InnoFIT "Leitfaden"

Forecast analysis and forecast improvement

Contact:
FH-Prof. Priv. Doz. DI (FH) Klaus Altendorfer PhD: klaus.altendorfer@fh-steyr.at
FH-Prof. Dr. Thomas Felberbauer, MSc: thomas.felberbauer@fhstp.ac.at

Overview

- InnoFIT project
- Forecast interrelation with production
- Forecast generation
- Customer provided demand data
- Forecast behaviour analysis
- Correlation and Covariance of Forecasts
- Clustering of products according to their forecast behaviour
- Decisions regarding forecast data
- Practical experience
- Webtool tutorials

Introduction - InnoFIT project

Information Technologies for Novel Forecasting Tools

InnoFIT Project - introduction

- Possible consequences of poor forecast quality and information uncertainty are:
  - Inefficient use of resources (machines, employees, materials…)
  - Overproduction, high inventory
  - Poor service level
  - Higher transport costs (e.g., express delivery)

- The forecast quality is rarely measured
- Existing forecast history data and additional sources of information could be used to improve forecast quality and to adapt demands.

InnoFIT Project - organization

- Project partners:
  - 3 research institutions
  - 4 industry partners (companies)
  - Research guided by FH OÖ and FH STP
- Duration: June 2018 – February 2022

Problems of incorrect forecasts

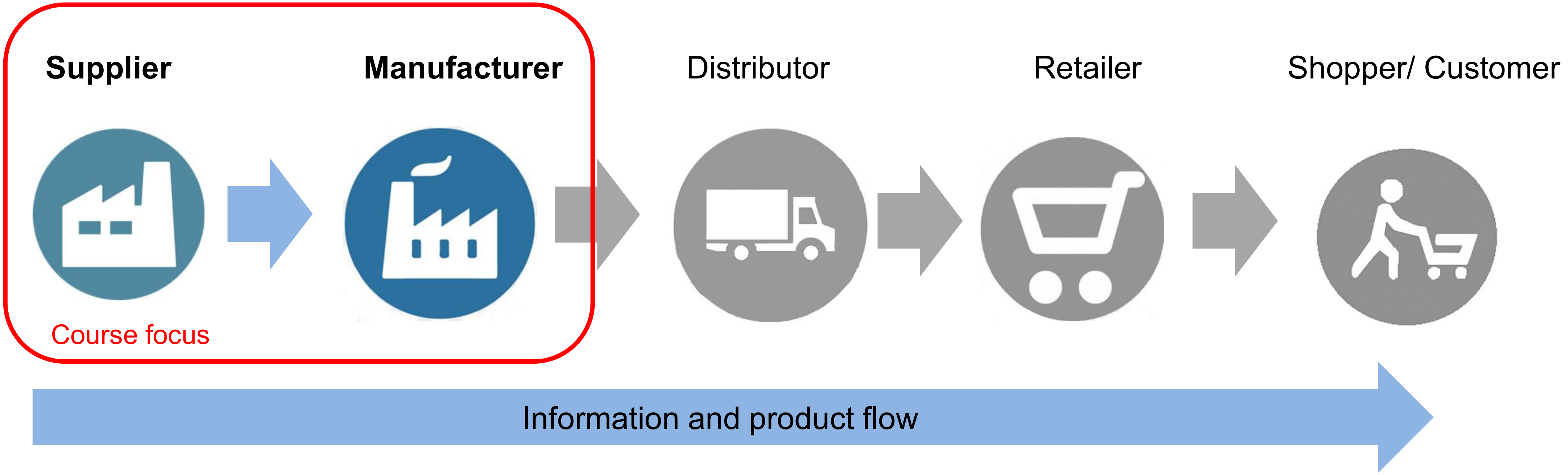
- **Groupwork:**
  - Discuss possible problems of incorrect forecasts
  - Describe the impact of these problems on relevant KPI’s
  - Give suggestions to counteract these problems
  - Create a flipchart to present the results
  - Time: 20 min

InnoFIT Project - background

- Major risks and problems resulting from unavailable or insufficient data on customer demand and customer orders (also poor forecasting):
  - Overproduction (stock, overtime), overbooking/underbooking, additional transport costs, poor delivery reliability, inefficient use of resources
- Historical forecast data is not stored in ERP Systems → not utilized
- Project goal:
  - Analysis of information uncertainties in forecasts through the use of mathematical and statistical tools and models
  - Enhancement of forecast methods and models
  - Simulation, visualization and clustering studies for innovative forecasting in production planning
  - Improvement of existing processes, providing recommendations for improved forecasting methods in companies

InnoFIT Project - background

– Supply chain components



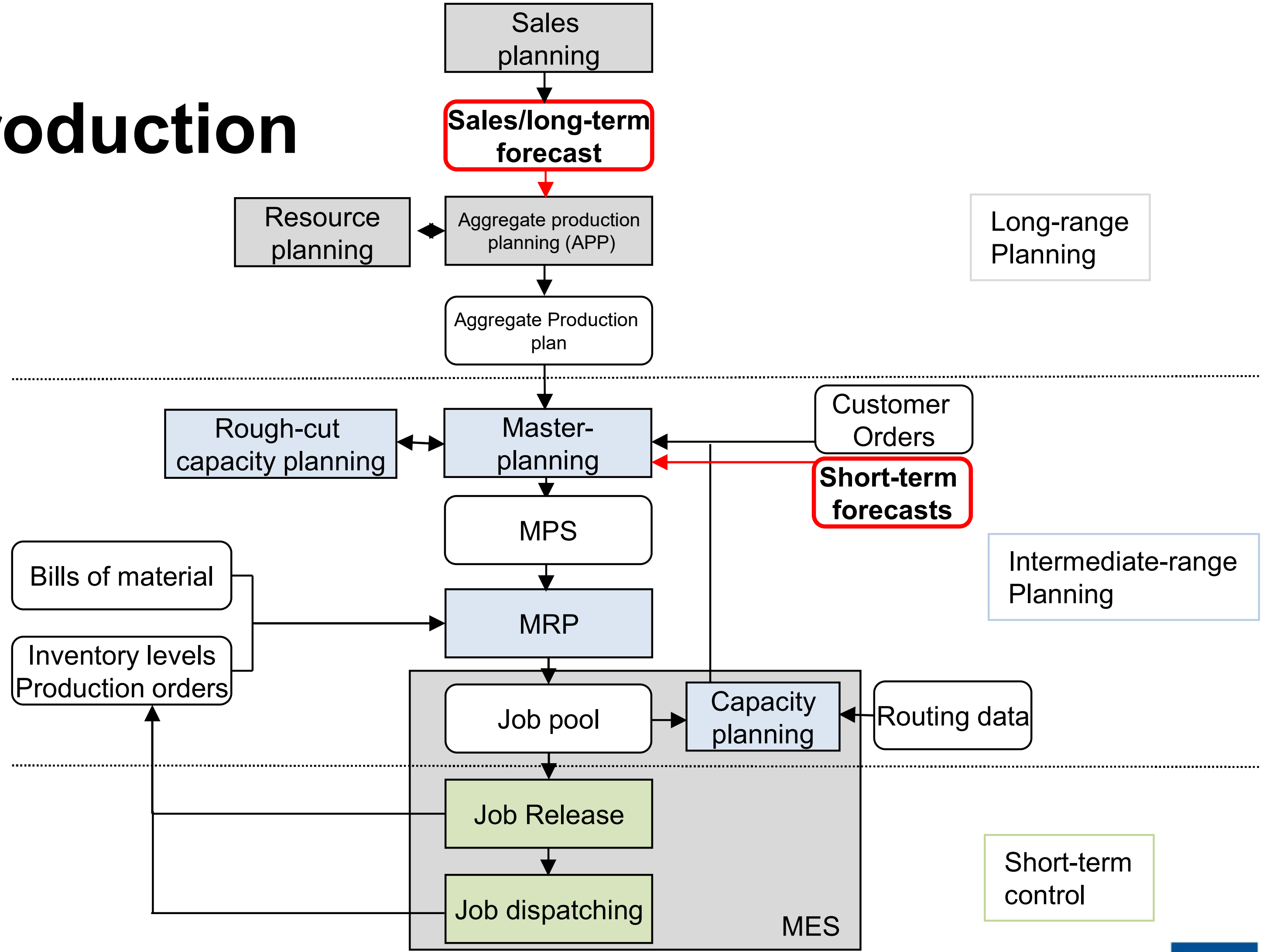
Forecast interrelation with production

Hierarchical planning (1/2)

- Functionality
  - Planning problems were dissolved in hierarchical sub-problems
  - Superordinate results are constraints for subordinate planning problems
  - Formulation of planning problems often as optimization problem
  - Mostly rolling planning with overlapping planning horizons
  - Planning mostly deterministic whereby reality is mostly stochastic

Hierarchical production planning (2/2)

MRP II Concept



Sales/Long-term forecasts

- First input in hierarchical production planning
- Strategic planning (yearly/quarterly)
- Expected sales volumes on a monthly basis
- Input on long term forecasts
  - Historical sales volume
  - customer contracts

- Possible link between long term and short-term forecasts
  - Long-term forecasts are updated by short-term forecasts

Short-term forecasts

- Input in master production scheduling (MPS)
- Often as rolling forecasts
- Lead to production orders in the next step (MRP)
- Long term forecasts can be updated by short-term forecasts

Example – Lot sizing based on forecast

- Wrong forecast leads to wrong production order
  - Lot sizing: Fixed order period = 2
  - Planned lead time = 1

Period	**1**	**2**	**3**
Final Order	0	30	30
Forecast		30	20
Production Order	50		
Inventory (start of period)	0	50	20
Inventory (end of period)	50	20	**-10**

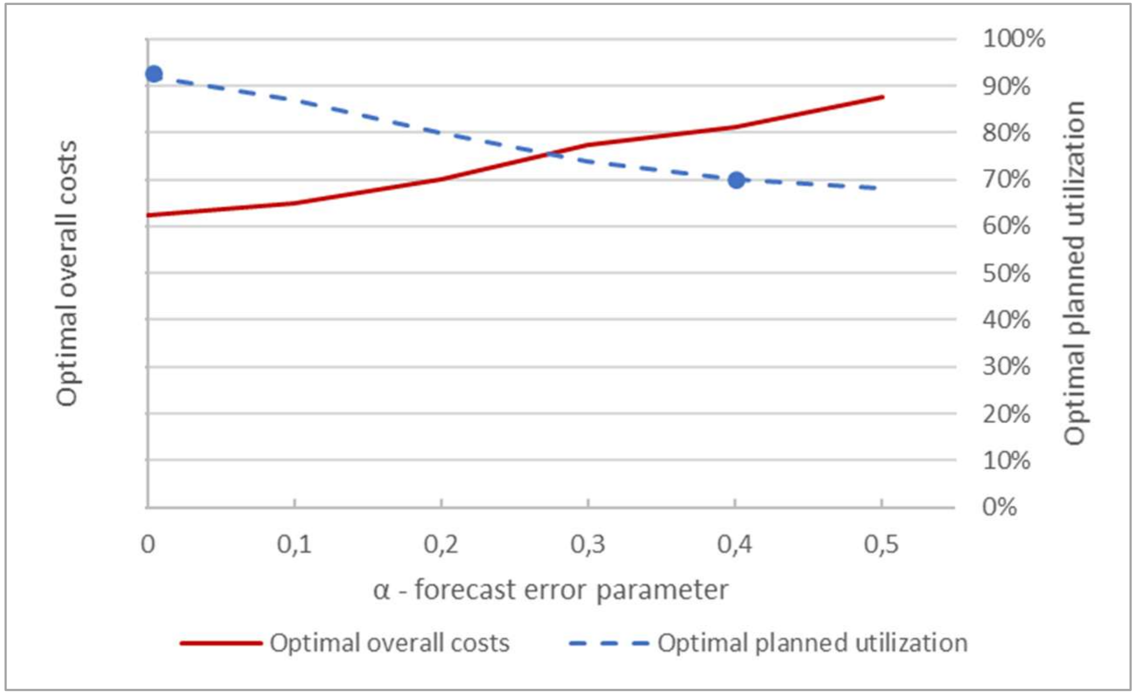
- Production order based on forecasts for period 2 and 3 = 50
- Final order for period 3 has increased by 10 pcs.
- Final Order for period 3 can not be served

Quantitative motivation from literature

- The performance of production systems is strongly influenced by the accuracy of the applied demand forecasts *Fildes and Kingsman (2011), Altendorfer et al. (2016), and Enns (2002)*
- Potential production cost savings of 10-30% through better customer information *Sanders and Graman (2009)*
- Reduction of forecast errors can lead to an increase of 10-30% in product profitability *Metters (1997)*
- Potential overall costs (inventory and tardiness costs) savings of 8-10% by moderate improvement of forecast quality *Altendorfer et al. (2016)*
- Inventory reduction potential of 5-20% through information exchange within the supply chain *Cui et al. (2015)*

Research Result: Effects of forecast errors on optimal utilization

- Higher long forecast errors lead to
  - > Increase in overall costs
  - > Lower optimal planned utilization
- Forecast error parameter
  - > α = 0 → 92% utilization
  - > α = 0,4 → 70% utilization
- Higher optimal utilization leads to higher throughput



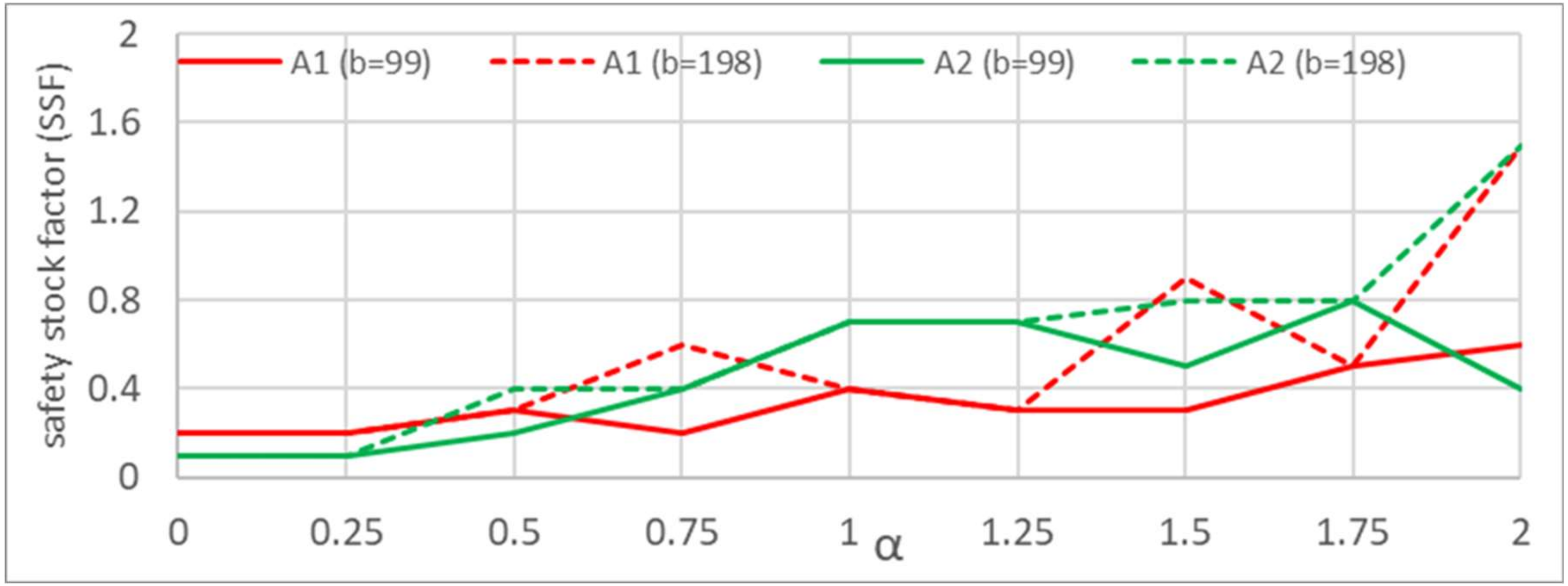
Source: Klaus Altendorfer, Thomas Felberbauer & Herbert Jodlbauer (2016): Effects of forecast errors on optimal utilisation in aggregate production planning with stochastic customer demand, International Journal of Production Research, DOI: 10.1080/00207543.2016.1162918

Negative impact of poor forecast quality

- Upper-level decisions are the constraints for lower-level decisions. Poor long-term forecasts lead to:
  - Inefficient use of resources (capacity, personnel, investments…)
  - Higher inventory
  - Higher backorder costs
  - Poor service level

Research Result: Effect of different forecast uncertainty levels on safety stock

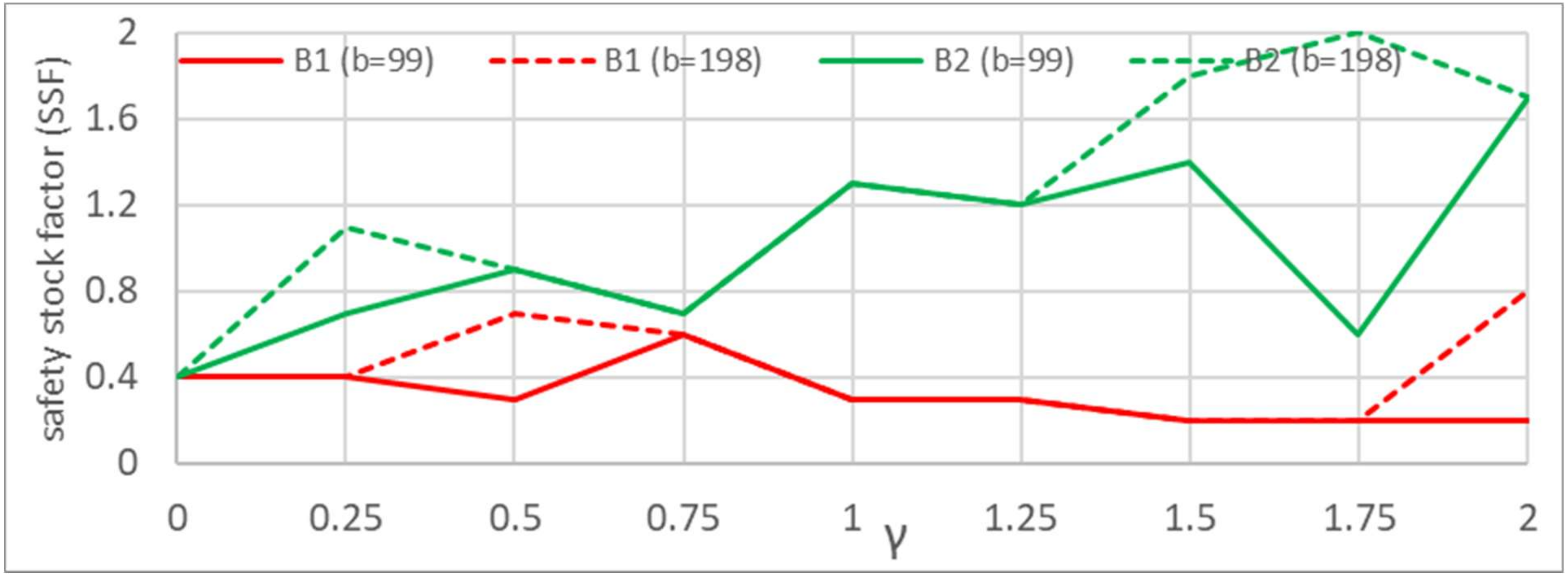
Unsystematic behavior: A1= Basic scenario; A2= Shorter forecast horizon; α= level of uncertainty



Source: Wolfgang Seiringer, Fabian Brockmann, Klaus Altendorfer & Thomas Felberbauer (2021): Influence of forecast error and forecast bias on safety stock on a MRP system with rolling horizon forecast updates, International Conference on Operations Research

Research Result: Effect of different forecast uncertainty levels on safety stock

Systematic behavior: B1= Overbooking; B2= Underbooking; γ= level of bias



Source: Wolfgang Seiringer, Fabian Brockmann, Klaus Altendorfer & Thomas Felberbauer (2021): Influence of forecast error and forecast bias on safety stock on a MRP system with rolling horizon forecast updates. International Conference on Operations Research

Research Result: Effect of different forecast uncertainty levels on safety stock

- Unbiased forecasts
  - Higher uncertainty leads to higher optimal safety stocks and higher overall costs
- Biased forecasts
  - Overbooking implies a certain buffer → low safety stocks
  - Underbooking → higher safety stocks

Difference between demand model and forecast model

- Forecast model*
  - Describes the way a forecast is created
  - Generates a single value as output

- Demand model
  - Describes how the demand is generated
  - Based on previous demands and forecasts
  - Creates a demand distribution which fits the demand profile

**(not the focus of the course)*

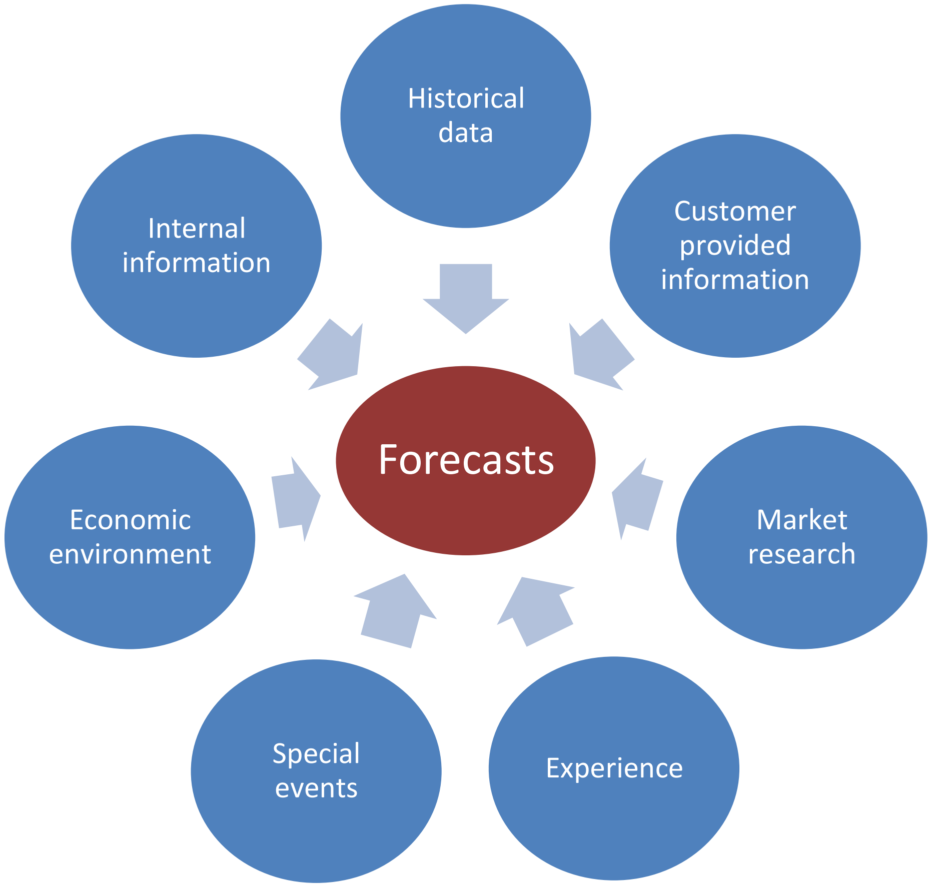
Forecast generation

Law of forecasting

- **First law of forecasting:** Forecast are always wrong!
- **Second law of forecasting:** Detailed forecasts are worse than aggregate forecasts!
- **Third law of forecasting:** The further into the future, the less reliable the forecast will be!

Forecasting

- Impact factors on forecasts for sales planning



How companies get forecast information?

- Mathematical models based on demand realizations from the past
  - Methods: Time series forecasts, Causal forecast methods, etc.
- Customer provided forecast with rolling horizon information updates
  - Methods: Forecast evolution
- Demand forecasts can be crucial for the efficiency of the production system (e.g., automotive industry)

Quantitative forecasting methods

- Based on the assumption that the future can be predicted by using numerical measures of the past in some kind of mathematical model
- Two basic classes of quantitative forecasting models
  - **Causal models** predict a future parameter as a function of other parameters (e.g., future demand for a product is predicted through growth in GDP)
  - **Time series models** predict a future parameter as a function of past values of that parameter (e.g., future demand for a product is predicted by historical demand data)

Sales Planning

- Quantitative methods
  - Moving average
  - Exponential smoothing
  - Approximation and extrapolation
  - Causal method
  - ARIMA model

Sales Planning

- Qualitative methods*
  > Sales estimate (e.g. by region)

  > Customer survey (data aggregation)

  > Expert appraisal (e.g. distributor)

  > Market research

  > Delphi method

**(not the focus of the course)*

Quantitative sales planning
Moving average

- Simple method (e.g., estimate annual amount)
- Problems with trends
- High variability in demands → good results in simulation comparisons

$$p_t = \frac{1}{n}\sum_{i=1}^{n} a_{t-i}$$

$p_t \ldots predicted\ value\ moving\ average\ at\ time\ t$

$a_t \ldots final\ order\ amount\ at\ time\ t$

$n \ldots number\ of\ periods\ for\ moving\ average$

- Example: 1a) Moving average

Moving average – Example 1a)

- Solution
- Number of periods: 4

18			
19	Period	Sales	Moving average
20	1	175	----
21	2	183	----
22	3	200	----
23	4	211	----
24	5	213	=AVERAGE(B20:B23)
25	6	207	201,75
26	7	208	207,75
27	8	178	209,75
28	9	192	201,5
29	10	183	196,25
30	11	178	190,25
31	12	192	182,75

19	Period	Sales	Moving average	1st ord
47	28	206	194,5	
48	29	216	197,5	
49	30	207	207,75	
50	31	207	213,75	
51	32	184	209	
52	33	189	203,5	
53	34	178	196,75	
54	35	174	189,5	
55	36	204	181,25	
56	37	----	186,25	
57	38	----	=AVERAGE($B$52:$B$55)	
58	39	----	AVERAGE(number1; [numbe	
59	40	----	186,25	
60	41	----	186,25	
61	42	----	186,25	
62	43	----	186,25	
63	44	----	186,25	

Quantitative sales planning
1st order exponential smoothing (1/2)

- Simple method (e.g., estimate annual amount)
- Problems with trends
- Greater weight on newer values

$$p_t = \alpha a_{t-1} + (1-\alpha)p_{t-1} = \sum_{i=1}^{t} \alpha(1-\alpha)^{i-1} a_{t-i}$$

$p_t \ldots predicted\ value\ 1st\ order\ exponential\ smoothing\ at\ time\ t$

$a_t \ldots final\ order\ amount\ at\ time\ t$

$\alpha \ldots smoothing\ parameter\ (\in [0,1])$

Quantitative sales planning
1st order exponential smoothing (2/2)

- The higher α, the more sensitive the forecast curve reacts → „newer“ values have higher impact
- The lower α, the more stable is the forecast curve → short term „outliers“ change the forecast little
- Optimal α-values between 0,1 and 0,3 (experience value)
- Example: 1b) 1st order exponential smoothing

1st order exponential smoothing – Example 1b)

- Solution
- E11 = alpha = 0,2

19	Period	Sales	Moving average	1st order	trend anc
20	1	175	----	175	
21	2	183	----	=$E$11*B20+((1-$E$11)*D20)	
22	3	200	----	177	
23	4	211	----	181	
24	5	213	192,25	187	
25	6	207	201,75	192	
26	7	208	207,75	195	
27	8	178	209,75	198	
28	9	192	201,5	194	

19	Period	Sales	Moving average	1st order	trend and seas
53	34	178	196,75	197	----
54	35	174	189,5	193	----
55	36	204	181,25	190	----
56	37	----	186,25	=$E$11*$B$55+((1-$E$11)*$D$55)	
57	38	----	186,25	192	
58	39	----	186,25	192	
59	40	----	186,25	192	
60	41	----	186,25	192	
61	42	----	186,25	192	
62	43	----	186,25	192	
63	44	----	186,25	192	

Quantitative sales planning
Exponential smoothing with trend and season (1/4)

- Multi-period forecast (e.g., on a monthly basis)
- Consideration of trend
- Consideration of any seasonality
- Results are difficult to understand
- Good for recurring effects, which are difficult to map functionally
- Constant part, trend and season were given a smoothing factor separately

Quantitative sales planning
Exponential smoothing with trend and season (2/4)

- Calculation

$$x_t = \alpha \frac{a_t}{i_{t-L}} + (1-\alpha)(x_{t-1} + b_{t-1})$$

$$b_t = \beta(x_t - x_{t-1}) + (1-\beta)b_{t-1}$$

$$i_t = \gamma \frac{a_t}{x_t} + (1-\gamma)i_{t-L}$$

$$p_{t+k} = (x_t - kb_t)i_{t-L+k}$$

Quantitative sales planning
Exponential smoothing with trend and season (3/4)

- Notation

$p_t \ldots predicted\ value\ exponential\ smoothing\ at\ time\ t$

$a_t \ldots final\ order\ amount\ at\ time\ t$

$x_t \ldots long\ term\ smoothing\ or\ average$

$b_t \ldots smoothing\ slope\ or\ trend$

$i_t \ldots smoothing\ of\ seasonal\ factor$

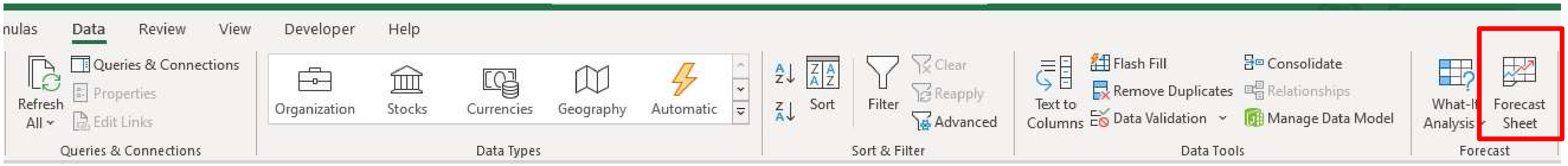
$\alpha, \beta, \gamma \ldots smoothing\ parameter (\in [0,1])$

$L \ldots period\ length$

Quantitative sales planning
Exponential smoothing with trend and season (4/4)

- Supported by MS Excel (detailed implementation cannot be checked)
- In Excel
  - Data → Forecast → Forecast Sheet

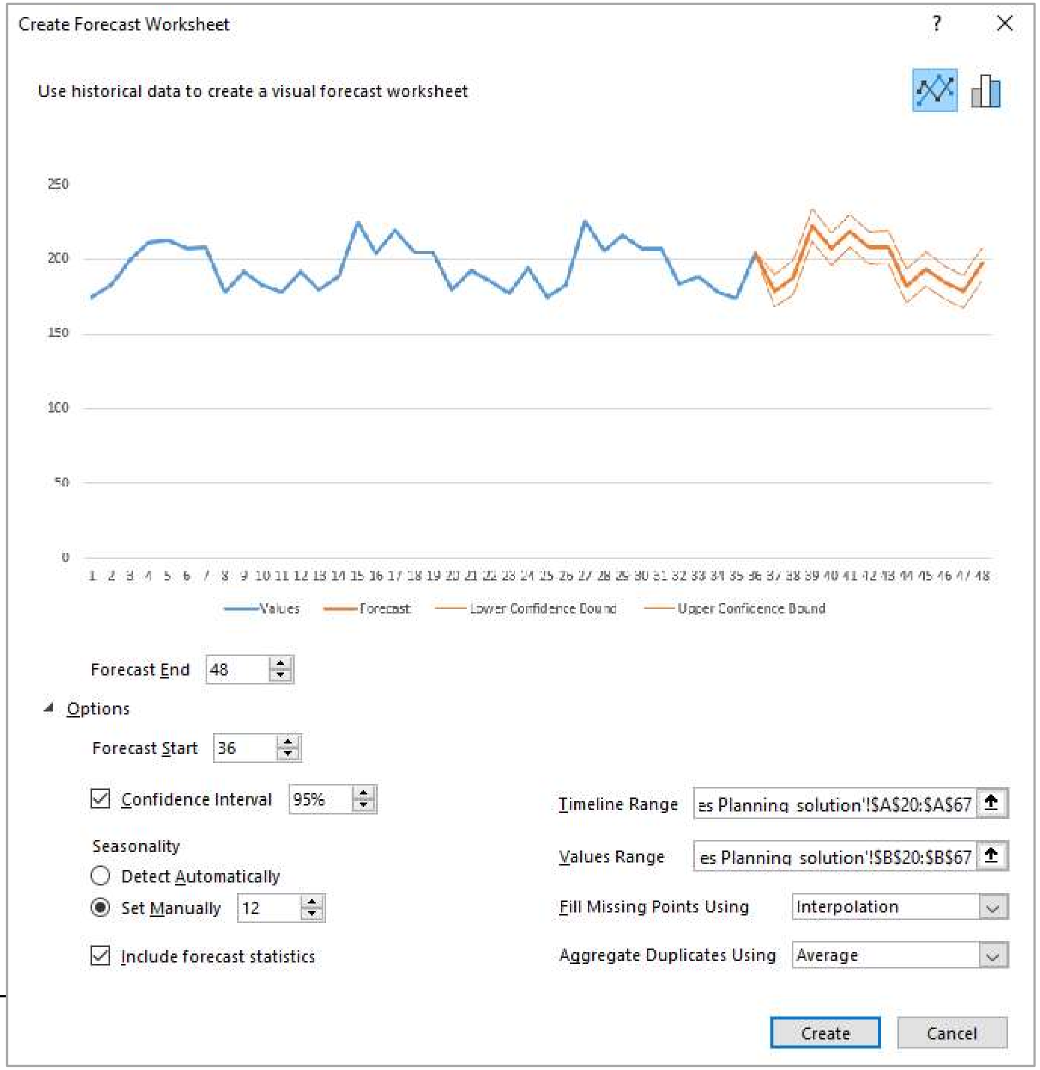
- Example 1c) Exponential smoothing with trend and season



Exponential smoothing with trend and season – Example 1c)

– Solution

17			Excel (Forecast Sheet)
18			Exp. smoothing
19	Period	Sales	trend and season
53	34	178	----
54	35	174	----
55	36	204	----
56	37	----	179,2123
57	38	----	187,7875
58	39	----	222,9904
59	40	----	206,8985
60	41	----	218,9557
61	42	----	207,8654
62	43	----	208,2947
63	44	----	182,1229
64	45	----	193,4943
65	46	----	184,4083
66	47	----	178,5242
67	48	----	197,2592



