

17.04.25

Introduction to Data Science and Machine Learning

INTRODUCTION

- **Introduction**
- **Organizational information**
- **Course structure**
- **What is Data Science?**
- **Tools for programming**

PERSONAL INTRODUCTION



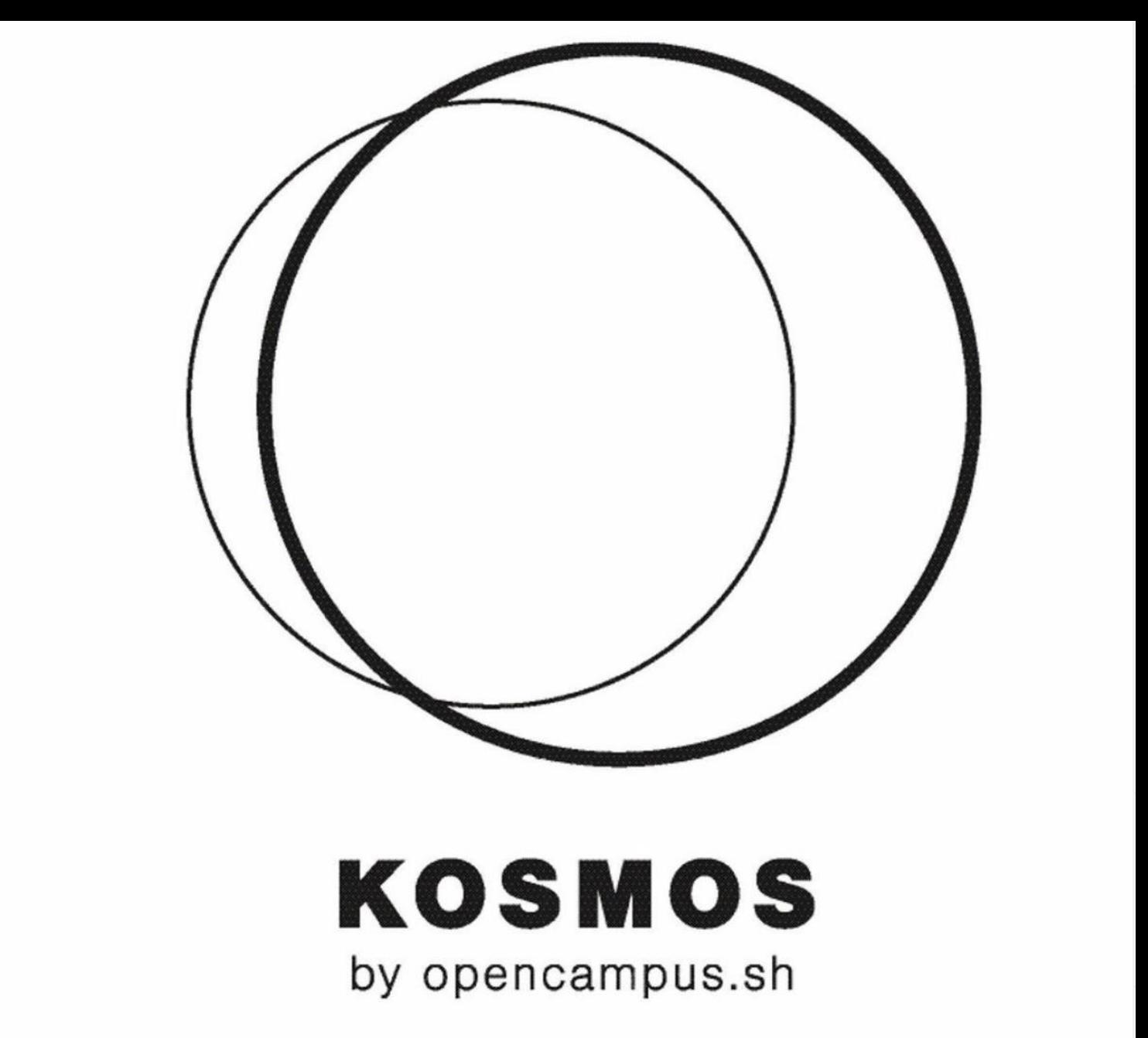
- Nonprofit organization which oversees a variety of initiatives
- Offering a wide range of educational opportunities, support, and networking for entrepreneurs, creatives, and anyone curious, regardless of age, educational background, or origin
- The services are open to everyone and mostly free.
- The goal is to support the entrepreneurial landscape, promote creative change processes, and contribute to innovative and sustainable future development.





