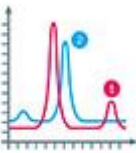


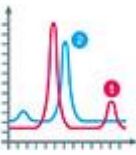
NHẬP MÔN LẬP TRÌNH KHOA HỌC DỮ LIỆU

Bài 8: Thư viện matplotlib

Nội dung

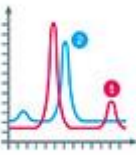


1. Giới thiệu và cài đặt matplotlib
2. Vẽ biểu đồ đơn giản
3. Một số loại biểu đồ thông dụng trong matplotlib
 1. Biểu đồ dạng đường (line plot)
 2. Biểu đồ dạng cột (bar plot)
 3. Biểu đồ bánh (pie chart)
 4. Một số chức năng hữu ích
4. Bài tập



Phần 1

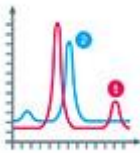
Giới thiệu và cài đặt matplotlib



Giới thiệu matplotlib

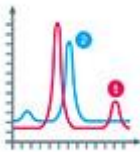
- “matplotlib” là thư viện chuyên về vẽ biểu đồ, mở rộng từ numpy
- Có mục tiêu đơn giản hóa tối đa công việc vẽ biểu đồ để “chỉ cần vài dòng lệnh”
- Hỗ trợ rất nhiều loại biểu đồ, đặc biệt là các loại được sử dụng trong nghiên cứu hoặc kinh tế như biểu đồ dòng, đường, tần suất (histograms), phổ, tương quan, errorcharts, scatterplots,...
- Cấu trúc của matplotlib gồm nhiều phần, phục vụ cho các mục đích sử dụng khác nhau

Giới thiệu matplotlib



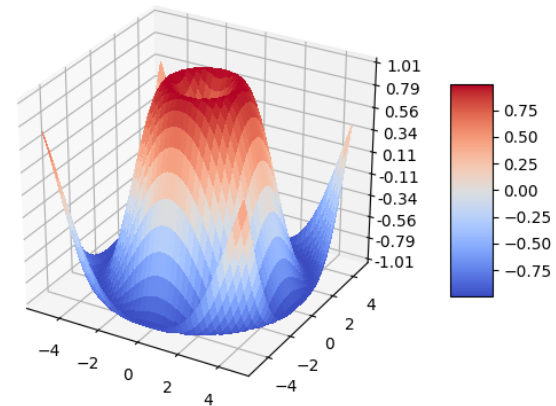
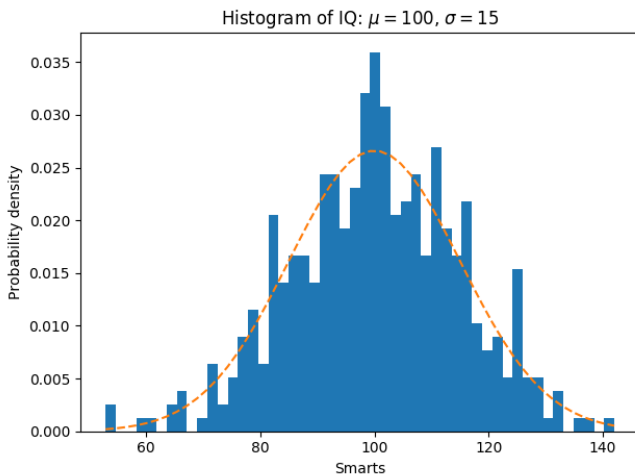
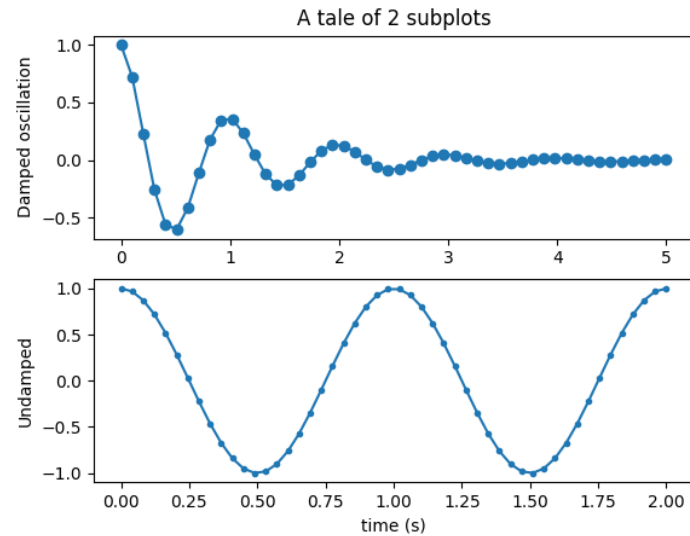
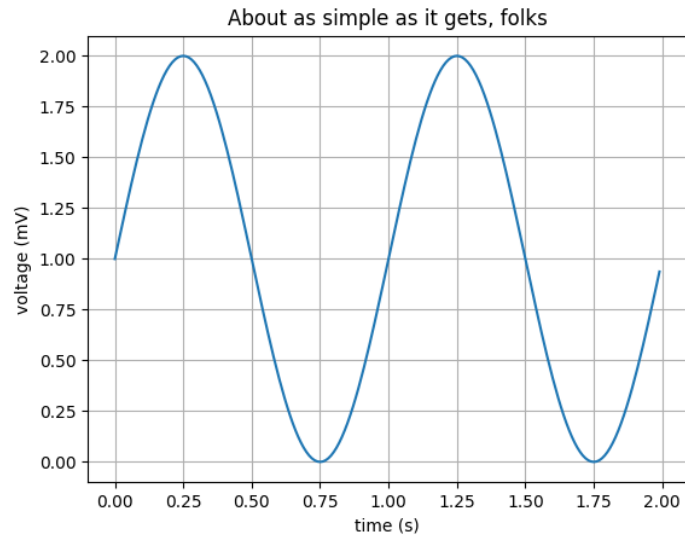
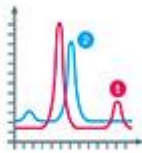
- Ngoài các API liên quan đến vẽ biểu đồ, matplotlib còn bao gồm một số interface: Object-Oriented API, The Scripting Interface (pyplot), The MATLAB Interface (pylab)
 - Các interface này giúp chúng ta thuận tiện trong việc thiết lập chỉ số trước khi thực hiện vẽ biểu đồ
 - Interface pylab hiện đã không còn được phát triển
 - Hầu hết các ví dụ trong slide này đều sử dụng pyplot
 - Sử dụng Object-Oriented API hoặc trực tiếp các API của matplotlib sẽ cho phép can thiệp sâu hơn vào việc vẽ biểu đồ (hầu hết project sẽ không có nhu cầu này)

Cài đặt: “pip install matplotlib”

