	9.545
2004	10.054
2005	9.654
2006	10.499
2007	10.248
2008	9.719
2009	10.203
2010	**10.165**
2011	10.127
2012	10.381
2013	10.248
2014	9.504
2015	10.167
2016	10.431
2017	10.025
2018	9.552
Impact	**27,9%**

$$D'_{2010} = (D_{2009}+D_{2011})/2 = (10203 + 10127)/2 = 10165$$

$$Impact_{2010} = \frac{13000 - 10165}{10165} * 100\ (\%) = \mathbf{27,9\%}$$

Outliers / Special Events & Actions
Observe Seasonal Outlier effect

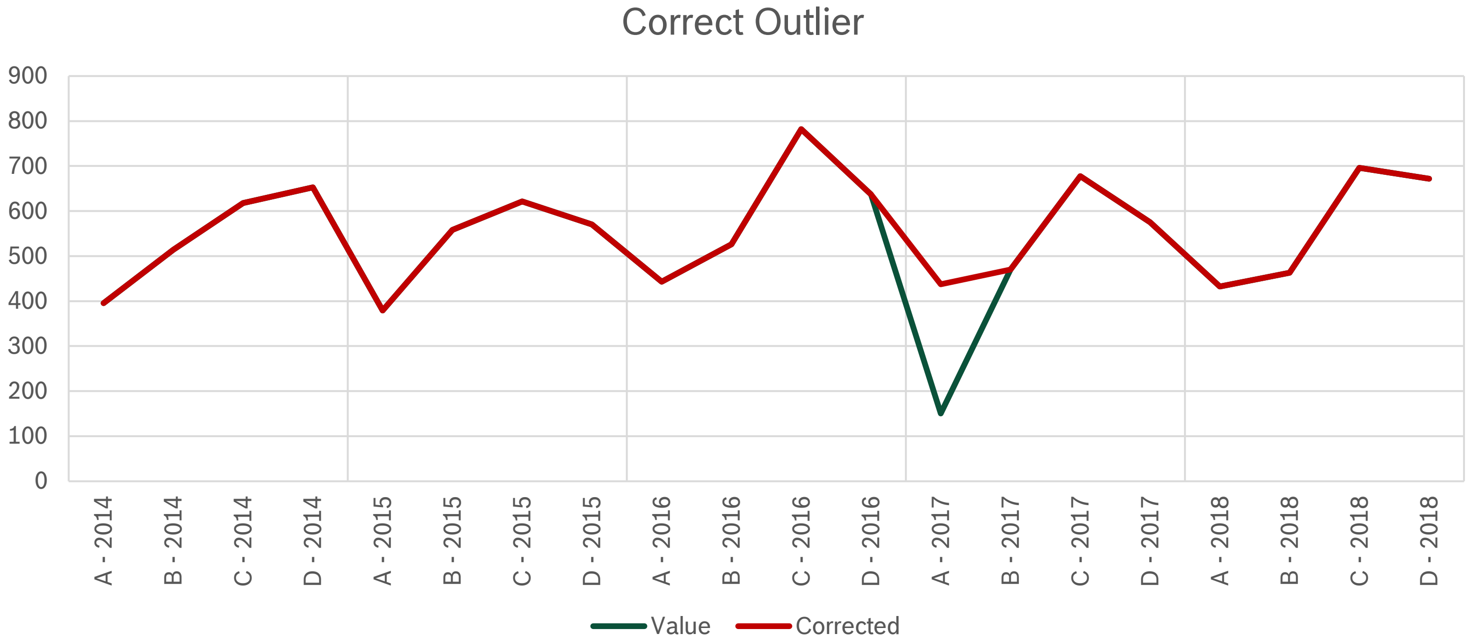
Data	
Year	**Value**
A - 2014	395,5
B - 2014	514,4
C - 2014	617,5
D - 2014	652,8
A - 2015	379,4
B - 2015	558,1
C - 2015	621,7
D - 2015	570,3
A - 2016	442,9
B - 2016	526,1
C - 2016	782,3
D - 2016	636,9
A - 2017	150,0
B - 2017	469,9
C - 2017	677,7
D - 2017	575,0
A - 2018	432,1
B - 2018	462,9
C - 2018	696,2
D - 2018	672,2



- What you can observe?
  - 20 data points, 5 years
  - Quarterly data
  - Quarter A 2017 → too low
    - Outlier due to SEA

Outliers / Special Events & Actions
Removing Seasonal Outlier

Data	
Year	Value
A - 2014	395,5
B - 2014	514,4
C - 2014	617,5
D - 2014	652,8
A - 2015	379,4
B - 2015	558,1
C - 2015	621,7
D - 2015	570,3
A - 2016	442,9
B - 2016	526,1
C - 2016	782,3
D - 2016	636,9
A - 2017	150,0
B - 2017	469,9
C - 2017	677,7
D - 2017	575,0
A - 2018	432,1
B - 2018	462,9
C - 2018	696,2
D - 2018	672,2



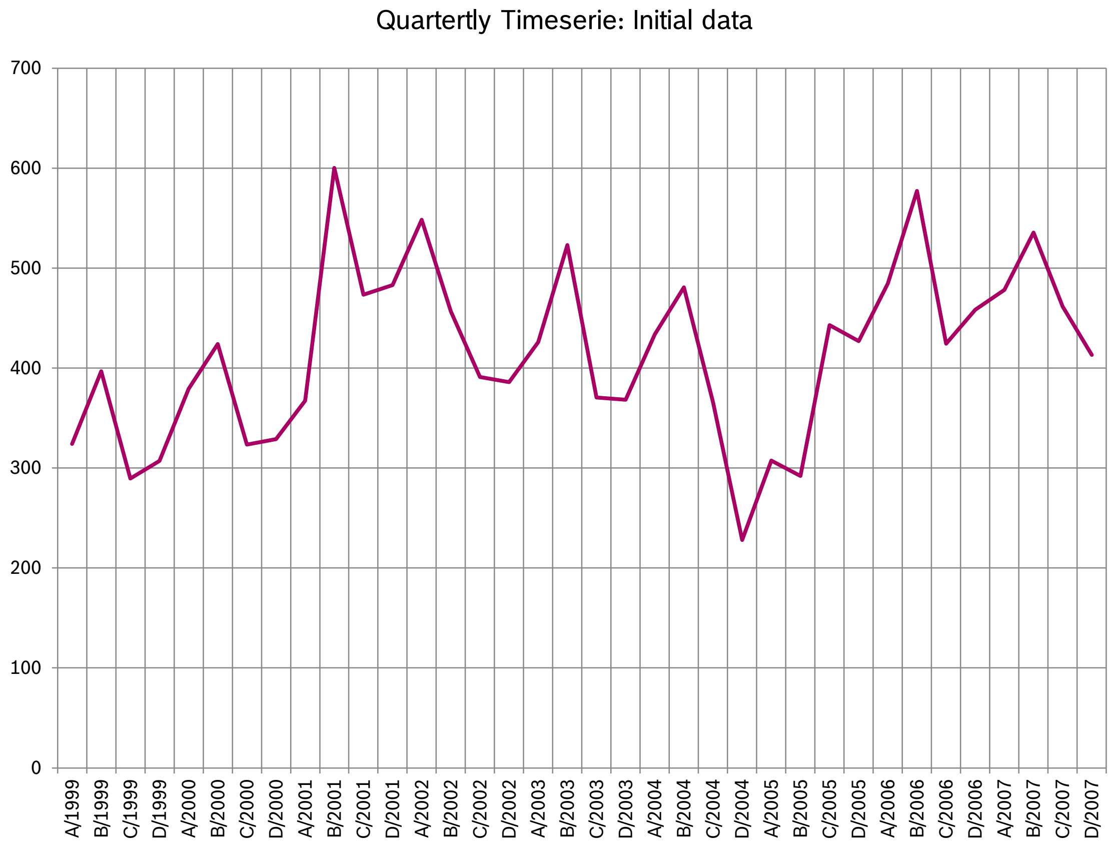
Data	
Year	Corrected
A - 2014	395,5
B - 2014	514,4
C - 2014	617,5
D - 2014	652,8
A - 2015	379,4
B - 2015	558,1
C - 2015	621,7
D - 2015	570,3
A - 2016	442,9
B - 2016	526,1
C - 2016	782,3
D - 2016	636,9
A - 2017	437,5
B - 2017	469,9
C - 2017	677,7
D - 2017	575,0
A - 2018	432,1
B - 2018	462,9
C - 2018	696,2
D - 2018	672,2
Impact	-65,7%

$$D'_{A\ 2007} = (D_{A\ 2006} + D_{A\ 2008})/2 = (442{,}9 + 432{,}1)/2 = 437{,}5$$

$$Impact_{A\ 2007} = \frac{150 - 437{,}5}{437{,}5} * 100\ (\%) = -65{,}7\%$$

Outliers / Special Events & Actions
Detection Methods

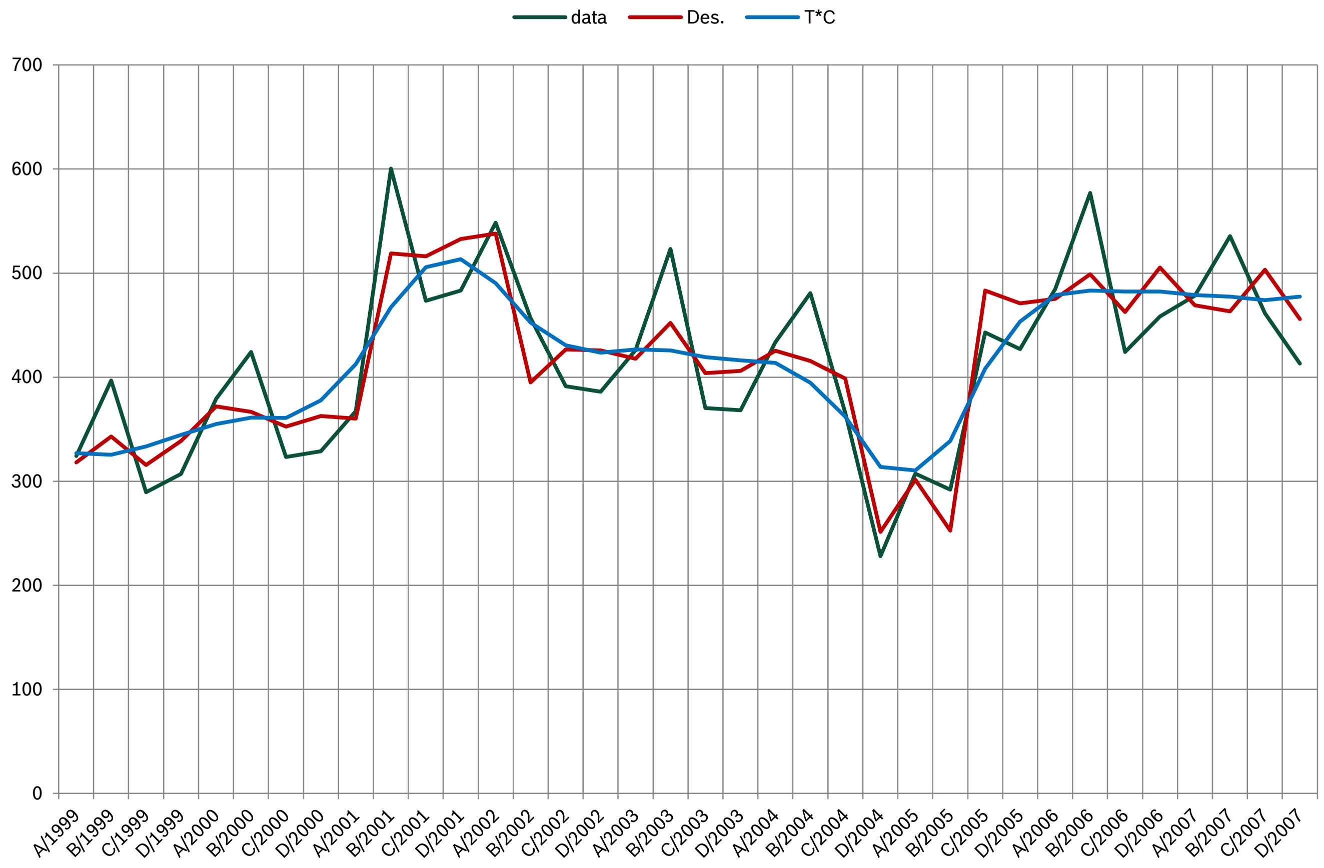
- Use case:
  - A quarterly timeserie with 36 data points, 9 years.
  - Do we see impact of SEA?
  - How we can locate them?
  - Can we use some detection methods?