 starter
kitchen.de







FABLAB KIEL



Social
Entrepreneurship
Education





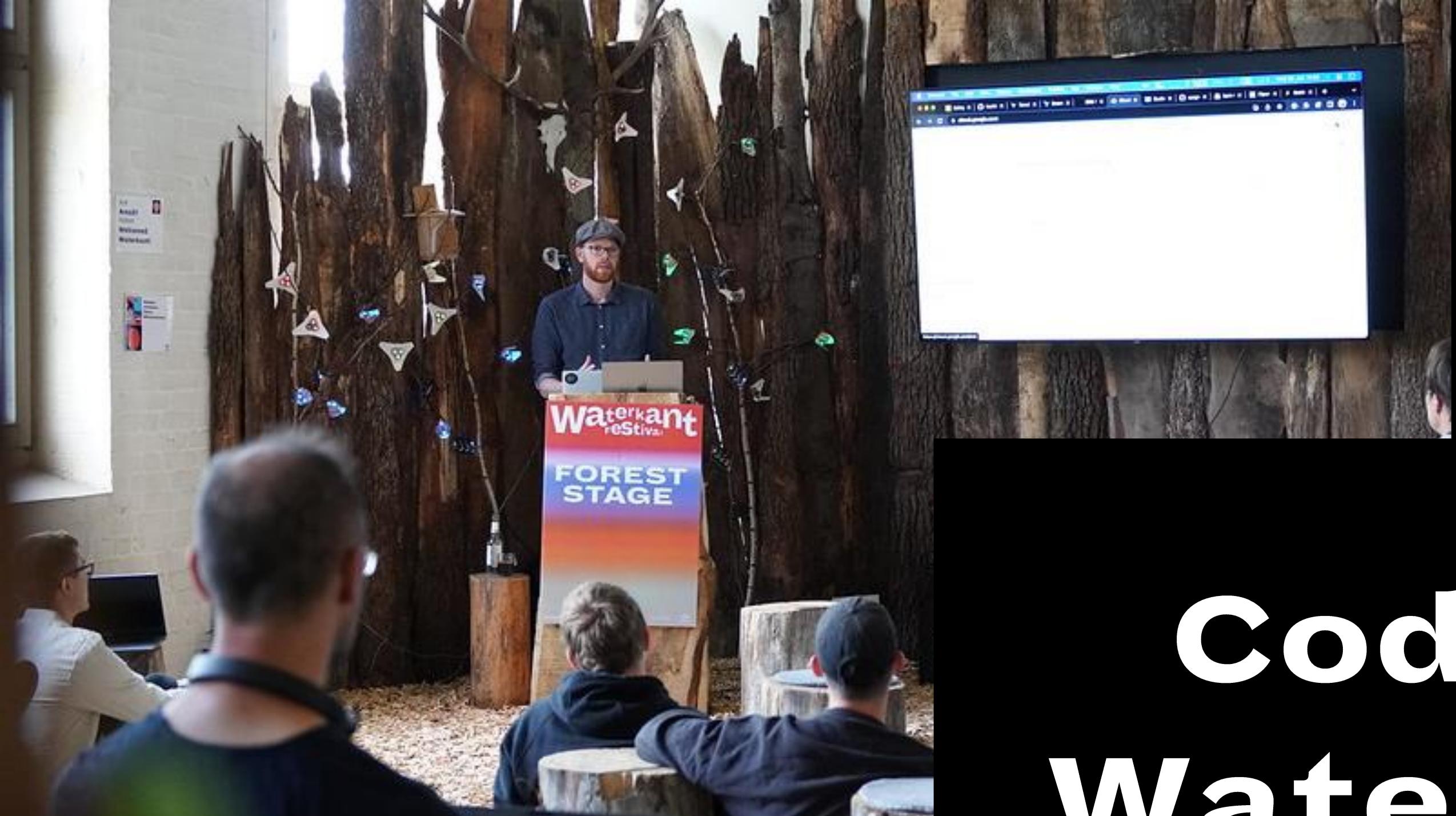
Diving into AI and Machine Learning at the Water's Edge.

steffen@opencampus.sh

SUBSCRIBE

<https://coding.waterkant.sh>

Coding. Waterkant 2025



Coding.Waterkant

2025

July 07 - 11

- **Work for four consecutive days:**
 - **on your own machine learning project or**
 - **support others in their project**
- **Take part in workshops**
- **Get input and feedback by invited experts**
- **Present your work to a larger audience.**
- **Take educational leave**
- **Get accommodation on site**

Im Rahmen der ständigen Verbesserungen unserer Plattform haben wir die E-Mail-Adresse deiner Gruppe in coding-waterkant-announce@messages.meetup.com geändert. Du kannst dies in den optionalen Einstellungen deiner Gruppe ansehen.

X



Coding.Waterkant

★★★★★ (126) [?](#)

Neu: Event-Feedbackübersicht

Tippe auf die Sterne, um alle deine Eventbewertungen in der Übersicht anzuzeigen.

[Weitere Informationen](#)

📍 Kiel, Deutschland

👤 981 Mitglieder · Öffentliche Gruppe [i](#)

👤 Organisiert von **opencampus.sh** and **6 others**

Teilen:

[Info](#)[Events](#)[Mitglieder](#)[Fotos](#)[Diskussionen](#)[Mehr](#)[Event erstellen](#)[Gruppe verwalten](#)

Über uns

Our meetup is organized with support of opencampus.sh and the Digitale Wirtschaft Schleswig-Holstein (DiWiSH).

[Mehr lesen](#)



Organizers



opencampus.sh and 6 others

[Nachricht](#)

Members (981)

[Alles ansehen](#)



ORGANIZATIONAL MATTERS

Attendance Registration:

Online

- **Use your full names in the zoom meetings!**
- **Turn your camera on!**

Presence

- **Scan the QR-Code if you participate in presence**

Please write us if you will not go on with the course!

MATTERMOST CHAT

- **Complete your profile in the Mattermost chat with your full name and a photo.**
- **Please, always ask any questions to us in the course chat!!**

MATTERMOST CHAT

The screenshot shows the Mattermost chat interface for the channel **C_Machine Learning With TensorFlow**. A yellow arrow points to the pinned post at the top of the channel page.

