Working with a real world data-set using SQL and Python

Estaimted time needed: **30** minutes

Objectives

After complting this lab you will be able to:

- Understand the dataset for Chicago Public School level performance
- Store the dataset in SQLite database.
- Retrieve metadata about tables and columns and query data from mixed case columns
- Solve example problems to practice your SQL skills including using built-in database functions

Chicago Public Schools - Progress Report Cards (2011-2012)

The city of Chicago released a dataset showing all school level performance data used to create School Report Cards for the 2011-2012 school year. The dataset is available from the Chicago Data Portal: https://data.cityofchicago.org/Education/Chicago-Public-Schools-Progress-Report-Cards-2011-/9xs2-f89t

This dataset includes a large number of metrics. Start by familiarizing yourself with the types of metrics in the database: https://data.cityofchicago.org/api/assets/AAD41A13-BE8A-4E67-B1F5-86E711E09D5F?download=true

NOTE:

Do not download the dataset directly from City of Chicago portal. Instead download a static copy which is a more database friendly version from this link.

Now review some of its contents.

Connect to the database

Let us now load the ipython-sql extension and establish a connection with the database

The syntax for connecting to magic sql using sqllite is

%sql sqlite://DatabaseName

where DatabaseName will be your **.db** file

In [1]:
```
import csv, sqlite3

con = sqlite3.connect("RealWorldData.db")
cur = con.cursor()
```

In [2]:
```
!pip install -q pandas==1.1.5
```

In [4]:
```
%load_ext sql
```

```
The sql extension is already loaded. To reload it, use:
  %reload_ext sql
```

In [5]:
```
%sql sqlite:///RealWorldData.db
```

Out[5]:
```
'Connected: @RealWorldData.db'
```

Store the dataset in a Table

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

using SQL, it first needs to be stored in the database.

We will first read the csv files from the given url into pandas dataframes

Next we will be using the df.to_sql() function to convert each csv file to a table in sqlite with the csv data loaded in it.

In [6]:
```
import pandas
df = pandas.read_csv("https://cf-courses-data.s3.us.cloud-object-storage.appdomain.
df.to_sql("CENSUS_DATA", con, if_exists='replace', index=False,method="multi")

df = pandas.read_csv("https://cf-courses-data.s3.us.cloud-object-storage.appdomain.
df.to_sql("CHICAGO_CRIME_DATA", con, if_exists='replace', index=False, method="mult

df = pandas.read_csv("https://cf-courses-data.s3.us.cloud-object-storage.appdomain.
df.to_sql("CHICAGO_PUBLIC_SCHOOLS_DATA", con, if_exists='replace', index=False, met
```

```
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages/pandas/core/generic.p
y:2615: UserWarning: The spaces in these column names will not be changed. In pandas
versions < 0.14, spaces were converted to underscores.
  method=method,
```

In [6]:
```
import pandas
df = pandas.read_csv("https://cf-courses-data.s3.us.cloud-object-storage.appdomain.
df.to_sql("CENSUS_DATA", con, if_exists='replace', index=False,method="multi")
```

```
df = pandas.read_csv("https://cf-courses-data.s3.us.cloud-object-storage.appdomain.
df.to_sql("CHICAGO_CRIME_DATA", con, if_exists='replace', index=False, method="mult

df = pandas.read_csv("https://cf-courses-data.s3.us.cloud-object-storage.appdomain.
df.to_sql("CHICAGO_PUBLIC_SCHOOLS_DATA", con, if_exists='replace', index=False, met
```

```
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages/pandas/core/generic.p
y:2615: UserWarning: The spaces in these column names will not be changed. In pandas
versions < 0.14, spaces were converted to underscores.
  method=method,
```

Double-click **here** for the solution.

Query the database system catalog to retrieve table metadata

You can verify that the table creation was successful by retrieving the list of all tables in your schema and checking whether the SCHOOLS table was created

In [7]:

```
# type in your query to retrieve list of all tables in the database
%sql SELECT name FROM sqlite_master WHERE type='table'
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[7]:

name
CENSUS_DATA
CHICAGO_CRIME_DATA
CHICAGO_PUBLIC_SCHOOLS_DATA

Double-click **here** for a hint

Double-click **here** for the solution.

Query the database system catalog to retrieve column metadata

The SCHOOLS table contains a large number of columns. How many columns does this table have?

In [8]:

```
# type in your query to retrieve the number of columns in the SCHOOLS table
%sql SELECT COUNT(name) FROM PRAGMA_TABLE_INFO ('CHICAGO_PUBLIC_SCHOOLS_DATA');
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[8]:

COUNT(name)
78

Double-click **here** for the solution.