Outliers / Special Events & Actions
Detection methods

Detection methods:

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



 Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

BOSCH

Outliers / Special Events & Actions
Detection Method A

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

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

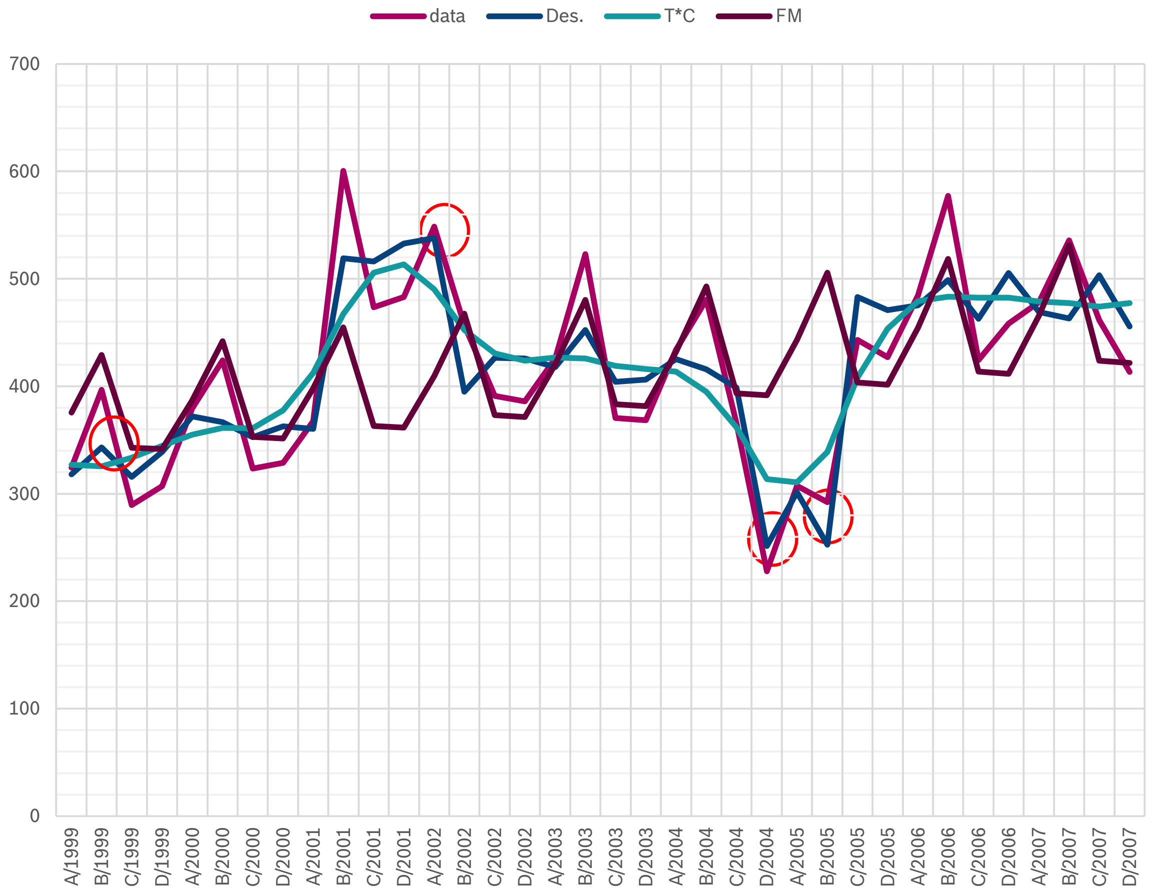
- Given that $t_a \leq 10$ and $t_\beta \leq 25$, the value of the period *t* of the timeserie is an unusual value when:

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

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

Outliers / Special Events & Actions
Detection Method A

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



Outliers / Special Events & Actions
Detection Method B

- 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:

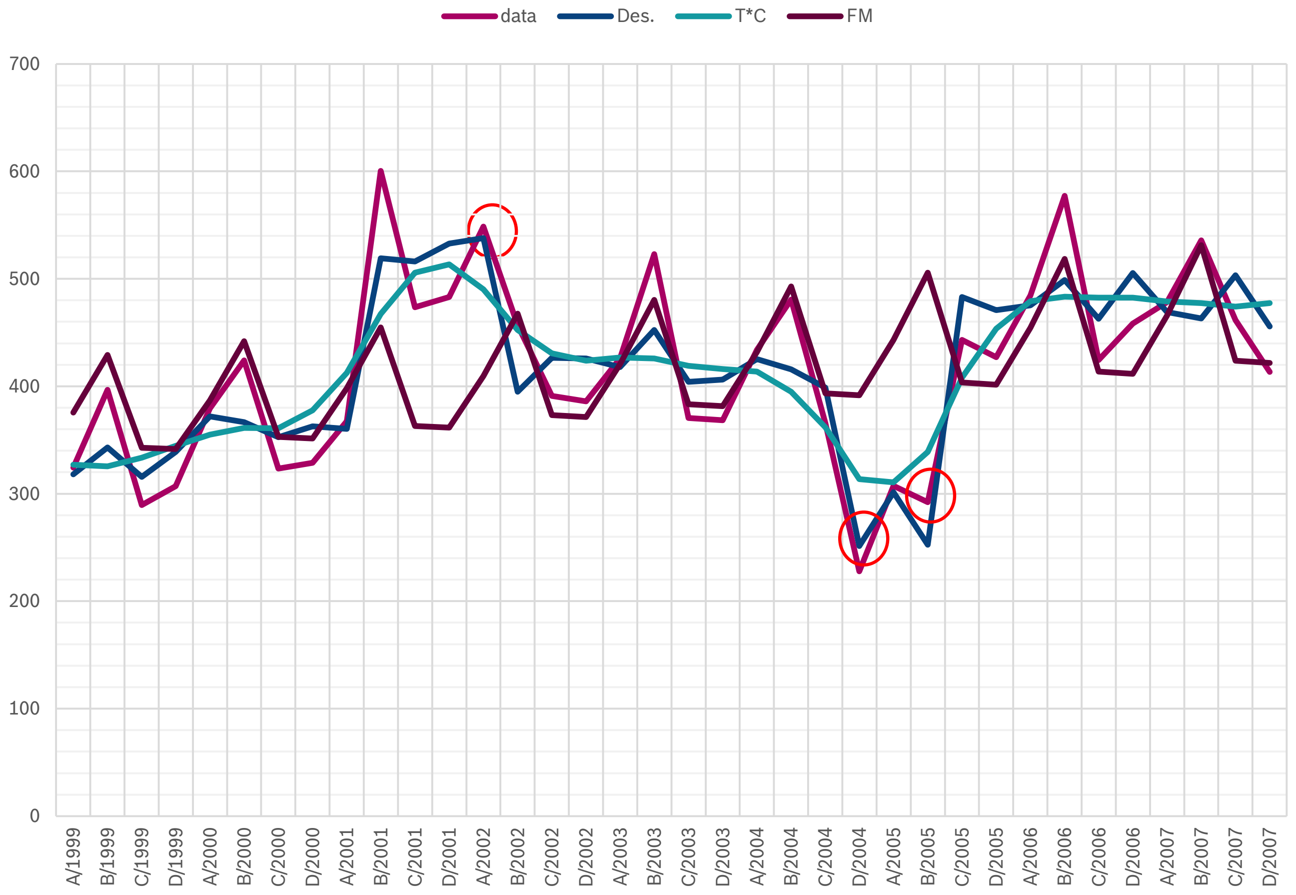
$$\begin{Bmatrix} D_t \geq \bar{D} + (3 - t_a) * StD_F \\ or \\ D_t \leq \bar{D} - (3 - t_a) * StD_F \end{Bmatrix}$$

Outliers / Special Events & Actions

Detection Method B

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

$$StD_F = 49{,}93$$



Outliers / Special Events & Actions
Detection Method C

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

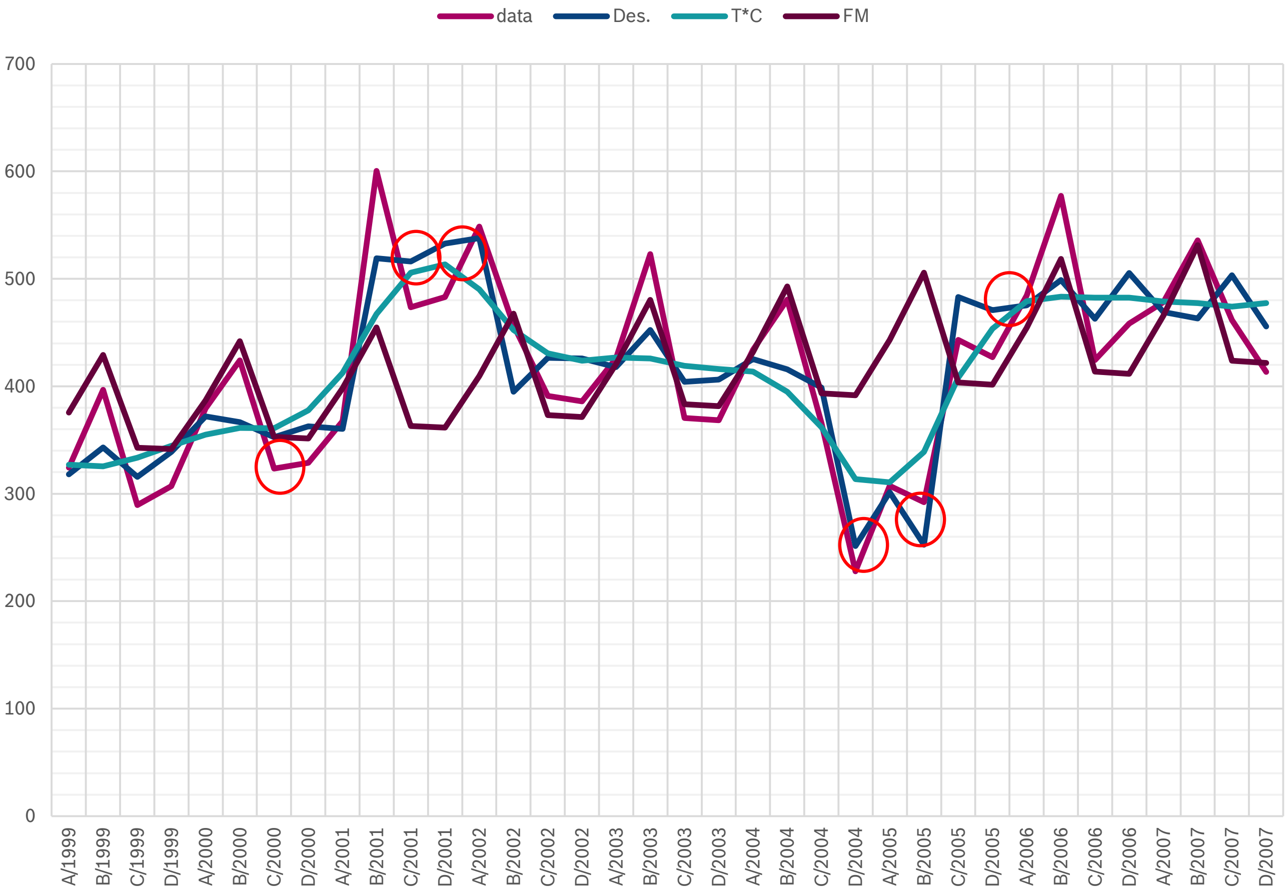
$$Ratio_t = \frac{CMA(7)_t^D}{CMA(5)_t^D}$$

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

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

Outliers / Special Events & Actions
Detection Method C

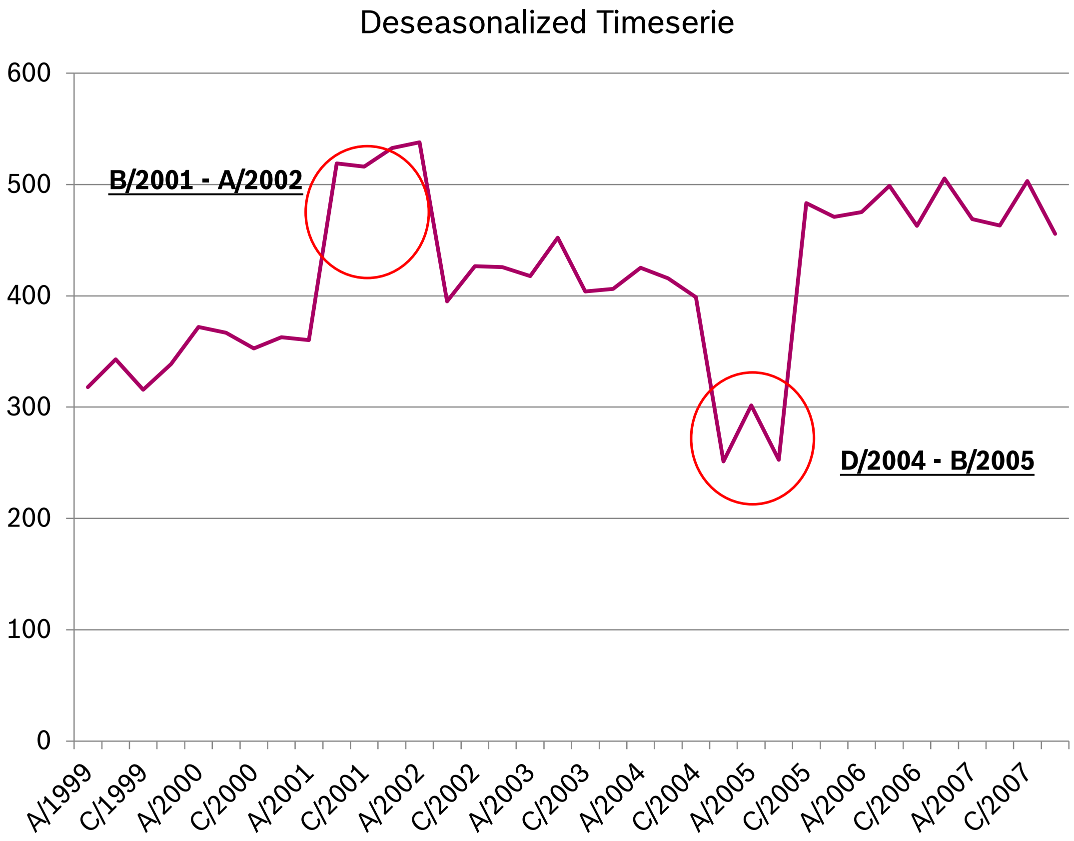
month	data	Deseas.	MA(5)	MA(7)	Ratio	is SEA?
A/1999	324,25	317,96				NO
B/1999	396,82	343,12				NO
C/1999	289,42	315,61	337,49			NO
D/1999	307,17	338,78	347,24	343,83	0,990	NO
A/2000	379,36	372,00	349,14	350,23	1,003	NO
B/2000	424,10	366,71	358,57	352,68	0,984	NO
C/2000	323,36	352,62	362,87	381,74	1,052	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



BOSCH

Outliers / Special Events & Actions
Identify

- 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.