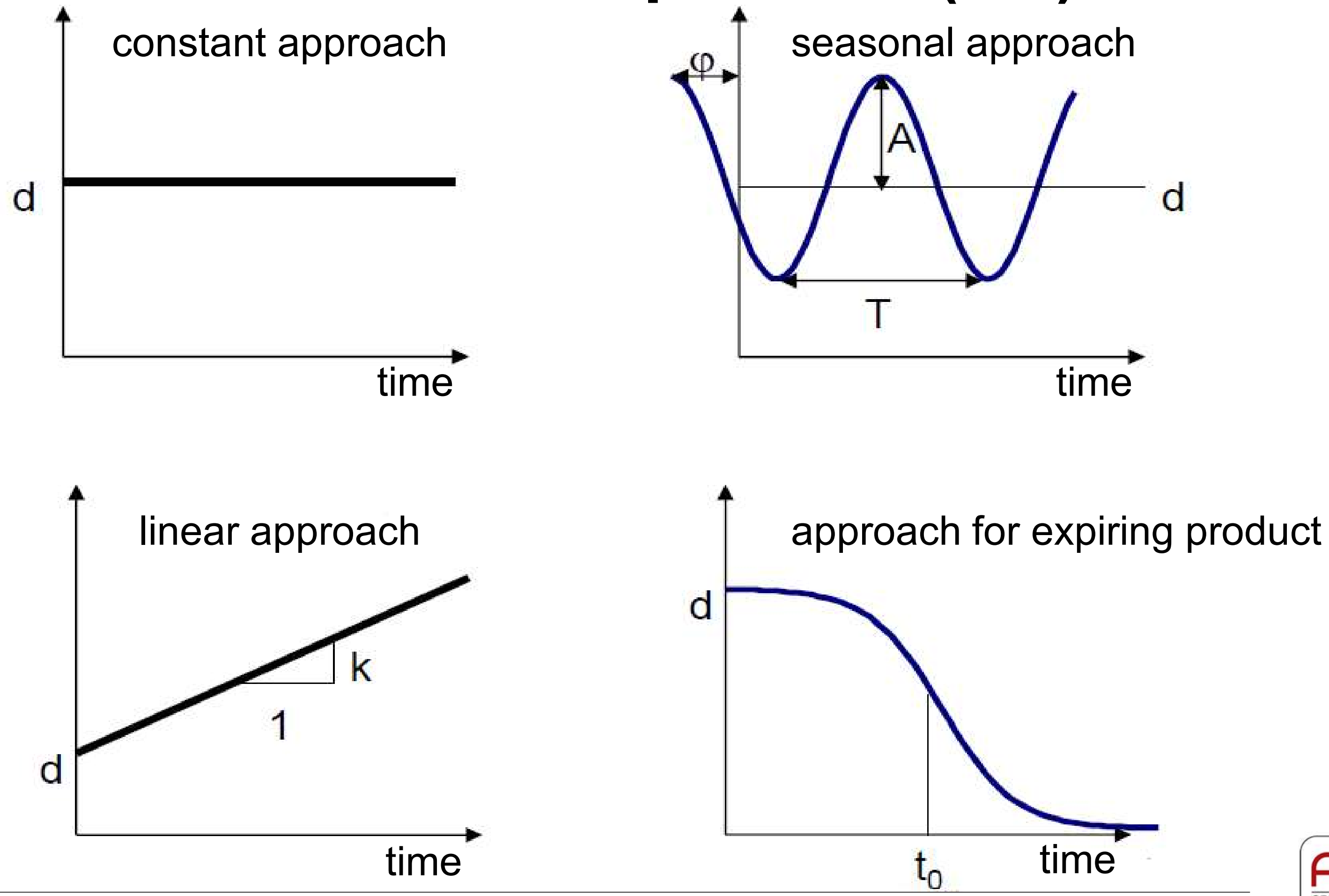
Quantitative sales planning
Approximation and extrapolation (1/6)

- 2-step process
- 1st step approximation:
  - Selection of suitable forecasting functions
  - Adjustment of the function parameters with least squares approach

- 2nd step extrapolation:
  - Continuing the approximated function in the future

- Attention: keep a close eye on trends

Quantitative sales planning
Approximation and extrapolation (2/6)



Quantitative sales planning
Approximation and extrapolation (3/6)

- Forecasting functions

$$p(t) = d$$ (constant approach)

$$p(t) = d + kt$$ (linear approach, approach with trend)

$$p(t) = d + A\cos\left(\frac{2\pi(t+\varphi)}{T}\right)$$ (seasonal approach)

$$p(t) = \frac{d}{1+\exp(v(t-t_0))}$$ (expiring product)

Quantitative sales planning
Approximation and extrapolation (4/6)

$d \ldots constant\ sales\ share$

$k \ldots trend\ of\ sales\ (slope\ of\ the\ best\ fit\ straight\ line)$

$A \ldots amplitude\ of\ seasonal\ variation$

$T \ldots duration\ of\ seasonal\ variation$

$\varphi \ldots phase\ displacement$

$t_0 \ldots time\ at\ which\ sales\ \frac{d}{2}\ \ is\ expected$

$v \ldots speed\ of\ expiration$

Quantitative sales planning
Approximation and extrapolation (5/6)

- Adjustment of the function parameters with least squares approach

$$\sum_{i=1}^{n} \left(p(t_i) - a_{t_i}\right)^2 \rightarrow Min$$

$t_i \ldots past\ points\ in\ time\ for\ which\ sales\ are\ known$

$a_{t_i} \ldots past\ sales\ at\ time\ t_i$

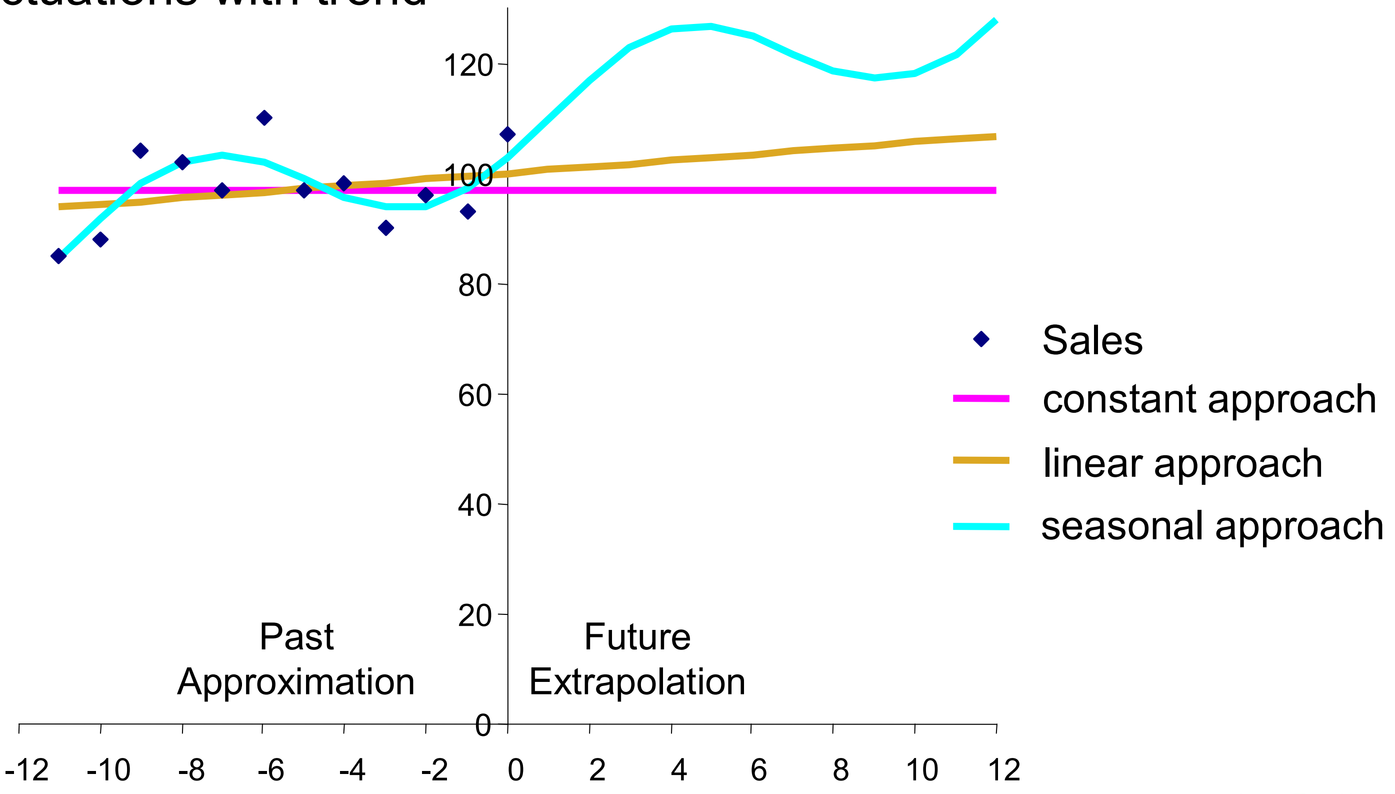
$p(t_i) \ldots selected\ forecast\ function, evaluated\ at\ time\ t_i$

$n \ldots number\ of\ past\ points\ in\ time$

- Example 1d) Approximation and extrapolation

Quantitative sales planning
Approximation and extrapolation (6/6)

– Seasonal fluctuations with trend



Approximation and extrapolation – Example 1d)

– Solution

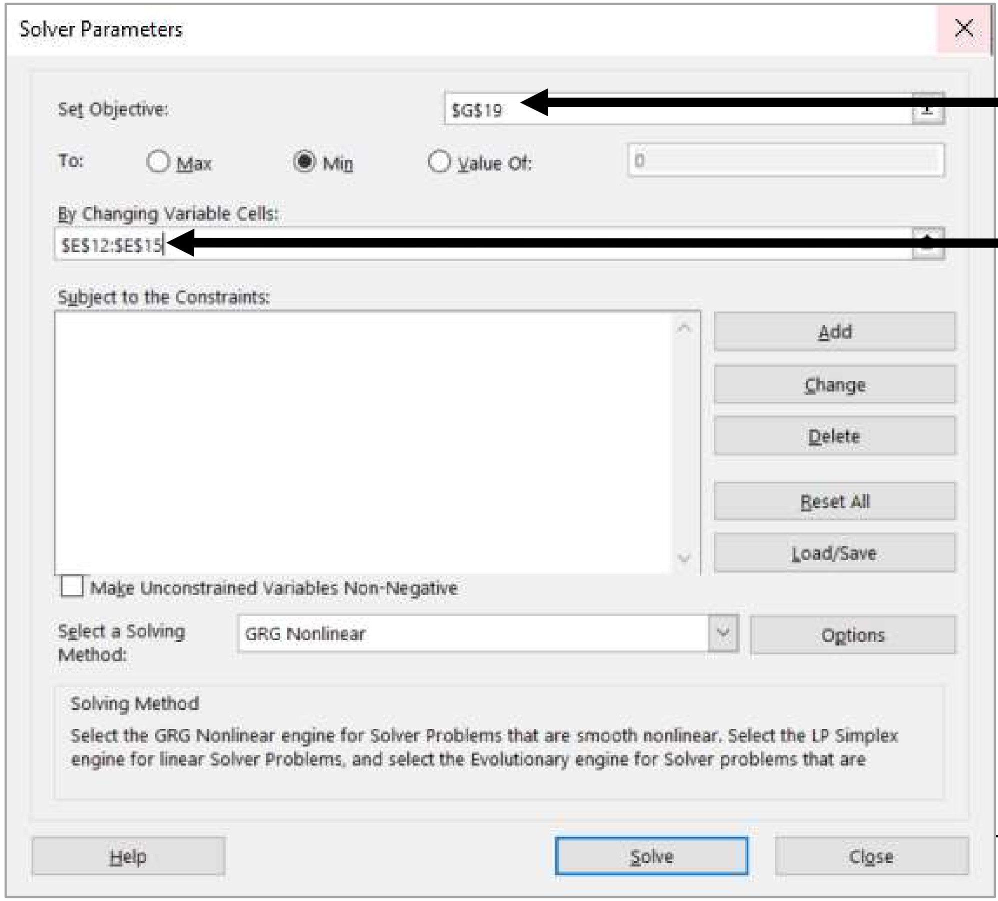
10	Moving average			number of periods	4		
11	Exponential smoothing			alpha	0,2		
12	Approximation and extrapolation			d	191,1252938		
13				k	0,216940599		
14				A	-16,96085545		
15				phi	1,222019332		
16							
17					Excel (Forecast Sheet)	Approx. and extrapolation	
18				Exp. smoothing	Exp. smoothing	P(t)	difference
19	Period	Sales	Moving average	1st order	trend and season	----	3236,863009
20	1	175	----	175	----	=$E$12+$E$14*COS(2*PI()*(A20+$E$15)/12)+$E$13*A20	
21	2	183	----	175	----	193,5264	110,8055282
22	3	200	----	177	----	201,903	3,621374021
23	4	211	----	181	----	207,5661	11,791854

17					Excel (Forecast Sheet)	Approx. and extrapolation	
18				Exp. smoothing	Exp. smoothing	P(t)	difference
19	Period	Sales	Moving average	1st order	trend and season	----	3236,863009
20	1	175	----	175	----	184,6227	=(F20-B20)^2

Approx. and extrapolation	
P(t)	difference
----	=SUM(G20:G55)
184,6227	92,59690029
193,5264	110,8055282
201,903	3,621374021
207,5661	11,791854
209,0564	15,55215391
206,0327	0,935652288
199,3634	74,59113519
190,8936	166,2442266

Approximation and extrapolation – Example 1d)

– Least square approach with Excel Solver to determine the parameters



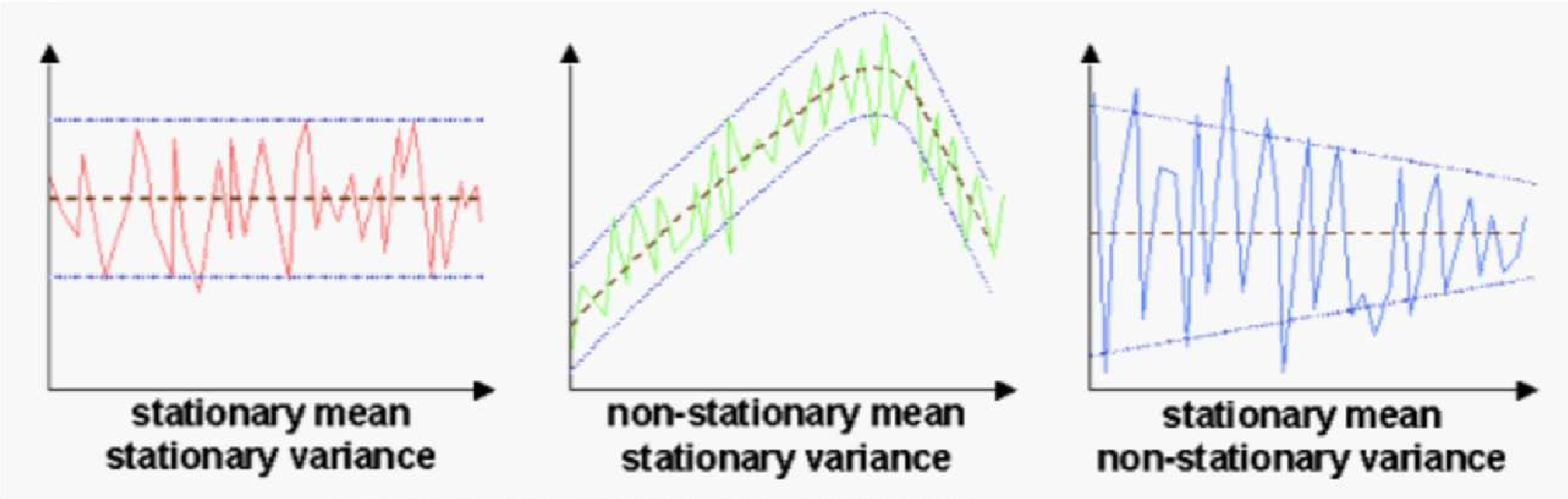
Approx. and extrapolation	
P(t)	difference
----	3236,863009

12	Approximation and extrapolation	d	191,1252938
13		k	0,216940599
14		A	-16,96085545
15		phi	1,222019332

17			Approx. and extrapolation	
18			P(t)	difference
19	Period	Sales	----	3236,863009
55	36	204	185,3294	348,5919036
56	37	----	192,4326	----
57	38	----	201,3363	----
58	39	----	209,7129	----
59	40	----	215,3759	----
60	41	----	216,8662	----
61	42	----	213,8426	----
62	43	----	207,1732	----
63	44	----	198,7034	----
64	45	----	190,7607	----
65	46	----	185,5315	----
66	47	----	184,4751	----
67	48	----	187,9327	----

Difference between stationary and non-stationary time series models

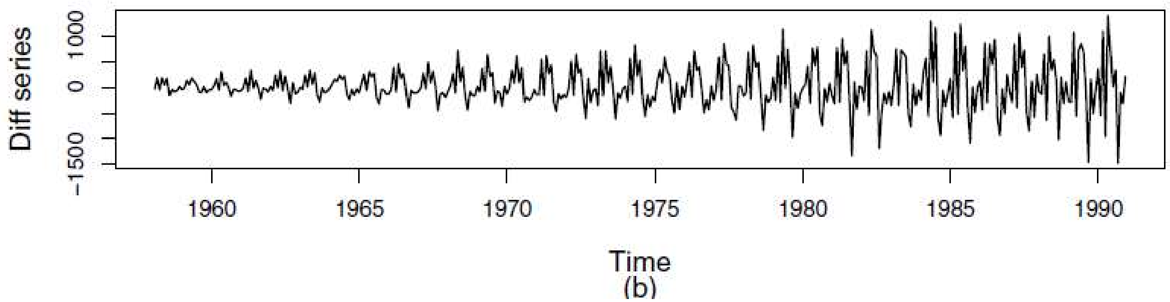
- Two statistical measures
  - Mean (red dotted line)
  - Variance (blue dotted line)



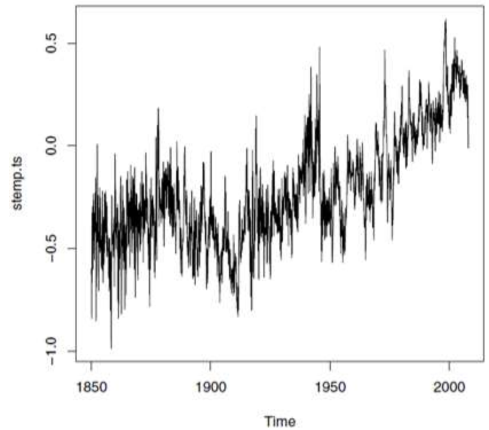
ARIMA (1/3)

- The random walk model with autoregressive and moving average terms. The underlying stochastic process (which aggregates the differenced series) to recover the original series is called **autoregressive integrated moving average** or **ARIMA**
- Seasonal and non-seasonal ARIMA models:
  - Non-seasonal ARIMA models (linear trend with white noise added, electricity production series; left)
  - Seasonal ARIMA models (temperature/climate scenarios; right)

Non-seasonal



Seasonal



Sources: Frances, Dijk & Opschoor, Time Series models for Business and Economic Forecasting, 2014; Cowpertwait & Metcalfe, Introductory Time Series with R, 2009

ARIMA (2/3)

- **Models with deterministic seasonality** imply a seasonal pattern that is stable over time, **models with seasonal unit roots** suggest the presence of changing and seasonal patterns.
- **Seasonal linear ARIMA (SARIMA)** models makes the correlation of y_t with explicit seasonal lags by augmenting ARIMA model with seasonal differencing filter and seasonal AR and MA components.
- The aspect of ARIMA demand model implies that the next demand occurrence depends on the last demand occurrence

Source: Frances, Dijk & Opschoor, Time Series models for Business and Economic Forecasting, 2014

ARIMA (3/3)

- Type of seasonality varies across time series
- SARIMA (seasonal random walk) model for forecasting applies when the seasonality in a time series yt is deterministic
- Deterministic seasonality can be accounted for by means of seasonal dummy variables, which presupposes a seasonal pattern that is stable over time

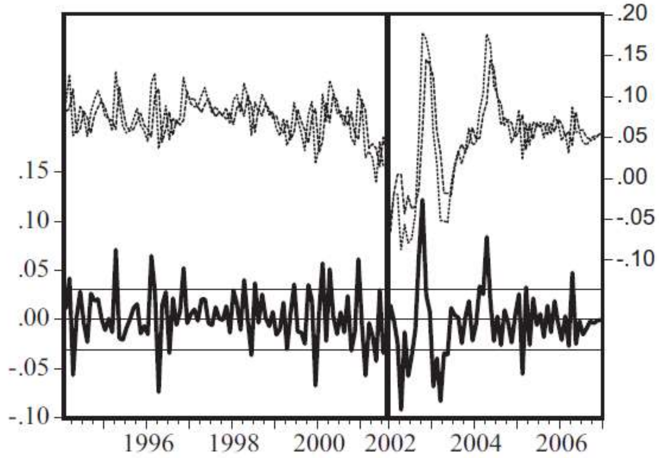


Figure 3.8: Typical fit of an AR time series model: estimation results of an AR(1) on $\Delta_{12} y_t$ with y_t the log monthly revenue-passenger revenue-passenger kilometres of European airlines, 1994.1–2001.8 and 2002.1–2006.12. The short-dashed and long-dashed lines correspond with the actual time series and fitted values, respectively. The solid line represents the residuals. The vertical line indicates the gap in the estimation sample due to the omission of the observations for 2001.9–2001.12.

Source: Frances, Dijk & Opschoor, Time Series models for Business and Economic Forecasting, 2014

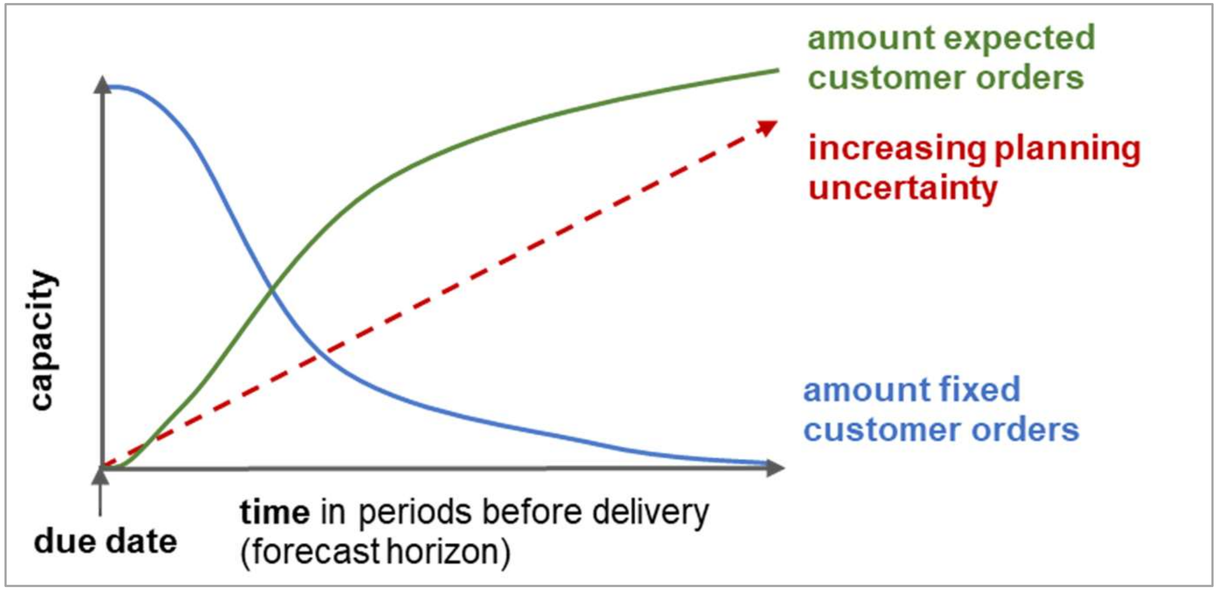
Customer provided demand data

Customer provided demand data

- Customers provide suppliers with demand data according to their future orders and forward information on:
  - Product type
  - Delivery date
  - Delivery quantity
  - Incoterms
  - …
- This information is input for sales and production planning
- Historical customer provided demand data can be used to improve forecasts

Customer order behaviour

- Planning uncertainty increases the further the forecasts are made in the future
  - Expected customer orders were transformed to fixed customer orders
  - Fixed customer orders are increasing close to the due date

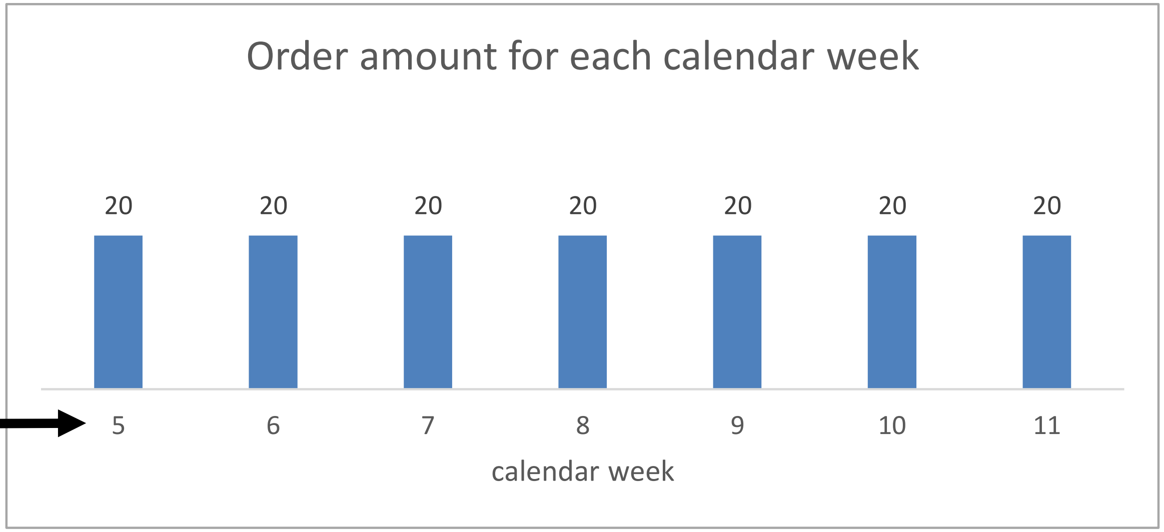


Change order amount and due date

- **Groupwork:**
  - Discuss possible problems of changes in orders and changes of delivery dates
  - Create a flipchart to present the results
  - Time: 20 min

Change in order quantity (1/3)

- Calendar week 5-11: order of 20 pcs. for each week



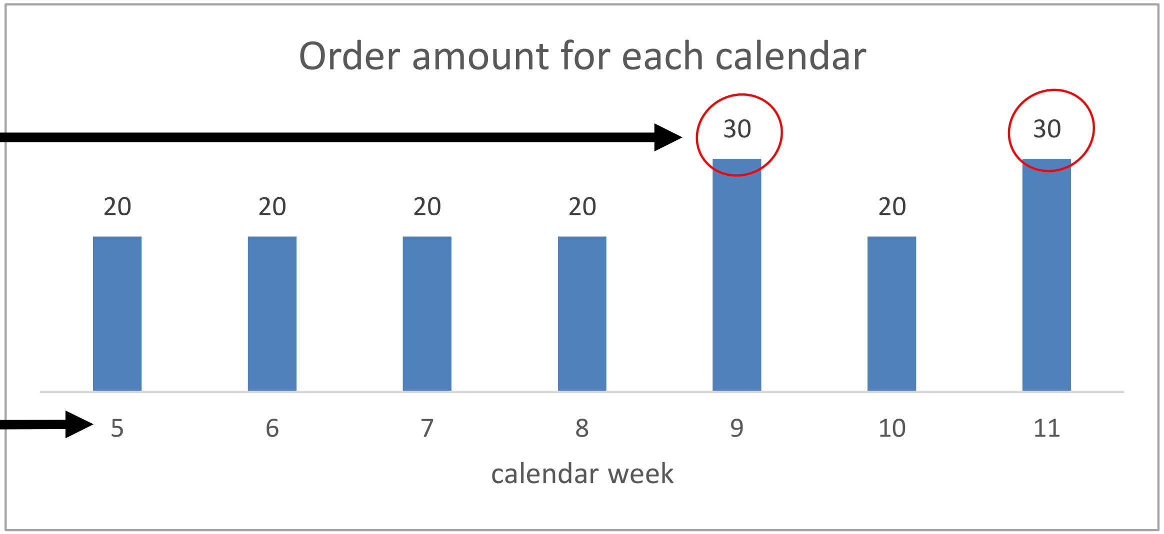
delivery weeks

Change in order quantity (2/3)

- Calendar week 5-11: order of 20 pcs. for each week
- Calendar week 9: increase in order quantity by 10 pcs.
- Calendar week 11: increase in order quantity by 10 pcs.

Increase in order quantity in week 9 and 11

delivery weeks



Change in order quantity (3/3)

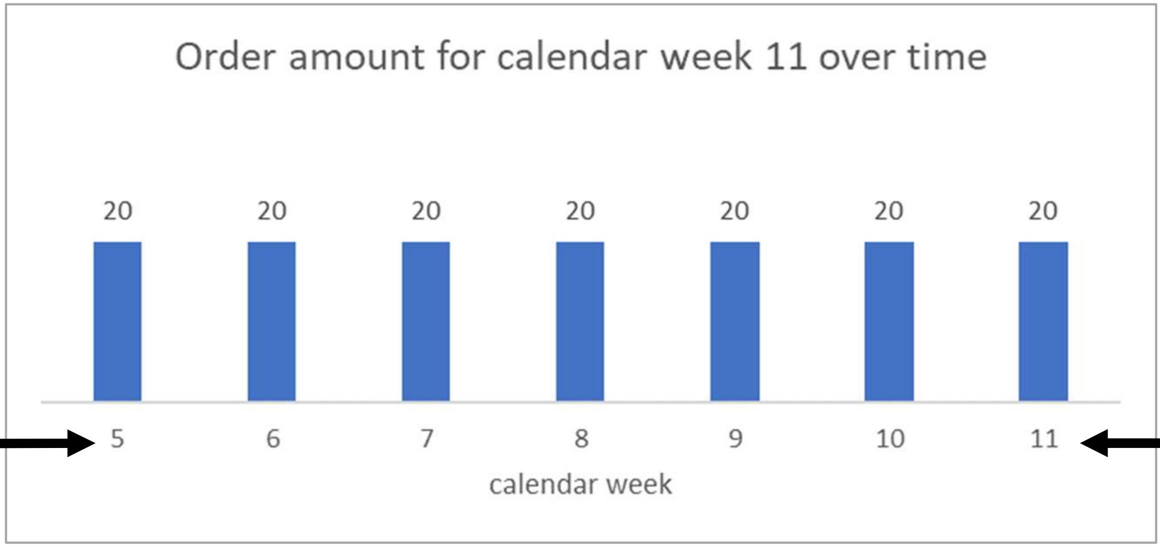
- Discussion → Impact of increase in order quantity
  - Additional capacity for 10 pcs. is required
  - Production planning must be adapted
  - Can lead to delay in delivery
  - Increase in revenue
  - …

- Discussion → Impact of decrease in order quantity

Increase in order quantity (1/4)

- Calendar week 5: order of 20 pcs. for calendar week 11

order week

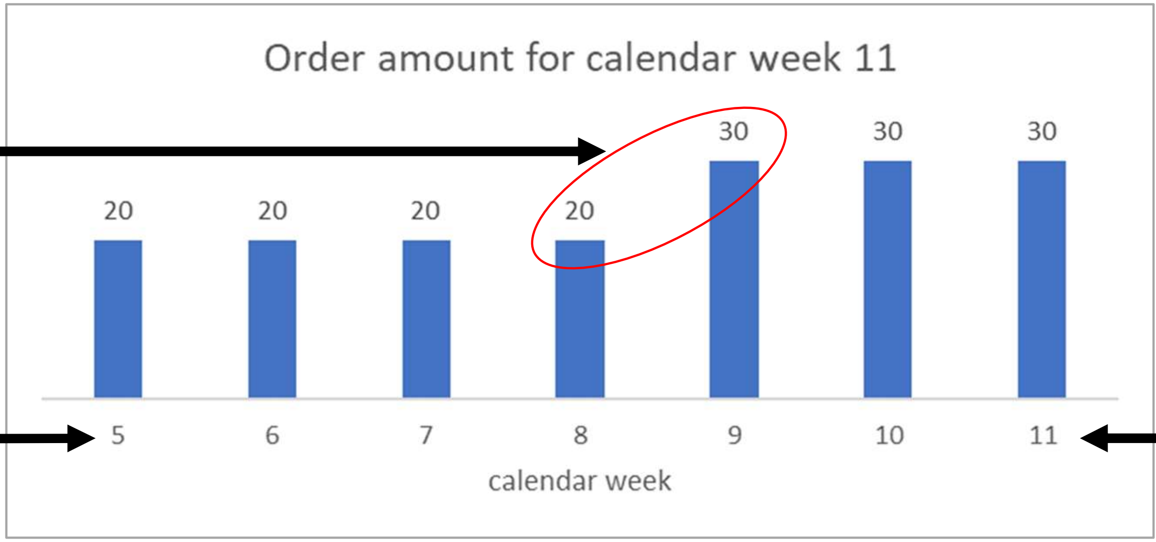


delivery week

Increase in order quantity (2/4)

- Calendar week 5: order of 20 pcs. for calendar week 11
- Calendar week 9: increase in order quantity by 10 pcs.

Increase in order quantity in week 9

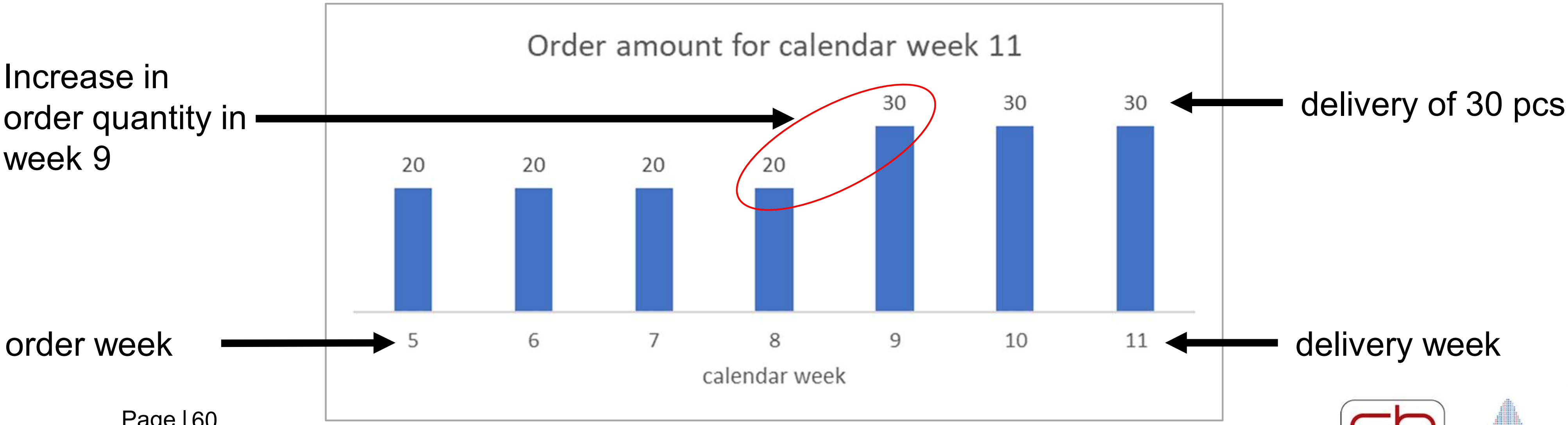


order week

delivery week

Increase in order quantity (3/4)

- Calendar week 5: order of 20 pcs. for calendar week 11
- Calendar week 9: increase in order quantity by 10 pcs.
- Calendar week 11: delivery of 30 pcs.



