L8: TIME SERIES FORECASTING

Aneta Havlínová, Michal Hakala

20th of June, 2023

TODAY'S LECTURE

1. What are time series, use cases
   - Examples
   - Stationarity
   - Decomposition
2. Things to consider before forecasting
3. Simple models: mean, naive, ..
4. Exponential Smoothing
5. ARIMA
6. Extra: Prophet
7. Explanatory variables in forecasting
8. Model evaluation

Anet

Michal

czechitas

Public

Panel Data

Time (t)	2018 (t=1)					2019 (t=2)					2020 (t=3)				
	Person (index)	Income	Education (in years)	Gender (male=0, female=1)	Czechitas Course (no=0, yes=1)	Person (index)	Income	Education (in years)	Gender (male=0, female=1)	Czechitas Course (no=0, yes=1)	Person (index)	Income	Education (in years)	Gender (male=0, female=1)	Czechitas Course (no=0, yes=1)
	1	61 000	17	1	0	1	66 000	18	1	0	1	69 000	19	1	1
	2	50 000	13	0	0	2	52 000	13	0	1	2	55 000	13	0	1
	…	…	…	…		…	…	…	…		…	…	…	…	
	n	60 000	21	1	0	n	64 000	22	1	0	n	68 000	22	1	0

- Individual indices $i = 1,2, \dots, n$
- Time indices $t = 1,2, \dots, T$
- Model example: $income_{it} = \beta_0 + \beta_1 education_{it} + \beta_2 gender_i + \beta_3 course_{it} + \epsilon_{it}$
- Multiple individuals are observed over multiple periods

Time Series

Time (t)	2018 (t=1)	2019 (t=2)	2020 (t=3)

Person (index)	Income	Education (in years)	Gender (male=0, female=1)	Czechitas Course (no=0, yes=1)
1	61 000	17	1	0

Person (index)	Income	Education (in years)	Gender (male=0, female=1)	Czechitas Course (no=0, yes=1)
1	66 000	18	1	0

Person (index)	Income	Education (in years)	Gender (male=0, female=1)	Czechitas Course (no=0, yes=1)
1	69 000	19	1	1

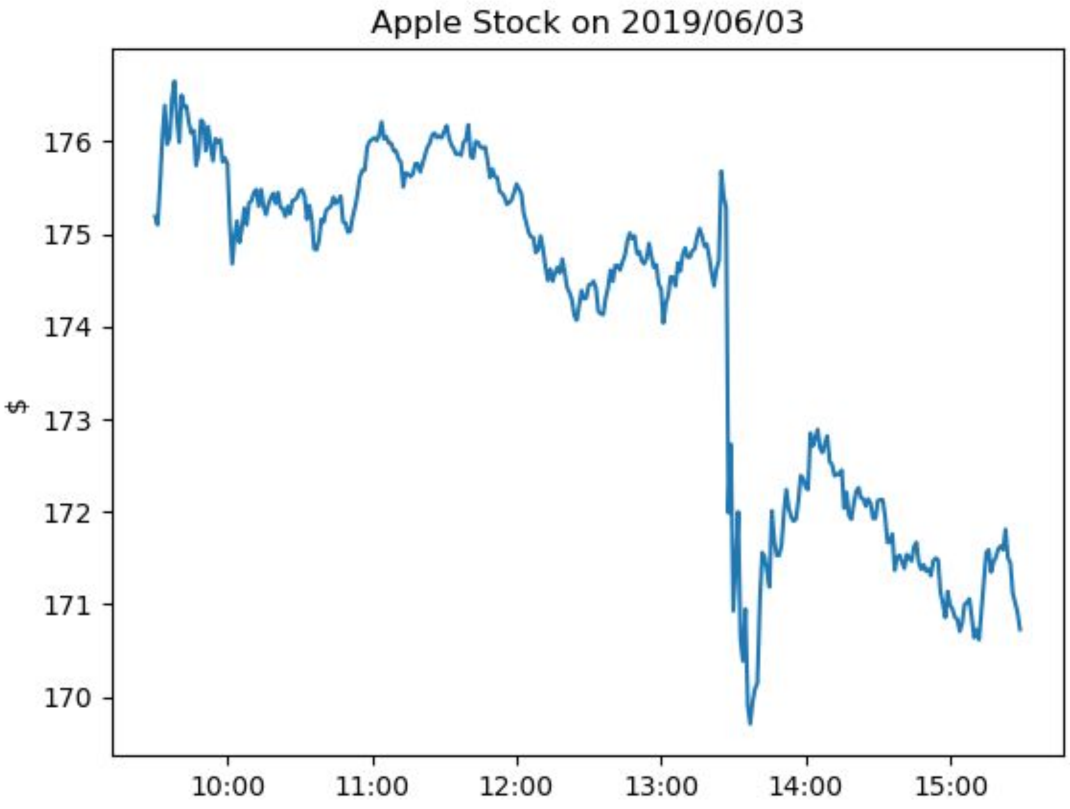
- One individual (= "one observations")
- Time indices $t = 1,2,\dots,T$
- Models mainly about dynamic properties (=how past affects future)

$$unemployment_t = \beta_0 + \beta_1 unemployment_{t-1} + \beta_2 crisis_{t-1} + \epsilon_t$$

- Extra assumption needed

czechitas

Time Series Example – Intraday Stock Prices



Time Series Tasks

Forecasting

- predicting future
- do not care about causal interpretation
- common task in practice
- Example: predict unemployment for the next month

Testing Hypothesis

- testing impact of one variable on another
- challenging to correctly specify model and use the right test
- Example: forward prices of commodities are predicting spot prices

Stationarity

Stationarity - Definition

Stationary Series

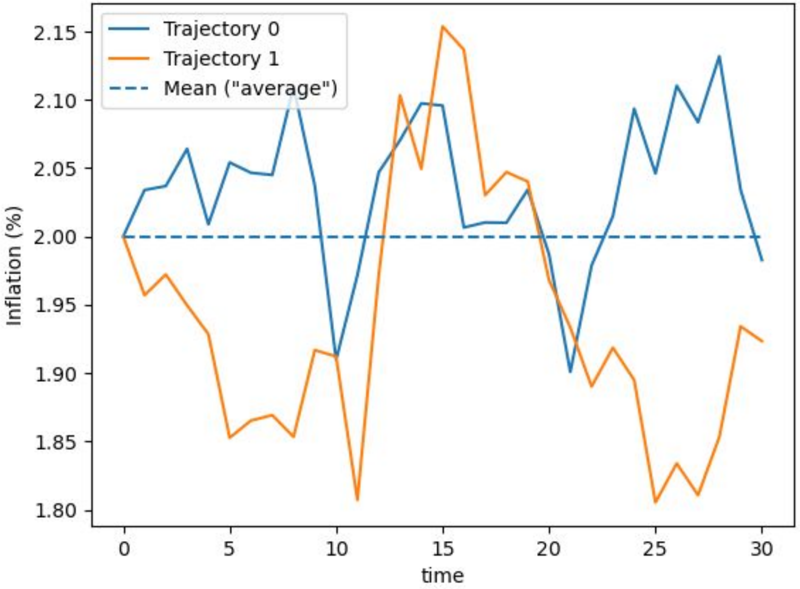
- long-term unconditional distribution of random variables is independent from time
- “values seems to be similar in nature in a long run”

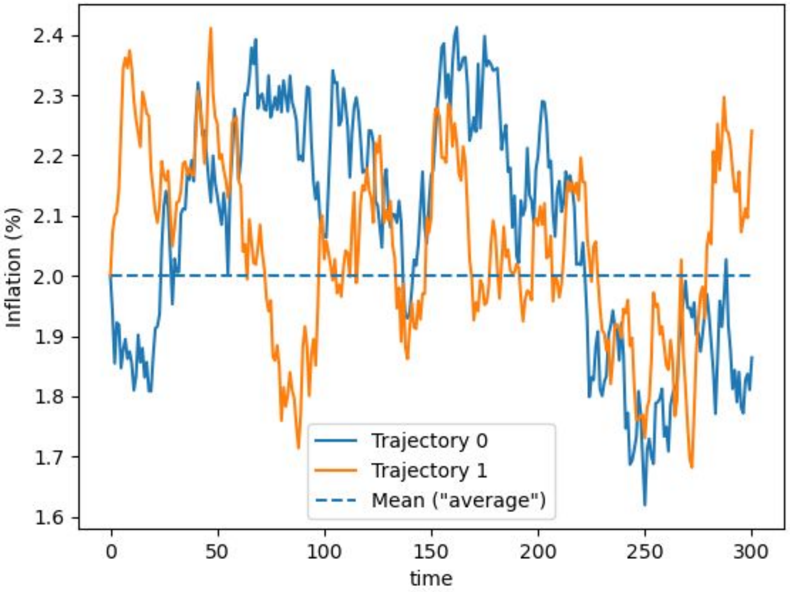
Variance/Covariance Stationary Series

- long-term mean (=”average”) is not changing in time
- long-term variance is not changing in time
- autocovariances (=covariance between current value and lagged value) are not changing in time

Mean Stationarity – Simulation Example

- Unconditional mean is not changing over time



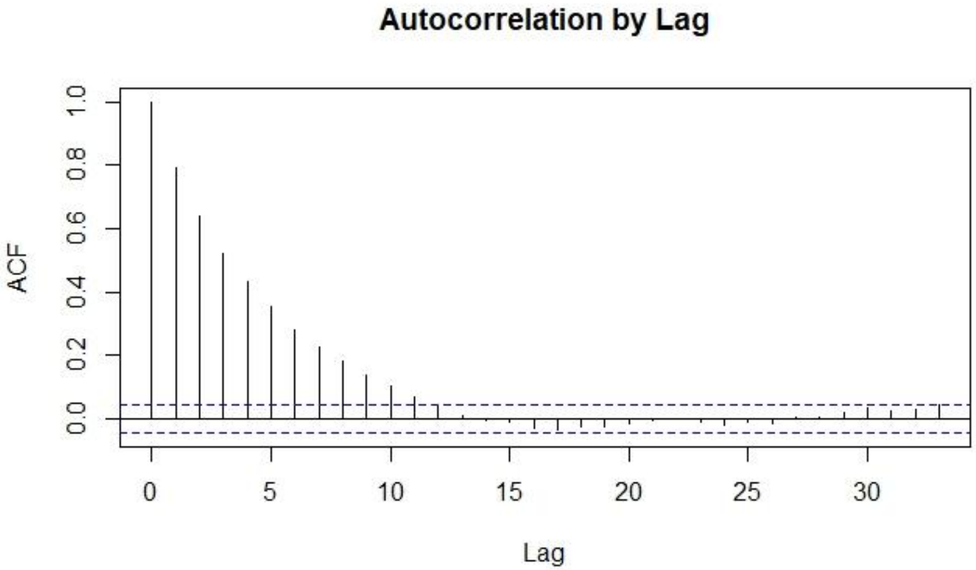


$$Inflation_t = \beta_o + \beta_1 Inflation_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma^2), \ |\beta_1| < 1$$

czechitas

Autocovariance and Autocorrelation Function

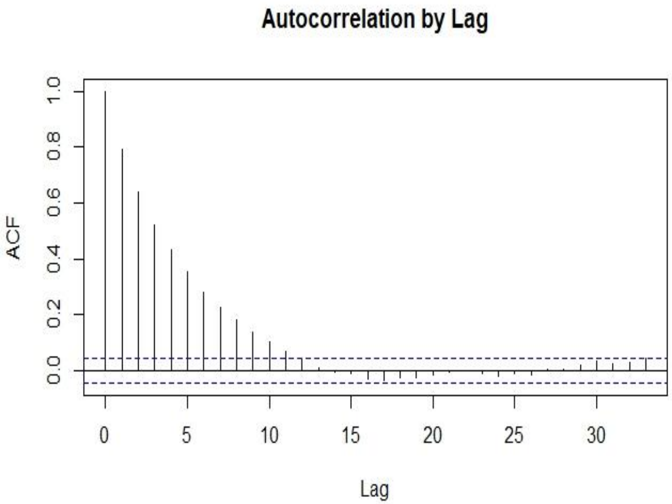
- Autocovariance $Cov(X_t; X_{t-k}) = \mathbb{E}[X_t X_{t-k}] - \mathbb{E}[X_t]\mathbb{E}[X_{t-k}]$ for any lag $\boldsymbol{k}$
- Captures linear dependence between values at different times
- Autocorrelation $Cor(X_t; X_{t-k}) = \frac{Cov(X_t; X_{t-k})}{\sqrt{Var(X_t)Var(X_{t-1})}}$

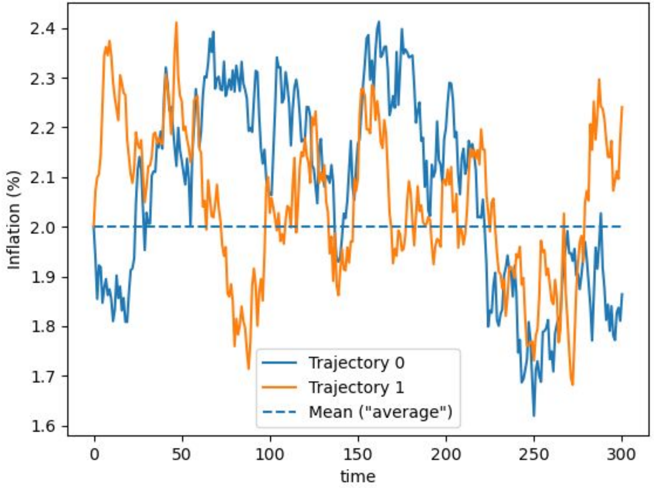


czechitas

Variance/Covariance Stationarity

- Variance $Var(X_t)$ is not changing over time
- Autocovariances $Cov(X_t; X_{t-k})$ (for any lag) are not changing over time





czechitas

Why Stationarity?

