

Aufgabenblatt #3: Transformation

Team Namen:

Start Datum: 02 Dezember 2024

Fälligkeitsdatum: 05 Januar 2025

1 Voraussetzungen

1. Gebt nur nötiges auf der Konsole aus
2. Benutzt das Konzept von Modulen
3. Benutzt immer Klassen und Funktionen, wenn möglich
4. PEP8 ab sofort immer zu nutzen

2 Hinweis

Jede externe Bibliothek ist ab sofort verwendbar, falls nicht ersichtlich ist, was euer Code macht, dann sollt ihr diesen bitte kommentieren.

3 Aufgaben

3.1 Aufgabe 1

Extrahiert dieselben Daten, wie auf Übungsblatt 2 (FinancialSample.csv).
Speichert die Daten dieses Mal in ein Pandas Dataframe und gebt die letzten 10 Elemente auf der Konsole aus.

3.2 Aufgabe 2

Transformiert die Spalte "Date" ins amerikanische Zeitformat (MM/DD/YYYY) mit Hilfe der datetime Bibliothek.

3.3 Aufgabe 3

Erzeugt ein **neues** Pandas Dataframe, welches nur die folgenden Spalten enthält und gebt die ersten 5 Einträge aus:

Product, Profit, COGS, Sales

3.4 Aufgabe 4

Lest die Spalten "Month Number", "Month Name" und "Year" in eine einzige Spalte mit einem kombinierten Datum aus allen drei Bestandteilen und löscht den Rest.

3.5 Aufgabe 5

Findet die Positionen der zehn größten lokalen Maxima. Ein lokales Maximum ist ein Wert, der von zwei kleineren Werten umschlossen wird. Zum Beispiel $[1, 3, 8, 5, 10, 4] \rightarrow 8$ und 10 sind lokale Maxima. Das Ergebnis lautet in dem Fall also Position 2 und Position 4.

3.6 Aufgabe 6

Erzeuge ein neues Pandas Dataframe, welches nur jede X-te Zeile aus den Original-Daten beinhaltet. Wobei X = Gruppennummer gilt. Hierfür muss Pandas zum Einlesen benutzt werden.

3.7 Aufgabe 7

Während ihr die .csv Datei einlest, ändert alle Werte in der "discount" Spalte und ändert die Werte folgendermaßen:

if "-" and $\leq 200 \rightarrow$ "Low"

elif > 200 and $< 2000 \rightarrow$ "Medium"

elif $\geq 2000 \rightarrow$ "High"

Benutzt hierfür speziell die pandas read csv Funktion und lest in der Dokumentation nach, ob ihr nützliche Parameter findet.