Outliers / Special Events & Actions
Estimate smoothed TS & Impact

- 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).

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

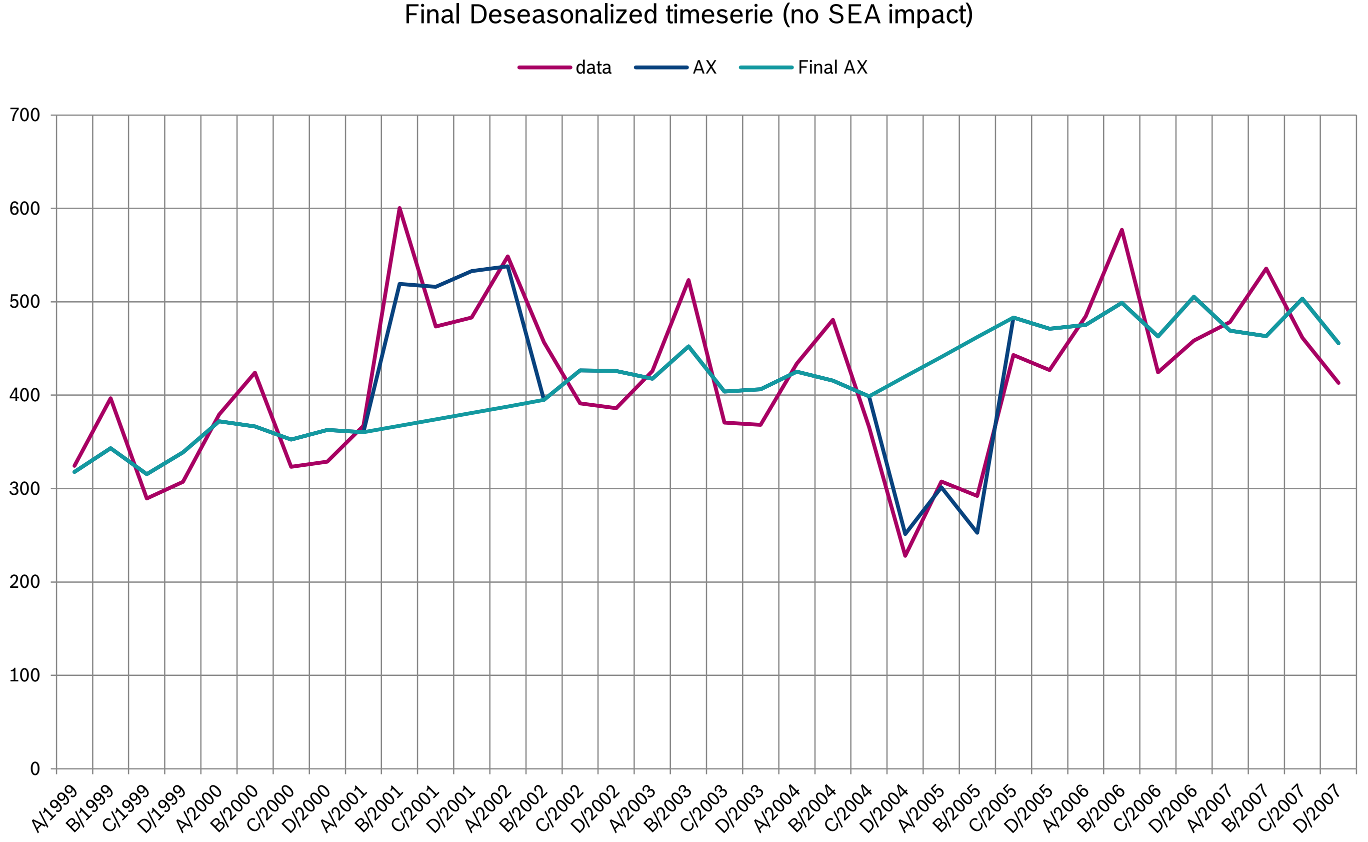
  - For data with ***strong trend***, the smoothed value of a period t (which has an unusual value due to a SEA) is calculated by a ***linear interpolation*** based on ***the preceding and the following period*** of the SEA.

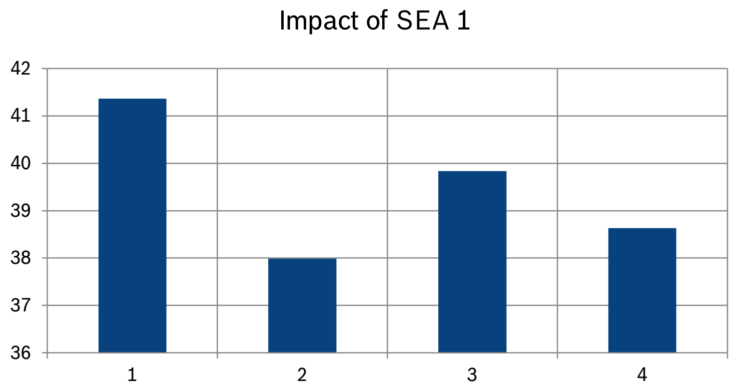
$$D'_t = (t - t_o) * \frac{D_{t_o + n+1} - D_{t_o}}{n+1} + D_{t_o}$$

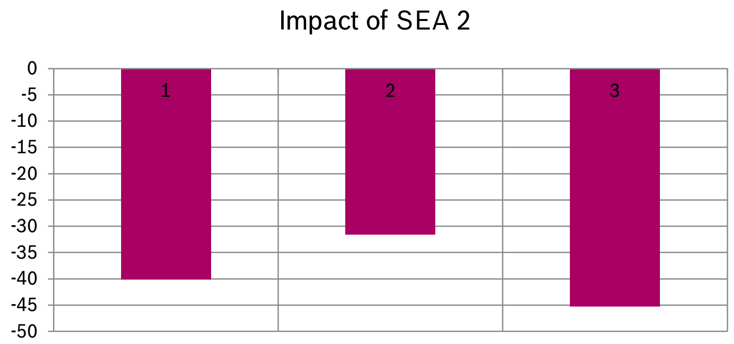
- The impact of an SEA for each period is calculated as the ***% ratio*** between the difference of the initial and the smoothed value and the smoothed value

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

Outliers / Special Events & Actions
Estimate smoothed TS & Impact







Outliers / Special Events & Actions
Practical Exercise

- Exercise steps:
  - Open files “**Forecasting Example - 70d - Outlier (Yearly) Exercise.xlsx**” and “Forecasting Example – 70e - Outlier (Quarterly) Exercise.xlsx”
  - Fill in your own Yearly or Quarterly data
    - Random input
    - A real example
  - Add manually an outlier:
    - Estimate cleaned data
    - Estimate impact
- Duration:
  - **5 minutes**
- Your Goal:
  - Get familiar with Outlier removal
  - Observe how data are cleaned
- Who wants to present his exercise results?

BOSCH

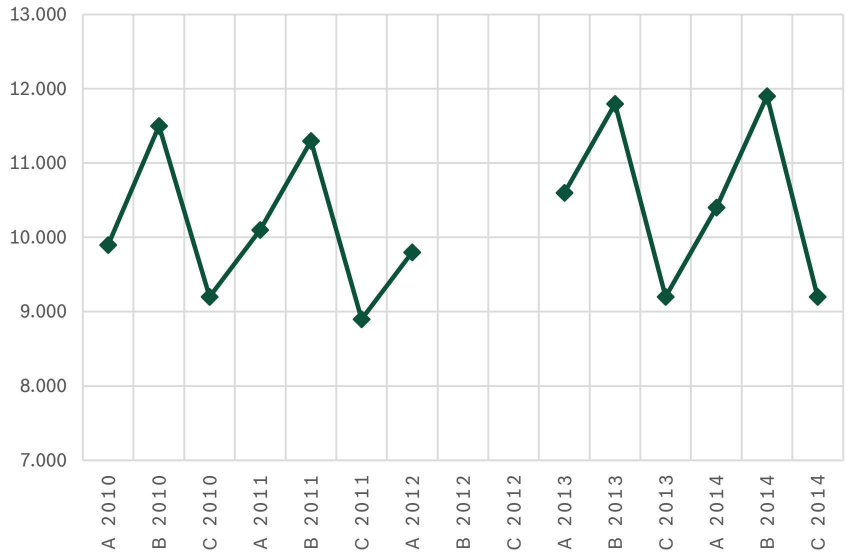
14. SPECIAL TOPIC

MISSING VALUES

BOSCH

Missing Values
Managing Fault/Missing values

- 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.



Missing Values
Managing Fault/Missing values

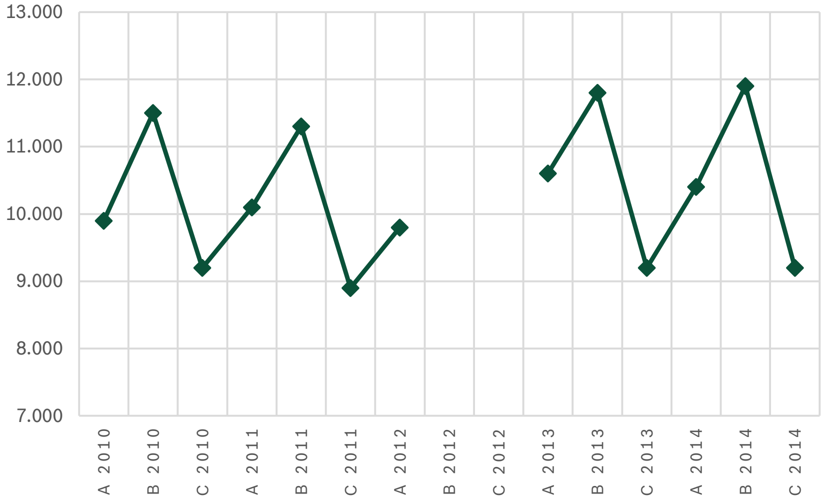
- If a timeserie is characterized by ***stagnation*** and has ***no seasonal behavior***:
  - the missing value can be defined as the ***average*** of the previous and the following values.
- If a timeserie has a clear ***seasonal behavior*** but ***no trend***:
  - the missing value can be defined as 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.

