Software Engineering + Data Wrangling with SQL – Combined Project

Cohort A19

Author: Shutima Potivorakun

Subject: Python application to accommodate SQL stored procedure and trigger algorithm

Date: 18th April 2020

Table of Contents

1. Project Overview

This project is aimed to verify the abilities to combine clean software engineering using Python as for implementation and Data Wrangling with SQL.

2. Deliverables

The final results of the project contain:

- This documentation description the work done for the project (this document)
- The completed Python applications: GetAllSurveyData.py, myScript.py (located in "SQL_PythonProject_ShutimaP" folder)
    - **GetAllSurveyData.py** – acts as module.py, a library that is being imported in myScript.py, it contains the class and all functions in the class
    - **myScript.py** – acts as the main program
- The CSV file of all survey data freshly extracted from the application

3. Database File

For this project, we use the database dump file "SurveySample_A19.bak" which can be downloaded from:
https://a19.moodle.dsti.institute/pluginfile.php/2752/question/questiontext/841/1/3605/Survey_Sample_A19.zip

4. Limitations / Possible Improvements

- There should be an improvement in the Python Class design. At the moment, there are some part in my Python class that is more of the library for the functions rather than a real object (or API).
- The database connection variables in SQLDatabase class need to be changed manually, for *DRIVER*, *SERVER* and *DATABASE*.

```
class SQLDatabase:

    def __init__(self, view_Name, export_file_path):
        self.sql_conn = odbc.connect(DRIVER='{ODBC Driver 17 for SQL Server}',
                                     SERVER='LAPTOP-NAD8U5G4',
                                     DATABASE='Survey_Sample_A19',
                                     Trusted_Connection='yes')
```

5. Getting Started

In order to run this program, you need to have Python 3.x minimum installed on your environment.

6. Operating System

This project was implemented on Windows 10 Professional Operating System using the following application:

- Microsoft SQL Server Management Studio 15.0.18206.0
- Visual Studio Code running on Python Virtual Environment with Python base version 3.7.7

7. Preparation and Execution Steps

Once successfully preparing virtual environment with Python base application of 3.x minimum, please follow the steps below to run the application.

- Copy the project folder “SQL_PythonProject_ShutimaP” into your machine
- Navigate to the project folder, in the file “GetAllSurveyData.py”, in the init section of SQLDatabase class, please modify the values for the following variables to match your system (*DRIVER*, *SERVER*, *DATABASE*).

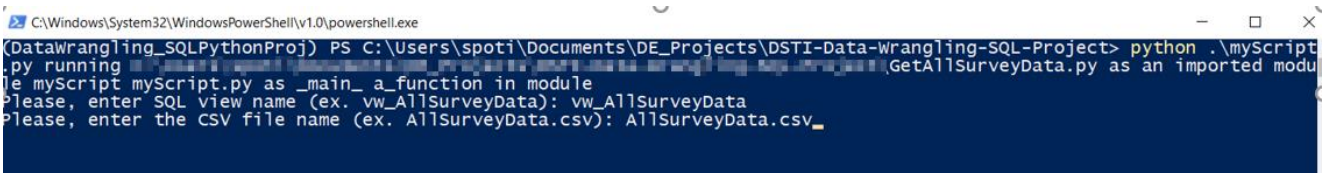
```
class SQLDatabase:

    def __init__(self, view_Name, export file path):
        self.sql_conn = odbc.connect(DRIVER='{ODBC Driver 17 for SQL Server}',
                                     SERVER='LAPTOP-NAD8U5G4',
                                     DATABASE='Survey_Sample_A19',
                                     Trusted_Connection='yes')
```

- Open command prompt (for example PowerShell), then navigate to the project folder
- Execute the following command to run the application:

