

# Quiz - Introduction to Python

Dieses Übungsblatt beinhaltet Aufgaben zu dem Kurs *Introduction to Python*, die euer Wissen zu den Kursinhalten basierend auf dem bereits verwendeten Gaming-Datensatz (*vgsales.csv*) abfragen. Bitte bearbeitet diese Aufgaben eigenständig und schickt eure Lösung, d.h. den dokumentierten Code mit Verweis auf die entsprechenden Fragen, bis zum **31. Januar 2023 23:59 Uhr** an [cestonaro@wiwi.uni-frankfurt.de](mailto:cestonaro@wiwi.uni-frankfurt.de) mit dem Betreff „Lösungen zum Python-Quiz“. Das Abschicken eurer Lösung ist Voraussetzung für den Erhalt eines Zertifikats für die Teilnahme am oben genannten Kurs.

Zur Bearbeitung der Fragen verwendet bitte folgende Python-Bibliotheken:

- os
- pandas
- numpy
- matplotlib
- seaborn

**Viel Erfolg und Spaß bei den Übungen!**

## Aufgaben zur eigenständigen Bearbeitung:

Bitte lese zunächst mit Hilfe von `pandas.read_csv()` den Datensatz *vgsales.csv* ein und beantworte anschließend folgende Fragen.

### 1. Globale Umsätze

- 1.1. Berechne welche Plattform im Jahr 2010 den größten Umsatz (*Global\_Sales*) ausweist?
- 1.2. Wie groß ist der größte Umsatz?
- 1.3. Berechne welche Plattform im Jahr 2010 den geringsten Umsatz (*Global\_Sales*) ausweist?
- 1.4. Wie groß ist der geringste Umsatz?  
Tipp: <https://docs.scipy.org/doc/numpy/reference/generated/numpy.argmax.html>

### 2. Neue Spiele

- 2.1. Wie viele neue Spiele gibt es im Jahr 2013, d.h. all jene Spiele, die es in den vorherigen Jahren nicht gab?  
Tipp: Bitte verwende `set` für diese Aufgabe.

### 3. Spielennamen

- 3.1. Wie heißt das Spiel mit dem längsten Namen?
- 3.2. Wie heißt das Spiel mit dem kürzesten Namen?

## 4. Umsatz

4.1. Überprüfe, ob für jedes Spiel gilt:  $Global\_Sales = NA\_Sales + EU\_Sales + JP\_Sales + Other\_Sales$ .

Tipp: Erstelle für den rechten Teil der Formel eine neue Variable *calcSales*.

4.2. In wie vielen Fällen ist der in 4.1 stehende Ausdruck nicht korrekt?

4.3. Erstelle einige zusammenfassende Statistiken zu der Differenz *calcSales - Global\_Sales*:

- Maximum
- Minimum
- Durchschnitt
- Varianz
- 5% Perzentil
- 95% Perzentil

Tipp:

<https://docs.scipy.org/doc/numpy/reference/generated/numpy.percentile.html>

## 5. Mario

5.1. Finde alle Spiele, die das Wort "Mario" beinhalten. Wie viele Spiele sind dies?

5.2. Erstelle einen neuen Ordner "Mariogames". Speichere die in 5.1 identifizierten Spiele mit den Spalten Name, Platform und Year in eine Datei "Mariogames/Mariogames.csv".

5.3. Erstelle ein Balkendiagramm ("bar plot"), um die Anzahl an Mariospielen pro Plattform zu zeigen.