Pandas

March 29, 2019

```
In [1]: !pip install pysqlite3

Collecting pysqlite3
  Downloading https://files.pythonhosted.org/packages/83/31/42af47b96c435d956e8eef5d65f0a2cc2120
Building wheels for collected packages: pysqlite3
  Building wheel for pysqlite3 (setup.py) ... done
  Stored in directory: /home/nbuser/.cache/pip/wheels/ac/33/d1/4419e015a2f61d53b9dd3721fa636b71a
Successfully built pysqlite3
Installing collected packages: pysqlite3
Successfully installed pysqlite3-0.2.1
```

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

```
In [2]: import sqlite3
        con = sqlite3.connect("database.db")
```

```
In [3]: df = pd.read_sql_query("SELECT * FROM employees", con)
        df
```

```
Out[3]:    EmployeeId  LastName FirstName                Title  ReportsTo  \
        0           1     Adams    Andrew      General Manager        NaN
        1           2   Edwards     Nancy       Sales Manager        1.0
        2           3   Peacock      Jane  Sales Support Agent        2.0
        3           4      Park  Margaret  Sales Support Agent        2.0
        4           5   Johnson     Steve  Sales Support Agent        2.0
        5           6  Mitchell   Michael          IT Manager        1.0
        6           7      King    Robert            IT Staff        6.0
        7           8  Callahan     Laura            IT Staff        6.0

                     BirthDate             HireDate                     Address  \
        0  1962-02-18 00:00:00  2002-08-14 00:00:00        11120 Jasper Ave NW
        1  1958-12-08 00:00:00  2002-05-01 00:00:00               825 8 Ave SW
        2  1973-08-29 00:00:00  2002-04-01 00:00:00               1111 6 Ave SW
        3  1947-09-19 00:00:00  2003-05-03 00:00:00           683 10 Street SW
        4  1965-03-03 00:00:00  2003-10-17 00:00:00               7727B 41 Ave
        5  1973-07-01 00:00:00  2003-10-17 00:00:00         5827 Bowness Road NW
        6  1970-05-29 00:00:00  2004-01-02 00:00:00  590 Columbia Boulevard West
```

```
7  1968-01-09 00:00:00  2004-03-04 00:00:00                    923 7 ST NW


           City State Country PostalCode              Phone                Fax  \
0      Edmonton    AB  Canada    T5K 2N1  +1 (780) 428-9482  +1 (780) 428-3457
1       Calgary    AB  Canada    T2P 2T3  +1 (403) 262-3443  +1 (403) 262-3322
2       Calgary    AB  Canada    T2P 5M5  +1 (403) 262-3443  +1 (403) 262-6712
3       Calgary    AB  Canada    T2P 5G3  +1 (403) 263-4423  +1 (403) 263-4289
4       Calgary    AB  Canada    T3B 1Y7   1 (780) 836-9987   1 (780) 836-9543
5       Calgary    AB  Canada    T3B 0C5  +1 (403) 246-9887  +1 (403) 246-9899
6    Lethbridge    AB  Canada    T1K 5N8  +1 (403) 456-9986  +1 (403) 456-8485
7    Lethbridge    AB  Canada    T1H 1Y8  +1 (403) 467-3351  +1 (403) 467-8772


                      Email
0     andrew@chinookcorp.com
1      nancy@chinookcorp.com
2       jane@chinookcorp.com
3   margaret@chinookcorp.com
4      steve@chinookcorp.com
5    michael@chinookcorp.com
6     robert@chinookcorp.com
7      laura@chinookcorp.com

In [8]: df.shape

Out[8]: (8, 15)

In [11]: df.iloc[0:5,:]

Out[11]:    EmployeeId LastName FirstName                 Title  ReportsTo  \
         0           1    Adams    Andrew       General Manager        NaN
         1           2  Edwards     Nancy         Sales Manager        1.0
         2           3  Peacock      Jane  Sales Support Agent        2.0
         3           4     Park  Margaret  Sales Support Agent        2.0
         4           5  Johnson     Steve  Sales Support Agent        2.0


                   BirthDate             HireDate             Address      City  \
         0  1962-02-18 00:00:00  2002-08-14 00:00:00  11120 Jasper Ave NW  Edmonton
         1  1958-12-08 00:00:00  2002-05-01 00:00:00          825 8 Ave SW   Calgary
         2  1973-08-29 00:00:00  2002-04-01 00:00:00          1111 6 Ave SW   Calgary
         3  1947-09-19 00:00:00  2003-05-03 00:00:00        683 10 Street SW   Calgary
         4  1965-03-03 00:00:00  2003-10-17 00:00:00           7727B 41 Ave   Calgary


           State Country PostalCode              Phone                Fax  \
         0    AB  Canada    T5K 2N1  +1 (780) 428-9482  +1 (780) 428-3457
         1    AB  Canada    T2P 2T3  +1 (403) 262-3443  +1 (403) 262-3322
         2    AB  Canada    T2P 5M5  +1 (403) 262-3443  +1 (403) 262-6712
         3    AB  Canada    T2P 5G3  +1 (403) 263-4423  +1 (403) 263-4289
         4    AB  Canada    T3B 1Y7   1 (780) 836-9987   1 (780) 836-9543
```

```
                           Email
         0     andrew@chinookcorp.com
         1      nancy@chinookcorp.com
         2       jane@chinookcorp.com
         3   margaret@chinookcorp.com
         4      steve@chinookcorp.com
```

In [13]: df.loc[:,["LastName","FirstName","City"]]

```
Out[13]:    LastName FirstName        City
         0     Adams    Andrew    Edmonton
         1   Edwards     Nancy     Calgary
         2   Peacock      Jane     Calgary
         3      Park  Margaret     Calgary
         4   Johnson     Steve     Calgary
         5  Mitchell   Michael     Calgary
         6      King    Robert  Lethbridge
         7  Callahan     Laura  Lethbridge
```

In [14]: df.columns

```
Out[14]: Index(['EmployeeId', 'LastName', 'FirstName', 'Title', 'ReportsTo',
                'BirthDate', 'HireDate', 'Address', 'City', 'State', 'Country',
                'PostalCode', 'Phone', 'Fax', 'Email'],
              dtype='object')
```

In [15]: df.isnull()

```
Out[15]:    EmployeeId  LastName  FirstName  Title  ReportsTo  BirthDate  HireDate  \
         0       False     False      False  False       True      False     False
         1       False     False      False  False      False      False     False
         2       False     False      False  False      False      False     False
         3       False     False      False  False      False      False     False
         4       False     False      False  False      False      False     False
         5       False     False      False  False      False      False     False
         6       False     False      False  False      False      False     False
         7       False     False      False  False      False      False     False

            Address   City  State  Country  PostalCode  Phone    Fax  Email
         0    False  False  False    False       False  False  False  False
         1    False  False  False    False       False  False  False  False
         2    False  False  False    False       False  False  False  False
         3    False  False  False    False       False  False  False  False
         4    False  False  False    False       False  False  False  False
         5    False  False  False    False       False  False  False  False
         6    False  False  False    False       False  False  False  False
         7    False  False  False    False       False  False  False  False
```

In [16]: df.isnull().sum()

```
Out[16]: EmployeeId    0
         LastName      0
         FirstName     0
         Title         0
         ReportsTo     1
         BirthDate     0
         HireDate      0
         Address       0
         City          0
         State         0
         Country       0
         PostalCode    0
         Phone         0
         Fax           0
         Email         0
         dtype: int64
```