

13.06.23

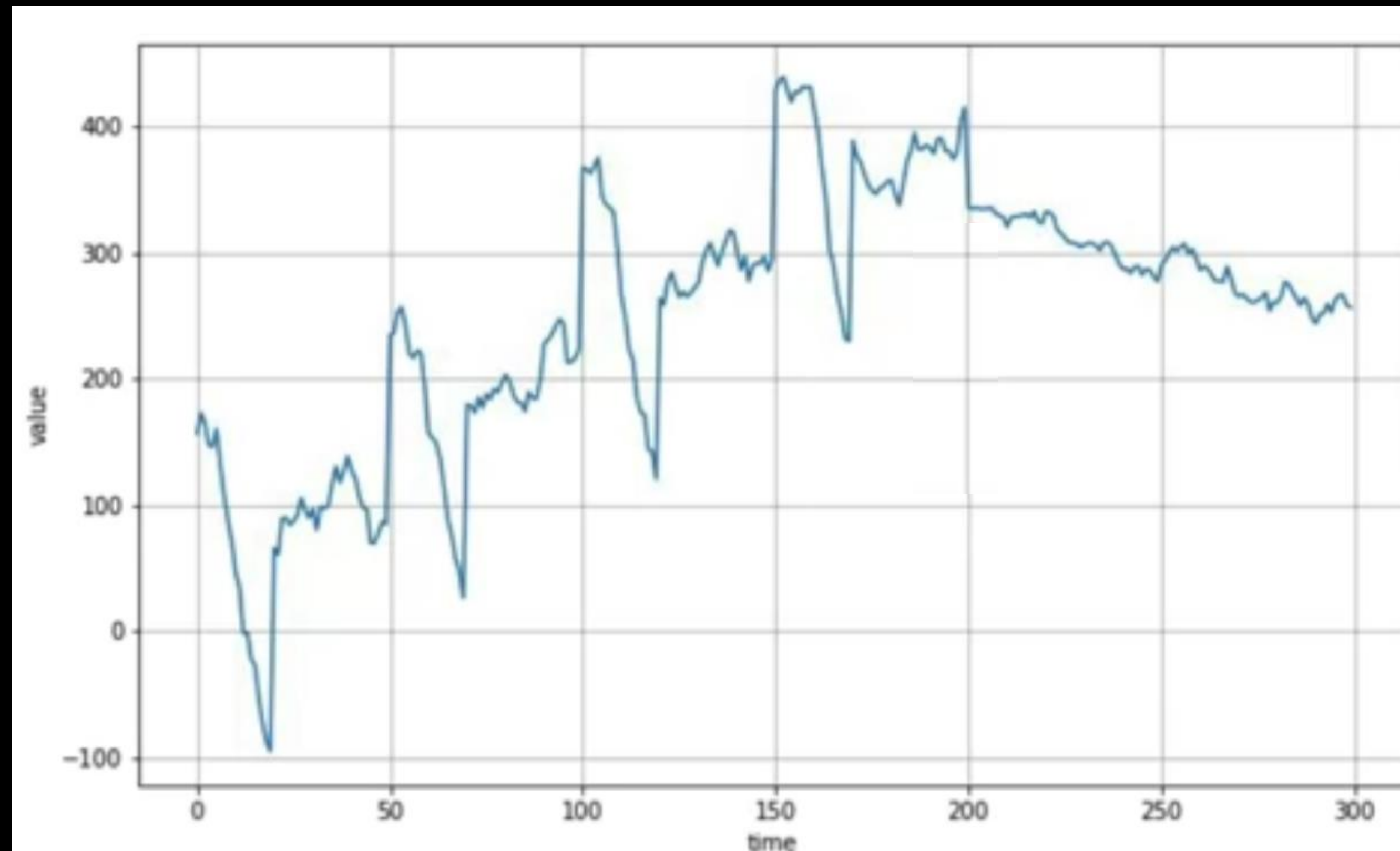
Einführung in Data Science und maschinelles Lernen