Problem

If parameters were changing in time, we would not be able to estimate them

- Stationarity makes parameters of our model constant in time

Question

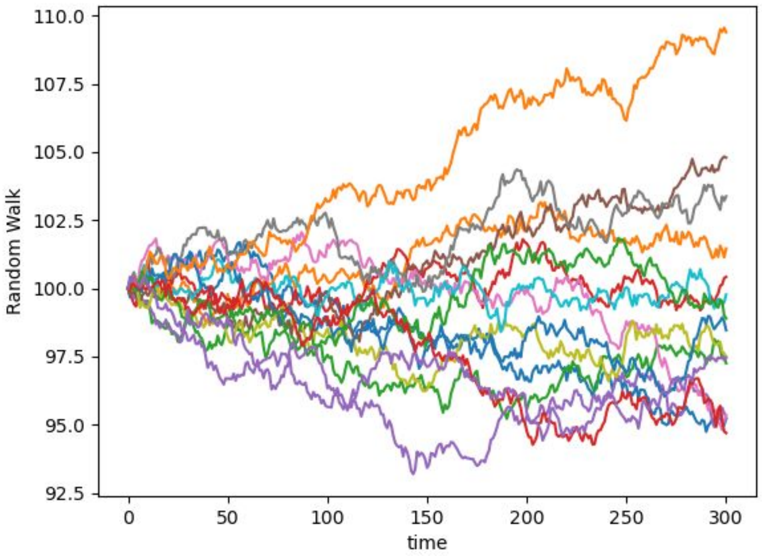
Does it mean that non-stationary series cannot be modeled?

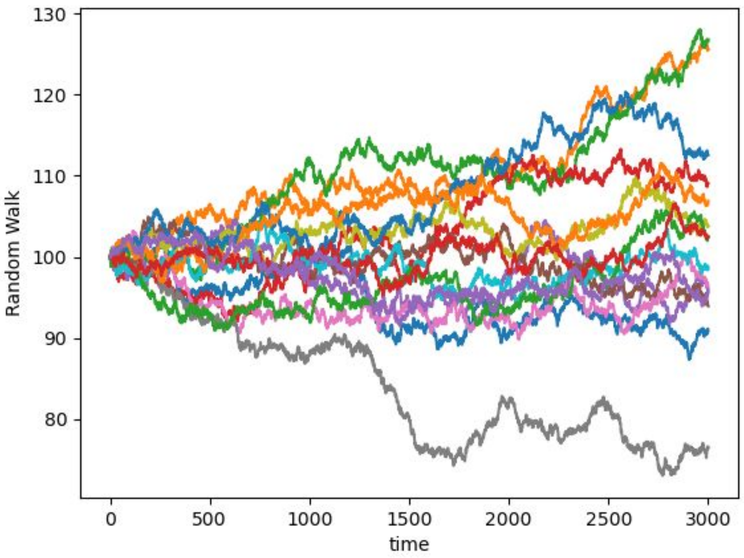
- No, but we have to specify form of non-stationarity

Example of Non-stationary Series – Unit Root Process

$$\dot{X}_t = X_{t-1} + \varepsilon_t, \;\; \varepsilon_t \sim iid, \mathbb{E}[\,\varepsilon_t] = 0,$$

Time (t)	0	1	2	3	4	5
		0,0425477	-0,1101036	0,00978842	0,10464312	-0,2383265
	100	100,042548	99,9324441	99,9422325	100,046876	99,8085491

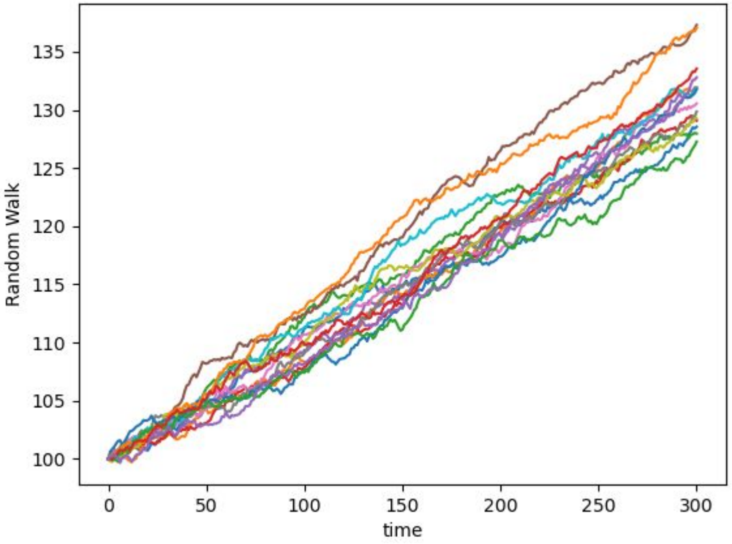


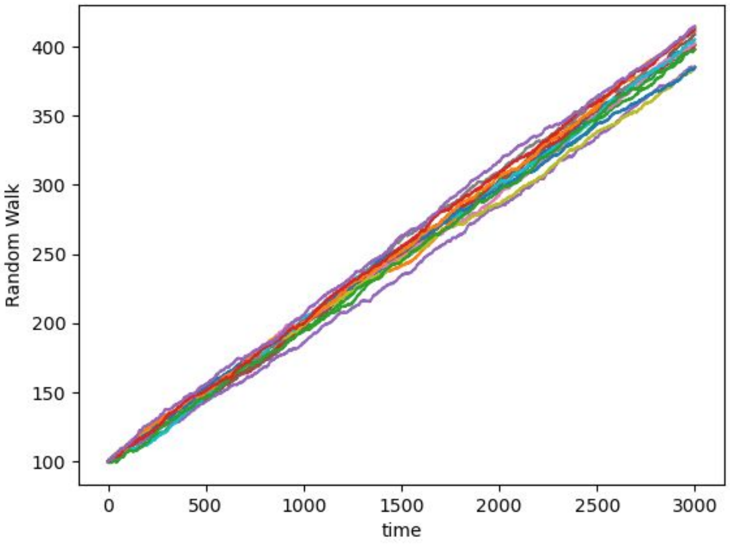


czechitas

Example of Non-stationary Series – Unit Root Process

$$\dot{X}_t = \mu + X_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim iid, \mathbb{E}[\,\varepsilon_t] = 0, \qquad \mu = 0.1$$





czechitas

Differencing – Unit Root Process

Non-stationary Process

$$X_t = \mu + X_{t-1} + \varepsilon_t = \mu + \underbrace{(\mu + X_{t-2} + \varepsilon_{t-1})}_{X_{t-1}} + \varepsilon_t = t * \mu + X_0 + \sum_{i=1}^{t} \varepsilon_i$$

- Mean $\mathbb{E}[X_t|X_0] = t * \mu + X_0$ [depends on t]
- Variance $Var[X_t|X_0] = Var\left[\sum_{i=1}^{t} \varepsilon_i\right] = tVar[\varepsilon_i]$ [depends on t]