Pinned Post: **C_Machine Learning With TensorFlow** (March 25)
Steffen Brandt 23:10
Welcome to the course "Machine Learning With TensorFlow"! In this course we will try to provide you with hands-on knowledge about how to train machine learning models with TensorFlow. An important part when working in the field of machine learning is networking and working together in a team. An important goal of the course is therefore that you get to know each other and work in a team on a project. I would therefore like to ask you to introduce yourself quickly here in the channel already. Maybe

Channel Header: C_Machine Learning With TensorFlow (Tuesday, 4-6 p.m.: Zoom; Course Handbook)

Left Sidebar: sose21 @steffen, Find channel, ANNELS, 00 - Announcements, 01 - Questions, C_Advanced Machine Lear..., C_Deep Learning from Scr..., C_Einführung in Data Scie..., C_Machine Learning für di..., C_Machine Learning With ..., Kursleitungen.

Right Sidebar: Pinned Posts, Search, @, ⌂, March 25, C_Machine Learning With ...

COURSE HANDBOOK



opencampus.sh Machine
Learning Program

EDU-Platform

Chat

Search...

opencampus.sh Machine Learning
Program

Course Kick-Off

How do I choose a course?

FAQ

COURSES

Einführung in Data Science und
maschinelles Lernen

Bedingungen für ein
Leistungszertifikat oder ECTS

Vorbereitung

Woche 1 - Einführung in Data
Science

Woche 2 - Grafische Darstellung
von Daten

Woche 3 - Versionierung mit git
(Teil 1) und Datenaufbereitung

Woche 1 - Einführung in Data Science

Diese Woche werdet Ihr...

eine Einführung zu den folgenden Themen bekommen:

- Was ist Data Science?
- R vs. Python vs. SPSS vs. ...
- Wozu RStudio?
- Datenstrukturen in R

Lernressourcen



221025_Einführung.pdf 6MB
PDF

ZOOM

- Try the different viewing modes:
 - Gallery View/ Active Speaker
 - Split Screen/ Full Screen Mode
- Maybe watch this video to get an idea:
<https://www.youtube.com/watch?v=v3IPAbpVjd4>

The screenshot shows the Zoom Meeting interface with the Settings window open. The Settings window has a sidebar with icons for General, Video, Audio, Share Screen, Chat, Background & Filters, Recording, Profile, Statistics, Keyboard Shortcuts, and Accessibility. The 'Share Screen' option is highlighted with a blue bar and a yellow arrow pointing to it from the left.

Window size when screen sharing:

- Fullscreen mode
- Maximize window
- Maintain current size

When I share my screen in a meeting

- Automatically share desktop
- Show all sharing options

When I share directly to a Zoom Room

- Automatically share desktop
- Show all sharing options

Advanced

Video Controls (Bottom Left):

- Steffen Brandt
- Select a Camera (Alt+N to switch)
 - NewTek NDI Video
 - ✓ Integrated Camera
- Choose Virtual Background...
- Choose Video Filter...
- Video Settings... (highlighted with a blue bar and a yellow arrow)

Meeting Controls (Bottom Bar):

- Mute
- Start Video (highlighted with a blue bar and a yellow arrow)
- Security
- Participants (1)

BREAKOUTS

For Online Participants:

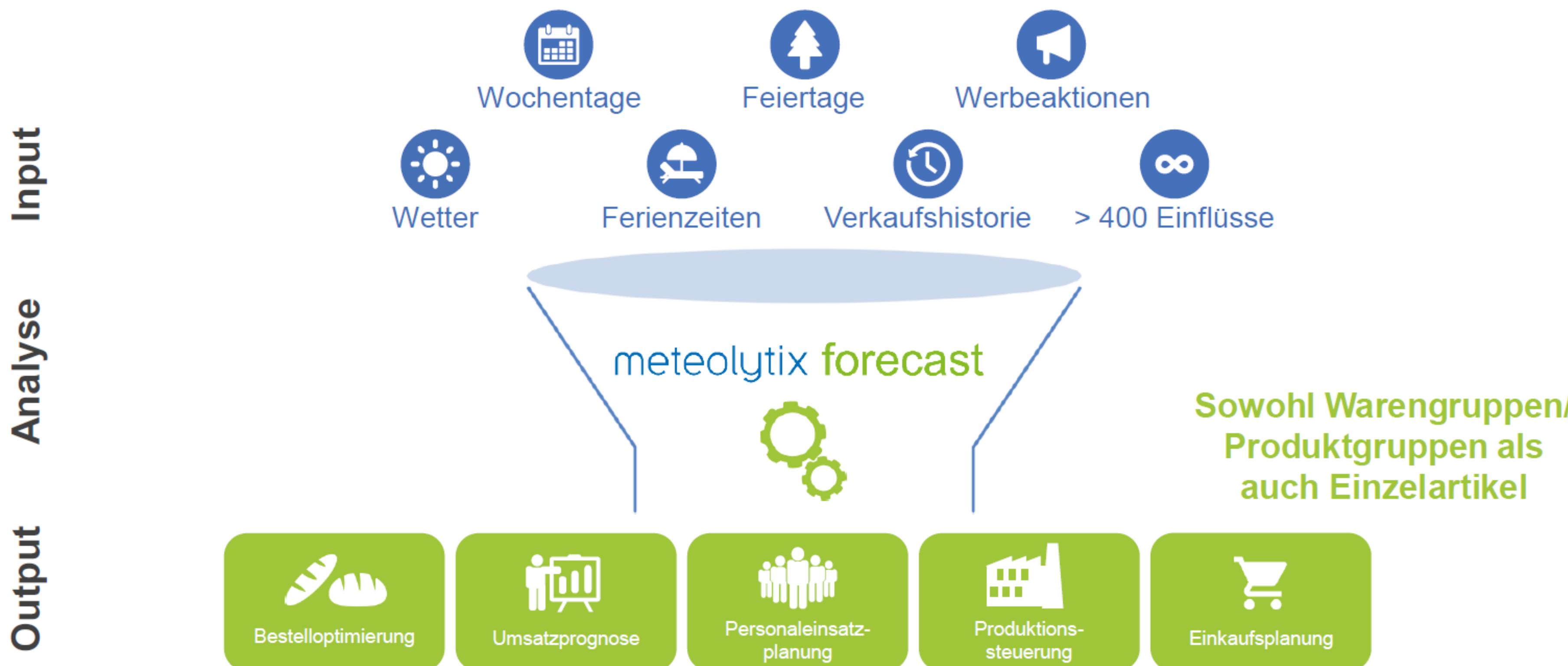
**Take a screenshot of the task before joining
breakouts!**

