Rutgers MQF Spring 2020 Exam 2

Return exam solution to Blackboard before the due time.

I. MC (20 points each 2 points)

1. If you define a column as an auto_increment column,

   a. a number is generated for that column whenever a row is added to the table
   b. you must provide a unique numeric value for that column whenever a row is added to the table
   c. you can't use the column as a primary key column
   d. you must also define the column with a default value

2. What uniquely identifies each row in a table?

   a. foreign key
   b. primary key
   c. cell
   d. field

3. To relate one table to another, a/an _____ in one table is used to point to the primary key in another table.

   a. index
   b. composite primary key
   c. primary key
   d. foreign key

4. What can you use to combine data from two or more tables into a single result set?

   a. view
   b. virtual table
   c. result set
   d. join

5. When you code a SELECT statement, you must code the four main clauses in the following order

   a. `SELECT, FROM, ORDER BY, WHERE`
   b. `SELECT, ORDER BY, FROM, WHERE`
   c. `SELECT, WHERE, ORDER BY, FROM`
   d. `SELECT, FROM, WHERE, ORDER BY`

6. In the statement `<a href = "address">`, `href` is a(n) _____ of the `a` element.

   a. anchor
   b. element
   c. label
   d. attribute

7. The `alt` attribute is used in a(n) _____ element to provide an alternate _____ description for the image.

   a. `img`, text.
   b. `a href`, address.
   c. `title`, text.
   d. `src`, image size.

8. Which statement is *false*?

   a. The `&code;` syntax is reserved for special characters.
   b. The `del` element is used to delete text.
   c. The `sup` and `sub` elements are used to denote superscript and subscript characters.
   d. The `hr` element introduces a horizontal rule.

9. An HTML5 table contains rows delimited by _____ elements, which contain cells delimited by _____ elements.

   a. `tr`, `td`
   b. `tr`, `tr`
   c. `table`, `td`
   d. `td`, `tr`

10. Which of the following accurately describes the `select` element?

    a. text input
    b. command button
    c. drop-down list
    d. check box

II. Short answers 30 points (each 5 points)

1. Given the dictionary dic, convert it into pandas DataFrame then output the top 2 numbers from each column
   dic = {'q1': [10, 15, 20, 30, 40, 50], 'q2': [41, 51, 16, 19, 36, 34], 'q3': [90, 45, 67, 132, 10, 41]}
   The answer should look like the following:

```
topmost 2 records within each group of a DataFrame:

For column q1
    q1  q2  q3
5  50  34  41
4  40  36  10

For column q2
    q1  q2  q3
1  15  51  45
0  10  41  90

For column q3
    q1  q2   q3
3  30  19  132
0  10  41   90
```

2. With a dictionary dic = {'q1': [10, 15, 20, 30, 40, 50], 'q2': [41, 51, 16, 19, 36, 34], 'q3': [90, 45, 67, 132, 10, 41]}. Write Python code using pandas to convert it into pandas DataFrame and then select column q1 and q3 from this DataFrame (you may answer this one using the DataFrame you created from the previous question).

	q1	q3
0	10	90
1	15	45
2	20	67
3	30	132
4	40	10
5	50	41

3. With a dictionary dic = {'q1': [10, 15, 20, 30, 40, 50], 'q2': [41, 51, 16, 19, 36, 34], 'q3': [90, 45, 67, 132, 10, 41]}. Write Python code using pandas to convert it into pandas DataFrame and then remove the first two rows from this DataFrame (you may answer this one using the DataFrame you created from the previous question).

```
The original DataFrame is
   q1  q2   q3
0  10  41   90
1  15  51   45
2  20  16   67
3  30  19  132
4  40  36   10
5  50  34   41

After removing first 2 rows the DataFrame is

   q1  q2   q3
2  20  16   67
3  30  19  132
4  40  36   10
5  50  34   41
```

