mod5_final_project

August 20, 2025

Assignment: Notebook for Graded Assessment

1 Introduction

Using this Python notebook you will:

1. Understand three Chicago datasets
2. Load the three datasets into three tables in a SQLIte database
3. Execute SQL queries to answer assignment questions

1.1 Understand the datasets

To complete the assignment problems in this notebook you will be using three datasets that are available on the city of Chicago's Data Portal:

1. Socioeconomic Indicators in Chicago
2. Chicago Public Schools
3. Chicago Crime Data

1.1.1 1. Socioeconomic Indicators in Chicago

This dataset contains a selection of six socioeconomic indicators of public health significance and a "hardship index," for each Chicago community area, for the years 2008 – 2012.

A detailed description of this dataset and the original dataset can be obtained from the Chicago Data Portal at:

https://data.cityofchicago.org/Health-Human-Services/Census-Data-Selected-socioeconomic-indicators-in-C/kn9c-c2s2

1.1.2 2. Chicago Public Schools

This dataset shows all school level performance data used to create CPS School Report Cards for the 2011-2012 school year. This dataset is provided by the city of Chicago's Data Portal.

A detailed description of this dataset and the original dataset can be obtained from the Chicago Data Portal at:

https://data.cityofchicago.org/Education/Chicago-Public-Schools-Progress-Report-Cards-2011-/9xs2-f89t

1.1.3 3. Chicago Crime Data

This dataset reflects reported incidents of crime (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2001 to present, minus the most recent seven days.

A detailed description of this dataset and the original dataset can be obtained from the Chicago Data Portal at:

https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2

1.1.4 Download the datasets

This assignment requires you to have these three tables populated with a subset of the whole datasets.

In many cases the dataset to be analyzed is available as a .CSV (comma separated values) file, perhaps on the internet.

Use the links below to read the data files using the Pandas library.

- Chicago Census Data

https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DB0201EN-SkillsNetwork/labs/FinalModule_Coursera_V5/data/ChicagoCensusData.csv?utm_medium=Exinflu SkillsNetwork-Channel-SkillsNetworkCoursesIBMDeveloperSkillsNetworkDB0201ENSkillsNetwork20127838-2021-01-01

- Chicago Public Schools

https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DB0201EN-SkillsNetwork/labs/FinalModule_Coursera_V5/data/ChicagoPublicSchools.csv?utm_medium=Exin SkillsNetwork-Channel-SkillsNetworkCoursesIBMDeveloperSkillsNetworkDB0201ENSkillsNetwork20127838-2021-01-01

- Chicago Crime Data

https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DB0201EN-SkillsNetwork/labs/FinalModule_Coursera_V5/data/ChicagoCrimeData.csv?utm_medium=Exinflu SkillsNetwork-Channel-SkillsNetworkCoursesIBMDeveloperSkillsNetworkDB0201ENSkillsNetwork20127838-2021-01-01

NOTE: Ensure you use the datasets available on the links above instead of directly from the Chicago Data Portal. The versions linked here are subsets of the original datasets and have some of the column names modified to be more database friendly which will make it easier to complete this assignment.

