QMB 6315: Python for Business Analytics
College of Business
University of Central Florida
Spring 2025

Assignment 3

Due Friday, April 18, 2025 at 11:59 PM
in digital form in your GitHub repository.

Instructions:

Complete this assignment on your private GitHub repo in a folder called `assignment_03`. In this folder, save your answers to Questions 1 and 2 in a file called `my_A3_queries.py`, following the script discussed in class in the file `PP_Ch_17B.py` in the course repository. When you are finished, submit it by uploading your files to your GitHub repo using any one of the approaches outlined in Question 3. You are free to discuss your approach to each question with your classmates but you must upload your own work.

Question 1:

The spreadsheet `US_state_pop_area.csv` is a listing of population figures and land area for US states and territories. In this exercise, you will create a table to store the values of population and land area and then use the table to retreive data in a variety of forms. In the sample script called `my_A3_queries.py`, the following steps have been completed.

a) Create a new database called `population_USA.db`.

b) Make a database table called `Density` that will hold the name of the state or territory (TEXT), the population (INTEGER), and the land area (REAL).

c) Insert the data from the table above.

Now, your task is to write Python code with SQL queries to do the following.

d) Retrieve the contents of the table.

e) Retrieve the populations.

f) Retrieve the states that have populations of less than one million.

g) Retrieve the states that have populations of less than one million or greater than five million.

h) Retrieve the states that *do not* have populations of less than one million or greater than five million.

Question 2:

The spreadsheet `US_cap_cities_pop.csv` is a listing of population figures and land area for the capital cities of US states. In this exercise, you will add a new table called `Capitals` to store the values of population and land area. `Capitals` has four columns: State (TEXT), Capital (TEXT), Area (REAL), and Population (INTEGER). In the sample script called `my_A3_queries.py`, the following steps have been completed.

a) Retrieve the contents of the table.

b) Retrieve the populations of the states and capitals (in a list of tuples of the form `[state_population, capital_population]`).

Now, your task is to write Python code with SQL queries to do the following.

c) Retrieve the land area of the states whose capitals that have populations greater than 100,000.

d) Retrieve the states with land densities greater than ten people per square mile and capital city populations more than 500,000.

e) Retrieve the total land area of the US.

f) Retrieve the average population of the capital cities.

g) Retrieve the lowest population of the capital cities.

h) Retrieve the highest population of the states or territories.

Question 3:

Upload your code to your GitHub repo using the interface in GitHub Desktop.

1. Save your file within the folder in your repository in GitHub Desktop.

2. When you see the changes in GitHub Desktop, add a description of the changes you are making in the bottom left panel.

3. Press the button "Commit to main" to commit those changes.

4. Press the button "Push origin" to push the changes to the online repository. After this step, the changes should be visible on a browser, after refreshing the page.