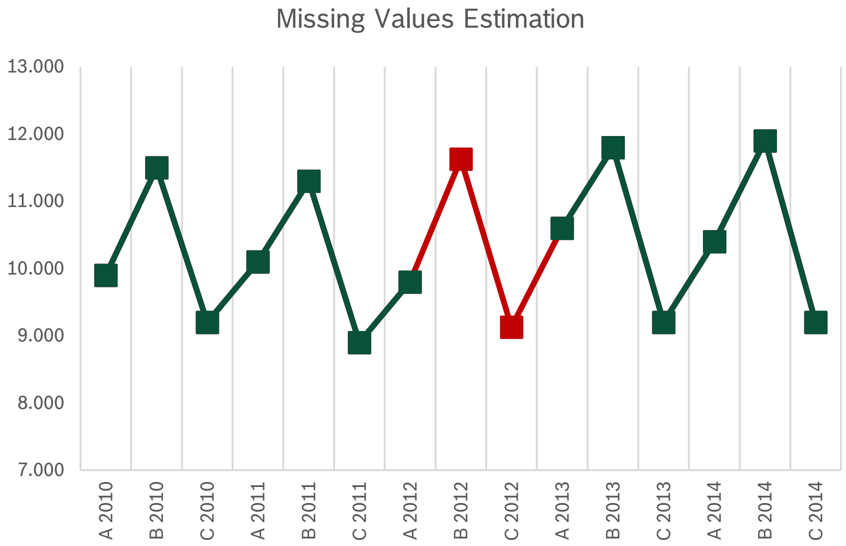
Missing Values

Managing Fault/Missing values

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

Period	Value
A 2010	9.900
B 2010	11.500
C 2010	9.200
A 2011	10.100
B 2011	11.300
C 2011	8.900
A 2012	9.800
B 2012	???
C 2012	
A 2013	10.600
B 2013	11.800
C 2013	9.200
A 2014	10.400
B 2014	11.900
C 2014	9.200

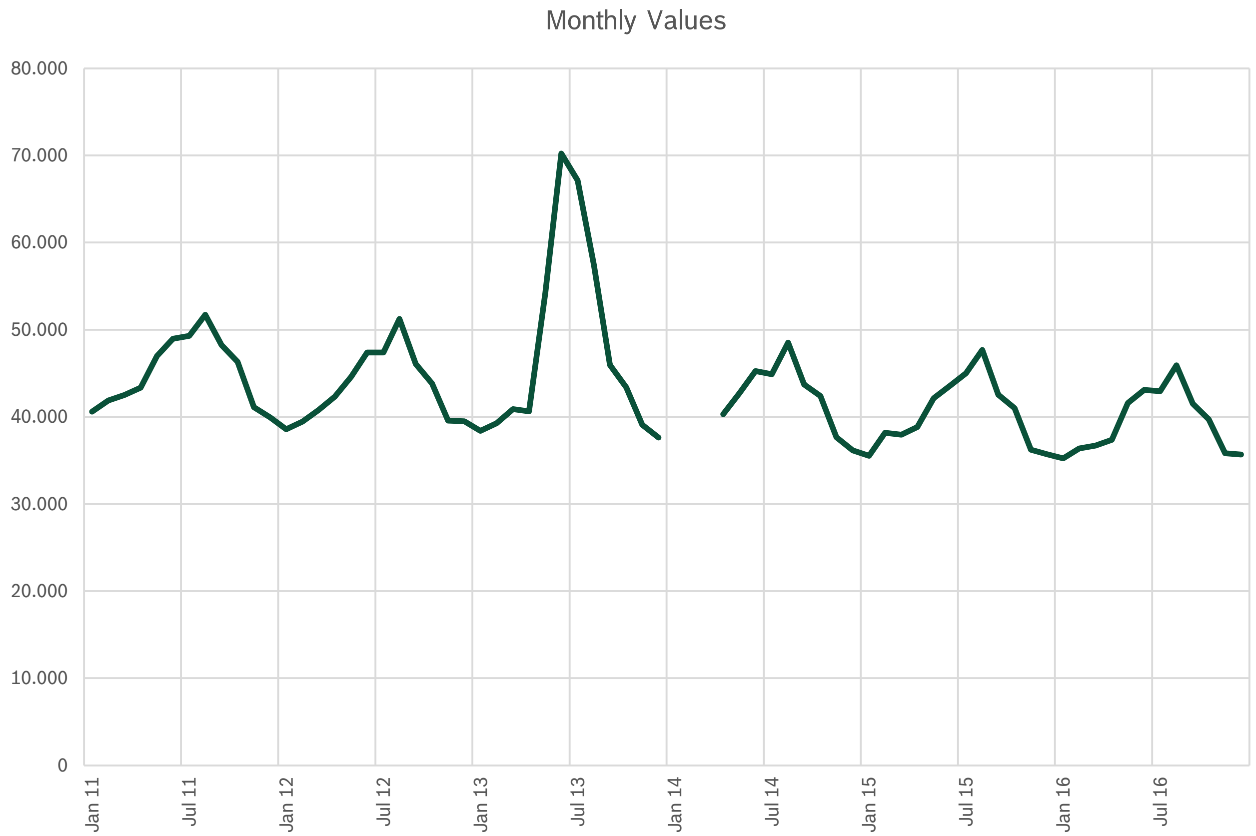




The timeserie has no trend, but has seasonality. Thus:

- 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

Missing Values
Observe missing values



- What you can observe?
  - 72 data points, Monthly data
  - Decreasing trend
  - Jan-Mar 2014 → missing

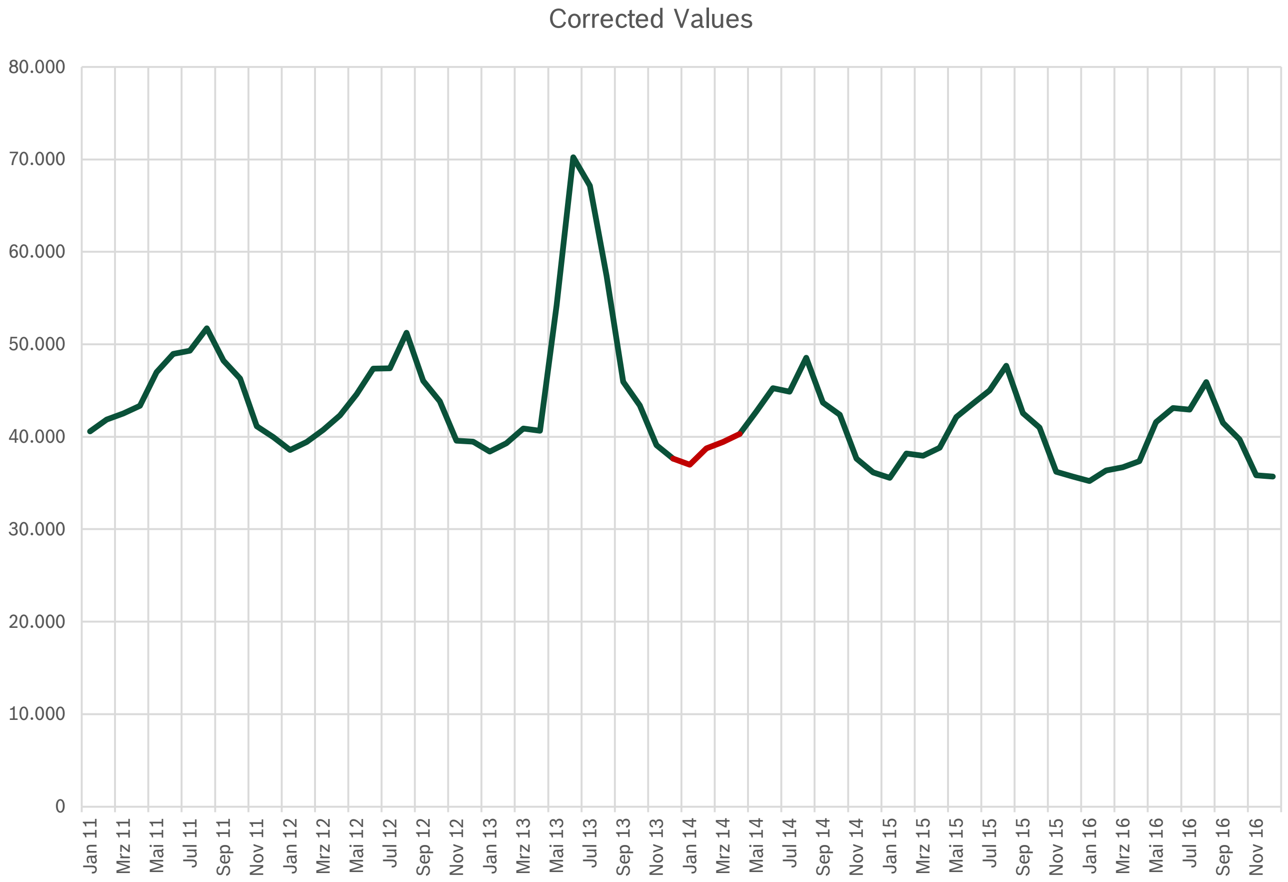
	Initial Data					
	2011	**2012**	**2013**	**2014**	**2015**	**2016**
Jan	40.583	38.581	38.381		35.540	35.219
Feb	41.866	39.451	39.287		38.187	36.357
Mar	42.507	40.766	40.889		37.953	36.696
Apr	43.363	42.299	40.633	40.313	38.819	37.364
Mai	46.978	44.596	54.224	42.734	42.136	41.575
Jun	48.976	47.369	70.233	45.260	43.578	43.102
Jul	49.300	47.402	67.146	44.876	45.010	42.950
Aug	51.724	51.251	57.408	48.539	47.678	45.929
Sep	48.246	46.065	45.944	43.712	42.551	41.512
Oct	46.301	43.827	43.390	42.396	41.003	39.700
Nov	41.119	39.557	39.099	37.640	36.205	35.829
Dec	39.938	39.479	37.634	36.159	35.699	35.690

Missing Values
Calculating

	Initial Data & Correction					
	2011	**2012**	**2013**	**2014**	**2015**	**2016**
Jan	40.583	38.581	38.381	36.961	35.540	35.219
Feb	41.866	39.451	39.287	38.737	38.187	36.357
Mar	42.507	40.766	40.889	39.421	37.953	36.696
Apr	43.363	42.299	40.633	40.313	38.819	37.364
Mai	46.978	44.596	54.224	42.734	42.136	41.575
Jun	48.976	47.369	70.233	45.260	43.578	43.102
Jul	49.300	47.402	67.146	44.876	45.010	42.950
Aug	51.724	51.251	57.408	48.539	47.678	45.929
Sep	48.246	46.065	45.944	43.712	42.551	41.512
Oct	46.301	43.827	43.390	42.396	41.003	39.700
Nov	41.119	39.557	39.099	37.640	36.205	35.829
Dec	39.938	39.479	37.634	36.159	35.699	35.690

$$Jan_{2014} = \frac{1}{2} * (Jan_{2013} + Jan_{2015}) = 36961$$

$$Feb_{2014} = \frac{1}{2} * (Feb_{2013} + Feb_{2015}) = 38737$$



Missing Values
Practical Exercise

- Exercise steps:
  - Open file “**Forecasting Example - 71c - Missing Values (Monthly) Exercise.xlsx**”
  - Fill in your own Monthly data
    - Random input
    - A real example
  - Leave 1-3 values empty!
  - Estimate missing values

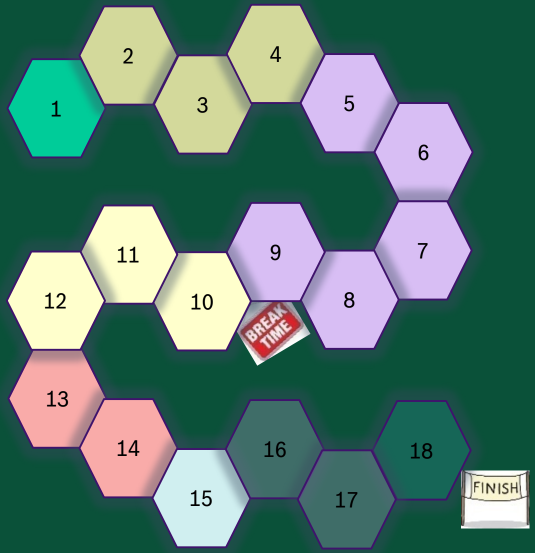
- Duration:
  - **5 minutes**

- Your Goal:
  - Get familiar with missing values
  - Observe how values are estimated

- Who wants to present his exercise results?

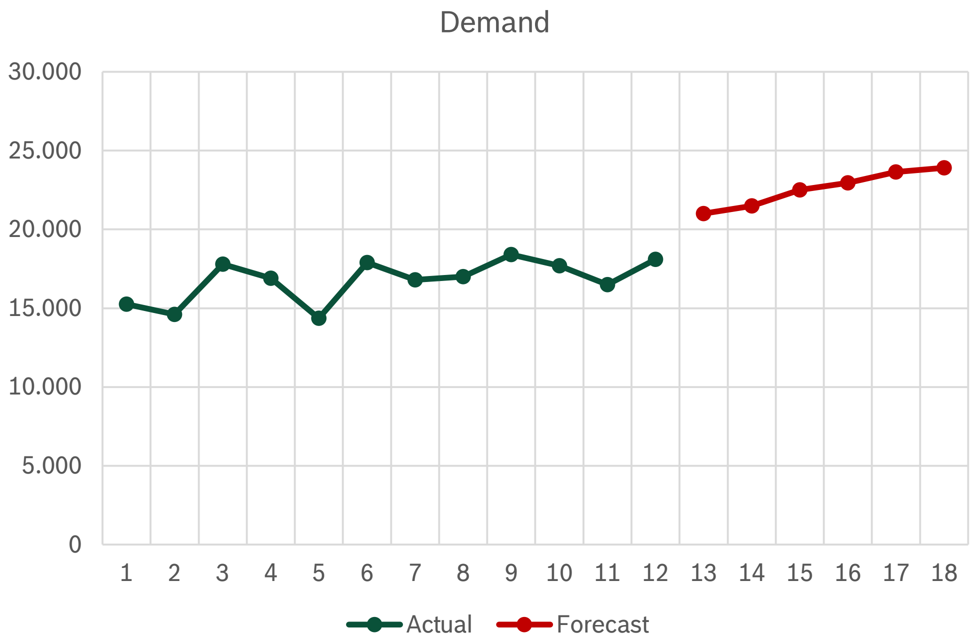
15. USE-CASE

DEMAND FORECASTING



Use-Case
Demand forecasting

- Current situation
  - Historical data are monthly sales (up to 36 months)
  - Goal: Forecast demand for a horizon of 18 months
  - Forecast on Item – Warehouse – Region – Key Account
  - Data have all kind of behavior
    - trends, sporadic, intermittent, weather driven, etc.
  - More than 10 million timeseries
  - Two types of forecast
    - Using best-fit algorithm (including outliers detection and seasonality) for statistically estimate forecasts
    - Adjust forecasts manually (due to launch and campaign demand)