ZEITREIHENANALYSEN

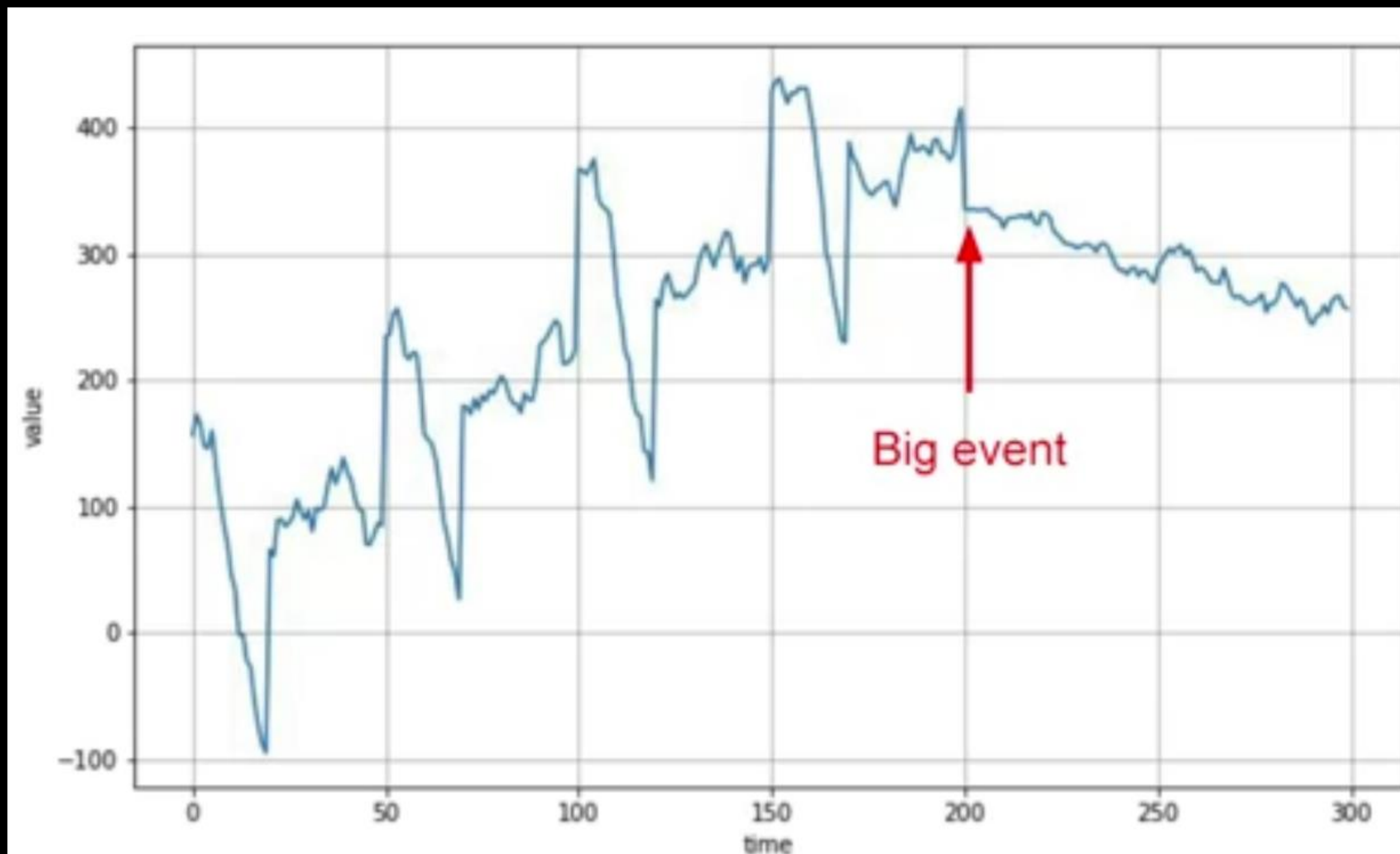
- **Muster in Zeitreihenanalysen**
- **Non-Stationarity**
- **Baseline Modelle und Naïve Forecasting**
- **Projektpräsentation**
- ***Und jetzt?***

DISKUSSION

- Welche verschiedenen Arten von Mustern kann man in der dargestellten Abbildung erkennen?



