

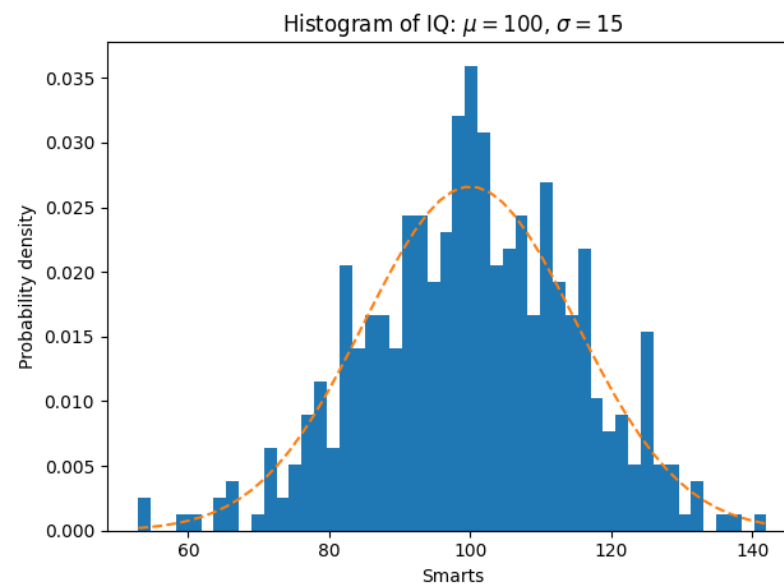
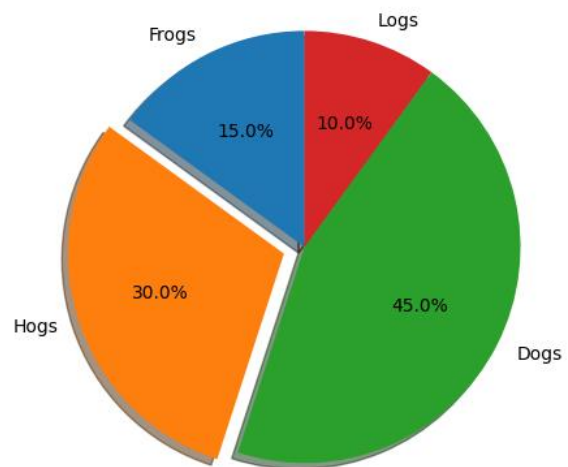
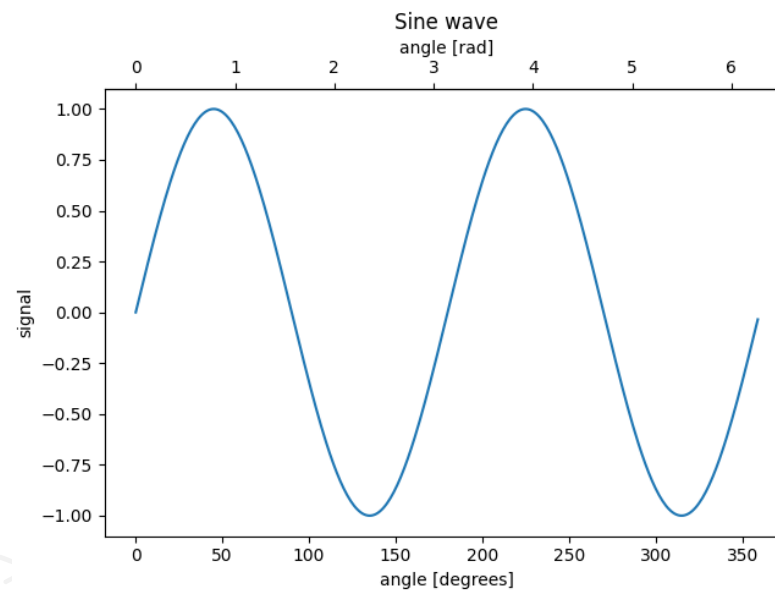
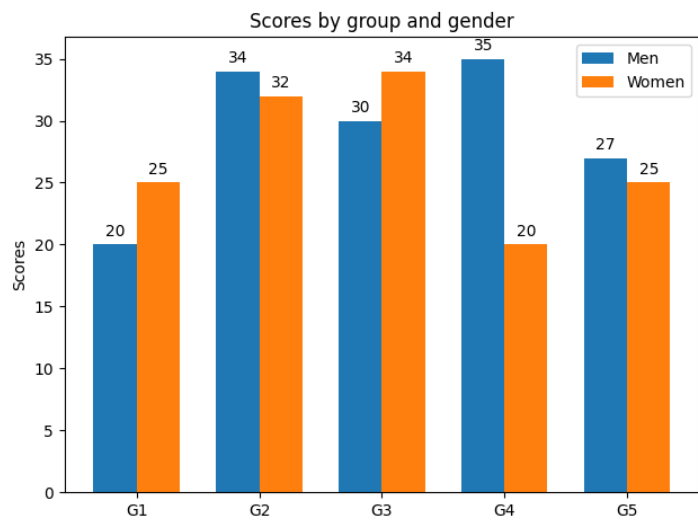
人工智能程序设计