Execute the below code cell to install the required libraries

[1]:
```
!pip install pandas
!pip install ipython-sql prettytable

import prettytable
```

```
prettytable.DEFAULT = 'DEFAULT'
```

```
Collecting pandas
  Downloading
pandas-2.3.1-cp312-cp312-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata
(91 kB)
Collecting numpy>=1.26.0 (from pandas)
  Downloading
numpy-2.3.2-cp312-cp312-manylinux_2_27_x86_64.manylinux_2_28_x86_64.whl.metadata
(62 kB)
Requirement already satisfied: python-dateutil>=2.8.2 in
/opt/conda/lib/python3.12/site-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in /opt/conda/lib/python3.12/site-
packages (from pandas) (2024.2)
Collecting tzdata>=2022.7 (from pandas)
  Downloading tzdata-2025.2-py2.py3-none-any.whl.metadata (1.4 kB)
Requirement already satisfied: six>=1.5 in /opt/conda/lib/python3.12/site-
packages (from python-dateutil>=2.8.2->pandas) (1.17.0)
Downloading
pandas-2.3.1-cp312-cp312-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (12.0
MB)
                         12.0/12.0 MB
124.2 MB/s eta 0:00:00
Downloading
numpy-2.3.2-cp312-cp312-manylinux_2_27_x86_64.manylinux_2_28_x86_64.whl (16.6
MB)
                         16.6/16.6 MB
143.2 MB/s eta 0:00:00
Downloading tzdata-2025.2-py2.py3-none-any.whl (347 kB)
Installing collected packages: tzdata, numpy, pandas
Successfully installed numpy-2.3.2 pandas-2.3.1 tzdata-2025.2
Collecting ipython-sql
  Downloading ipython_sql-0.5.0-py3-none-any.whl.metadata (17 kB)
Collecting prettytable
  Downloading prettytable-3.16.0-py3-none-any.whl.metadata (33 kB)
Requirement already satisfied: ipython in /opt/conda/lib/python3.12/site-
packages (from ipython-sql) (8.31.0)
Requirement already satisfied: sqlalchemy>=2.0 in
/opt/conda/lib/python3.12/site-packages (from ipython-sql) (2.0.37)
Collecting sqlparse (from ipython-sql)
  Downloading sqlparse-0.5.3-py3-none-any.whl.metadata (3.9 kB)
Requirement already satisfied: six in /opt/conda/lib/python3.12/site-packages
(from ipython-sql) (1.17.0)
Requirement already satisfied: ipython-genutils in
/opt/conda/lib/python3.12/site-packages (from ipython-sql) (0.2.0)
Requirement already satisfied: wcwidth in /opt/conda/lib/python3.12/site-
packages (from prettytable) (0.2.13)
Requirement already satisfied: greenlet!=0.4.17 in
```

```
/opt/conda/lib/python3.12/site-packages (from sqlalchemy>=2.0->ipython-sql)
(3.1.1)
Requirement already satisfied: typing-extensions>=4.6.0 in
/opt/conda/lib/python3.12/site-packages (from sqlalchemy>=2.0->ipython-sql)
(4.12.2)
Requirement already satisfied: decorator in /opt/conda/lib/python3.12/site-
packages (from ipython->ipython-sql) (5.1.1)
Requirement already satisfied: jedi>=0.16 in /opt/conda/lib/python3.12/site-
packages (from ipython->ipython-sql) (0.19.2)
Requirement already satisfied: matplotlib-inline in
/opt/conda/lib/python3.12/site-packages (from ipython->ipython-sql) (0.1.7)
Requirement already satisfied: pexpect>4.3 in /opt/conda/lib/python3.12/site-
packages (from ipython->ipython-sql) (4.9.0)
Requirement already satisfied: prompt_toolkit<3.1.0,>=3.0.41 in
/opt/conda/lib/python3.12/site-packages (from ipython->ipython-sql) (3.0.50)
Requirement already satisfied: pygments>=2.4.0 in
/opt/conda/lib/python3.12/site-packages (from ipython->ipython-sql) (2.19.1)
Requirement already satisfied: stack_data in /opt/conda/lib/python3.12/site-
packages (from ipython->ipython-sql) (0.6.3)
Requirement already satisfied: traitlets>=5.13.0 in
/opt/conda/lib/python3.12/site-packages (from ipython->ipython-sql) (5.14.3)
Requirement already satisfied: parso<0.9.0,>=0.8.4 in
/opt/conda/lib/python3.12/site-packages (from jedi>=0.16->ipython->ipython-sql)
(0.8.4)
Requirement already satisfied: ptyprocess>=0.5 in
/opt/conda/lib/python3.12/site-packages (from pexpect>4.3->ipython->ipython-sql)
(0.7.0)
Requirement already satisfied: executing>=1.2.0 in
/opt/conda/lib/python3.12/site-packages (from stack_data->ipython->ipython-sql)
(2.1.0)
Requirement already satisfied: asttokens>=2.1.0 in
/opt/conda/lib/python3.12/site-packages (from stack_data->ipython->ipython-sql)
(3.0.0)
Requirement already satisfied: pure_eval in /opt/conda/lib/python3.12/site-
packages (from stack_data->ipython->ipython-sql) (0.2.3)
Downloading ipython_sql-0.5.0-py3-none-any.whl (20 kB)
Downloading prettytable-3.16.0-py3-none-any.whl (33 kB)
Downloading sqlparse-0.5.3-py3-none-any.whl (44 kB)
Installing collected packages: sqlparse, prettytable, ipython-sql
Successfully installed ipython-sql-0.5.0 prettytable-3.16.0 sqlparse-0.5.3
```