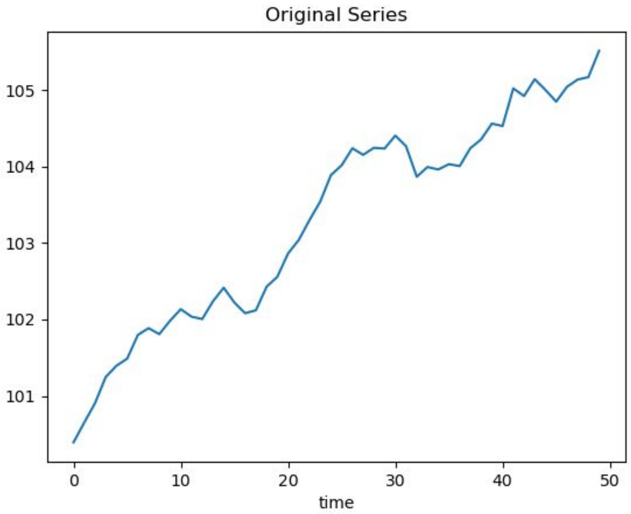
Stationary Process - Differenced Series

$$\Delta X_t = X_t - X_{t-1} = \mu + \varepsilon_t$$

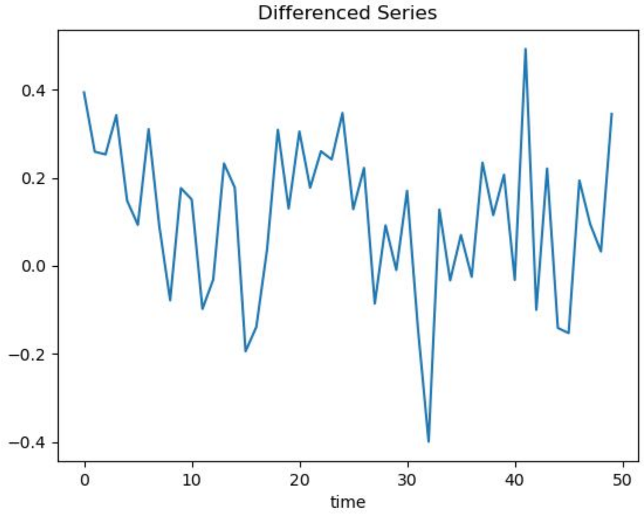
- Mean and variance time independent

Differencing

- Removes “stochastic” trend
- Sometimes repeated differencing is needed

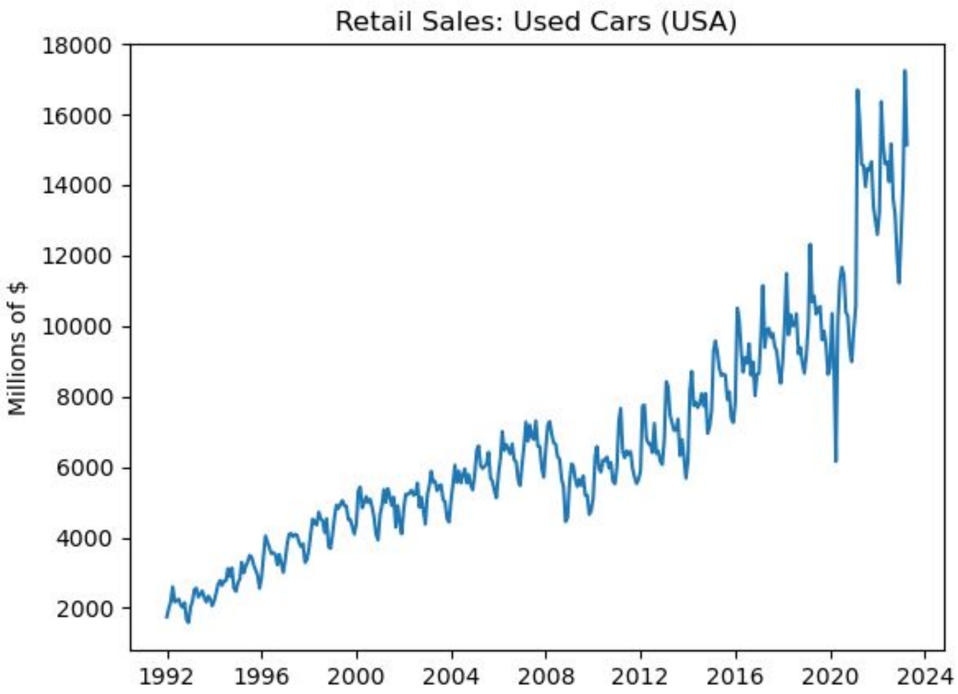


Differencing



czechitas

Example of Non-stationary Series - Seasonality



Seasonality/Diurnal Patterns

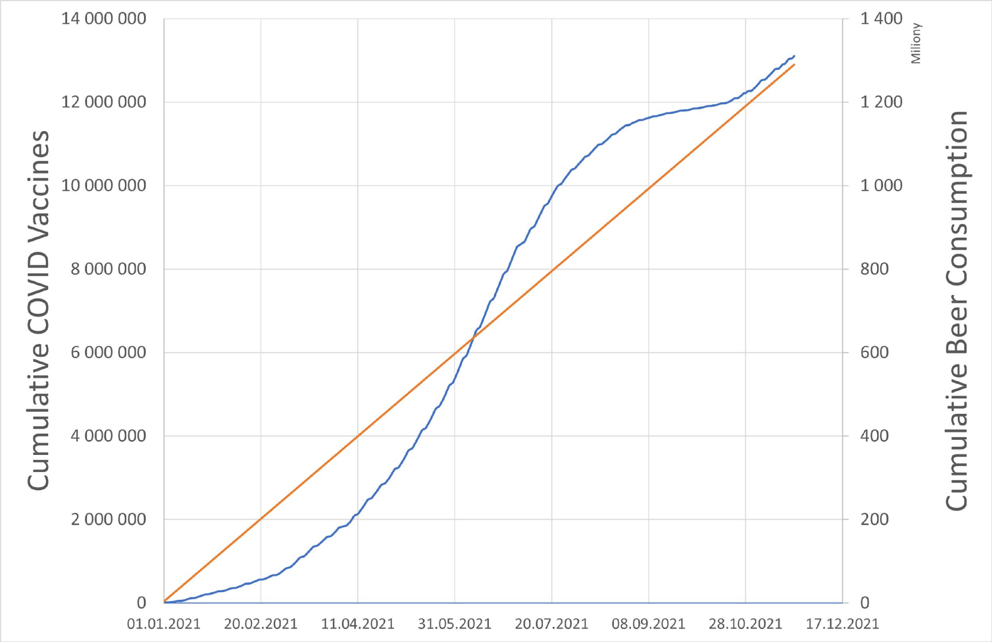
- Repeated patterns within year or day
- Ignoring leads to invalid estimates and imprecise forecasts

- Differencing with respect to years (SARIMA)
- Seasonal adjustment (decomposition based models)

czechitas

Example of Non-stationary Series – “Deterministic” Trend

Correlation 0.98



Deterministic Trend

- Trend being function of time
- Ignoring leads to invalid estimates and “spurious dependencies”

- Differencing (ARIMA or VAR/ECM)
- Detrending (decomposition based models)

Non-stationary Components

Source	When Ignored	Solution
• Deterministic function of time • ”Stochastic” trend • Seasonality	• Invalid estimates • Incorrect forecasts • ”Spurious” Relationship	• Differencing • Decomposition

Decomposition

Time Series Decomposition

- Assumes time series can be break down into 4 components

Time Series

Trend T_t
- Long-term growth or decline

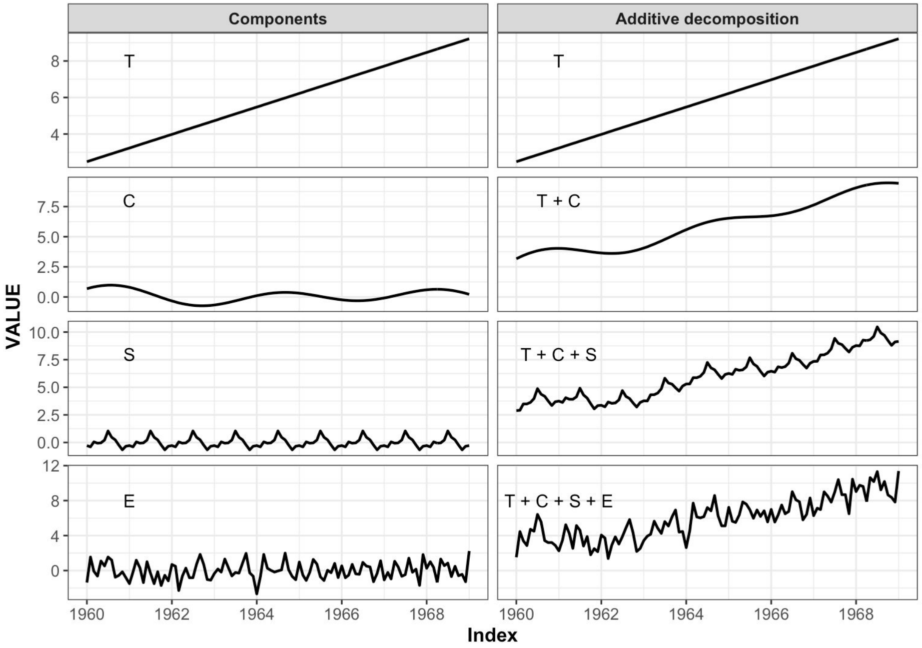
Cycle C_t
- Regular changes in a long-run