SALES FORECAST PROJECT

- **Real sales data from various product groups of a bakery branch over a period of approximately 3 years**
- **Weather data for the corresponding period (and beyond)**

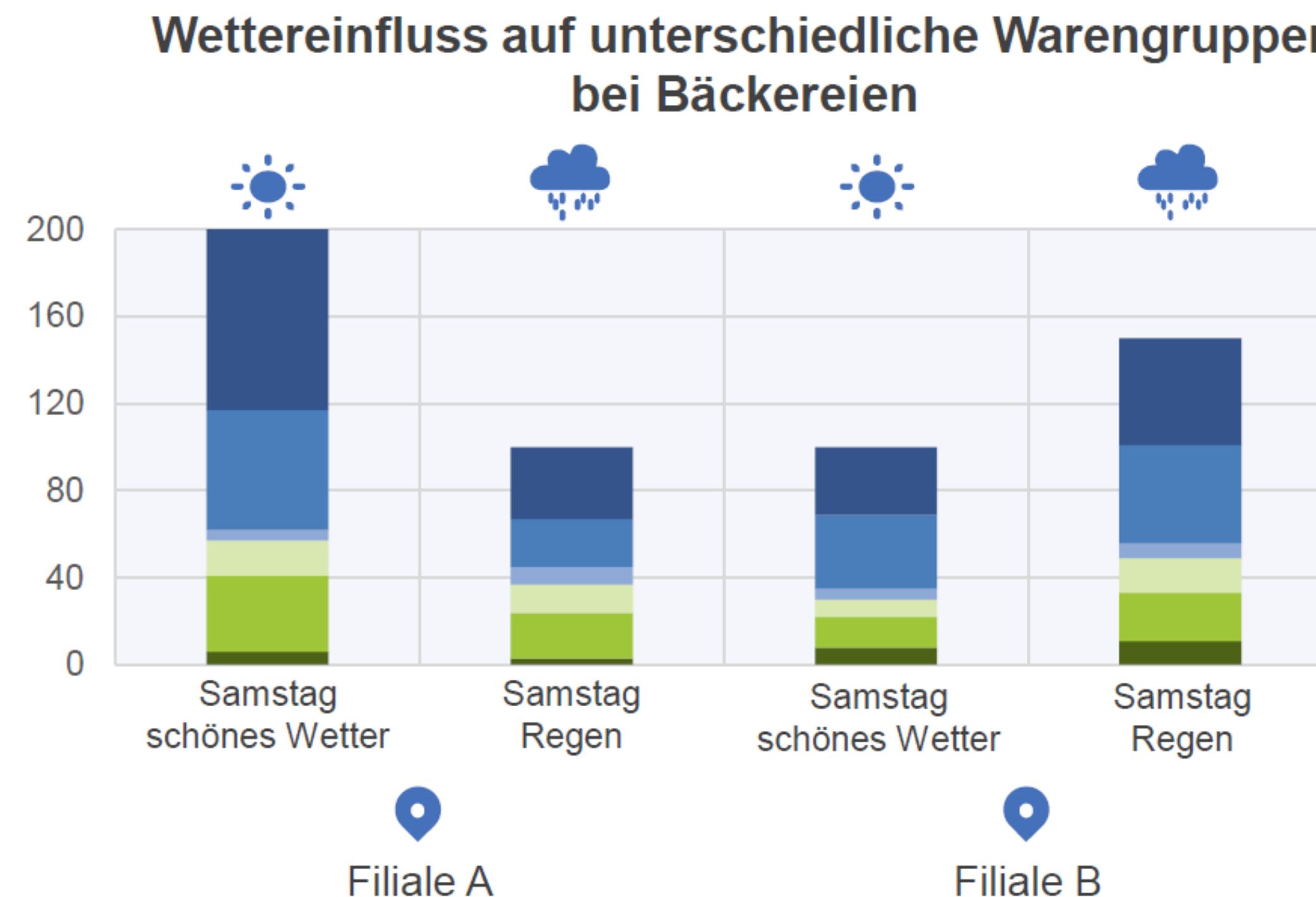
meteolytix forecast analysiert die Datenzusammenhänge von mehr als 400 Einflussfaktoren und liefert Absatzprognosen für viele Einsatzfelder.

