

# Pandas zur Datenanalyse Einführung

# Was ist Pandas?

- „Pandas“ = „*panel data*“
- “*Pandas*” = “*Python Analysis Data*”
- **Bibliothek zur Verarbeitung und Analyse von Datenmengen**
- Effiziente **Auswertung** von **Datentabellen** und **Zeitreihen**

```
?pandas
```

**Type:** module

**String form:** <module 'pandas' from 'C:\\Users\\Fabio\\anaconda3\\lib\\site-packages\\pandas\\\_\_init\_\_.py'>

**File:** c:\\users\\fabio\\anaconda3\\lib\\site-packages\\pandas\\\_\_init\_\_.py

**Docstring:**

pandas - a powerful data analysis and manipulation library for Python

=====

**\*\*pandas\*\*** is a Python package providing fast, flexible, and expressive data structures designed to make working with "relational" or "labeled" data both easy and intuitive. It aims to be the fundamental high-level building block for doing practical, **\*\*real world\*\*** data analysis in Python. Additionally, it has the broader goal of becoming **\*\*the most powerful and flexible open source data analysis / manipulation tool available in any language\*\***. It is already well on its way toward this goal.

# Welche Daten mit Pandas?

- **CSV-Dateien**
- **Excel-Tabellen**
- **JSON**
- **SQL-Datenbanken**
- **...**

# Einsatzbereiche

- **Sortierung und Filterung**
- **Datenbereinigung und Anpassung**
- **Aggregation und Gruppierung der Daten**
- **Zusammenführung und Verknüpfung der Daten**
- **Berechnungen von Spalten und Transformation**

# Literaturverzeichnis

- Pandas Pydata (2023): About Pandas, Online verfügbar unter:  
<https://pandas.pydata.org/about/>