M2 科学计算与数据分析基础

2.4 Matplotlib与可视化

张 莉





Matplotlib 及可视化

1. 绘图基本方法
2. 图形属性控制
3. 基于pandas的绘图
4. 常见类型图举例
5. 其他绘图模块例

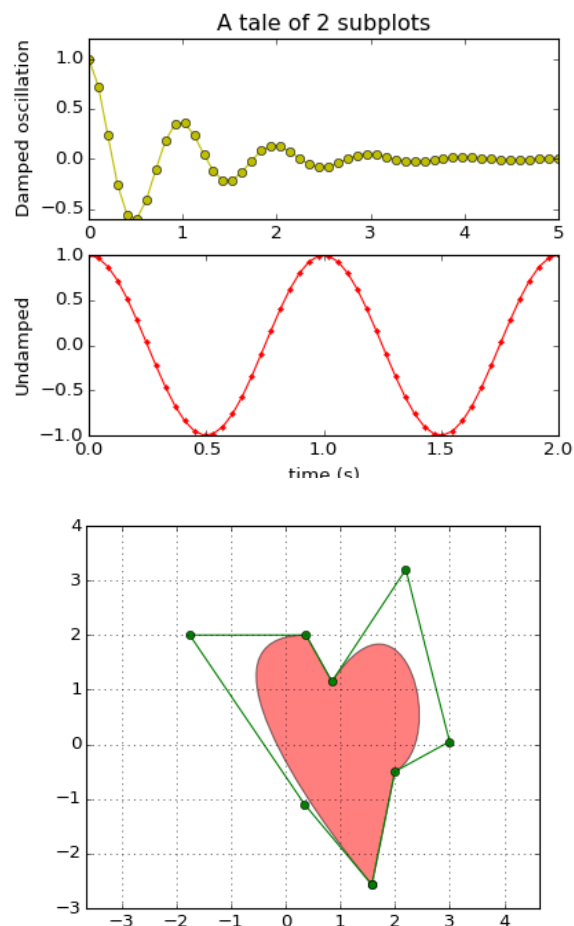


人工智能程序设计

MATPLOTLIB绘图基本方法

南京大学本科《人工智能程序设计》课程专用

Matplotlib绘图



- **Matplotlib绘图**

著名Python绘图库，主要用于二维绘图

— 画图质量高

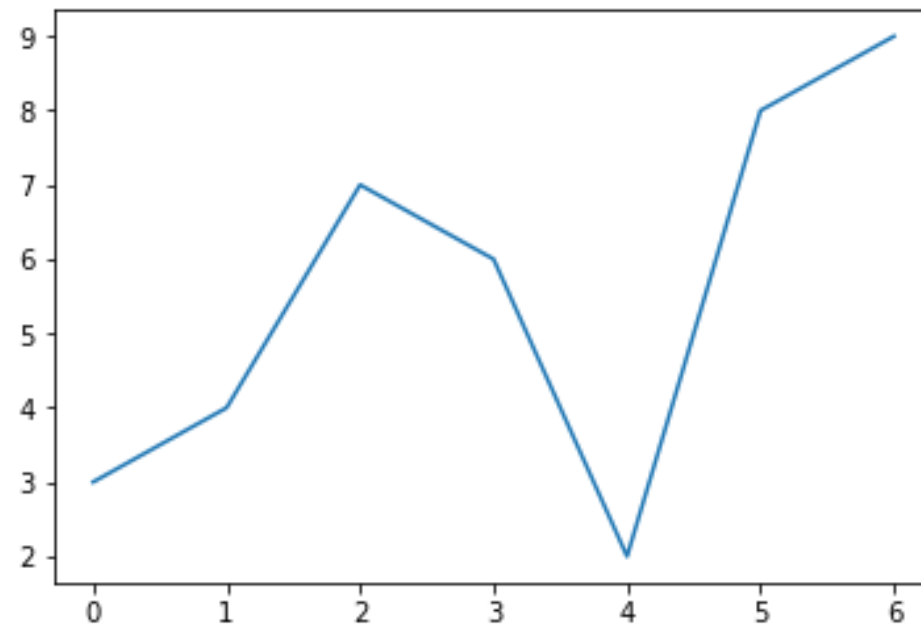
— 方便快捷的绘图模块

- 绘图API——pyplot模块

折线图

Source

```
>>> import matplotlib.pyplot as plt  
>>> plt.plot([3, 4, 7, 6, 2, 8, 9])  
>>> plt.show()
```