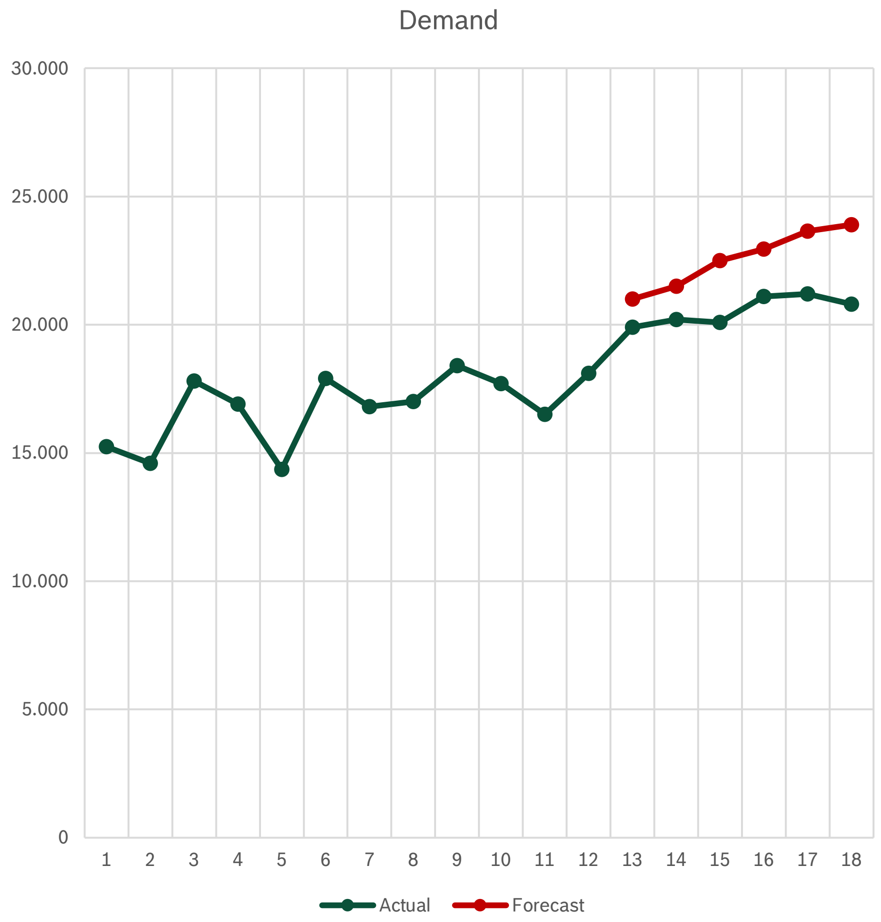


 Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

BOSCH

Use-Case
Demand forecasting

- Current problem
  - Most of the forecasts are very positive biased
  - Forecasts of demands are higher than actual sales
  - Overstocked items!
  - Increased cost!
  - Can we identify the source of the problem?

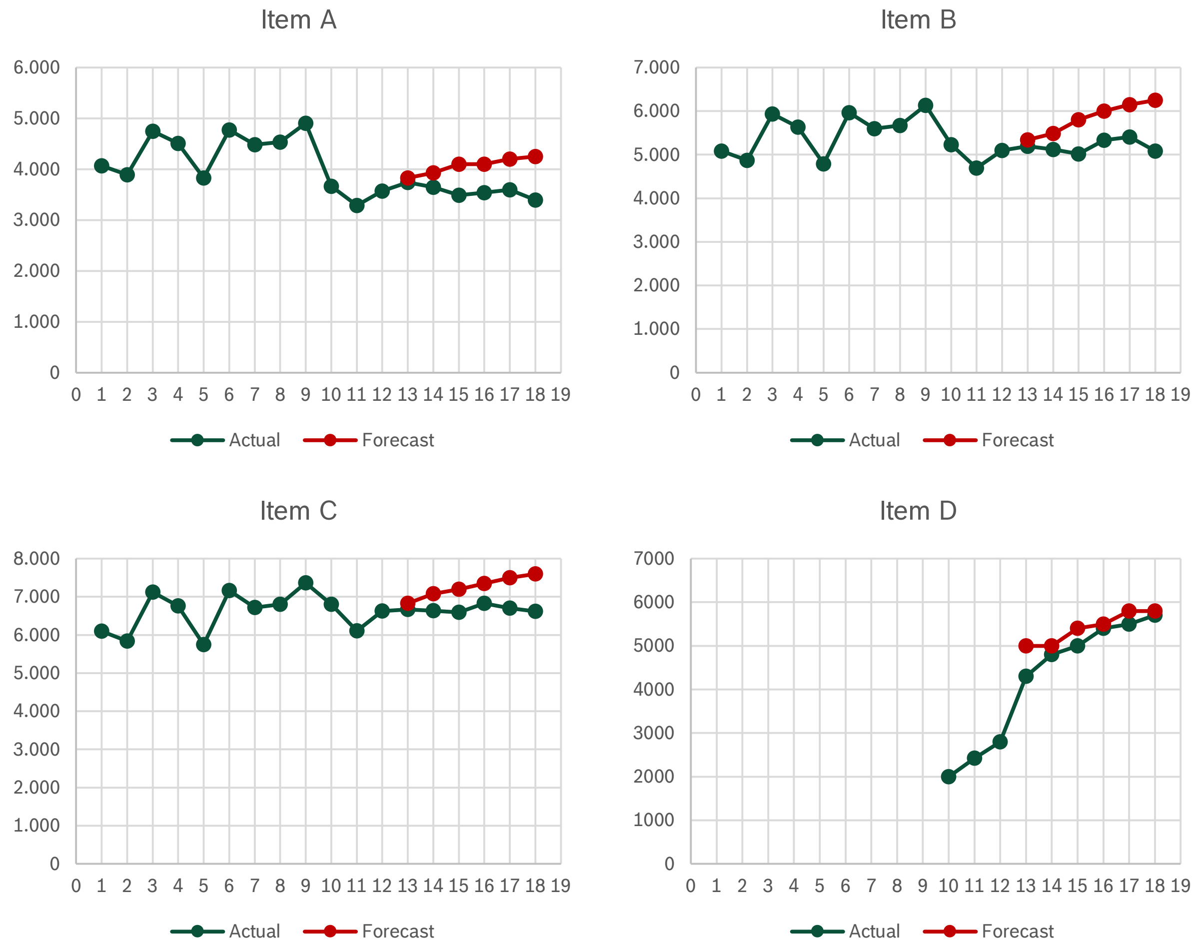


	Actual	Forecast
1	15.250	
2	14.600	
3	17.800	
4	16.900	
5	14.360	
6	17.900	
7	16.800	
8	17.000	
9	18.400	
10	17.700	
11	16.500	
12	18.100	
13	19.896	21.000
14	20.199	21.500
15	20.094	22.500
16	21.095	22.950
17	21.202	23.650
18	20.798	23.900

BOSCH

Use-Case
Demand forecasting

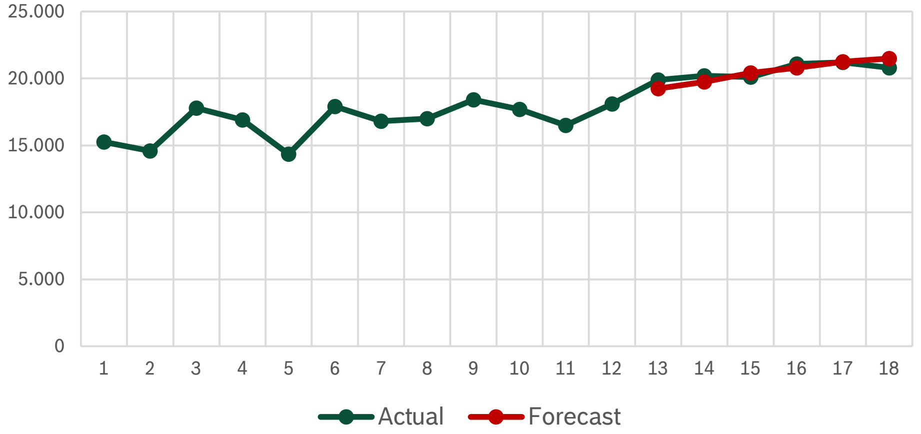
- Identifying the source of the problem
  - Item Group forecasts:
    - Are the sum of each item forecasts
  - Item D (newest model)
    - Forecast: mostly manually (sales expected to go up)
  - Items A, B, C (older models)
    - Forecast: statistically only
    - They are not adjusted it manually (sales should be lower, since Item D has bigger market share)

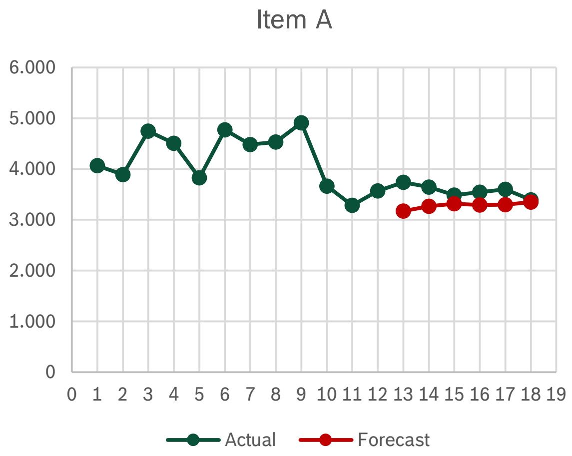


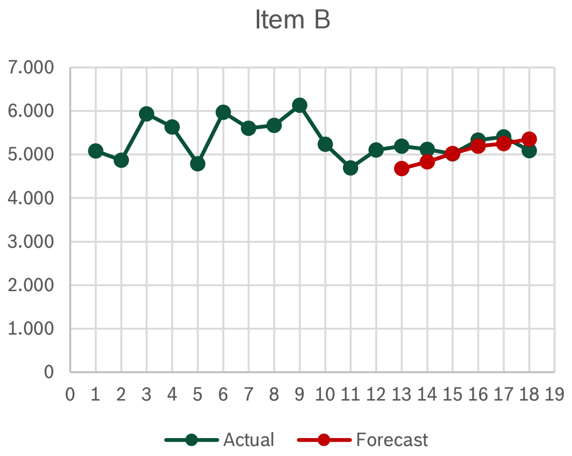
BOSCH

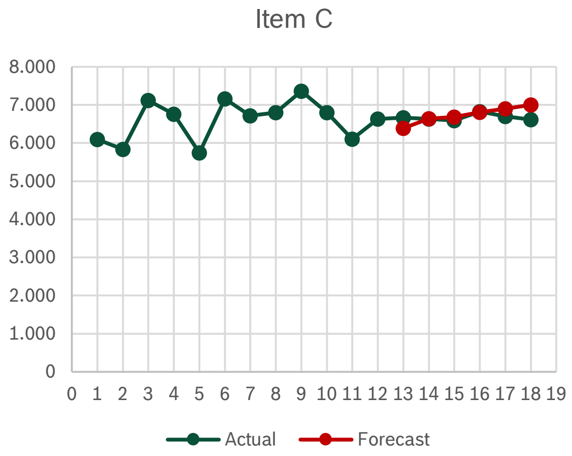
Use-Case
Demand forecasting

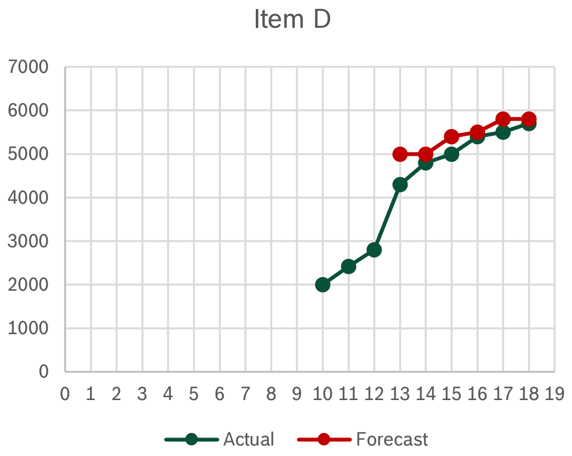
- Solving the problem
  - Forecast manually for Item D
  - Statistically forecast Items A, B, C
  - Adjust statistical forecast manually, in order to keep total forecast in an acceptable level.