WAS WIR MACHEN



Die Stärke des Wettereffekts variiert von Ort zu Ort und wird jeweils filialindividuell berücksichtigt.

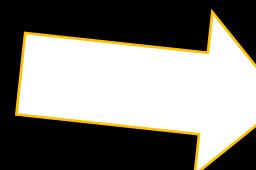
WAS WIR MACHEN



PROJECT GOALS

- **Programming with AI support in VS Code**
- **Working in a team with version control (Git)**
- **Preparation of variables**
- **Use of graphical representations**
- **Use of regressions to identify relevant variables**
- **Handling missing values**
- **Optimization of a neural network**

Definition of the Teams



17.04.2025	Introduction to Data Science
16:00 - 18:00	Starterkitchen, Kuhnkestr. 6, 24118 Kiel + ONLINE
24.04.2025	Data Visualization
16:00 - 18:00	Starterkitchen, Kuhnkestr. 6, 24118 Kiel + ONLINE
08.05.2025	Version Control with Git and GitHub
16:00 - 18:00	Starterkitchen, Kuhnkestr. 6, 24118 Kiel + ONLINE
15.05.2025	Data Preparation
16:00 - 18:00	Starterkitchen, Kuhnkestr. 6, 24118 Kiel + ONLINE
22.05.2025	Overfitting and Model Evaluation
16:00 - 18:00	Starterkitchen, Kuhnkestr. 6, 24118 Kiel + ONLINE
05.06.2025	Introduction to Machine Learning
16:00 - 18:00	Starterkitchen, Kuhnkestr. 6, 24118 Kiel + ONLINE
12.06.2025	Neural Networks
16:00 - 18:00	Starterkitchen, Kuhnkestr. 6, 24118 Kiel + ONLINE
19.06.2025	Handling Missing Values
16:00 - 18:00	Starterkitchen, Kuhnkestr. 6, 24118 Kiel + ONLINE
26.06.2025	Time Series Analysis
16:00 - 18:00	Starterkitchen, Kuhnkestr. 6, 24118 Kiel + ONLINE
03.07.2025	Project Presentations
16:00 - 18:00	Starterkitchen, Kuhnkestr. 6, 24118 Kiel + ONLINE

**„Education is the only thing that humans try to
get less out of than what they pay for.“**

FIRST BREAKOUT

- **15 Minutes**
- **Present yourself**
- **Discussion Question:**
How is Artificial Intelligence defined?
 - **Try to agree on a definition**
 - **Share your definition in the course channel in Mattermost**

WHAT IS AI?

- Chess computer
- Google Maps
- ChatGPT

“What's the difference between data science, machine learning, and artificial intelligence?”

<http://varianceexplained.org/r/ds-ml-ai>

Data science produces insights.

Machine learning produces predictions.

Artificial intelligence produces actions.

DATA SCIENCE

Application of statistical analysis methods to generate results for a report or article.

- **Descriptive statistics**
- **Statistical inferences**
- **Data visualization**
- **Design of experiments**

MACHINE LEARNING

The use of statistical optimization techniques to estimate model parameters for a prediction function validated with unseen data.

Typically, model performance is continuously improved by incorporating new data, allowing the system to adapt and refine its predictions over time.

- Predicting future states
- Inferring unobserved states
- Estimating current states

ARTIFICIAL INTELLIGENCE

„an autonomous agent executes or recommends actions“

(Poole, Mackworth, & Goebel, 1998)

„Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals.“

(European Commission, 2018)

„By Artificial Intelligence, we understand highly developed software systems that are capable of learning and can be trained to handle complex tasks.“

(AI Strategy of Schleswig-Holstein, 2019)

TOOLS TO CREATE AND EXECUTE CODE

- Programming language + function libraries
- Execution environment
- Development environment

PROGRAMMING LANGUAGE: PYTHON

- **Open source, huge community**
 - large amounts of existing example code
 - very good AI support
- **Industry and (slowly getting) academia standard for data analysis**
- **Primary ML language**
- **Fastest-growing language with wide application range including productive settings**
- **Extensive function libraries for:**
 - **Web Applications (z.B. Django und Flask)**
 - **Data Analysis (NumPy, Pandas)**
 - **Visualizations (Matplotlib, Seaborn)**
 - **Neural Networks (TensorFlow, PyTorch)**

PYTHON CHARACTERISTICS

- **Interpreter language (executes line-by-line, similar to a calculator)**
- **Console-based (no GUI)**
- **Evolved from software engineering (vs. R from statistics)**

EXECUTION ENVIRONMENT:

Jupyter Notebook:

- Browser-based execution of code via a GUI**
- Popular in ML and Data Science**

Python Script:

- Execution of code from a text file**
- Used particularly in production**

File Edit View Insert Cell Kernel Help

Cell Toolbar: None

Simple spectral analysis

An illustration of the [Discrete Fourier Transform](#) using windowing, to reveal the frequency content of a sound signal.