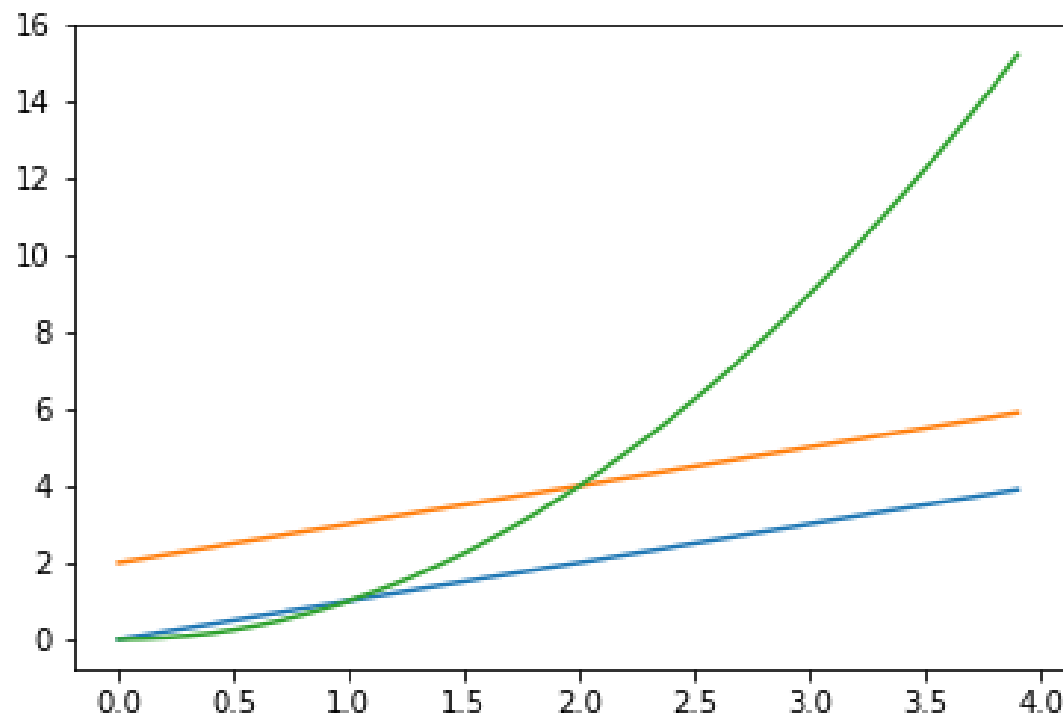
```
plt.plot(range(7), [3, 4, 7, 6, 2, 8, 9])
```

折线图-绘制多组数据

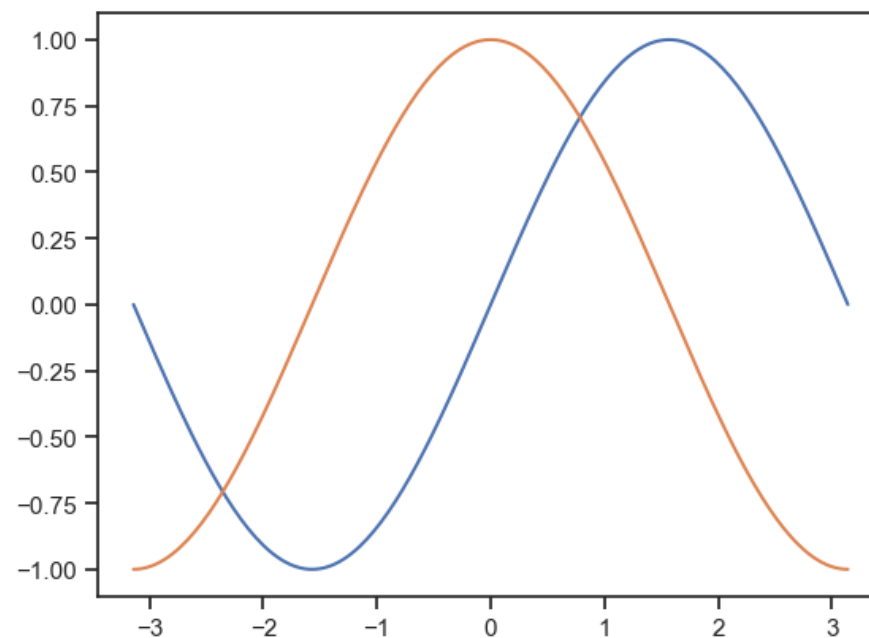
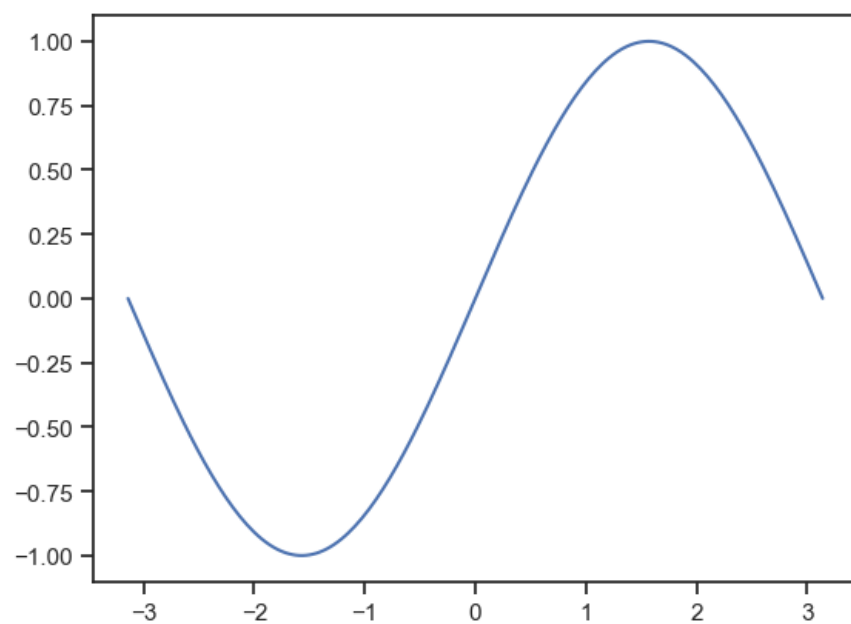
- NumPy数组也可以作为Matplotlib的参数
- 多组成对数据绘图