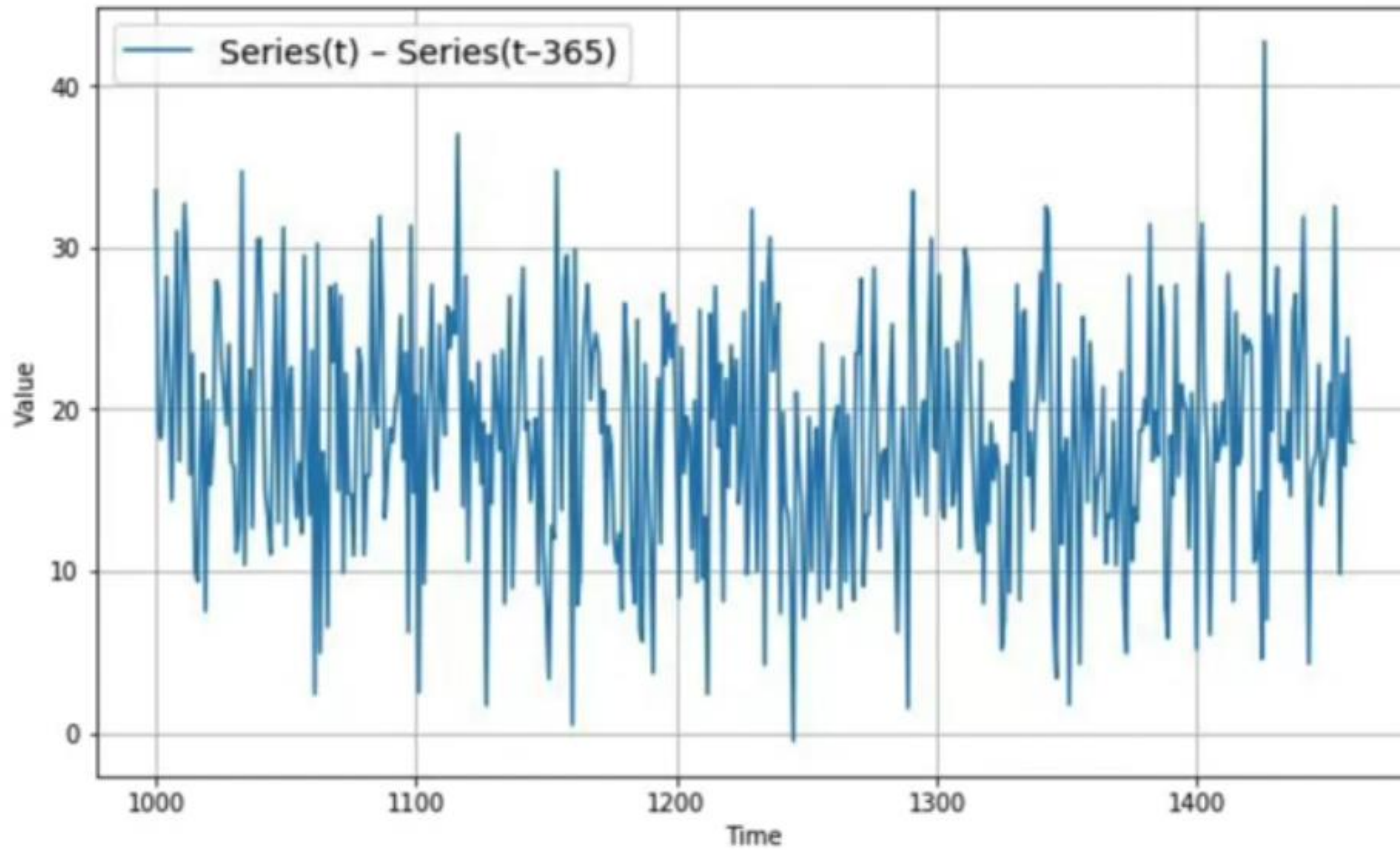
NON-STATIONARITY



MUSTER IN ZEITREIHENANALYSEN

- **Trends**
- **„Jahresgang“ (Seasonality)**
- **Rauschen (Noise)**
- **Autokorrelation**

DIFFERENCING



DIFFERENCING

- **Subtraktion der vorherigen Beobachtung von der aktuellen Beobachtung**
- **Methode um die Zeitabhängigkeit von Zeitreihen zu entfernen bzw. zu mindern.**
- **Versuch „stationäre“ Zeitreihen zu erhalten**