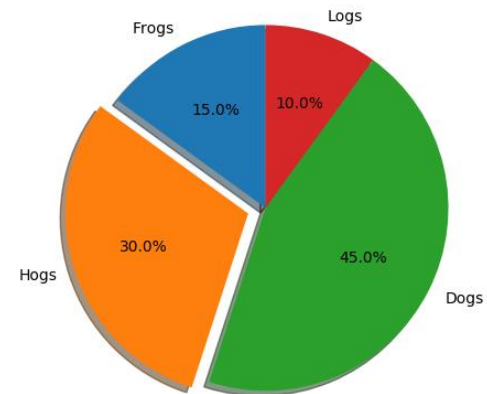
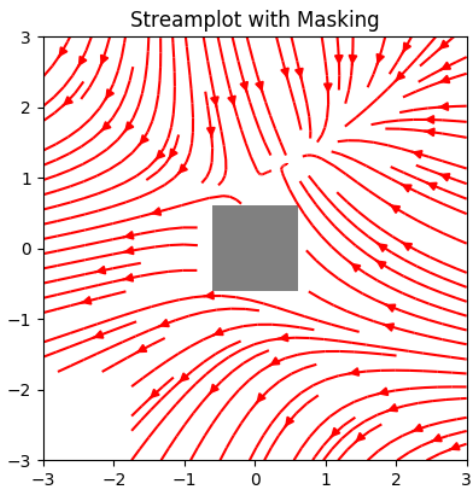
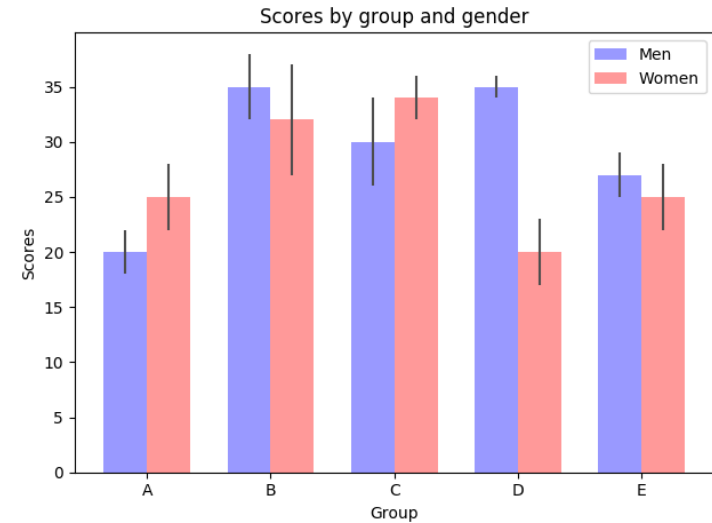
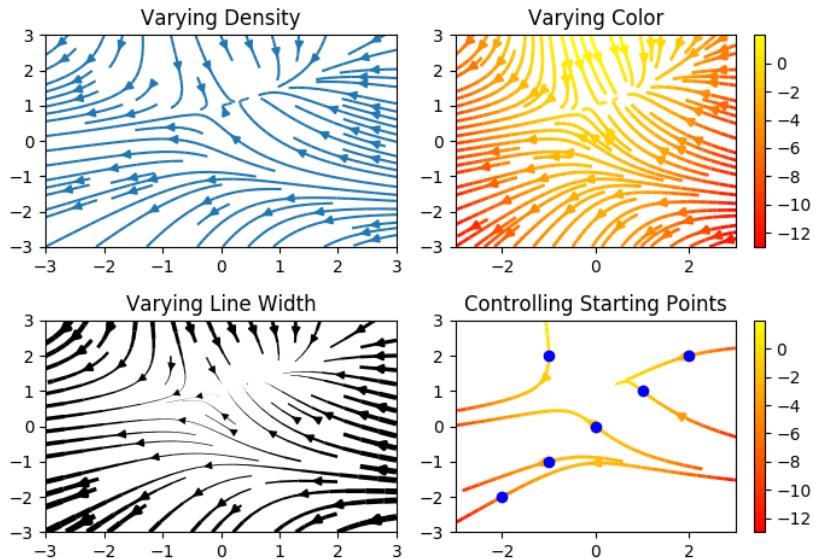
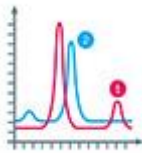


```
C:\Users\XuanNam\Downloads>pip install matplotlib
Collecting matplotlib
  Downloading https://files.pythonhosted.org/packages/bf/b9/485032835e979ee11d514bb3b9b0543a928b8b96c099c178aeabd2ba861/matplotlib-2.2.2-cp36-cp36m-win_amd64.whl (8.7MB)
    100% |████████████████████████████████████████| 8.7MB 3.3MB/s
Requirement already satisfied: pyparsing!=2.0.4,!=2.1.2,!=2.1.6,>=2.0.1 in c:\dev\python36\lib\site-packages (from matplotlib) (2.2.0)
Collecting cycler>=0.10 (from matplotlib)
  Downloading https://files.pythonhosted.org/packages/f7/d2/e07d3ebb2bd7af696440ce7e754c59dd546ffef1bbe732c8ab68b9c834e61/cyclor-0.10.0-py2.py3-none-any.whl
Collecting kiwisolver>=1.0.1 (from matplotlib)
  Downloading https://files.pythonhosted.org/packages/44/72/16630c3392eba03788ad87949390516bbc488e8e118047a3b824631d21a6/kiwisolver-1.0.1-cp36-none-win_amd64.whl (57kB)
    100% |████████████████████████████████████████| 61kB 4.0MB/s
Requirement already satisfied: python-dateutil>=2.1 in c:\dev\python36\lib\site-packages (from matplotlib) (2.7.2)
Requirement already satisfied: six>=1.10 in c:\dev\python36\lib\site-packages (from matplotlib) (1.11.0)
Requirement already satisfied: numpy>=1.7.1 in c:\dev\python36\lib\site-packages (from matplotlib) (1.14.3)
Requirement already satisfied: pytz in c:\dev\python36\lib\site-packages (from matplotlib) (2018.4)
Requirement already satisfied: setuptools in c:\dev\python36\lib\site-packages (from kiwisolver>=1.0.1->matplotlib) (28.8.0)
Installing collected packages: cycler, kiwisolver, matplotlib
Successfully installed cycler-0.10.0 kiwisolver-1.0.1 matplotlib-2.2.2
```

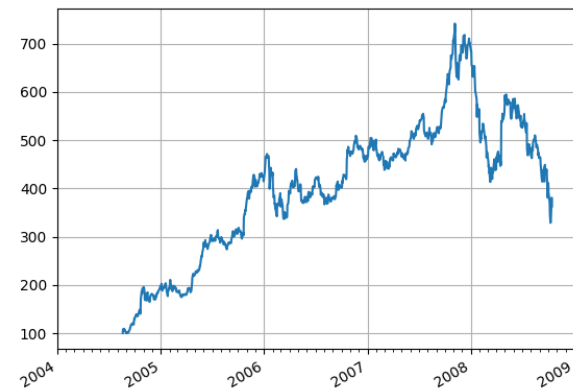
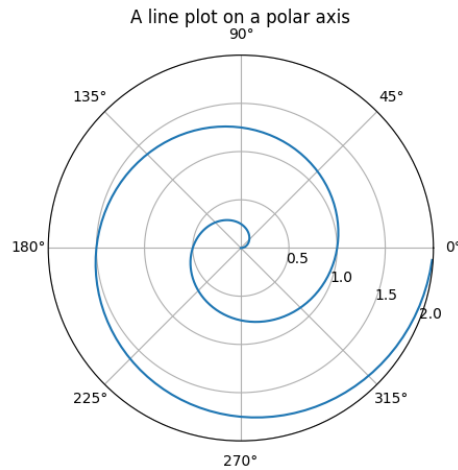
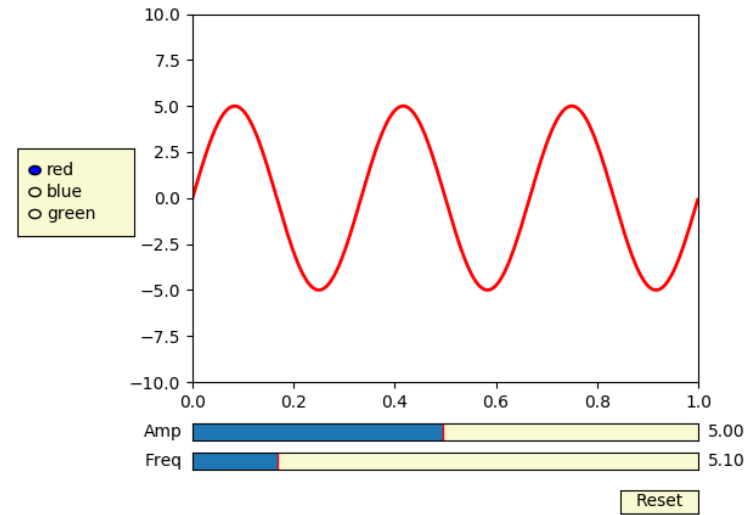
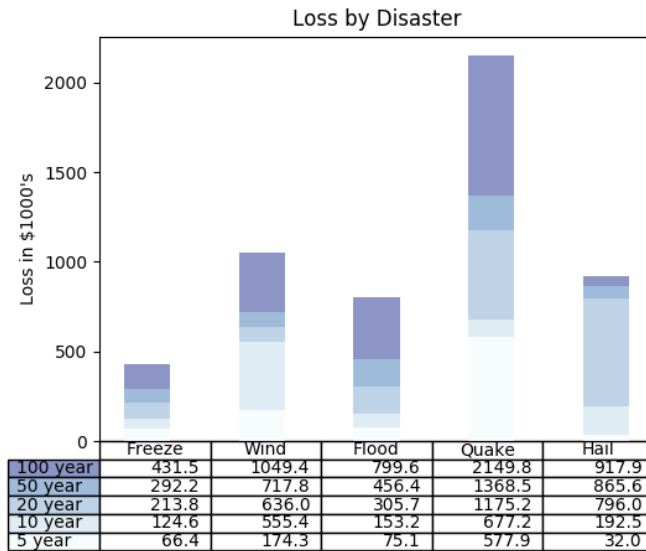
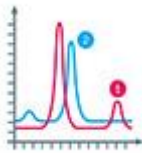
Một số biểu đồ vẽ bằng matplotlib

