Mock - Final Exam

Department of E-Commerce Engineering

I. Multiple Choice Questions:

1. We have the instruction as following:

```
employees = pd.read_table('./employee.csv', sep='|', index_col='user_id')
```

1.1 ***employees*** is :

a. DataFrame b. csv c. string

1.2 What is each column separated by:

a. comma b. vertical bar c. full stop

2. What will be the output of the following Python code?

```
>>>t=(1,2,4,3)
>>>t[1:-1]
```

a. (1, 2) b. (1, 2, 4) c. (2, 4) d. (2, 4, 3)

3. Tuple is mutable.

a. True b. False

4. Set is immutable.

a. True b. False

II. Comprehension Questions:

1. Explain the differences of discrete data and continuous data with examples.
2. What are the most common operations of data aggregation?
3. What are the examples of the visualization charts used in data visualization? Give 3 examples and explain what each chart is used for?
4. What can be derived from highlighting the data? Why we need to highlight the data?

III. Basic Data Analysis:

1. We have the sample data of the lemonade sold in a month as shown in the following table. Answer each question below:

Date	Day	Temperature	Rainfall	Flyers	Price	Sales
01/01/2017	Sunday	27	2.00	15	0.3	10
02/01/2017	Monday	28.9	1.33	15	0.3	13
03/01/2017	Tuesday	34.5	1.33	27	0.3	15
04/01/2017	Wednesday	44.1	1.05	28	0.3	17
05/01/2017	Thursday	42.4	1.00	33	0.3	18
06/01/2017	Friday	25.3	1.54	23	0.3	11
07/01/2017	Saturday	32.9	1.54	19	0.3	13
08/01/2017	Sunday	37.5	1.18	28	0.3	15
09/01/2017	Monday	38.1	1.18	20	0.3	17
10/01/2017	Tuesday	43.4	1.05	33	0.3	18

1.1 Find the sum of the flyers used in the lemonade sale in the table.

1.2 We create one column “Revenue” to record the total income of each day as following. Fill in the cell with the correct value.

Date	Revenue
01/01/2017	
02/01/2017	
03/01/2017	
04/01/2017	
05/01/2017	
06/01/2017	
07/01/2017	
08/01/2017	
09/01/2017	
10/01/2017	

1.3 Find the total revenue of the lemonade sale in the table.

1.4 Draw the graph illustrating the date and its revenue as following:

IV. Python:

Write the following answer in Python.

1. Create the Tuple (5, 6, 7, 8, 9) and assign it to the variable **A.**
2. Find the first two elements of the tuple **A.**
3. We have the list as following:

B = [“ECom”, 1 ,“loves”, “NIPTICT”, “and”]

   a. Find the first three elements of the list **B.**
   b. Change the second element of the list to **“really”**.
   c. Change the list **B** into a set **C**.
   d. Add **“MPTC”** to the set **C.**
   e. Print all the elements of the set **C.**

4. We have the sample DataFrame taken from the Employee dataset as following:

user_id	age	gender	occupation	zip_code
1	24	M	technician	85711
2	53	F	other	94043
3	23	M	writer	32067
4	24	M	technician	43537
5	33	F	other	15213
6	42	M	executive	98101
7	57	M	administrator	91344
8	36	M	administrator	05201
9	29	M	student	01002
10	53	M	lawyer	90703

By using **Pandas Library** in Python, answer the following questions:

1. Display the first 5 rows of the dataset.
2. Find the number of employees in this dataset.
3. Display the names of all attributes in this dataset.
4. Display only the “age” and “gender” column.
5. Display the five occupations with highest number of employees.