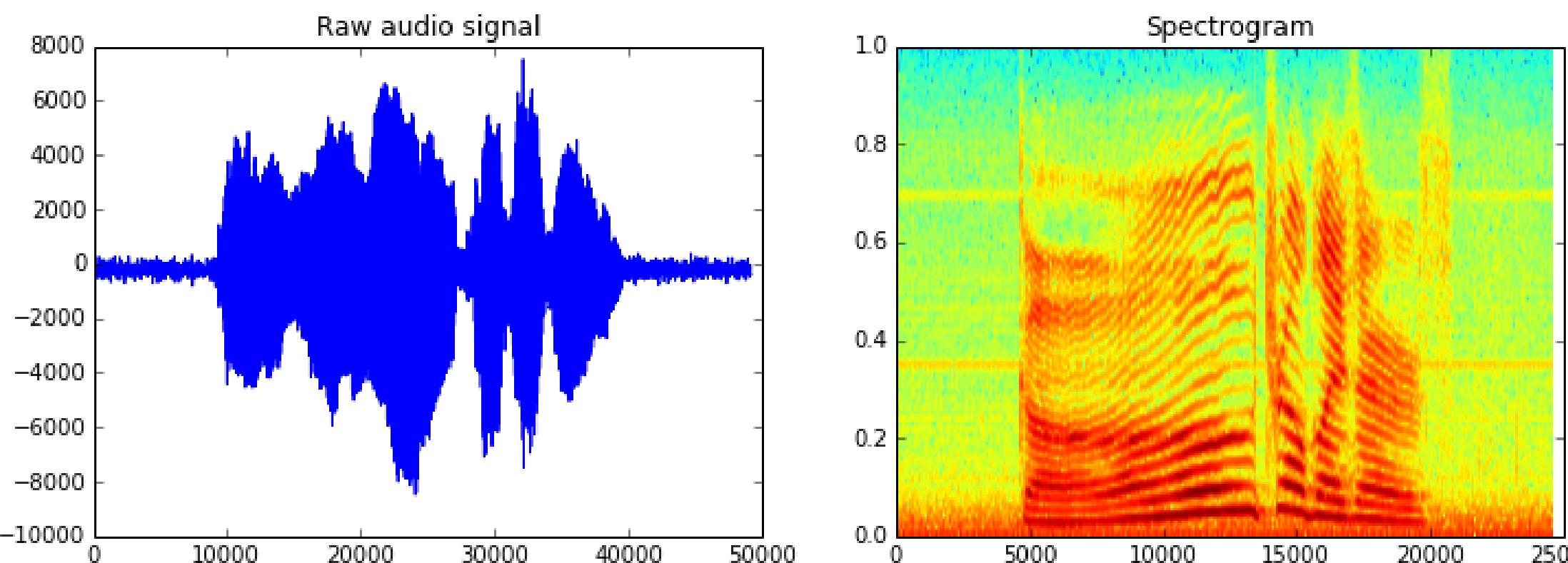
$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} \quad k = 0, \dots, N - 1$$

We begin by loading a datafile using SciPy's audio file support:

```
In [1]: from scipy.io import wavfile  
rate, x = wavfile.read('test_mono.wav')
```

And we can easily view its spectral structure using matplotlib's builtin specgram routine:

```
In [2]: %matplotlib inline  
from matplotlib import pyplot as plt  
fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(12, 4))  
ax1.plot(x); ax1.set_title('Raw audio signal')  
ax2.specgram(x); ax2.set_title('Spectrogram');
```



COLAB

<https://colab.research.google.com/>

The screenshot shows the Google Colaboratory interface. The top navigation bar includes the Colab logo, the title "Willkommen bei Colaboratory", and menu items: Datei, Bearbeiten, Anzeige, Einfügen, Laufzeit, Tools, and Hilfe. On the far right are buttons for Teilen, Einstellungen, and Anmelden. The left sidebar, titled "Inhalt", lists categories: Erste Schritte, Data Science, Maschinelles Lernen, Weitere Ressourcen, and Vorgestellte Beispiele. A button "+ Bereich" is located at the bottom of this sidebar. The main content area displays the text "Willkommen bei Colab!" and "(Neu) Gemini API testen". Below this, a bulleted list provides links to: Generate a Gemini API key, Talk to Gemini with the Speech-to-Text API, Gemini API: Quickstart with Python, Gemini API code sample, Compare Gemini with ChatGPT, and More notebooks. At the bottom of the content area, there is a descriptive text about Colab features and a video thumbnail titled "3 Cool Google Colab Features" featuring a man speaking.

Willkommen bei Colaboratory

Datei Bearbeiten Anzeige Einfügen Laufzeit Tools Hilfe

Inhalt

Erste Schritte

Data Science

Maschinelles Lernen

Weitere Ressourcen

Vorgestellte Beispiele

+ Bereich

Willkommen bei Colab!

(Neu) Gemini API testen

- [Generate a Gemini API key](#)
- [Talk to Gemini with the Speech-to-Text API](#)
- [Gemini API: Quickstart with Python](#)
- [Gemini API code sample](#)
- [Compare Gemini with ChatGPT](#)
- [More notebooks](#)

Wenn Sie Colab schon kennen, erfahren Sie in diesem Video mehr zu interaktiven Tabellen, zur Verlaufsansicht für ausgeführten Code und zur Befehlspalette.

