csv_to_sqlite

January 7, 2025

1 Loading CSV file into a SQLite database

[1]:
```
import pandas as pd
```

[2]:
```
df = pd.read_csv("./data/students.csv")
```

[3]:
```
# Preview the first few rows of the DataFrame
df.head()
```

[3]:

	Student ID	Name	Age	Subject	Year of Study	Country of Origin
0	2703f3f0	Mr Clifford Watson	25.0	English Literature	1.0	Saint Barthelemy
1	a8040287	Elliott Ward	25.0	Computer Science	4.0	Guinea
2	d8da5486	Miss Pauline Dunn	22.0	Engineering	4.0	Afghanistan
3	3ac1b74d	Mr Dominic Mason	22.0	Physics	1.0	Palau
4	67850858	Mrs Melanie Brown	18.0	English Literature	3.0	Algeria

[4]:
```
import sqlite3
```

1.1 Create (or update) the database

[5]:
```
# Connect to SQLite database (or create it)
conn = sqlite3.connect("./data/students.db")

# Load the data into the SQLite database
df.to_sql("students", conn, if_exists="replace", index=False)

# Close the database connection
conn.close()
```

1.2 Load the first record from the db and display

[6]:
```
# Connect to the SQLite database
conn = sqlite3.connect("./data/students.db")

# Query the first record from the table
query = "SELECT * FROM students LIMIT 1"
first_record = conn.execute(query).fetchone()

# Close the database connection
conn.close()
```

[7]:
```
# Display the first record
print("First record:", first_record)
```

```
First record: ('2703f3f0', 'Mr Clifford Watson', 25.0, 'English Literature',
1.0, 'Saint Barthelemy')
```

1.3 Loading an Excel file to a DataFrame

[8]:
```
# For this to work you need to have installed openpyxl, e.g. with pip or conda
df = pd.read_excel('./data/Students.xlsx', sheet_name='students')
df.head()
```

[8]:

	Student ID	Name	Age	Subject	Year of Study	Country of Origin
0	2703f3f0	Mr Clifford Watson	25.0	English Literature	1.0	Saint Barthelemy
1	a8040287	Elliott Ward	25.0	Computer Science	4.0	Guinea
2	d8da5486	Miss Pauline Dunn	22.0	Engineering	4.0	Afghanistan
3	3ac1b74d	Mr Dominic Mason	22.0	Physics	1.0	Palau
4	67850858	Mrs Melanie Brown	18.0	English Literature	3.0	Algeria