Now retrieve the the list of columns in SCHOOLS table and their column type (datatype) and length.

In [9]:
```
# type in your query to retrieve all column names in the SCHOOLS table along with t
%sql SELECT name,type,length(type) FROM PRAGMA_TABLE_INFO('CHICAGO_PUBLIC_SCHOOLS_D
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[9]:

name	type	length(type)
School_ID	INTEGER	7
NAME_OF_SCHOOL	TEXT	4
Elementary, Middle, or High School	TEXT	4
Street_Address	TEXT	4
City	TEXT	4
State	TEXT	4
ZIP_Code	INTEGER	7
Phone_Number	TEXT	4
Link	TEXT	4
Network_Manager	TEXT	4
Collaborative_Name	TEXT	4
Adequate_Yearly_Progress_Made_	TEXT	4
Track_Schedule	TEXT	4
CPS_Performance_Policy_Status	TEXT	4
CPS_Performance_Policy_Level	TEXT	4
HEALTHY_SCHOOL_CERTIFIED	TEXT	4
Safety_Icon	TEXT	4
SAFETY_SCORE	REAL	4
Family_Involvement_Icon	TEXT	4
Family_Involvement_Score	TEXT	4
Environment_Icon	TEXT	4
Environment_Score	REAL	4
Instruction_Icon	TEXT	4
Instruction_Score	REAL	4
Leaders_Icon	TEXT	4
Leaders_Score	TEXT	4
Teachers_Icon	TEXT	4
Teachers_Score	TEXT	4
Parent_Engagement_Icon	TEXT	4
Parent_Engagement_Score	TEXT	4

name	type	length(type)
Parent_Environment_Icon	TEXT	4
Parent_Environment_Score	TEXT	4
AVERAGE_STUDENT_ATTENDANCE	TEXT	4
Rate_of_Misconducts__per_100_students_	REAL	4
Average_Teacher_Attendance	TEXT	4
Individualized_Education_Program_Compliance_Rate	TEXT	4
Pk_2_Literacy__	TEXT	4
Pk_2_Math__	TEXT	4
Gr3_5_Grade_Level_Math__	TEXT	4
Gr3_5_Grade_Level_Read__	TEXT	4
Gr3_5_Keep_Pace_Read__	TEXT	4
Gr3_5_Keep_Pace_Math__	TEXT	4
Gr6_8_Grade_Level_Math__	TEXT	4
Gr6_8_Grade_Level_Read__	TEXT	4
Gr6_8_Keep_Pace_Math_	TEXT	4
Gr6_8_Keep_Pace_Read__	TEXT	4
Gr_8_Explore_Math__	TEXT	4
Gr_8_Explore_Read__	TEXT	4
ISAT_Exceeding_Math__	REAL	4
ISAT_Exceeding_Reading__	REAL	4
ISAT_Value_Add_Math	REAL	4
ISAT_Value_Add_Read	REAL	4
ISAT_Value_Add_Color_Math	TEXT	4
ISAT_Value_Add_Color_Read	TEXT	4
Students_Taking__Algebra__	TEXT	4
Students_Passing__Algebra__	TEXT	4
9th Grade EXPLORE (2009)	TEXT	4
9th Grade EXPLORE (2010)	TEXT	4
10th Grade PLAN (2009)	TEXT	4
10th Grade PLAN (2010)	TEXT	4

name	type	length(type)
Net_Change_EXPLORE_and_PLAN	TEXT	4
11th Grade Average ACT (2011)	TEXT	4
Net_Change_PLAN_and_ACT	TEXT	4
College_Eligibility__	TEXT	4
Graduation_Rate__	TEXT	4
College_Enrollment_Rate__	TEXT	4
COLLEGE_ENROLLMENT	INTEGER	7
General_Services_Route	INTEGER	7
Freshman_on_Track_Rate__	TEXT	4
X_COORDINATE	REAL	4
Y_COORDINATE	REAL	4
Latitude	REAL	4
Longitude	REAL	4
COMMUNITY_AREA_NUMBER	INTEGER	7
COMMUNITY_AREA_NAME	TEXT	4
Ward	INTEGER	7
Police_District	INTEGER	7
Location	TEXT	4

Double-click **here** for the solution.

Questions

1. Is the column name for the "SCHOOL ID" attribute in upper or mixed case?
2. What is the name of "Community Area Name" column in your table? Does it have spaces?
3. Are there any columns in whose names the spaces and paranthesis (round brackets) have been replaced by the underscore character "_"?

Problems

Problem 1

How many Elementary Schools are in the dataset?