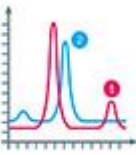


Một số biểu đồ vẽ bằng matplotlib



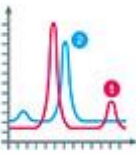
Một số biểu đồ vẽ bằng matplotlib





Phần 2

Vẽ biểu đồ đơn giản

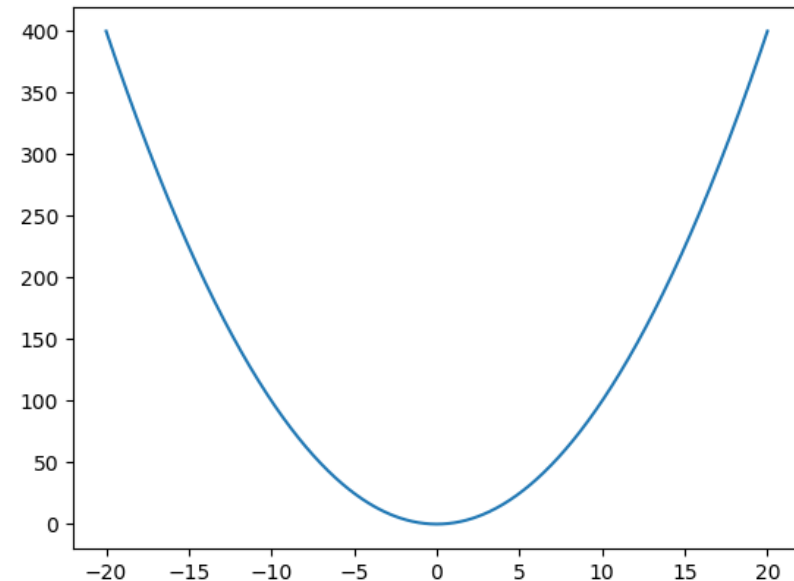


Ví dụ: vẽ biểu đồ $y = x^2$

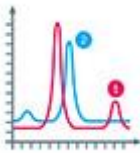
```
import numpy as np          # thư viện numpy
import matplotlib.pyplot as plt  # thư viện pyplot

# chia đoạn từ -20 đến 20 thành 1000 đoạn
x = np.linspace(-20, 20, 1000)
# tính y
y = x * x

# vẽ biểu đồ tương quan giữa x và y
plt.plot(x, y)
# hiển thị biểu đồ
plt.show()
```



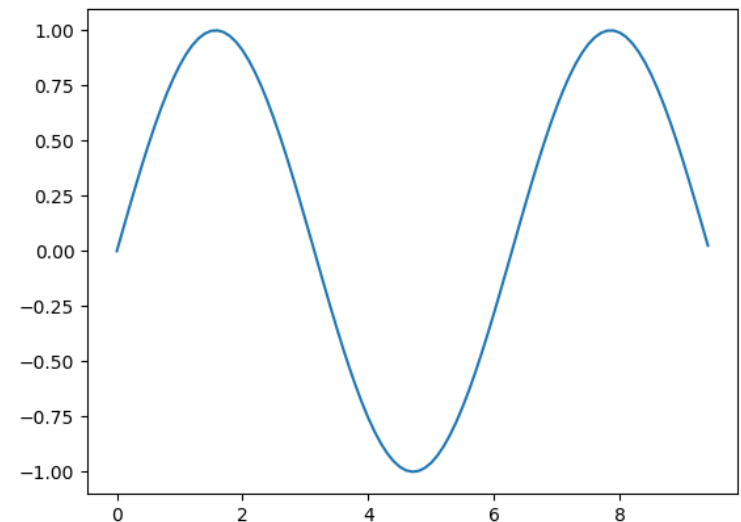
Ví dụ: vẽ biểu đồ hình sin



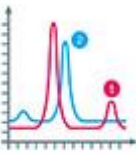
```
import numpy as np                    # thư viện numpy
import matplotlib.pyplot as plt       # thư viện pyplot

# chia đoạn từ 0 đến 3π thành các đoạn con 0.1
x = np.arange(0, 3 * np.pi, 0.1)
# tính sin tương ứng với từng phần tử của x
y = np.sin(x)

# vẽ biểu đồ tương quan giữa x và y
plt.plot(x, y)
# hiển thị biểu đồ
plt.show()
```



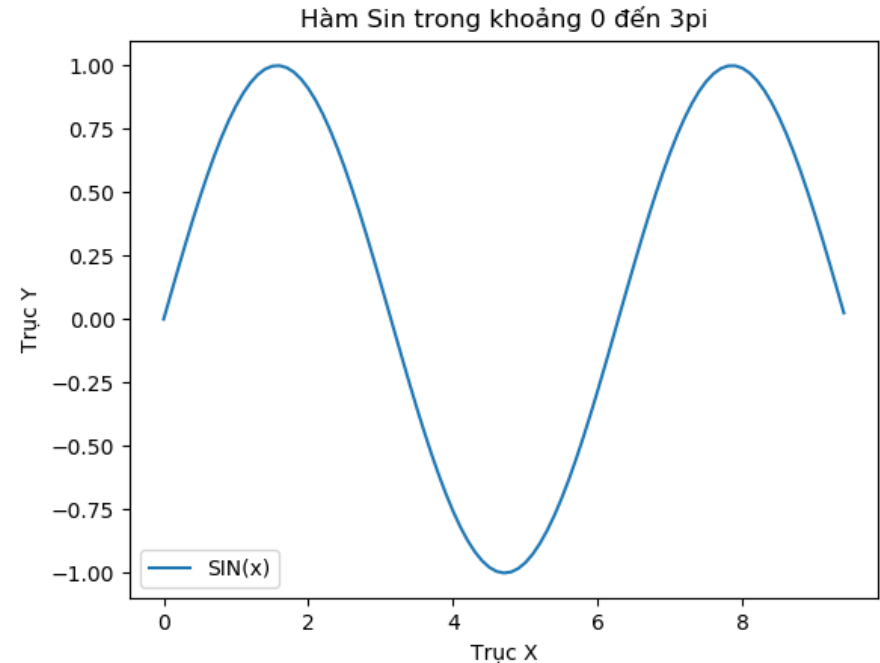
Ví dụ: biểu đồ hình sin đầy đủ hơn



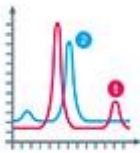
```
import numpy as np
import matplotlib.pyplot as plt

x = np.arange(0, 3 * np.pi, 0.1)
y = np.sin(x)

plt.plot(x, y)
# các thông tin bổ sung cho biểu đồ
plt.xlabel('Trục X')
plt.ylabel('Trục Y')
plt.title('Hàm Sin trong khoảng 0 đến 3pi')
plt.legend(['SIN(x)'])
plt.show()
```



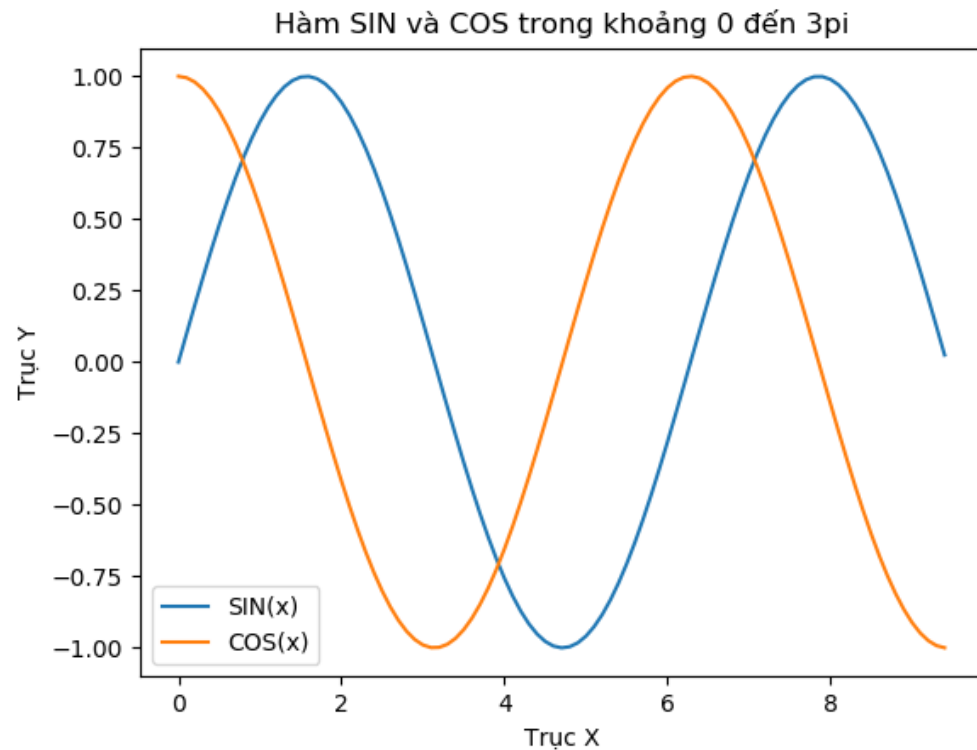
Ví dụ: biểu đồ cả SIN và COS



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

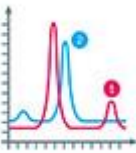
x = np.arange(0, 3 * np.pi, 0.1)
y_sin = np.sin(x)
y_cos = np.cos(x)