LAG-DIFFERENZ

- **Die Differenz zwischen aufeinanderfolgenden Beobachtungen wird als Lag-1-Differenz bezeichnet.**
- **Die Lag-Differenz kann an die spezifische zeitliche Struktur angepasst werden.**
- **Bei Zeitreihen mit einer saisonalen Komponente kann man davon ausgehen, dass die Verzögerung der Periode (Breite) der Saisonalität entspricht.**

BEISPIELAUSWERTUNGEN

```
1
2 ▾ #####
3 ▸ ### Preparation of the Environment #####
14 ▾ #####
15 ### Reading the data file
16 ▸ #####
23 ▾ #####
24 ### Prepare data
25 ▸ #####
30 ▾ #####
31 # Pedestrians hourly
32
33 # Basic plot
34 ggplot(pedestrians_hourly) +
35   geom_line(aes(x=datetime, y=`pedestrians count`), color="#69b3a2") +
36   xlab("") +
37   theme_ipsum() +
38   theme(axis.text.x=element_text(angle=60, hjust=1))
39
40 # Time frame specific plot
41 ggplot(pedestrians_hourly) +
42   geom_line(aes(x=datetime, y=`pedestrians count`), color="#69b3a2") + |
43   xlab("") +
44   theme_ipsum() +
45   theme(axis.text.x=element_text(angle=60, hjust=1)) +
46   scale_x_datetime(limit=c(as.POSIXct("2021-10-01"), as.POSIXct("2021-11-01")))
47
48
49 ▾ #####
50 # Pedestrians daily
51
52 # Basic plot
53 ggplot(pedestrians_daily) +
54   geom_line(aes(x=date, y=`pedestrians count`), color="#69b3a2") +
55   xlab("") +
```

BASELINE MODELLE

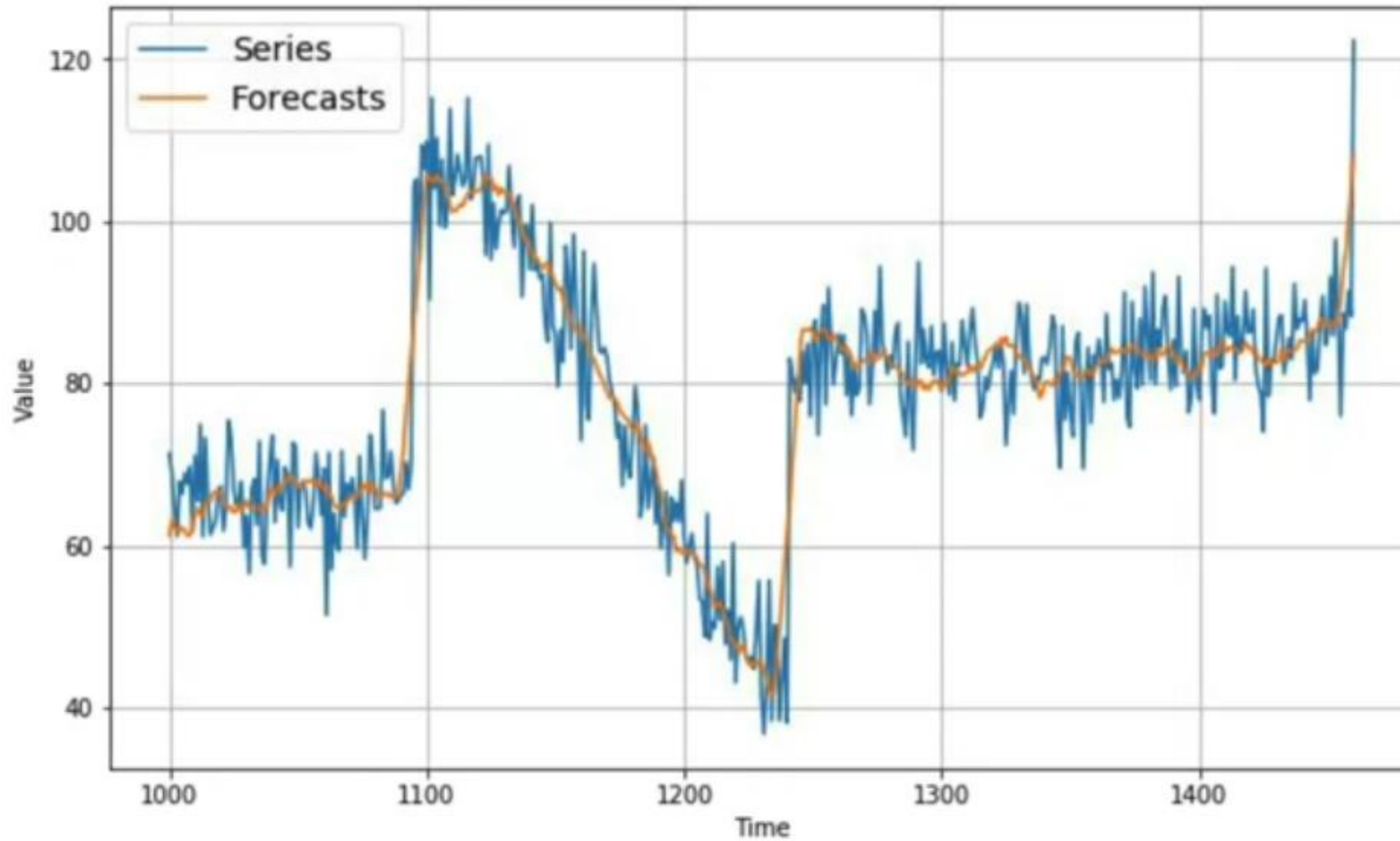
- **Allein an Metriken wie MAPE oder RMSE kann man häufig schlecht abschätzen, wie viel das eigene Modell gelernt hat.**
- **Es ist daher wichtig, die Ergebnisse anderer Modelle als „Baseline“ bzw. Referenz zu nutzen.**