1.1.5 Store the datasets in database tables

To analyze the data using SQL, it first needs to be loaded into SQLite DB. We will create three tables in as under:

1. **CENSUS_DATA**
2. **CHICAGO_PUBLIC_SCHOOLS**

3. CHICAGO_CRIME_DATA

Load the `pandas` and `sqlite3` libraries and establish a connection to `FinalDB.db`

[12]:
```
import pandas as pd
import sqlite3

conn = sqlite3.connect('FinalDB.db')

df_sql = pd.read_sql_query("SELECT * FROM sqlite_master WHERE type='table'",␣
 ↪conn )

df_sql.head()
```

```
   COMMUNITY_AREA_NUMBER COMMUNITY_AREA_NAME  PERCENT_OF_HOUSING_CROWDED  \
0                    1.0         Rogers Park                         7.7
1                    2.0          West Ridge                         7.8
2                    3.0              Uptown                         3.8
3                    4.0      Lincoln Square                         3.4
4                    5.0        North Center                         0.3

   PERCENT_HOUSEHOLDS_BELOW_POVERTY  PERCENT_AGED_16__UNEMPLOYED  \
0                              23.6                          8.7
1                              17.2                          8.8
2                              24.0                          8.9
3                              10.9                          8.2
4                               7.5                          5.2

   PERCENT_AGED_25__WITHOUT_HIGH_SCHOOL_DIPLOMA  \
0                                          18.2
1                                          20.8
2                                          11.8
3                                          13.4
4                                           4.5

   PERCENT_AGED_UNDER_18_OR_OVER_64  PER_CAPITA_INCOME  HARDSHIP_INDEX
0                              27.5              23939            39.0
1                              38.5              23040            46.0
2                              22.2              35787            20.0
3                              25.5              37524            17.0
4                              26.2              57123             6.0
   School_ID                                  NAME_OF_SCHOOL  \
0     610038                  Abraham Lincoln Elementary School
1     610281  Adam Clayton Powell Paideia Community Academy …
2     610185                Adlai E Stevenson Elementary School
3     609993                    Agustin Lara Elementary Academy
4     610513                      Air Force Academy High School
```

```
  Elementary, Middle, or High School         Street_Address     City State  \
0                                 ES        615 W Kemper Pl  Chicago    IL
1                                 ES  7511 S South Shore Dr  Chicago    IL
2                                 ES     8010 S Kostner Ave  Chicago    IL
3                                 ES     4619 S Wolcott Ave  Chicago    IL
4                                 HS        3630 S Wells St  Chicago    IL

   ZIP_Code    Phone_Number  \
0     60614  (773) 534-5720
1     60649  (773) 535-6650
2     60652  (773) 535-2280
3     60609  (773) 535-4389
4     60609  (773) 535-1590

                                                Link  \
0  http://schoolreports.cps.edu/SchoolProgressRep…
1  http://schoolreports.cps.edu/SchoolProgressRep…
2  http://schoolreports.cps.edu/SchoolProgressRep…
3  http://schoolreports.cps.edu/SchoolProgressRep…
4  http://schoolreports.cps.edu/SchoolProgressRep…

                     Network_Manager  … Freshman_on_Track_Rate__  \
