

Dokumentation der Python Applikation CSV-Importer

https://github.com/nilakanthaLE/AWD-sommer-2019/tree/master/Aufgabe%201/1_b

Die Datei `csvImporter.py` enthält den relevanten Python-Code. Weiterhin wird auf Code einer Datei mit Namen `helper.py` zurückgegriffen. Am Ende von `csvImporter.py` sind Parameter für drei CSV-Dateien vorgegeben. Im Testaufruf wird die Methode `csvImport` mit einem der drei Parameter-Tupel aufgerufen. Mit dem Befehl `python csvImporter.py` wird das Programm gestartet.

Der Nutzer sieht eine Überschrift mit dem Namen der eingelesenen CSV. Darunter befindet sich der Datenblock. Eventuell sind darin Spalten oder Zeilen enthalten, die nicht zu den Daten, sondern zur Beschreibung der Spalten bzw. Zeilen gehören. Der Nutzer wird gefragt, ob die Daten bereits in Ordnung sind. Andernfalls wird er gebeten, die Anzahl der überflüssigen (Beschreibungs-) Spalten und Zeilen anzugeben.

Sobald die Daten korrekt sind und der Nutzer dies mit ‚j‘ bestätigt, steht das Array in der geforderten Form zur Verfügung. Die Eingaben des Nutzers werden für beliebig viele CSV-Dateien gespeichert. Sie stehen damit auch nach einem erneuten Programmstart zur Verfügung.

Die Parameter bestimmen, ob eventuell Zelleneinträge vorhanden sind, die keine Zahlen sind und die dennoch bewahrt werden sollen. Zudem ist es möglich Daten direkt zu Beginn von ungewünschten Zeichen zu bereinigen. Dies ist im Fall der `PendlerCSV` notwendig, weil hier die Trennung der Tausender durch ein merkwürdiges Leerzeichen erfolgt.

Nach mehreren Ansätzen eines völlig automatisierten Einlesens, habe ich mich zuletzt für diese Variante einer Interaktion mit dem User entschieden. Die Daten der verschiedenen CSVs waren zu heterogen, wodurch eine völlige Automatisierung immer komplexer wurde. Trotz dieser Heterogenität war es dennoch mein Anspruch eine einheitliche, hinreichend flexible Lösung für den Import unterschiedlicher CSVs zu entwickeln.

Auf eine weitere Extraktion der Beschreibung für Zeilen und Spalten habe ich verzichtet, weil dies nicht in der Aufgabenstellung gefordert wurde.

Die Dokumentation der Funktionsweise der einzelnen verwendeten Methoden und Klassen sind im Quellcode als Kommentare eingefügt.