Database for Beginners

September 3, 2020

Compiled by Ebude Yolande

0.1 Introduction

Nowadays, most companies use database to store information. There are different kind of database but in this workshop we will handle Structured Query Language (SQL) type: SQLite. To start download this dataset SqLite file

0.2 Content

- Import different database types
- Modify database

```
[1]: import sqlite3 as sq
     import pandas as pd
     import warnings
     warnings.filterwarnings('ignore')
```

0.3 Import different database types

0.3.1 SQlite

```
[2]: conn=sq.connect('GeneratorRefilling.sqlite')
     cur=conn.cursor()
     query="SELECT * FROM RefillingGen2018"
     data=pd.read_sql(query,conn)
     data.head(3)
```

```
[2]:   Generator Buyer_Name  Purchase_Quantity Purchase_Date Purchase_Time  \
     0   Genset1       Andy              756.0      02/01/15         15:25
     1   Genset5       Kate              163.5      04/01/15         10:45
     2   Genset9      Louis              860.0      06/01/15         15:45

        Payment      Mass  Generator_Amount Refilling_Name  Refilling_Quantity  \
     0   453600   628.992             756.0           Andy               756.0
     1    98100   136.032             163.5           Kate               163.5
     2   516000   715.520             860.0           Kate               800.0
```

```
     Refilling_Date Refilling_Time  Check
0         03/01/15          07:10      1
1         03/01/15          17:10      1
2         06/01/15          17:40      1
```

0.3.2   Dat file

```python
[3]: tab = pd.read_table('3055.dat')
     temp=tab.columns
     tab.columns=['num','val','Firstname','Lastname','Email',
                  'val1','unit','Date','Datetime','Verify'] # change the header
     tab.iloc[-1]=temp #assign header to last row
     tab.tail(3)
```

```
[3]:      num val Firstname  Lastname                             Email val1 unit  \
     596  598   1      Wade  Delvalle    wade.delvalle@sakilacustomer.org  604    t
     597  599   2    Austin   Cintron  austin.cintron@sakilacustomer.org  605    t
     598  524   1     Jared       Ely       jared.ely@sakilacustomer.org  530    t

                Date                 Datetime Verify
     596  2006-02-14  2013-05-26 14:49:45.738      1
     597  2006-02-14  2013-05-26 14:49:45.738      1
     598  2006-02-14  2013-05-26 14:49:45.738    1.1
```

0.4   Modify database

Create a table

```python
[5]: conn=sq.connect('GeneratorRefilling.sqlite')
     cur=conn.cursor()
     query='CREATE TABLE People (id integer PRIMARYKEY, name char(25), time Date)'
     conn.execute(query)
```

```
[5]: <sqlite3.Cursor at 0x29c5b4e6b90>
```

Create new table of an existing dataframe

```python
[6]: tab.to_sql('Tab3055',conn)
```

```python
[7]: query="SELECT * FROM Tab3055"
     tab3055=pd.read_sql(query,conn)
     tab3055.head(3)
```

```
[7]:    index num  val Firstname  Lastname                               Email  \
     0      0   1  1.0      Mary     Smith        mary.smith@sakilacustomer.org
     1      1   2  1.0  Patricia   Johnson  patricia.johnson@sakilacustomer.org
     2      2   3  1.0     Linda  Williams    linda.williams@sakilacustomer.org

        val1 unit        Date                 Datetime Verify
     0   5.0    t  2006-02-14  2013-05-26 14:49:45.738    1.0
     1   6.0    t  2006-02-14  2013-05-26 14:49:45.738    1.0
```

```
2  7.0    t  2006-02-14  2013-05-26 14:49:45.738     1.0
```

Add rows in new table

[8]:
```
query = "INSERT INTO People(id, name, time) VALUES ('1','Penina','12/01/2019')"
conn.execute(query)
query="SELECT * FROM People"
Peop=pd.read_sql(query,conn)
Peop.head()
```

[8]:
```
   id    name       time
0   1  Penina  12/01/2019
```

Delete rows in table

[9]:
```
query = "DELETE FROM Tab3055 WHERE Firstname=='Mary'"
conn.execute(query)
```

[9]: `<sqlite3.Cursor at 0x29c5b81a2d0>`

Change a value in the table

[10]:
```
query = "UPDATE People SET name='Angela' WHERE id='1'"
conn.execute(query)
```

[10]: `<sqlite3.Cursor at 0x29c5b81a570>`

[11]:
```
query="SELECT * FROM People"
Peop=pd.read_sql(query,conn)
Peop.head()
```

[11]:
```
   id    name       time
0   1  Angela  12/01/2019
```

Close the database

[12]:
```
conn.close()
```

0.5 Challenge

1. Import "GeneratorRefilling.sqlite" and "3057.dat" and change the header of "3057.dat" to ['num','name','surname','date']

2. Create a table in the database for "3057.dat" called Tab3057

3. Delete the "Tab3055" table

4. Add a row to Tab3057 table

5. Delete the 3rd row of Tab3057

6. Add column 'Verify' to Tab3057 and assign 1 is num>50 and 0 else