Seasonality S_t
- Regular changes within year

Residuals E_t
- Random variations

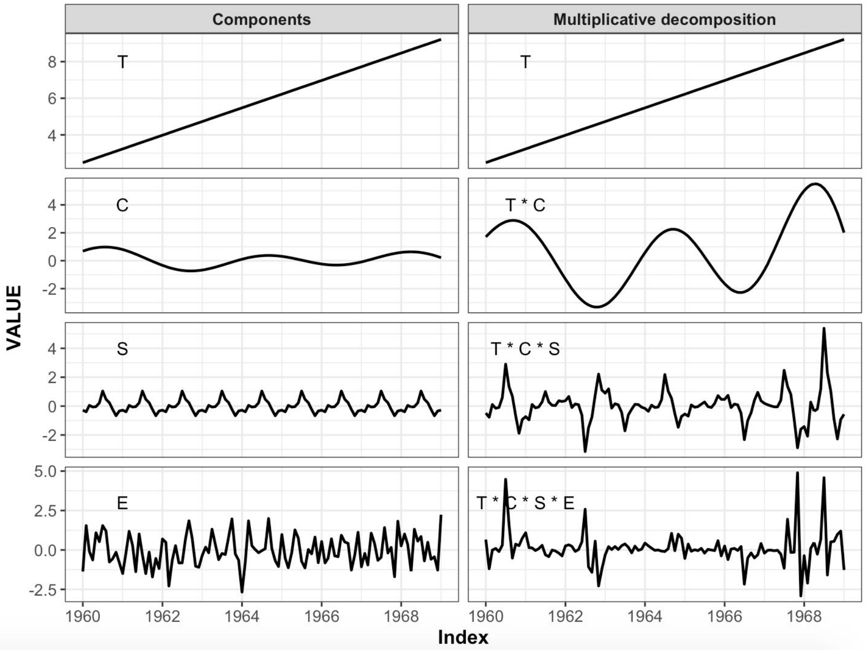
Additive decomposition

$$T + C + S + E$$



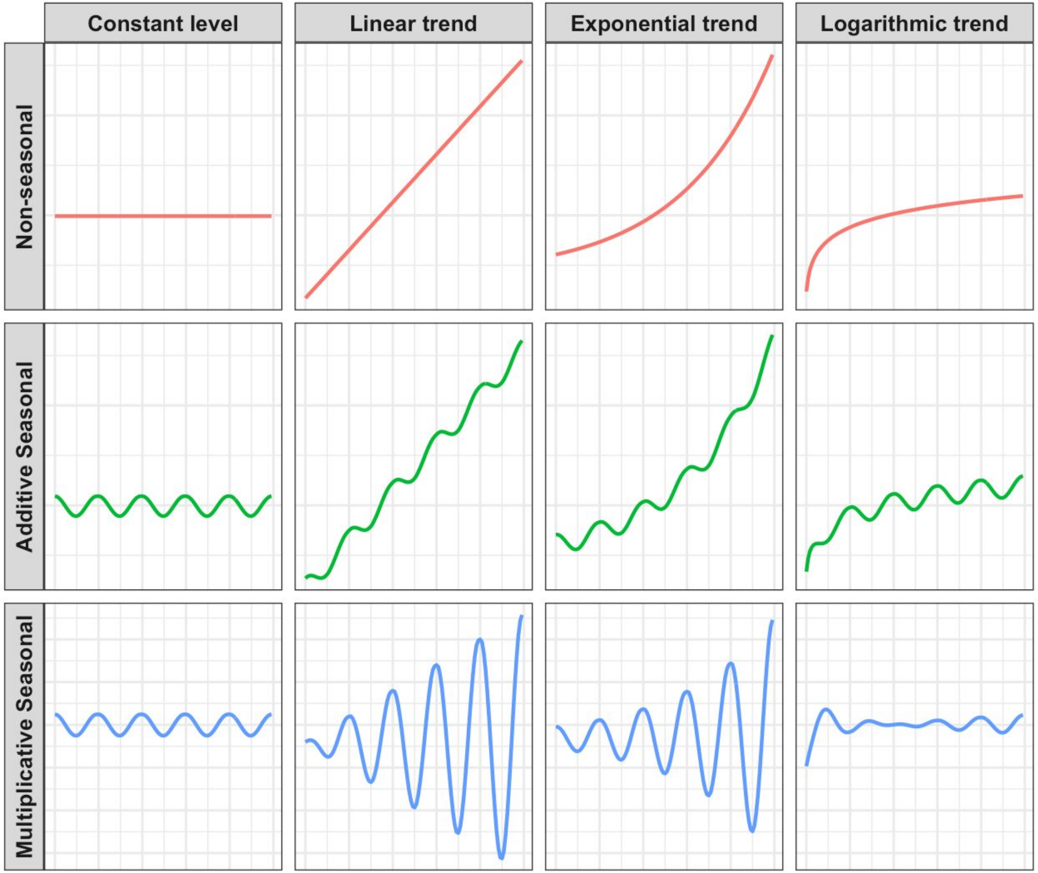
Multiplicative decomposition

$$T * C * S * E$$



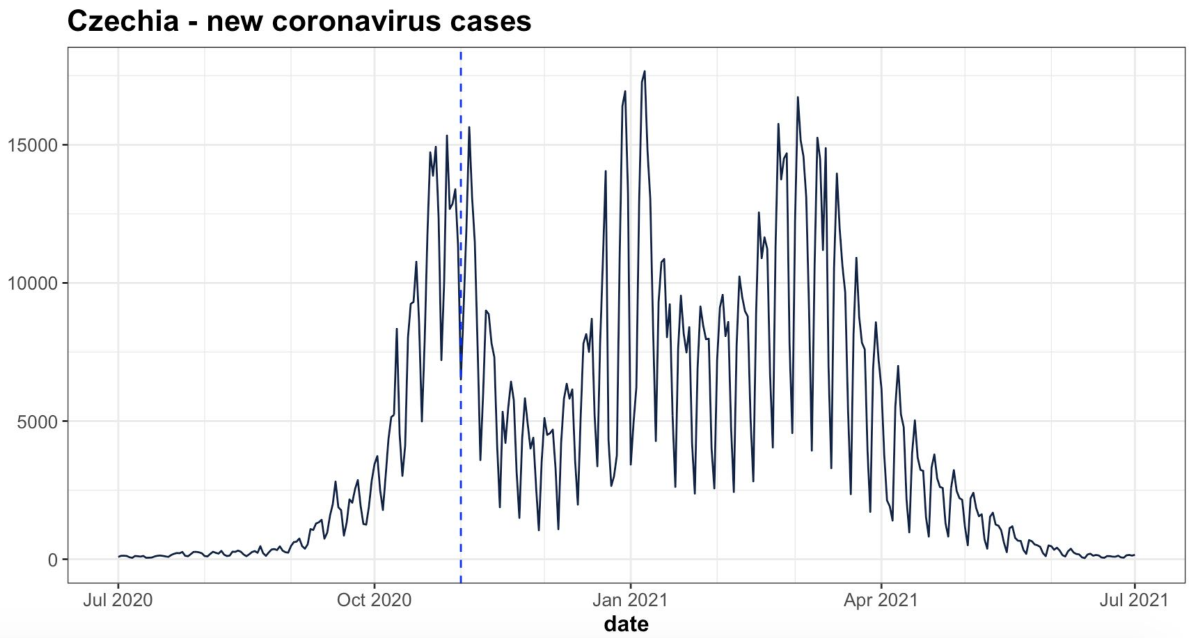
czechitas

Examples of different trends and seasonality



Danger of Deterministic Trends

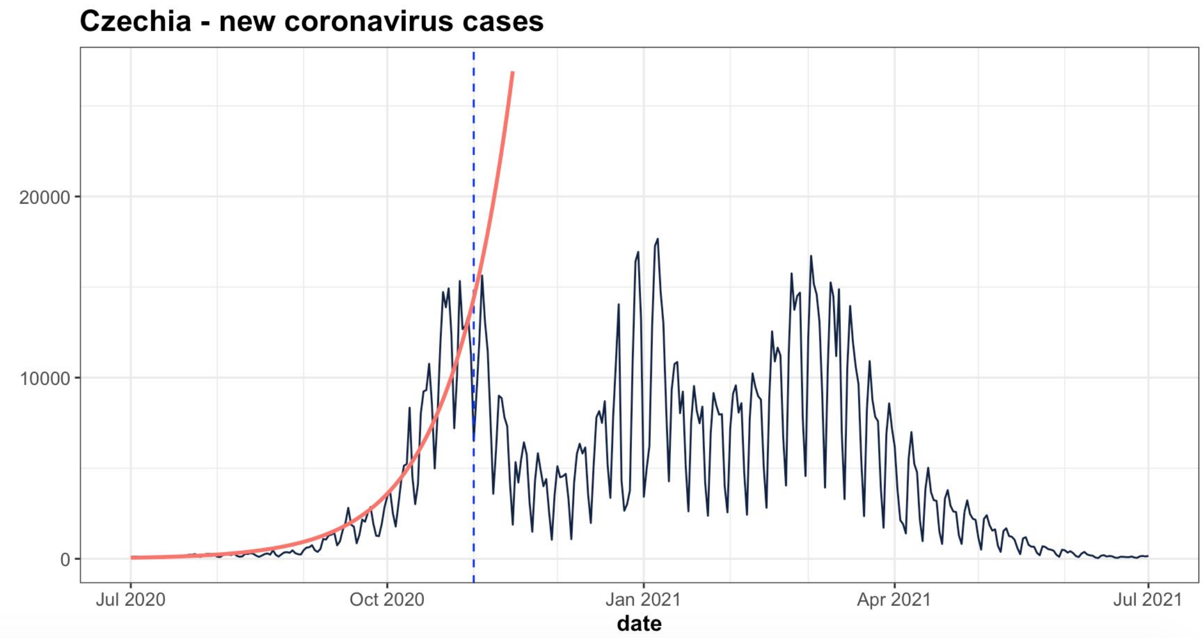
Be careful if you detect strong trend in your data. If you know that your data can't cross some threshold in the near future, then try to avoid linear or exponential trend.



czechitas

Danger of Deterministic Trends

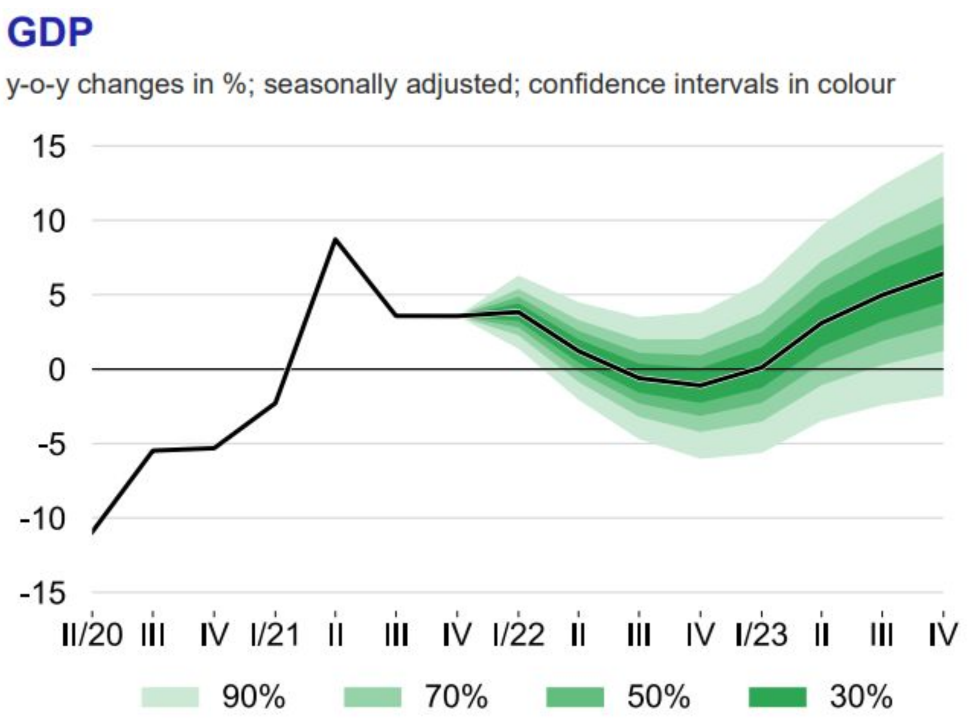
Be careful if you detect strong trend in your data. If you know that your data can't cross some threshold in the near future, then try to avoid linear or exponential trend.



czechitas

THINGS TO CONSIDER BEFORE FORECASTING

Forecasting



Before Forecasting

1. What horizon do I need?
2. What frequency of data do I have?
3. Is the series predictable?

Objects of Forecasting

1. Point Predictions
2. Interval Predictions

czechitas

Point Forecasts

- One-Period-Ahead

$$\hat{X}_{T+1|T} = E[X_{T+1}|X_T, X_{T-1}, X_{T-2}, \dots]$$

Example: $X_t = 0.5 + 0.2X_{t-1} + \varepsilon_t$

$$E[X_{T+1}|X_T, X_{T-1}, X_{T-2}, \dots] = E[0.5 + 0.2X_T + \varepsilon_t|X_T] = 0.5 + 0.2X_T + 0$$

$$\hat{X}_{T+1|T} = 0.5 + 0.2 * \boxed{0.42} + 0 = 0.584$$

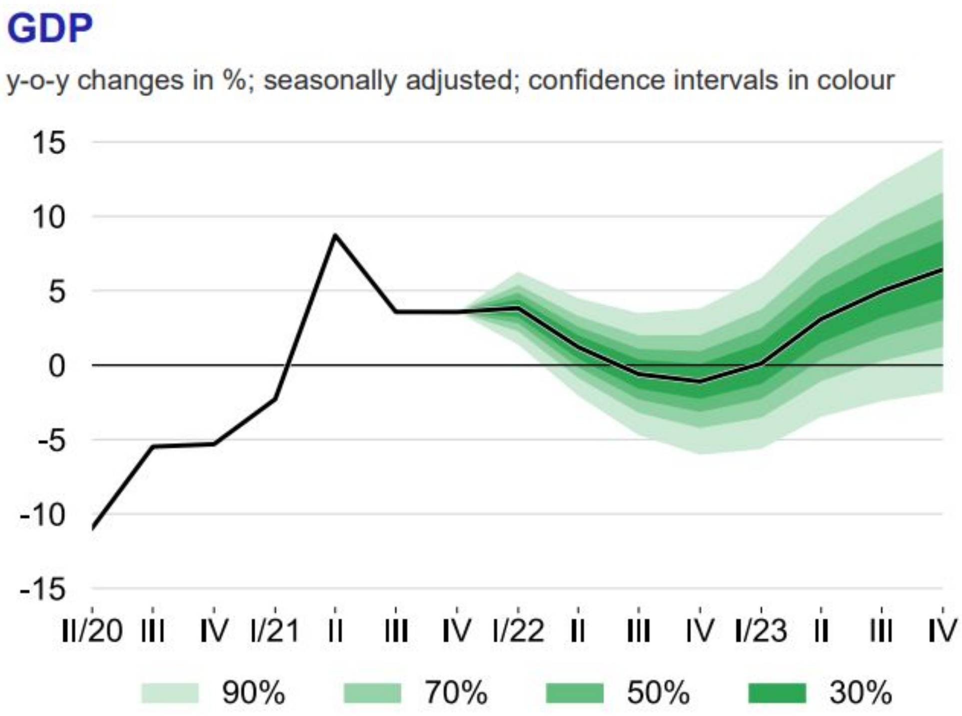
0	1	T=2	3
0,4	0,45	0,42	0,584

Historical Data

Forecast

1) Prediction Horizon

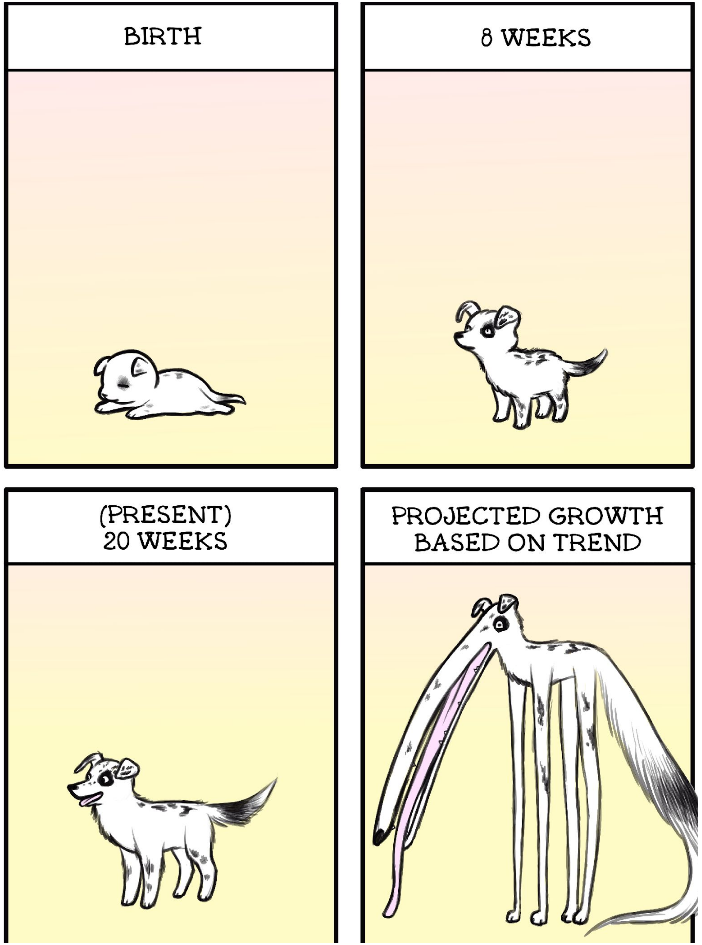
Horizon



- Longer horizon is harder to predict

Prediction Horizon

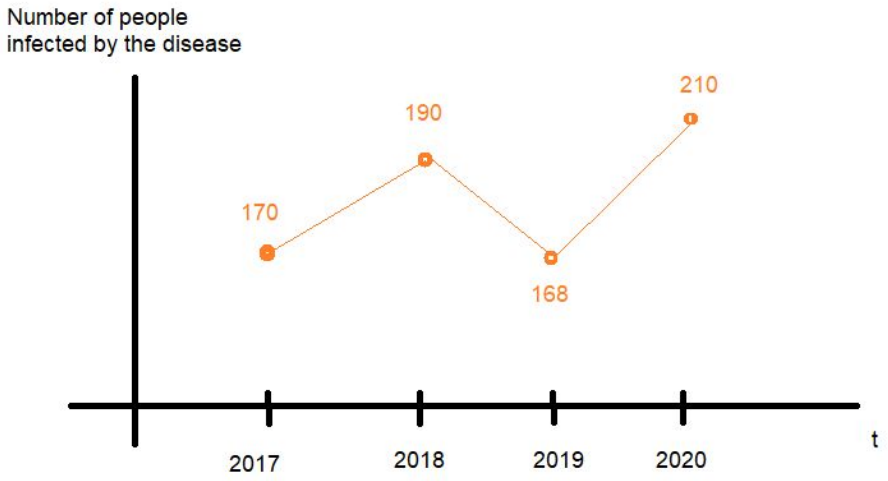
- We are making strong assumptions about trend
- Assumptions may be correct in short-run but wrong in a long-run