plt.plot(x, y_sin)
plt.plot(x, y_cos)
plt.xlabel('Trục X')
plt.ylabel('Trục Y')
plt.title('Hàm SIN và COS trong khoảng 0 đến 3pi')
plt.legend(['SIN(x)', 'COS(x)'])
plt.show()
```



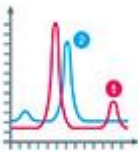
Các bước vẽ biểu đồ với matplotlib

- Điều kiện cần: đã có sẵn dữ liệu
- Có thể có 4 bước cơ bản:
 1. **Chọn loại biểu đồ phù hợp**
 - Tùy thuộc rất nhiều vào loại dữ liệu
 - Tùy thuộc vào mục đích sử dụng của người dùng
 2. **Thiết lập các thông số cho biểu đồ**
 - Thông số của các trục, ý nghĩa, tỉ lệ chia,...
 - Các điểm nhấn trên bản đồ
 - Góc nhìn, mẫu tô, màu và các chi tiết khác
 - Các thông tin bổ sung
 3. **Vẽ biểu đồ**
 4. **Lưu ra file**



Phần 3

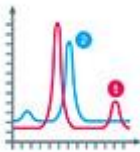
Một số loại biểu đồ thông dụng trong matplotlib



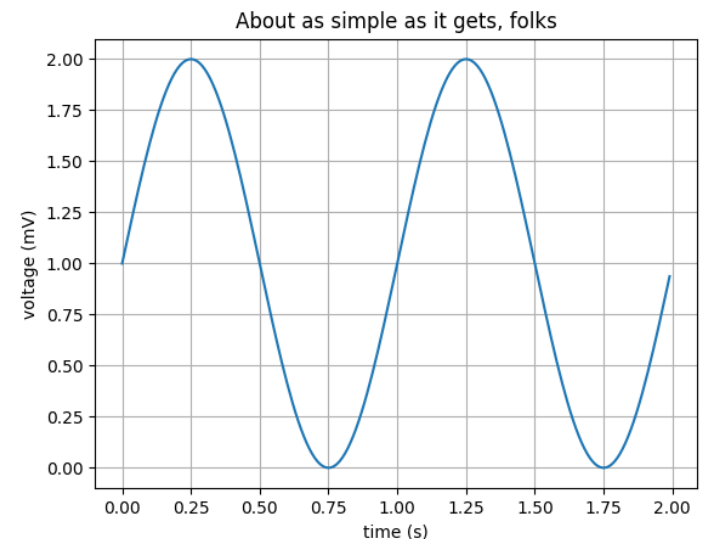
Phần 3.1

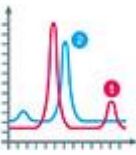
Biểu đồ dạng đường (line plot)

Line plot



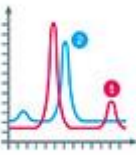
- Biểu đồ thể hiện tương quan giữa X và Y
- Cú pháp:
 - `plot([x], y, [fmt], data=None, **kwargs)`
 - `plot([x], y, [fmt], [x2], y2, [fmt2], ..., **kwargs)`
- “fmt” là quy cách vẽ đường
- “data” là nhãn của dữ liệu
- `**kwargs`: tham số vẽ đường
- Vẽ nhiều lần trên một biểu đồ
- Kết quả trả về là một list các đối tượng Line2D





Line plot: fmt

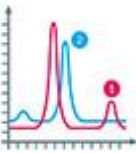
- “fmt” gồm 3 phần `fmt = '[color][marker][line]'`
- `[color]` – viết tắt tên màu:
 - ‘b’ – blue
 - ‘g’ – green
 - ‘r’ – red
 - ‘c’ – cyan
 - ‘m’ – magenta
 - ‘y’ – yellow
 - ‘b’ – black
 - ‘w’ – white
 - `#rrggbb` – chỉ ra mã màu theo hệ RGB



Line plot: fmt

- [marker] – cách đánh dấu dữ liệu:
 - ‘o’ – hình tròn
 - ‘v’ – tam giác xuống (‘^’, ‘<’, ‘>’)
 - ‘*’ – ngôi sao
 - ‘.’ – chấm
 - ‘p’ – ngũ giác
 - ...
- [line] – cách vẽ đường:
 - ‘-’ – nét liền
 - ‘--’ – nét đứt
 - ‘-.’ – gạch chấm
 - ‘:’ – đường chấm

Ví dụ



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

# chia đoạn 0-5 thành các bước 0.2
t = np.arange(0., 5., 0.2)
```

```
# Vẽ 3 đường:
```

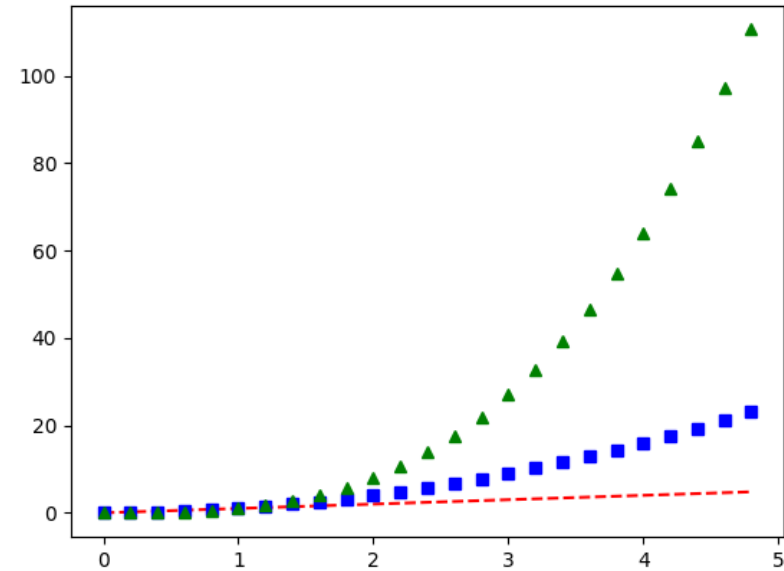
```
# - màu đỏ nét đứt:  $y = x$ 
```

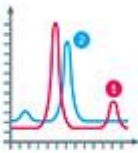
```
# - màu xanh dương, đánh dấu ô vuông:  $y = x^2$ 
```

```
# - màu xanh lá, đánh dấu tam giác:  $y = x^3$ 
```

```
plt.plot(t, t, 'r--', t, t**2, 'bs', t, t**3, 'g^')
```

```
plt.show()
```

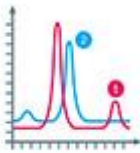




Phần 3.2

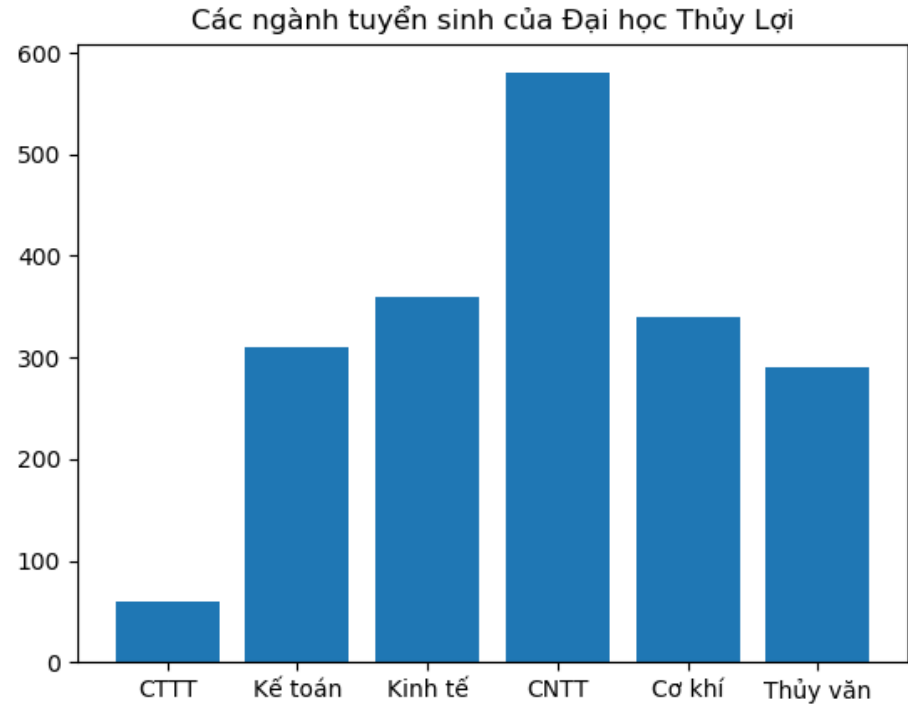
Biểu đồ dạng cột (bar plot)

Biểu đồ dạng cột chiều đứng



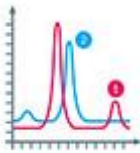
```
import matplotlib.pyplot as plt
```

```
D = { 'CTTT': 60,  
      'Kế toán': 310,  
      'Kinh tế': 360,  
      'CNTT': 580,  
      'Cơ khí': 340,  
      'Thủy văn': 290 }
```



