



## Übung zu Digitalisierung und Statistik

**Abgabe & Bearbeitung:** Freitag, 10. Jänner

### Aufgabe 11.1 (6 Punkte)

Gegeben sei die Excel-Datei *Daten-Freitag.xlsx*. Laden Sie das erste Datenblatt mittels Pandas in Python ein (wenn Sie ein Jupyter-Notebook nutzen, müssen Sie die Datei ins Jupyter-Datenverzeichnis hochladen). Berechnen Sie für die beiden ersten Spalten die Parameter  $m$ ,  $b$  einer linearen Regression, das Bestimmungsmaß  $R^2$  sowie den Pearson-Koeffizienten  $r_{x,y}$ . Sie können gerne hierbei Ihren Code aus den bisherigen Übungen wiederverwenden. Geben Sie alle berechneten Werte in der Konsole aus.

### Aufgabe 11.2 (4 Punkte)

Laden Sie das zweite Datenblatt in Python ein. Hier sehen Sie die geleakte Tabelle des Inventars des nichtexistenten Online-Händlers *danube.at* ;-). Erstellen Sie mit der `groupby`-Funktion:

- eine Tabelle, die für jede Kategorie die Summe aller vorhandenen Artikel anzeigt (Aggregationsfunktion: `sum()`)
- eine Tabelle, die für jede Kategorie den Preis des günstigsten Artikels anzeigt

Geben Sie beide Tabellen in der Konsole aus.