Increase in order quantity (4/4)

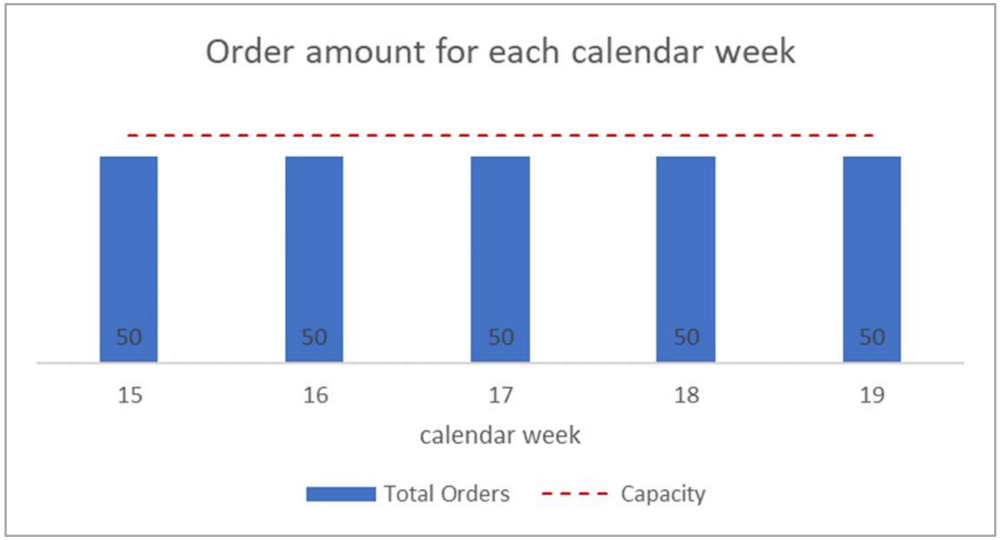
- Discussion → Impact of increase in order quantity
  - Additional capacity for 10 pcs. is required
  - Production planning must be adapted
  - Can lead to delay in delivery
  - Increase in revenue
  - …

- Discussion → Impact of decrease in order quantity

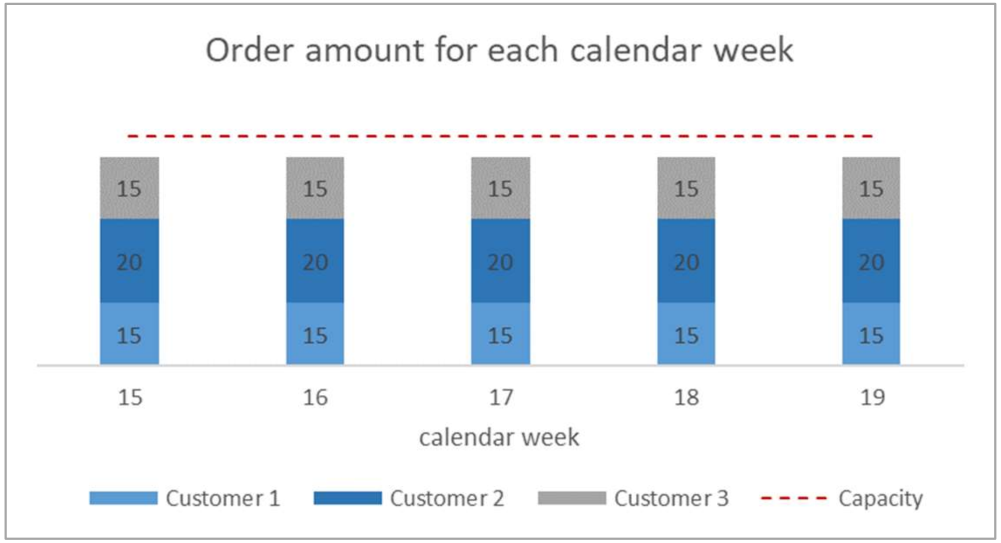
Change of delivery date (1/3)

- Capacity for each calendar week = 55 pcs.
- Customer orders of 50 pcs. for each week

No single customer orders available



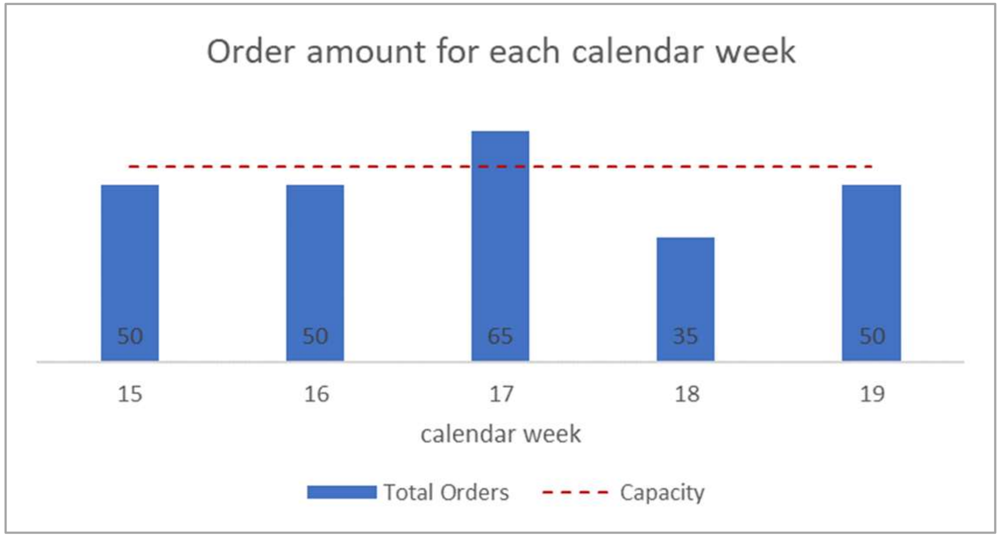
Single customer orders are available



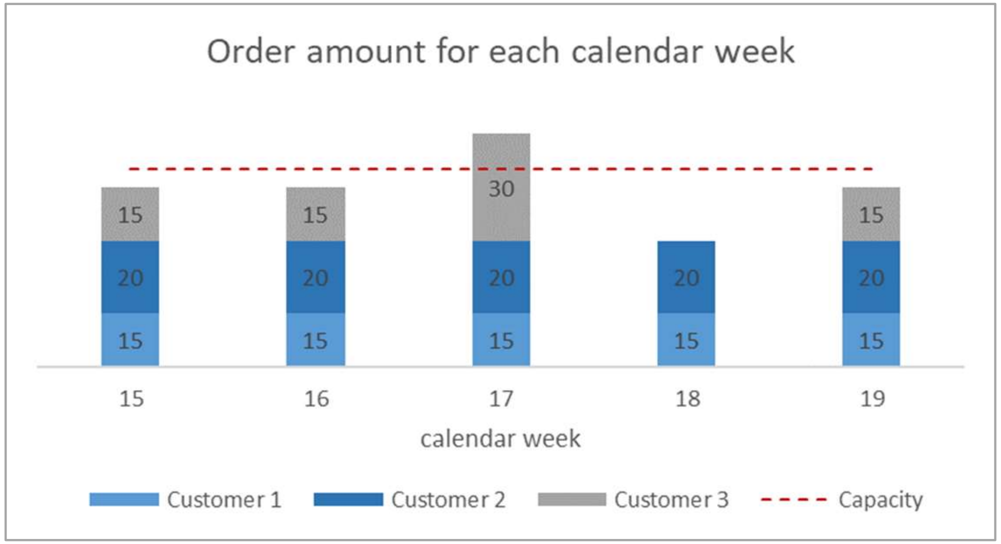
Change of delivery date (2/3)

- Capacity for each calendar week = 55 pcs.
- Customer 3 changes the delivery date for 15 pcs. from calendar week 18 to 17

No single customer orders available



Single customer orders are available



Change of delivery date (3/3)

- Discussion → Impact of change of delivery date
  - Amount and due date shifts are simple to track if single orders are known
  - Shift in capacity consumption for 15 pcs.
  - Production planning must be adapted
  - Can cause capacity peaks → inefficient use of resources
  - Can lead to delayed delivery caused by bottlenecks in production
  - …

Rolling forecasts

- Forecasts for future demands change every period (each month an order update)
- Rolling planning often at single material level
- Rolling planning → overwriting previous values
- Example: rolling forecast based on moving average

	period					
	1	2	3	4	5	6
demand	330	270				
forecast			300	300	300	300

Forecast after 2 periods
(330+270) / 2 = 300

	period					
	1	2	3	4	5	6
demand	330	270	150			
forecast				210	210	210

Forecast after 3 periods
(330+270+150) / 3 = 210

Forecasts with rolling horizon information updates

- Customers provide a long-term forecast (delivery plans)
- Forecast updates are provided each period until delivery
- Changes in the forecasts can be analyzed
  - Systematic effects (e.g., systematic overbooking)
  - Uncertainty/variance of the forecast
  - Effects of the forecast on planning
- History of the delivery plans is necessary (historical forecast data)

Forecast evolution based on MMFE

- Martingale Model for Forecast Evolution (MMFE) – assumes that forecasts represent the conditional expectations of demands given all available information
- MMFE Model describes the evolution of forecasts over time, the process of forecasting and uncertainty updates (forecasts are updated at the beginning of each period → ε)
- Uncertainty reduces closer to the demand realization due to successive forecast updates
- Forecast evolution can describe the rolling horizon information updates

$$F_{t+1,s} = F_{t,s} + \varepsilon_{t,s}$$

$$s \geq t$$

$F_{1,\,s}$ ... $initial\ forecast\ in\ period\ s\ for\ each\ s \geq 1$

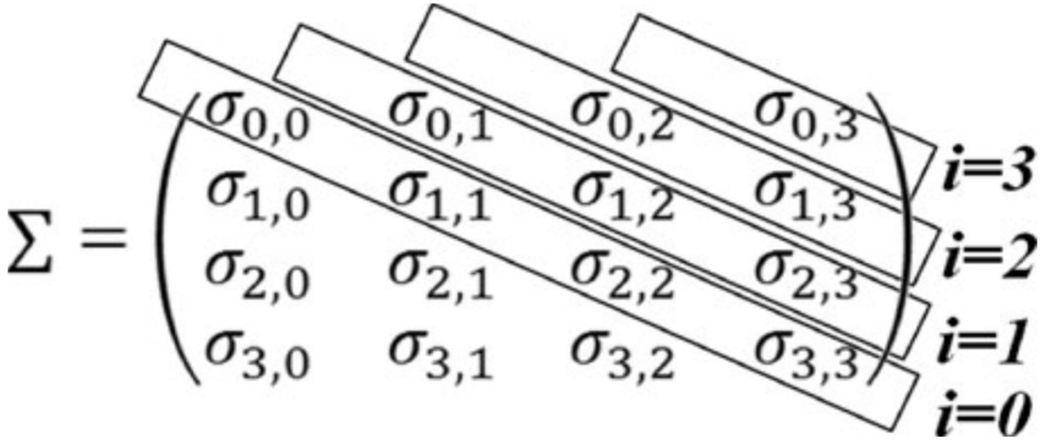
$F_{t,\,s}$ ... $forecast\ at\ the\ beginning\ of\ period\ t$

$\varepsilon_{t,\,s}$ ... $random\ variable\ (forecast\ update)\ in\ period\ t$

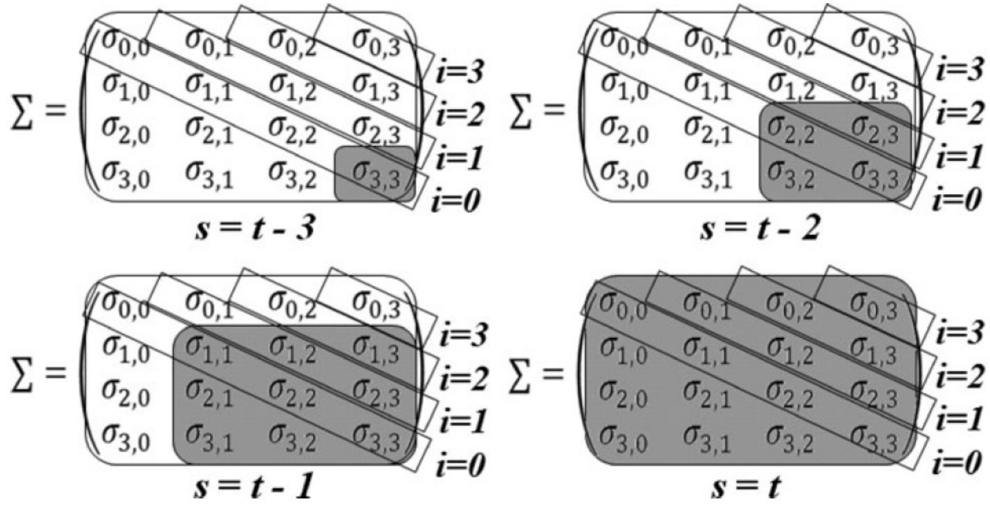
Source: Norouzi & Uzsoy, Modeling the evolution of dependency between demands, with application to inventory planning, 2013

Forecast evolution based on MMFE

Unconditional covariance of demand



Conditional covariance evolution



For periods $s >= t$

s = *current period (when forecast is made)*

t= *forecast period*

Source: Norouzi & Uzsoy, Modeling the evolution of dependency between demands, with application to inventory planning, 2013

Forecast behaviour analysis

Forecast behaviour

- Forecast behaviour analysis only for forecasts with rolling horizon forecast updates
- Describes the behaviour of the forecast values compared to the final order within the historical forecast streams of one product
- Forecast behaviour should be described over time (e.g. periods before delivery)
- Forecast behaviours have to be expressed by numbers (e.g. forecast error measures)
- The Forecast behaviour can be related to different entities (customers, products, product groups,…)

Modification matrix (rolling forecasts)

- Company provides its own demand on a recurring basis
- Example: forecast in week 1 for every period until week 10 in a rolling horizon planning (each week an order update for the following weeks for one material)
  - > Final order amount for calendar week 1 and the forecast for each week until week 10

		delivery plans
		1
forecast delivery plans	1	**620**
	2	640
	3	814
	4	525
	5	761
	6	583
	7	**645**
	8	577
	9	572
	10	473

Current week

final order amount (frozen order amount)

forecasts

At week 1 the forecast for week 7 is 645 pcs.

Modification matrix

- Previous values are not overwritten → saved as historical forecast data

forecast delivery plans	delivery plans 1	2	3	4	5	6	7	8	9	10
1	620									
2	640	520								
3	814	710	550							
4	525	533	470	370						
5	761	885	797	802	590					
6	583	583	738	649	620	470				
7	645	756	725	961	862	831	620			
8	577	644	644	689	818	812	694	560		
9	572	550	616	638	668	847	781	693	550	
10	473	473	486	495	531	563	675	630	585	450

Diagonal – final orders (frozen order amount)

Final order amount for week 5 is 590 pcs.

Modification matrix

- Previous values are not overwritten → saved as historical forecast data

forecast delivery plans		delivery plans									
		1	2	3	4	5	6	7	8	9	10
	1	620									
	2	640	520								
	3	814	710	550							
	4	525	533	470	370						
	5	761	885	797	802	590					
	6	583	583	738	649	620	470				
	7	645	756	725	961	862	831	620			
	8	577	644	644	689	818	812	694	560		
	9	572	550	616	638	668	847	781	693	550	
	10	473	473	486	495	531	563	675	630	585	450

Forecaststream for week 5

In week 1 (4 weeks before delivery) the forecast for week 5 was 761 pcs.

Forecast error measure: PE

- PE (percentage error)
- Deviation in percent from final order

$$PE_{i,j} = \frac{X_{i,j} - X_{0,j}}{X_{0,j}}$$

$X_{i,\,j}$ ... $forecast\ for\ period\ j\ predicted\ i\ periods\ before$

$X_{0,\,j}$ ... $final\ order\ for\ period\ j$

Modification matrix

- Percentage Error (PE)

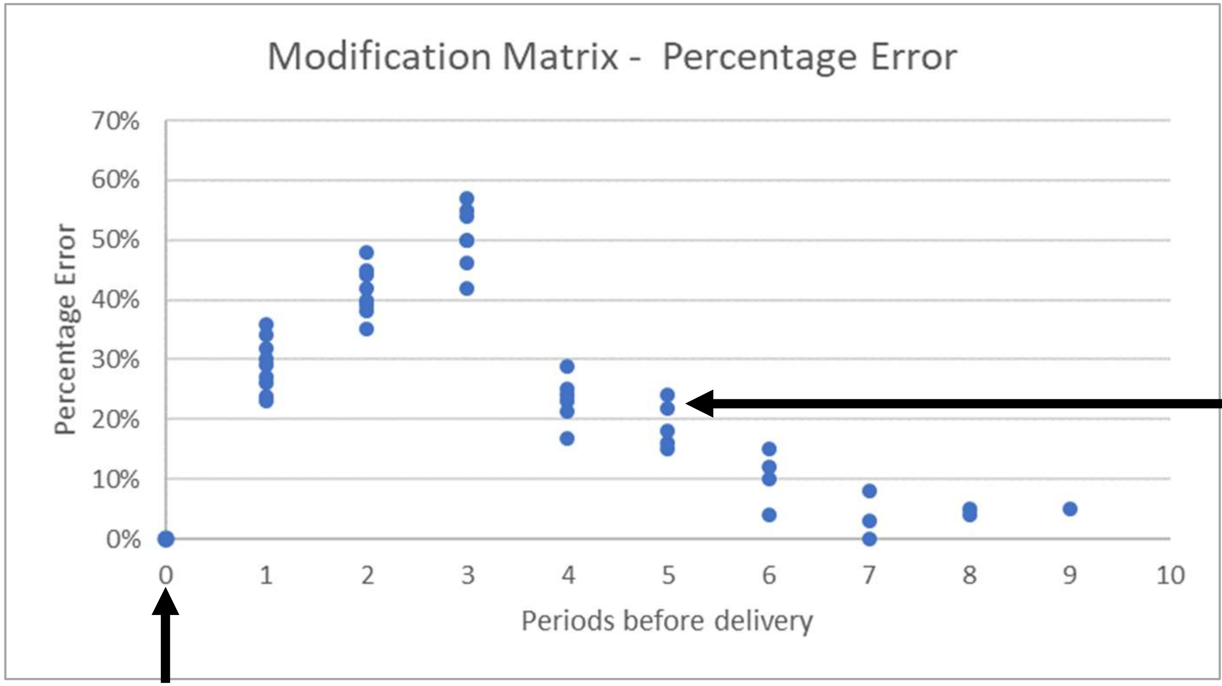
		delivery plans									
		1	2	3	4	5	6	7	8	9	10
forecast delivery plans	1	0,00									
	2	0,23	0,00								
	3	0,48	0,29	0,00							
	4	0,42	0,44	0,27	0,00						
	5	0,29	0,50	0,35	0,36	0,00					
	6	0,24	0,24	0,57	0,38	0,32	0,00				
	7	0,04	0,22	0,17	0,55	0,39	0,34	0,00			
	8	0,03	0,15	0,15	0,23	0,46	0,45	0,24	0,00		
	9	0,04	0,00	0,12	0,16	0,21	0,54	0,42	0,26	0,00	
	10	0,05	0,05	0,08	0,10	0,18	0,25	0,50	0,40	0,30	0,00

In week 1 (4 weeks before delivery) the order for week 5 was 29% overbooked (761/590 - 1)

Final order
0% deviation

Forecast uncertainty and cancellation behavior

- Analysis: systematic forecast behavior with the PE (percentage Error)



5 periods before delivery, the forecast was approx. 15 – 25% above the final order

Final order

Systematic Forecast behaviour

- Over- or underbooking of Forecasts:
  - Forecast values are in average to high/low compared to the final order amount
  - With Respect to periods before delivery
  - Can be described with the forecast error measure MPE
- Possible reasons for systematic forecast behaviour could be
  - Unconsidered demand changes
  - Capacity reservation of customers
  - General uncertainties

Forecast error measure: MPE

- MPE (mean percentage error)
- Systematic forecast behavior
  - MPE > 0 → Overbooking
  - MPE < 0 → Underbooking

$$MPE_i = \frac{\sum_{j=1}^{N}(X_{i,j} - X_{0,j})}{\sum_{j=1}^{N} X_{0,j}}$$

$X_{i,\,j}$ … $forecast\ for\ period\ j\ predicted\ i\ periods\ before$

$X_{0,\,j}$ … $final\ order\ for\ period\ j$

N … $number\ of\ forecast\ streams$

Periods before delivery (PBD) Matrix (1/6)

- Other form of presentation
- Transformation: modification matrix to PBD Matrix

		delivery plans			
		1	2	3	4
forecast delivery plans	1	620			
	2	640	520		
	3	814	710	550	
	4	525	533	470	370
	5	761	885	797	802
	6	583	583	738	649
	7	645	756	725	961
	8	577	644	644	689
	9	572	550	616	638
	10	473	473	486	495

		delivery plans (calendar weeks)			
		1	2	3	4
Periods before delivery	0				
	1				
	2				
	3				
	4				
	5				
	6				
	7				
	8				
	9				

Periods before delivery (PBD) Matrix (2/6)

- Final orders → Periods before delivery = 0

		delivery plans			
		1	2	3	4
forecast delivery plans	1	620			
	2	640	520		
	3	814	710	550	
	4	525	533	470	370
	5	761	885	797	802
	6	583	583	738	649
	7	645	756	725	961
	8	577	644	644	689
	9	572	550	616	638
	10	473	473	486	495

		delivery plans (calendar weeks)			
		1	2	3	4
Periods before delivery	0	620	520	550	370
	1				
	2				
	3				
	4				
	5				
	6				
	7				
	8				
	9				

Periods before delivery (PBD) Matrix (3/6)

- Forecast for period 2 in period 1 = 640 → PBD = 1

forecast delivery plans		delivery plans			
		1	2	3	4
	1	620			
	2	640	520		
	3	814	710	550	
	4	525	533	470	370
	5	761	885	797	802
	6	583	583	738	649
	7	645	756	725	961
	8	577	644	644	689
	9	572	550	616	638
	10	473	473	486	495

Periods before delivery		delivery plans (calendar weeks)			
		1	2	3	4
	0	620	520	550	370
	1		640		
	2				
	3				
	4				
	5				
	6				
	7				
	8				
	9				

Periods before delivery (PBD) Matrix (4/6)

- Forecasts PBD = 1

forecast delivery plans	delivery plans 1	2	3	4
1	620			
2	640	520		
3	814	710	550	
4	525	533	470	370
5	761	885	797	802
6	583	583	738	649
7	645	756	725	961
8	577	644	644	689
9	572	550	616	638
10	473	473	486	495

Periods before delivery	delivery plans (calendar weeks) 1	2	3	4
0	620	520	550	370
1		640	710	470
2				
3				
4				
5				
6				
7				
8				
9				

Periods before delivery (PBD) Matrix (5/6)

- Forecasts for period 4

forecast delivery plans	delivery plans 1	2	3	4
1	620			
2	640	520		
3	814	710	550	
4	525	533	470	370
5	761	885	797	802
6	583	583	738	649
7	645	756	725	961
8	577	644	644	689
9	572	550	616	638
10	473	473	486	495

Periods before delivery	delivery plans (calendar weeks) 1	2	3	4
0	620	520	550	370
1		640	710	470
2			814	533
3				525
4				
5				
6				
7				
8				
9				

Periods before delivery (PBD) Matrix (6/6)

- Periods before delivery (PBD) = 0 = Final Order Amount
- Only calendar week 10 has a complete historical forecast stream with 9 PBD

	delivery plans (calendar weeks)									
Periods before delivery	1	2	3	4	5	6	7	8	9	10
0	620	520	550	370	590	470	620	560	550	450
1		640	710	470	802	620	831	694	693	585
2			814	533	797	649	862	812	781	630
3				525	885	738	961	818	847	675
4					761	583	725	689	668	563
5						583	756	644	638	531
6							645	644	616	495
7								577	550	486
8									572	473
9										473

Forecast uncertainty – Example 2

- Calculate the PE (percentage error) from the forecast values to the final order amount
  - Data set 2 (Product 10 and 60)

- Which product and period before delivery has the highest overbooking on average in data set 2?
- Visualize the results in a scatterplot over periods before delivery
- What is the difference between these two data sets regarding forecast uncertainty?

Forecast history data

– Product 10

> Random forecast errors

Product 10		delivery plans (calendar weeks)									
		1	2	3	4	5	6	7	8	9	10
Periods before delivery	0	832	899	789	644	703	838	874	836	879	849
	1	796	910	831	640	684	855	857	839	908	852
	2	788	879	833	645	683	851	879	831	900	824
	3	803	876	835	665	664	854	880	814	922	820
	4	813	866	832	671	673	844	855	801	884	808
	5	813	852	827	681	674	840	861	794	888	801
	6	828	835	790	684	688	846	846	790	839	817
	7	786	835	803	730	683	834	832	808	837	846
	8	786	808	808	748	697	816	801	826	812	808
	9	795	791	791	741	717	845	798	794	809	798
	10	822	800	782	772	787	821	774	806	816	789

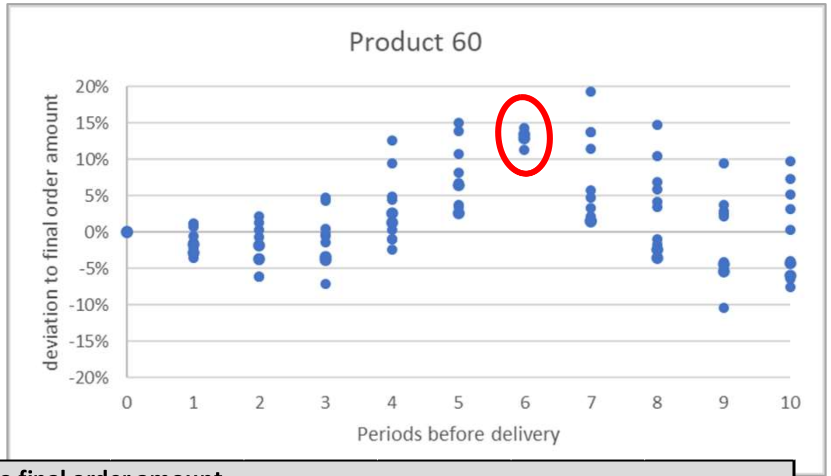
Forecast history data

- Product 60
  - Systematic bias (overbooking and cancellation)

Product 60		delivery plans (calendar weeks)									
		1	2	3	4	5	6	7	8	9	10
Periods before delivery	0	752	843	817	869	758	854	814	810	733	766
	1	748	829	826	838	742	830	820	796	723	775
	2	762	827	811	816	712	822	817	779	749	753
	3	784	814	815	837	704	821	802	806	768	769
	4	847	865	853	861	751	865	816	791	802	803
	5	856	897	905	892	784	876	869	840	843	829
	6	939	952	992	984	844	969	920	926	955	947
	7	855	855	910	919	783	867	853	827	874	871
	8	783	813	873	854	803	834	842	802	841	846
	9	780	797	841	778	778	816	781	773	802	783
	10	807	807	820	803	782	803	782	758	804	806

Forecast uncertainty – Example 2

– Solution



Product 60		deviation to final order amount										
		1	2	3	4	5	6	7	8	9	10	mean
Periods before delivery	0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	1	-0,01	-0,02	0,01	-0,04	-0,02	-0,03	0,01	-0,02	-0,01	0,01	-0,01
	2	0,01	-0,02	-0,01	-0,06	-0,06	-0,04	0,00	-0,04	0,02	-0,02	-0,02
	3	0,04	-0,03	0,00	-0,04	-0,07	-0,04	-0,01	0,00	0,05	0,00	-0,01
	4	0,13	0,03	0,04	-0,01	-0,01	0,01	0,00	-0,02	0,09	0,05	0,03
	5	0,14	0,06	0,11	0,03	0,03	0,03	0,07	0,04	0,15	0,08	0,07
	6	0,25	0,13	0,21	0,13	0,11	0,13	0,13	0,14	0,30	0,24	**0,18**
	7	0,14	0,01	0,11	0,06	0,03	0,02	0,05	0,02	0,19	0,14	0,08
	8	0,04	-0,04	0,07	-0,02	0,06	-0,02	0,03	-0,01	0,15	0,10	0,04
	9	0,04	-0,05	0,03	-0,10	0,03	-0,04	-0,04	-0,05	0,09	0,02	-0,01
	10	0,07	-0,04	0,00	-0,08	0,03	-0,06	-0,04	-0,06	0,10	0,05	0,00

Unsystematic Forecast behaviour

- General forecast quality with no systematic effects:
  - > With Respect to periods before delivery
  - > Can be described with the forecast error measure NRMSE
  - > Forecast quality should become better near the due date
- Possible reasons for unsystematic forecast behaviour could be
  - > General uncertainties
  - > Demand fluctuation
  - > Information quality
  - > Wrong forecasting method

Forecast error measure: RMSE

- RMSE (root mean squared error)
- Unsystematic forecast behavior
- Similar to standard deviation from statistics
- Lower RMSE is better than a higher one

$$RMSE_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(X_{i,j} - X_{0,\,j})^2}$$

$X_{i,\,j}$ ... *forecast for period j predicted i periods before*

$X_{0,\,j}$ ... *final order for period j*

N ... *number of forecast streams*

Forecast error measure: NRMSE

- NRMSE (normalized root mean squared error)
- Unsystematic forecast behavior
- Similar to variation coefficient from statistics
- Lower RMSE is better than a higher one

$$NRMSE_i = \frac{RMSE_i}{\frac{1}{N}\sum_{j=1}^{N} X_{0,j}}$$

$X_{i,\,j}$ ... $forecast\ for\ period\ j\ predicted\ i\ periods\ before$

$X_{0,\,j}$ ... $final\ order\ for\ period\ j$

N ... $number\ of\ forecast\ streams$

NRMSE – Example 3

- Calculate the NRMSE (normalized root mean squared error) with respect to periods before delivery
  - > Data set 2 (Product 10 and 60)

- Interpret the results according to forecast uncertainty
- Visualize the results in a line plot over periods before delivery

NRMSE – Example 3

- Solution

Number of forecast streams	10
Total final orders	8143

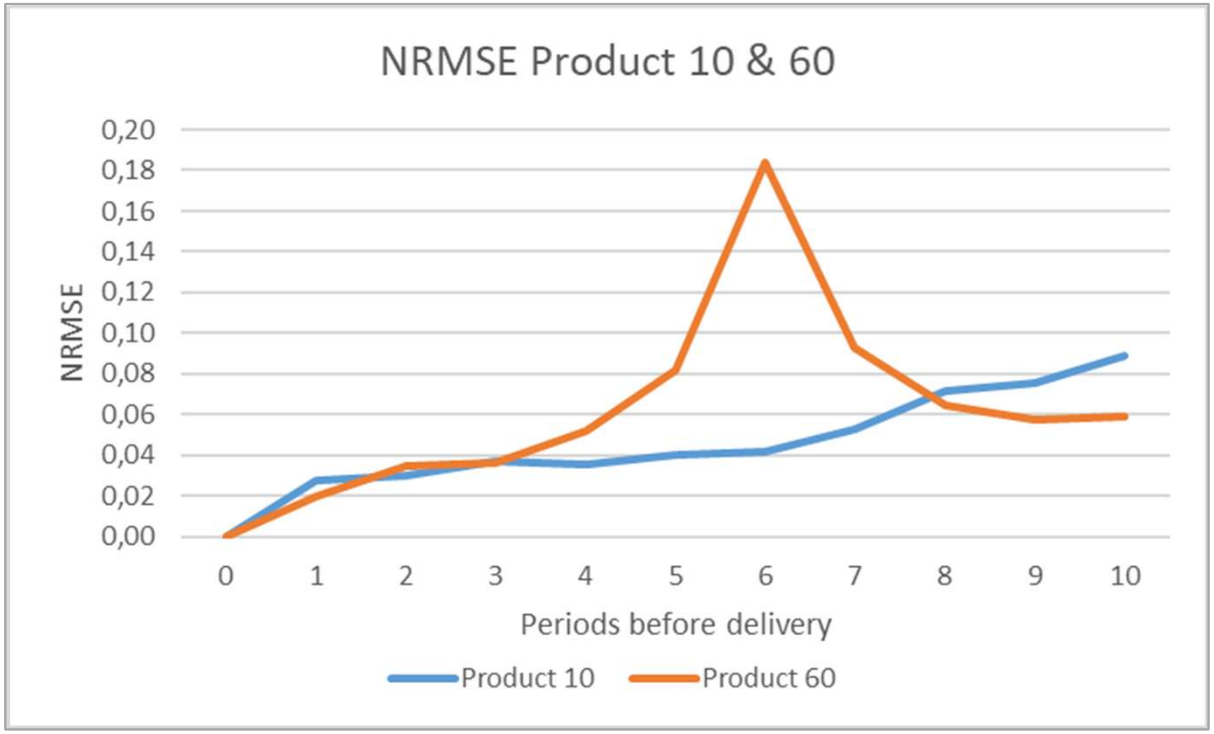
Product 10		RMSE	NRMSE
Periods before delivery	0	0,00	0,00
	1	22,35	0,03
	2	24,41	0,03
	3	29,86	0,04
	4	28,73	0,04
	5	32,47	0,04
	6	33,95	0,04
	7	43,06	0,05
	8	58,05	0,07
	9	61,43	0,08
	10	72,21	0,09

Number of forecast streams	10
Total final orders	8016

Product 60		RMSE	NRMSE
Periods before delivery	0	0,00	0,00
	1	15,81	0,02
	2	27,81	0,03
	3	28,80	0,04
	4	41,87	0,05
	5	65,70	0,08
	6	147,67	0,18
	7	74,44	0,09
	8	51,54	0,06
	9	45,94	0,06
	10	47,15	0,06

NRMSE – Example 3

- Solution



Forecast error measure: MAD

- MAD (mean absolute deviation)
- Unsystematic forecast behavior
- Not suitable for comparison two products
- Lower MAD is better than a higher one

$$MAD_i = \frac{1}{N}\sum_{j=1}^{N}\left|X_{i,j} - X_{0,\,j}\right|$$

$X_{i,\,j}$ ... $forecast\ for\ period\ j\ predicted\ i\ periods\ before$

$X_{0,\,j}$ ... $final\ order\ for\ period\ j$

N ... $number\ of\ forecast\ streams$

Forecast error measure: MAPE

- MAPE (mean absolute percentage error)
- Unsystematic forecast behavior

$$MAPE_i = \frac{\sum_{j=1}^{N} \left| X_{i,j} - X_{0,j} \right|}{\sum_{j=1}^{N} X_{0,j}}$$

$X_{i,\,j}$ ... $forecast\ for\ period\ j\ predicted\ i\ periods\ before$

$X_{0,\,j}$ ... $final\ order\ for\ period\ j$

N ... $number\ of\ forecast\ streams$

Forecast error measure: MFB

- MFB (mean forecast bias)
- Systematic forecast behavior
  - MFB > 1 → Overbooking
  - MFB < 1 → Underbooking

$$MFB_i = \frac{\sum_{j=1}^{N} X_{i,j}}{\sum_{j=1}^{N} X_{0,j}}$$

$X_{i,\,j}$ ... $forecast\ for\ period\ j\ predicted\ i\ periods\ before$

$X_{0,\,j}$ ... $final\ order\ for\ period\ j$

N ... $number\ of\ forecast\ streams$

Problems of Outliers

- **Groupwork:**
  - Discuss possible problems of outliers in final orders and forecast streams
  - Create a flipchart to present the results
  - Time: 20 min

Outliers (1/3)

- An Outlier is a value which is way too high or low compared to other values in the same numerical series
- Can caused by unpredictable situations, one-off effects, human mistakes,…
- Outliers can be detected through statistical approaches
  - Interquartile ranges (e.g. 1,5 * IQR-Rule)
  - Number of standard deviations (e.g. 2-σ-Method)
  - Outlier tests (e.g. Grubbs test)
- Outliers according to forecasting and customer demand are mostly values which are too high

Outliers (2/3)