czechitas

2) Possibility to Forecast

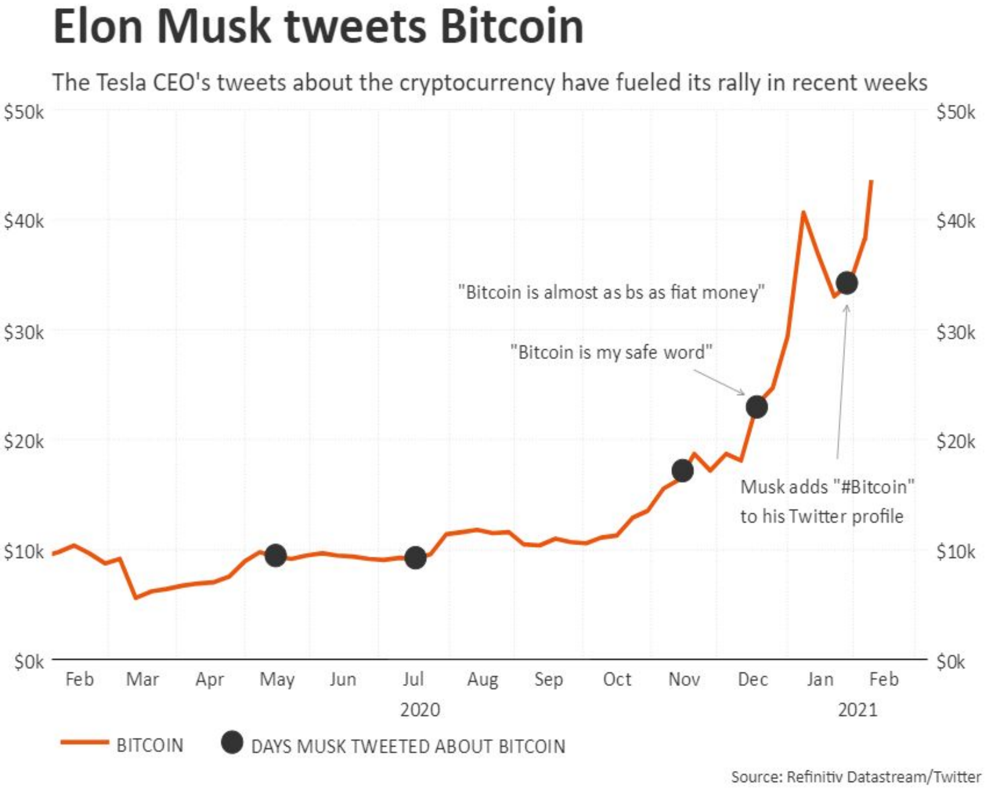
Amount of data available is usual limiting factor



Problems with Data Availability

1. Short history of data
2. Small frequency of available data

Occurrence of major random (non predictable) events influencing predicted variable

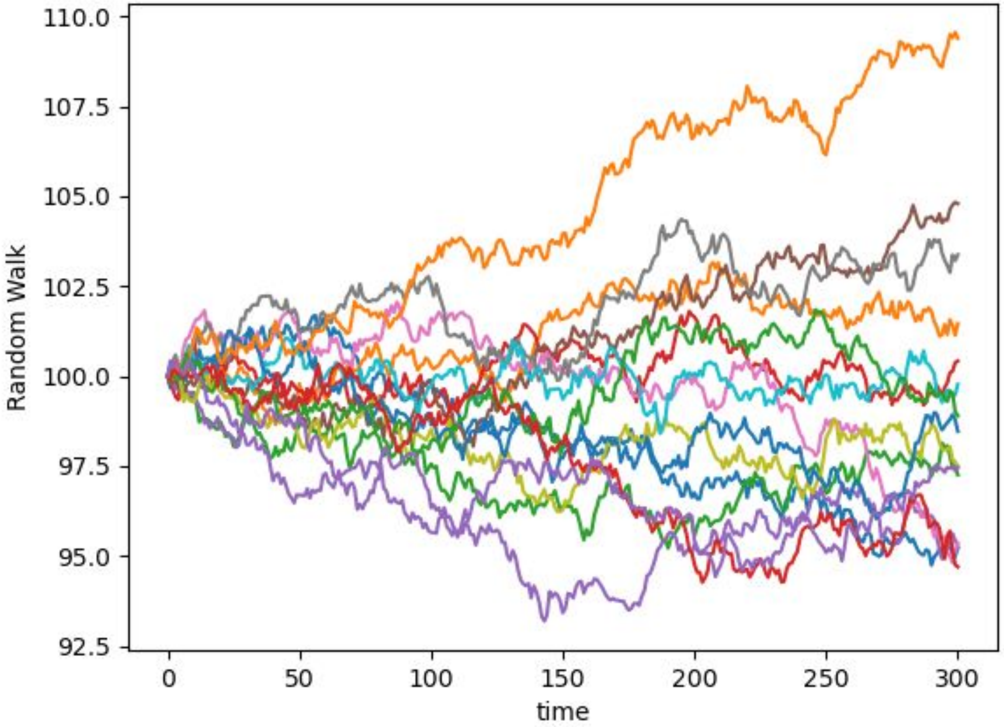


Unpredictable Rare Events

1. Human behavior
2. Forces of nature
3. …

czechitas

Variable to predict is not predictable



Random Walk

$$X_t = X_{t-1} + \varepsilon_t$$

- Forecast $\hat{X}_{t+1|t} = X_t$
- Randomly moves up and down – do not know what will happen next

Examples

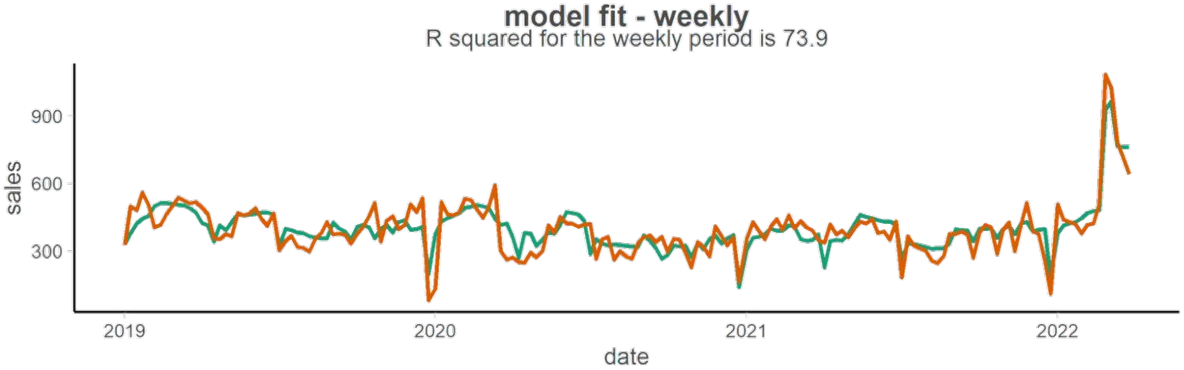
- Stock prices
- Brownian motion – random movement of particles

czechitas

3) Frequency Required for the Forecast

Example 2: Comparing fit of the weekly vs monthly model
Why has the monthly aggregation higher R^2?

WEEKLY

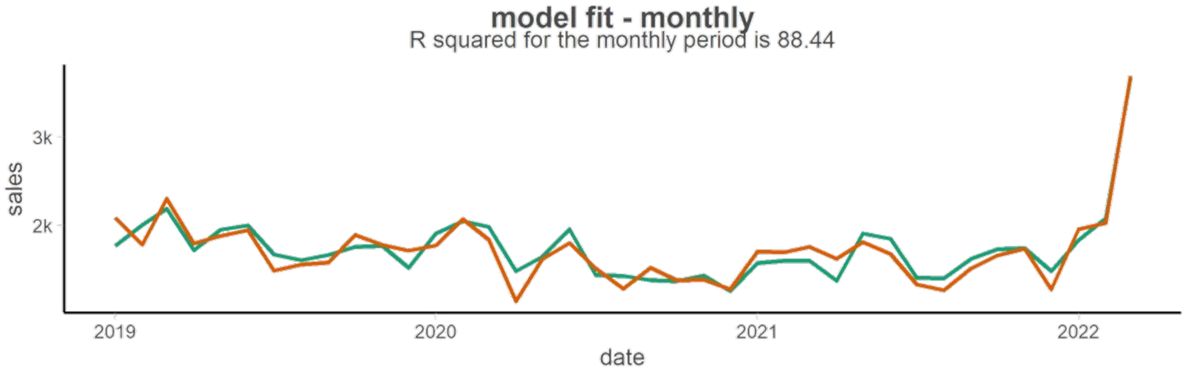


model R^2 = 0.74

Predicted

Actual

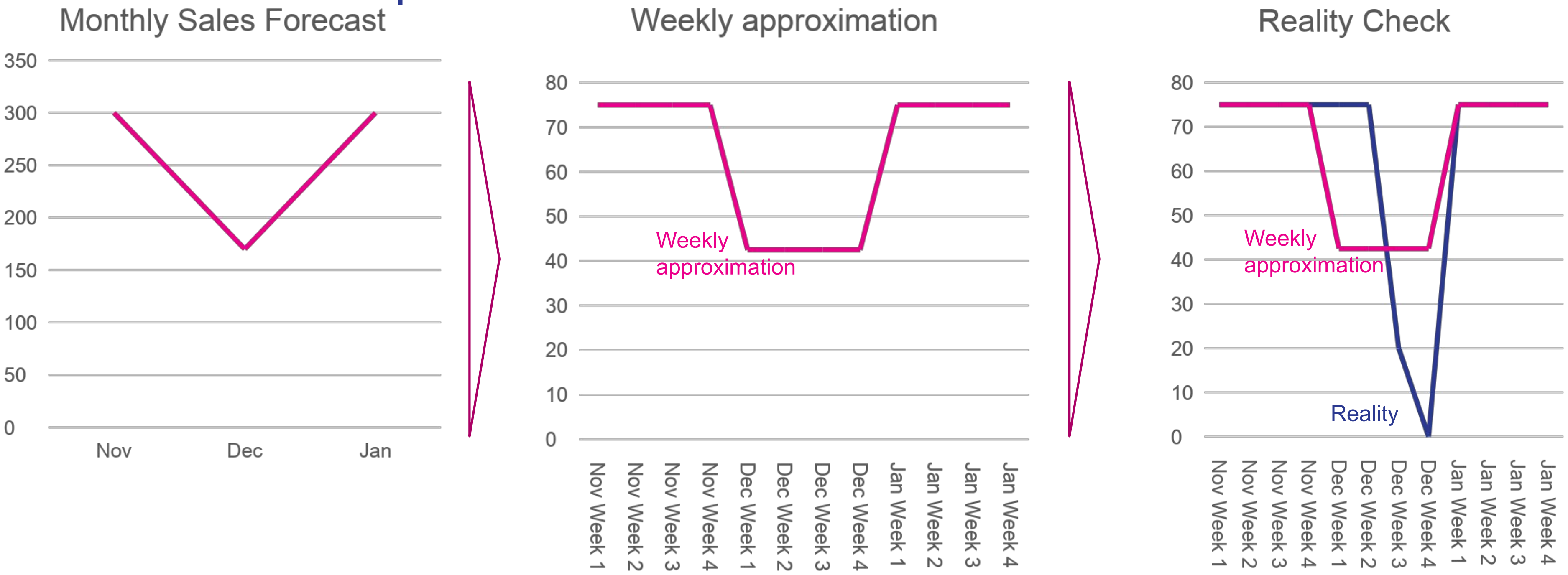
MONTHLY



model R^2 = 0.88

czechitas

Example 3: Getting weekly forecast out of the monthly forecast – It is possible, but what are the drawbacks?