0        Fullerton Elementary Network  …                      NDA
1           Skyway Elementary Network  …                      NDA
2           Midway Elementary Network  …                      NDA
3         Pershing Elementary Network  …                      NDA
4  Southwest Side High School Network  …                     91.8

  X_COORDINATE Y_COORDINATE   Latitude  Longitude COMMUNITY_AREA_NUMBER  \
0  1171699.458  1915829.428  41.924497 -87.644522                    7
1  1196129.985  1856209.466  41.760324 -87.556736                   43
2  1148427.165  1851012.215  41.747111 -87.731702                   70
3  1164504.290  1873959.199  41.809757 -87.672145                   61
4  1175177.622  1880745.126  41.828146 -87.632794                   34

  COMMUNITY_AREA_NAME  Ward Police_District                    Location
0        LINCOLN PARK    43              18  (41.92449696, -87.64452163)
1         SOUTH SHORE     7               4  (41.76032435, -87.55673627)
2             ASHBURN    13               8  (41.74711093, -87.73170248)
3            NEW CITY    20               9    (41.8097569, -87.6721446)
4       ARMOUR SQUARE    11               9  (41.82814609, -87.63279369)

[5 rows x 78 columns]
         ID CASE_NUMBER        DATE                     BLOCK IUCR  \
0   3512276    HK587712  2004-08-28        047XX S KEDZIE AVE  890
1   3406613    HK456306  2004-06-26  009XX N CENTRAL PARK AVE  820
2   8002131    HT233595  2011-04-04        043XX S WABASH AVE  820
3   7903289    HT133522  2010-12-30      083XX S KINGSTON AVE  840
```

```
4  10402076    HZ138551  2016-02-02           033XX W 66TH ST  820

  PRIMARY_TYPE                    DESCRIPTION          LOCATION_DESCRIPTION  \
0        THEFT                   FROM BUILDING            SMALL RETAIL STORE
1        THEFT                 $500 AND UNDER                         OTHER
2        THEFT                 $500 AND UNDER  NURSING HOME/RETIREMENT HOME
3        THEFT  FINANCIAL ID THEFT: OVER $300                     RESIDENCE
4        THEFT                 $500 AND UNDER                         ALLEY

   ARREST  DOMESTIC  …  DISTRICT  WARD  COMMUNITY_AREA_NUMBER  FBICODE  \
0   False     False  …         9  14.0                   58.0        6
1   False     False  …        11  27.0                   23.0        6
2   False     False  …         2   3.0                   38.0        6
3   False     False  …         4   7.0                   46.0        6
4   False     False  …         8  15.0                   66.0        6

  X_COORDINATE  Y_COORDINATE  YEAR   LATITUDE  LONGITUDE  \
0    1155838.0     1873050.0  2004  41.807440 -87.703956
1    1152206.0     1906127.0  2004  41.898280 -87.716406
2    1177436.0     1876313.0  2011  41.815933 -87.624642
3    1194622.0     1850125.0  2010  41.743665 -87.562463
4    1155240.0     1860661.0  2016  41.773455 -87.706480

                      LOCATION
0    (41.8074405, -87.703955849)
1  (41.898279962, -87.716405505)
2  (41.815933131, -87.624642127)
3  (41.743665322, -87.562462756)
4  (41.773455295, -87.706480471)

[5 rows x 21 columns]
```

[12]:

	type	name	tbl_name	rootpage	sql
0	table	CENSUS_DATA	CENSUS_DATA	2	CREATE TABLE "CENSUS_DATA" (\n"index" INTEGER,…
1	table	CHICAGO_PUBLIC_SCHOOLS	CHICAGO_PUBLIC_SCHOOLS	5	CREATE TABLE "CHICAGO_PUBLIC_SCHOOLS" (\n"inde…
2	table	CHICAGO_CRIME_DATA	CHICAGO_CRIME_DATA	105	CREATE TABLE "CHICAGO_CRIME_DATA" (\n"index" I…