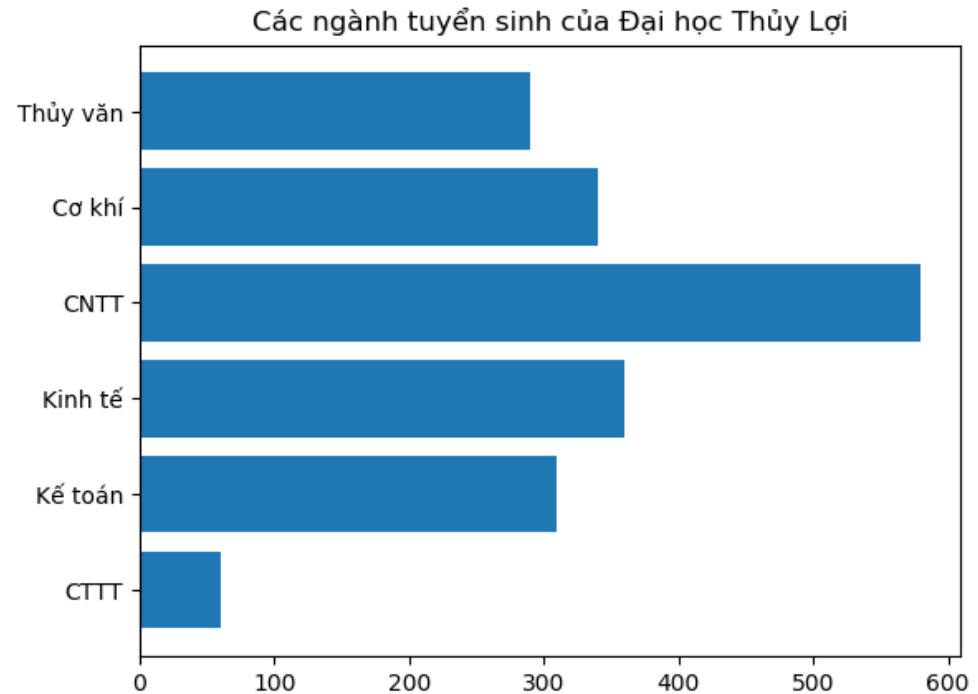
```
plt.bar(range(len(D)), D.values(), align='center')  
plt.xticks(range(len(D)), D.keys())  
plt.title('Các ngành tuyển sinh của Đại học Thủy Lợi')  
plt.show()
```

Biểu đồ dạng cột chiều ngang



```
import matplotlib.pyplot as plt
```

```
D = { 'CTTT': 60,  
      'Kế toán': 310,  
      'Kinh tế': 360,  
      'CNTT': 580,  
      'Cơ khí': 340,  
      'Thủy văn': 290 }
```



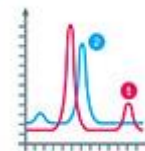
```
plt.barh(range(len(D)), list(D.values()))
```

```
plt.yticks(range(len(D)), D.keys())
```

```
plt.title('Các ngành tuyển sinh của Đại học Thủy Lợi')
```

```
plt.show()
```


Có thể ghép 2 biểu đồ



```
import matplotlib.pyplot as plt
```

```
plt.bar([1,3,5,7,9],[5,2,7,8,2], label="One")
```

```
plt.bar([2,4,6,8,10],[8,6,2,5,6], label="Two", color='g')
```

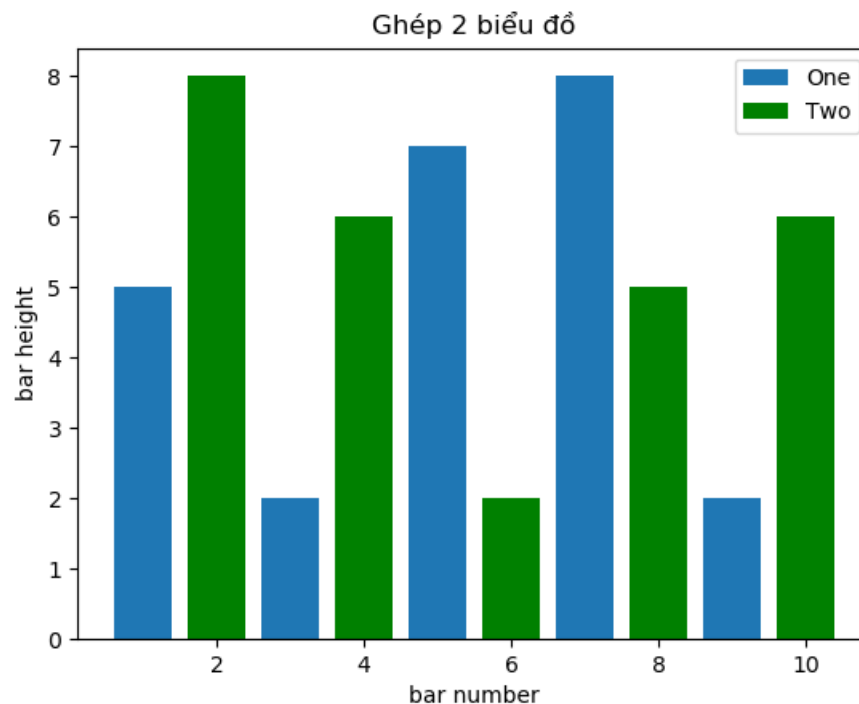
```
plt.legend()
```

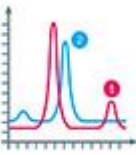
```
plt.xlabel('bar number')
```

```
plt.ylabel('bar height')
```

```
plt.title('Ghép 2 biểu đồ')
```

```
plt.show()
```





Phần 3.3

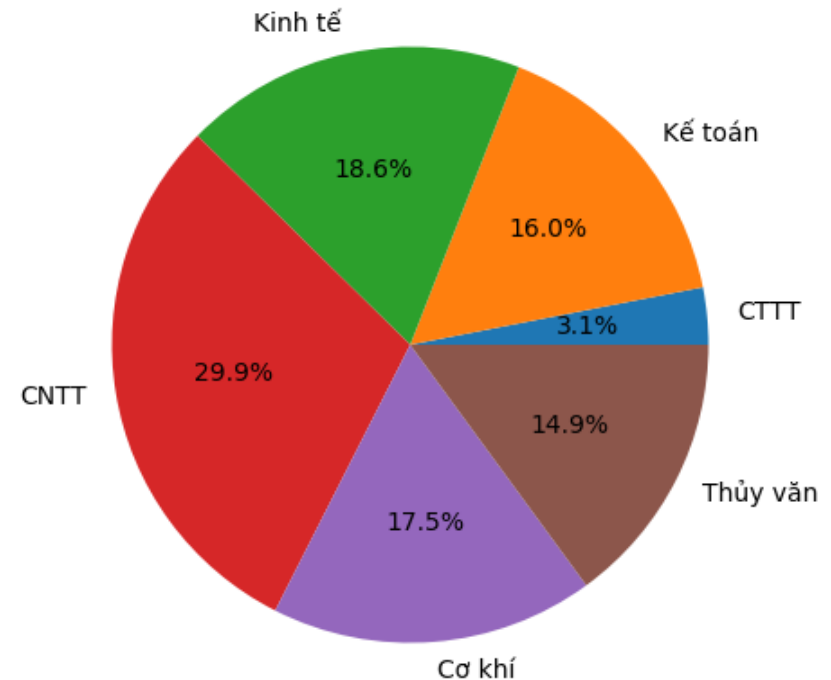
Biểu đồ bánh (pie chart)

Làm rõ tương quan chi tiết / tổng thể



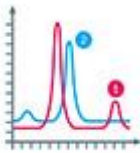
```
import matplotlib.pyplot as plt
```

```
D = { 'CTTT': 60,  
      'Kế toán': 310,  
      'Kinh tế': 360,  
      'CNTT': 580,  
      'Cơ khí': 340,  
      'Thủy văn': 290 }
```



```
plt.pie(D.values(), labels=D.keys(), autopct='%1.1f%%')  
plt.axis('equal') # trục x = trục y
```

```
plt.show()
```



Biểu đồ bánh đầy đủ hơn

```
import matplotlib.pyplot as plt
```

```
# dữ liệu
```

```
labels = 'Python'
```

```
sizes = [215,
```

```
colors = ['gold'
```

```
explode = (0.1
```