When moving to more frequent approximation form aggregated data, you can not estimate higher frequency noise, you can just average it

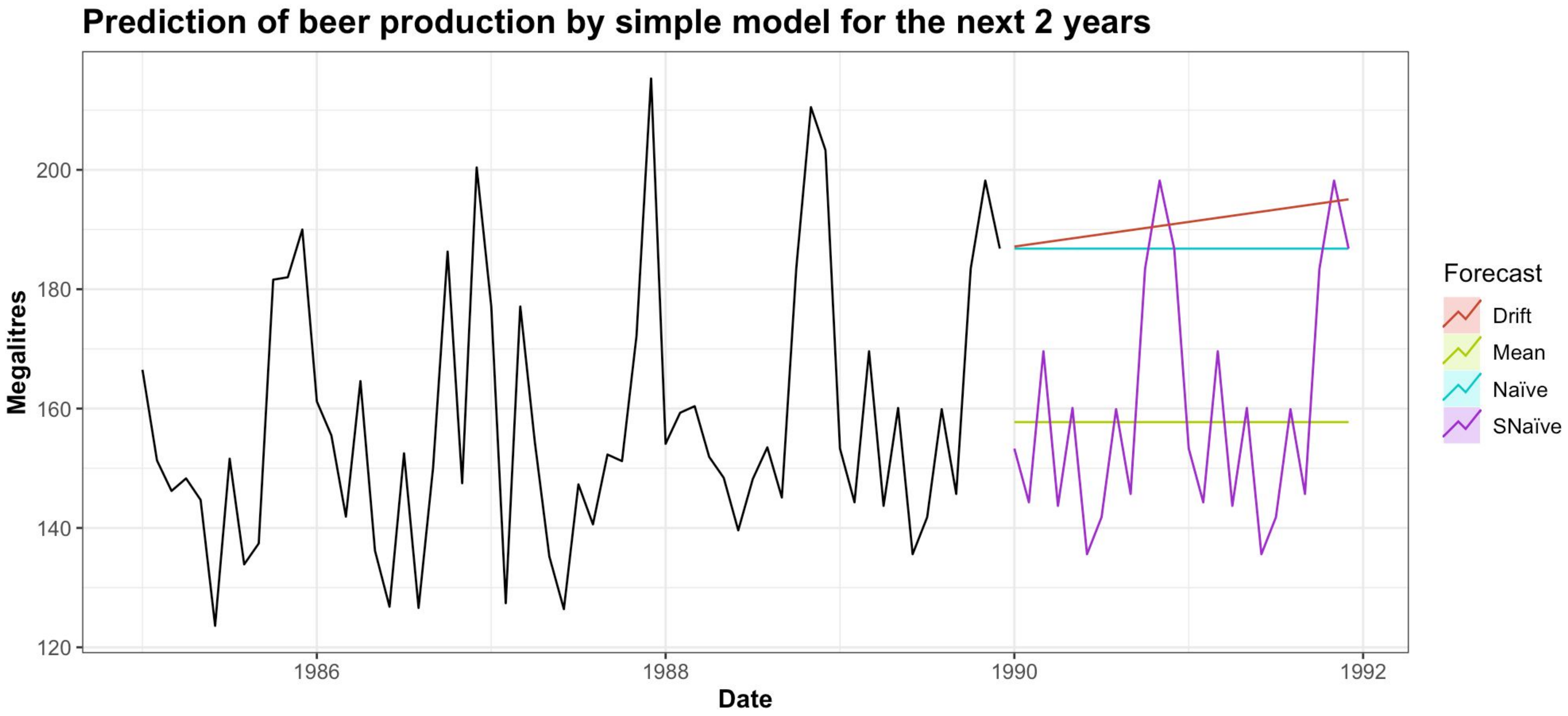
FORECASTING: THE SIMPLEST MODELS

Simple techniques

- used as benchmarks
- used when not much data available

Model	Formula
Naive model	$\hat{y}_{T+h\|T} = y_T$
Snaive model (s=seasonal)	$\hat{y}_{T+h\|T} = y_{T+h-m(k+1)}$, where m is the seasonal period, and k is the integer part of $(h-1)/m$.
Mean model	$\hat{y}_{T+h\|T} = \bar{y} = (y_1 + \cdots + y_T)/T$
Median model	$\hat{y}_{T+h\|T} =$ "middle" data value of an ordered data set $\{y_1, \ldots, y_T\}$
Drift model	$\hat{y}_{T+h\|T} = y_T + h\left(\frac{y_T - y_1}{T-1}\right)$

Simple techniques



ETS - EXPONENTIAL SMOOTHING MODELS

(Error, Trend, Seasonal)

Adaptive Methods

Simple Mean Model

- Forecast $\hat{X}_{T+h|T} = \bar{X} = \frac{X_1+X_2+\cdots+X_T}{T}$
- Fails when any trend is present

Adaptive Methods

- Similar to simple methods
- Calculated only from a few newest values, e.g., average over last 5 observations

2. ETS Exponential Smoothing Models (Error, Trend, Seasonal)

- puts larger weights to recent observations than to observations from the more distant past

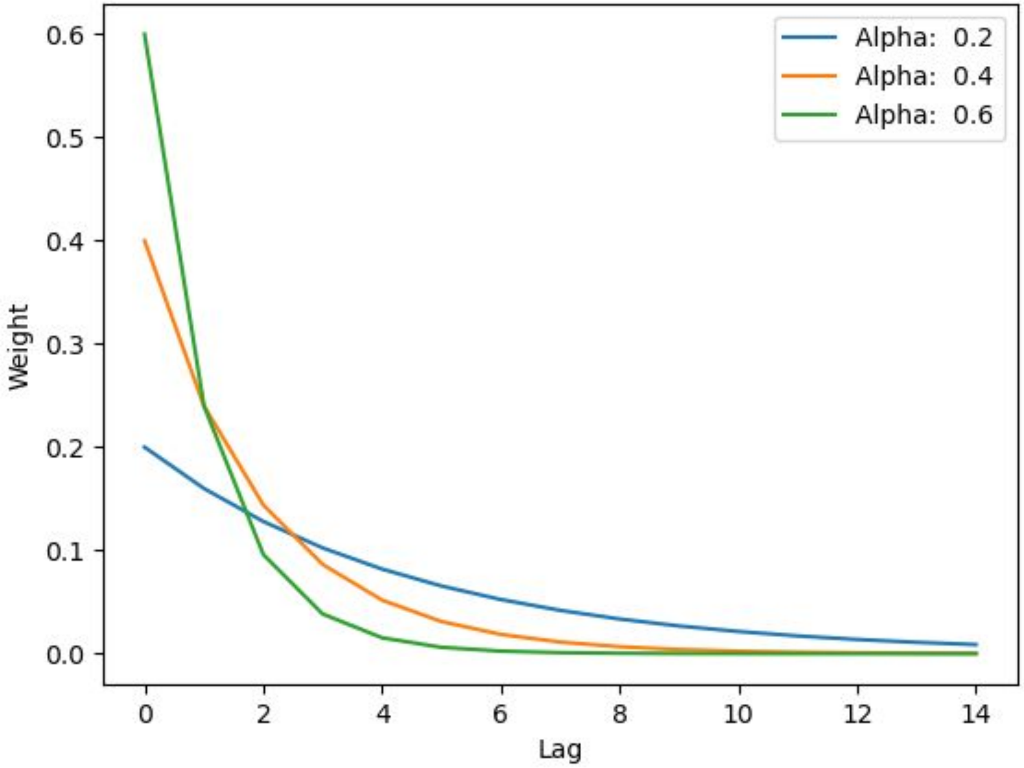
Trend exponential smoothing method

- ETS without seasonality

$$\hat{X}_{T+1|T} = \alpha X_T + \alpha(1-\alpha)X_{T-1} + \alpha(1-\alpha)^2 X_{T-2} + \cdots$$
$$0 \leq \alpha \leq 1$$

Interpretation of Exponential Smoothing

- $$\hat{X}_{T+1|T} = \boxed{\alpha} X_T + \boxed{\alpha(1-\alpha)} X_{T-1} + \alpha(1-\alpha)^2 X_{T-2} + \cdots = \boxed{w_1} X_T + \boxed{w_2} X_{T-1} + w_3 X_{T-2} + \cdots$$



Smoothing Parameter

Higher alpha puts more weight on more recent observations, less on distant observations

czechitas

ETS Model - Demo

ARIMA MODEL

ARIMA

AR - Autoregressive

I - Integrated

MA - Moving Average

ARIMA

Assumptions:

- **Stationary data** (= the model is **integrated**) → to achieve this, we usually need to apply data differencing
- **No seasonality** (adjusted approach needed for this case)

Components:

A) Auto Regressive part – AR

B) Moving Average part – MA

PART I: AUTOREGRESSIVE MODEL (AR)

We forecast the variable of interest using **past value of that variable**.

Day	Share price	Share price (t-1)	Share price (t-2)
01.01.2022	50	-	-
02.01.2022	52	50	-
03.01.2022	53.5	52	50
04.01.2022	52.75	53.5	52
05.01.2022	51.5	52.75	53.5

Dependent variable

czechitas

PART I: AUTOREGRESSIVE MODEL (AR)

We forecast the variable of interest using **past value of that variable**.

Autoregressive model of order 1:

$$y_t = ß_0 + ß_1 y_{t-1} + \varepsilon_t$$

y_t = share price on Wednesday 5th January
$ß_1$ = correlation between share price on Wednesday and Tuesday
ε_t = white noise (random error, residual)

czechitas

PART I: AUTOREGRESSIVE MODEL (AR)

We forecast the variable of interest using **past value of that variable**.

Autoregressive model of order p:

$$y_t = ß_0 + ß_1 y_{t-1} + ß_2 y_{t-2} + \cdots + ß_p y_{t-p} + \varepsilon_t$$

y_t = share price on Wednesday 5th January
$ß_1$ = correlation between share price on Wednesday and Tuesday
$ß_2$ = correlation between share price on Wednesday and Monday
ε_t = white noise (random error, residual)

PART II: MOVING AVERAGE (MA)

We add **present and past residuals** to the model

→ we try to capture the **unknown factors** influencing our variable

Moving average model of order 1:

$$y_t = \theta_0 + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

y_t = share price on Wednesday 5th January
ε_t = white noise (random error) on 5.1.
ε_{t-1} = white noise (random error) on 4.1.

czechitas

Public

PART II: MOVING AVERAGE (MA)

We add **present and past residuals** to the model

→ we try to capture the **unknown factors** influencing our variable

Moving average model of order q:

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$$

y_t = share price on Wednesday 5th January
ε_t = white noise (random error) on 5.1.
ε_{t-1} = white noise (random error) on 4.1.
etc.

czechitas

ARIMA

NOW WE CAN PUT IT ALL TOGETHER!!

AR - Autoregressive

I - Integrated

MA - Moving Average

ARIMA

Combination of autoregressive and moving average model is **ARIMA(p, d, q):**

$$y'_t = \boxed{ß_0 + ß_1 y'_{t-1} + \cdots + ß_p y'_{t-p}} + \boxed{\theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t}$$

p = Number of autoregressive terms

q = Number of moving average terms

czechitas

Public

ARIMA

Combination of autoregressive and moving average model is **ARIMA(p, d, q):**

$$y'_t = \boxed{ß_0 + ß_1 y'_{t-1} + \cdots + ß_p y'_{t-p}} + \boxed{\theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t}$$

Day	Share price	Share price (t-1)	Share price (t-2)	etc.
01.01.2022	50	-	-	.
02.01.2022	52	50	-	
03.01.2022	53.5	52	50	
04.01.2022	52.75	53.5	52	

Dependent variable

czechitas

ARIMA

Combination of autoregressive and moving average model is **ARIMA(p, d, q):**

$$y'_t = \underbrace{\boxed{ß_0 + ß_1 y'_{t-1} + \cdots + ß_p y'_{t-p}}}_{\text{AR (p)}} + \underbrace{\boxed{\theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t}}_{\text{MA (q)}}$$

ATTENTION! y'_t is the differenced series

p = Number of autoregressive terms

d = degree of differencing involved

q = Number of moving average terms

ARIMA

Combination of autoregressive and moving average model is **ARIMA(p, d, q):**

$$y'_t = \underbrace{\boxed{\beta_0 + \beta_1 y'_{t-1} + \cdots + \beta_p y'_{t-p}}}_{\text{AR (p)}} + \underbrace{\boxed{\theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t}}_{\text{MA (q)}}$$