Mögliche Baselines:

- **Ergebnisse bisher genutzter Modelle für den gleichen Datensatz**
- **Ergebnisse von Modellen auf artverwandten Datensätzen**
- **Speziell bei Zeitreihen: Ergebnisse basierend auf Naïve Forecasting**

NAÏVE FORECASTING

- **Mögliches Baseline Modell für Zeitreihenanalysen**
- **Vorhersage entspricht dem jeweils letzten beobachteten Wert**
- **Saisonal Naïve Forecasting:**
Vorhersage entspricht dem letzten Wert mit der gleichen Saisonalität.



Forecasts = trailing moving average of differenced series + centered moving average of past series ($t - 365$)

BEISPIELBERECHNUNG VON LAG-DATEN

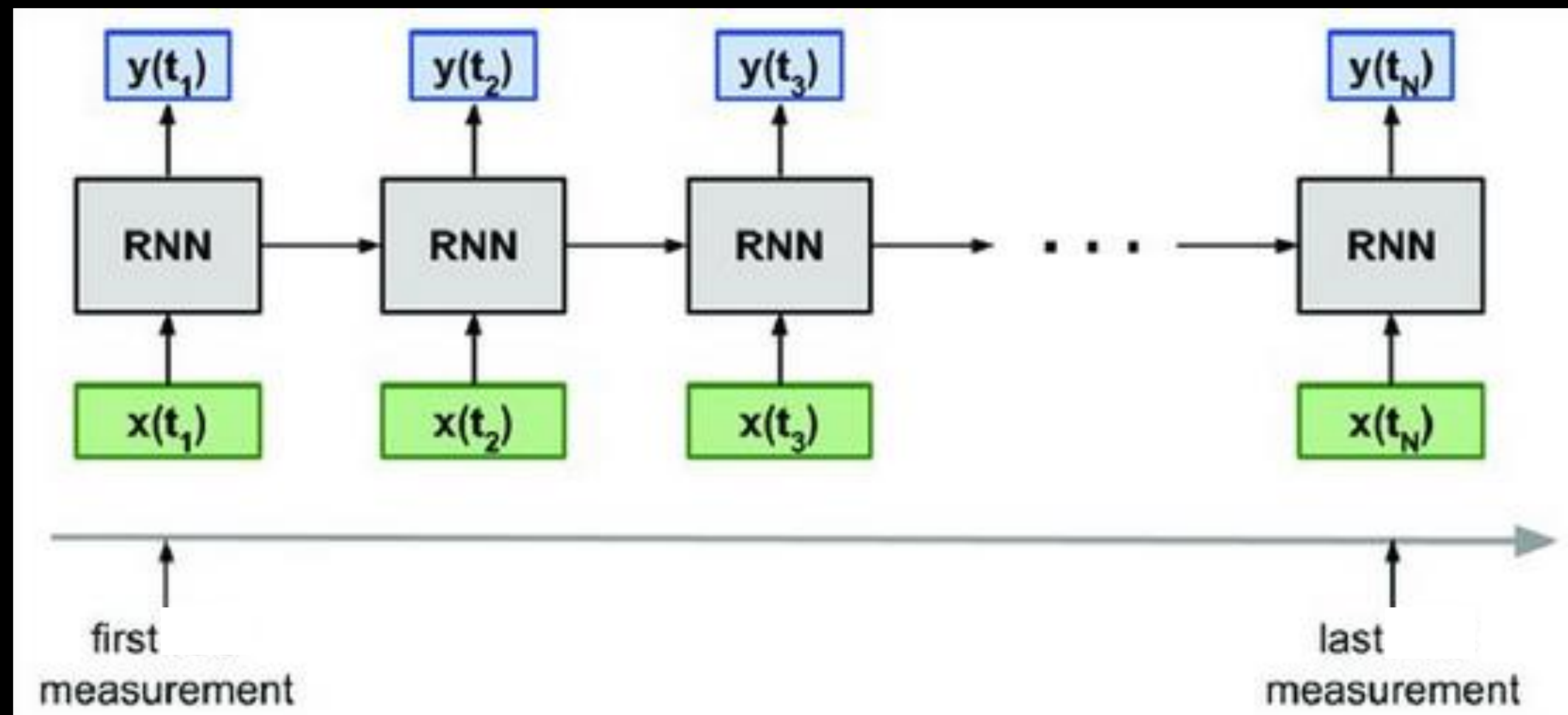
```
1 library(ggplot2)
2 library(dplyr)
3 
4 # Create some example data
5 ts_data <- data.frame(date = seq(from = as.Date("2022-01-01"), to = as.Date("2022-12-31"), by = "day"),
6   |               |       product = 1,
7   |               | value = rnorm(365, mean = 100, sd = 10))
8 
9 # Plot the data using ggplot
10 ggplot(ts_data, aes(x = date, y = value)) +
11   geom_line() +
12   ggtitle("Time Series Data") +
13   xlab("Date") +
14   ylab("Value")
15 
16 # Add variable including the value of the day before
17 ts_data_with_lag <- ts_data %>%
18   arrange(date) %>%
19   mutate(value_prev_day = lag(value, default = NA))
20 
21 
22 # Example data, in which several values (labels) are given for each day
23 multiple_ts_data <- ts_data %>%