4. Use numpy arange and reshape to create three dimension numpy array like the following

```
array([[[ 0,  1,  2,  3],
        [ 4,  5,  6,  7],
        [ 8,  9, 10, 11]],

       [[12, 13, 14, 15],
        [16, 17, 18, 19],
        [20, 21, 22, 23]]])
```

5. Write a NumPy program to generate 20 random integers between 1 and 13.
6. Given a numpy array x contains of the following. Write python statements to print its median and mean.

```
array([[ 3, 14,  5,  6, 17,  8],
       [ 9, 10, 31, 12, 13,  4],
       [ 1, 16, 16, 18, 19, 20],
       [11, 22, 33, 44, 25,  2]])
```

III. Answer the following Questions (Total 100 points)

1. Giving the file iris.csv use pandas to read in this file. Write Python program using pandas module and other modules if needed. Answer the following questions. Please write your answers of this question using python notebook. **(52 points. Each sub-question 4 points)**

   a. Write python statements using pandas to output the following

	sepal.length	sepal.width	petal.length	petal.width	variety
0	5.1	3.5	1.4	0.2	Setosa
1	4.9	3.0	1.4	0.2	Setosa
2	4.7	3.2	1.3	0.2	Setosa
3	4.6	3.1	1.5	0.2	Setosa
4	5.0	3.6	1.4	0.2	Setosa
5	5.4	3.9	1.7	0.4	Setosa
6	4.6	3.4	1.4	0.3	Setosa

   b. Use pandas statement to output the last three records.

	sepal.length	sepal.width	petal.length	petal.width	variety
147	6.5	3.0	5.2	2.0	Virginica
148	6.2	3.4	5.4	2.3	Virginica
149	5.9	3.0	5.1	1.8	Virginica

   c. Change the column names of iris (replace . with _ and change variety to species. Then output the last 7 records with the newly modified column names.

	sepal_length	sepal_width	petal_length	petal_width	species
143	6.8	3.2	5.9	2.3	Virginica
144	6.7	3.3	5.7	2.5	Virginica
145	6.7	3.0	5.2	2.3	Virginica
146	6.3	2.5	5.0	1.9	Virginica
147	6.5	3.0	5.2	2.0	Virginica
148	6.2	3.4	5.4	2.3	Virginica
149	5.9	3.0	5.1	1.8	Virginica

   d. Print the records from index 50 to 70

   e. Print the number of records which species name is Versicolor
   f. Print the number of records which species name is Virginica
   g. Print the number of records which species name is Setosa

   h. Create four new records as the following

	sepal_length	sepal_width	petal_length	petal_width	species
0	4.9	3.5	1.4	0.3	Setosa
1	6.4	2.9	4.5	1.4	Versicolor
2	5.9	2.3	4.2	1.1	Versicolor
3	4.8	3.4	1.4	0.2	Setosa

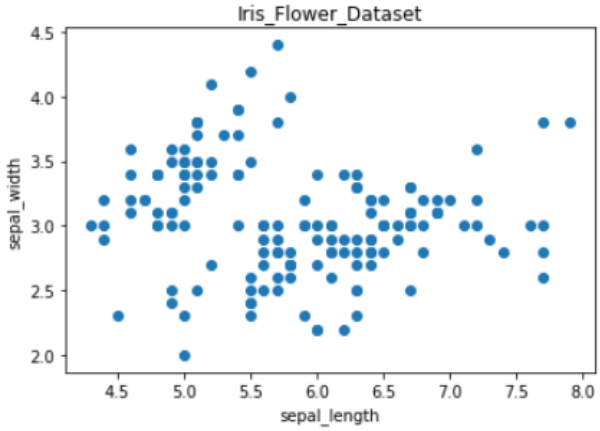
   i. Insert these four record at the bottom of the original iris DatFrame.
   j. Verify by print the new length of the DataFrame. It should show 154

```
154
```