Source

```
>>> import numpy as np
>>> import matplotlib.pyplot as plt
>>> t=np.arange(0.,4.,0.1)
>>> plt.plot(t, t, t, t+2, t, t**2)
```



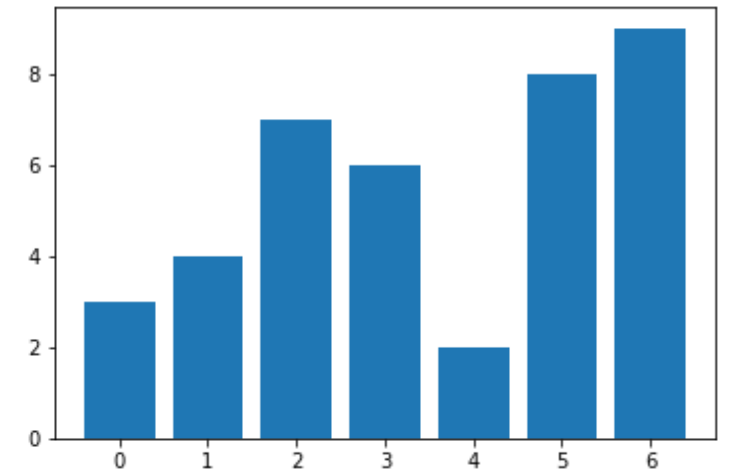
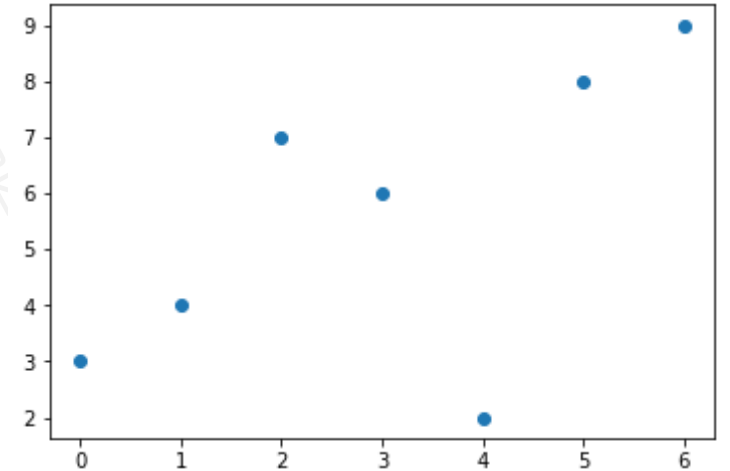
绘图 — 1组数据&2组数据



绘制其他类型的图

Source

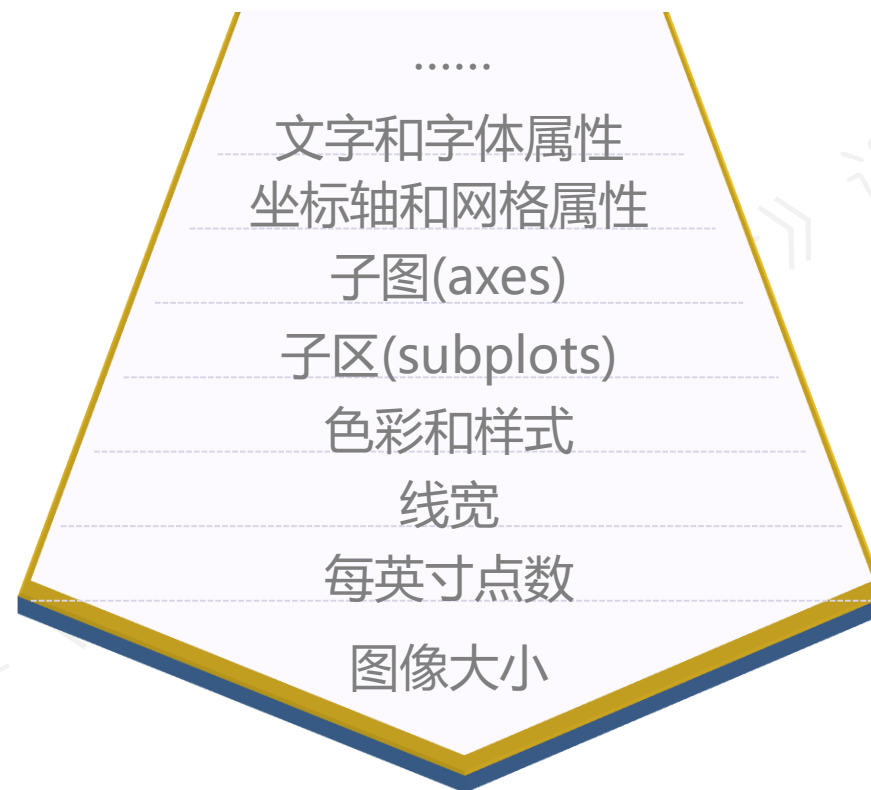
```
>>> import matplotlib.pyplot as plt  
>>> plt.scatter(range(7), [3, 4, 7, 6, 2, 8, 9])  
>>> plt.bar(range(7), [3, 4, 7, 6, 2, 8, 9])
```