24   rbind(data.frame(date = seq(from = as.Date("2022-01-01"), to = as.Date("2022-12-31"), by = "day"),
25     |               |       product=2,
26     |               | value = rnorm(365, mean = 100, sd = 10)))
27 
28 
29 multiple_ts_data_with_lag <- multiple_ts_data %>%
30   arrange(date, product) %>%
31   group_by(product) %>%
32   mutate(value_prev_within_day = lag(value, default = NA)) %>%
33   ungroup()
```

BREAKOUT

- **Versucht den Vorjahreswert als unabhängige Variable zu Eurem Modell hinzuzufügen.**
- **Welche weiteren Variablen basierend auf vergangenen Umsätzen lassen in den Datensatz einfügen?**

RECURRENT NEURAL NETS

- **Neurales Netz bei dem die Knoten einer Schicht untereinander verknüpft sind.**



INHALT DER PROJEKTPRÄSENTATION

- **Dauer: ca. 8 Minuten pro Team**
- **Eure Namen auf der Titelseite**
- **Auflistung und kurze Beschreibung der selbst erstellten Variablen**
- **Balkendiagramme mit Konfidenzintervallen für zwei selbst erstellte Variablen**
- **Optimierung eines linearen Modells:**
 - (a) Modellgleichung**
 - (b) Adjusted r^2**
 - (c) MAPE für den Zeitraum vom 9.6.2019 bis 30.7.2019**
- **Optimierung eines neuronalen Netzes:**
 - (a) Source Code zur Definition des neuronalen Netzes**
 - (b) Darstellung der Loss-Funktionen für Trainings- und Validierungsdatensatz**
 - (c) MAPE für den Zeitraum vom 9.6.2019 bis 30.7.2019**