ATTENTION! y'_t is the differenced series

p = Number of autoregressive terms

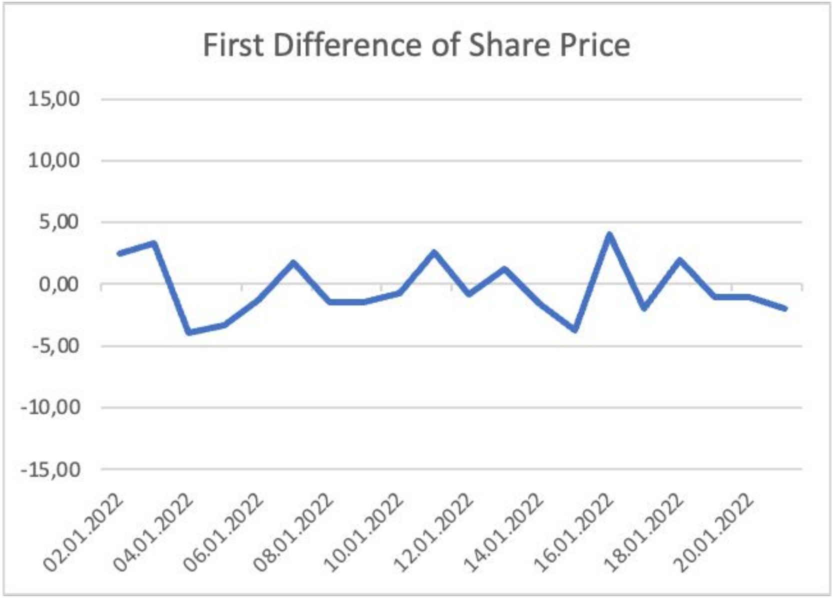
d = degree of differencing involved

q = Number of moving average terms

ARIMA - First order differencing

Date	Share Price	Lagged Price	First Difference
01.01.2022	100.00	-	-
02.01.2022	102.50	100.00	2.5
03.01.2022	105.75	102.50	3.25
04.01.2022	101.80	105.75	-3.95
05.01.2022	98.50	101.80	-3.30

ARIMA - First order differencing



czechitas

Public

ARIMA - Second order differencing

Date	Share Price	Lagged Price	First Difference	Second Difference
01.01.2022	100.00	-	-	-
02.01.2022	102.50	100.00	2.5	-
03.01.2022	105.75	102.50	3.25	0.75
04.01.2022	101.80	105.75	-3.95	-7.20
05.01.2022	98.50	101.80	-3.30	0.65

We difference until we get stationary series.

Stationarity can be tested by **Augmented Dickey-Fuller** test.

ARIMA

Combination of autoregressive and moving average model is **ARIMA(p, d, q):**

$$y'_t = \underbrace{\boxed{ß_0 + ß_1 y'_{t-1} + \cdots + ß_p y'_{t-p}}}_{\text{AR (p)}} + \underbrace{\boxed{\theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t}}_{\text{MA (q)}}$$

where y'_t is the differenced series.

p = Number of autoregressive terms

d = Degree of differencing involved

q = Number of moving average terms

But how do we decide which p and q we should use??

ARIMA: Information Criteria

How well does the model explain the data, without being unnecessarily complex?

Akaike's information criterion (AIC) $= -2\log(L) + 2(p + q + k + 1)$
Bayesian information criterion (BIC) $= \text{AIC} + (\log(n) - 2)(p + q + k + 1)$

→ We choose p and q so that **information criteria are minimized.**

Quiz

Which statement is **true**?

a) ARIMA has autoregressive and moving average components, both can have multiple lags.

a) ARIMA works well also with non-stationary series.

a) When we are choosing ARIMA specification, we maximize Information Criteria.

czechitas

ARIMA PRACTICAL EXAMPLE DEMO

FORECASTING WITH EXPLANATORY VARIABLES

ARIMA

We were forecasting stock prices using only stocks’ past values and residuals:

$$y'_t = ß_0 + ß_1 y'_{t-1} + \cdots + ß_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

ARIMAX (AutoRegressive Integrated Moving Average with eXogenous variables)

We can add more than just dependent variable to the model:

$$Sales_t = c + \varphi * Sales_{t-1} + \theta * \varepsilon_{t-1} + \beta_1 * AdvertisingCost_t + \varepsilon_t$$

ARIMAX (AutoRegressive Integrated Moving Average with eXogenous variables)

More than just dependent variable in the model:

$$Sales_t = c + \varphi * Sales_{t-1} + \theta * \varepsilon_{t-1} + \beta_1 * AdvertisingCost_t + \varepsilon_t$$

Strengths

- Increased **accuracy**
- **Ability to explain** what is **driving** forecast

czechitas

ARIMAX (AutoRegressive Integrated Moving Average with eXogenous variables)

More than just dependent variable in the model:

$$Sales_{t} = c + \varphi * Sales_{t-1} + \theta * \varepsilon_{t-1} + \beta_1 * AdvertisingCost_{t} + \varepsilon_t$$

Strengths

- Increased **accuracy**
- **Ability to explain** what is **driving** forecast

Weaknesses

- **Need to have prediction for external variables** (!!!)
- **Hard to find** factors influencing forecasting variables and gather all the data

EXTRA: PROPHET METHOD

PROPHET

- an **additive model** - decomposes series into the **trend** component, **seasonal** component and **holidays** component

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

- works well with missing data and outliers

PROPHET

1. **Automatic Seasonality Detection:**

   It can handle multiple seasonal patterns (daily, weekly, monthly, yearly)
   → suitable for series with complex and overlapping seasonal effects.

1. **Flexible Trend Modeling:**

   It can capture both linear and non-linear trends
   → effective for series with trend shifts or irregular patterns.

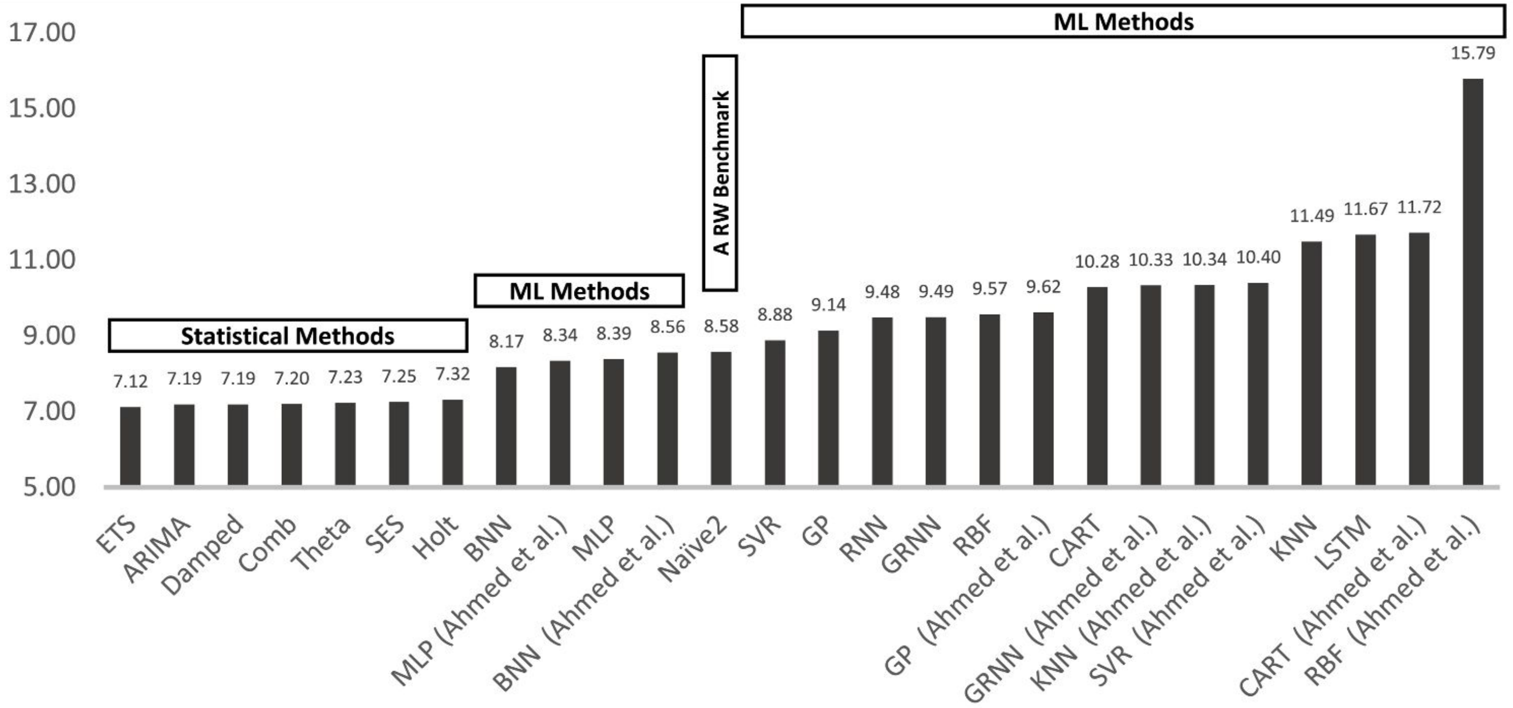
1. **Holiday Effects:**

   It has an option to include custom-defined holiday events
   → beneficial for series affected by recurring or one-time events.

czechitas

FORECASTING MODEL EVALUATION

Statistical methods vs. ML methods performance



Forecasting performance (sMAPE) of the ML and statistical methods included in the study.
The results are reported for one-step-ahead forecasts having applied the most appropriate preprocessing alternative. 1045 monthly time series used in the M3 Competition. Link to the article

Train and Test sets: principles

Train set

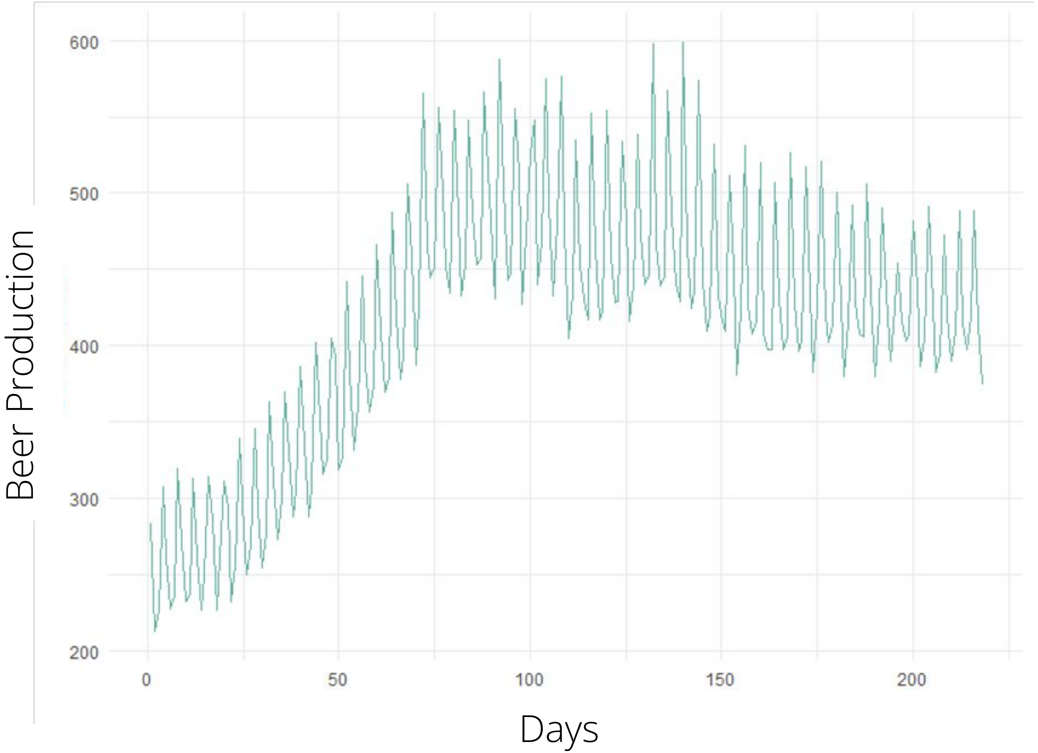
- Long **enough to catch seasonality** if there is one
- Some models might need more data to train than others