人工智能程序设计

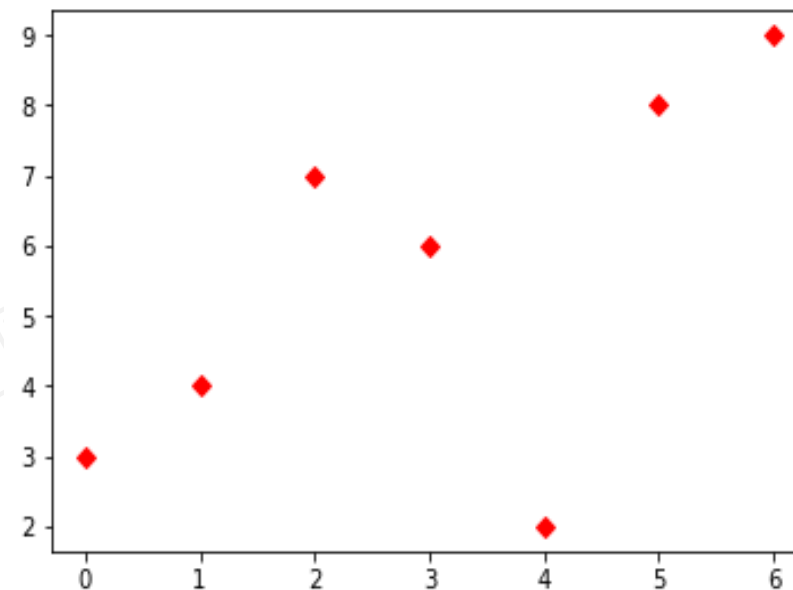
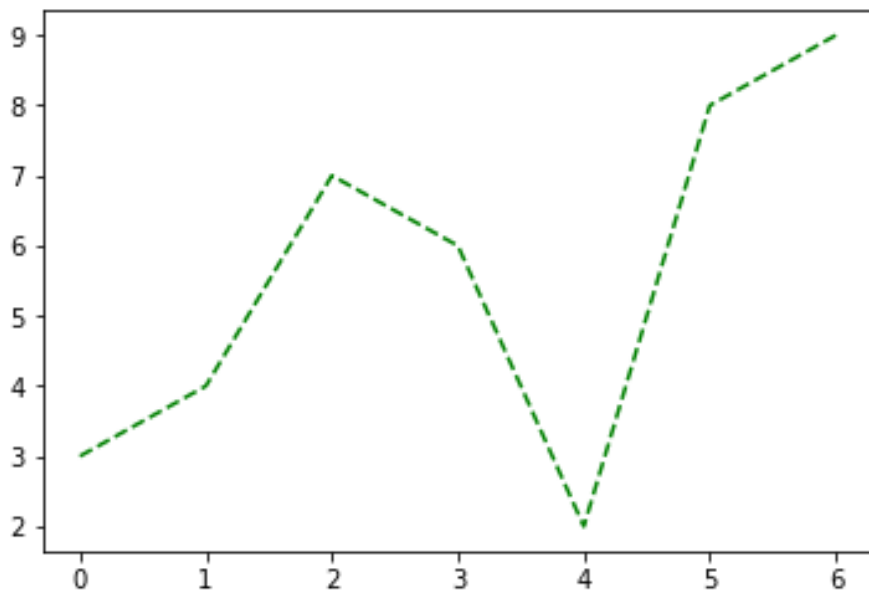
2 MATPLOTLIB图形属性控制

Matplotlib属性



Matplotlib可以控制的默认属性

色彩和样式



```
plt.plot(range(7), [3, 4, 7, 6, 2, 8, 9], 'g--' )  
plt.plot(range(7), [3, 4, 7, 6, 2, 8, 9], 'rD' )
```

色彩和样式

符号	颜色
b	blue
g	green
r	red
c	cyan
m	magenta
Y	yellow
k	black
w	white

线型	描述
'-'	solid
'--'	dashed
'-.'	dash_dot
':'	dotted
'None'	draw nothing
''	draw nothing
''	draw nothing

标记	描述
"o"	circle
"v"	triangle_down
"s"	square
"p"	pentagon
"*"	star
"h"	hexagon1
"+"	plus
"D"	diamond
...	...

