

Datengewinnung und -analyse

# 4. Datenanalyse mit Pandas

Markus Schanta



# pandas: ein Überblick



- Programmbibliothek für Python zur Verarbeitung, Analyse und Darstellung von Daten
- 2008 erstmals von Wes McKinney veröffentlicht
- Seit ca. 2012-2015 der de-facto Standard für Datenanalyse in Python

# pandas Import

- `pandas` kann in Python wie folgt importiert werden:

# pandas Import

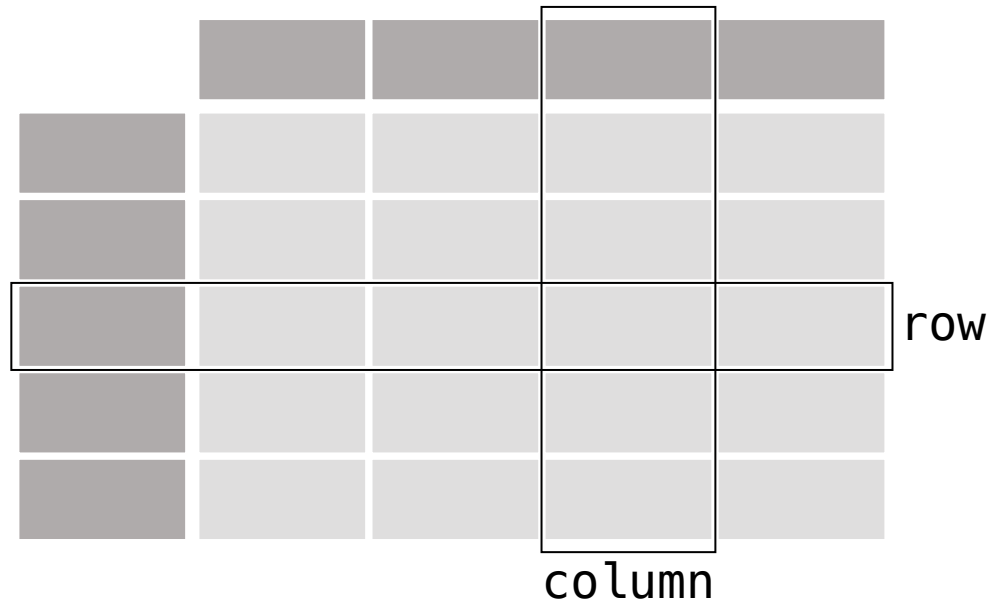
- `pandas` kann in Python wie folgt importiert werden:

```
import pandas as pd
```

# pandas DataFrames

- Daten werden in **pandas** hauptsächlich in sogenannten **DataFrames** gespeichert.

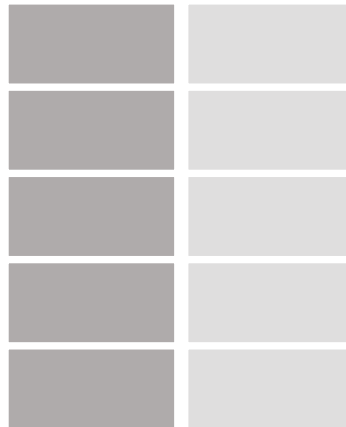
## DataFrame



# pandas Series

- Die Zeilen und Spalten eines **DataFrames** bestehen aus sogenannten **Series**.

Series



# pandas DataFrames

- Ein einfaches Beispiel zur Erstellung eines **DataFrames**:

# pandas DataFrames

- Ein einfaches Beispiel zur Erstellung eines **DataFrames**:

```
import pandas as pd

data = {
    'Name': ['Alice', 'Bob', 'Charlie'],
    'Alter': [25, 30, 35],
    'Stadt': ['Wien', 'Graz', 'Linz']
}

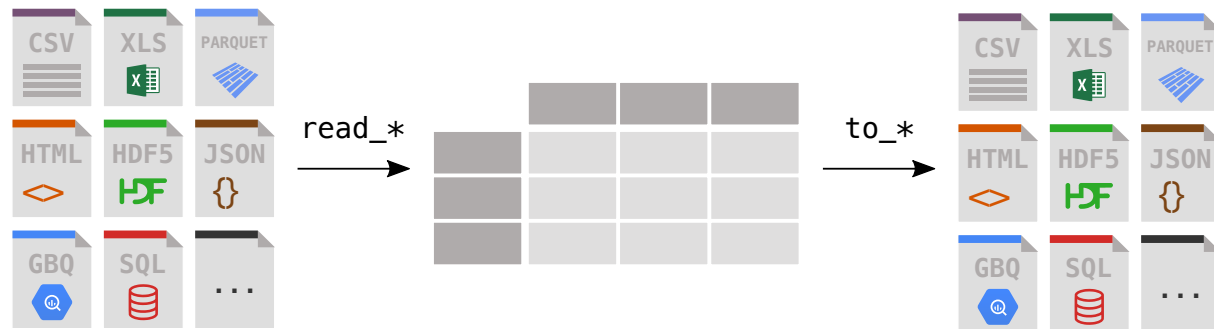
df = pd.DataFrame(data)

print(df)
```