Load the SQL magic module

[13]:
```
pip install ipython-sql
```

```
Requirement already satisfied: ipython-sql in /opt/conda/lib/python3.12/site-
packages (0.5.0)
Requirement already satisfied: prettytable in /opt/conda/lib/python3.12/site-
```

```
packages (from ipython-sql) (3.16.0)
Requirement already satisfied: ipython in /opt/conda/lib/python3.12/site-
packages (from ipython-sql) (8.31.0)
Requirement already satisfied: sqlalchemy>=2.0 in
/opt/conda/lib/python3.12/site-packages (from ipython-sql) (2.0.37)
Requirement already satisfied: sqlparse in /opt/conda/lib/python3.12/site-
packages (from ipython-sql) (0.5.3)
Requirement already satisfied: six in /opt/conda/lib/python3.12/site-packages
(from ipython-sql) (1.17.0)
Requirement already satisfied: ipython-genutils in
/opt/conda/lib/python3.12/site-packages (from ipython-sql) (0.2.0)
Requirement already satisfied: greenlet!=0.4.17 in
/opt/conda/lib/python3.12/site-packages (from sqlalchemy>=2.0->ipython-sql)
(3.1.1)
Requirement already satisfied: typing-extensions>=4.6.0 in
/opt/conda/lib/python3.12/site-packages (from sqlalchemy>=2.0->ipython-sql)
(4.12.2)
Requirement already satisfied: decorator in /opt/conda/lib/python3.12/site-
packages (from ipython->ipython-sql) (5.1.1)
Requirement already satisfied: jedi>=0.16 in /opt/conda/lib/python3.12/site-
packages (from ipython->ipython-sql) (0.19.2)
Requirement already satisfied: matplotlib-inline in
/opt/conda/lib/python3.12/site-packages (from ipython->ipython-sql) (0.1.7)
Requirement already satisfied: pexpect>4.3 in /opt/conda/lib/python3.12/site-
packages (from ipython->ipython-sql) (4.9.0)
Requirement already satisfied: prompt_toolkit<3.1.0,>=3.0.41 in
/opt/conda/lib/python3.12/site-packages (from ipython->ipython-sql) (3.0.50)
Requirement already satisfied: pygments>=2.4.0 in
/opt/conda/lib/python3.12/site-packages (from ipython->ipython-sql) (2.19.1)
Requirement already satisfied: stack_data in /opt/conda/lib/python3.12/site-
packages (from ipython->ipython-sql) (0.6.3)
Requirement already satisfied: traitlets>=5.13.0 in
/opt/conda/lib/python3.12/site-packages (from ipython->ipython-sql) (5.14.3)
Requirement already satisfied: wcwidth in /opt/conda/lib/python3.12/site-
packages (from prettytable->ipython-sql) (0.2.13)
Requirement already satisfied: parso<0.9.0,>=0.8.4 in
/opt/conda/lib/python3.12/site-packages (from jedi>=0.16->ipython->ipython-sql)
(0.8.4)
Requirement already satisfied: ptyprocess>=0.5 in
/opt/conda/lib/python3.12/site-packages (from pexpect>4.3->ipython->ipython-sql)
(0.7.0)
Requirement already satisfied: executing>=1.2.0 in
/opt/conda/lib/python3.12/site-packages (from stack_data->ipython->ipython-sql)
(2.1.0)
Requirement already satisfied: asttokens>=2.1.0 in
/opt/conda/lib/python3.12/site-packages (from stack_data->ipython->ipython-sql)
(3.0.0)
Requirement already satisfied: pure_eval in /opt/conda/lib/python3.12/site-
```

```
packages (from stack_data->ipython->ipython-sql) (0.2.3)
Note: you may need to restart the kernel to use updated packages.
```

Use `Pandas` to load the data available in the links above to dataframes. Use these dataframes to load data on to the database `FinalDB.db` as required tables.