```
In [12]: %sql SELECT 'Elementary_Middle_or_High_School' from PRAGMA_TABLE_INFO ('CHICAGO_PUB
```

```
 * sqlite:///RealWorldData.db
(sqlite3.OperationalError) no such column: Elementary_Middle_or_High_School
[SQL: SELECT Elementary_Middle_or_High_School from PRAGMA_TABLE_INFO ('CHICAGO_PUBLI
C_SCHOOLS_DATA') WHERE (select COUNT("Elementary_Middle_or_High_School") = 'ES');]
(Background on this error at: http://sqlalche.me/e/13/e3q8)
```

Double-click **here** for a hint

Double-click **here** for another hint

Double-click **here** for the solution.

Problem 2

What is the highest Safety Score?

```
In [13]: %sql select MAX(Safety_Score) from CHICAGO_PUBLIC_SCHOOLS_DATA ;
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[13]:

MAX(Safety_Score)
99.0

Double-click **here** for a hint

Double-click **here** for the solution.

Problem 3

Which schools have highest Safety Score?

```
In [24]: %sql select Name_of_School, Safety_Score from CHICAGO_PUBLIC_SCHOOLS_DATA where \
           Safety_Score= (select MAX(Safety_Score) from CHICAGO_PUBLIC_SCHOOLS_DATA);
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[24]:

NAME_OF_SCHOOL	SAFETY_SCORE
Abraham Lincoln Elementary School	99.0
Alexander Graham Bell Elementary School	99.0
Annie Keller Elementary Gifted Magnet School	99.0
Augustus H Burley Elementary School	99.0
Edgar Allan Poe Elementary Classical School	99.0
Edgebrook Elementary School	99.0
Ellen Mitchell Elementary School	99.0
James E McDade Elementary Classical School	99.0
James G Blaine Elementary School	99.0
LaSalle Elementary Language Academy	99.0
Mary E Courtenay Elementary Language Arts Center	99.0
Northside College Preparatory High School	99.0
Northside Learning Center High School	99.0
Norwood Park Elementary School	99.0
Oriole Park Elementary School	99.0
Sauganash Elementary School	99.0
Stephen Decatur Classical Elementary School	99.0
Talman Elementary School	99.0
Wildwood Elementary School	99.0

Double-click **here** for the solution.

Problem 4

What are the top 10 schools with the highest "Average Student Attendance"?

In [27]:

```
%sql select Name_of_School, Average_Student_Attendance from CHICAGO_PUBLIC_SCHOOLS_
  order by Average_Student_Attendance desc nulls last LIMIT 10;
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[27]:

NAME_OF_SCHOOL	AVERAGE_STUDENT_ATTENDANCE
John Charles Haines Elementary School	98.40%
James Ward Elementary School	97.80%
Edgar Allan Poe Elementary Classical School	97.60%
Orozco Fine Arts & Sciences Elementary School	97.60%
Rachel Carson Elementary School	97.60%
Annie Keller Elementary Gifted Magnet School	97.50%
Andrew Jackson Elementary Language Academy	97.40%
Lenart Elementary Regional Gifted Center	97.40%
Disney II Magnet School	97.30%
John H Vanderpoel Elementary Magnet School	97.20%

Double-click **here** for the solution.

Problem 5

Retrieve the list of 5 Schools with the lowest Average Student Attendance sorted in ascending order based on attendance

In [30]:

```
%sql select Name_of_School, Average_Student_Attendance from CHICAGO_PUBLIC_SCHOOLS_
  order by Average_Student_Attendance nulls last LIMIT 5;
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[30]:

NAME_OF_SCHOOL	AVERAGE_STUDENT_ATTENDANCE
Richard T Crane Technical Preparatory High School	57.90%
Barbara Vick Early Childhood & Family Center	60.90%
Dyett High School	62.50%
Wendell Phillips Academy High School	63.00%
Orr Academy High School	66.30%

Double-click **here** for the solution.

Problem 6

Now remove the '%' sign from the above result set for Average Student Attendance column

In [32]:

```
%sql SELECT Name_of_School, REPLACE(Average_Student_Attendance, '%','')  \
     from CHICAGO_PUBLIC_SCHOOLS_DATA \
     order by Average_Student_Attendance \
     LIMIT 5
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[32]:

NAME_OF_SCHOOL	REPLACE(Average_Student_Attendance, '%','')
Velma F Thomas Early Childhood Center	None
Richard T Crane Technical Preparatory High School	57.90
Barbara Vick Early Childhood & Family Center	60.90
Dyett High School	62.50
Wendell Phillips Academy High School	63.00

Double-click **here** for a hint

Double-click **here** for the solution.