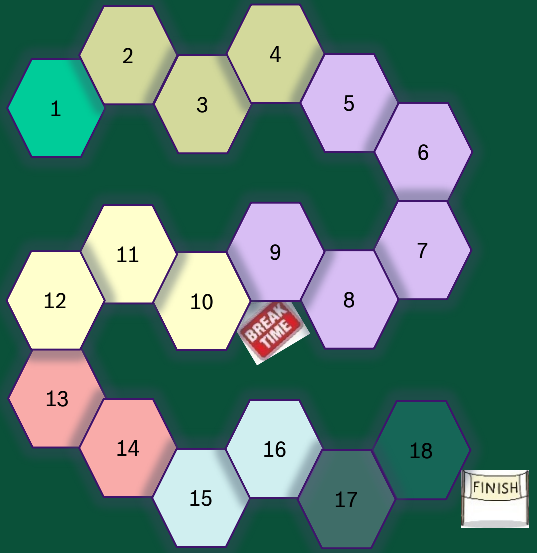






16. USE-CASE

FORECASTING ENERGY CONSUMPTION OF BUILDINGS



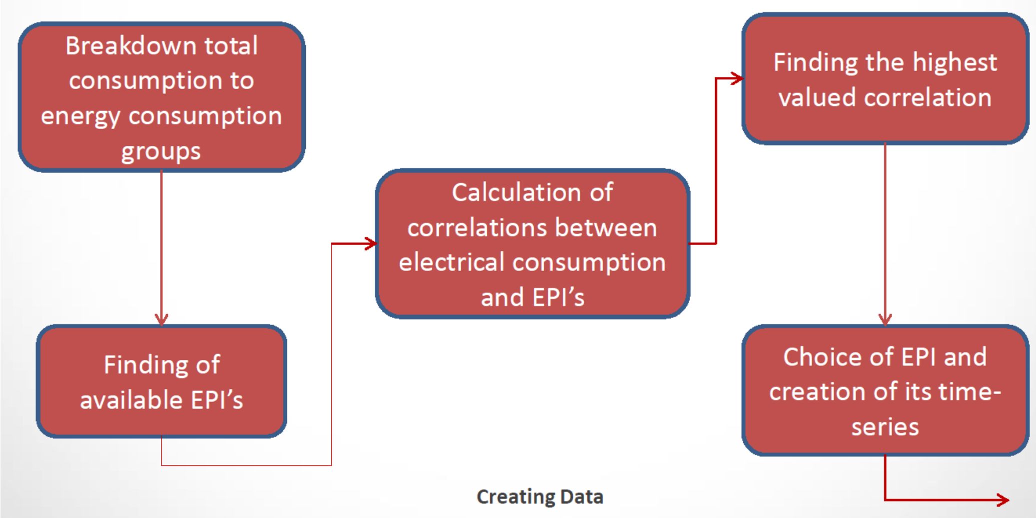
BOSCH

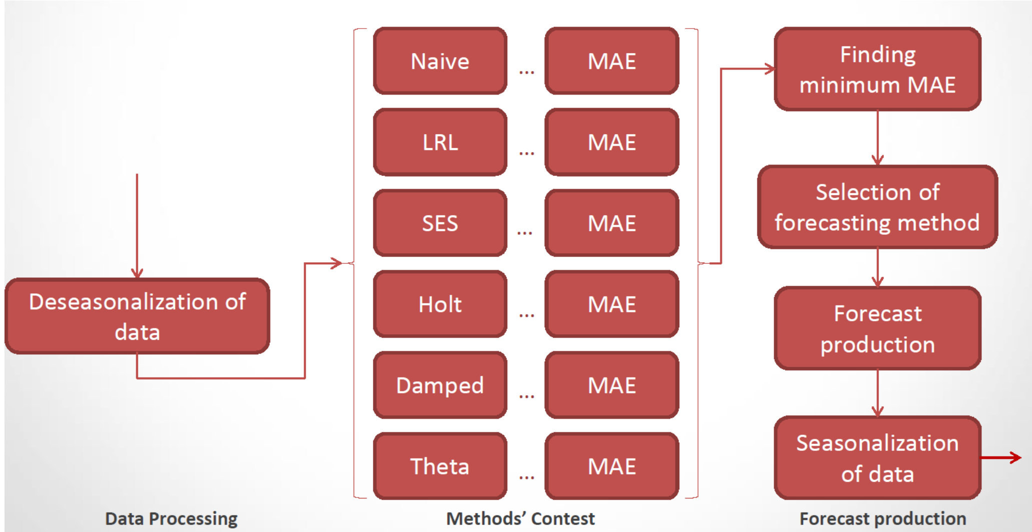
Use-Case
Forecasting Energy Consumption of Buildings (1)

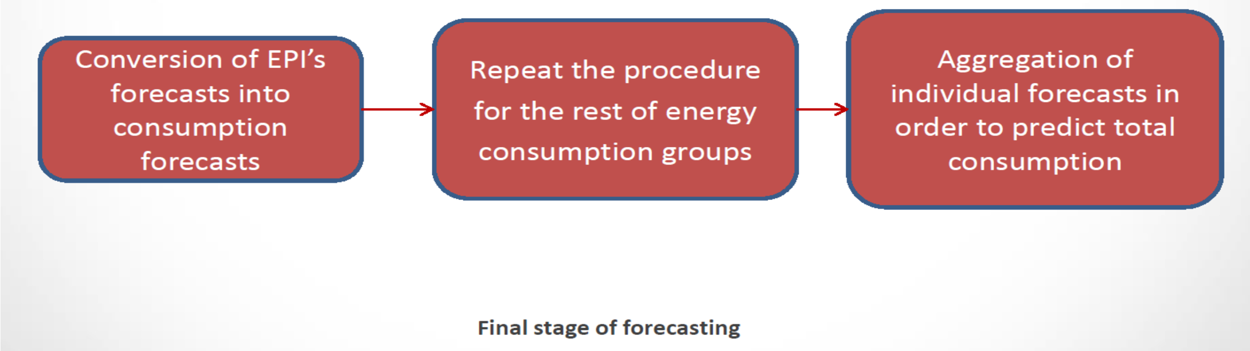
- ***Scope***: Use performance indicators for forecasting consumption in energy-intensive buildings.
  - Three types of indicators to investigate:
    - ***Physical Indicators***: Relation between energy consumption and a physical measured object.
    - ***Value Based Indicators***: Relation between 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)*

Use-Case
Forecasting Energy Consumption of Buildings (2)







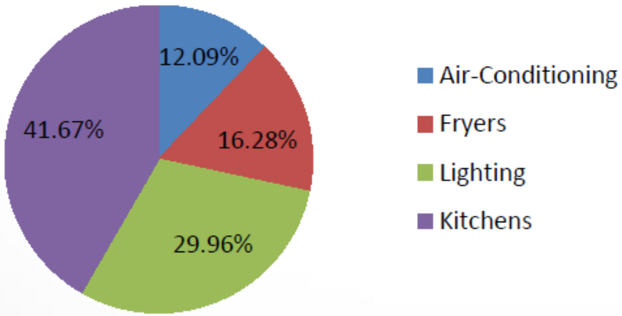
Use-Case
Forecasting Energy Consumption of Buildings (3)

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

Energy Consumption Groups:

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

Energy Consumption Groups



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

Use-Case
Forecasting Energy Consumption of Buildings (4)

- Find correlations in the energy consumption groups

Results for kitchen consumption group(line 1)

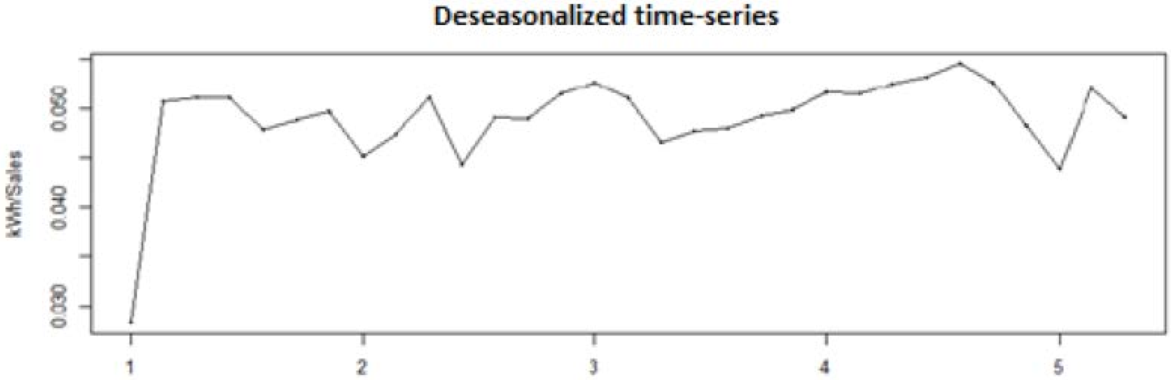
EPI's denominator	Correlation Value
O.H.	-0.194071
% Consumption	-0.239891
Inside Temperature	-0.275588
Outside Temperature	0.530509
Differential Temperature	-0.310188
Sales	0.590446
Receipts	0.340027

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

Use-Case
Forecasting Energy Consumption of Buildings (5)

- Starting Forecast Production (KWh/Sales)



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

BOSCH

Use-Case
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

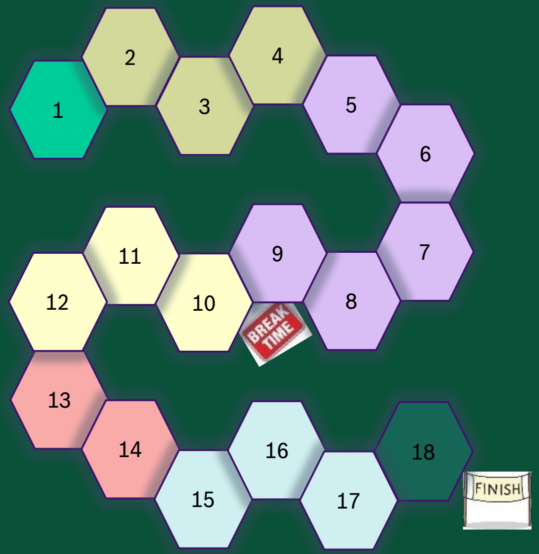
Total accuracy of method

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.

17. USE-CASE

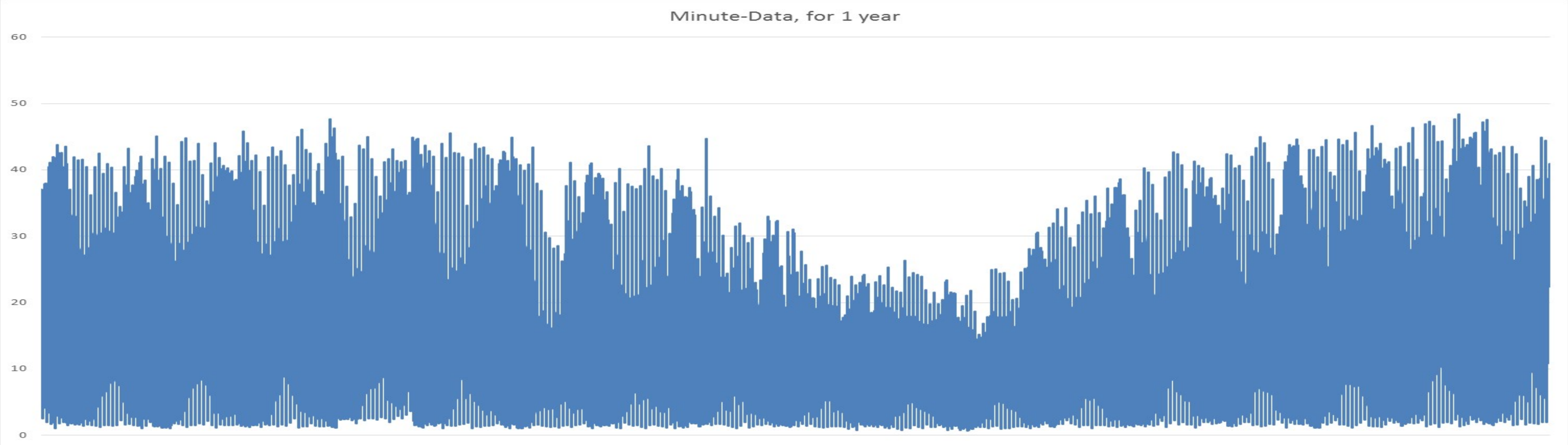
FORECASTING WITH MINUTE DATA



BOSCH

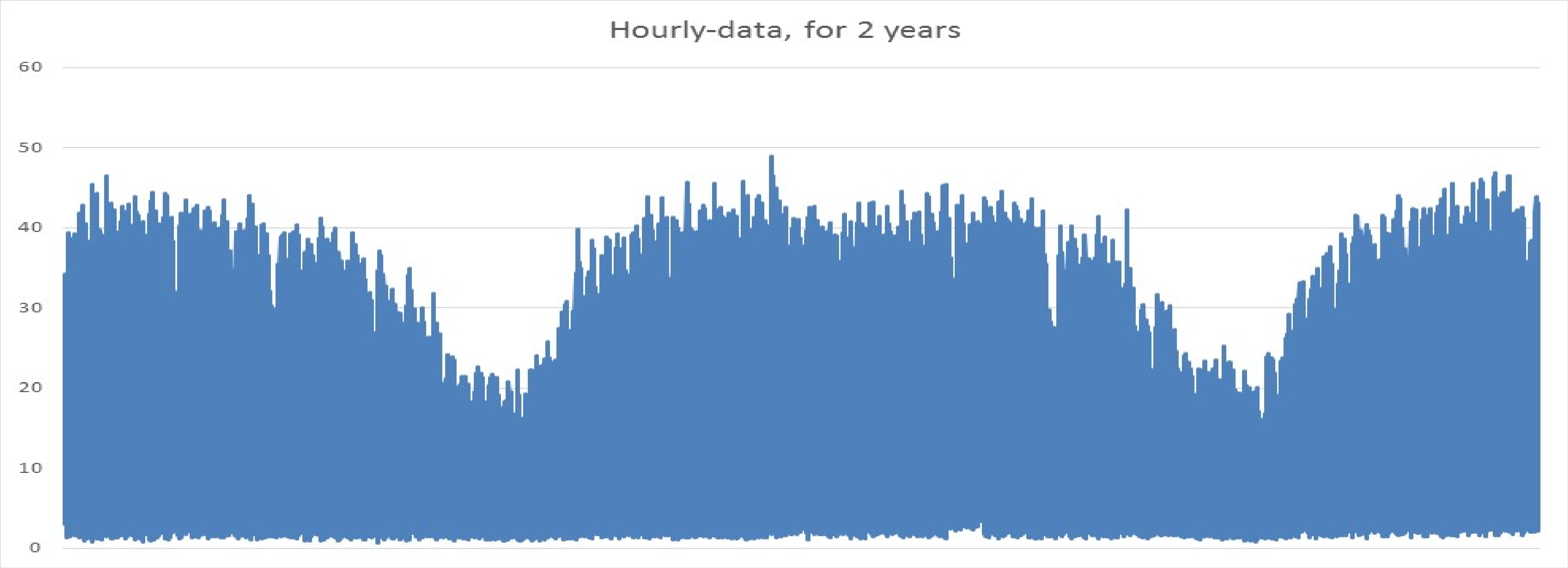
Use-Case
Forecasting with minute data (1)

- 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?



