Data loading and processing

Overview

In this exercise, you will load data from a json file and store it into an sqlite database.

Afterwards, you will use sql or Python to calculate values and write the update values into a table.

Preliminaries

You do have a

- virtual environment for Python, e.g., in VSCode
- you installed pandas, sqlalchemy
- you loaded the example data from moodle
  - SQL script for the creation of sales.db
  - SalesRecords.json

First part: Create and analyze the database

Create the database successively by executing the script parts. Alternatively, write a short Python program that does this for you.

You should now have a database called sales.db with 2 tables: account and customer

The *customer* table includes information on customers of your company (you are selling sneakers).

The *account* table includes the account information for your customers. The balance is the internal balance of the customer with our company resulting from buying our products.

Account and customer tables are linked through a foreign key.

(a) Which ones are the referring fields?

Second part: load the sales data

You will now be using Python for loading the Sales Data into a Pandas Dataframe. Have a look at "SalesRecords.json"

(b) What is special about the structure? Can it be inserted as is into a database (table)?

Have a look at the following Web page, wich will give you some examples for "normalizing" json data. https://towardsdatascience.com/how-to-best-work-with-json-in-python-2c8989ff0390

Also, have a look at the pandas cheat sheet given in the folder.

Third part: Write the normalized data into the database

For better handling of the data, try to write the normalized data into a table. The table should be named "jsondata".

Use SQLAlchemy for writing the data. Have a look at the following Web page, showing how to do this: https://techoverflow.net/2021/04/27/how-to-export-pandas-dataset-to-sqlite-database/

Fourth part: Create an SQL statement

Write an SQL that provides the following information:

1. the Customer Name (Surname) from both, the jsondata and the customer table
2. the amount spent for buying a product for
3. the balance of the customer's account

The result should look like this:

Customer Name	Sales Amount	Account Balance
Gruene	3.99	400.00

Final part: Update the Account table

Finally, update the account table and withdraw the sales amount from the account balance.

Run an SQL that shows that the amount was updated.