HINWEISE

- **Ihr könnt ein R-Markdown Notebook für die Präsentation nutzen, aus dem Ihr z.B. eine Powerpoint generiert.**
- **Styling und Dokumentierung des Codes mithilfe von ChatGPT.**
- **Berichtet von dem Fehler, der Euch am meisten Zeit gekostet hat, oder der Sache, die ihr mit viel Zeitaufwand versucht habt, die aber nicht geklappt hat.**

R and OOP anti-patterns

June 9, 2023 | Bob Carpenter

Thomas Lumley just dropped a blog post, Blank cheque inheritance and statistical objects, which begins as follows. One of the problems with object-oriented programming for statistical methods is that inheritance is backwards. Everything is fine for data structures, and Bioconductor ... Continue reading → [Read more...]

'Advanced Shiny Development' the hands-on workshop

June 9, 2023 | Mirai Solutions

Best practices for a robust and maintainable shiny app, a hand on workshop on 21/06. Do you know how to build a basic Shiny web application, but would you like to bring your Shiny development to the next level? Learn from professional experts how to...

[Read more...]



Building a basic Shiny app with Golem – Part I (Video)

June 7, 2023 | pacha.dev/blog

R and Shiny Training: If you find this blog to be interesting, please note that I offer personalized and group-based training sessions that may be reserved through Buy me a Coffee. Additionally, I provide training services in the Spanish language ... [Read more...]

Understanding the file.info() Function in R: Listing Files by

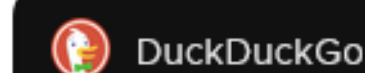
Search R-bloggers..

Go

Your e-mail here

Subscribe

52793 readers



DuckDuckGo hat diesen Inhalt blockiert, um zu verhindern, dass Facebook dich trackt

Wir haben Facebook daran gehindert, dich zu tracken, als die Seite geladen wurde. Wenn du die Blockierung für diesen Inhalt aufhebst, kennt Facebook deine Aktivitäten. [Mehr erfahren](#)

Blockierung aufheben

Most viewed posts (weekly)

How to improve your storytelling with R
PCA vs Autoencoders for Dimensionality Reduction

How to install (and update!) R and RStudio
Update to Data Science Software Popularity
Why GLMs should be a priority when teaching statistics



KI in der Praxis

6. Juni 2023

Apple spricht nicht über KI, packt sie aber trotzdem überall rein



DEEP MINDS Podcast

Podcast über Künstliche
Intelligenz und Wissenschaft



Künstliche Intelligenz und Robotik | DEEP MINDS #15

Verfügbar bei [Youtube](#), [Soundcloud](#),
[Spotify](#), [Apple](#), [Google](#) und [Amazon](#)



KI-Forschung

KI und Gesellschaft

KI in der Praxis

[Write](#)

★ Get unlimited access to all of Medium. [Become a member](#)



Machine Learning

[Follow](#)[Start writing](#)

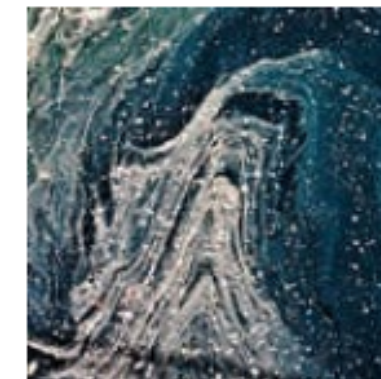
Trending Latest Best



Merve Noyan · 17 hours ago

Complete Guide on Deep Learning Architectures Part 2: Autoencoders

Autoencoder: Basic Ideas Autoencoder is the type of a neural network that reconstructs an input from the output. The basic idea here is tha...



Machine Learning 5 min read



Bex T.in Towards Data Science · 18 hours ago ★ Member-only

10 Confusing XGBoost Hyperparameters and How to Tune Them Like a Pro in 2023

XGBoost hyperparameters done with style and visuals — Intro Today, I



251K

Stories

92K

Writers



Related Topics

[Data Science](#)[Artificial Intelligence](#)[Deep Learning](#)[Python](#)[AI](#)[Technology](#)[Programming](#)[NLP](#)[Neural Networks](#)

[See more topics](#)

Top Writers



The PyCoach

10M+ Views on Medium ||

Make money by writing about...