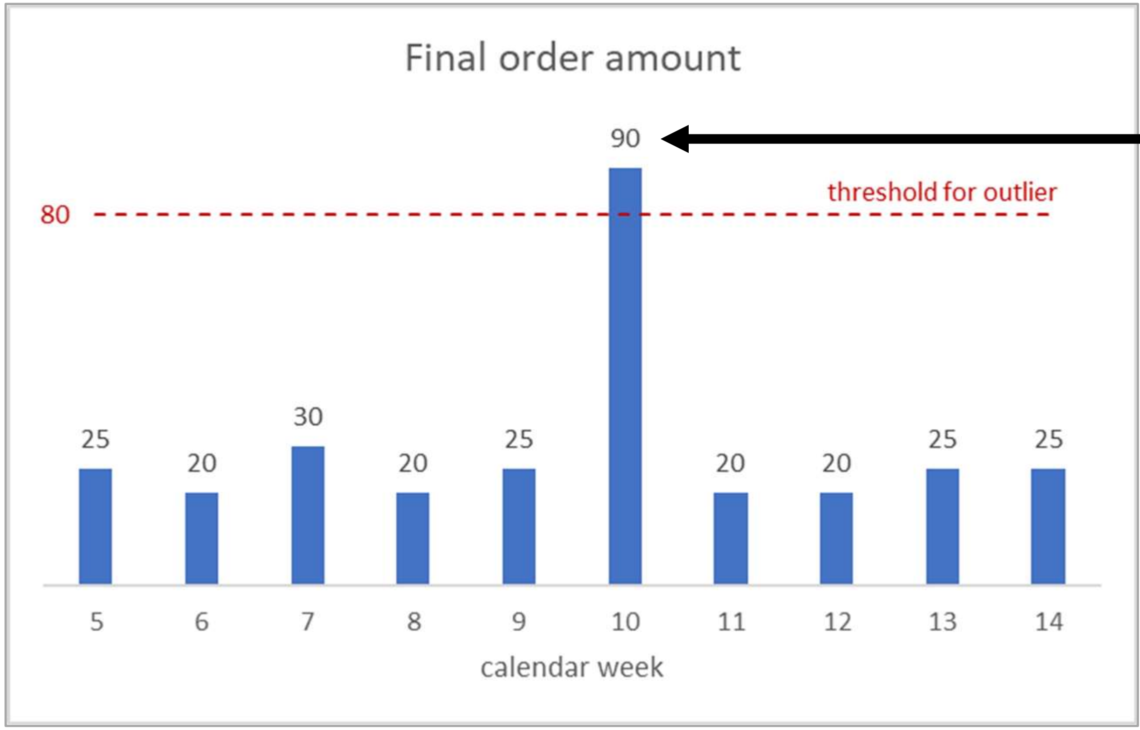
- Outliers in previous final orders
  - Can distort future forecasts (e.g. moving average)
  - Negative impact on forecast error measures
  - Can lead to stock out situations
- Outliers in forecast stream
  - Is the forecast value correct or is this a mistake?
  - Is there a plausible explanation?
  - Can lead to big problem in production planning due to high needed capacity

Outliers (3/3)

- Outliers in previous final orders
  - Final order amounts from one period compared to other periods
- Outliers in forecast stream
  - Focus on one forecast stream
  - Forecast value within the forecast stream compared to the other forecast values
  - Compare the changes within the forecast stream (e.g., difference of forecast values between 4 and 5 periods before delivery)

Outlier in final Orders

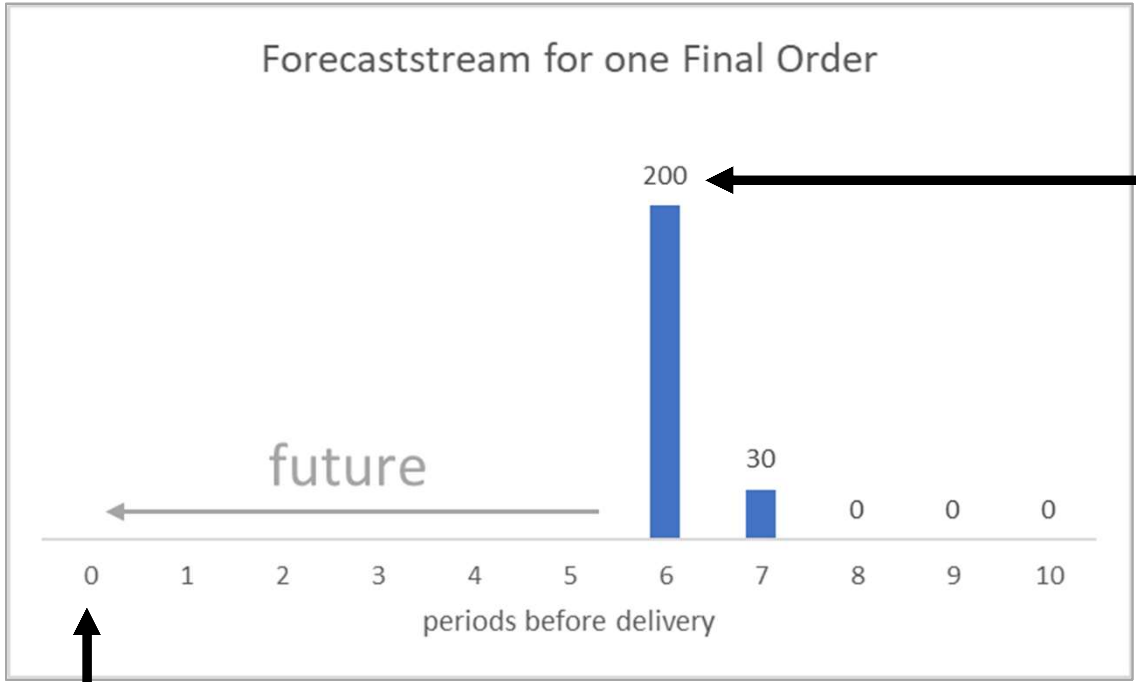
– Threshold value for outlier is determined with 80 pcs



Outlier in calendar week 10 detected

Outlier in Forecast Stream (1/4)

– Consideration 6 periods before delivery

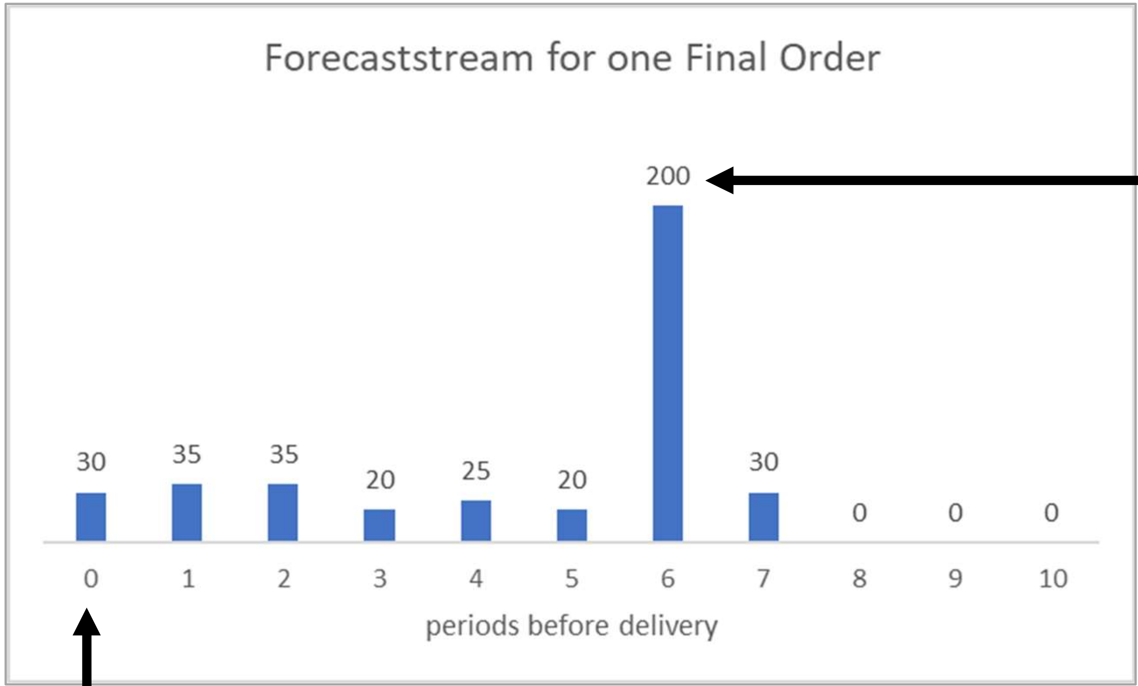


Outlier yes/no?
Could also be an increase in demand

Due date

Outlier in Forecast Stream (2/4)

- Outlier in forecast stream 6 periods before delivery

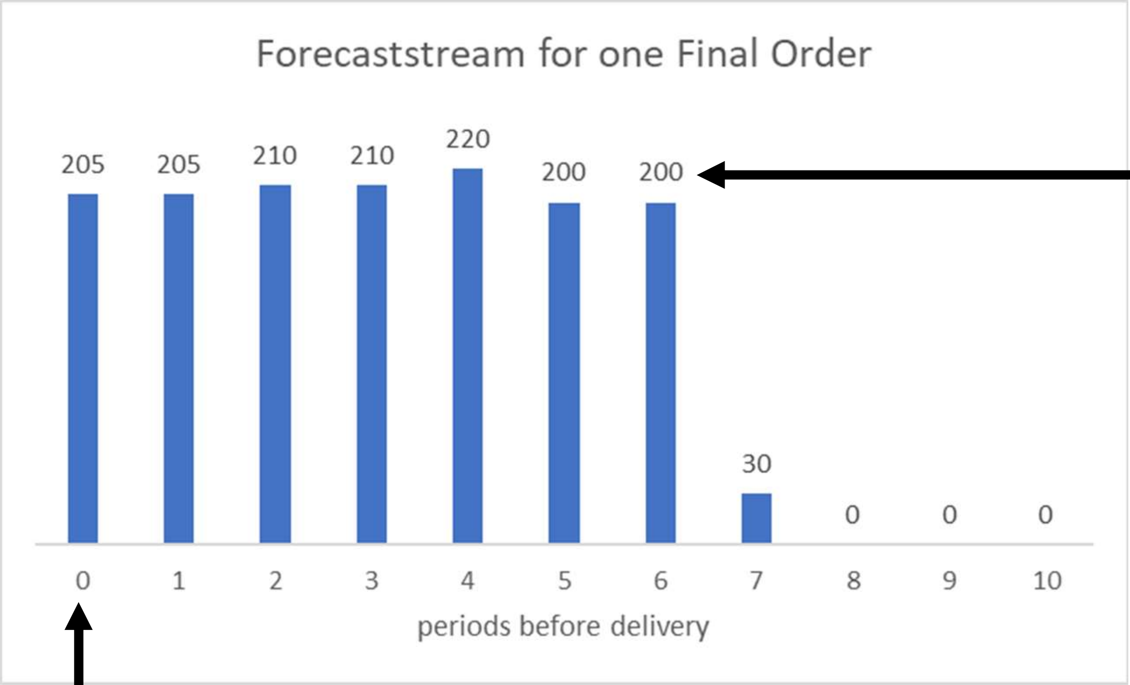


Outlier 6 periods before delivery detected

Due date

Outlier in Forecast Stream (3/4)

– No Outlier in forecast stream 6 periods before delivery

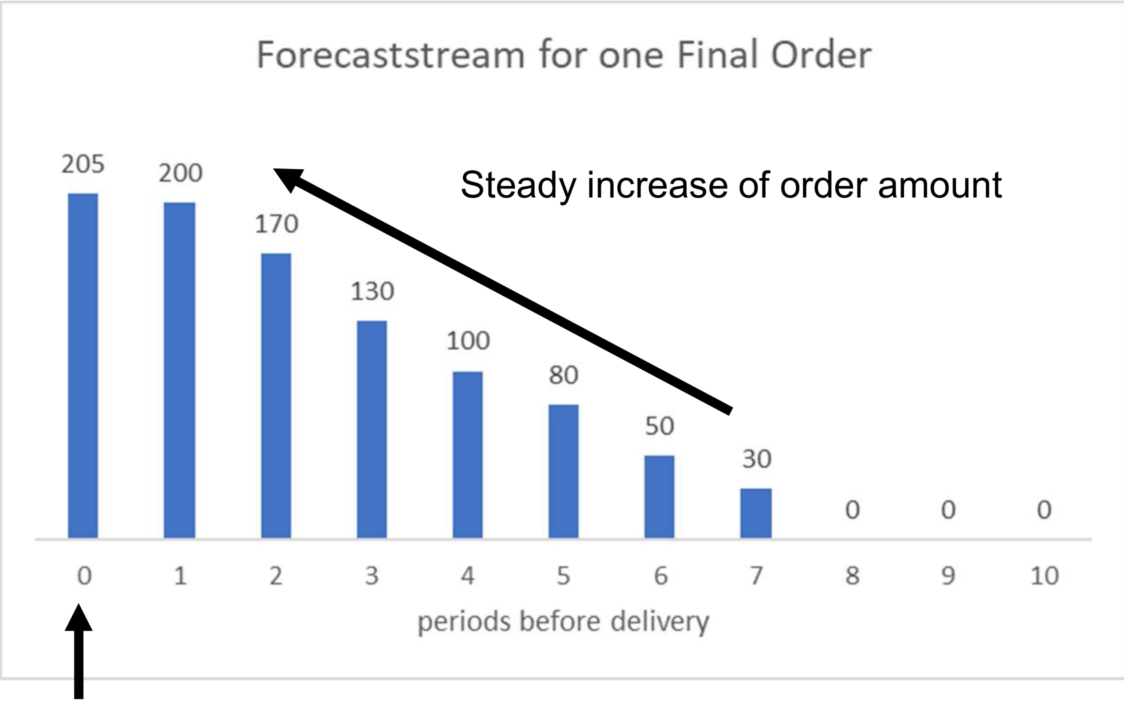


Increase in quantity 6 periods before delivery

Due date

Outlier in Forecast Stream (4/4)

- No Outlier in forecast stream, forecast is steadily increasing

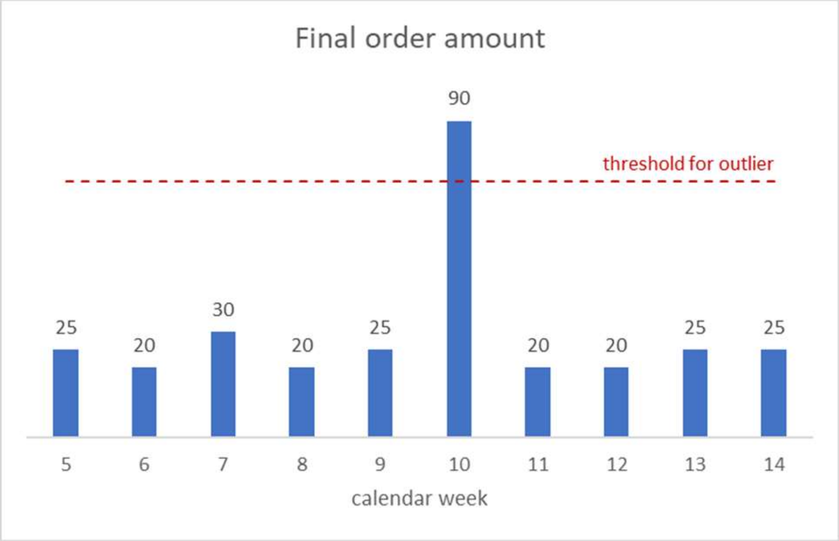


Due date

Outlier detection: Standard deviations

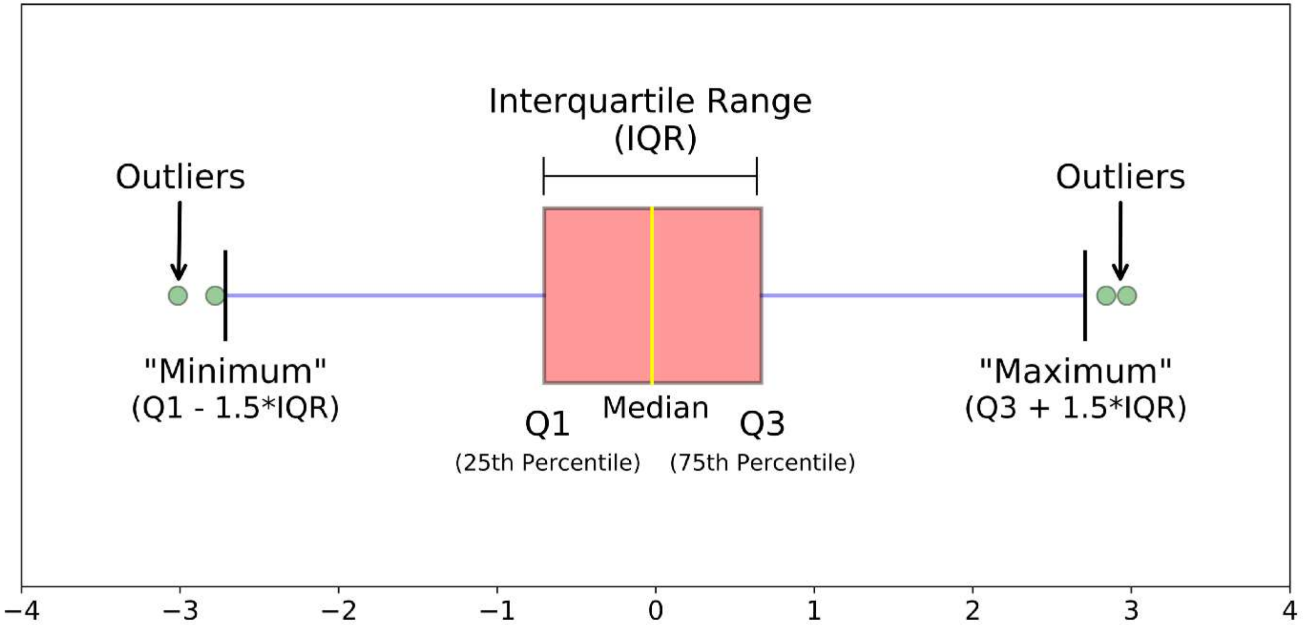
- Calculation of threshold value based on number of standard deviation
  - Determine the number of standard deviation
  - Calculation of standard deviation and average of the dataset
  - Threshold value for outlier = average + number of stdev * stdev

Average	30
Standard deviation	21
Number of standard deviations	2
Threshold value	73



Outlier detection: Interquartile range (1/2)

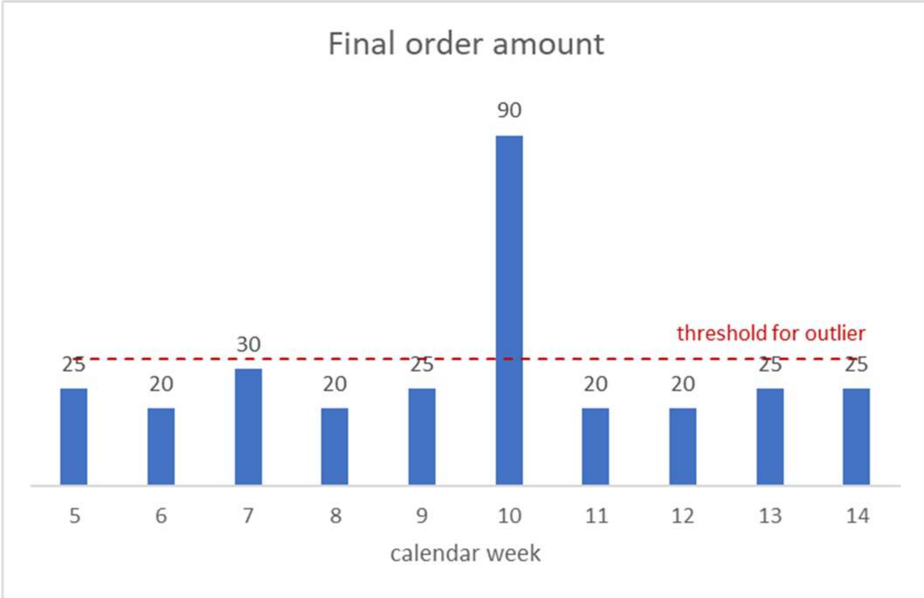
- Outliers can be visualized in a boxplot



Source: https://towardsdatascience.com/understanding-boxplots-5e2df7bcbd51 (17.01.2022)

Outlier detection: Interquartile range (2/2)

- IQR (Interquartile range) = Q3 – Q1
- Calculation in Excel
  - Q1 = PERCENTILE.INC(Array;0,25)
  - Q3 = PERCENTILE.INC(Array;0,75)
- Outlier Threshold value
  - Q3 + 1,5*IQR
  - Q1 – 1,5*IQR



First Quartil	20
Third Quartil	25
Interquartil Range	5
Threshold value	33

Outlier detection: Grubbs Test (1/4)

- Statistical approach to detect outliers
  - Test statistic = G
  - Threshold value = $G_{critical}$
  - Significance level = α (e.g. 0,05)
- If G is higher than $G_{critical}$, die maximum value in the data set is an outlier. After removing this value, the test must be repeated till there is no outlier in the data set anymore

Outlier detection: Grubbs Test (2/4)

$$G = \frac{max\, X_j - \bar{X}}{S_n}$$

X_j ... *final order for period j*

$\bar{X}$ ... *average of final orders*

S_n ... *standard deviation of n final orders*

$$G_{critical} = \frac{(n-1)t_{critical}}{\sqrt{[n(n-2+t^2{}_{critical})]}}$$

$t_{critical}$ ... *value of the t distribution*: $t_{\frac{\alpha}{n},n-2}$

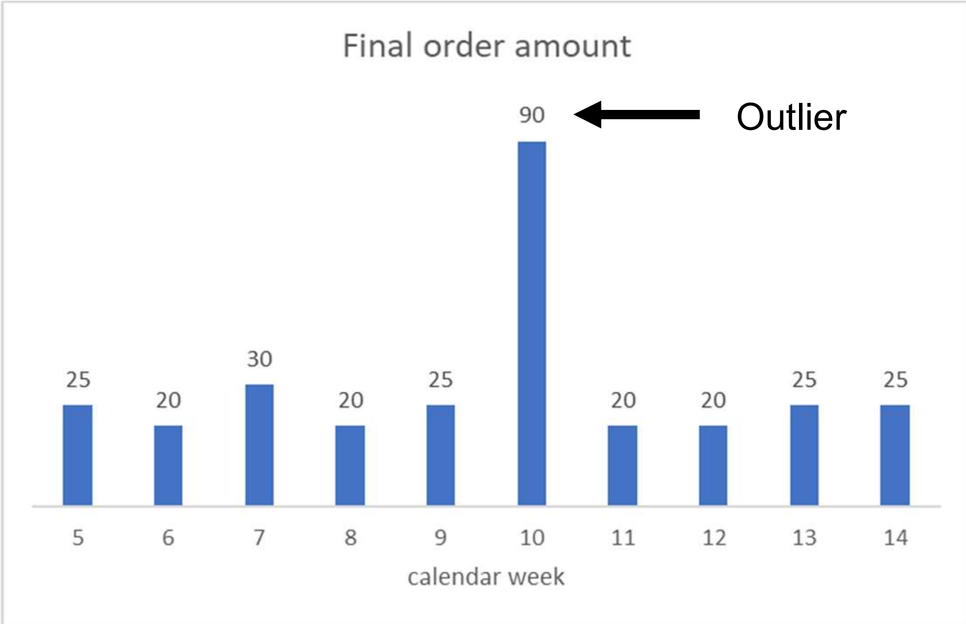
n ... *number of final orders*

α ... *significance level*

Outlier detection: Grubbs Test (3/4)

- Calculation Excel
  - > t_critical = T.INV(1-0,005;8)

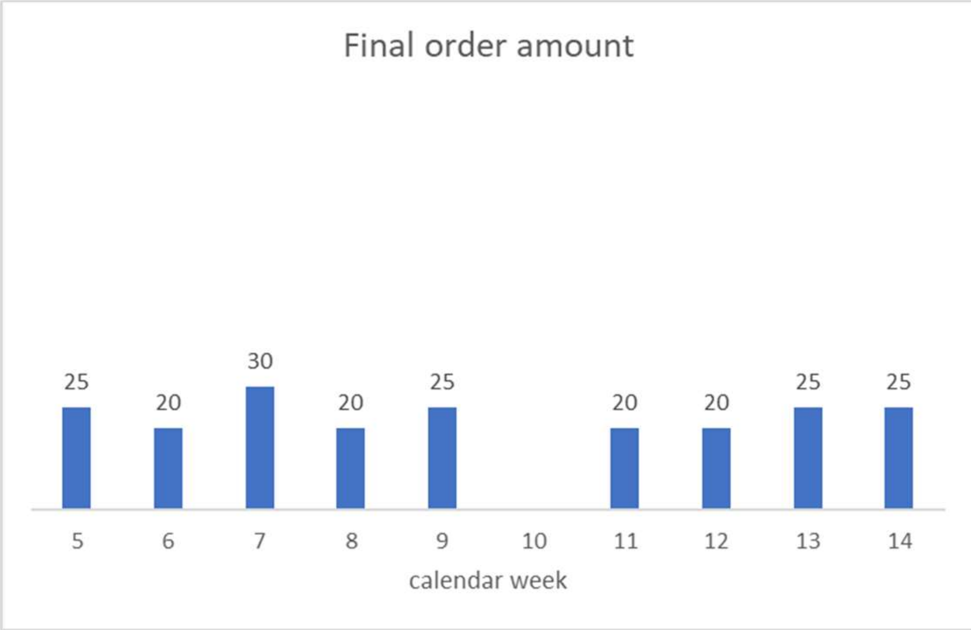
Maximum	90
Average	30
Standard Deviation	21
G	2,811
Significance Level (alpha)	0,05
Number of final orders	10
Significance Value (alpha/n)	0,005
Degrees of Freedom (n-2)	8
t_critical	3,355
G_critical	2,176
Outlier (G>G_critical)	**Yes**



Outlier detection: Grubbs Test (4/4)

- Outlier is removed and the test is repeated
- No outlier → Grubbs test is finished
- Calculation Excel

Maximum	30
Average	23
Standard Deviation	4
G	1,886
Significance Level (alpha)	0,05
Number of final orders	9
Significance Value (alpha/n)	0,006
Degrees of Freedom (n-2)	7
t_critical	3,422
G_critical	2,110
Outlier (G>G_critical)	**No**



Outlier detection – Example 4

- Identify possible outliers within the data set by using a threshold based on a certain number of standard deviations
  - > Data set 3 (Product 30)

- Which final order is identified as an outlier and how much standard deviations did you use for calculating the threshold?
- Visualize the final order amounts in a bar chart
- Did you recognize further outliers in the data set?

Forecast history data

- Product 30
  - > Random forecast errors + outliers

Product 30		delivery plans (calendar weeks)									
		1	2	3	4	5	6	7	8	9	10
Periods before delivery	0	561	549	438	780	173	530	1150	2517	787	1132
	1	578	547	241	721	37	320	1054	2262	874	1321
	2	438	654	278	754	0	427	1034	1242	893	1261
	3	487	775	242	856	13	639	850	1156	828	1253
	4	425	796	532	673	117	788	699	1165	579	1173
	5	597	781	713	725	430	1013	705	1063	645	1262
	6	670	531	4580	1005	522	1190	854	1053	325	1251
	7	627	365	553	798	753	1211	885	1028	471	1108
	8	559	483	737	789	898	791	917	911	445	1332
	9	752	790	624	822	693	848	994	913	674	1085
	10	807	741	806	992	766	768	795	830	739	970

Outlier detection – Example 4

- Solution (1/2)
  - Calculating mean and standard deviation from final orders
  - Compare Threshold value (Mean + St dev.*number of σ) with final orders

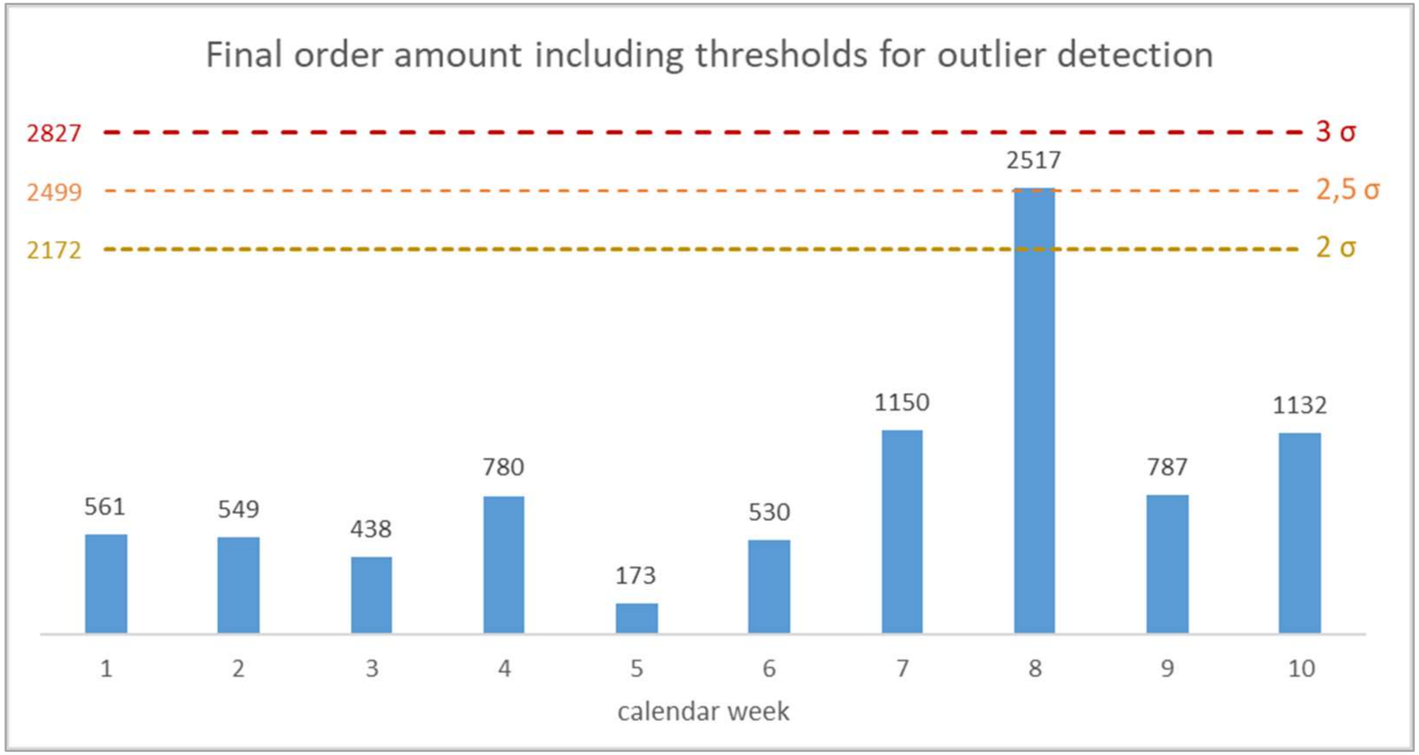
Product 30		Threshold for Outlier		
Week	Final Order	2 σ	2,5 σ	3 σ
1	561			
2	549			
3	438			
4	780			
5	173			
6	530			
7	1150			
8	2517	Outlier	Outlier	
9	787			
10	1132			

Mean	861,7
St.dev.	654,9391914

Threshold	
2 σ	2172
2,5 σ	=$B$36+$B$37*2,5
3 σ	2827

Outlier detection – Example 4

- Solution (2/2)



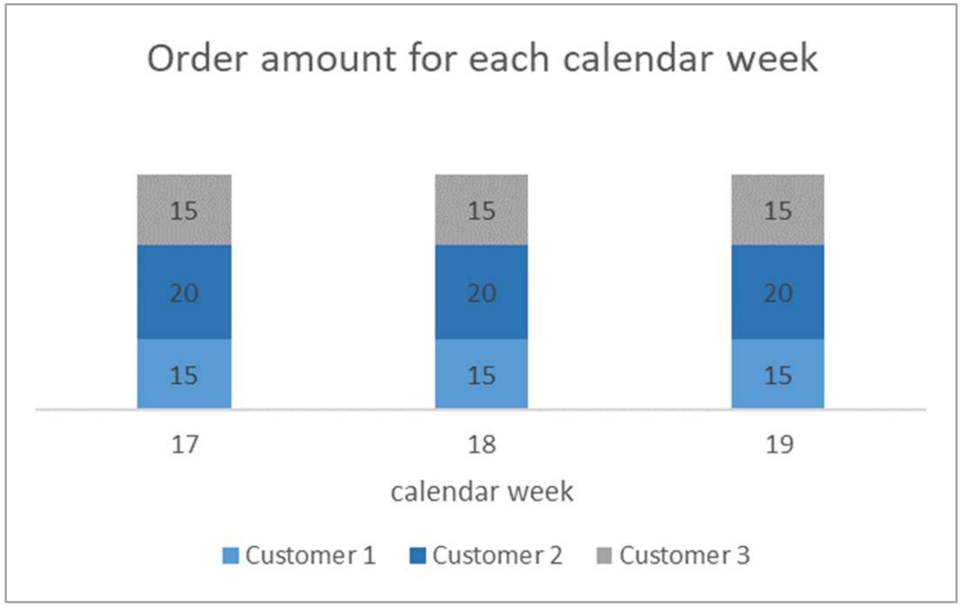
Amount shifting behaviour

- Specific type of change in delivery date, where the customer is shifting quantities from an existing order from one due date to another due date
- Amount shifting can occur in both directions (shifting forward or backwards)
- Shifted quantities can also be shifted back to the originally due date
- Hardly to identify by forecast error measures
- Possible reasons for unsystematic forecast behaviour could be
  - Demand change of the customer
  - Bundle up several orders to one delivery

Amount shifting in forecasts (1/3)

- Consideration 5 periods before calendar week 17 (=calendar week 12)
  - Customer 2 forecasted 20 pcs for calendar week 17, 18 and 19

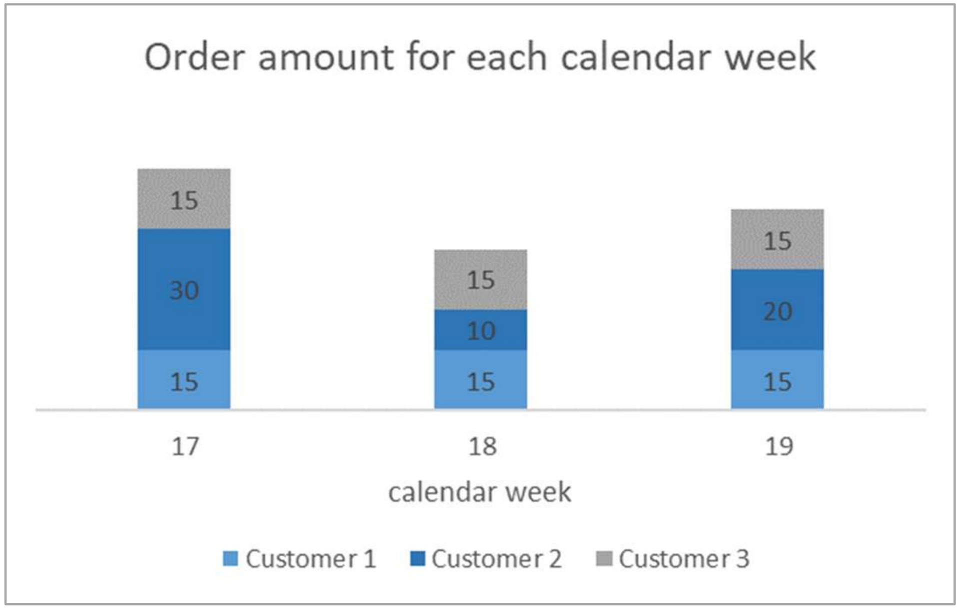
Customer 2		calendar weeks		
		17	18	19
Periods before delivery	0	future		
	1			
	2			
	3			
	4			
	5	20		
	6	20	20	
	7	20	20	20
	8	20	20	20



Amount shifting in forecasts (2/3)

- Consideration 4 periods before calendar week 17 (=calendar week 13)
  - Shift (forward) of 10 pcs from week 18 to 17

Customer 2		calendar weeks		
		17	18	19
Periods before delivery	0			
	1			
	2		future	
	3			
	4	30	10 pcs.	
	5	20	10	
	6	20	20	20
	7	20	20	20
	8	20	20	20

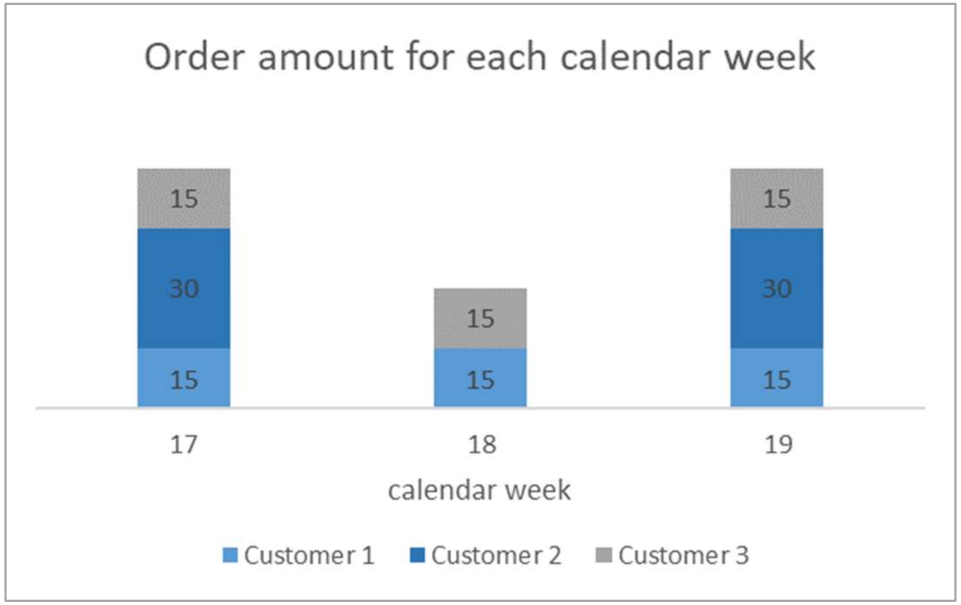


Amount shifting in forecasts (3/3)

- Consideration 1 periods before calendar week 17 (=calendar week 16)
  - Shift (backwards) of 10 pcs from week 18 to 19

Customer 2		calendar weeks 17	18	19
Periods before delivery	0			
	1	30		
	2	30	0	
	3	30	10	30
	4	30	10	20
	5	20	10	20
	6	20	20	20
	7	20	20	20
	8	20	20	20

10 pcs.