Problem 7

Which Schools have Average Student Attendance lower than 70%?

In [37]:

```
%sql SELECT Name_of_School, Average_Student_Attendance  \
    from CHICAGO_PUBLIC_SCHOOLS_DATA \
    where CAST ( REPLACE(Average_Student_Attendance, '%', '') AS DOUBLE ) < 70 \
    order by Average_Student_Attendance
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[37]:

NAME_OF_SCHOOL	AVERAGE_STUDENT_ATTENDANCE
Richard T Crane Technical Preparatory High School	57.90%
Barbara Vick Early Childhood & Family Center	60.90%
Dyett High School	62.50%
Wendell Phillips Academy High School	63.00%
Orr Academy High School	66.30%
Manley Career Academy High School	66.80%
Chicago Vocational Career Academy High School	68.80%
Roberto Clemente Community Academy High School	69.60%

Double-click **here** for a hint

Double-click **here** for another hint

Double-click **here** for the solution.

Problem 8

Get the total College Enrollment for each Community Area

In [41]:

```
%sql SELECT COMMUNITY_AREA_NAME, College_Enrollment_Rate__  \
     from CHICAGO_PUBLIC_SCHOOLS_DATA \
     where (SELECT sum(College_Enrollment_Rate__) FROM CHICAGO_PUBLIC_SCHOOLS_DATA)
     GROUP BY COMMUNITY_AREA_NAME;
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[41]:

COMMUNITY_AREA_NAME	College_Enrollment_Rate__
ALBANY PARK	NDA
ARCHER HEIGHTS	61.3
ARMOUR SQUARE	NDA
ASHBURN	NDA
AUBURN GRESHAM	NDA
AUSTIN	51.3
AVALON PARK	NDA
AVONDALE	NDA
BELMONT CRAGIN	NDA
BEVERLY	NDA
BRIDGEPORT	NDA
BRIGHTON PARK	NDA
BURNSIDE	NDA
CALUMET HEIGHTS	NDA
CHATHAM	NDA
CHICAGO LAWN	NDA
CLEARING	NDA
DOUGLAS	69.3
DUNNING	NDA
EAST GARFIELD PARK	64.2
EAST SIDE	NDA
EDGEWATER	NDA
EDISON PARK	NDA
ENGLEWOOD	NDA
FOREST GLEN	NDA
FULLER PARK	NDA
GAGE PARK	NDA
GARFIELD RIDGE	52.9
GRAND BOULEVARD	NDA
GREATER GRAND CROSSING	NDA

COMMUNITY_AREA_NAME	College_Enrollment_Rate__
HEGEWISCH	NDA
HERMOSA	NDA
HUMBOLDT PARK	NDA
HYDE PARK	NDA
IRVING PARK	40.4
JEFFERSON PARK	NDA
KENWOOD	NDA
LAKE VIEW	NDA
LINCOLN PARK	NDA
LINCOLN SQUARE	NDA
LOGAN SQUARE	NDA
LOOP	86.9
LOWER WEST SIDE	48.8
MCKINLEY PARK	NDA
MONTCLARE	NDA
MORGAN PARK	NDA
MOUNT GREENWOOD	NDA
NEAR NORTH SIDE	NDA
NEAR SOUTH SIDE	NDA
NEAR WEST SIDE	NDA
NEW CITY	NDA
NORTH CENTER	79.8
NORTH LAWNDALE	NDA
NORTH PARK	78.5
NORWOOD PARK	NDA
OAKLAND	NDA
OHARE	NDA
PORTAGE PARK	40.9
PULLMAN	NDA
RIVERDALE	60.5

COMMUNITY_AREA_NAME	College_Enrollment_Rate__
ROGERS PARK	NDA
ROSELAND	NDA
SOUTH CHICAGO	NDA
SOUTH DEERING	NDA
SOUTH LAWNDALE	NDA
SOUTH SHORE	NDA
UPTOWN	NDA
WASHINGTON HEIGHTS	NDA
WASHINGTON PARK	NDA
WEST ELSDON	NDA
WEST ENGLEWOOD	NDA
WEST GARFIELD PARK	NDA
WEST LAWN	NDA
WEST PULLMAN	NDA
WEST RIDGE	NDA
WEST TOWN	NDA
WOODLAWN	NDA

Double-click **here** for a hint

Double-click **here** for another hint

Double-click **here** for the solution.

