



Data visualization

(Zaouche Djaouida)

2019–2020

1. The most fundamental plotting function of the matplotlib.pyplot package is plot. In principle, it needs two sets of numbers.

```
import numpy as np

import matplotlib.pyplot as plt

y=np.random.rand(20)

print(y)

x=np.arange(len(y))

plt.plot(x,y)

plt.show()
```

2. To better readability, we can add other information to your figures. Comment the following code.

```
import numpy as np
y=np.array([i for i in np.arange(100)])
import matplotlib.pyplot as plt
plt.figure(figsize=(3,3))
print(y.cumsum())
plt.plot(y.cumsum(),'r',lw=4)
plt.plot(y.cumsum(),'b')
plt.xlabel('index')
plt.ylabel('value')
plt.legend(['1st','2nd'])
plt.title('A simple Plot')
```

3. By using subplots function of the matplotlib.pyplot package, we can create a figure and a set of subplots. It corresponds to a figure that can group several sub-figures, modeled by an object Figure.

- a. Comment the following code.

```
import numpy as np
import matplotlib.pyplot as plt
```

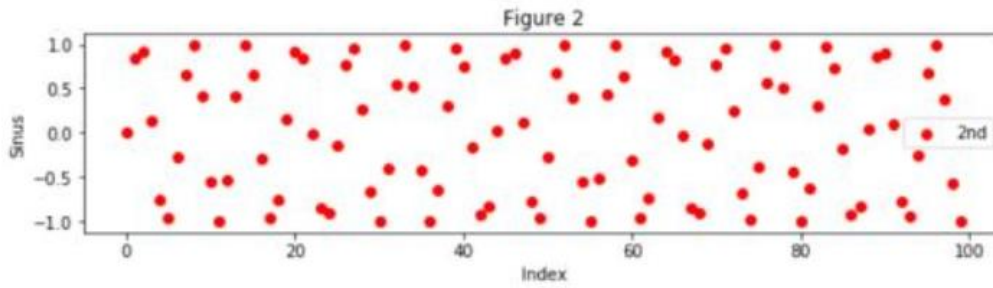
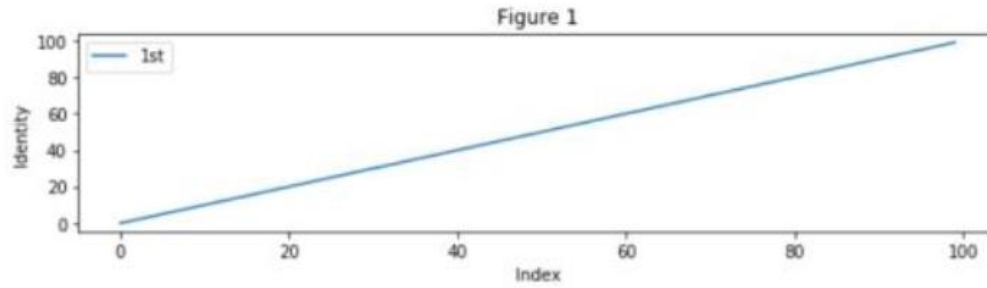
```

y=np.array([i for i in np.arange(100)])
x=np.arange(len(y))
y2=y*y
y3=np.sin(y)
y4=np.cos(y)
plt.figure(figsize=(10,10))
plt.subplot(221)
plt.plot(x,y)
plt.xlabel('Index')
plt.ylabel('Identity')
plt.title('Figure 1')
plt.legend(['1st'])
plt.subplot(222)
plt.plot(x,y2)
plt.xlabel('Index')
plt.ylabel('Square')
plt.title('Figure 2')
plt.legend(['2nd'])
plt.subplot(223)
plt.xlabel('Index')
plt.ylabel('Sinus')
plt.title('Figure 3')
plt.legend(['3rd'])
plt.plot(x,y3)
plt.subplot(224)
plt.ylabel('Cosinus')
plt.legend(['4th'])
plt.xlabel('Index')
plt.plot(x,y4)
plt.show()

```

For better readability, use the `subplots_adjust` function to add space between figures.

b. Consider the figure shown below. Give the code to display it, use the `subplot` function to put figure 1 and figure 2 inside a same figure. Use the `subplots_adjust` function to add space between the first figure and the second one.



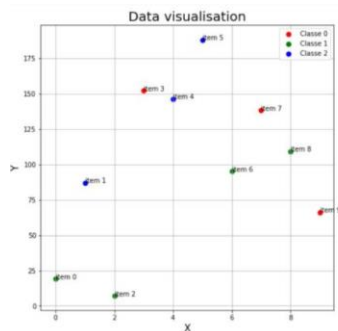
4. Build a dataframe as shown below :

```
[78]:
```

	X	Y	C
item 0	0	19	1
item 1	1	87	2
item 2	2	7	1
item 3	3	152	0
item 4	4	146	2
item 5	5	188	2
item 6	6	95	1
item 7	7	138	0
item 8	8	109	1
item 9	9	66	0

It contains 10 rows and 3 columns X, Y and C. Each row *i* is indexed by the string *item i*. Column X contains values 0, ..., 9. Column Y contains random values between 1 and 200, columns C contains random values between 0 and 2.

5. Write a script to draw the following plot :



Each dot is colored according to its attribute C and labeled by its index row.

