Homework 1: RDB and CSV Datatables

(10 points) Donald Ferguson

Fall 2019 COMS W4111

Due 9/22/19 at 11:59pm

Please post all clarification questions about this homework on the class Piazza under the “hw1” folder. You may post your question privately to the instructors if you wish.

This is an individual assignment. Although you may discuss the questions with other students, you should not discuss the answers with other students. All code and written answers must be entirely your own.

Late policy: You may use late days on this assignment. However, you MUST include as a comment in your submission on courseworks how many late days you are using. If you don’t tell us, or have used all your late days, 10% per late day will be deducted from the homework grade.

1 General Instructions

Please see the HW1 set-up guide here and course website here

The homework only requires programming knowledge that you learn in 1004 or 1007.

The end of lecture 1 and Lecture 2 provides a more detailed explanation and discussion of the HW. The HTML versions can be found on the website.

There is extra credit for this homework assignment that can be found in Lecture 1. Only those who attend or watch Lecture 1 will implement this portion.

Solving the HW requires:

- Writing several "helper" functions each of which does a simple task.
- Putting the smaller pieces together into the larger methods in the class.

The complexity stems from the fact that we do not sufficiently teach you how to write larger programs through decomposition and assembly. All of your past programming assignments have not taught you to develop and test this way.

You are responsible for:

- Implementing and **testing** three classes: RDBDataTable, CSVDataTable, BaseDataTable.

The purpose of this homework is to teach you the behavior of SQL Databases by asking you to implement functions that will model the behavior of a real database with CSVDataTable

2 Implementation

The lectures sufficiently explain the different steps necessary for completion of the homework assignment. To put it concisely, you are first responsible for:

- Open and loading files.
- Connecting to databases (SQL)
- Matching rows to templates.

Secondly, you are required to implement the methods in these files:

- BaseDataTable
- RDBDataTable, CSVDataTable
    - find_by_template(self, template, field_list=None, limit=None, offset=None, order_by=None)
    - delete_by_template(self, template)
    - update_by_template(self, template, new_values)
    - insert(self, new_record)

  *by_template is core to three of the methods

- You must test all methods

If the underlying data is the same, CSVDataTable and RDBDataTable behave **exactly the same.**

3 Logistics

- Submission format is a zip file of the format
    - Data
        - Baseball
            - .csv files
    - src
        - BaseDataTable.py

        - CSVDataTable.py
        - RDBDataTable.py
    - tests
        - unit_test.py
        - csv_table_tests.py (Test code)
        - csv_table_test.txt (Test output)
        - rdb_table_tests.py (Test code)
        - rdb_table_test.txt (Test output)
- ReadMe.txt
    - Explain how your code works and why
    - Use concepts you learned about relational datatables in Lectures 1 and 2.

- We will use Piazza for clarifications and discussions.
- We may be asking for git commit as well - stay tuned. For now, courseworks will suffice.