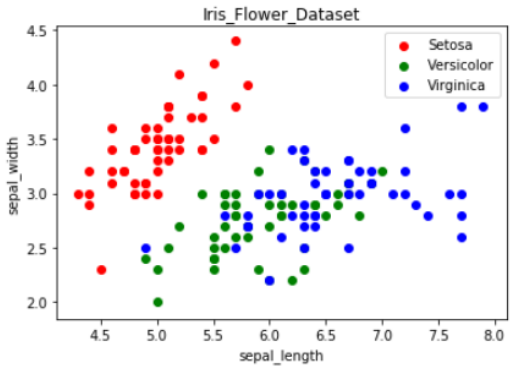
   k. Print the last 7 records again and it should show the following

	sepal_length	sepal_width	petal_length	petal_width	species
147	6.5	3.0	5.2	2.0	Virginica
148	6.2	3.4	5.4	2.3	Virginica
149	5.9	3.0	5.1	1.8	Virginica
150	4.9	3.5	1.4	0.3	Setosa
151	6.4	2.9	4.5	1.4	Versicolor
152	5.9	2.3	4.2	1.1	Versicolor
153	4.8	3.4	1.4	0.2	Setosa

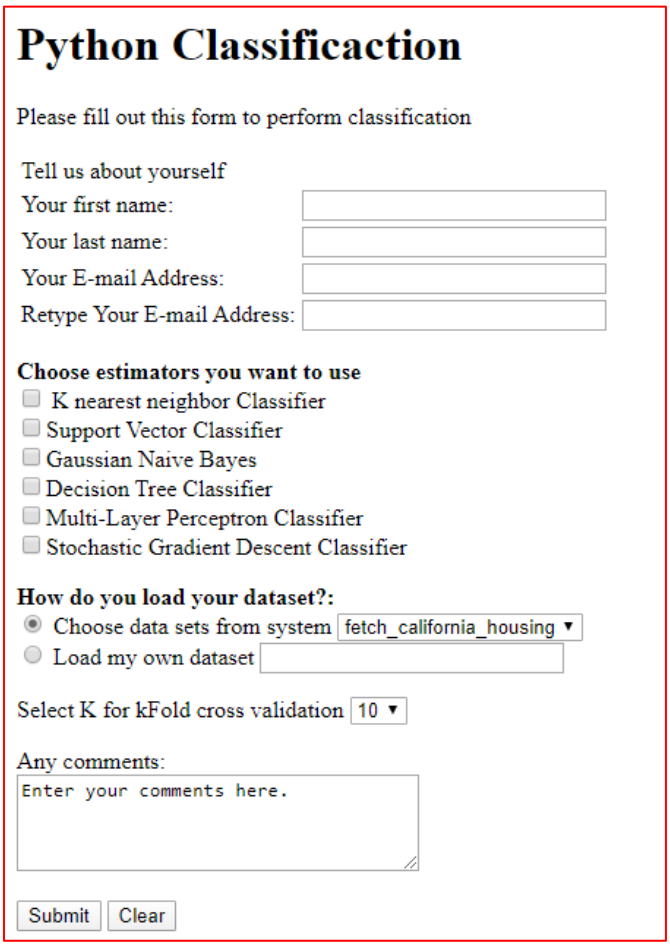
l. Draw the following scatter diagram using all the records with sepal_width as y axis, and sepal_length as x axis. The title should say Iris_Flower_Dataset. It should look like the following



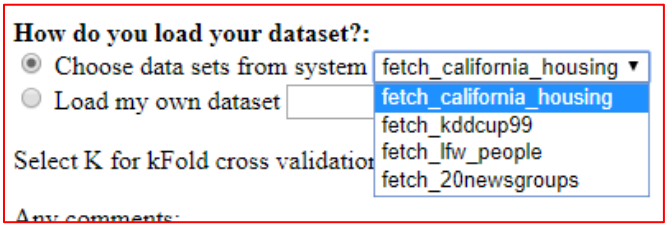
m. Now change the above scatter diagram to the following. Note different flower should have different color. Note legend on the top right corner.