多种属性

File

Filename: 2.py

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

```
import numpy as np
```

```
plt.figure(figsize=(8,6),dpi=100)
```

```
t=np.arange(0.,4.,0.1)
```

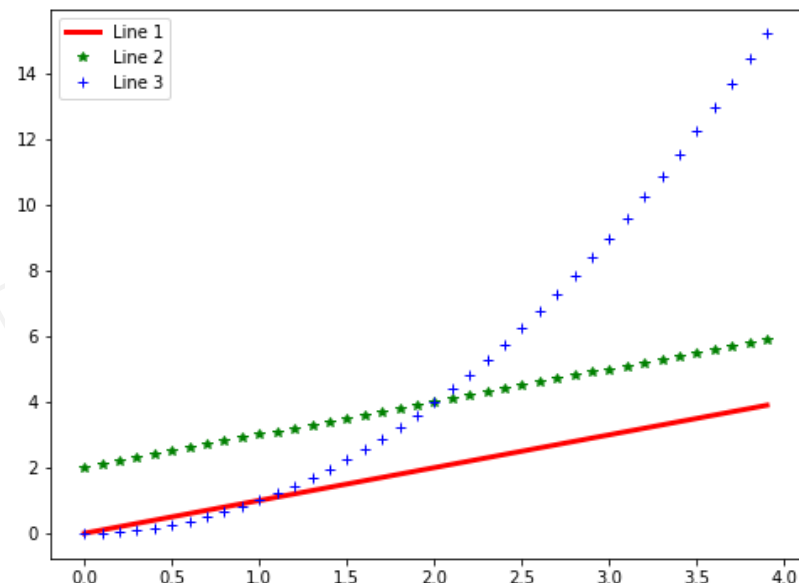
```
plt.plot(t,t,color='red',linestyle='-',linewidth=3,label='Line 1')
```

```
plt.plot(t,t+2,color='green',linestyle='',marker='*',linewidth=3,label='Line 2')
```

```
plt.plot(t,t**2,color='blue',linestyle='',marker='+',linewidth=3,label='Line 3')
```

```
plt.legend(loc='upper left')
```

```
plt.show()
```