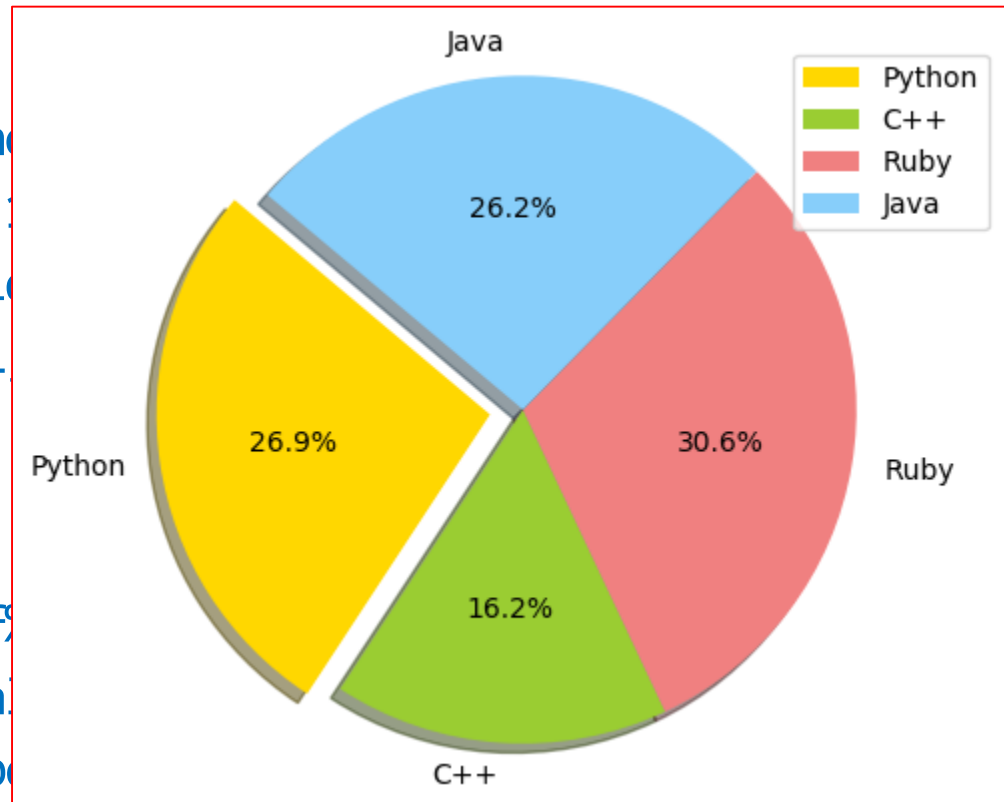
```
# vẽ
```

```
plt.pie(sizes,  
autopct='%1.1f'
```

```
plt.axis('equal'
```

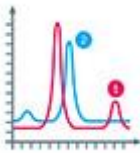
```
plt.legend(labels'
```

```
plt.show()
```



```
kyblue']
```

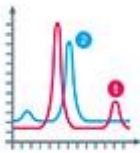
```
colors,
```



Phần 3.4

Một số chức năng hữu ích

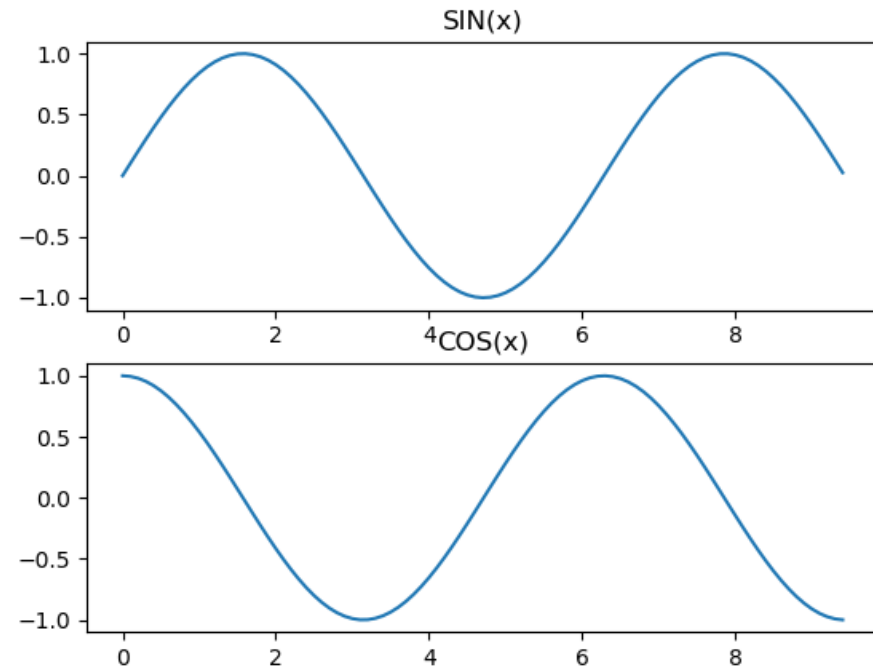
Chia thành các biểu đồ con



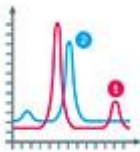
```
import numpy as np
import matplotlib.pyplot as plt

x = np.arange(0, 3 * np.pi, 0.1)
y_sin = np.sin(x)
y_cos = np.cos(x)
```

```
plt.subplot(2, 1, 1)      # biểu đồ 1
plt.plot(x, y_sin)
plt.title('SIN(x)')
plt.subplot(2, 1, 2)    # biểu đồ 2
plt.plot(x, y_cos)
plt.title('COS(x)')
plt.show()
```



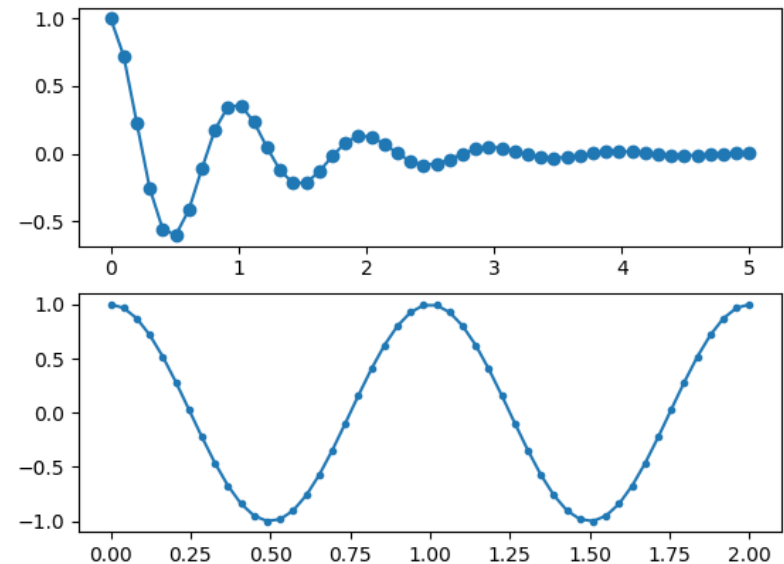
Chia thành các biểu đồ con

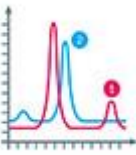


```
import numpy as np
import matplotlib.pyplot as plt

x1 = np.linspace(0.0, 5.0)
x2 = np.linspace(0.0, 2.0)
y1 = np.cos(2 * np.pi * x1) * np.exp(-x1)
y2 = np.cos(2 * np.pi * x2)

plt.subplot(2, 1, 1)
plt.plot(x1, y1, 'o-')
plt.subplot(2, 1, 2)
plt.plot(x2, y2, '-.')
plt.show()
```





Chia thành các biểu đồ con

```
import matplotlib.pyplot as plt
```

```
import numpy as np
```

```
np.random.seed(1234)
```

```
data = np.random.randn(1000)
```

```
fig, axs = plt.subplots(2, 2)
```

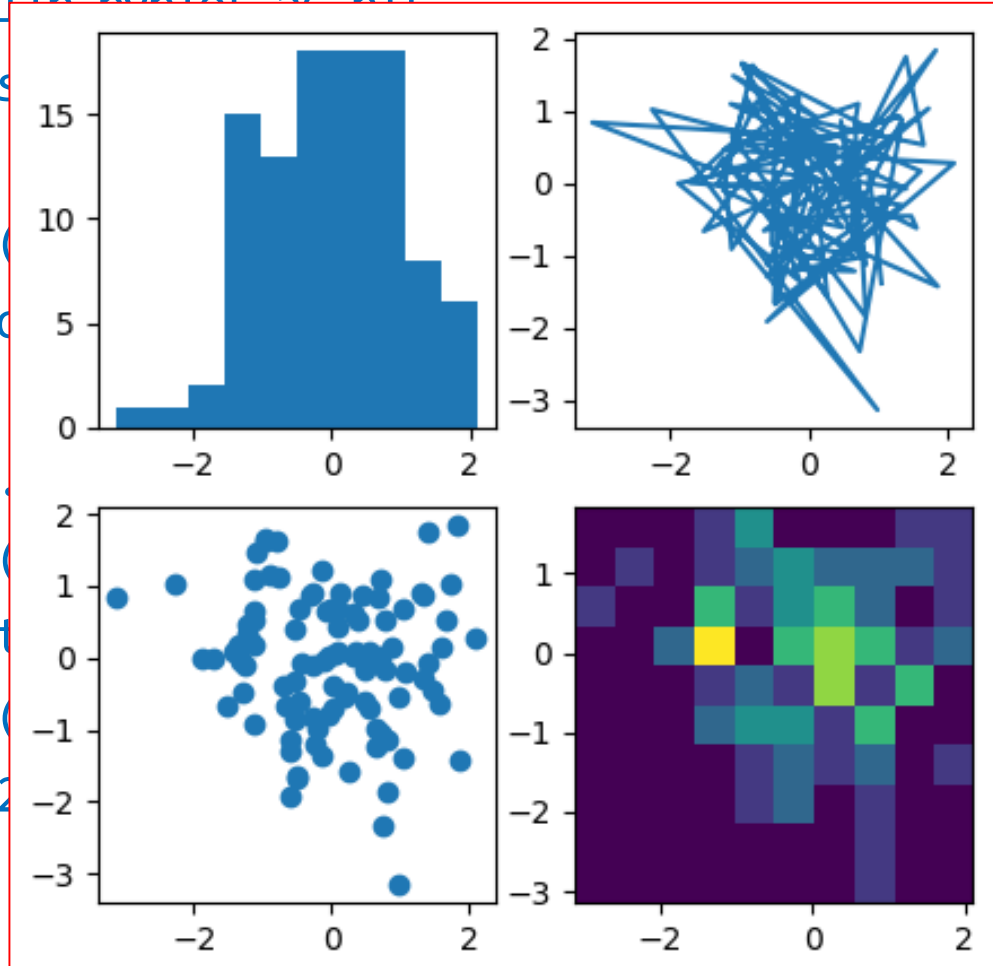
```
axs[0, 0].hist(data)
```

```
axs[1, 0].scatter(data, data)
```