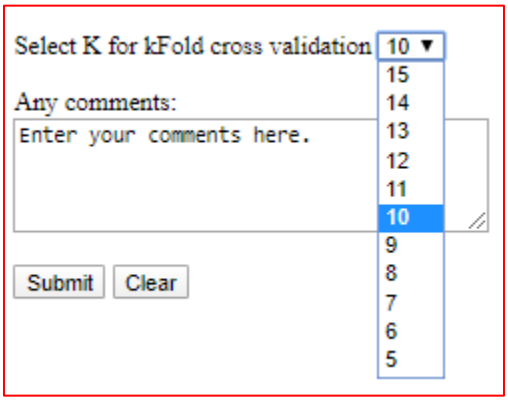
2. Write an HTML document which looks like the following (**24 points**)



There are four items in the drop down list



For kFold drop down list it contains 11 values from 5 to 15 with 10 selected by default



3. Write SQL statements to do the following (**24 points, each sub-question 6 points**)

   a. Create a MySql database and call it ***MQF2020*** which contains a table ***studentGrade*** which contains four fields as the following (all fields cannot be null). Student ID is the primary key and can be automatically generated by the system if omitted when insert records.
   studentID, firstName, lastName, GPA

   b. Insert 5 records as the following using insert statement.

13579	Karen	White	3.8
24680	Susan	Johnson	3.9
12345	Jessy	King	3.4
67890	Helen	Huang	4.0
99100	Peter	Smith	3.9

   c. Update Jessy King's GPA from 3.4 to 3.7

   d. Assuming there are five thousand records in studentGrade table and you are asked to list the top 3 students' names (first name then space then last name). The criteria to rank is that higher GPA student will be ranker higher. If there is a tie then use studentID to resolve. We will list student with small studentID before large studentID.

Submission Instructions:

Step 1: Create a Word file or PDF file and call it ***yourname*-exam2-sol,** where ***yourname*** is yourFirstnameLastname.

Step 2: Write your name on the first line of the first page of the file from Step1

Step 3: Put "MC answers" on a separate line

Step 4: Create a table form like the following. Put your answer under the column 'Answer'

Question	Answer
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Step 5: For Short answer questions, start a new page on your ***yourname*-exam2-sol** file and mark carefully SA1, SA2,.. SA6 and put the answers under each. Make sure you answer them according to the order. I.e., Answer SA1, SA2, SA3, SA4, SA5, SA6 in that order.

Note: Since these short answers are Python programs so create a Python notebook and name it **SA-*yourname*.ipynb** and then put **all** the answers of six questions into this notebook, where ***yourname*** is yourfirstNameLastName

Step 6: Create a folder and call it ***SRC-yourname*** and put the python notebook of Step 5 into this folder.

Step 7: For "Answer the following Questions", start a new page on your ***yourname*-exam2-sol** file. Put the answers of each question in that order. Since Q1 has sub questions from *a* to *m* so mark them carefully as Q1.a, Q1.b,…Q1.m and answer them in that order. (Similar for Q3, mark them as Q3.a, Q3.b, Q3.c and Q3.d and answer them in that order)

Step 8: Create a Python notebook for Q1 and put each sub question into this notebook. Call this notebook as **Q1-*yourname*.ipynb**. Put Q1-*yourname*.ipynb into folder ***SRC-yourname*** created at Step 6.

Step 9: Create an html document and call it **Q2-*yourname*.html** which contains the answer of Q2. Put **Q2-*yourname*.html** into folder ***SRC-yourname*** created at Step 6.

Step 10: Create four (4) **.sql** files and name them Q3-a-*yourname*.sql, Q3-b-*yourname*.sql, Q3-c-*yourname*.sql, Q3-d-*yourname*.sql which contains answer for Q3.a, Q3.b, Q3.c, and Q3.d. Put these four files into folder ***SRC-yourname*** created at Step 6.

Step 11: Zip both ***yourname*-exam2-sol** file and ***SRC-yourname*** folder into ***yourname*.zip** and submit the ***yourname*.zip** to Blackboard before the due time. Note Blackboard will be automatically closed so budget extra time for submission process.