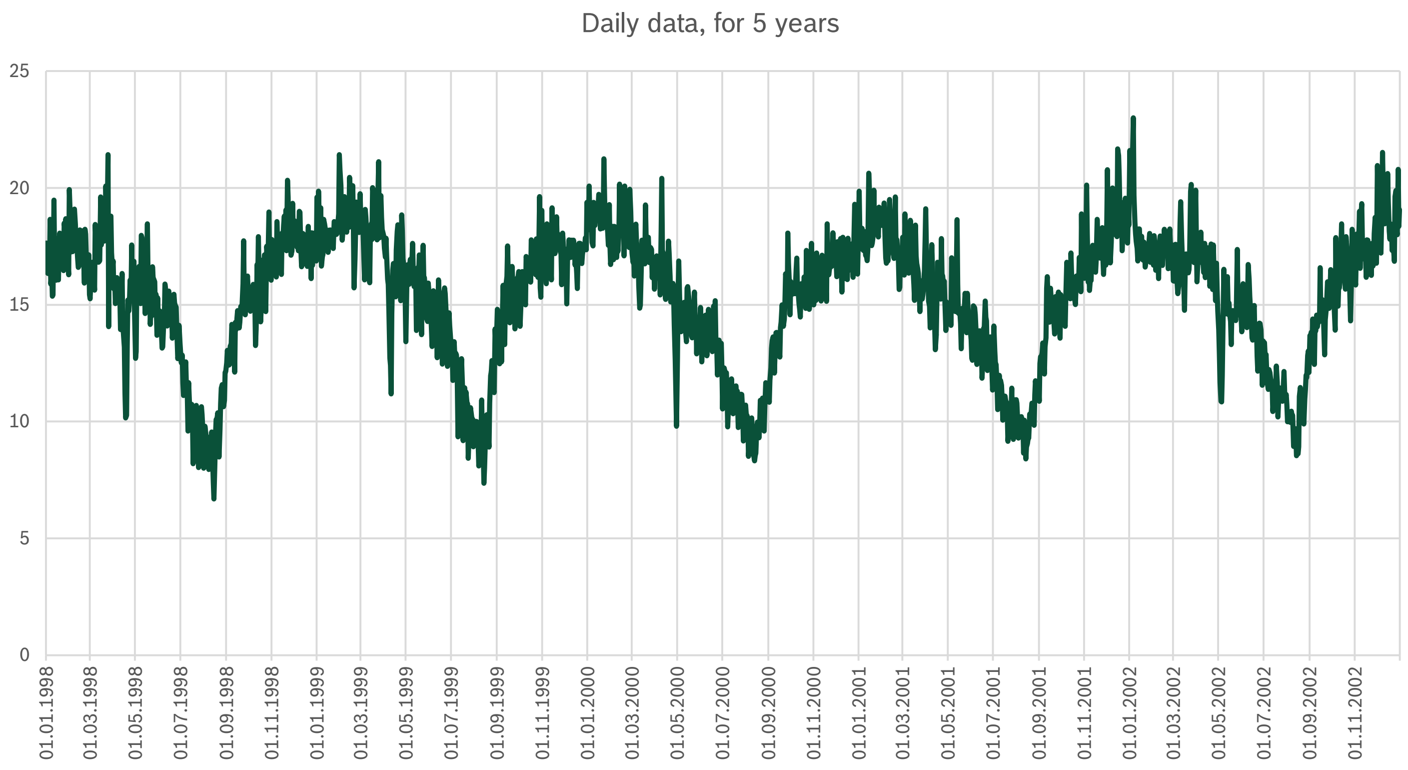
Use-Case
Forecasting with minute data (2)

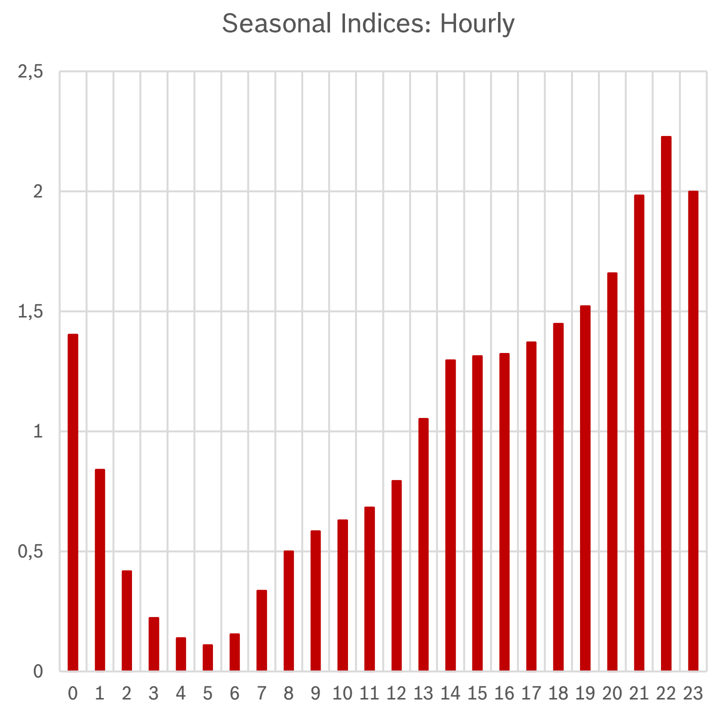
- ***Step 1***: Aggregate minute-data to Hourly-data.



Use-Case
Forecasting with minute data (3)

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

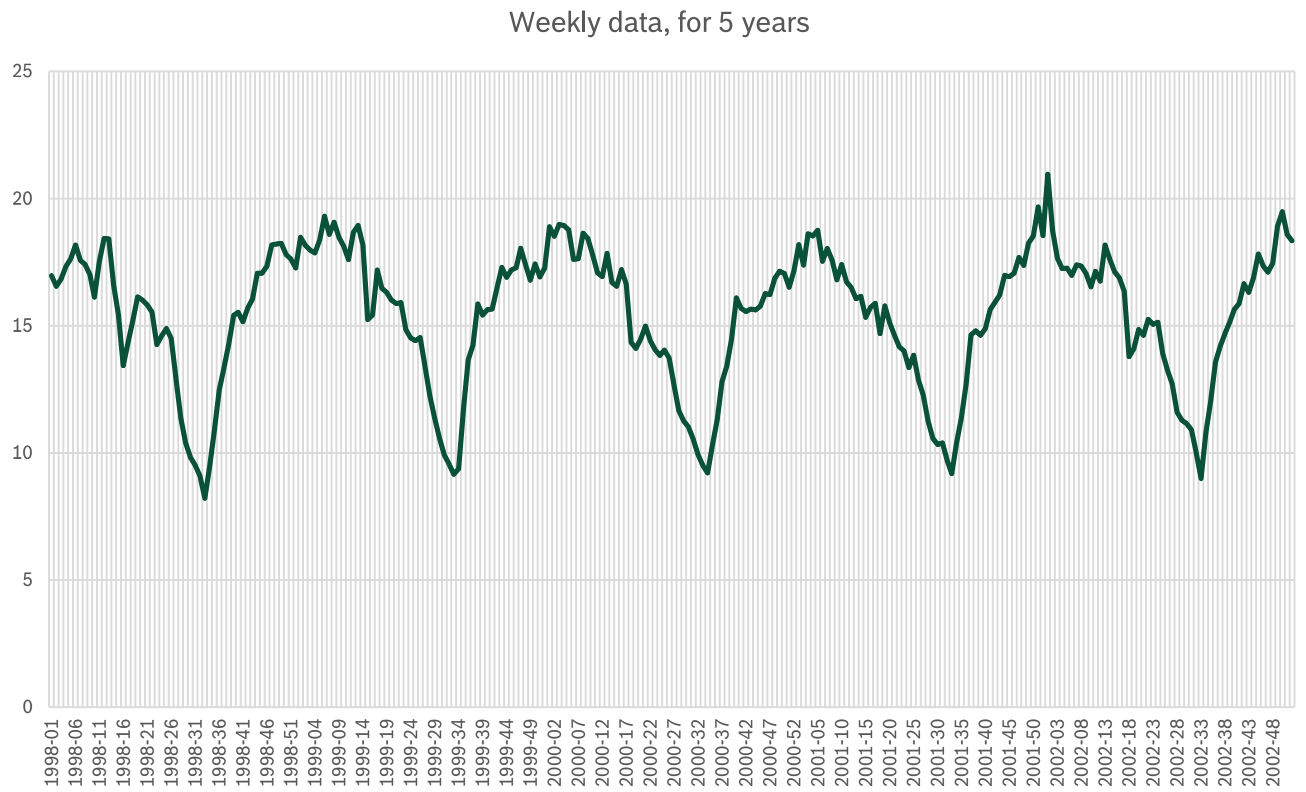


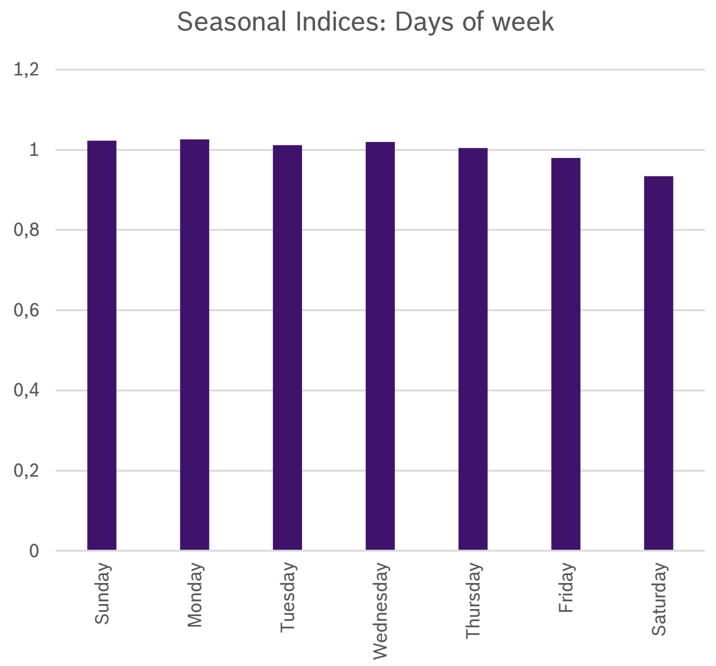


Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

Use-Case
Forecasting with minute data (4)