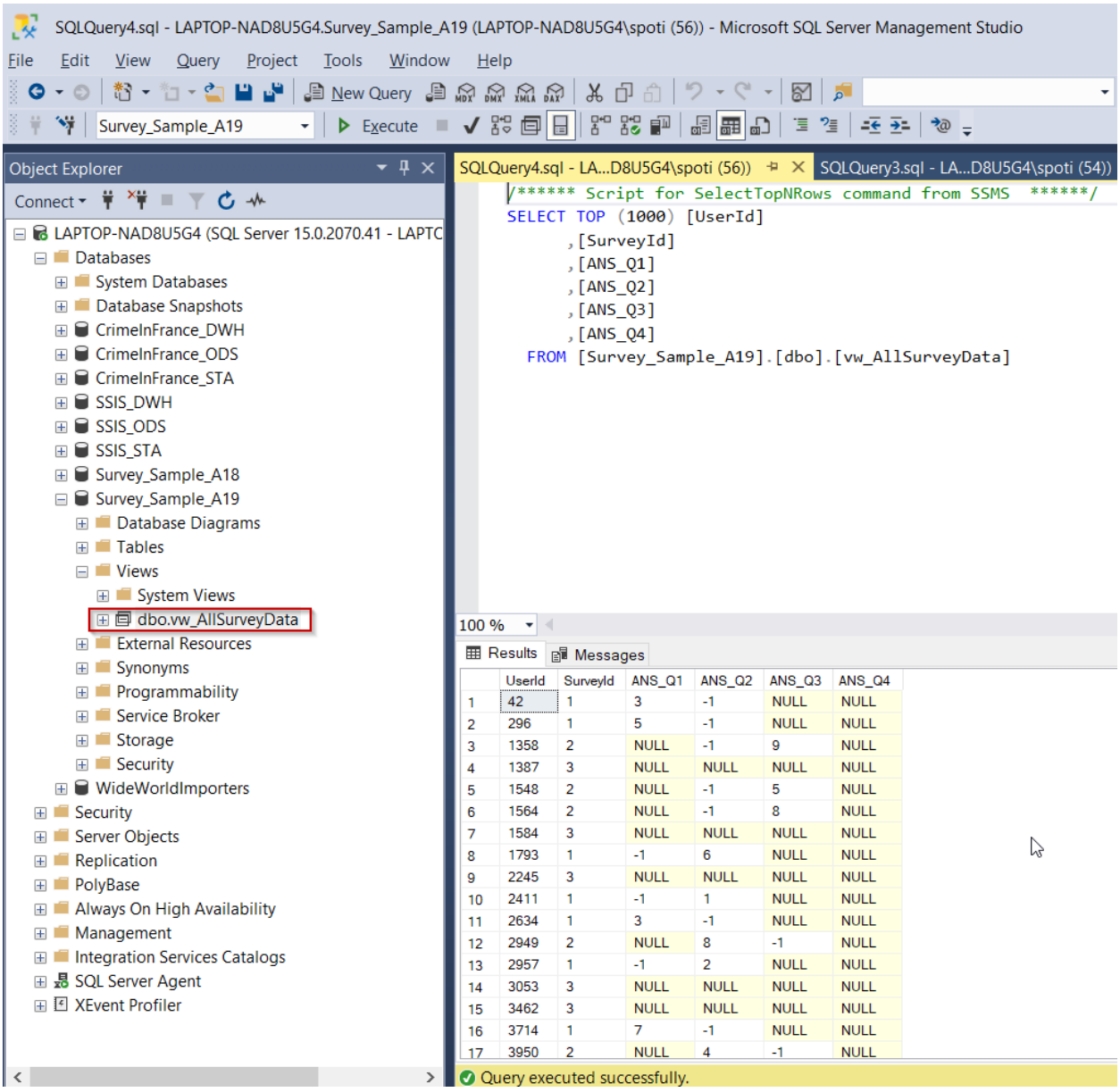
```
>python .\myScript.py running [PATH_TO_FOLDER]\
SQL_PythonProject_ShutimaP\GetAllSurveyData.py as an imported module
myScript myScript.py as _main_ a_function in module
```