[ ]:
```
df_census = pd.read_csv('ChicagoCensusData.csv')

print( df_census.head() )

df_census.to_sql( "CENSUS_DATA", conn, if_exists='replace' )


df_school = pd.read_csv('ChicagoPublicSchools.csv')

print( df_school.head() )

df_school.to_sql( "CHICAGO_PUBLIC_SCHOOLS", conn, if_exists='replace' )

df_crime = pd.read_csv('ChicagoCrimeData.csv')

print( df_crime.head() )

df_crime.to_sql( "CHICAGO_CRIME_DATA", conn, if_exists='replace' )
```

Establish a connection between SQL magic module and the database `FinalDB.db`

[17]:
```
# %load_ext sql

%reload_ext sql

%config SqlMagic.displaycon=False
%config SqlMagic.feedback=False
%config SqlMagic.autopandas=True

%sql sqlite:////root/.ipython/profile_default/history.sqlite
```

```
Traceback (most recent call last):
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/engine/base.py", line
146, in __init__
    self._dbapi_connection = engine.raw_connection()
                             ^^^^^^^^^^^^^^^^^^^^^^^

  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/engine/base.py", line
3298, in raw_connection
    return self.pool.connect()
           ^^^^^^^^^^^^^^^^^^^

  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
449, in connect
    return _ConnectionFairy._checkout(self)
```

```
          ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
1263, in _checkout
    fairy = _ConnectionRecord.checkout(pool)
            ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
712, in checkout
    rec = pool._do_get()
          ^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/impl.py", line
179, in _do_get
    with util.safe_reraise():
         ^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/util/langhelpers.py",
line 146, in __exit__
    raise exc_value.with_traceback(exc_tb)
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/impl.py", line
177, in _do_get
    return self._create_connection()
           ^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
390, in _create_connection
    return _ConnectionRecord(self)
           ^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
674, in __init__
    self.__connect()
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
900, in __connect
    with util.safe_reraise():
         ^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/util/langhelpers.py",
line 146, in __exit__
    raise exc_value.with_traceback(exc_tb)
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
896, in __connect
    self.dbapi_connection = connection = pool._invoke_creator(self)
                                         ^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/engine/create.py",
line 646, in connect
    return dialect.connect(*cargs, **cparams)
           ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/engine/default.py",
line 622, in connect
    return self.loaded_dbapi.connect(*cargs, **cparams)
           ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
sqlite3.OperationalError: unable to open database file
```

```
The above exception was the direct cause of the following exception:

Traceback (most recent call last):
  File "/opt/conda/lib/python3.12/site-packages/sql/magic.py", line 196, in
execute
    conn = sql.connection.Connection.set(
           ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sql/connection.py", line 70, in
set
    cls.current = existing or Connection(descriptor, connect_args, creator)
                              ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sql/connection.py", line 55, in
__init__
    self.internal_connection = engine.connect()
                               ^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/engine/base.py", line
3274, in connect
    return self._connection_cls(self)
           ^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/engine/base.py", line
148, in __init__
    Connection._handle_dbapi_exception_noconnection(
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/engine/base.py", line
2439, in _handle_dbapi_exception_noconnection
    raise sqlalchemy_exception.with_traceback(exc_info[2]) from e
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/engine/base.py", line
146, in __init__
    self._dbapi_connection = engine.raw_connection()
                             ^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/engine/base.py", line
3298, in raw_connection
    return self.pool.connect()
           ^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
449, in connect
    return _ConnectionFairy._checkout(self)
           ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
1263, in _checkout
    fairy = _ConnectionRecord.checkout(pool)
            ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
712, in checkout
    rec = pool._do_get()
          ^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/impl.py", line
179, in _do_get
    with util.safe_reraise():
```

```
        ^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/util/langhelpers.py",
line 146, in __exit__
    raise exc_value.with_traceback(exc_tb)
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/impl.py", line
177, in _do_get
    return self._create_connection()
           ^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
390, in _create_connection
    return _ConnectionRecord(self)
           ^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
674, in __init__
    self.__connect()
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
900, in __connect
    with util.safe_reraise():
         ^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/util/langhelpers.py",
line 146, in __exit__
    raise exc_value.with_traceback(exc_tb)
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/pool/base.py", line
896, in __connect
    self.dbapi_connection = connection = pool._invoke_creator(self)
                                         ^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/engine/create.py",
line 646, in connect
    return dialect.connect(*cargs, **cparams)
           ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/opt/conda/lib/python3.12/site-packages/sqlalchemy/engine/default.py",
line 622, in connect
    return self.loaded_dbapi.connect(*cargs, **cparams)
           ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
sqlalchemy.exc.OperationalError: (sqlite3.OperationalError) unable to open
database file
(Background on this error at: https://sqlalche.me/e/20/e3q8)

Connection info needed in SQLAlchemy format, example:
               postgresql://username:password@hostname/dbname
               or an existing connection: dict_keys([])
```

