This homework is used to test your understanding of data collection (crawling) using Python and Python-MySQL connector. It includes 2 tasks:

**Task 1: Collecting data using url**

* Get data from IMBD: <https://www.imdb.com/chart/top>

Data should be a table and include columns: rank, title, year, rating. It looks like a sample table:

|  |  |  |  |
| --- | --- | --- | --- |
| **rank** | **title** | **year** | **rating** |
| 1 | The Shawshank Redemption | 1994 | 9.2 |
| 2 | The Godfather | 1972 | 9.1 |

* You can save the data into a pandas DataFrame.

**Task 2: Store data into MySQL using pymysql package**. Please use Python to finish the first tasks below.

* Execute the create table query "CREATE TABLE imdb (rank INT, title TEXT, year INT, rating DOUBLE)"
* Insert each records using query "INSERT INTO imdb VALUES (rankv, titlev, yearv, ratingv)" by replacing rankv, titlev, yearv, ratingv with actual value for each row. You may need to make a loop and do the insert operation iteratively.
* Go to your MySQL and check the imdb table to see whether all data records are inserted correctly.

**Submission**:

* Please submit your python code - or a clear screenshot of your code for both tasks.
* Please submit a screenshot of your MySQL database showing that you have data in the imdb table.