3 Cool Google Colab Features

[]

CHARACTERISTICS OF JUPYTER NOTEBOOKS

- **Open Source application, run in a web browser or locally**
- **Convenient editing and execution of code in the form of code cells**
- **Display of results directly in the notebook below the executed code cell - including graphical output**
- **Insertion of formatted comments and explanations via Markdown cells**

MARKDOWN IN MATTERMOST

The screenshot shows the Mattermost desktop application. On the left, there's a sidebar with sections for '23W', '24S', 'COURSE ORGANIZATION', 'CHANNELS' (with a list item '24W | Python Programming...'), and 'DIRECT MESSAGES'. A '+' icon is also visible. The main area shows a message from user 'ph1uib-VfD_izZO2Y5sC3UdQE5x8XNnxUO1qJLaRUGL3qWeTjomUBn_ET6FuvDIHg9dZIKaEWwQU;tDcic6-g4ECMhWPIf90KIUpGUFAJEm8k&sa=X&ved=2ahUKEwibmtru6pqJAxVG9rsIHRMhKCoQjive&vld=cid:5c617867,vid:YVFa5VljCDY,st:0' containing a long URL. A floating rich text editor window is overlaid on the message, containing the following text:

****fett** oder *kursiv* oder als
Überschrift**

The rich text editor has a toolbar with icons for bold (B), italic (I), strikethrough (S), heading (H), a quote symbol (‘ ’), a double quote symbol (“ ”), a list icon (≡), a table icon (≡), and a help icon (ⓘ).

EXAMPLES FOR MARKDOWN SYNTAX

Markdown	Preview
bold text	bold text
<i>*italicized text*</i> or <u>_italicized text_</u>	<i>italicized text</i>
<code>`Monospace`</code>	Monospace
~~strikethrough~~	strikethrough
[A link](https://www.google.com)	A link
![An image](https://www.google.com/images/rss.png)	

DEVELOPMENT ENVIRONMENT: VISUAL STUDIO CODE (VS CODE)

- **Open-source development environment that supports the creation of programming code**
- **Standard development environment for numerous programming languages in software development**
- **Many extensions and additional features available via plugins**

TOOLS TO CREATE AND EXECUTE CODE

Software:

- **Programming Language**
→ Python und Python Packages
- **Execution Environment**
→ Python Environments or Jupyter Notebooks
- **Development Environment**
→ VS Code

TOOLS TO CREATE AND EXECUTE CODE

Hardware:

- Previously, typically only the local computer
- Due to increasing computational requirements, code is now often executed in the cloud
- Development environments are also increasingly used in the cloud

OUR APPROACH

- **Local installation of the VS Code development environment**
- **Execution of the code in an environment in the cloud using GitHub Codespaces**
- **Linking the local development environment with the cloud-based execution environment**

WHY GITHUB CODESPACES?

- **No local Python installation required**
→ Avoids potential compatibility and configuration issues across different operating systems
- **Consistent Execution Environments**
→ Everyone uses the same Python versions and code examples work the same for all
- **Integration with GitHub and GitHub Copilot**
→ No additional setup required
- **Live Sharing**
→ Share development environments for real-time collaboration
- **Automatic Backups**

EXAMPLE USAGE

- 1. Create a GitHub repository (a cloud-based project workspace)**
- 2. Add a GitHub Codespace to this repository**
- 3. Open the Codespace with your local VS Code**



steffen74

 Type / to search[Overview](#)[Repositories 16](#)[Projects](#)[Packages](#)[Stars 18](#)

Steffen Brandt

steffen74

[Edit profile](#)

23 followers · 1 following

<https://github.com/new> Find a repository...

Type ▾

Language ▾

Sort ▾

[New](#)

workshop-ressources Public

Example data sets and programm code for the workshop at the EALTA 2019

● Jupyter Notebook Updated on Jun 28

[Star](#)

GenAI-powered-Literature-Analysis Public

Forked from [ApaydinK/GenAI-powered-Literature-Analysis](#)

● Python Updated on Jun 14

[Star](#)

bakery_sales_prediction Public

Updated on May 9

[Star](#)

ConstitutionalAiTuning Public

Create a new repository

A repository contains all project files, including the revision history. Already have a project repository elsewhere?

[Import a repository.](#)

Required fields are marked with an asterisk ().*

Repository template

No template ▾

Start your repository with a template repository's contents.

Owner *



steffen74 ▾

Repository name *

test-repo



test-repo is available.

Great repository names are short and memorable. Need inspiration? How about [shiny-carnival](#) ?

Description (optional)

Public

Anyone on the internet can see this repository. You choose who can commit.

Private

You choose who can see and commit to this repository.

Initialize this repository with:

Add a README file

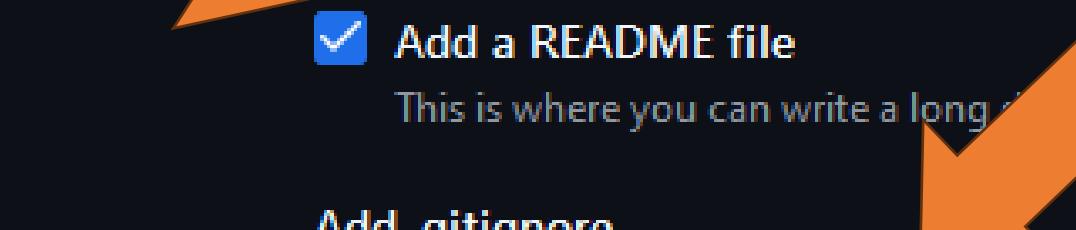
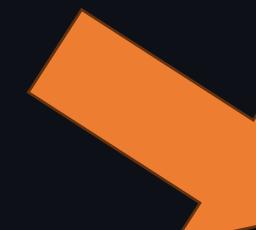
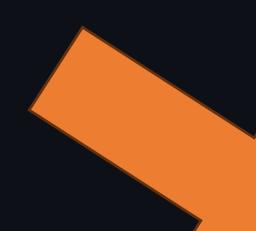
This is where you can write a long description for your project. [Learn more about READMEs.](#)

Add .gitignore

.gitignore template: Python ▾

Choose which files not to track from a list of templates. [Learn more about ignoring files.](#)

Choose a license



steffen74 / test-repo

Type / to search | + ⚙️ 🌐 🔍 📎 📧

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

test-repo Public

main 1 Branch Tags

Go to file Add file Code

Local Codespaces

Codespaces Your workspaces in the cloud +

No codespaces You don't have any codespaces with this repository checked out Create codespace on main Learn more about codespaces... Codespace usage for this repository is paid for by steffen74.