Problem 9

Get the 5 Community Areas with the least total College Enrollment sorted in ascending order

In [43]:

```
%sql select Community_Area_Name, sum(College_Enrollment) AS TOTAL_ENROLLMENT \
   from CHICAGO_PUBLIC_SCHOOLS_DATA \
   group by Community_Area_Name  \
   ORDER BY sum (College_Enrollment) \
   LIMIT 5
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[43]:

COMMUNITY_AREA_NAME	TOTAL_ENROLLMENT
OAKLAND	140
FULLER PARK	531
BURNSIDE	549
OHARE	786
LOOP	871

Double-click **here** for a hint

Double-click **here** for the solution.

Problem 10

List 5 schools with lowest safety score.

```
In [46]: %sql select Name_of_School, Safety_Score \
            from CHICAGO_PUBLIC_SCHOOLS_DATA where Safety_Score!= 'None' \
            order by Safety_Score asc \
            LIMIT 5
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[46]:

NAME_OF_SCHOOL	SAFETY_SCORE
Edmond Burke Elementary School	1.0
Luke O'Toole Elementary School	5.0
George W Tilton Elementary School	6.0
Foster Park Elementary School	11.0
Emil G Hirsch Metropolitan High School	13.0

Double-click **here** for the solution.

Problem 11

Get the hardship index for the community area which has College Enrollment of 4368

```
In [55]: %sql select hardship_index from Census_data ce, CHICAGO_PUBLIC_SCHOOLS_DATA ch \
             where ce.Community_area_number = ch.Community_area_number \
             and College_Enrollment = 4368;
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[55]:

HARDSHIP_INDEX
6.0

Double-click **here** for the solution.

Problem 12

Get the hardship index for the community area which has the highest value for College Enrollment

In [56]:

```
%sql select hardship_index from Census_data ce, CHICAGO_PUBLIC_SCHOOLS_DATA ch \
    where ce.Community_area_number = ch.Community_area_number \
    select Community_area_number MAX(College_Enrollment);
```