You can now proceed to the the following questions. Please note that a graded assignment will follow this lab and there will be a question on each of the problems stated below. It can be from the answer you received or the code you write for this problem. Therefore, please keep a note of both your codes as well as the response you generate.

1.2 Problems

Now write and execute SQL queries to solve assignment problems

1.2.1 Problem 1

Find the total number of crimes recorded in the CRIME table.

[26]:
```
df_crimeCount = pd.read_sql_query("SELECT COUNT(*) FROM CHICAGO_CRIME_DATA",␣
↪conn )

df_crimeCount
```

[26]:

	COUNT(*)
0	533

1.2.2 Problem 2

List community area names and numbers with per capita income less than 11000.

[27]:
```
df_income = pd.read_sql_query("SELECT COMMUNITY_AREA_NAME FROM CENSUS_DATA␣
↪WHERE PER_CAPITA_INCOME < 11000", conn )

df_income
```

[27]:

	COMMUNITY_AREA_NAME
0	West Garfield Park
1	South Lawndale
2	Fuller Park
3	Riverdale

1.2.3 Problem 3

List all case numbers for crimes involving minors?(children are not considered minors for the purposes of crime analysis)

[35]:
```
# df_caseNums = pd.read_sql_query("SELECT *FROM CHICAGO_CRIME_DATA ", conn )
df_caseNums = pd.read_sql_query("SELECT CASE_NUMBER,PRIMARY_TYPE,DESCRIPTION␣
↪FROM CHICAGO_CRIME_DATA WHERE PRIMARY_TYPE LIKE '%minor%' OR DESCRIPTION␣
↪like '%minor%'", conn )

df_caseNums
```

[35]:

	CASE_NUMBER	PRIMARY_TYPE	DESCRIPTION
0	HL266884	LIQUOR LAW VIOLATION	SELL/GIVE/DEL LIQUOR TO MINOR
1	HK238408	LIQUOR LAW VIOLATION	ILLEGAL CONSUMPTION BY MINOR

1.2.4 Problem 4

List all kidnapping crimes involving a child?

[105]:
```

```

```
df_kidnappings = pd.read_sql_query("SELECT CASE_NUMBER,PRIMARY_TYPE,DESCRIPTION␣
↪FROM CHICAGO_CRIME_DATA WHERE PRIMARY_TYPE LIKE '%kid%' OR DESCRIPTION like␣
↪'%child%'", conn )

df_kidnappings
```

[105]:

	CASE_NUMBER	PRIMARY_TYPE	DESCRIPTION
0	HN567387	OFFENSE INVOLVING CHILDREN	AGG SEX ASSLT OF CHILD FAM MBR
1	HR391350	OFFENSE INVOLVING CHILDREN	SEX ASSLT OF CHILD BY FAM MBR
2	HN144152	KIDNAPPING	CHILD ABDUCTION/STRANGER

1.2.5 Problem 5

List the kind of crimes that were recorded at schools. (No repetitions)

[46]:
```
df_schoolcrime = pd.read_sql_query("SELECT DISTINCT PRIMARY_TYPE FROM␣
↪CHICAGO_CRIME_DATA WHERE LOCATION_DESCRIPTION like '%school%'", conn )

df_schoolcrime
```

[46]:

	PRIMARY_TYPE
0	BATTERY
1	CRIMINAL DAMAGE
2	NARCOTICS
3	ASSAULT
4	CRIMINAL TRESPASS
5	PUBLIC PEACE VIOLATION

1.2.6 Problem 6

List the type of schools along with the average safety score for each type.