About

No description, website, or topics provided.

Readme Activity 0 stars 1 watching 0 forks

Releases No releases published Create a new release

Packages No packages published Publish your first package

© 2024 GitHub, Inc. Terms Privacy Security Status Docs Contact Manage cookies Do not share my personal information

steffen74 / test-repo

Type / to search | + ⚙️ 🌐 🔍 📎 📧

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

test-repo Public

main 1 Branch Tags

Go to file Add file Code

Local Codespaces

Codespaces Your workspaces in the cloud +

No codespaces You don't have any codespaces with this repository checked out Create codespace on main Learn more about codespaces... Codespace usage for this repository is paid for by steffen74.

About

No description, website, or topics provided.

Readme Activity 0 stars 1 watching 0 forks

Releases

No releases published Create a new release

Packages

No packages published Publish your first package

Pin Unwatch Fork 0 Star 0

test-repo

test-repo

© 2024 GitHub, Inc. Terms Privacy Security Status Docs Contact Manage cookies Do not share my personal information

The screenshot shows a GitHub repository page for 'test-repo'. The 'Code' tab is active. In the top right, there are buttons for 'Pin', 'Unwatch', 'Fork', and 'Star'. An orange arrow points from the 'Unwatch' button to the 'Create codespace on main' button in the 'Codespaces' sidebar. Another orange arrow points from the 'About' section to the 'No codespaces' message in the sidebar. The sidebar also includes sections for 'Local', 'Codespaces', 'Readme', 'Activity', '0 stars', '1 watching', '0 forks', 'Releases', and 'Packages'.

steffen74 / test-repo

Type / to search | + ⚙️ 🌐 🔍 📧

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

test-repo Public

main 1 Branch Tags

Go to file Add file Code

Local Codespaces

Codespaces Your workspaces in the cloud +

On current branch

turbo xylophone main No changes Active •••

Codespace usage for this repository is paid for by steffen74.

About

No description, website, or topics provided.

Readme Activity 0 stars 1 watching 0 forks

Releases No releases published Create a new release

Packages No packages published Publish your first package

test-repo

© 2024 GitHub, Inc. Terms Privacy Security Status Docs Contact Manage cookies Do not share my personal information

steffen74 / test-repo

Type / to search | + ⚙️ 🌐 🔍 📎 📧

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

test-repo Public

main 1 Branch Tags

Go to file Add file Code About

Local Codespaces

Codespaces Your workspaces in the cloud +

On current branch

turbo xylophone main No changes Active ...

Codespace usage for this repository is paid for by steffen74

No description, website, or topics provided.

Readme Activity 0 stars 1 watching

Rename Export changes to a branch Change machine type Stop codespace

Auto-delete codespace

Open in Browser Open in Visual Studio Code Open in JetBrains Gateway Beta Open in JupyterLab Beta Delete

Do not share my pe

© 2024 GitHub, Inc. Terms Privacy Security Status Docs Contact Manage cookies

The screenshot shows a GitHub repository page for 'test-repo'. A context menu is open over a 'Codespaces' entry. Three orange arrows highlight specific items: 1) The 'Unwatch' button in the top right corner of the header. 2) The 'Active' status indicator next to the 'turbo xylophone' entry in the Codespaces list. 3) The 'Delete' option at the bottom of the expanded menu.

BREAKOUT

Calculate the mean of the following seven numbers using Python:

1

2

3

4

5

6

9,87654321

SUPPORT TOOLS

- **Code Completion**
(GitHub Copilot, Supermaven / Cursor Tab, ...)
- **Specific Code Assistants**
(GitHub Copilot, Cursor Composer, Qodo Gen, ...)
- **Chatbots**
(Claude, Gemini, ChatGPT, ...)
- **Google**
- **Stackoverflow**
- **Dokumentation der Funktionen**

LEARNING RESOURCES

- Watch [this video](#) on working with strings (16 minutes)
- Watch [this snippet](#) on working with numbers (“Numbers” and “Working With Numbers”; 5 minutes)
- Watch the first four chapters of [this video](#) on functions (12 minutes)
- Watch [this video](#) on working with lists in Python (18 minutes)
- Watch [this video](#) on setting up VS Code, specifically for Python and Data Science

TASKS

- Import the dataset "umsatzdaten_gekuerzt.csv", which you can download via the following link:
<https://raw.githubusercontent.com/opencampus-sh/einfuehrung-in-data-science-und-ml/main/wetter.csv>
- Calculate the overall average temperature.
- Calculate the average temperature for the month of July.
- Compare whether the months of July and May differ significantly in their average temperature.

REGISTRATION FOR GITHUB COPILOT

- Register for GitHub Copilot as described [here](#).

As a student or teacher you get free access to GitHub Copilot Pro following the instructions given [here](#).

Next week, we will go into more detail on possible ways of using Ai assistance.