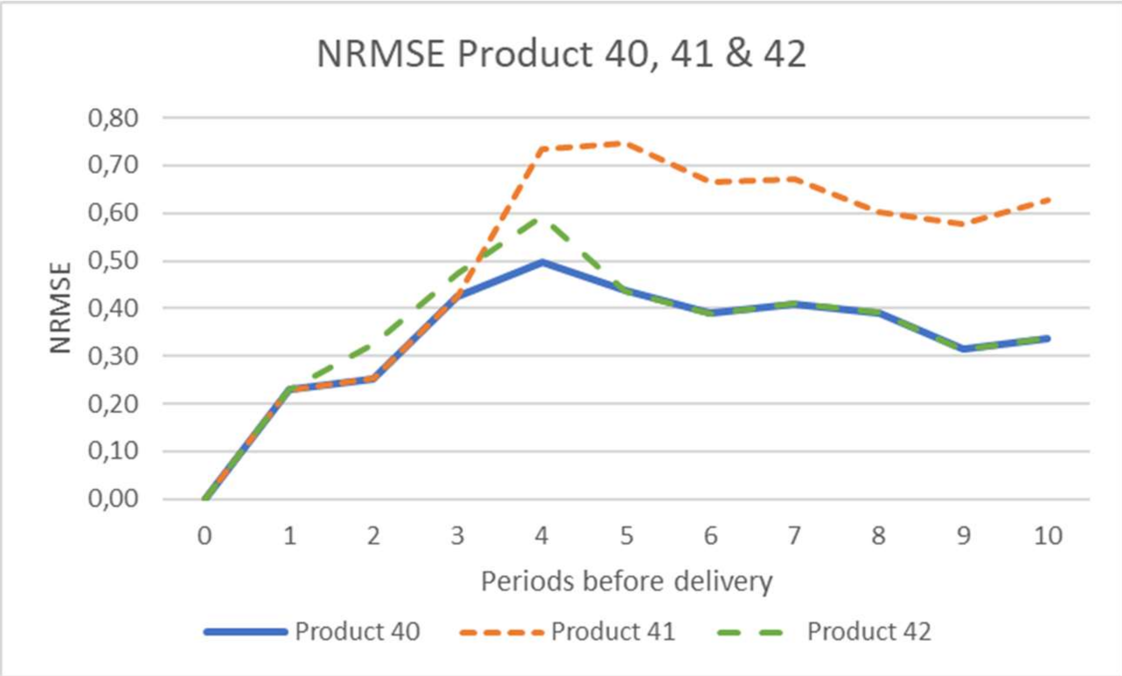


Amount shifting in forecasts – Example 5

- Calculate and analyze the NRMSE with respect to periods before delivery for the Products 40, 41 and 42.
  - Data set 4
    - Product 40 (no amount shifting)
    - Product 41 (amount shifting)
    - Product 42 (amount shifting and shifting back)
- Try to mark the shifting amounts in the data sets
- Hint: use example 3 as template to calculate the NRMSE

Amount shifting in forecasts – Example 5

- Solution



Cancelling of orders

- Customers are cancelling orders on a regular basis
  - With respect to periods before delivery
  - Related to specific customers
  - Can be described with the KPI confirmation safety
- Possible reasons for cancelling orders could be
  - Capacity reservation of customers
  - Long term forecasts were not updated till the order is cancelled
  - Demand change of the customer

Confirmation Safety

- KPI which describes the relative frequency, that a customer provided demand forecast leads to a final order with respect to periods before delivery (PBD)
- Useful if forecasts were cancelled on a regular basis
- Confirmation safety can be linked to customer forecast behaviour

$$CS_i = \frac{X_i}{N_i}$$

CS_i ... $Confirmation\ Safety\ i\ periods\ before\ delivery$

X_i ... $number\ of\ forecasts\ i\ periods\ before\ delivery\ , which\ lead\ to\ a\ final\ order$

N_i ... $number\ of\ forecasts\ i\ periods\ before\ delivery$

Overview of different forecast behaviours

- Unsystematic forecast behaviour (forecast quality)
- Systematic forecast (under- overbooking behaviour)
- Outliers in final orders and forecast streams
- Amount shifting in forecasts
- Cancelling of orders

Correlation and Covariance of Forecasts

Covariance and Correlation

- Are both measures to investigate the linear association between two variables
- Linear association between variables does not imply causality
- Correlation is calculated by standardizing the covariance
- Used before selecting the underlying forecasting model:
  - To test the application of a general linear model.
- The correlogram (Extension of the correlation calculation) is used to identify autoregression in historical data
  - Autoregression need a special forecasting method e.g. the ARIMA model

Covariance and Correlation

- Covariance

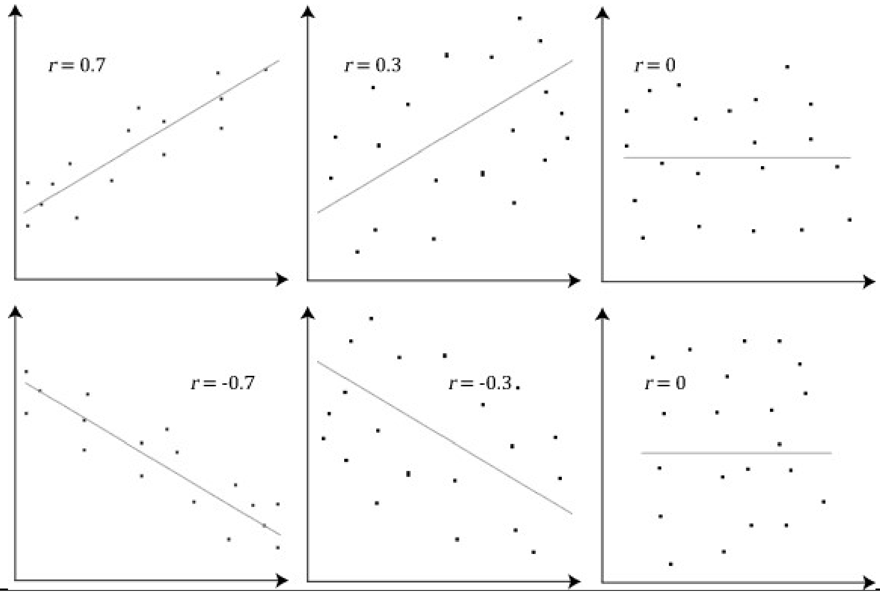
$$\sigma_{XY} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

- Correlation

$$r = \frac{\sigma_{XY}}{\sigma_x \sigma_y}$$

Covariance and Correlation

- Value range
  - $-1 < r < 0$: negative correlation, negative slope
  - $r = 0$: no correlation
  - $0 < r < 1$: positive correlation, positive slope



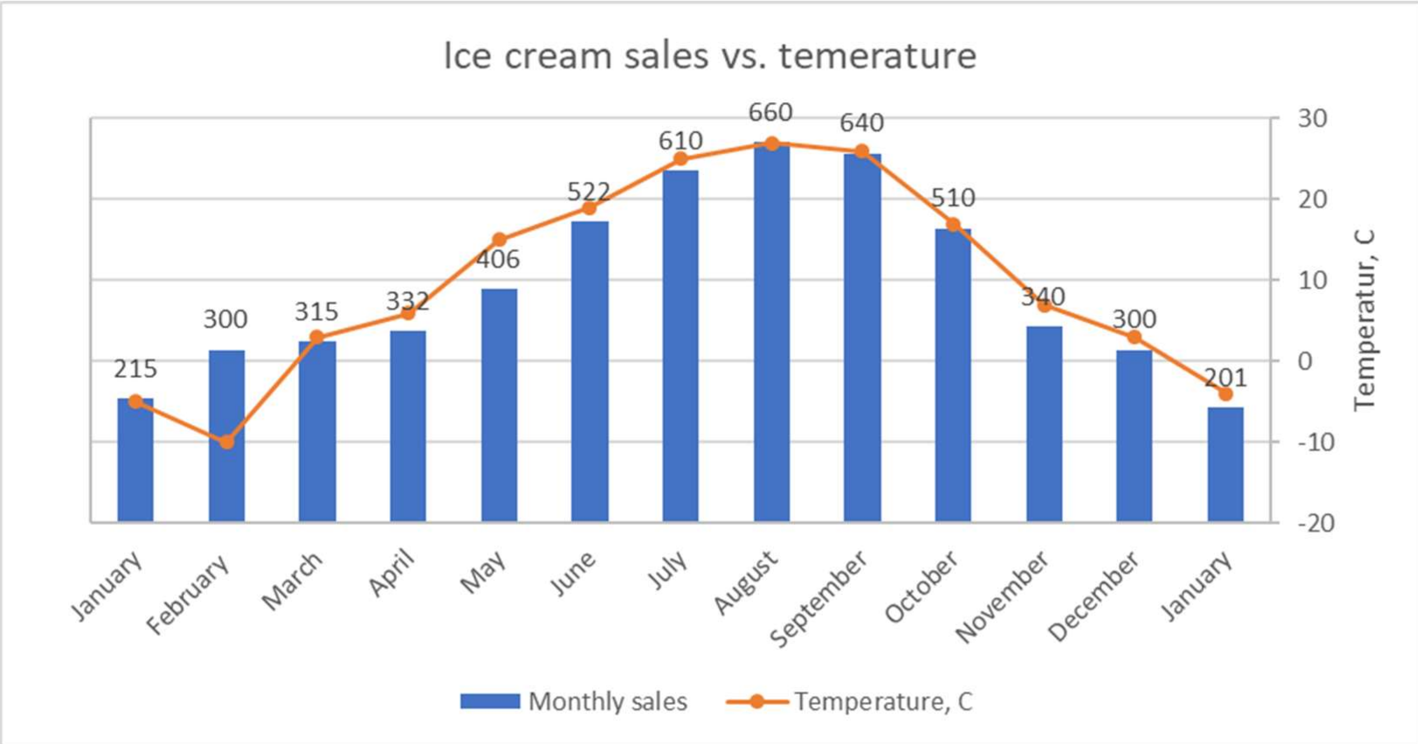
Source Wikipedia

Correlation

- Example: there is a data set of one year data of temperature change and sales of ice cream.
- Calculate the correlation of the data set between the temperature change and sales increase.

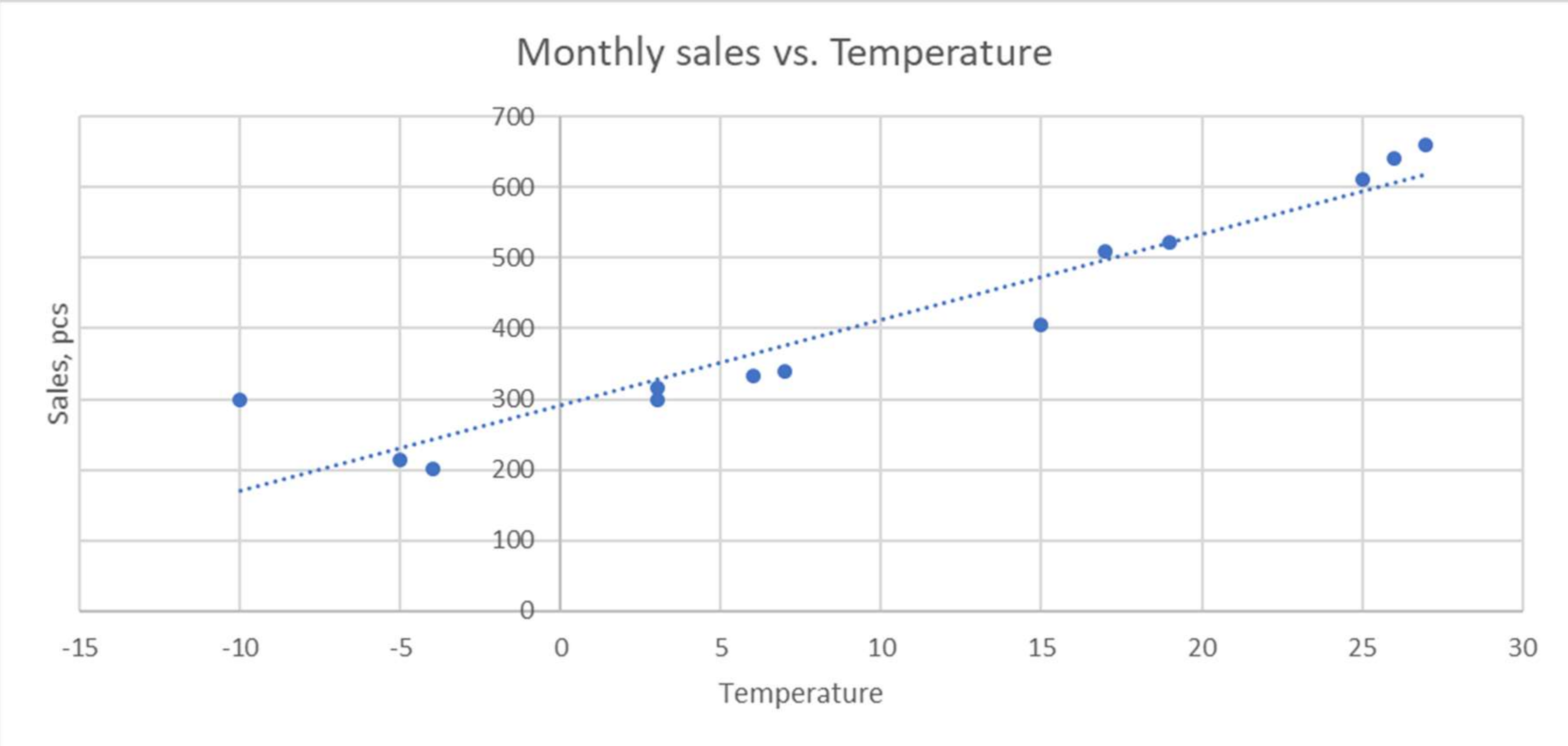
Month	Temperature, C	Monthly sales
January	-5	215
february	-10	300
March	3	315
April	6	332
May	15	406
June	19	522
July	25	610
August	27	660
September	26	640
October	17	510
November	7	340
December	3	300
January	-4	201

Correlation: ice cream sales vs. temperature



Distribution of sales (pcs) and temperature (C)

Correlation: ice cream sales vs. temperature

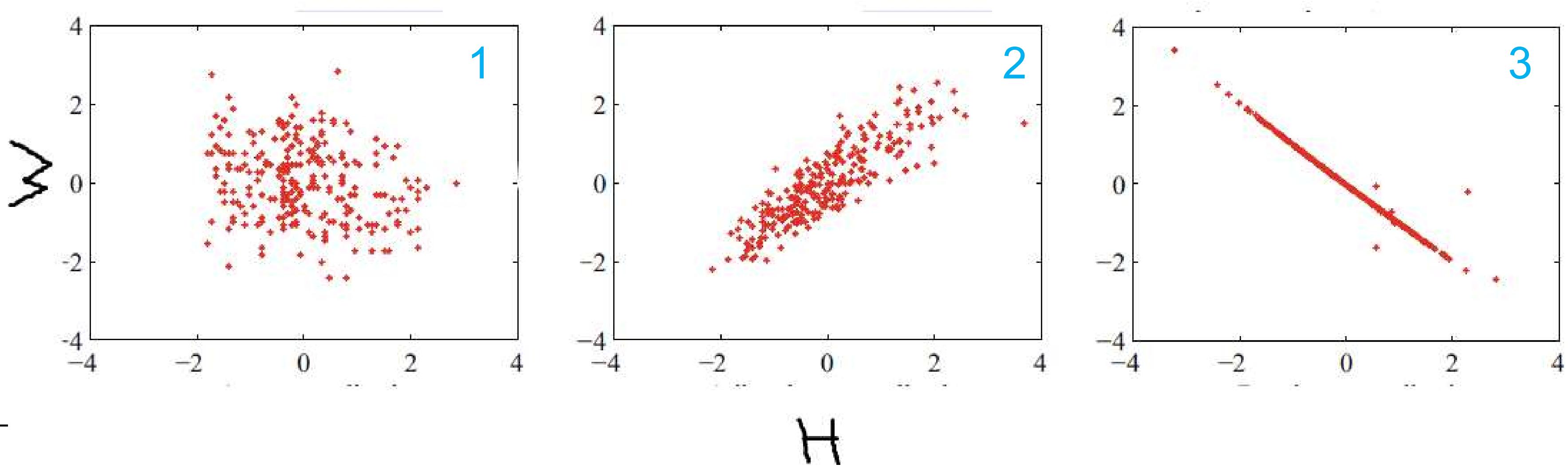


correlation	0,950

<u>Result</u>:
Positive correlation. Meaning the increase of temperature leads or correlates with the increase of sales.

Correlation between products

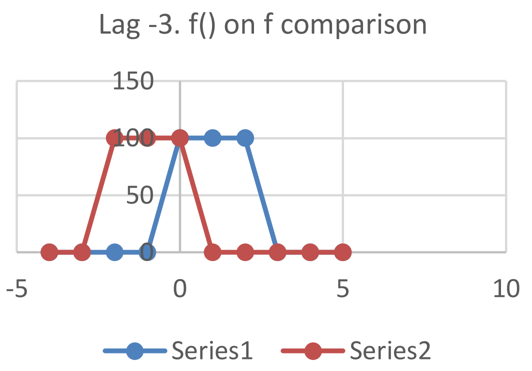
- The correlation *corr ({x,y})* is the **extent of relationship existence** between two components (also products or variables) *x* and *y* in a data set. **Zero correlation** means there is no relationship; **positive correlation** means larger $\dot{x}$ values tend to appear with larger $\dot{y}$ values; and **negative correlation** means that larger $\dot{x}$ values tend to appear with smaller $\dot{y}$ values.
- Say, we want to compare two products using a correlation, with x and y coordinates as heights (x) and weights (y), we have the following scatterplots depicting different correlations: (1) no relationship; (2) positive; (3) negative

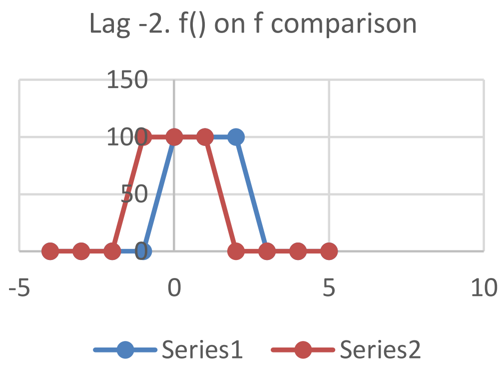


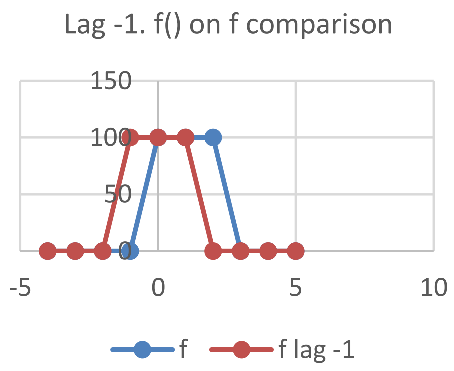
Forsyth, 2018

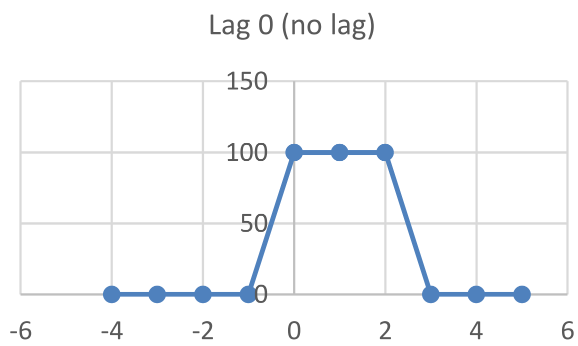
Cross-correlation

- Cross-correlation is the application of one function over the other in a step-by-step lag-based comparison (mapping of two functions, e.g., sliding dot product or sliding inner-product), applied in signal processing, to measure similarity of two series
- Correlation coefficient can be between -1 and 1, where 1 is the perfect positive or “optimal” correlation
- Mapping of a function over itself is called autocorrelation

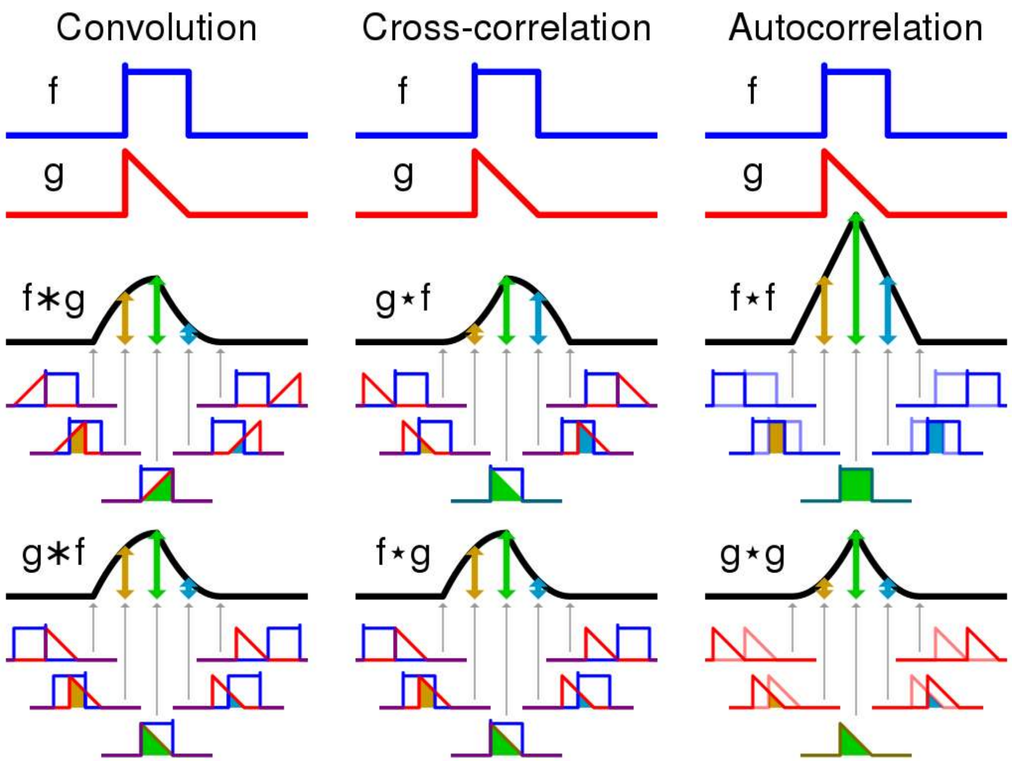








Convolution, cross-correlation and autocorrelation



– Source: wikipedia

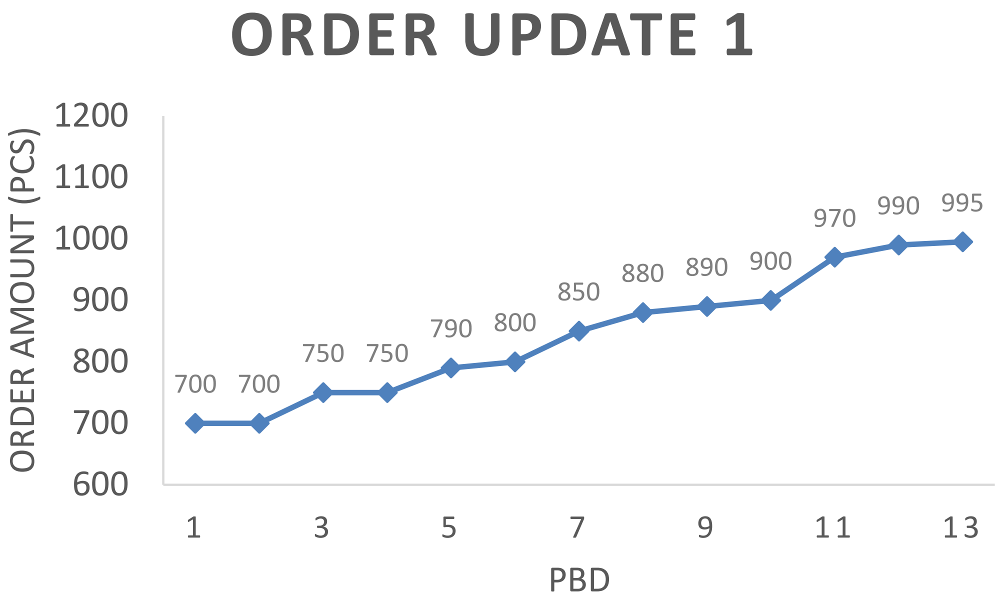
Correlation between updates

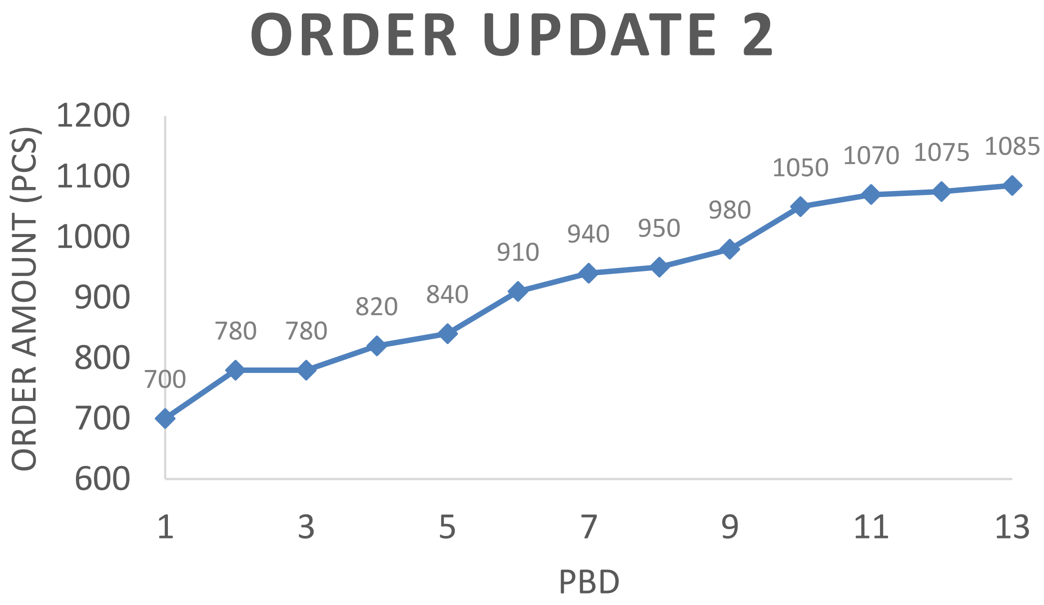
PBD	Order Update 1	Order Update 2
1	700	700
2	700	780
3	750	780
4	750	820
5	790	840
6	800	910
7	850	940
8	880	950
9	890	980
10	900	1050
11	970	1070
12	990	1075
13	995	1085

Example: there is a data set of two different order updates per each period before delivery (PBD). There is a dependency between Order Update 1 and Order Update 2, e.g., with 1 PBD delay, Order Update 2 increases in the same extent with respect to PBD. Calculate the correlation of two order updates with respect to PBD

Correlation between updates

- Positive correlation. Graph overview of two order updates:





Clustering of products according to their forecast behaviour

Definition Clustering

- A cluster is a group of data objects with similar properties
- Target of clustering is to identify different clusters in a dataset which have similar properties according to the selected attributes

- Clustering regarding forecast behavior based on
  - different forecast error measures with respect to periods before delivery
  - customer order behavior
  - other relevant attributes

Clustering Methods

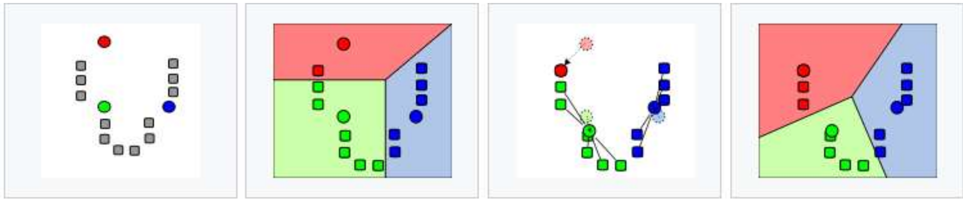
- Machine learning algorithms
  - K-Means
  - Agglomerative clustering
  - Affinity Propagation
- Least square approach with predefined clusters
- Clustering by experts

K-Means (1/2)

- Uses unsupervised learning to identify known clustering issues
- Divides dataset into predefined number of clusters
  - > Depending on the minimum distance, data points are allocated to defined number of clusters
  - > Measure the distance
  - > Assign datapoint to the nearest cluster
  - > Calculate the cluster mean (centroid) using the new point
  - > Repeat until the centroids are stable
- Requires:
  - > Number of clusters (cluster centroid), for example: 3 clusters in a training set => **K=3**

K-Means (2/2)

- Basic clustering algorithm to identify number of clusters based on the squared Euclidean distance to a nearest mean or centroid
- Cluster number is predefined by user
- Represents each cluster by a single mean vector



1. *k* initial "means" (in this case *k*=3) are randomly generated within the data domain (shown in color).

2. *k* clusters are created by associating every observation with the nearest mean. The partitions here represent the Voronoi diagram generated by the means.

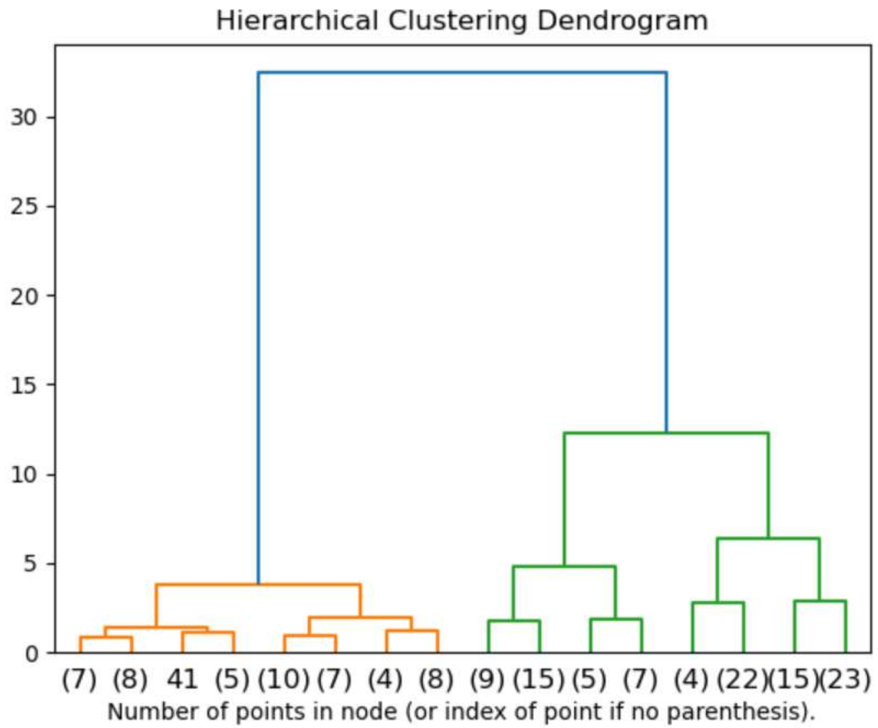
3. The centroid of each of the *k* clusters becomes the new mean.

