Programming Assignment: SQLite Interaction

Dr. Youmi Suk

DS 2001: Proramming for Data Science

Brandon Ou (bro9tn)

23 February 2022

Text

Description automatically generated1. (1 PT) In SQLite, create a database called *mydata.db*  
List the databases available in SQLite. Take a screenshot to show the database was created, and include it in your results document.

2. (1 PT) Create a dataset with five rows and three columns. One of the columns must include categorical data (e.g., industry, city). The format needs to be a list of tuples.

3. (1 PT) Create a table in mydata.db called *mytable* and pass a schema. End the transaction with a commit.

4. (1 PT) Insert your dataset into *mytable*. End the transaction with a commit.

Text

Description automatically generated5. (1 PT) Write and execute a query that returns all of the inserted records. Print the result.

6. (1 PT) Write and execute a query that filters out some of the inserted records based on a condition. Print the result.

Graphical user interface

Description automatically generated

7. (1 PT) Write and execute a query that uses a GROUP BY clause to aggregate on the categorical field. For example, you might compute the number of records for each level of industry (e.g., TECH 2, CONSUMER GOODS 3). Print the result.

Graphical user interface, text, application

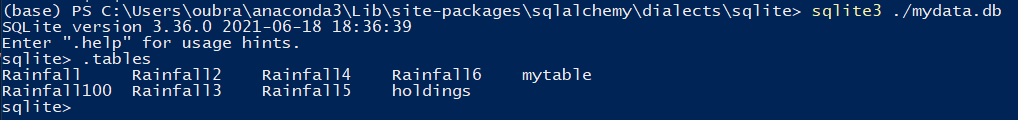
Description automatically generated

8. (1 PT) Based on the results from one of the queries, load the results into a pandas dataframe. Print the dataframe.

Graphical user interface, text, application, email

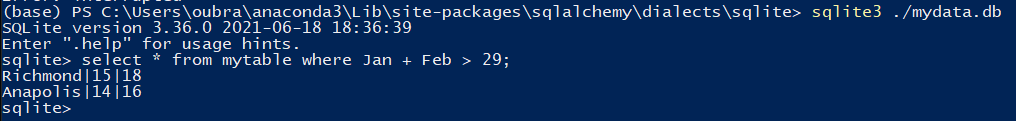
Description automatically generated

9. (1 PT) From SQLite, list the tables. You should see *mytable* listed. Take a screenshot to show the table was created, and include it in your results document.



* mytable (rightmost element in the first row) is in the list of all tables in mydata.db

10. (1 PT) From SQLite, write and execute a query on *mytable*. Take a screenshot to include the query and results in your results document.



**The Python code for problems 2-8 is shown below:**

#!/usr/bin/env python

# coding: utf-8

import sqlite3

# 2 - Table stores the January and February total rainfall in inches

data = [

    ('New York City', 11, 12),

    ('Los Angeles', 10, 8),

    ('Washington DC', 13, 14),

    ('Richmond', 15, 18),

    ('Anapolis', 14, 16),

]

path\_to\_db = "C:/Users/oubra/anaconda3/Lib/site-packages/sqlalchemy/dialects/sqlite/mydata.db"

conn = sqlite3.connect(path\_to\_db)

cur = conn.cursor()

# 3

cur.execute('create table mytable (City, Jan rainfall, Feb rainfall)')

conn.commit()

# 4

cur.executemany('insert into mytable values (?,?,?)', data)

conn.commit()

# 5

for row in cur.execute('select \* from mytable'):

    print(row)

# 6

for row in cur.execute('select \* from mytable where Feb >= 16'):

    print(row)

# 7

# Calculates total rainfall in January and February

for row in cur.execute('select City, (Jan + Feb) from mytable GROUP BY City'):

    print(row)

# 8

# https://pandas.pydata.org/pandas-docs/stable/user\_guide/io.html#reading-tables

import pandas as pd

df = pd.read\_sql\_query("select City, Jan, Feb from mytable GROUP BY City", conn)

df.head(5)

**The same code is screenshotted to show outputs within a Jupyter Notebook:**

Graphical user interface, text, application, email

Description automatically generated