[104]:
```
df_safetyscore = pd.read_sql_query("SELECT \"Elementary, Middle, or High␣
↪School\" AS Type, AVG(SAFETY_SCORE) FROM CHICAGO_PUBLIC_SCHOOLS GROUP BY␣
↪Type", conn )

df_safetyscore
```

[104]:

	Type	AVG(SAFETY_SCORE)
0	ES	49.520384
1	HS	49.623529
2	MS	48.000000

1.2.7 Problem 7

List 5 community areas with highest % of households below poverty line

[66]:
```

```

```
df_bpl = pd.read_sql_query("SELECT␣
↪COMMUNITY_AREA_NUMBER,COMMUNITY_AREA_NAME,PERCENT_HOUSEHOLDS_BELOW_POVERTY␣
↪FROM CENSUS_DATA ORDER BY PERCENT_HOUSEHOLDS_BELOW_POVERTY DESC LIMIT 5",␣
↪conn )

df_bpl
```

[66]:

	COMMUNITY_AREA_NUMBER	COMMUNITY_AREA_NAME	PERCENT_HOUSEHOLDS_BELOW_POVERTY
0	54.0	Riverdale	56.5
1	37.0	Fuller Park	51.2
2	68.0	Englewood	46.6
3	29.0	North Lawndale	43.1
4	27.0	East Garfield Park	42.4

1.2.8 Problem 8

Which community area is most crime prone? Display the coumminty area number only.

[78]:
```
df_crimeprone = pd.read_sql_query("SELECT COMMUNITY_AREA_NUMBER FROM(SELECT␣
↪COUNT(PRIMARY_TYPE) AS CNT,COMMUNITY_AREA_NUMBER FROM CHICAGO_CRIME_DATA␣
↪GROUP BY COMMUNITY_AREA_NUMBER ORDER BY CNT DESC LIMIT 1)", conn )

df_crimeprone
```

[78]:

	COMMUNITY_AREA_NUMBER
0	25.0

Double-click **here** for a hint

1.2.9 Problem 9

Use a sub-query to find the name of the community area with highest hardship index

[97]:
```
df_hardship = pd.read_sql_query("SELECT COMMUNITY_AREA_NAME FROM(SELECT␣
↪MAX(HARDSHIP_INDEX),COMMUNITY_AREA_NAME FROM CENSUS_DATA)", conn )

df_hardship
```

[97]:

	COMMUNITY_AREA_NAME
0	Riverdale

1.2.10 Problem 10

Use a sub-query to determine the Community Area Name with most number of crimes?

[103]:
```

```

```
df_mostcrimes = pd.read_sql_query("SELECT DISTINCT COMMUNITY_AREA_NAME FROM␣
 ↪CHICAGO_PUBLIC_SCHOOLS WHERE COMMUNITY_AREA_NUMBER =(SELECT␣
 ↪COMMUNITY_AREA_NUMBER FROM(SELECT COUNT(PRIMARY_TYPE) AS␣
 ↪CNT,COMMUNITY_AREA_NUMBER FROM CHICAGO_CRIME_DATA GROUP BY␣
 ↪COMMUNITY_AREA_NUMBER ORDER BY CNT DESC LIMIT 1))", conn )

df_mostcrimes
```

[103]:

	COMMUNITY_AREA_NAME
0	AUSTIN

1.3 Author(s)

Hima Vasudevan

Rav Ahuja

Ramesh Sannreddy

1.4 Contribtuor(s)

Malika Singla

Abhishek Gagneja

##

© IBM Corporation 2023. All rights reserved.