4. Steps 2 and 3 are repeated until convergence has been reached.

Image source: wikipedia

Agglomerative Clustering

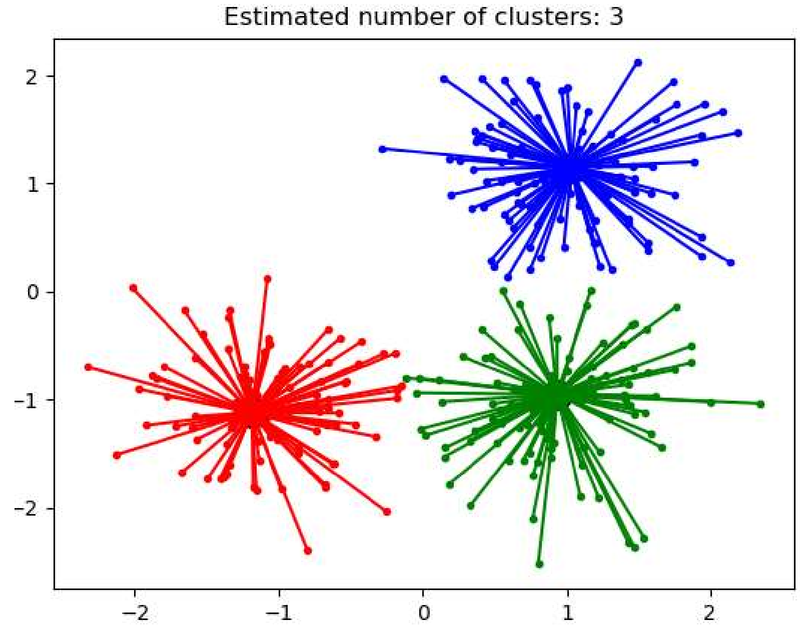
- Hierarachical clustering (bottom-up) method to merge pairs of clusters until all clusters are merged into one cluster
- Sensitive to data outliers and noise
- Has various metrics (euclidean, cosine, manhattan) and linkage criteria (ward, max/average/single linkage)
- Also requires the user to predefine the number of clusters



Source: scikit-learn.org

Affinity Propagation

- Finds exemplars of the input data that are representing the clusters → nearest neighbor graph
- **No need to predefine the cluster number**
- Requires a damping factor (0.5 - 1) which is the extent to which current value is maintained relative to incoming values. Default damping factor = 0.5



Source: scikit-learn.org

Least Square Approach with predefined Clusters

- Number of clusters has to be defined
- Each cluster has to be described with
  - Forecast error measures
  - Respect to periods before delivery
- Calculation of the difference between the forecast error measure from the product and the cluster for each period before delivery
- Squared errors have to be calculated for each cluster for every product → automation is necessary for many products
- Each product is assigned to the cluster which shows the least squared error (sum over all periods before delivery) in overall

Clustering with predefined clusters – Example 6

- Apply the least square approach with predefined clusters to cluster the products based on MPE with respect to periods before delivery
- Define clusters which show the following forecast behaviour
  - > Constant overbooking of 10%
  - > Constant underbooking of 20%
  - > No systematic behaviour
- Try to apply this clustering approach with the NRMSE (3 Clusters)
- MPE and NRMSE with respect to periods before delivery are already calculated

Clustering with predefined clusters – Example 6

– Solution (1/3)

> Predefining Clusters

		Input: MPE - periods before delivery									
		1	2	3	4	5	6	7	8	9	10
	Input Forecast error measurs here -->	0,059	0,060	0,076	0,067	0,069	0,056	0,044	0,047	0,064	0,077
Cluster number	**Cluster description**	**Predefined Cluster**									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	Constant overbooking of 10%	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1
2	Constant underbooking of 20%	-0,2	-0,2	-0,2	-0,2	-0,2	-0,2	-0,2	-0,2	-0,2	-0,2
3	No systematic behaviour	0	0	0	0	0	0	0	0	0	0

Clustering with predefined clusters – Example 6

- Solution (2/3)
  - Calculation of the least squared error for each product

Difference										
1	2	3	4	5	6	7	8	9	10	11
-0,041	-0,040	-0,024	-0,033	-0,031	-0,044	-0,056	-0,053	-0,036	-0,023	0,000
0,259	0,260	0,276	0,267	0,269	0,256	0,244	0,247	0,264	0,277	0,000
0,059	0,060	0,076	0,067	0,069	0,056	0,044	0,047	0,064	0,077	0,000

Squared difference (^2)											Sum squared errors
1	2	3	4	5	6	7	8	9	10	11	
0,002	0,002	0,001	0,001	0,001	0,002	0,003	0,003	0,001	0,001	0,000	0,016
0,067	0,067	0,076	0,071	0,072	0,066	0,059	0,061	0,070	0,077	0,000	0,687
0,003	0,004	0,006	0,005	0,005	0,003	0,002	0,002	0,004	0,006	0,000	0,039

Least squard error	0,016
Cluster number	1
Cluster description	Constant overbooking of 10%

Clustering with predefined clusters – Example 6

- Solution (3/3)

		MPE - periods before delivery									
Product	**Cluster description**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
61	Constant overbooking of 10%	0,059	0,060	0,076	0,067	0,069	0,056	0,044	0,047	0,064	0,077
62	Constant underbooking of 20%	-0,317	-0,327	-0,324	-0,328	-0,318	-0,277	-0,236	-0,189	-0,124	-0,089
63	Constant overbooking of 10%	0,051	0,067	0,083	0,081	0,090	0,099	0,108	0,108	0,125	0,125
64	No systematic behaviour	-0,007	-0,004	0,000	-0,006	0,017	0,023	0,025	0,030	0,040	0,038
65	No systematic behaviour	0,014	-0,018	-0,024	-0,031	-0,022	-0,012	-0,001	0,010	0,033	0,033
66	Constant overbooking of 10%	0,142	0,086	0,060	0,068	0,058	0,060	0,065	0,071	0,101	0,114
67	No systematic behaviour	0,035	0,022	0,012	-0,016	-0,017	-0,022	-0,034	-0,032	-0,015	-0,010
68	Constant underbooking of 20%	-0,257	-0,284	-0,290	-0,320	-0,280	-0,254	-0,203	-0,159	-0,107	-0,065

Qualitative clustering by Experts

- Clustering of products with similar behaviour by production or sales planners who have experience and market knowledge concerning the products
- Experts maybe know possible existing clusters
- Clustering of products based on
  - Forecasting method
  - Markets
  - Branches
  - Production technologies
- Prioritization of products
- Verification of clustering results from other approaches (algorithms)

Clustering according Forecast behaviour

- Clustering algorithms are necessary to carry out the classification of products and customers
- Different Forecast behaviours were described by different measures and can be clustered as follow:
  - Forecast quality: NRMSE
  - Systematic forecast behaviour: MPE, MAD
  - Cancellation of orders: Confirmation Safety
- For outliers and amount shifting behaviours in forecasts, clustering is difficult to apply

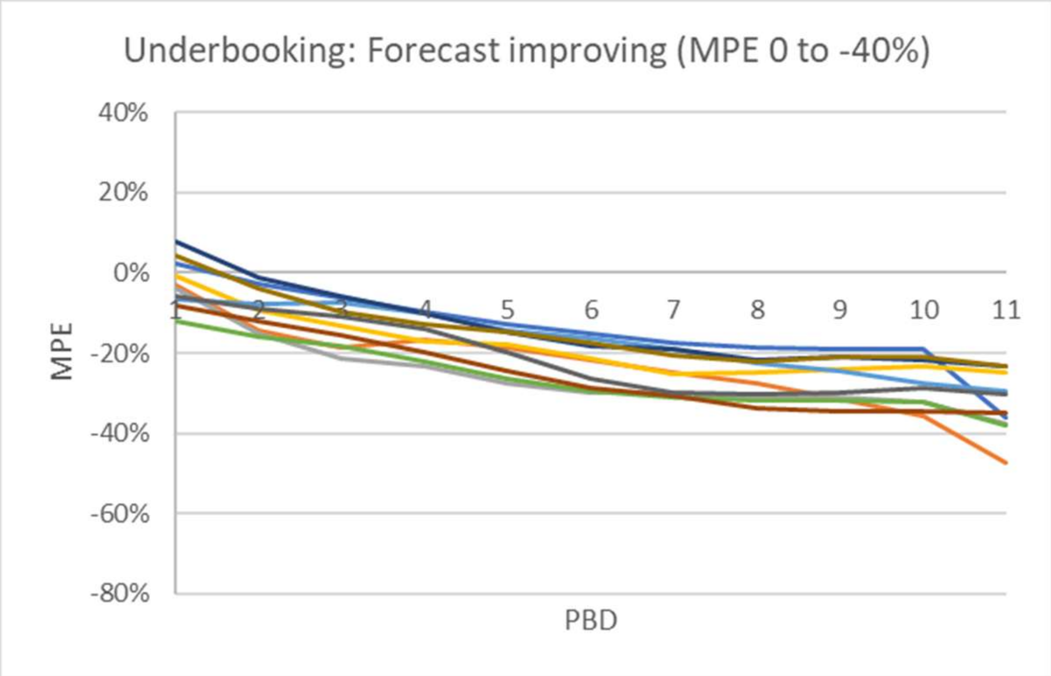
Clustering – Practical Example (1/2)

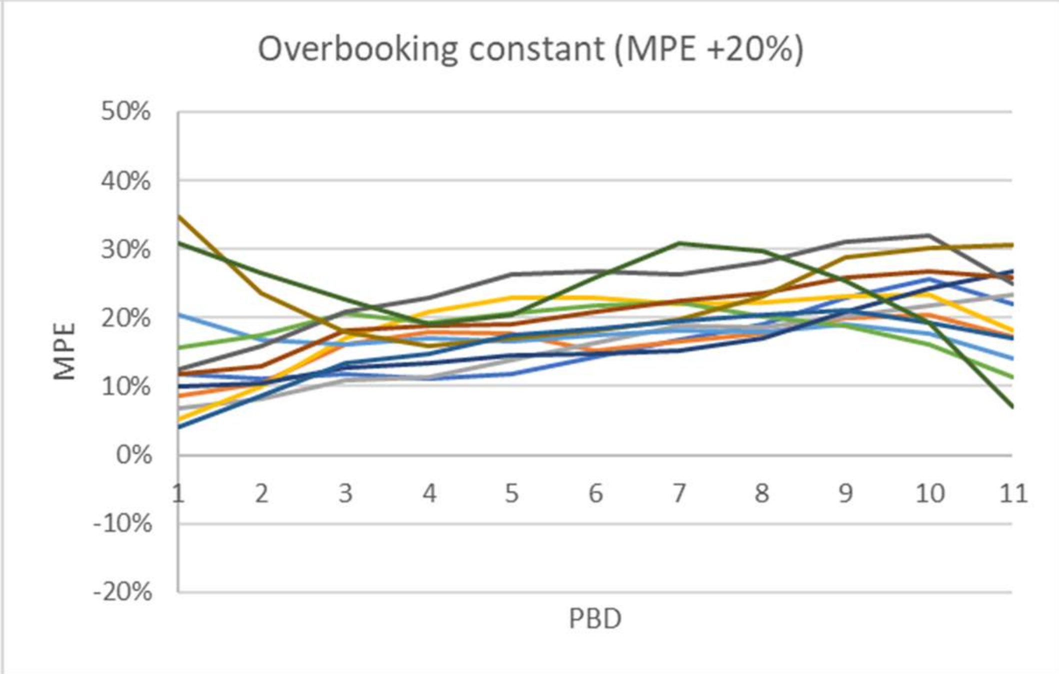
- Clustering with predefined clusters based on MPE
  - Number of clusters: 11
  - Forecasts stream with 11 periods before delivery
  - Excel Macro were used for automation of the calculation

Cluster number	Cluster description	Predefined Cluster										
		1	2	3	4	5	6	7	8	9	10	11
1	Overbooking constant (MPE +10%)	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1
2	Overbooking constant (MPE +20%)	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2
3	Overbooking constant (MPE +30%)	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3
4	Underbooking constant (MPE -10%)	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1
5	Underbooking constant (MPE -20%)	-0,2	-0,2	-0,2	-0,2	-0,2	-0,2	-0,2	-0,2	-0,2	-0,2	-0,2
6	Underbooking constant (MPE -30%)	-0,3	-0,3	-0,3	-0,3	-0,3	-0,3	-0,3	-0,3	-0,3	-0,3	-0,3
7	No systematic behaviour (MPE 0%)	0	0	0	0	0	0	0	0	0	0	0
8	Overbooking: Forecast improving (MPE 0 bis +40%)	0	0,04	0,08	0,12	0,16	0,2	0,24	0,28	0,32	0,36	0,4
9	Underbooking: Forecast improving (MPE 0 to -40%)	0	-0,04	-0,08	-0,12	-0,16	-0,2	-0,24	-0,28	-0,32	-0,36	-0,4
10	First overbooking, than underbooking (MPE -20% to +20%)	-0,2	-0,16	-0,12	-0,08	-0,04	0	0,04	0,08	0,12	0,16	0,2
11	First underbooking, than overbooking (MPE +20% to -20%)	0,2	0,16	0,12	0,08	0,04	0	-0,04	-0,08	-0,12	-0,16	-0,2

Clustering – Practical Example (2/2)

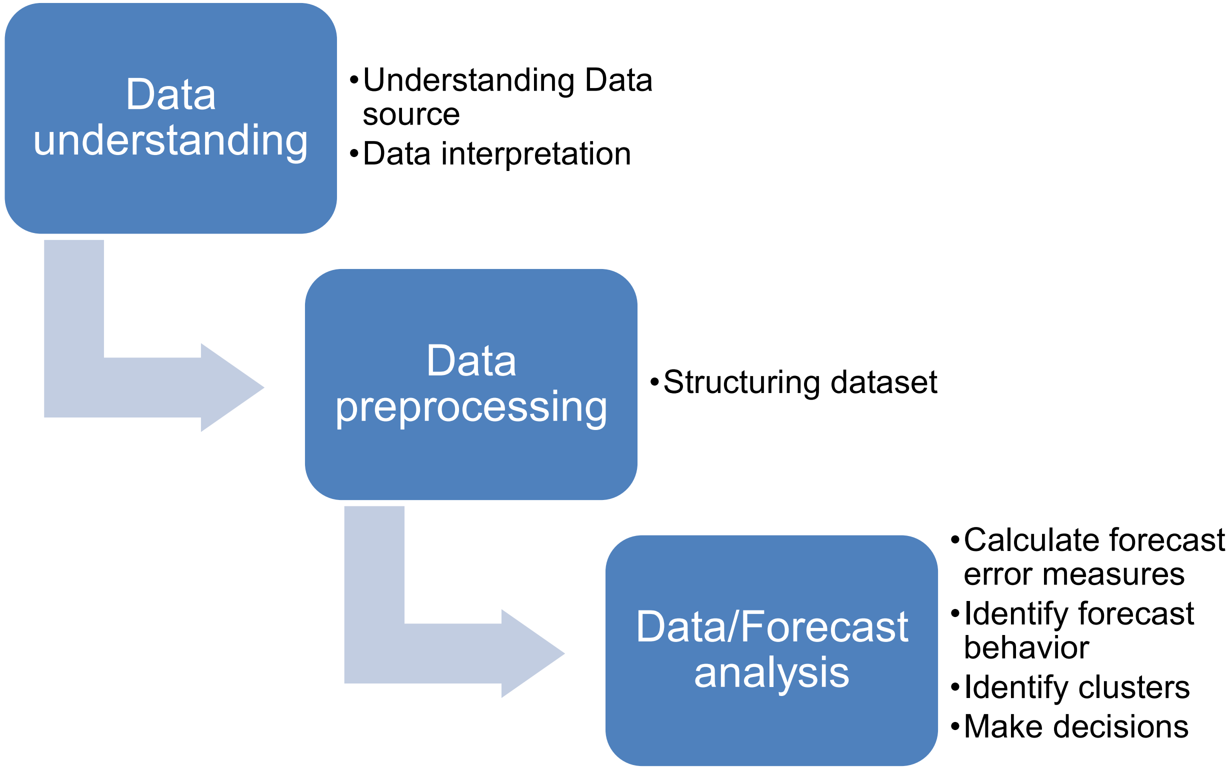
- Visualization of clusters





Decisions regarding forecast data

Process of forecast analysis



Forecast improvement

Forecast analysis

- Classification Matrix
- Outlier detection

Forecasting process

- Improve forecast generation process
- Usage of customer provided forecast data
- Apply different planning methods

Forecast improvement

Demand adaptation / Forecast correction

- Correction/adaptation of existing forecasts to improve forecast quality
- Apply outlier correction methods

Talk to customers

- Improve forecast quality by discussing the forecast behavior with customers

Classification Matrix

- Overview to classify products/customers concerning the forecast behaviour
- Structure:
  - Category (describes the forecast behaviour for the classification)
  - Criteria (figure/measure which describes the forecast behaviour)
  - Classification method (method/algorithm for classification within the category)
  - Classes/Cluster (describes the clusters within the category)

Classification Matrix

Category	Criteria	Classification method	Classes/Cluster
Forecast quality	NRMSE over periods before delivery	• Clustering algorithms • Clustering with predefined clusters • Based on threshold values and conditions	• Forecast evolution (forecast is improving till the due date) • Constant forecast quality (no improvement of forecast quality over periods before delivery)
Under/overbooking behaviour	MPE over periods before delivery		• Constant under/overbooking • No systematic behaviour • Under/overbooking with trend (improving forecast quality)
Cancelling of orders	Confirmation safety over periods before delivery		• No cancelling (final order will occur) • Cancelling is likely (delivery plan is cancelled at a specific probability)

Classification Matrix

Category	Criteria	Classification method	Classes/Cluster
Outliers in final orders	Final order amount	• Grubbs test • 1,5 * IQR-rule • 2-σ-method • Based on threshold values and conditions	• Outlier • No Outlier
Outliers in forecast stream	Forecast values → Changes of forecast values within the forecast stream		• Outlier • No Outlier
Amount shifting behaviour	Forecast values → Changes of forecast values from different forecast streams	• Variance-Covariance-Matrix* • Manually by experts (Sales department, Key account managers)	• Shifting behaviour (forward/backwards/both) • Total quantity/part of quantity • No shifting behaviour

**(More a theoretical approach and not the focus of the course)*

Self generated forecast vs. customer provided forecast

- Self generated forecasts → Forecast prediction (chapter forecast generation)
- Customer provided forecasts → Demand forecasts in a rolling horizon
- If NRMSE of a customer provide forecast is higher than a self generated forecast (e.g., moving average), there is no information advantage in using these customer provide forecast
  - > Applying the self generated forecasts would improve production system performance in that case
  - > For unsystematic forecast behaviour, customer provided forecasts mostly lead to higher forecast quality

Usage of customer provided forecast data

- Evaluate the quality of forecast data recording to the forecast error measures
- Analyze forecast behavior
  - Systematic/unsystematic errors
  - Amount shifting/cancelling
  - Outliers
- Quality of customer provided forecast data
  - Good → use it for PPC
  - Bad → apply forecast correction methods or search for other planning methods or inventory policies (MTS, MTA)

Improve forecasting process

- Usage of different data sources
  - Single orders
  - Frame contracts
  - Sales data
  - …
- Adjustment of long-term quantities (not forecast relevant orders in the past e.g. bargain sale, were not considered for future forecasts)
- Automatic outlier detection → production planner can focus on outliers in forecast streams and final orders
- Continuous monitoring and controlling of forecast error measures is necessary to held and increase forecast quality.

Change planning method

- Change the planning method or inventory policy if forecast quality is insufficient
- Application of MTA or MTS
- Usage of self generated forecasts (chapter forecast generation)
- Demand driven planning methods → no forecasts needed anymore
  - Reorder point policy

Application of MTA policy

- Preproduction is not customer-related
- Only the final assembly steps were scheduled with customer provide forecast data
- Forecasts for final orders have to be provided not too far in the future → better forecast quality near to the due date
- Depends on the production system and the products → MTA policy possible?
  - Number of Variants
  - Variant formation point
  - Product complexity

Reorder point policy (1/2)

- Useful for products with poor forecast quality and high demand variance
- Demand driven planning method → no forecasts needed anymore
- Reorder point should cover the demand during the replacement time
- Order is triggered if the inventory position come under the reorder point
  - Inventory Position = physical stock + open supplies (– reservations) – outstanding orders

Reorder point policy (2/2)

- Parameters
  - Reorder point
  - Lot size
- Important information needed
  - Replenishment lead time
  - Demand of the product
  - Variance of demand during the replacement time

Forecast correction

- Forecast correction based on forecast analysis
- Improve forecast quality by adapting forecasts using
  - historical data
  - customer order behavior
  - mathematical forecast correction methods

- Target: Improvement of forecast error measures (NRMSE, MPE) and production order accuracy

Forecast bias correction

- Forecast bias means systematic forecast behaviour (over/underbooking)
- Correction based on latest known MPE → rolling adaption of forecast streams
- Most simple forecast bias correction method → forecast is corrected automatically

$$\hat{X}_{i,j} = \frac{X_{i,j}}{1 + MPE_i}$$

$\hat{X}_{i,\,j}$ … $corrected\ forecast\ for\ period\ j\ predicted\ i\ periods\ before$

$X_{i,\,j}$ … $forecast\ for\ period\ j\ predicted\ i\ periods\ before$

$MPE_i(m)$ … $latest\ MPE\ for\ i\ periods\ before\ delivery\ over\ m\ historical\ forecasts$

Forecast correction – Example 7

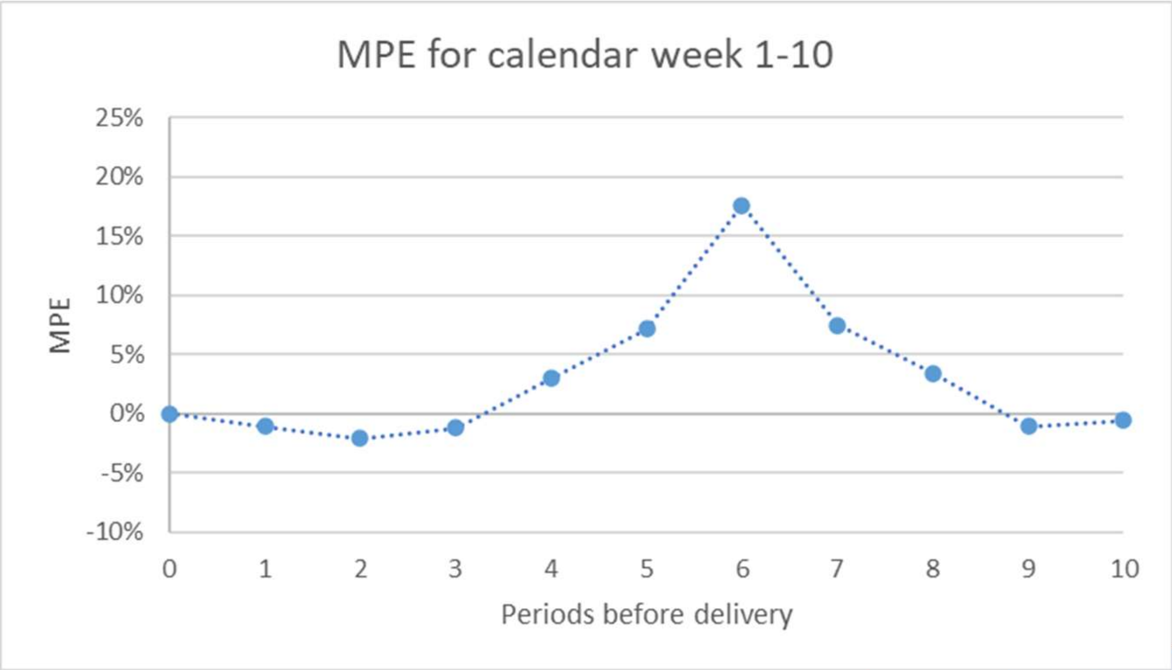
- Apply the forecast bias correction for the delivery plans for calendar weeks 11-14 based on the calculated MPE for the calendar weeks 1-10
- Assumption: you are in calendar week 10 and you can only apply the forecast bias correction on the known forecasts (for calculating MPE and NRMSE you have to use the "future" values)
- Calculate MPE and NRMSE separately for the following time horizons
  - Calendar week 1-10
  - Calendar week 11-14 without correction
  - Calendar week 11-14 with forecast bias correction
- Data set 2 → Product 60
- Forecasts for calendar weeks 11-14 (future) are stated in the Excel file

Forecast correction – Example 7

– Solution (1/3)

 > Calculate Forecast error measures for week 1-10

Product 60		MPE	RMSE	NRMSE
Periods before delivery	0	0%	0,00	0,00
	1	-1%	15,81	0,02
	2	-2%	27,81	0,03
	3	-1%	28,80	0,04
	4	3%	41,87	0,05
	5	7%	65,70	0,08
	6	18%	147,67	0,18
	7	7%	74,44	0,09
	8	3%	51,54	0,06
	9	-1%	45,94	0,06
	10	-1%	47,15	0,06



Forecast correction – Example 7

– Solution (2/3)
  > Apply forecast correction on week 11-14 (only known values
  > C56 → MPE for 1 PBD calculated for week 1-10

future values --> not known yet

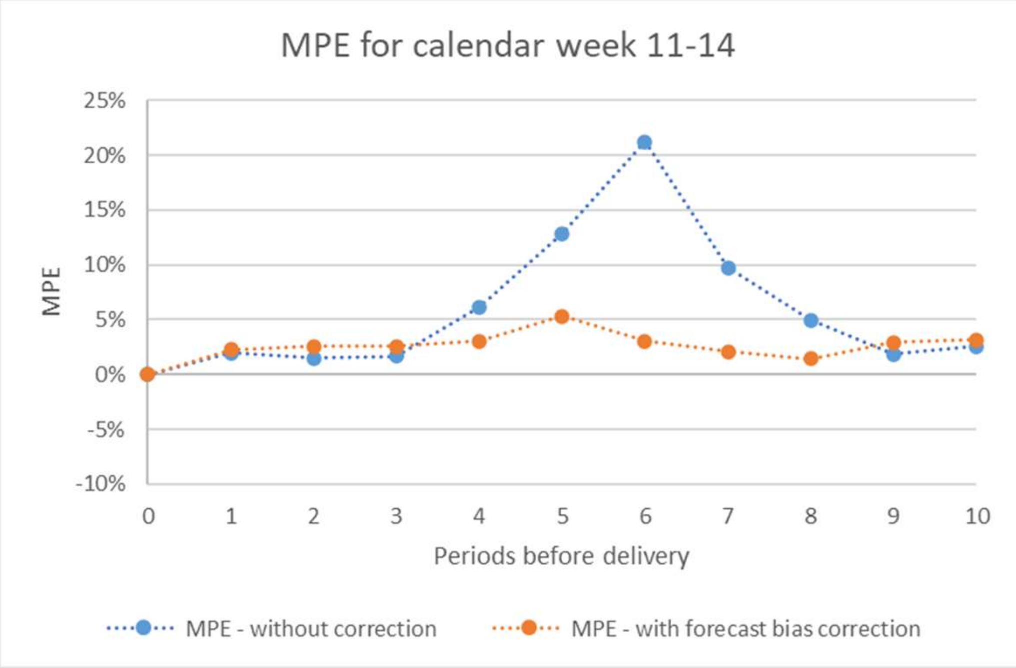
Product 60		delivery plans (calendar weeks)			
		11	12	13	14
Periods before delivery	0	798	727	682	905
	1	810	760	709	895
	2	813	757	706	882
	3	811	770	707	877
	4	830	826	760	887
	5	880	878	835	920
	6	949	949	917	958
	7	848	888	825	854
	8	798	859	789	820
	9	768	819	768	814
	10	772	813	791	817

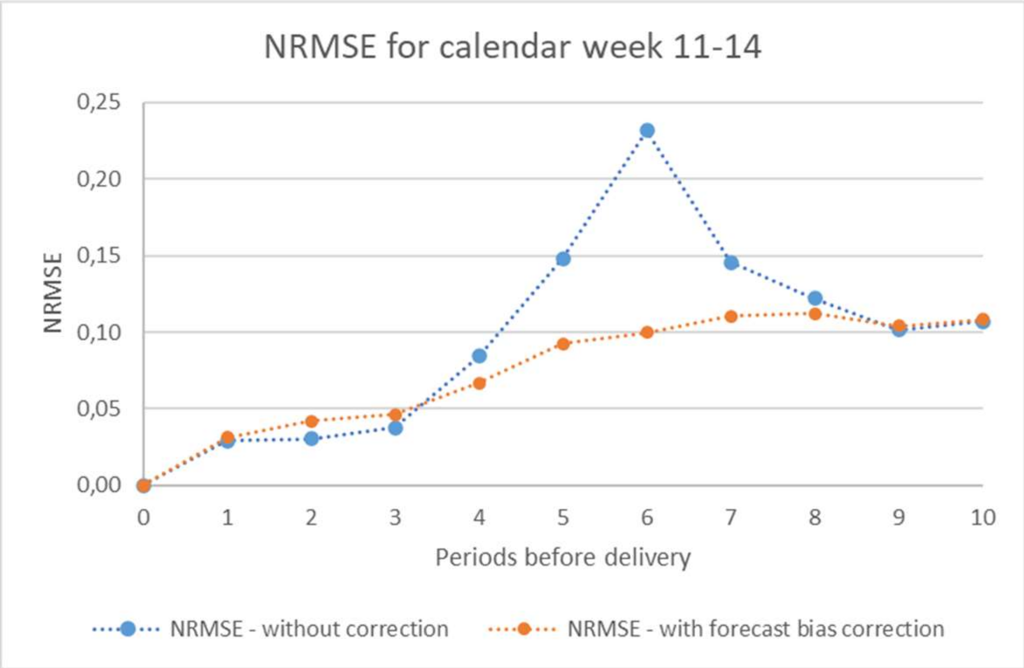
future values --> not known yet

delivery plans (calendar weeks) corrected			
11	12	13	14
		682	905
=Q11/(1+C56)		709	895
830	773	706	882
821	779	716	877
806	802	738	861
821	819	779	858
807	807	780	815
789	826	768	795
772	831	763	793
776	828	776	823
776	817	795	822

Forecast correction – Example 7

- Solution (3/3)
  - Calculate Forecast error measures for week 11-14
  - Without correction and with forecast bias correction





Outlier correction (1/3)

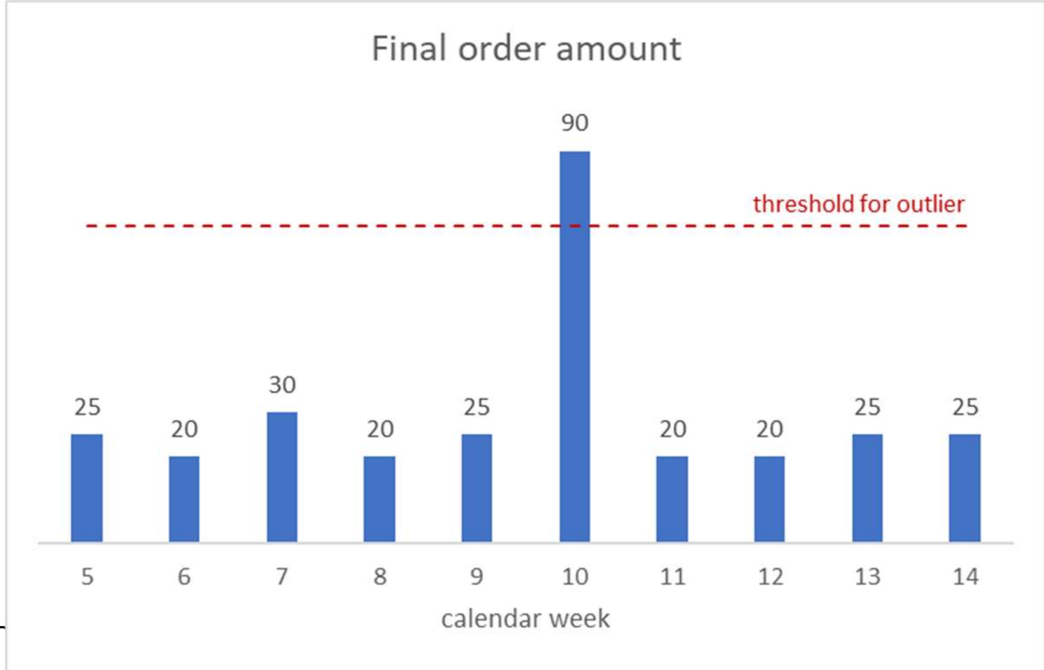
- 1st step: Outlier detection (chapter forecast behaviour analysis)
- 2nd step: Analyze outliers (where do the outliers come from?)
- 3rd step: Decision with which value the outlier should be corrected (threshold value, mean, no correction, determined value…)
- 4th step: Apply outlier correction to the dataset

- Target of outlier correction
  - Correct future demand data, to get a better forecast quality
  - Correct historical demand date, for generating better forecasts

Outlier correction (2/3)

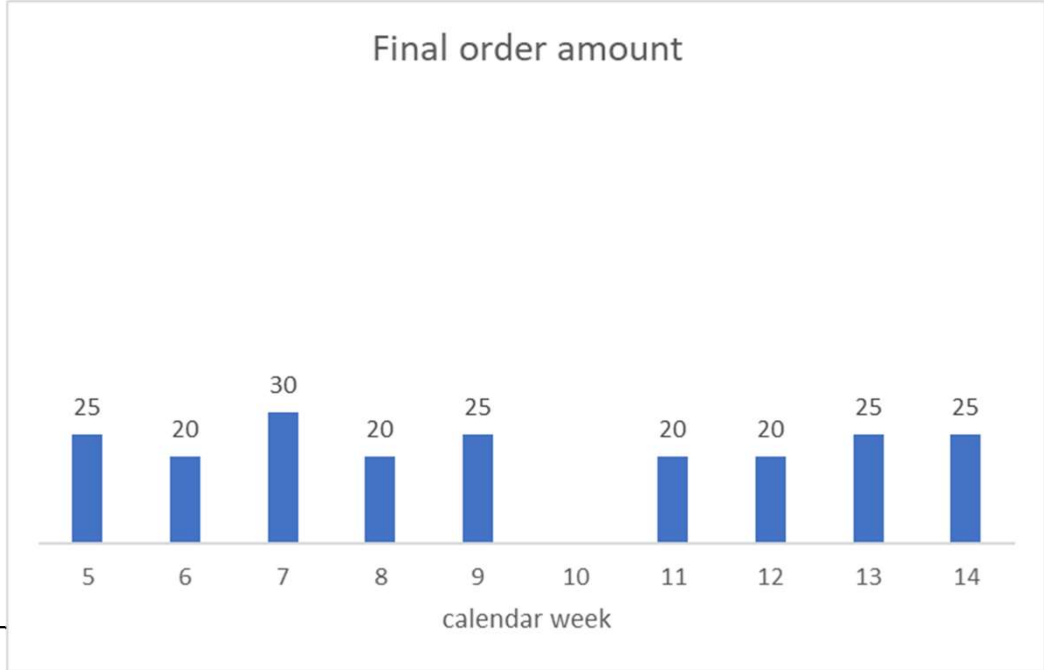
– Example:
  > Outlier in final order amount → outlier detection (1st step)
  > Analyze outlier → one-off effect (2nd step)
  > Moving average = 30

Average	30
Standard deviation	21
Number of standard deviations	2
Threshold value	73



Outlier correction (3/3)

- Example:
  - Remove outlier for forecast generation (moving average) → decision (3rd step)
  - Calculate new moving average for forecast → apply outlier correction (4th step)
  - Corrected moving average = 23
  - Quantity reduction by 23%



Talk to customers

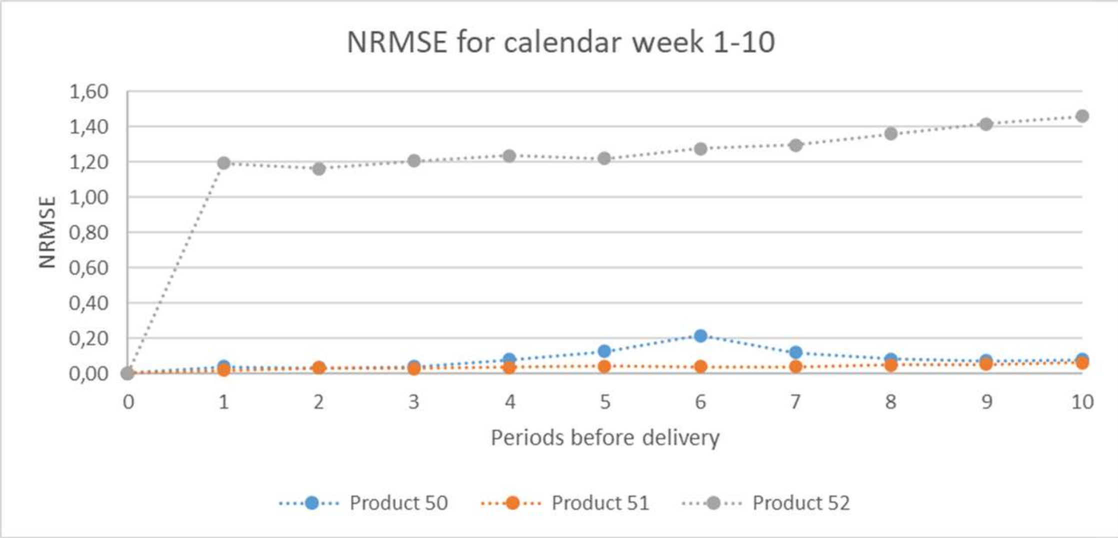
- Increase quality of customer provided forecast data by showing and discussing specific forecast behaviours with the customer
- Analyse and visualize the forecast behaviours with the mentioned tools in the course
- Arrange a frozen zone, so that the customer is not able to change the order amount within this time
- Share demand data, delivery plans were updated automatically

