

Pratikum Blatt 2 (4,5 Punkte)

Datenanalyse und Einführung in Maschinelles Lernen
WS 2025/26

Datenanalyse mit Python
Abgabe am 29./23.10.2025

Dozentin: Grit Behrens
mailto: grit.behrens@hsbi.de



Studiengang Informatik Fachbereich Campus Minden

Aufgabe 1 (1,5 Punkte – je 0,5)

a) Erstellen Sie ein DataFrame, der wie folgt aussieht:

```
Wien    country    Österreich
        area        414.6
        population  1805681
Berlin  country    Deutschland
        area        891.85
        population  3562166
Zürich  country    Schweiz
        area        87.88
        population  378884
dtype: object
```

b) Sortieren Sie nach dem Index und erzeugen Sie folgende Ausgabe:

```
Berlin  area        891.85
        country    Deutschland
        population  3562166
Wien    area        414.6
        country    Österreich
        population  1805681
Zürich  area        87.88
        country    Schweiz
        population  378884
dtype: object
```

Aufgabe 1

c) Vertauschen Sie die Indizes, sortieren Sie neu und erzeugen Sie die folgende Ausgabe:

```
area      Berlin      891.85
          Wien        414.6
          Zürich      87.88
country   Berlin      Deutschland
          Wien        Österreich
          Zürich      Schweiz
population Berlin      3562166
          Wien        1805681
          Zürich      378884
dtype: object
```

Aufgabe 2 (3 Punkte, je 0,5)

- a) Erzeugen Sie einen DataFrame mit einem Index (nicht nachträglich setzen, Index direkt erzeugen), der aus Vornamen besteht, sowie einer Spalte für das gewicht und einer Spalte für die Körpergröße, siehe auch die Abbildung.

	Gewicht	Größe
Henry	65	179
Sarah	58	165
Elke	58	172
Lulu	45	154
Vera	43	150
Toni	99	189
Maria	68	176
Chris	60	175

- b) Extrahieren Sie alle Zeilen, deren BMI im normalen Bereich liegt, d.h. zwischen 18.5 und 25. Es gilt: $BMI = \text{Gewicht in kg} / (\text{Größe in m})^2$.
Siehe auch die Abbildung:

	Gewicht	Größe
Henry	65	179
Sarah	58	165
Maria	68	176

Aufgabe 2 (3 Punkte)

c) Geben Sie alle Zeilen aus, in deren Namen ein kleines „e“ vorkommt, siehe auch die Abbildung.

	Gewicht	Größe
Henry	65	179
Elke	58	172
Vera	43	150

d) Fügen Sie eine Spalte mit dem BMI an.

	Gewicht	Größe	BMI
Henry	65	179	20.286508
Sarah	58	165	21.303949
Elke	58	172	19.605192
Lulu	45	154	18.974532
Vera	43	150	19.111111
Toni	99	189	27.714790
Maria	68	176	21.952479
Chris	60	175	19.591837

Aufgabe 2 (3 Punkte)

e) Sortieren Sie nach dem BMI

	Gewicht	Größe	BMI
Toni	99	189	27.714790
Maria	68	176	21.952479
Sarah	58	165	21.303949
Henry	65	179	20.286508
Elke	58	172	19.605192
Chris	60	175	19.591837
Vera	43	150	19.111111
Lulu	45	154	18.974532

f) Geben Sie nun alle Zeilen aus, deren BMI-Werte unter 20 liegen und deren Name ein kleines „e“ enthält:

	Gewicht	Größe	BMI
Elke	58	172	19.605192
Vera	43	150	19.111111