رقم الطالب في كشف الحضور:		نموذج رقم:
عدد صفحات الامتحان:		اسم المساق: لغات برمجة علم البيانات (عملي)
اسم الطالب/ة رباعي:		رقم المساق: DSAI 1103
الرقم الجامعي:		مدة الامتحان: ساعتين
رقم الشعبة:	الامتحانات النهائية للفصل	وقت الامتحان: 12:00 – 10:00
اسم المدرس: فاطمة عبد الكريم عبد العزيز	الدراسي الثاني 2022/2021	تاريخ الامتحان: 2022/6/1

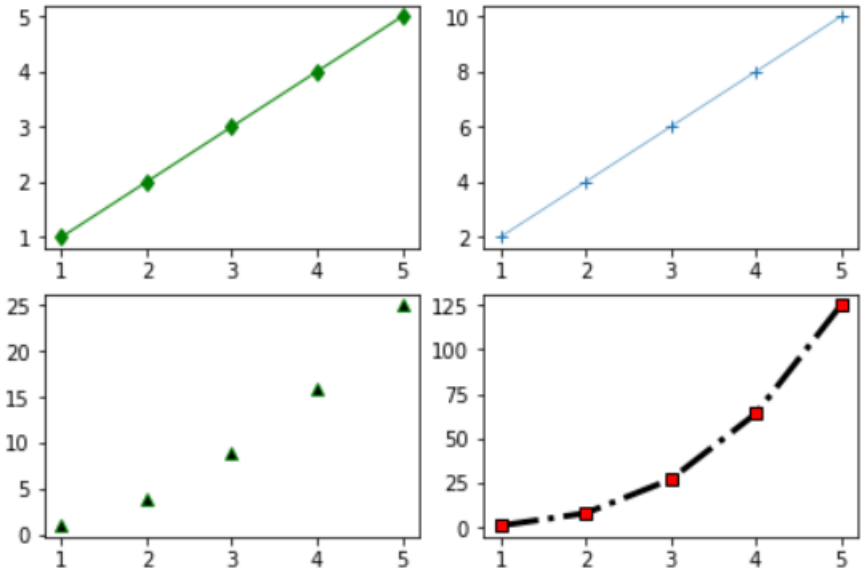
رصد درجات الطالب

السؤال الأول: (5 علامة)

show the following plots:

input:

```
x=np.array([1, 2, 3, 4, 5])
```



السؤال الثاني: (5 علامة)

Convert the following to data frame then sort the data according to the id in descending order and if they are equals sort them according to name then rank the GPA as following output

Input:

```
data = {'names': ['saeed', 'mohammed', 'omar', 'bahaa', 'loay', 'jawad', 'mahmmoud'],
 'ids': ['120215568', '120216874', '120219874', '120214532', '120212258', '120214521',
'120217452'],
 'GPAs': [89, 77, 52, 60, 58, 54, 92]}
```

Output:

```
0    6.0
1    5.0
2    1.0
3    4.0
4    3.0
5    2.0
6    7.0
Name: GPAs, dtype: float64
```

السؤال الثالث: **(5 علامة)**

Write a program to compare the elements of the two Pandas Series

Input:

```
ds1 = pd.Series([3, 7, 9, 11, 15])
ds2 = pd.Series([2, 7, 11, 9, 6])
```

output:

```
Equals:
0    False
1     True
2    False
3    False
4    False
dtype: bool
Greater than:
0     True
1    False
2    False
3     True
4     True
dtype: bool
Less than:
0    False
1    False
2     True
3    False
4    False
dtype: bool
```

السؤال الرابع: **(5 علامة)**

create data frame as shown then change anyone GPA between 70 and 80 to C

input:

```
data = {'names': ['ahmed', 'mohammed', 'anas', 'foad', 'loay', 'gamal', 'saed'],
 'ids': ['120215568', '120216874', '120219874', '120214532', '120212258',
'120214521', '120217452'],
 'GPAs': [89, 77, 52, 60, 75, 54, 92]}
```

output:

	names	ids	GPAs
0	ahmed	120215568	89
1	mohammed	120216874	C
2	anas	120219874	52
3	foad	120214532	60
4	loay	120212258	C
5	gamal	120214521	54
6	saed	120217452	92

السؤال الخامس: **(5 علامة)**

concatenate the following arrays horizontally (by the columns) then split it as shown

input:

```
arr1 = np.array([[1, 2, 3], [4, 5, 6]])
arr2 = np.array([[7, 8, 9], [10, 11, 12]])
```

output:

```
first: [[1 2 3 7 8 9]]
second: [[ 4  5  6 10 11 12]]
```

السؤال السادس: **(5 علامة)**

for the following array, repeat the first element twice, the second element three times, and the third element four times (10, 3) then use tile method to get this shape (6, 6) of the array

input:

```
arr = np.random.randn(3,3)
```

السؤال السابع: **(5 علامة)**

for the given Array print the cumulative product of the rows then the cumulative sum of the columns then finds the minimum and the maximum value of the cumulative sum array then print the number of the values that is greater than 50 in the cumulative sum array

input:

```
values = np.arange(20).reshape((5,4))
```

السؤال الثامن: (5 علامة)

Get the highlighted elements

[[0 1 2 3]
[4 5 6 7]
[8 9 10 11]
[12 13 14 15]]

input:
arr = np.arange(16).reshape((4, 4))

السؤال التاسع: (5 علامة)

For the following dataframe prepare the data for the data science task (fill the null data and remove the noisy data)
Input:
data = {'Department': ['Computer Science', 'Software Engineering', 'Computer Science', 'Software Engineering', 'Software Engineering', 'Computer Science', 'Software Engineering'],
'Age': [33, 43, 5, 35, 97, 26, 27],
'Salary': [2350, 2800, np.nan, 2700, np.nan, 3000, np.nan]}
frame = pd.DataFrame(data)

نياتي لكـم بالتوفيق والنجاح. .