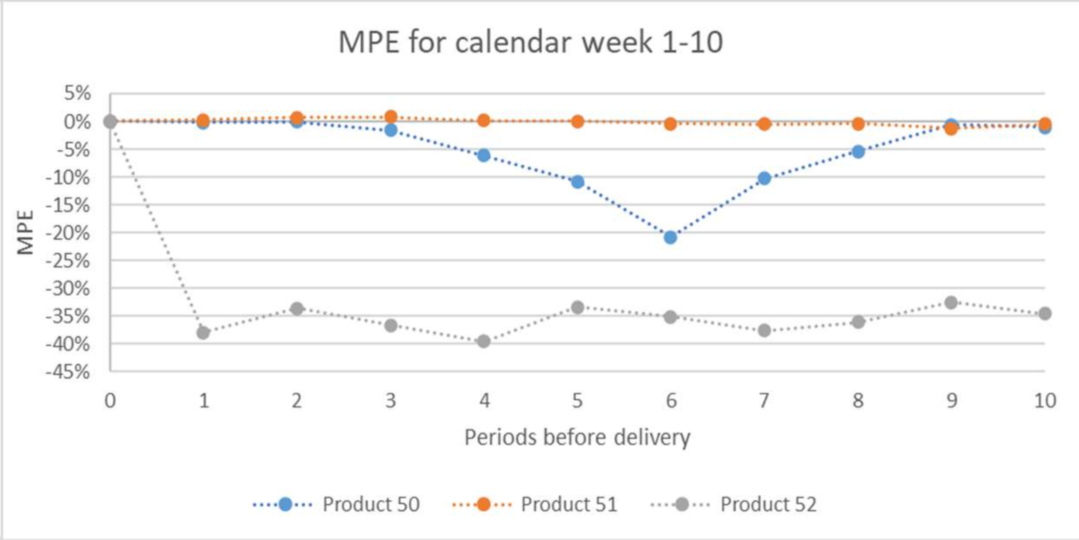
Case Study – Example 8

- Predict the final order amount (forecast value) for calendar week 11-14 in calendar week 10 for each product
- Target: Minimize the Percentage Error for each forecast value
  - > Take the latest forecast as forecast values for the product with the highest forecast quality
  - > Apply an outlier detection method to detect outlier/s (too high) in the final order amounts of one of the products → remove the outlier/s from the data set and use the average to generate the forecast values
  - > For the product which shows a systematic forecast behaviour apply forecast bias correction method on the latest forecast values to get a forecast value
  - > Interpretate the results

Case Study – Example 8

- Solution (1/3)





Case Study – Example 8

- Solution (2/3)
  - Product 50: systematic forecast behaviour → forecast bias correction is applied on latest forecast
  - Product 51: good forecast quality → latest forecast is taken
  - Product 52: outlier in final orders → average without outlier is taken as forecast

Case Study – Example 8

- Solution (3/3)

Predicted oder amounts in calendar week 10				
	11	12	13	14
Product 50	768	759	741	775
Product 51	1703	1806	1434	1727
Product 52	147	147	147	147

Final Order amount (future)				
	11	12	13	14
Product 50	780	751	767	751
Product 51	1664	1740	1514	1578
Product 52	360	195	252	59

Percentage Error				
	11	12	13	14
Product 50	-2%	1%	-3%	3%
Product 51	2%	4%	-5%	9%
Product 52	-59%	-25%	-42%	149%

Forecast improvement - Practical example (1/6)

- Target: Improve forecast quality by applying different forecast methods
- Calculation of NRMSE and MPE to analyse forecast quality and behaviour
- Clustering of products based on forecast error measures of original forecast
- Applying optimal forecast method on each products to improve NRMSE
- 15 products
- 3 forecast methods
  - Original Forecast based on moving average + manual adaptions (FC1)
  - Moving average + customer provided forecast information (FC2)
  - Moving average + customer provided forecast information + MPE correction (FC3)

Forecast improvement - Practical example (2/6)

- Forecast error measures were calculated for FC1 for 11 periods before delivery
- Clustering with predefined clusters were applied twice
  - > MPE Clustering
    - Strong/Light Underbooking (-0,3/-0,1 MPE)
    - No systematic Forecast behaviour (0,0 MPE)
    - Strong/Light Overbooking (0,3/0,1 MPE)
  - > NRMSE Clustering
    - Bad/Very bad Forecast quality (1,0/1,5 NRMSE)
    - Medium Forecast quality (0,5 NRMSE)

Forecast improvement - Practical example (3/6)

– Results FC1 (assorted ascending on Average MPE)

Product	Average MPE	Cluster MPE	Average NRMSE	Cluster NRMSE
1	-0,92	Strong Underbooking	1,2	Bad Forecast quality
2	-0,43	Strong Underbooking	0,8	Bad Forecast quality
3	-0,37	Strong Underbooking	0,9	Bad Forecast quality
4	-0,15	Light Underbooking	1,7	Very bad Forecast quality
5	-0,13	Light Underbooking	1,6	Very bad Forecast quality
6	-0,04	No systematic Forecast behaviour	0,6	Medium Forecast quality
7	-0,03	No systematic Forecast behaviour	0,7	Medium Forecast quality
8	-0,03	No systematic Forecast behaviour	0,5	Medium Forecast quality
9	-0,01	No systematic Forecast behaviour	0,7	Medium Forecast quality
10	-0,01	No systematic Forecast behaviour	0,6	Medium Forecast quality
11	0,00	No systematic Forecast behaviour	0,7	Medium Forecast quality
12	0,03	No systematic Forecast behaviour	0,7	Medium Forecast quality
13	0,07	Light Overbooking	1,7	Very bad Forecast quality
14	0,94	Strong Overbooking	1,3	Very bad Forecast quality
15	1,17	Strong Overbooking	1,7	Very bad Forecast quality

Forecast improvement - Practical example (4/6)

- Decisions regarding forecast analysis results
  - Bad/Very bad forecast quality → change forecast method (FC2 or FC3)
  - Good/Medium forecast quality → keep original forecast method (FC1)
  - Strong over/underbooking → apply MPE correction (FC3)
  - Light over/underbooking → no MPE correction necessary (FC2)

- MPE correction was only applied on FC2 (FC2 + MPE correction = FC3) because new forecast methods should not require the original forecast (FC1)

Forecast improvement - Practical example (5/6)

- Results after applying new forecast methods on some products

Product	Cluster MPE	Cluster NRMSE	Forecast method	Average MPE	Average NRMSE	Improvement NRMSE
1	Strong Underbooking	Bad Forecast quality	FC3	0,03	0,8	37%
2	Strong Underbooking	Bad Forecast quality	FC3	-0,14	0,6	15%
3	Strong Underbooking	Bad Forecast quality	FC3	-0,17	0,8	12%
4	Light Underbooking	Very bad Forecast quality	FC2	0,20	1,2	28%
5	Light Underbooking	Very bad Forecast quality	FC2	0,14	1,1	34%
6	No systematic Forecast behaviour	Medium Forecast quality	FC1	-0,04	0,6	0%
7	No systematic Forecast behaviour	Medium Forecast quality	FC1	-0,03	0,7	0%
8	No systematic Forecast behaviour	Medium Forecast quality	FC1	-0,03	0,5	0%
9	No systematic Forecast behaviour	Medium Forecast quality	FC1	-0,01	0,7	0%
10	No systematic Forecast behaviour	Medium Forecast quality	FC1	-0,01	0,6	0%
11	No systematic Forecast behaviour	Medium Forecast quality	FC1	0,00	0,7	0%
12	No systematic Forecast behaviour	Medium Forecast quality	FC1	0,03	0,7	0%
13	Light Overbooking	Very bad Forecast quality	FC2	0,26	1,2	30%
14	Strong Overbooking	Very bad Forecast quality	FC3	-0,12	1,1	17%
15	Strong Overbooking	Very bad Forecast quality	FC3	-0,04	1,0	40%

Forecast improvement - Practical example (6/6)

- Summary
  - Forecast quality (NRMSE) has improved by 14% on average by using different forecast methods
  - No forecast method is suitable for every product
  - Decisions can be applied on whole product groups
  - Systematic forecast behavior should be corrected
  - Automatic generated forecasts including customer provided information (+correction methods) show a high potential

Practical experience

Data preparation

- Can be time-consuming
- Understand the data structure first
- Datasets can be incomplete, or data is missing → full forecast streams are mandatory for good analysis results
- Focus on the most important product groups/customers (ABC-analysis in advance)
- Try to implement the data preparation in the data query, e.g., MS Excel queries or MS Access

Data preparation in MS Access

- Generate a database to manipulate the data set with MS Access queries
- Prepare the data set for the suitable format (Web tool format)
- Define the conditions in the "Design mode" by generating a query (see screenshots). Run the query and save the file.

Field:	Product	ActualDate	ForecastDate	OrderAmount	PBD
Table:	24-11	24-11	24-11	24-11	24-11
Sort:	Ascending		Ascending		Ascending
Show:	☑	☑	☑	☑	☑
Criteria:					<11 And >=0
or:					

Data preparation in Excel

- Excel algorithm for data preprocessing to simplify further calculations
  - No empty data entries
  - No negative values for order amount
  - Full forecast streams (for a certain period, e.g., week or month and PBD, no forecast dates are missing)
  - Every product has the same number of periods before delivery (PBD)
  - After preparation, data set can be used for further calculations and analysis

Forecast error measures

- Usage of two forecast error measures:
  - > NRMSE for forecast quality
  - > MPE for systematic forecast behaviour
- Calculate other statistical measures like
  - > Mean of final orders
  - > Variation coefficient of final orders
- Confirmation safety is useful to analyse different customers which often cancel their orders in advance

NKE forecasting Report (1/6)

- Data query from ERP System → import in Excel Report
- Overall view and article view
- Includes:
  - Final order amounts from the last year
  - Outliers in final orders and impact on yearly average
  - Forecast error measures (MPE, NRMSE)
  - Statistical measures
  - Previous forecasts for the future
  - Future order amounts from different data sources
  - Suggestion for new forecast

NKE forecasting Report (2/6)

- Overall view
  - Selection of date and number of standard deviations for outlier threshold value
  - Link to article view → selection of one article number
  - Gives a quick overview about all articles in the report

Analyse Forecastqualität der vergangenen 12 Monate

Datum	2020-04-30
Stabw Ausreißer	2,0

Artikel für Artikelübersicht	13316
	zur Artikelübersicht

Artikel	Mittelwert Abrufmenge	Streuung Abrufmenge	Variationskoeffizient Abrufmenge	Anzahl Ausreißer Final Order Amount	Höherer Mittelwert durch Ausreißer in %	NRMSE	MPE
11121	49,5	53,2	1,07	1	35%	1,14	41,8%
13316	631,1	514,0	0,81	1	19%	0,82	-24,6%
18928	26,2	26,3	1,00	0	0%	1,11	23,5%
71018	78,5	59,2	0,75	0	0%	0,79	-14,1%

NKE forecasting Report (3/6)

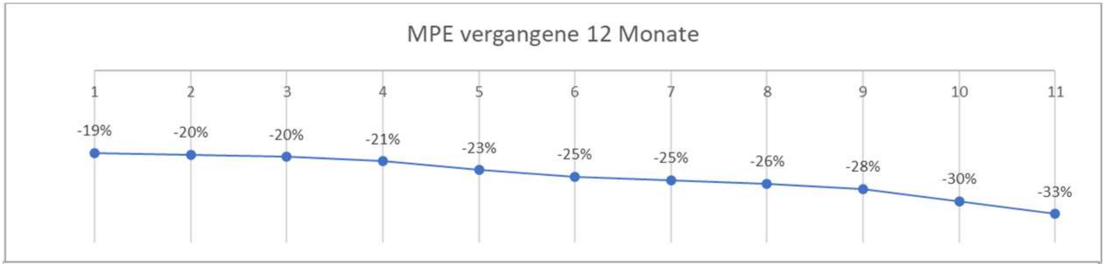
– Article view (1/4)

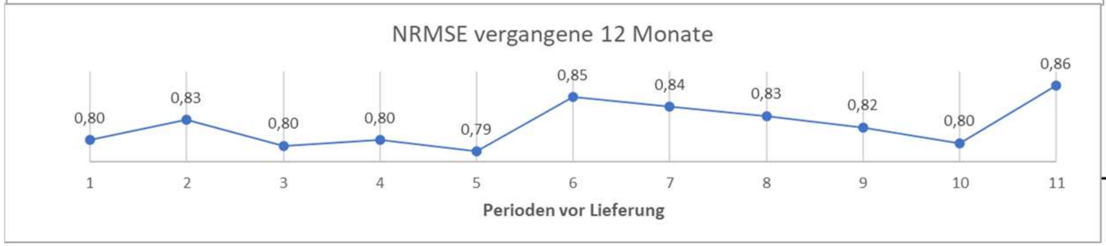
Datum	2020-04-30
Grenzwert Ausreißer	2,0

Artikel	13316
zur Gesamtübersicht	

Artikel	Mittelwert	Streuung	Variationskoeffizient	Anzahl Ausreißer Final Order Amount	Höherer Mittelwert durch Ausreißer in %	NRMSE	MPE
13316	631,1	514,0	0,81	1	19%	0,821	-24,6%

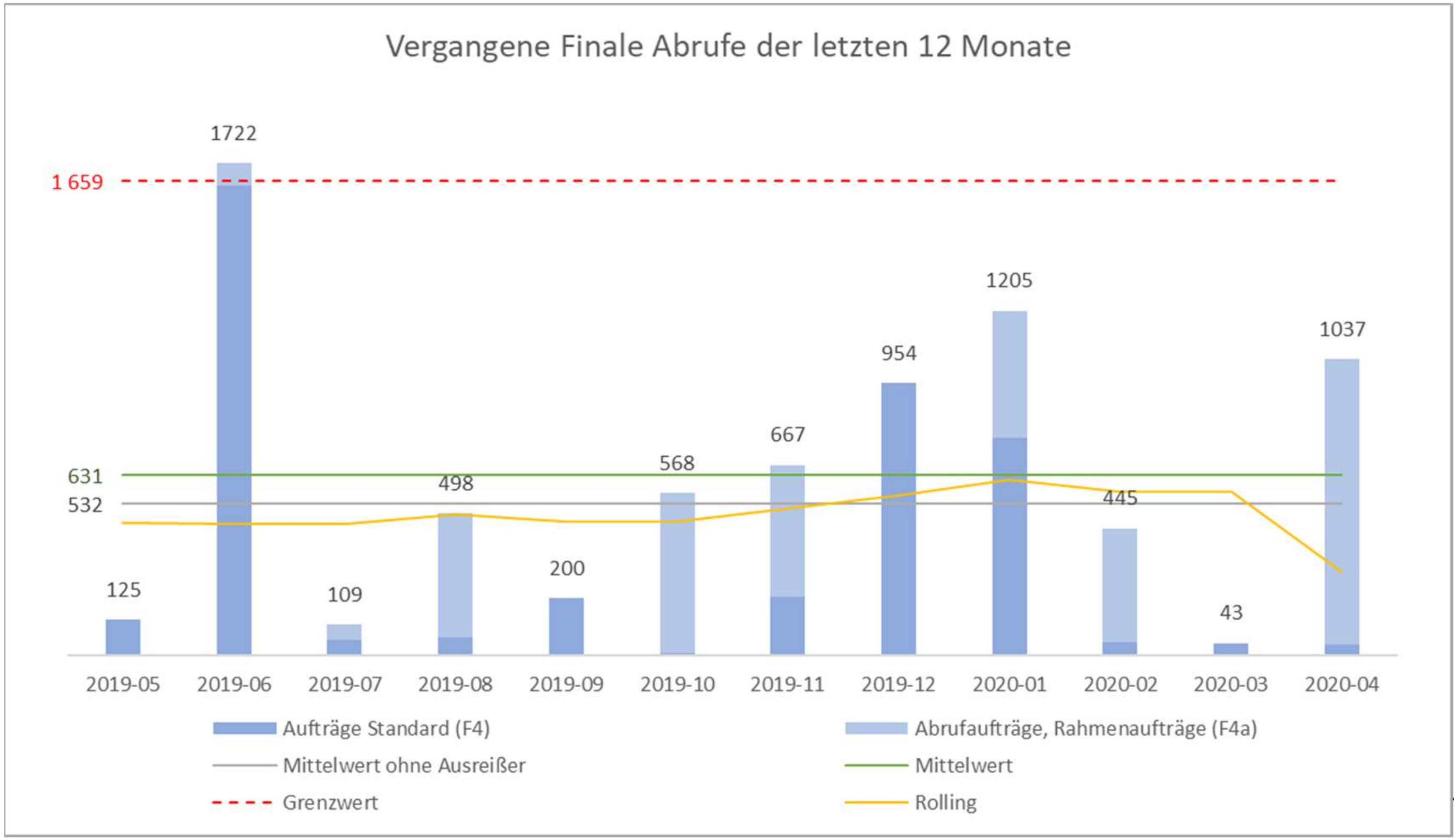
	pro Monat	pro Jahr
Planbedarf anhand Mittelwert ohne Ausreißer	**532**	**6 383**
Planbedarf anhand Mittelwert inkl. Ausreißer	**631**	**7 573**





NKE forecasting Report (4/6)

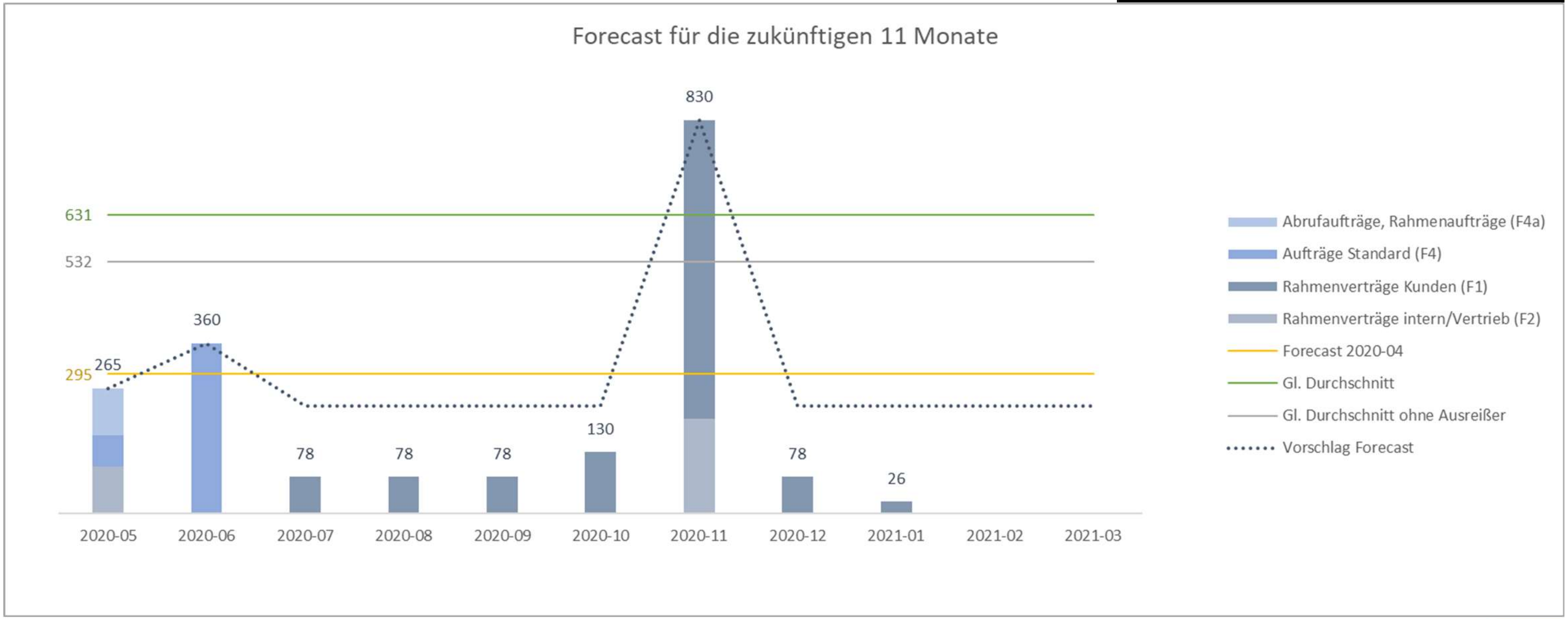
- Article view (2/4)



NKE forecasting Report (5/6)

– Article view (3/4)

Vorschalg Forecast	
Ausreißer bereinigen	Ja
Rahmen bereinigen	Ja
Mindestwert Forecast	0



NKE forecasting Report (6/6)

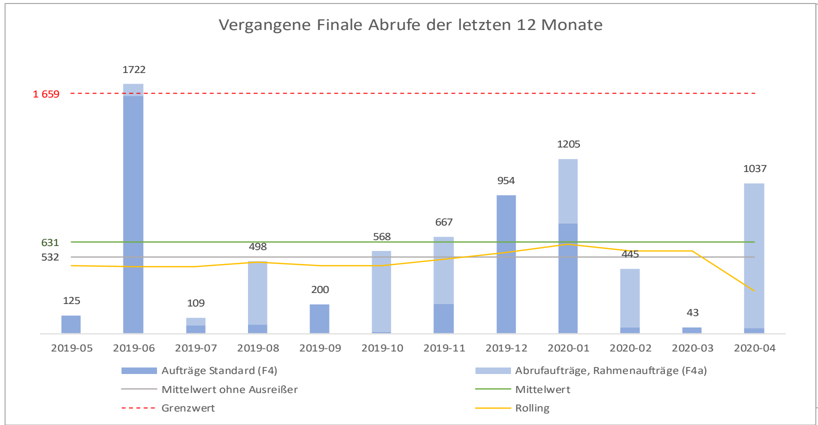
– Article view (4/4)

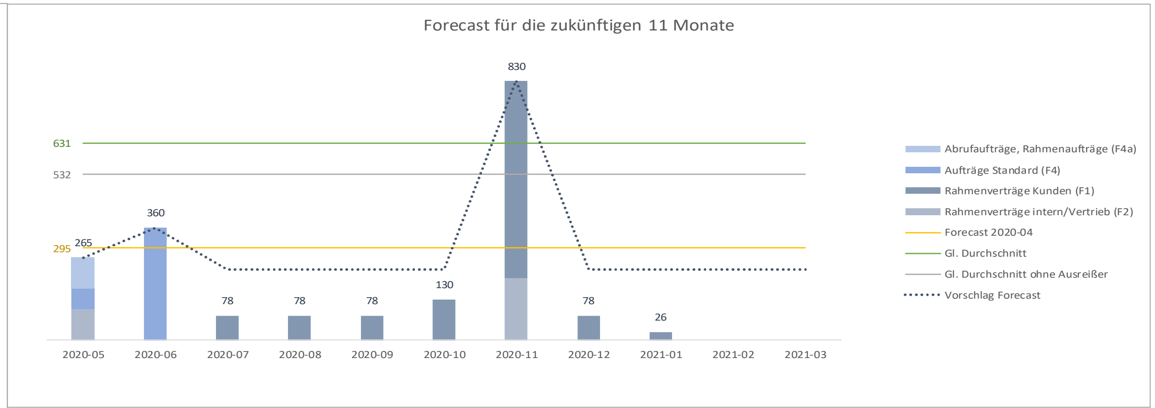
Datum	2020-04-30
Grenzwert Ausreißer	2,0

Artikel	13316
	zur Gesamtübersicht

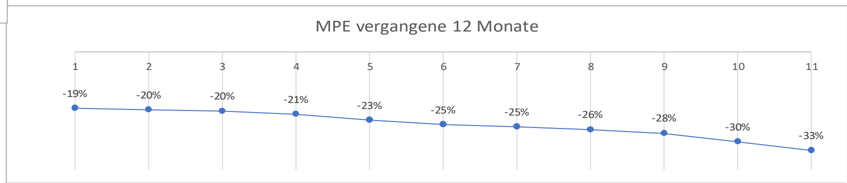
Vorschalg Forecast	
Ausreißer bereinigen	Ja
Rahmen bereinigen	Ja
Mindestwert Forecast	0

Artikel	Mittelwert	Streuung	Variationskoeffizient	Anzahl Ausreißer Final Order Amount	Höherer Mittelwert durch Ausreißer in %	NRMSE	MPE
13316	631,1	514,0	0,81	1	19%	0,821	-24,6%





	pro Monat	pro Jahr
Planbedarf anhand Mittelwert ohne Ausreißer	532	6 383
Planbedarf anhand Mittelwert inkl. Ausreißer	631	7 573



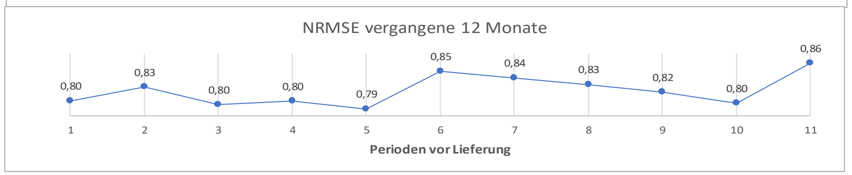
Nützliche Informationen:

NRMSE verbessert sich zum Lieferabruf → Indiz für funktionierenden Forecastprozess

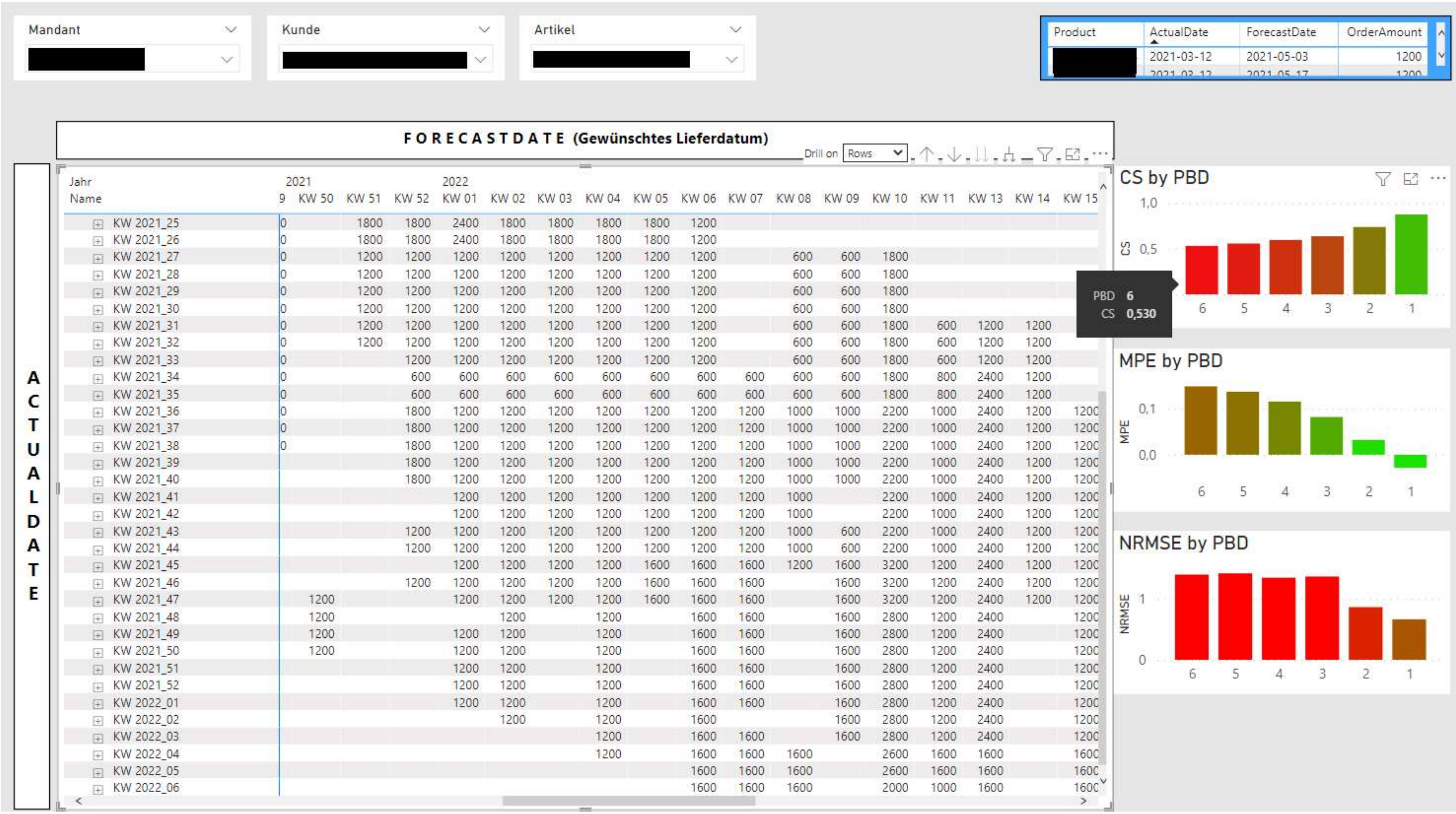
NRMSE größer als 1,0 → Forecast sehr schlecht

MPE negativ → Forecast geringer als Abrufmenge

MPE positiv → Forecast höher als Abrufmenge



Dashboard for forecast quality in MS Power BI



Clustering

- Decide first, which forecast behaviour should be used for clustering and how to use the results
- Calculate required forecast error measures in advance (use Web Tool for calculation and export the data)
- Visualize the created clusters over periods before delivery
- Try to standardize and automate the process from data pre-processing to clustering
- Web Tool is not designed for large data sets (more than 1MB) and many products (>100 products)
- Apart from the web tool, clustering with many different products can be carried out in Python or other programs

Improving forecasting process

- First, understand the existing forecasting processes
- There is no one fits all solution
- Use forecast error measures and clustering methods to identify how good/bad the forecasts are
- Focus on the most important products/customers (ABC-Analysis)
- Use clustering methods to identify forecast behaviour groups
- Based on the identified forecast behaviour groups, define suitable approaches to tackle the behaviour with inventory models and production strategies, e.g., MTS, MTO, MTA, etc.

Webtool tutorials

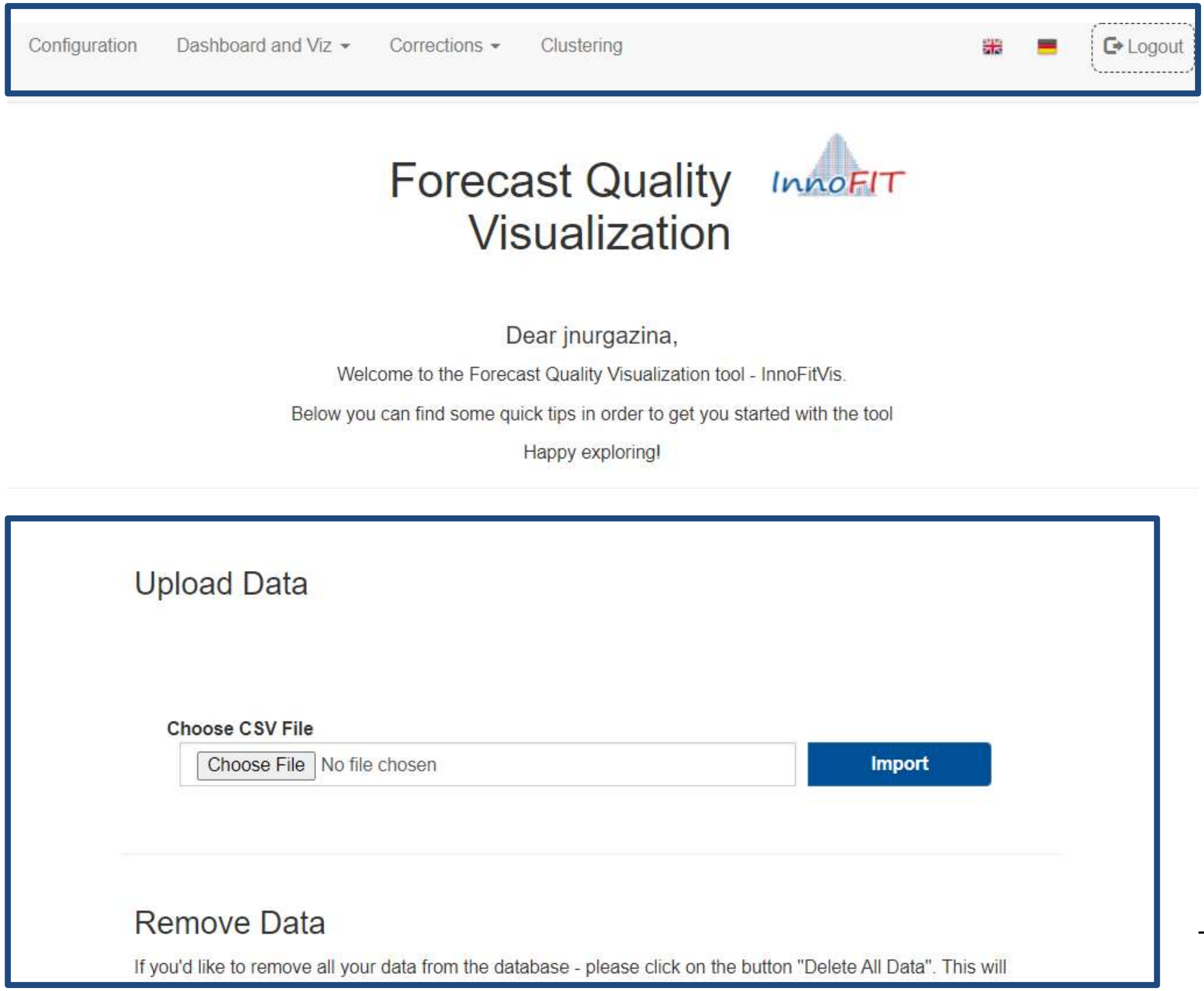
Web Tool

- Web tool overview
- Data structure and data preparation
- Tutorial on classical order analysis
- Tutorial error measures with respect to PBD
- Tutorial on clustering

Web Tool Overview

- Access link: http://innofitvis.fhstp.ac.at/
- Limitation of the data storage: 1 MB per user
- **Unrestricted home page** (with no login credentials required) is available with some details: project information, link to Leitfaden and contact information
- **Functionality**:
  - (1) basic order analysis
  - (2) dashboard
  - (3) forecast error measures
  - (4) clustering
  - (5) filtering on the *Configuration* page
  - (6) uploading or deleting data

Web Tool Overview



Main top navigation bar

Data length overview can be seen and reset on *Configuration* page

All available visualizations on *Dashboard and Viz* menu

Clustering and *Corrections* pages

Section to upload or remove data

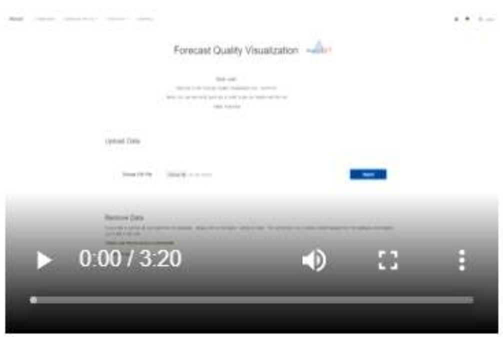
Web Tool Overview

Data format:

Please upload your data in .csv (comma-separated value) format in the defined data structure, please keep comma **(,)** as a delimiter. The required data structure is shown in the image below. Set the dates (ActualDate and ForecastDate) format as: **YYYY-MM-DD (year-month-day)** and keep the format as **a date**. Please keep OrderAmount values as **integers**.