# Daten mit **pandas** laden und speichern

- **pandas** bietet Funktionen zum Laden (`read_*`) und Speichern (`to_*`) von Daten in verschiedenen Formaten, z.B. CSV, Excel, SQL-Datenbanken und mehr.



# Selektieren von Spalten



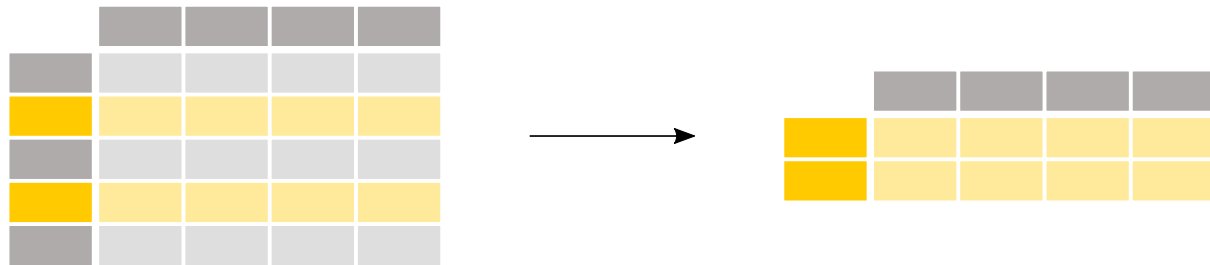
# Selektieren von Spalten



- Über den Spaltennamen als Attribut oder als Schlüssel:

```
# Als Attribut  
selection = df.Name  
  
# Als Schlüssel  
selection = df['Name']
```