[Follow](#)



r/MachineLearning



r/MachineLearning

Reddit durchsuchen

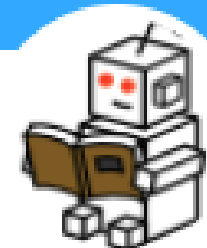


Gratis



matanzino

1 Karma



Machine Learning

r/MachineLearning

Beigetreten



Beiträge



Beitrag erstellen



Heiß



Neu



Top



6



VON EINEM MOD ANGEHEFTET

Gepostet von u/AutoModerator vor 1 Tag

Discussion

[D] Simple Questions Thread



20 Kommentare



Auszeichnen



Teilen



Merken



4



Gepostet von u/ML_WAYR_bot vor 22 Stunden

Discussion

[D] Machine Learning - WAYR (What Are You Reading) - Week 128



1 Kommentar



Auszeichnen



Teilen



Merken



120



Gepostet von u/jayalammar vor 5 Stunden

Research

[R] The Illustrated Retrieval Transformer (GPT3 performance at 4% the size)

Hi [r/MachineLearning](#),

Über diese Community



Willkommen in MachineLearning

2.2m

Mitglieder

488

Online



Am 29. Juli 2009 erstellt

Beitrag erstellen

COMMUNITY-EINSTELLUNGEN



Nach Flair filtern

Discussion

Research

Project

Twitter

@clashML

LINKEDIN

- **Philipp Schmid**
- **Lior Sinclair**
- ...



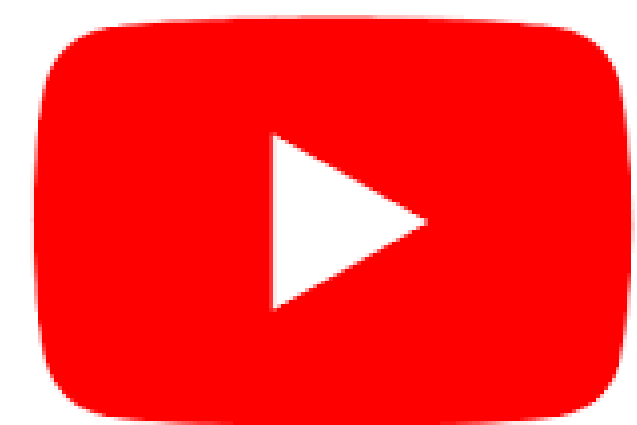
coursera



UDACITY



Udemy



YouTube



KI-Campus

Die Lernplattform
für Künstliche Intelligenz

{ MACHINE # LEARNING > DEGREE } BY OPENCAMPUS.SH |

GET A SOLID UNDERSTANDING OF
MACHINE LEARNING AND LEARN
HOW TO IMPLEMENT YOUR OWN
STATE OF THE ART MACHINE
LEARNING PROJECTS.

→ JOIN OUR COURSES



GET READY TO REALIZE YOUR
OWN STATE OF THE ART
MACHINE LEARNING PROJECTS



BECOME PART OF NETWORK OF
MACHINE LEARNING
ENTHUSIASTS



COMBINE WORLD CLASS
LEARNING CONTENT WITH
IMPLEMENTING PROJECTS

Machine Learning with TensorFlow

DIENSTAG 16:00 - 17:45

Get hands-on experience in applying machine learning techniques with TensorFlow.

Die Bewerbungsfrist ist leider abgelaufen.

Du wirst lernen

- ✓ Best Practices für TensorFlow, ein populäres Open-Source-Framework für maschinelles Lernen, um neuronale Netzwerke zu trainieren
- ✓ Umgang mit Bilddaten aus der realen Welt und Erkundung von Strategien zur Vermeidung von Overfit, einschließlich Augmentation und Drop-Out
- ✓ Erstellung eines Systems zur Verarbeitung natürlicher Sprache
- ✓ Anwendung von RNNs, GRUs und LSTMs zum Training dieser Lernmodelle unter Verwendung von Text- und Zeitreihendaten

DIENSTAG

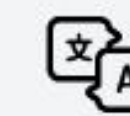
16:00 -
17:45



ONLINE +
KIEL

ECTS

5



ENGLISCH