

## Wirtschaftsinformatik 2

### Wochenaufgabe - Woche 7

In dieser Woche beschäftigen wir uns weiter mit den Möglichkeiten von Pandas' DataFrames. Dazu kommt insbesondere das Filtern von Daten, was sich sehr gut mit den bisherigen Möglichkeiten kombinieren lässt, Statistiken zu berechnen. Das Filtern erlaubt jetzt insbesondere die Berechnung von *bedingten Statistiken*, d.h. Statistiken für Teile der Daten.

#### Aufgabe 1 (Statistiken mit Pandas DataFrame)

In dieser Aufgabe geht es um Daten von Fahrradkäufern. Unter der URL

<https://data.hsbo.de/bike-sales.csv>

findet sich eine Tabelle mit möglichen Kunden eines Fahrradvertriebs. Neben Eigenschaften wie *Einkommen* und *Bildung* finden sich Spalten über den *Beruf*, *Hausbesitz*, *Region* oder die *Pendel-Entfernung*.

Aus Marketing-Sicht ist es hilfreich, derartige Kundendaten zu untersuchen, um z.B. Personen zu identifizieren, die als mögliche Käufer angeschrieben oder mit Werbung kontaktiert werden können.

- Lesen Sie die Daten in einen DataFrame ein.
- Ermitteln Sie die Anzahl der Datensätze (Zeilen)
- Wie ist das durchschnittliche Alter und Einkommen von Kunden, die ein Fahrrad gekauft haben?
- Wieviel verheiratete Männer mit Kindern haben ein Fahrrad gekauft?
- Im Marketing wird häufig der Begriff *Buyer Persona* benutzt um *typische Käufer* zu erfassen und für Marketingzwecke zu nutzen. Entwickeln Sie eine eigene Buyer Persona indem Sie vier charakterisierende Eigenschaften auswählen und ermitteln Sie, wieviele Kunden in dem Datensatz dieser Persona entsprechen. Wieviel davon haben ein Fahrrad gekauft?

Als Beispiel:

Geschlecht	weiblich
Alter	30-50
Einkommen	> 40000
Autos	< 2