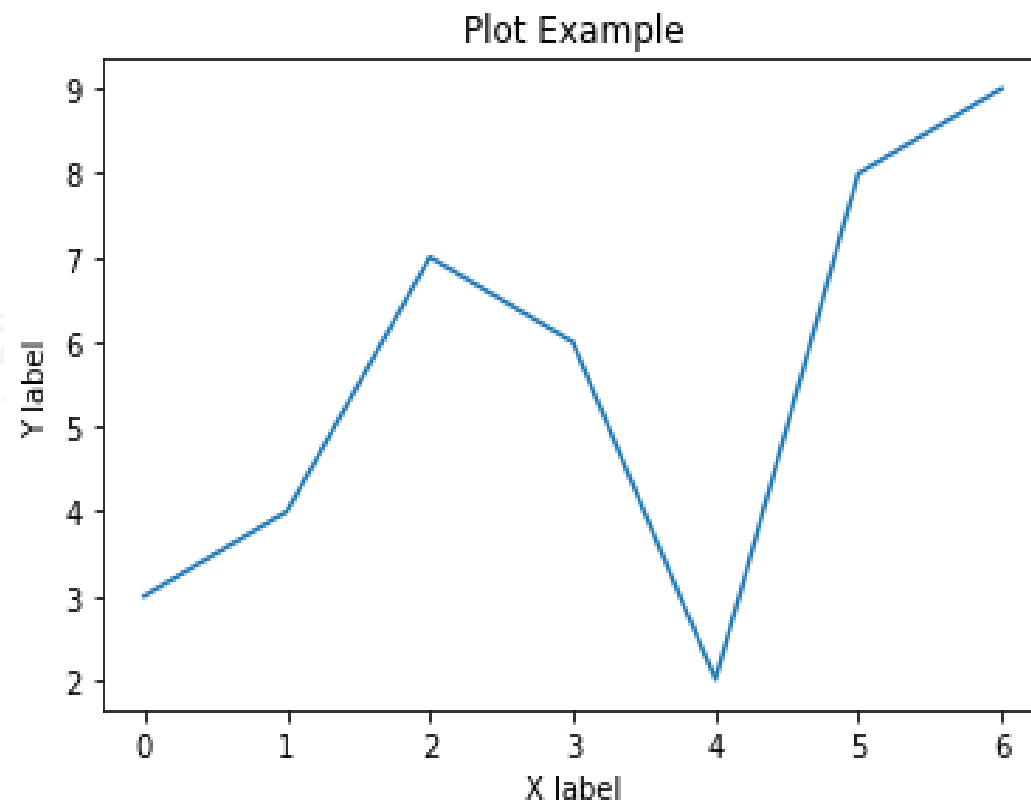
文字

加标题：图、横轴和纵轴

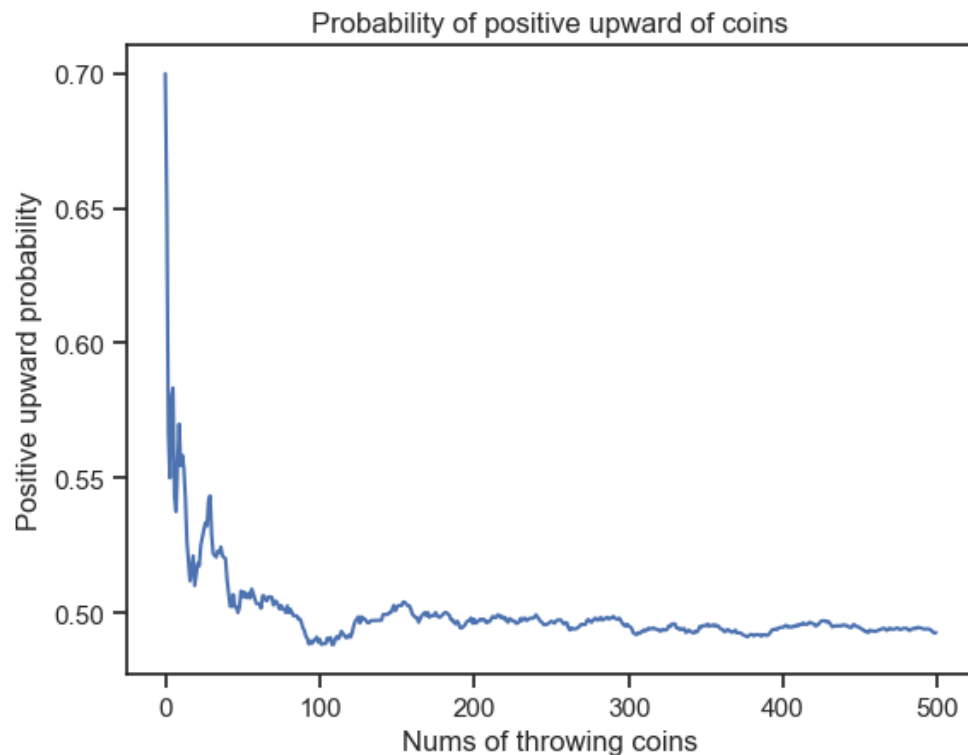
F_{ile}

Filename: 3.py

```
import matplotlib.pyplot as plt
plt.title('Plot Example')
plt.xlabel('X label')
plt.ylabel('Y label')
plt.plot(range(7), [3, 4, 7, 6, 2, 8, 9])
```



投掷硬币—正面朝上的累计概率趋势图



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

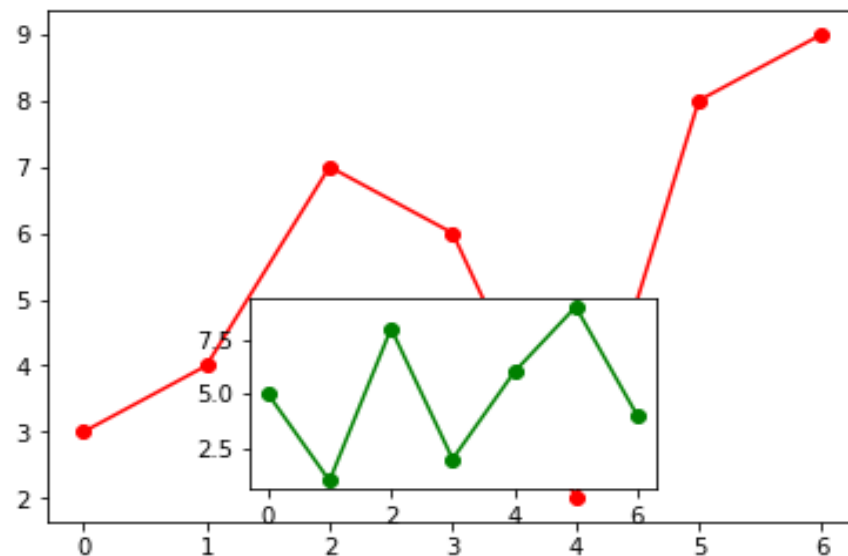
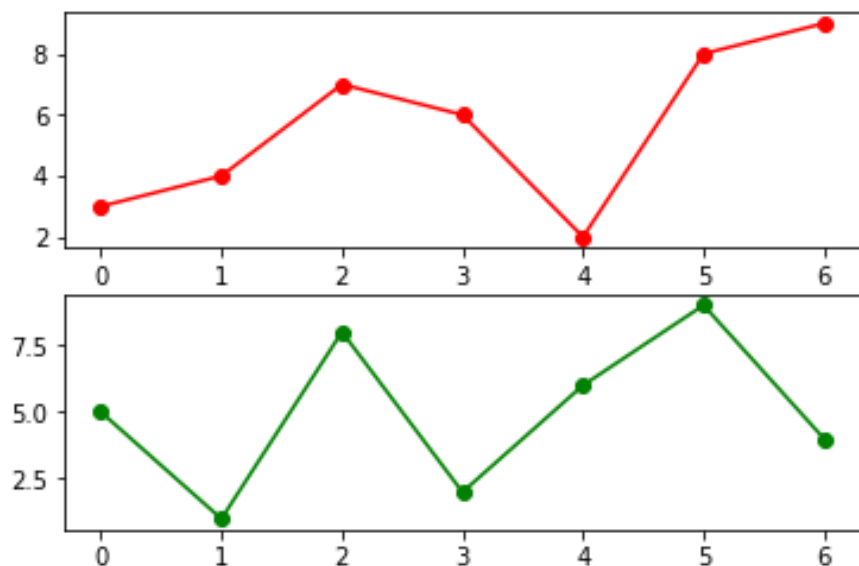
```
batch = 500
result, result_mean = [], []
```

```
for i in range(batch):
    for k in range(10):
        result.append(random.randint(0, 1))
        result_mean.append(np.mean(result))
```