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

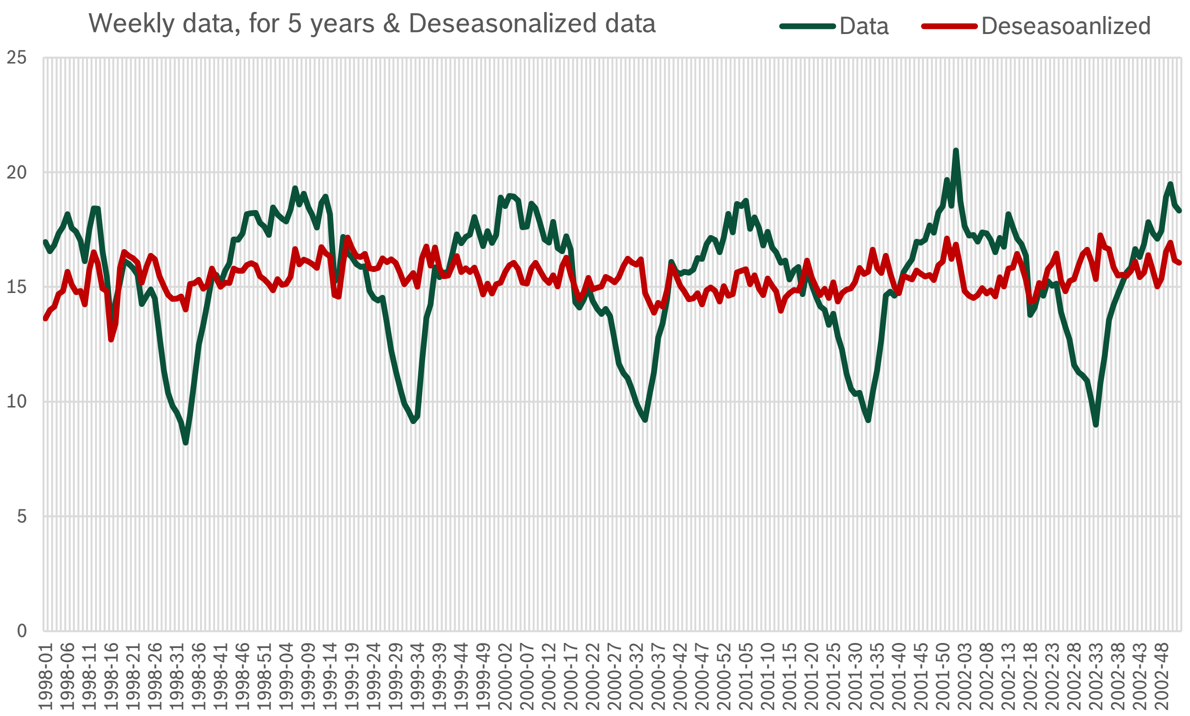


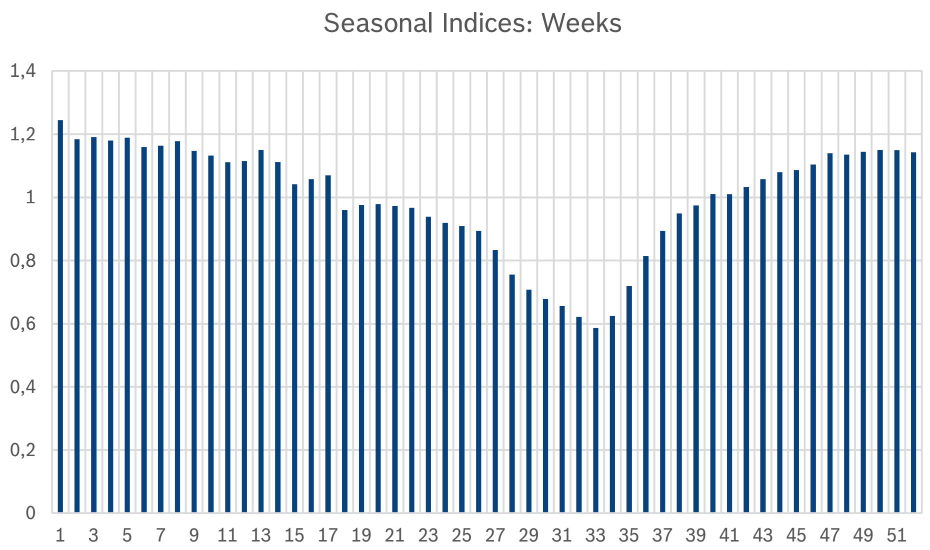


Internal | Patelis Alexandros - CR/AEU2 | 2019-07-01

Use-Case
Forecasting with minute data (5)

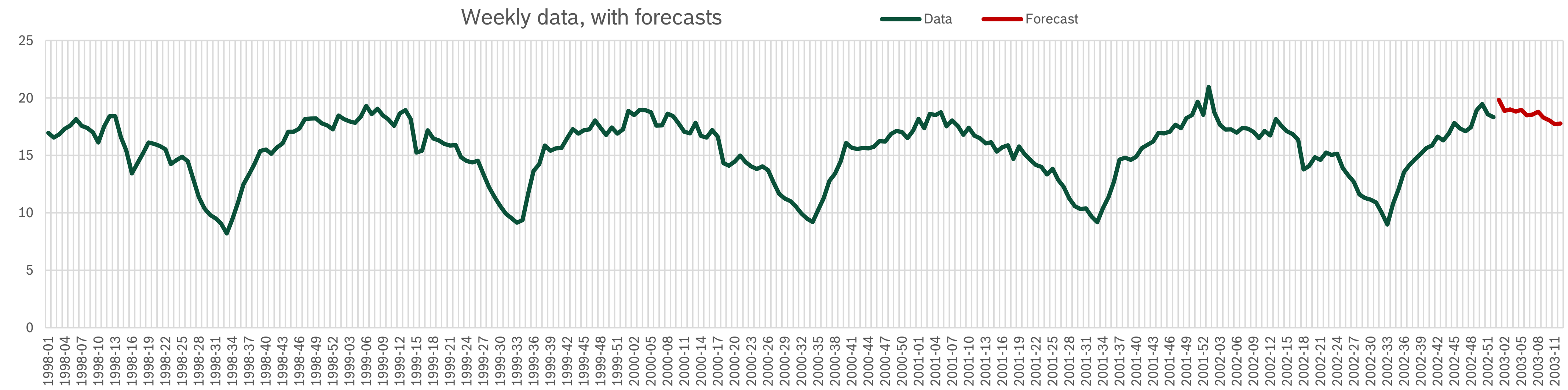
- ***Step 4***: Remove also weekly seasonality, and estimate Weekly seasonal indices.





Use-Case
Forecasting with minute data (6)

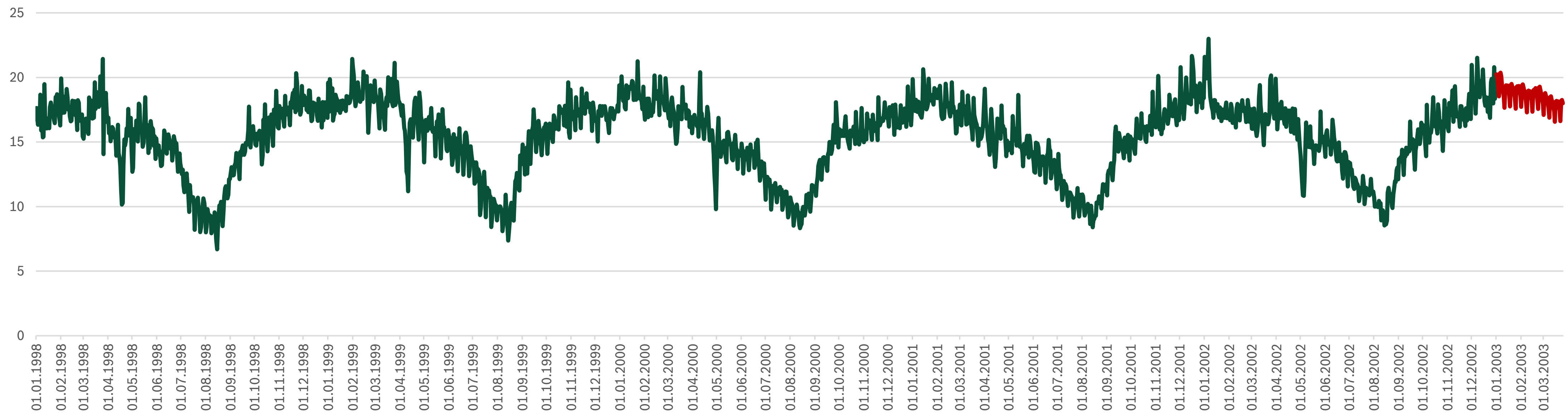
- ***Step 5***:
  - Estimate forecast(s) for next week(s).
  - Re-seasonalize the forecasts, by using the seasonal indexes for weeks.



Use-Case
Forecasting with minute data (7)

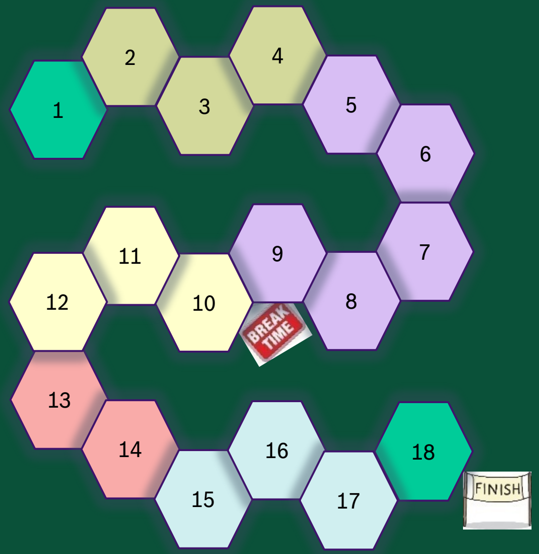
- ***Step 6***:
  - Disaggregate forecasts, from weekly to daily, by using the seasonal indexes for days of week.
  - Disaggregate forecasts, from daily to hourly, by using the seasonal indexes for hours of day.

Daily data, with forecasts



18. MORE?

TIMESERIES FORECASTING (ADVANCED)



Monday, September 23rd, 13:00 – 16:00

Timeseries Forecasting: Advanced

Agenda

1 Recap Beginners

2 Preparation & Analysis

3 Errors & Accuracy (More)

4 Decomposition

5 Theta Classical

6 Theta Alternative

7 Averaging

8 Arima

9 Multiple Regression

10 Limits of Predictability

11 Neural Networks

12 Special Topic Working Days Adj.

13 Special Topic Structural Changes

14 Special Topic Intermittent Demand

15 Use Case Inventory Demand

16 Use Case Energy Management

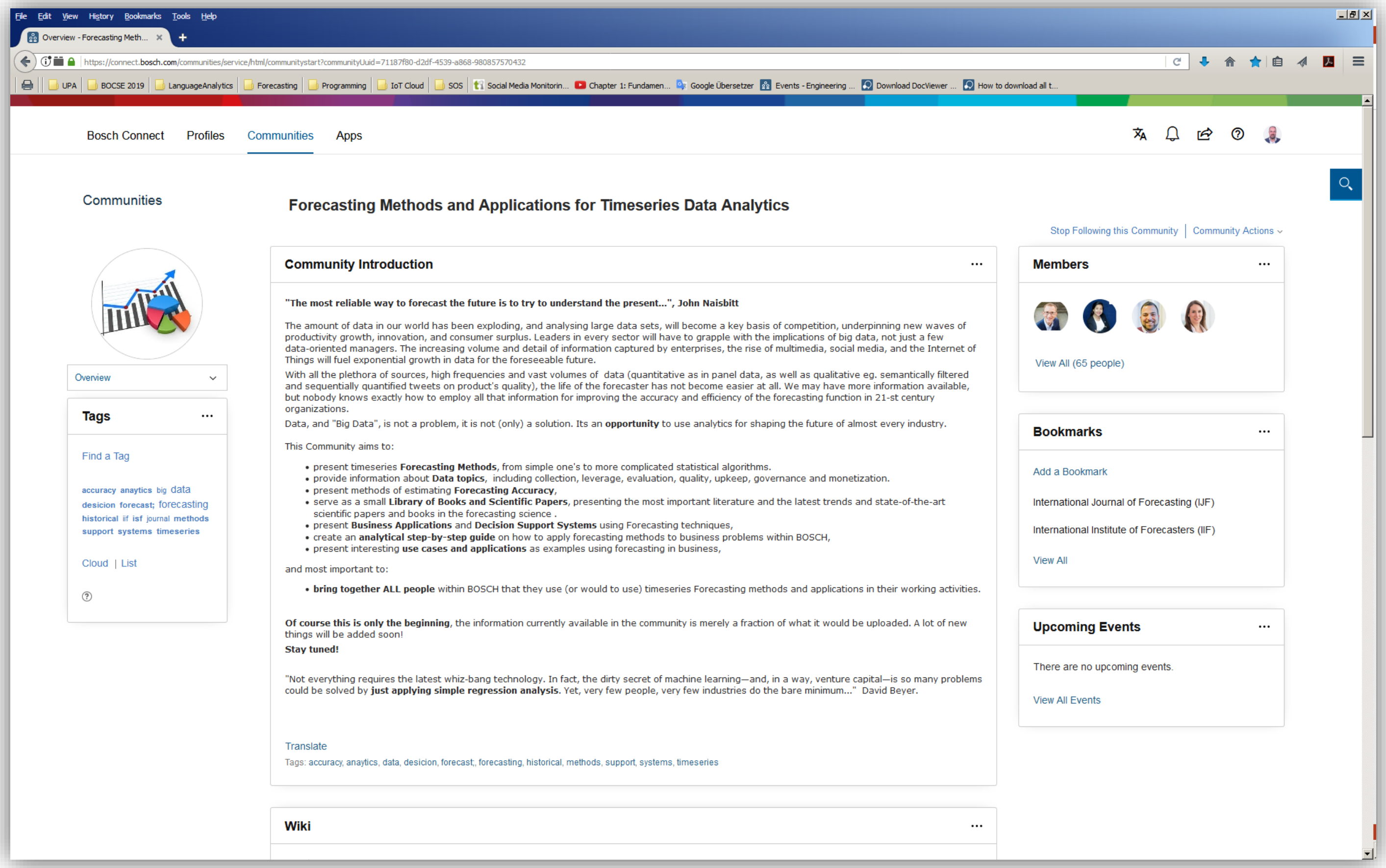
17 Use Case Social Media

18 More… ?

BOSCH

PROPOSED LINKS & LITERATURE

Timeseries Forecasting: Methods & Applications
Bosch Community



http://bos.ch/Cfd

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

- K. Nikolopoulos, D. Thomakos (2019): *“Forecasting with the Theta Method: Theory and Applications”*
- 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

"The mind is not a vessel to be filled, but a fire to be kindled."

Plutarch, Greek Historian