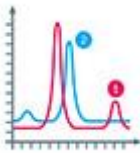
```
axs[0, 1].plot(data)
```

```
axs[1, 1].hist2d(data, data)
```

```
plt.show()
```



Nạp ảnh (png)



```
import matplotlib.pyplot as plt
import matplotlib.image as mpimg
```

```
image = mpimg.imread("txnam.png")
```

```
fig, axs = plt.subplots(2, 2, figsize=(5, 5))
```

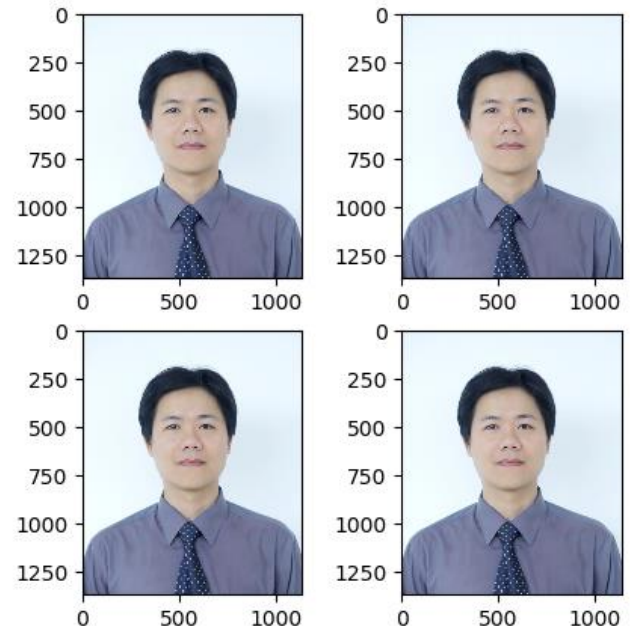
```
axs[0, 0].imshow(image)
```

```
axs[1, 0].imshow(image)
```

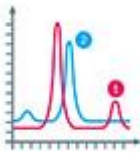
```
axs[0, 1].imshow(image)
```

```
axs[1, 1].imshow(image)
```

```
plt.show()
```



Lưu biểu đồ ra file



```
import matplotlib.pyplot as plt
import matplotlib.image as mpimg
```

```
image = mpimg.imread("txnam.png")
```

```
fig, axs = plt.subplots(2, 2, figsize=(5, 5))
```

```
axs[0, 0].imshow(image)
```

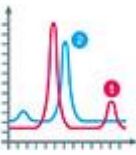
```
axs[1, 0].imshow(image)
```

```
axs[0, 1].imshow(image)
```

```
axs[1, 1].imshow(image)
```

```
plt.savefig('1.png')
```

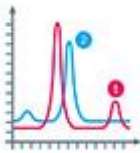
```
plt.savefig('1.pdf')
```



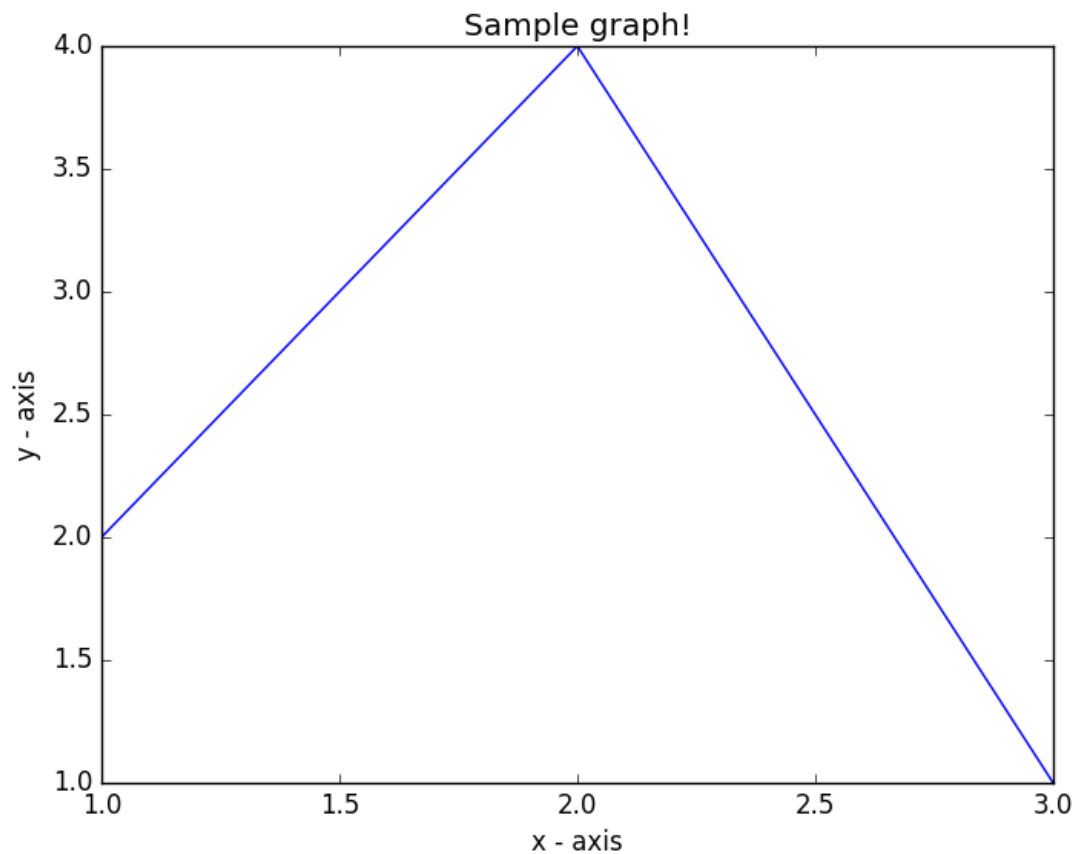
Phần 4

Bài tập

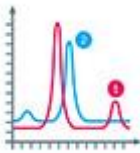
Bài tập



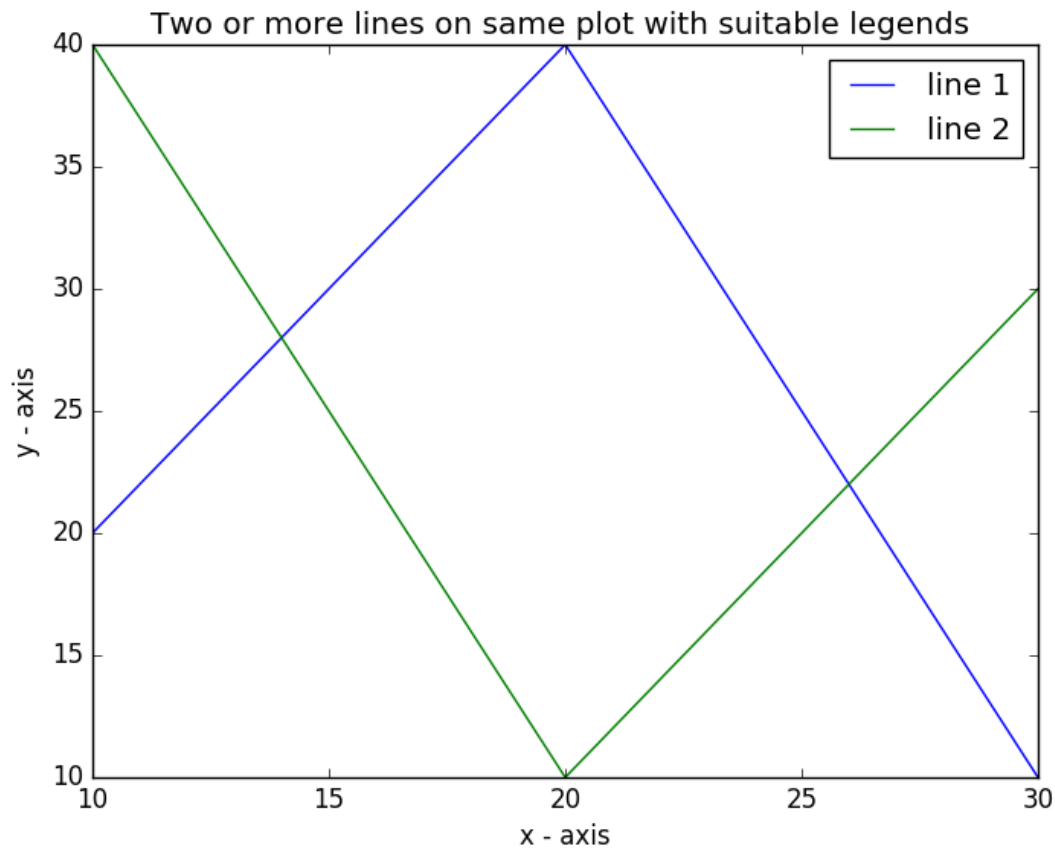
1. Viết đoạn chương trình python vẽ biểu đồ sau



