

Quiz – Introduction to Python

Dieses Übungsblatt beinhaltet Aufgaben zu dem efl-Kurs *Introduction to Python*, die euer Wissen zu den Kursinhalten basierend auf dem bereits verwendeten Gaming-Datensatz abfragen. Bitte bearbeitet diese Aufgaben eigenständig und erstellt eure Lösung in Form von dokumentiertem Code, in dem ihr eure Vorgehensweise beschreibt und auf die konkrete Frage im Quiz Bezug nehmt. Ihr könnt gerne das Code-Skeleton (`quiz_introduction_skeleton.py`) als Ausgangsbasis verwenden.

Sendet eure Lösungen bis zum 19. Januar 2025 23:59 Uhr an **dscourses@eflab.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:

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?

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?
 - 3.1.1. Wieso definieren wir `longest_name = ""`?
- 3.2. Wie heißt das Spiel mit dem kürzesten Namen?
 - 3.2.1. Wieso definieren wir `shortest_name = game_names[0]`? Wieso definieren wir hier anders als bei `longest_name`?

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 Spalte *Calc_Sales*.

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

4.3. Erstelle einige zusammenfassende Statistiken zu der Differenz *Calc_Sales - Global_Sales*:

- Maximum
- Minimum
- Durchschnitt
- Varianz
- 25% Perzentil
- 75% Perzentil

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.

6. Lieblingsspiele

6.1. Wähle dir deine 3 Lieblingsspiele aus.

6.2. Prüfe, ob deine Spiele in unsere Datenbank enthalten sind. Schreibe dafür eine Funktion *check_game*, die jeweils auf der Konsole ausgibt, ob das Spiel enthalten ist oder nicht.

6.2.1. Die Funktion nimmt den Namen des Spiels entgegen.

6.2.2. Teste deine Funktion mit den gegebenen Beispielsvariablen.

6.3. Prüfe, ob deine 3 Lieblingsspiele in unserer Datenbank enthalten sind.

6.4. Erstelle dir drei Variablen und speichere jeweils dein Review zu deinen Lieblingsspielen ab, z.B. „das Spiel macht Spaß“

6.5. Speichere in einem neuen Dataframe dein Ranking ab, also die Namen deiner Lieblingsspiele und die dazugehörigen Reviews.

6.6. Kombiniere dein Ranking mit den Daten aus unserer Datenbank.

6.6.1. Was tun wir in der folgenden Zeile:

```
filtered_df = df.loc[(df["Name"] == my_games[0]) | (df["Name"] == my_games[1]) | (df["Name"] == my_games[2])]
```