```
 * sqlite:///RealWorldData.db
Done.
```

Out[56]:

HARDSHIP_INDEX
39.0
39.0
39.0
39.0
39.0
39.0
46.0
46.0
46.0
46.0
46.0
46.0
46.0
46.0
46.0
20.0
20.0
20.0
20.0
20.0
20.0
20.0
17.0
17.0
17.0
17.0
17.0
6.0
6.0
6.0

HARDSHIP_INDEX
6.0
6.0
6.0
6.0
5.0
5.0
5.0
5.0
5.0
5.0
5.0
5.0
5.0
5.0
5.0
2.0
2.0
2.0
2.0
2.0
2.0
2.0
1.0
1.0
1.0
1.0
1.0
1.0
1.0
8.0

HARDSHIP_INDEX
8.0
21.0
21.0
21.0
21.0
21.0
21.0
21.0
21.0
25.0
25.0
11.0
11.0
11.0
33.0
33.0
33.0
33.0
33.0
53.0
53.0
53.0
53.0
53.0
53.0
53.0
53.0
35.0
35.0
35.0

HARDSHIP_INDEX
35.0
35.0
35.0
35.0
34.0
34.0
34.0
34.0
34.0
34.0
34.0
34.0
34.0
28.0
28.0
28.0
28.0
28.0
28.0
50.0
70.0
70.0
70.0
70.0
70.0
70.0
70.0
70.0
70.0
70.0

HARDSHIP_INDEX
70.0
70.0
71.0
71.0
71.0
71.0
42.0
42.0
42.0
42.0
23.0
23.0
23.0
23.0
23.0
23.0
23.0
23.0
23.0
23.0
23.0
85.0
85.0
85.0
85.0
85.0
85.0
85.0
85.0
85.0

HARDSHIP_INDEX
85.0
85.0
85.0
85.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
10.0
73.0
73.0
73.0
73.0
73.0
73.0

HARDSHIP_INDEX
73.0
73.0
73.0
73.0
73.0
73.0
73.0
73.0
73.0
73.0
73.0
73.0
73.0
73.0
73.0
73.0
73.0
92.0
92.0
92.0
92.0
92.0
92.0
92.0
92.0
83.0
83.0
83.0
83.0
83.0

HARDSHIP_INDEX
83.0
83.0
83.0
83.0
83.0
83.0
83.0
83.0
15.0
15.0
15.0
15.0
15.0
15.0
15.0
15.0
15.0
15.0
15.0
15.0
15.0
15.0
15.0
15.0
87.0
87.0
87.0
87.0
87.0
87.0

HARDSHIP_INDEX
87.0
87.0
87.0
87.0
87.0
87.0
87.0
87.0
87.0
87.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0
96.0

HARDSHIP_INDEX
96.0
96.0
76.0
76.0
76.0
76.0
76.0
76.0
76.0
76.0
76.0
76.0
76.0
3.0
7.0
7.0
7.0
82.0
82.0
82.0
47.0
47.0
47.0
47.0
47.0
47.0
47.0
47.0
47.0
47.0

HARDSHIP_INDEX
47.0
78.0
97.0
97.0
57.0
57.0
57.0
57.0
57.0
57.0
57.0
57.0
26.0
26.0
26.0
26.0
26.0
26.0
26.0
88.0
88.0
88.0
88.0
88.0
14.0
14.0
14.0
14.0
58.0
58.0

HARDSHIP_INDEX
58.0
58.0
58.0
58.0
58.0
58.0
55.0
55.0
55.0
55.0
55.0
55.0
55.0
55.0
55.0
60.0
60.0
60.0
60.0
60.0
60.0
60.0
60.0
60.0
41.0
41.0
41.0
75.0
75.0
75.0

HARDSHIP_INDEX
75.0
75.0
75.0
75.0
79.0
38.0
38.0
38.0
38.0
38.0
38.0
52.0
52.0
52.0
52.0
52.0
52.0
52.0
52.0
52.0
52.0
52.0
52.0
52.0
51.0
51.0
51.0
51.0
51.0
65.0

HARDSHIP_INDEX
65.0
65.0
65.0
64.0
64.0
64.0
64.0
64.0
62.0
62.0
62.0
62.0
62.0
62.0
62.0
62.0
62.0
62.0
98.0
98.0
98.0
98.0
44.0
44.0
32.0
32.0
32.0
32.0
32.0
67.0

HARDSHIP_INDEX
67.0
84.0
84.0
84.0
84.0
84.0
84.0
84.0
84.0
61.0
61.0
61.0
61.0
43.0
43.0
43.0
43.0
43.0
91.0
91.0
91.0
91.0
91.0
91.0
91.0
91.0
91.0
91.0
91.0
91.0

HARDSHIP_INDEX
91.0
69.0
69.0
69.0
93.0
93.0
93.0
93.0
93.0
93.0
93.0
93.0
93.0
93.0
29.0
29.0
29.0
29.0
56.0
56.0
56.0
56.0
80.0
80.0
80.0
80.0
80.0
80.0
80.0
89.0

HARDSHIP_INDEX
89.0
89.0
89.0
89.0
89.0
89.0
89.0
89.0
89.0
89.0
89.0
89.0
94.0
94.0
94.0
94.0
94.0
94.0
94.0
94.0
94.0
94.0
94.0
94.0
94.0
94.0
94.0
94.0
94.0
66.0

HARDSHIP_INDEX
66.0
66.0
66.0
66.0
66.0
66.0
66.0
66.0
66.0
37.0
37.0
37.0
37.0
37.0
37.0
37.0
37.0
74.0
74.0
74.0
74.0
74.0
74.0
74.0
74.0
74.0
74.0
12.0
12.0
12.0

HARDSHIP_INDEX
12.0
48.0
48.0
48.0
48.0
48.0
48.0
48.0
48.0
48.0
16.0
16.0
16.0
16.0
30.0
30.0
30.0
30.0
30.0
24.0
19.0
19.0
19.0
19.0
19.0
19.0

Double-click **here** for the solution.

Summary

In this lab you learned how to work with a real word dataset using SQL and Python. You learned how to query columns with spaces or special characters in their names and with

mixed case names. You also used built in database functions and practiced how to sort, limit, and order result sets, as well as used sub-queries and worked with multiple tables.

Author

Rav Ahuja

Change Log

Date (YYYY-MM-DD)	Version	Changed By	Change Description
2022-03-04	2.2	Lakshmi Holla	Made changes in markdown cells
2020-11-27	2.1	Sannareddy Ramesh	Modified data sets and added new problems
2020-08-28	2.0	Lavanya	Moved lab to course repo in GitLab