The detailed description of data format requirements can be found here: Data Format Requirements

A quick guide on how to use the tool:



The format of the table headings and structure:

Intepret Error Measures:

The guidelines on how to interpret forecast error measures visualized in this tool can be found here (you will be redirected to another page): **How to Interpret Error Measures**

Section for introducing data format requirements

Data format guidelines on *Data Format Requirements* page

Introduction video

Section for details about forecast error measures

FH OBERÖSTERREICH InnoFIT fh/// st.pölten

Data structure and preparation (1/2)

- Web tool is a web interface (website) for visualizing customer-provided historical forecast data as well as fictive (and simulation) data
- Data should be provided in a .**CSV** format in a tabular form:
  - Column names should be defined exactly as shown here

Product	ActualDate	ForecastDate	OrderAmount
Up to 35 characters (integers and letters)	yyyy-mm-dd (year-month-day)	yyyy-mm-dd (year-month-day)	Integer

```
Product,ActualDate,ForecastDate,OrderAmount
Product1,2019-06-10,2019-06-10,90
Product1,2019-06-10,2019-06-17,0
Product1,2019-06-10,2019-06-24,108
Product1,2019-06-10,2019-07-01,108
Product1,2019-06-10,2019-07-08,108
Product1,2019-06-10,2019-07-15,126
Product1,2019-06-10,2019-07-22,0
Product1,2019-06-10,2019-07-29,234
Product1,2019-06-10,2019-08-05,90
Product1,2019-06-10,2019-08-12,0
```

Data structure and preparation (2/2)

- While preparing the data in a suitable format, please consider the following aspects during preparation:
  - Use Chrome, Brave or Firefox web browsers for better visualization quality (Internet Explorer: no)
  - Do not overwrite data (no duplicates)
  - Date format as YYYY-MM-DD (year-month-day)
  - No spacing, empty cells or empty rows
  - No negative OrderAmount values
  - Data upload limit: 1MB per user
  - You will need credentials to start using the web tool

  If you‘re planning to use clustering please additionally consider the following aspects:
  - For each Product & ForecastDate combination there must be a final order defined
  - For each Product & ForecastDate combination there should be equal number of periods before delivery (PBD) (PBD = Forecast date – Actual date)

Web Tool: Tutorial

- Web tool has the following functionalities (modules):
  - Basic order analysis
  - Dashboard and Forecast error measures
  - Clustering
- Please use **configuration page** to filter for: Product, ActualDate or ForecastDate selection
- Please refer to page „**How to interpret Error Measures**“ on the home page for explanation of visualized forecast error measures and how to interpret them

How to Interpret Error Measures

Web Tool: Tutorial on Basic order analysis

- Open the web tool: https://innofitvis.fhstp.ac.at
- Log in using the credentials: username: `tutorial`; password: `Mpass12345!`
- Select the "*Dashboard and Viz*" on the top navigation bar
- Basic order analysis includes the following visualizations:
  - Final order amount
    - The graph visualizes final orders only (as a scatterplot visualization). X-axis represents due date and Y-axis represents order amount
  - Delivery plans
    - The graph visualizes forecasts and final orders (final orders are shown as blue stars).
  - Delivery plans matrix
    - The same as delivery plans but in a matrix form (chromatic composition matrix)
  - Percentage error
  - Delivery plans matrix with percentage error

Web Tool: Tutorial on Basic order analysis

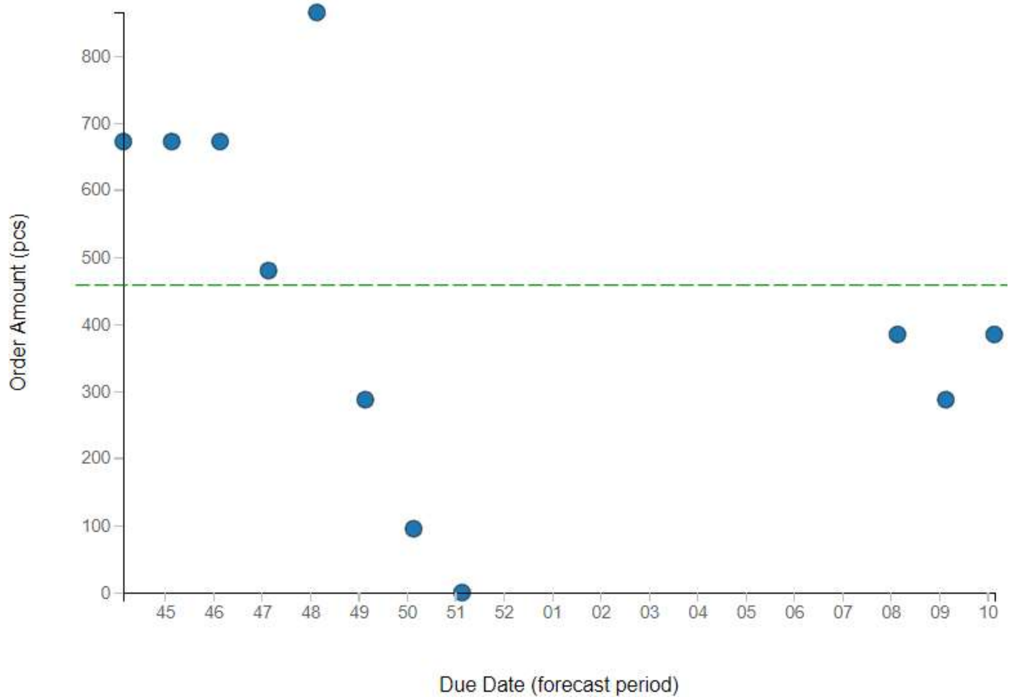
- To start the tutorial, click on the "*Configuration*" on the top navigation bar and select the following products:
- ern_1
- ern_2
- ern_3
- Analyze the products together and individually by using the configuration page (Product filter)
- Prepare the summaries of the analysis. What conclusions can be elaborated?
- Please do not remove or upload data

Visualizations in the basic order analysis:

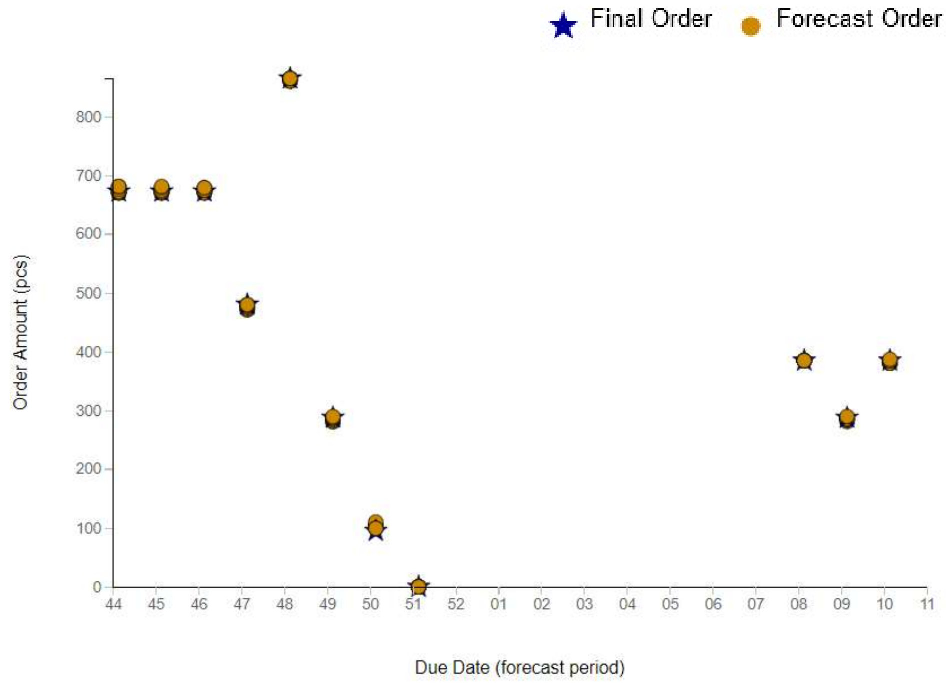
- Final order amount (scatterplot)
  - screenshot
- Delivery plans (scatterplot)
  - screenshot
- Delivery plans matrix
  - screenshot
- Percentage error (scatterplot)
- Delivery plans matrix with percentage error
  - screenshot

Tutorial on Basic order analysis: visualizations

Final order amount (scatterplot)

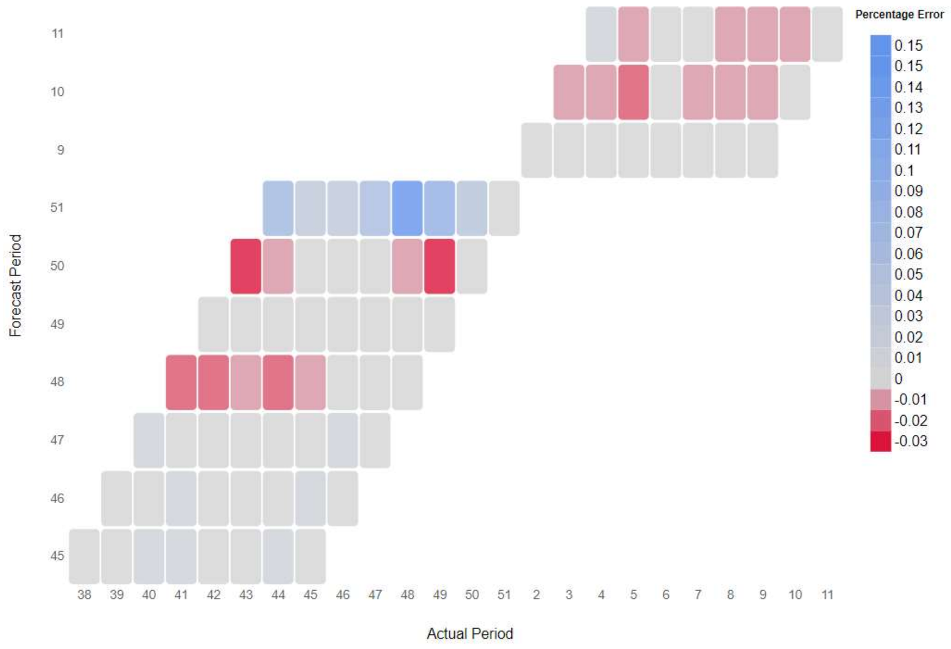


Delivery plans (scatterplot)

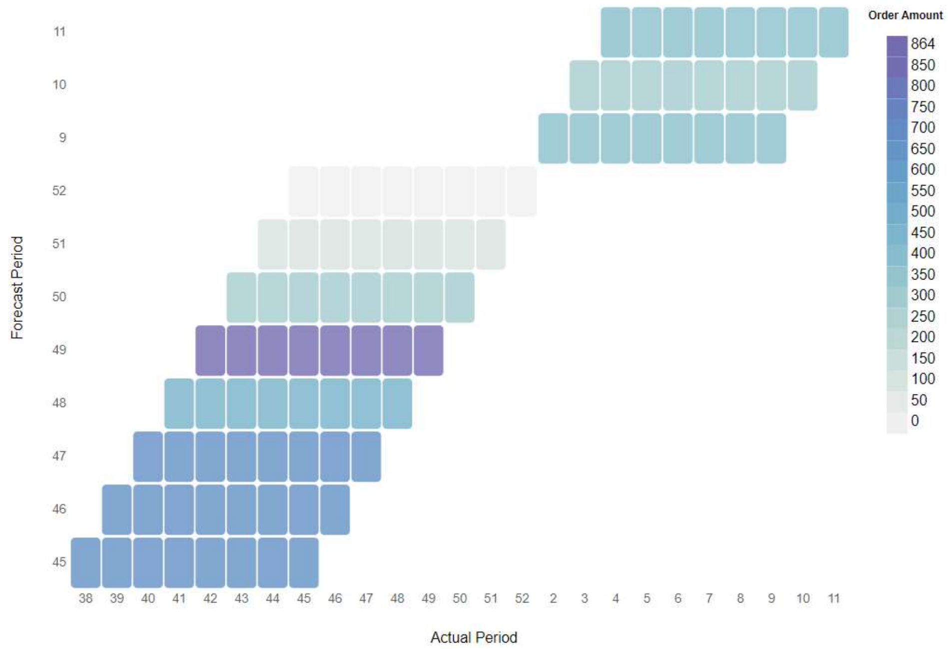


Tutorial on Basic order analysis: visualizations

Delivery plans matrix with percentage error



Delivery plans matrix



Web Tool: Tutorial on error measures

- The tutorial on the error measures using fictive data set:
  - Analyze data of three products: ern_1, ern_2 and ern_3
  - Detect over- and underbooking behavior using MPE
  - Forecast accuracy should be measured using NRMSE

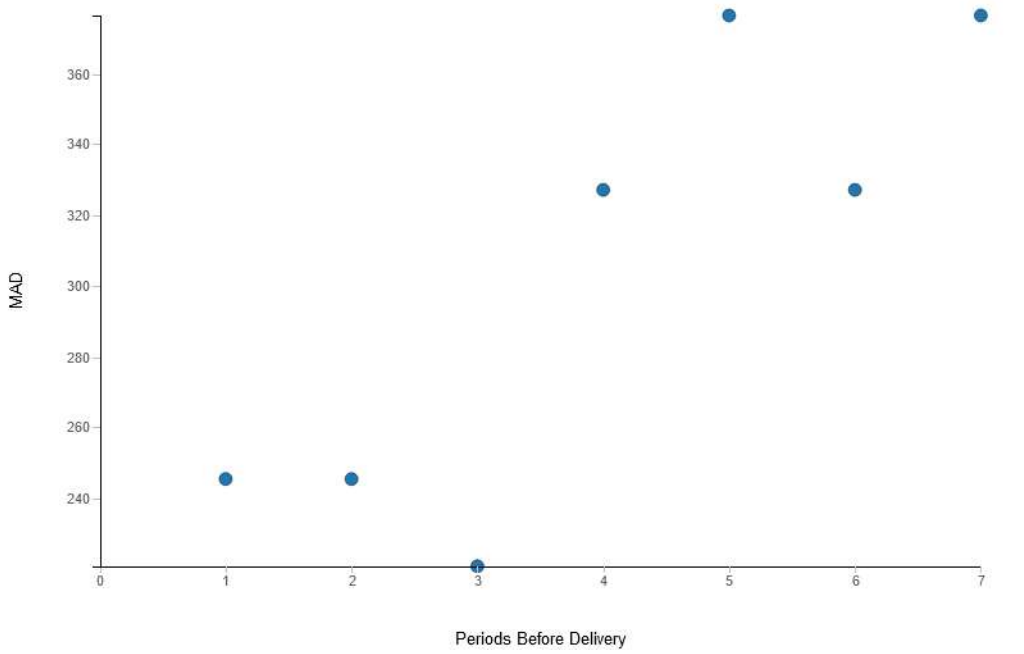
The web tool provides an overview of 8 forecast error measures. Each of the error measure has different approach and use.

Web Tool: Forecast Error Measures

– ern_1. Mean Absolute Deviation (or MAD):

Forecast Error Measures

- **Mean Absolute Deviation (MAD)**
- Mean Deviation (MD)
- Mean Square Error (MSE)
- Root Mean Square Error (RMSE)
- Normalized Root Mean Square Error (RMSE*)
- Mean Percentage Error (MPE)
- Mean Absolute Percentage Error (MAPE)
- Mean Forecast Bias (MFB)



$$MAD_j = \frac{1}{n}\sum_{i=1}^{n} |x_{i,j} - x_{i,0}|$$

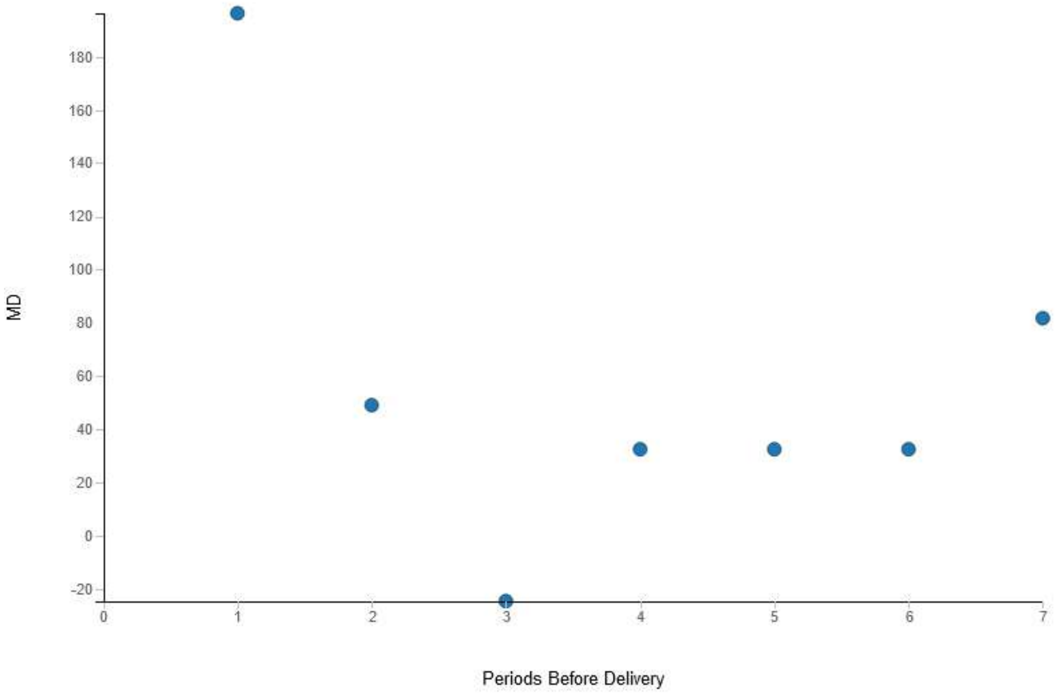
Where forecasts are $\mathbf{x_{i,j}}$, and final orders are $\mathbf{x_{i,0}}$

MAD represents the average absolute error of forecast deviation from the final order. Since the estimation is in absolute values, **it does not show whether there was over- or under-booking in the forecast.**

Web Tool: Forecast Error Measures

Forecast Error Measures
- Mean Absolute Deviation (MAD)
- Mean Deviation (MD)
- Mean Square Error (MSE)
- Root Mean Square Error (RMSE)
- Normalized Root Mean Square Error (RMSE*)
- Mean Percentage Error (MPE)
- Mean Absolute Percentage Error (MAPE)
- Mean Forecast Bias (MFB)

– ern_1. Mean deviation (MD)



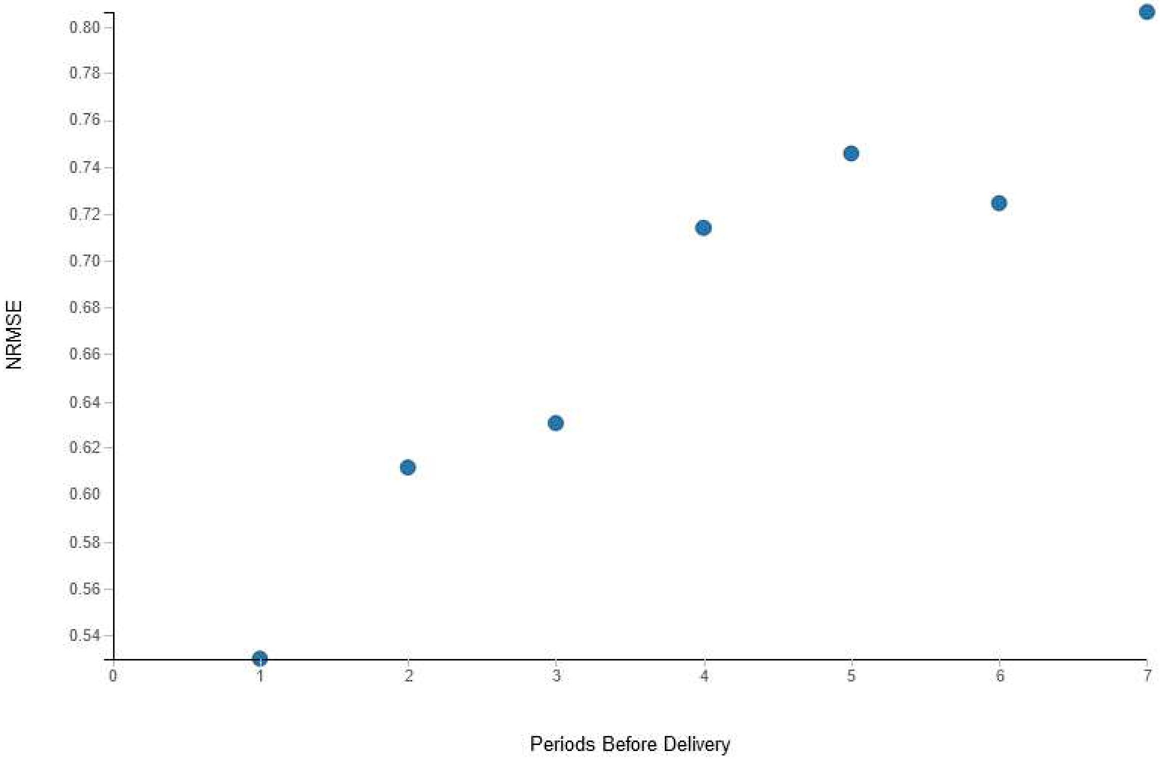
$$MD_j = \frac{1}{n}\sum_{i=1}^{n}(x_{i,j} - x_{i,0})$$

Where forecasts are $\mathbf{x_{i,j}}$, and final orders are $\mathbf{x_{i,0}}$

MD represents the average error of forecast deviation from the final order. Unlike MAD, **it can show whether there was over- or under-booking in the forecast.** Negative value represents underbooking (when final orders are bigger than forecasts)

Web Tool: Forecast Error Measures

- ern_1. Normalized RMSE (NRMSE or RMSE*):



Forecast Error Measures
Mean Absolute Deviation (MAD)
Mean Deviation (MD)
Mean Square Error (MSE)
Root Mean Square Error (RMSE)
Normalized Root Mean Square Error (RMSE*)
Mean Percentage Error (MPE)
Mean Absolute Percentage Error (MAPE)
Mean Forecast Bias (MFB)

$$RMSE*_j = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{i,j} - x_{i,0})^2}}{\frac{1}{n}\sum_{i=1}^{n} x_{i,0}}$$

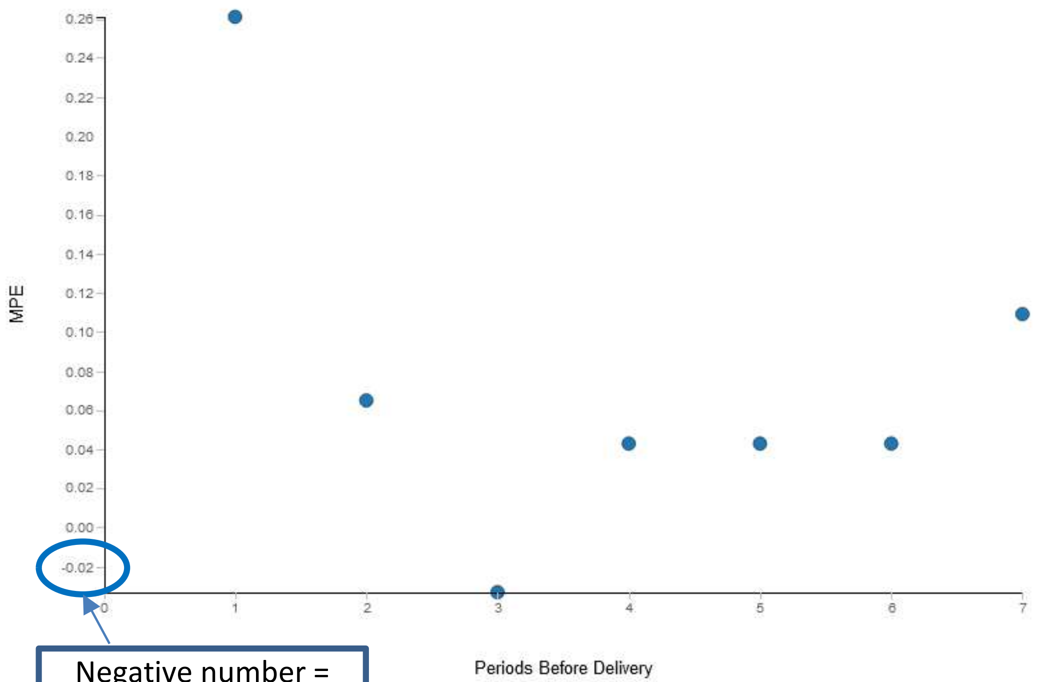
Where forecasts are $\mathbf{x}_{i,j}$, and final orders are $\mathbf{x}_{i,0}$

Represents the forecasting accuracy with respect to PBD – **zero means the perfect score.**

Web Tool: Forecast Error Measures

– ern_1. Mean Percentage error (MPE)

Forecast Error Measures
Mean Absolute Deviation (MAD)
Mean Deviation (MD)
Mean Square Error (MSE)
Root Mean Square Error (RMSE)
Normalized Root Mean Square Error (RMSE*)
Mean Percentage Error (MPE)
Mean Absolute Percentage Error (MAPE)
Mean Forecast Bias (MFB)



Negative number = underbooking

$$MPE_j = \frac{\sum_{i=1}^{n} x_{i,j} - x_{i,0}}{\sum_{i=1}^{n} x_{i,0}}$$

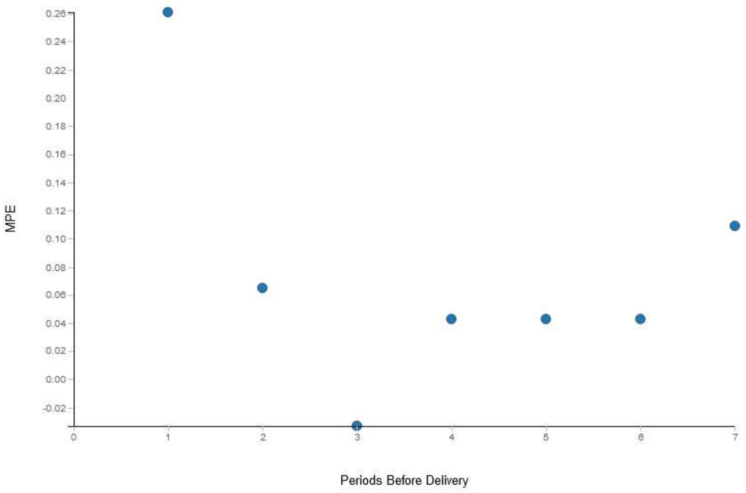
Where forecasts are $\mathbf{x}_{i,j}$, and final orders are $\mathbf{x}_{i,0}$

Represents the weighted forecasting accuracy with respect to PBD. Defined by the bias in the forecast behaviour. Positive or negative estimation represents whether there was **over- or under-booking in forecasts**.

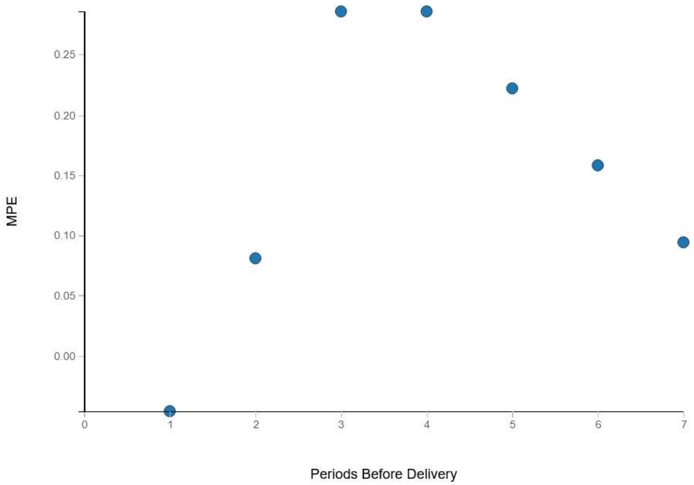
Web Tool: Forecast Error Measures

– Mean Percentage error (MPE) comparison

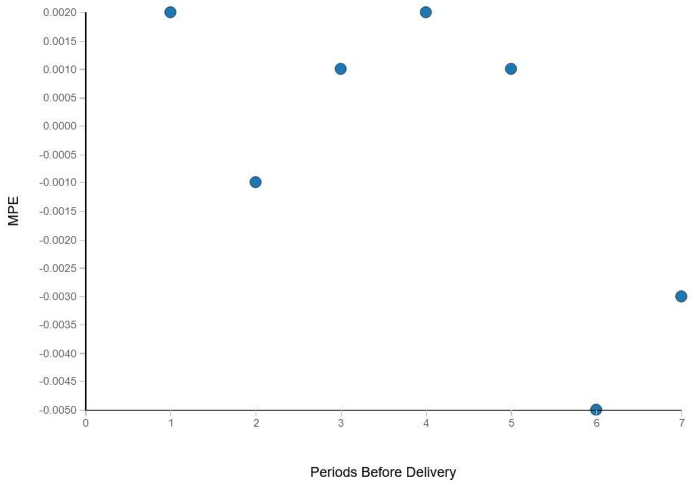
ern_1



ern_2



ern_3

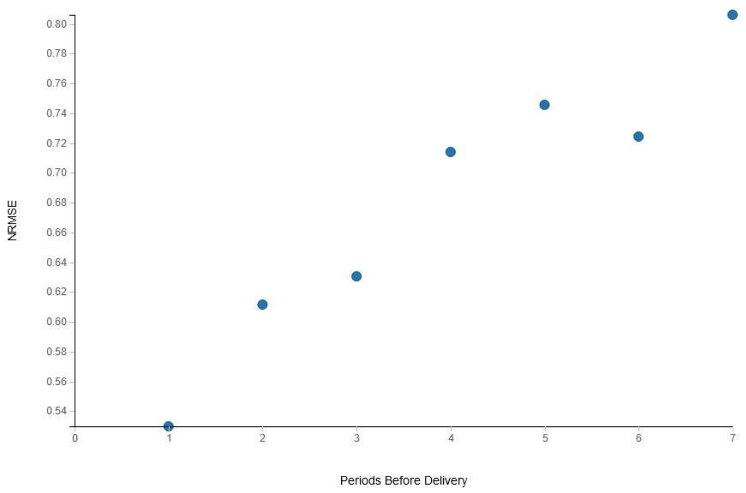


Which of the products has a better forecast accuracy?
Which one has underbooking?
Why?

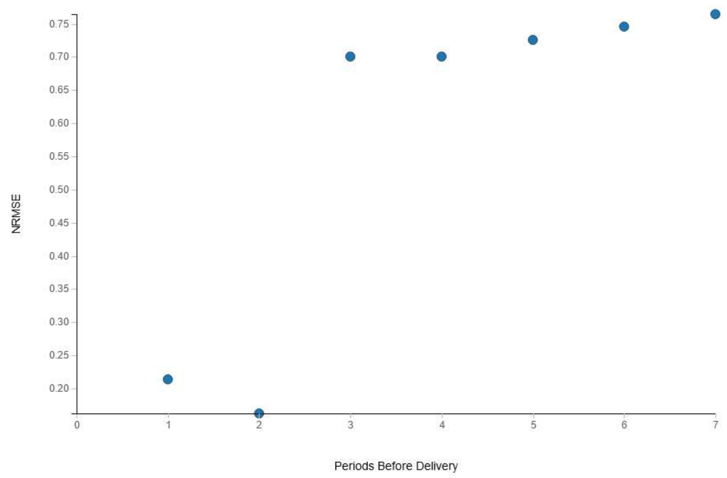
Web Tool: Forecast Error Measures

– NRMSE or RMSE* comparison

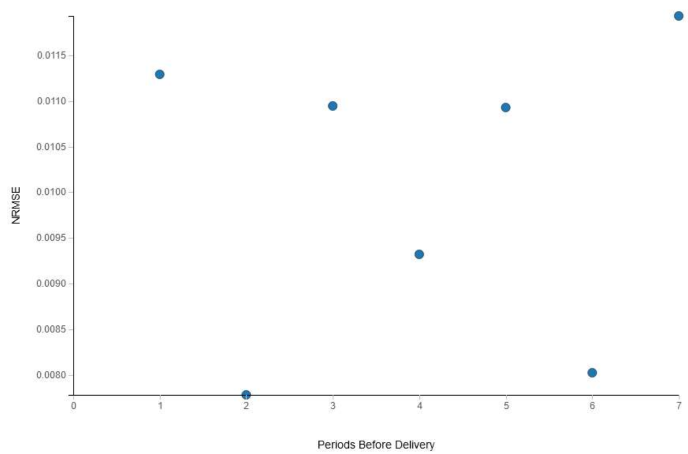
ern_1



ern_2



ern_3



Which of the products has a better forecast accuracy?
Why?

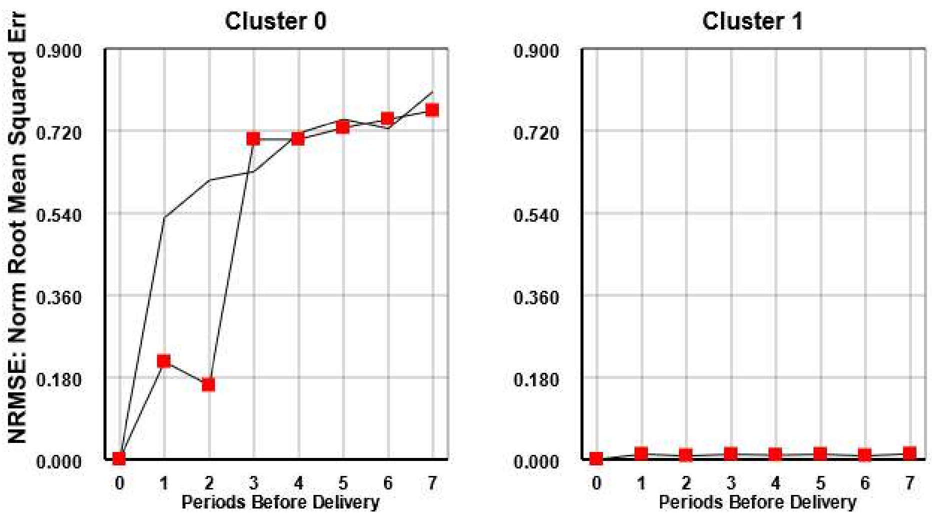
Web Tool: Tutorial on clustering

- Select all products: ern_1, ern_2 and ern_3
- Go to clustering page and apply the Affinity propagation clustering for NRMSE parameter (damping factor 0.5).
- Use configuration page to select all three products
- Wait for 2 minutes to load the clustering result

- Present your clustering result. How many clusters are allocated? Which products are allocated to what clusters?

Web Tool: Tutorial on clustering: Results

What kind of conclusions can be elaborated from this analysis?



Product_1
Product_2

Product_3

RESULTS: NRMSE: Norm Root Mean Squared Error and Affinity

Number Clusters: 2
Davies-Bouldin Score: 0.1649
Silhouette-Score: 0.4488

Appendix

Funding Information

This Leitfaden was created in the research project InnoFIT (No. 867471) funded by the Austrian Research Promotion Agency (FFG) in the initiative “Produktion der Zukunft”.

Contact for further information:

FH-Prof. Priv. Doz. DI (FH) Klaus Altendorfer PhD: klaus.altendorfer@fh-steyr.at

FH-Prof. Dr. Thomas Felberbauer, MSc: thomas.felberbauer@fhstp.ac.at

Examples in the slides

For example 1-8, there are data and MS Excel solutions available on request from:

klaus.altendorfer@fh-steyr.at or thomas.felberbauer@fhstp.ac.at