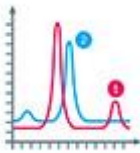
Bài tập



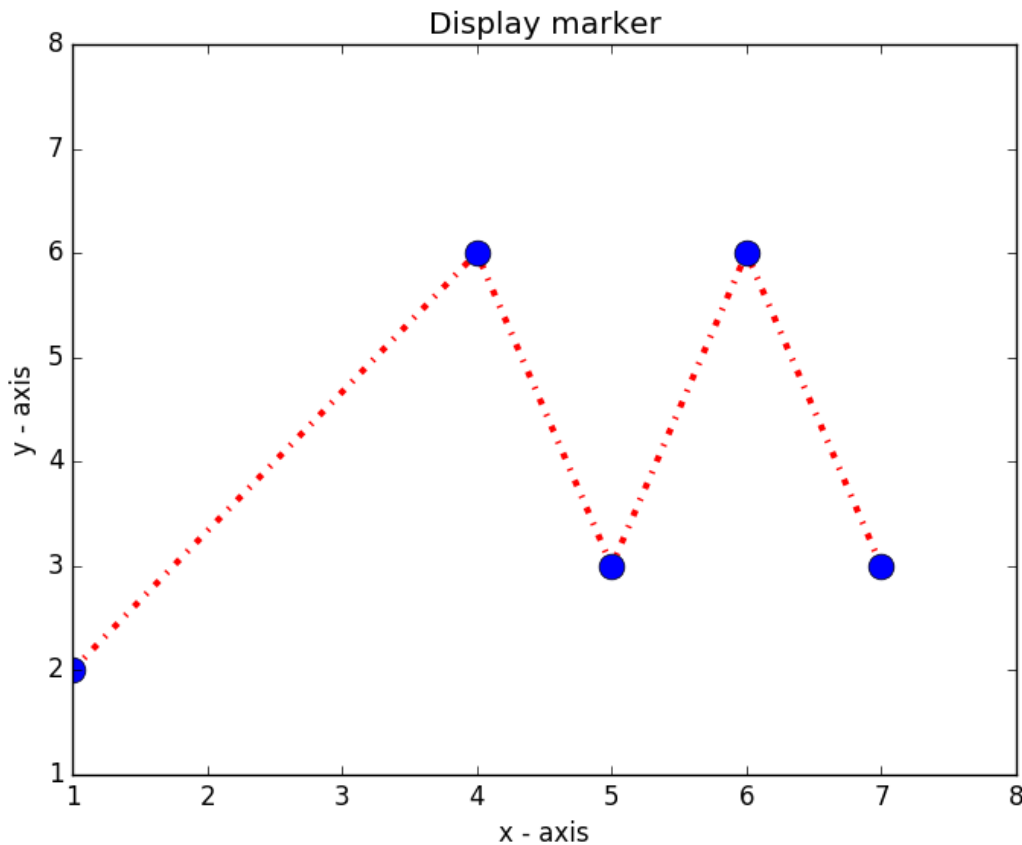
2. Viết đoạn chương trình python vẽ biểu đồ sau



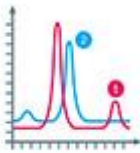
Bài tập



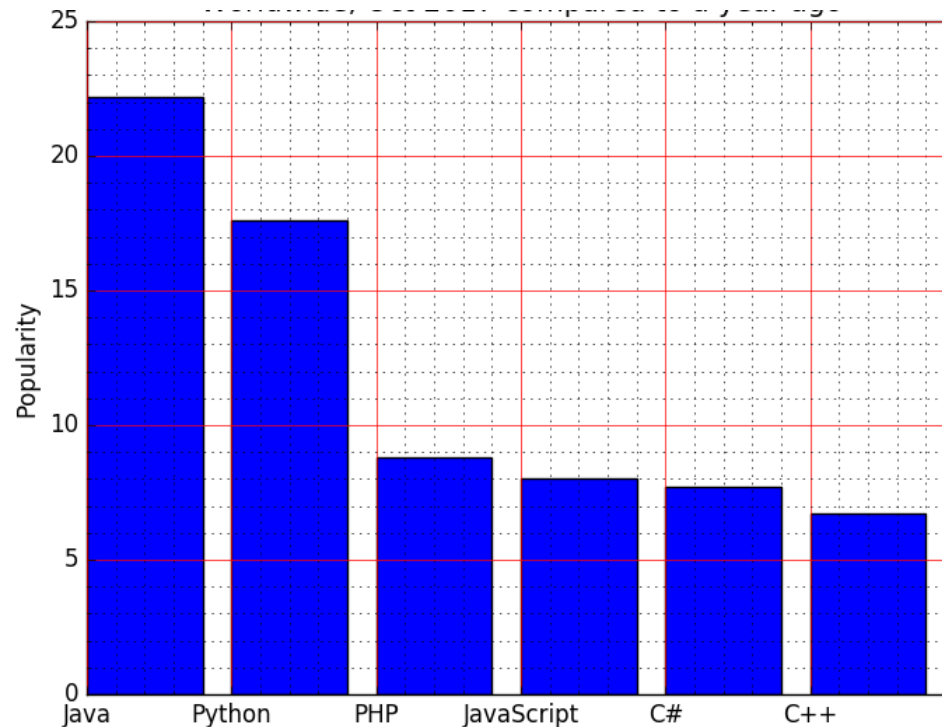
3. Viết đoạn chương trình python vẽ biểu đồ sau



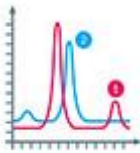
Bài tập



4. Vẽ biểu đồ như hình dưới ứng với bộ dữ liệu sau
- Languages: Java, Python, PHP, JavaScript, C#, C++
 - Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

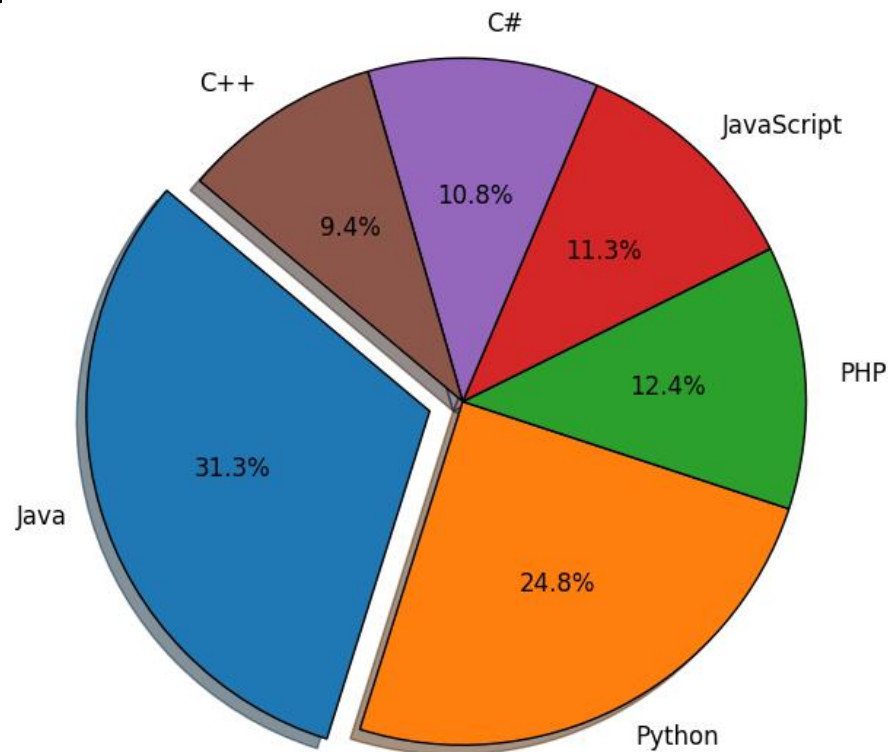


Bài tập

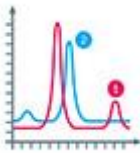


5. Vẽ biểu đồ như hình dưới ứng với bộ dữ liệu sau

- Languages: Java, Python, PHP, JavaScript, C#, C++
- Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7



Chữa bài 1



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

```
x = (1, 2, 3)
```

```
y = (2, 4, 1)
```

```
plt.plot(x, y)
```

```
plt.axis([1,3,1,4])
```

```
plt.xticks(np.arange(1,3.1,0.5))
```

```
plt.xlabel('x-axis')
```

```
plt.ylabel('y-axis')
```

```
plt.title('Sample graph!')
```

```
plt.show()
```