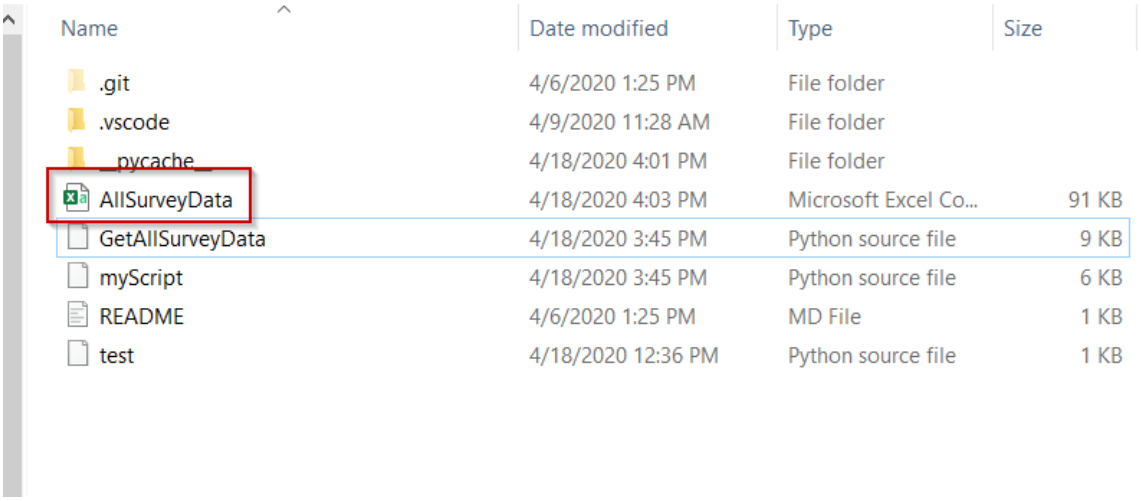
- The script will prompt to enter the SQL view name and export CSV file name as followed:



- If the new Survey Structure is different than the saved one, the trigger will be activated and the SQL view will be created (if not already existed), and the CSV file with all survey data will be generated in the current project directory.

```
Survey IDs List: [1, 2, 3]
Question IDs List: [1, 2, 3, 4]
Survey ID and Question ID In and NOT In Survey Structure table:
    SurveyId  QuestionId  InSurvey
0          1           1         1
1          1           2         1
2          1           3         0
3          1           4         0
4          2           1         0
5          2           2         1
6          2           3         1
7          2           4         0
8          3           1         0
9          3           2         0
10         3           3         0
11         3           4         0
Survey ID and Question ID that are in Survey Structure: [[1, 1], [1, 2], [2, 2], [2, 3]]
New SurveyStructure is different than saved one, need to trigger view
Successfully export CSV file in the current directory
(DataWrangling_SQLPythonProj) PS C:\Users\spoti\Documents\DE_Projects\DSTI-Data-Wrangling-SQL-Project>
```





Name	Date modified	Type	Size
.git	4/6/2020 1:25 PM	File folder	
.vscode	4/9/2020 11:28 AM	File folder	
__pycache__	4/18/2020 4:01 PM	File folder	
AllSurveyData	4/18/2020 4:03 PM	Microsoft Excel Co...	91 KB
GetAllSurveyData	4/18/2020 3:45 PM	Python source file	9 KB
myScript	4/18/2020 3:45 PM	Python source file	6 KB
README	4/6/2020 1:25 PM	MD File	1 KB
test	4/18/2020 12:36 PM	Python source file	1 KB

8. Final Note

Thank you very much for his Python/SQL Final project. Thanks to this project, I have really learnt a lot about Python coding with real practice. This project helps me to find out the area that I need more practice in coding which I will be improving myself for sure. That will be all for this documentation and thank you for reading. Please, feel free to give me any comment on my code 😊