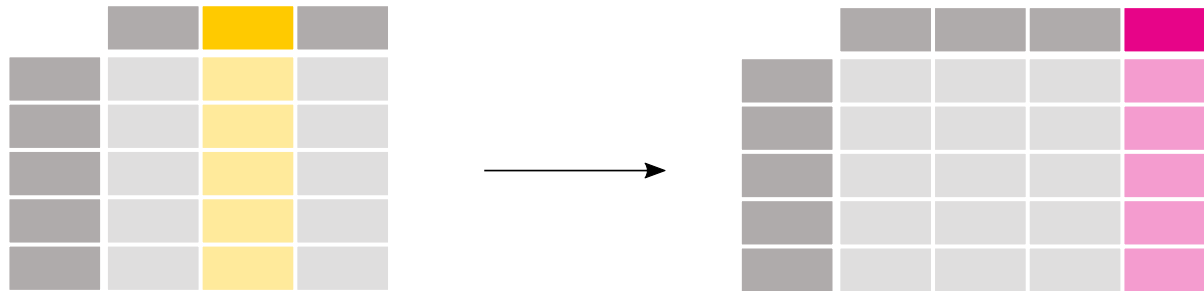
# Selektieren von Zeilen



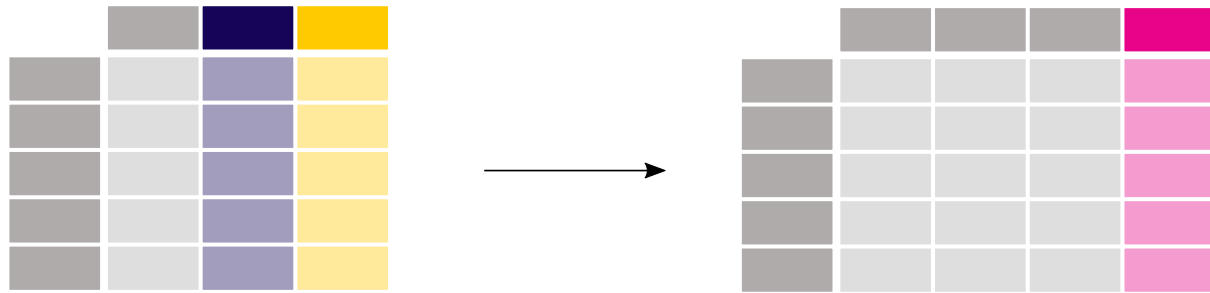
# Selektieren von Zeilen und Spalten



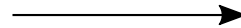
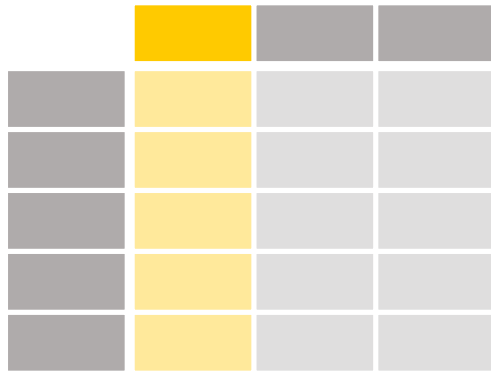
# Neue Spalten aus bestehenden Spalten erstellen



# Neue Spalten aus mehreren bestehenden Spalten erstellen

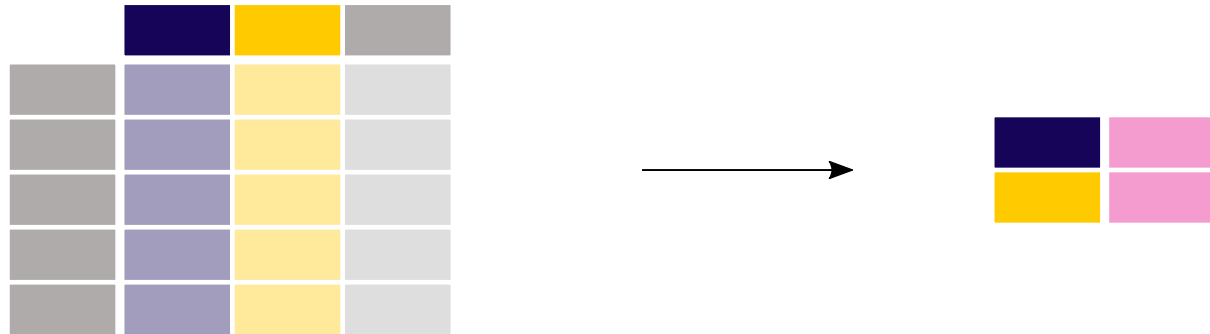


# Einzelne Spalten aggregieren

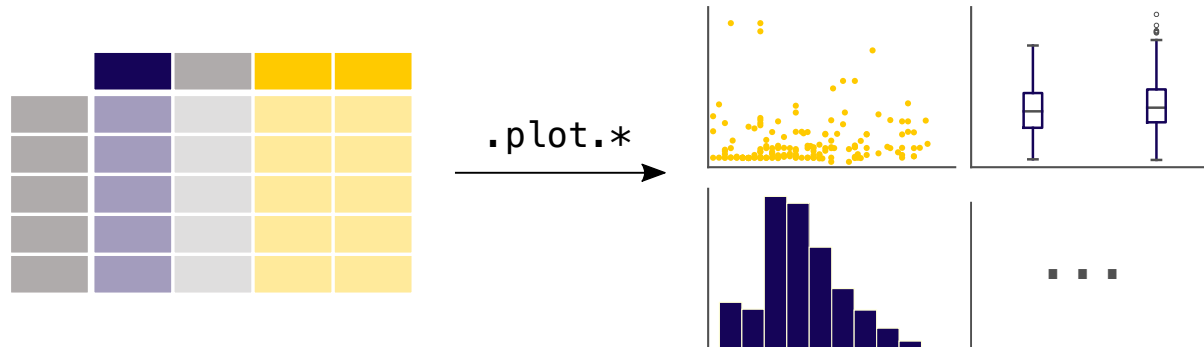




# Mehrere Spalten aggregieren



# Umfangreiche Methoden zur Visualisierung von Daten



**Fragen?**

# Nächste Schritte

- Review Folien & Materialien der heutigen Einheit
- 1. Gruppenübung:
  - Angabe diese Woche auf Moodle verfügbar
  - Abgabe bis So. 11.01.2026
- Nächste Einheit: Mo. 14.01.2026 18:30

**Danke und schönen Abend!**