Test set

- It should be the **size of your forecasting horizon**

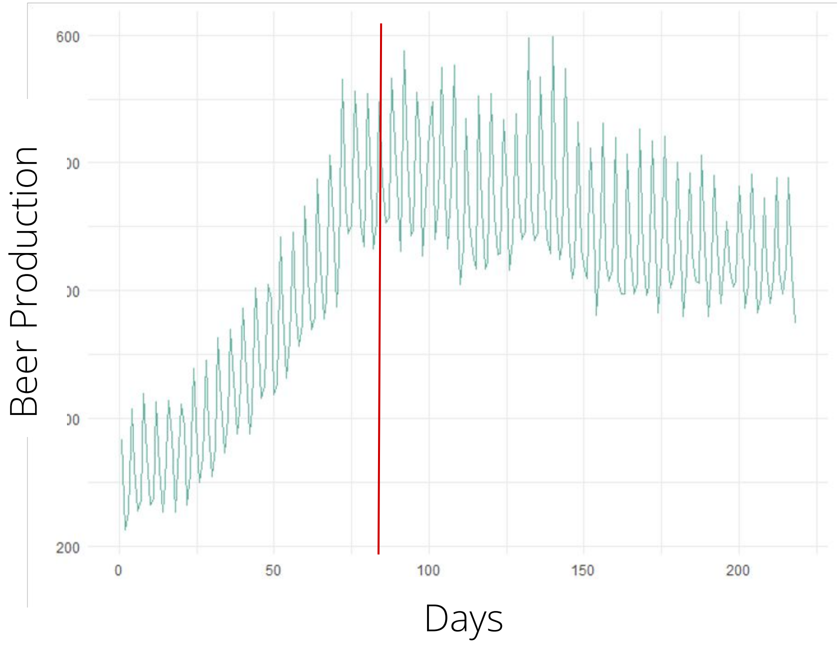
czechitas

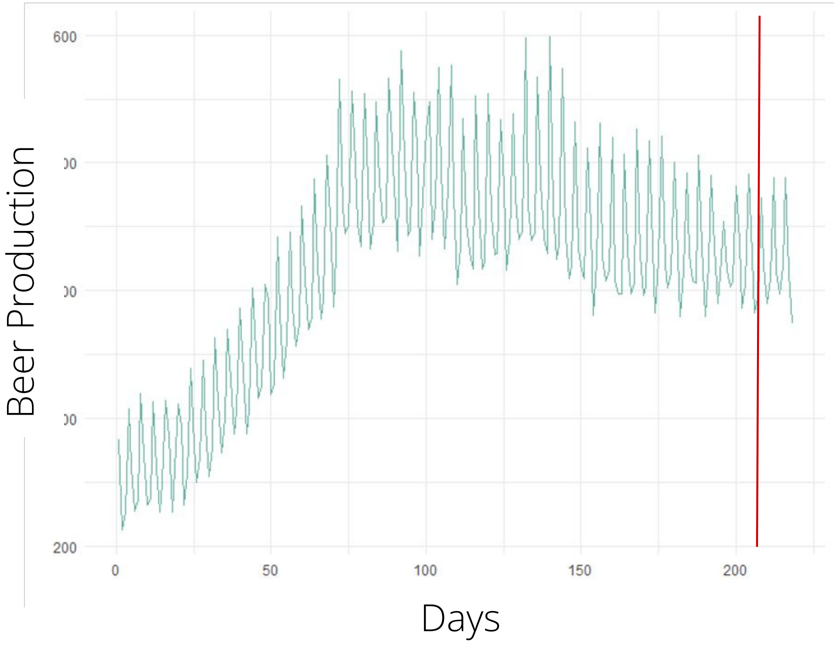
Q: We want 30 days forecast. How to split train/test?



czechitas

Q: We want 30 days forecast. How to split train/test?

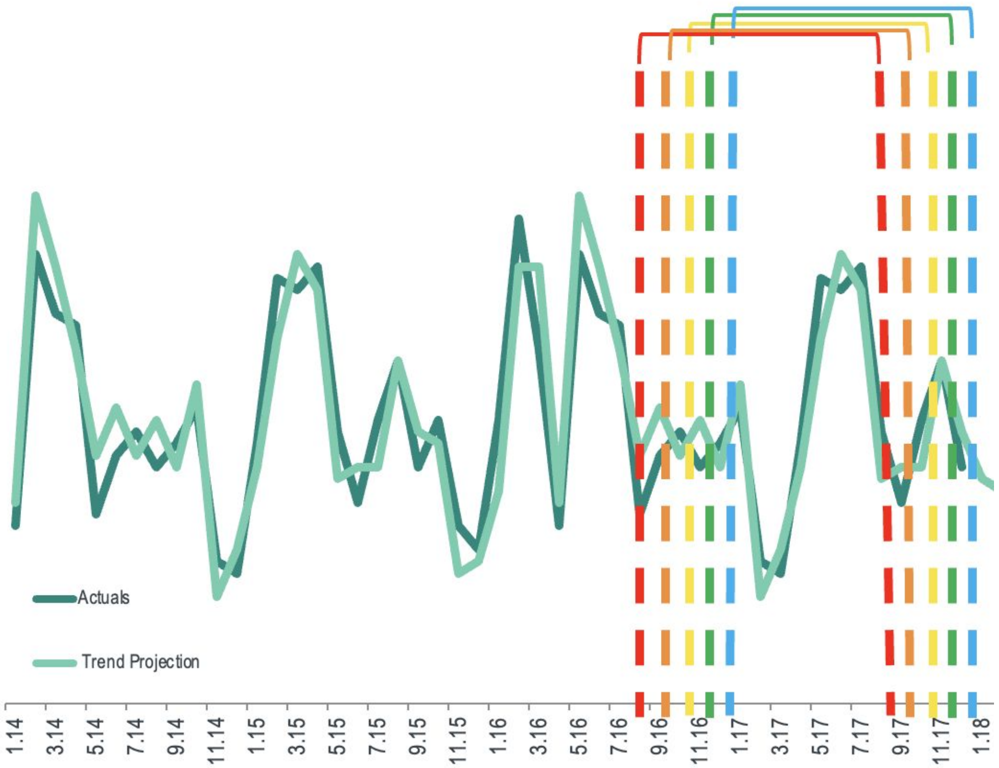




X Not catching the recent trend

X Not long enough testing window

Is our validation general enough?



Only one test sample

→ result might be accidentally good or bad.

Compare different models: metrics

Mean absolute error – average size of errors.

$$\text{MAE} = \frac{1}{n}\sum_{j=1}^{n} |y_j - y_j|$$

Mean absolute percentage error – percentage average errors. Easy to explain to business.

$$\text{MAPE} = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

RMSE, AIC, BIC, …

Compare different models

Comparing models among each other

Try ARIMA, ETS, Prophet, etc.
→ Then see which one gives the lowest error.

Comparing to naive forecast

How much model is better than assuming that yesterday's value will hold tomorrow.

czechitas

Example: Compare MAE for discussed methods

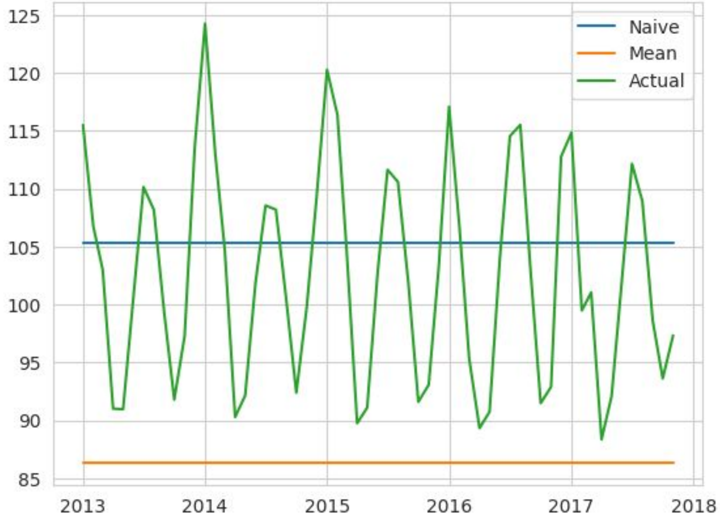
Model	MAE
ARIMA	3,476
ETS	3,214
Naive	8,222
Snaive	3,665
Simple mean	16,282

$$MAE = \frac{\sum_{i=1}^{n} |y_{T+i} - \hat{y}_{T+i|T}|}{n}$$

T - number of observations in training dataset
n - number of observations in testing dataset

Question: Why is naïve and simple mean performing so poorly?

a. Naïve and mean ignore trend
b. Naïve ignores seasonality and mean ignores seasonality and trend
c. Just a coincidence



Quiz

Which statement is **true**?

a) If we have relatively long data history, it is not useful to try simple forecasting methods (mean, last observation, etc.).

a) Validation with multiple train/test splits is computationally costly and should be rather avoided.

a) Test set length should be at least as long as our forecasting horizon.

CASE STUDIES

Overview

Many Time Series to forecast

ETS
AUTO ARIMA
ML

Without exogenous

ETS
AUTO ARIMA
ML

Without exogenous

Need for automation

One time need to forecast

Forecast needed regularly

ARIMA/ETS with manual setup
ARIMAX
OLS
ML

With exogenous

ARIMA/ETS with manual setup
ARIMAX
OLS
ML

With exogenous

Need for automation

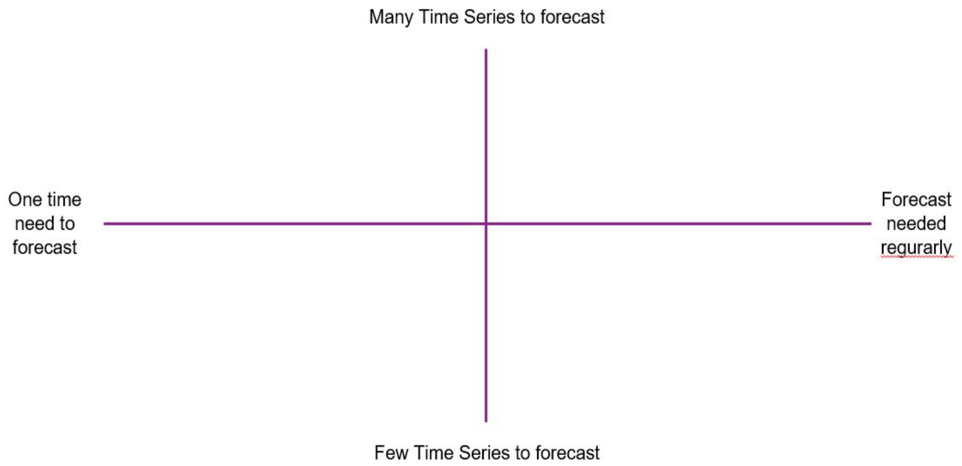
Few Time Series to forecast

czechitas

Case study 1: Czech National Bank – GDP forecasting

Description

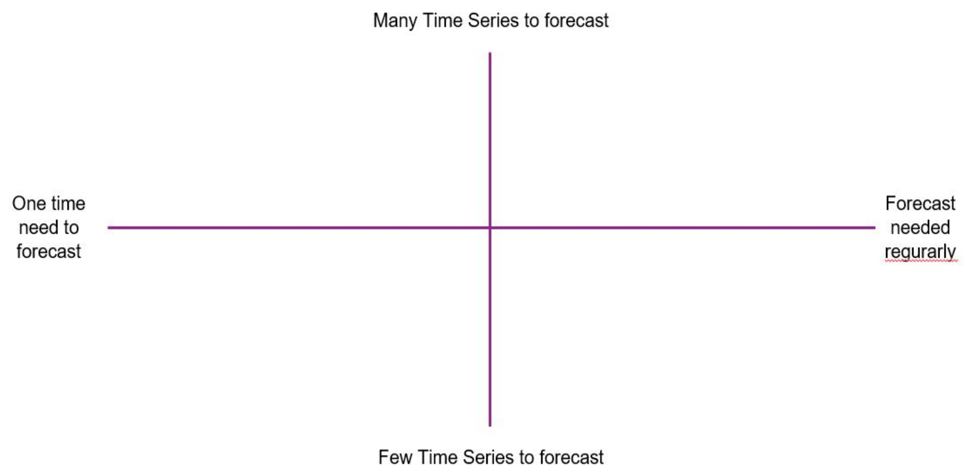
- Delivered Quarterly
- Need to know what are the factors driving GDP growth
- Need to know what are the causes of changes from last quarter



Case study 2: Demand for Animal Health drugs in MSD

Description

- Delivered monthly
- 1000+ products to be forecasted
- Need to deliver forecast in couple days after data release



SUMMARY

Time Series Summary

Concepts: Stationarity, decomposition: trend, seasonality, cycle, residuals

Think about: feasibility, data availability, horizon

Methods: simple (mean, etc.), ETS, ARIMA, Prophet

Evaluation: train/test, metrics (MAE, AIC, ..)

czechitas

Thank you for your attention!