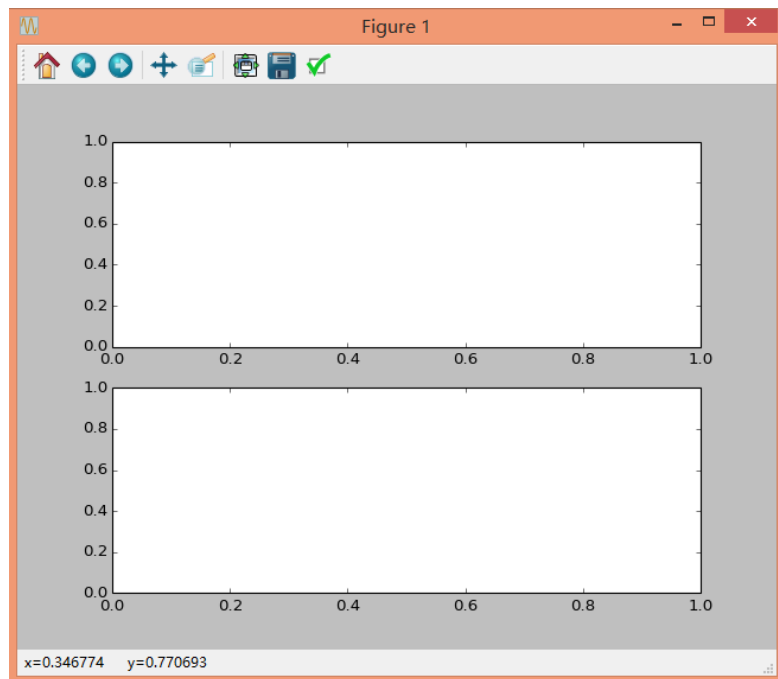
```
plt.plot(np.arange(batch), result_mean)
plt.xlabel('Nums of throwing coins')
plt.ylabel('Positive upward probability')
plt.title('Probability of positive upward of coins')
plt.show()
```


绘制子图

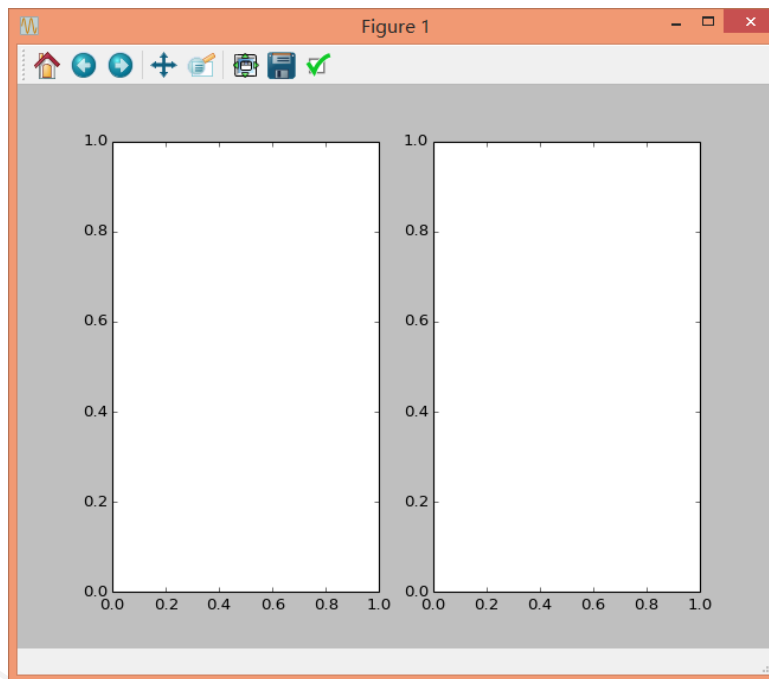
- 在Matplotlib中绘图在当前图形（figure）和当前坐标系（axes）中进行，默认在编号为1的figure中绘图，可以在一个图的多个区域分别绘图
- 使用subplot()函数和axes()函数



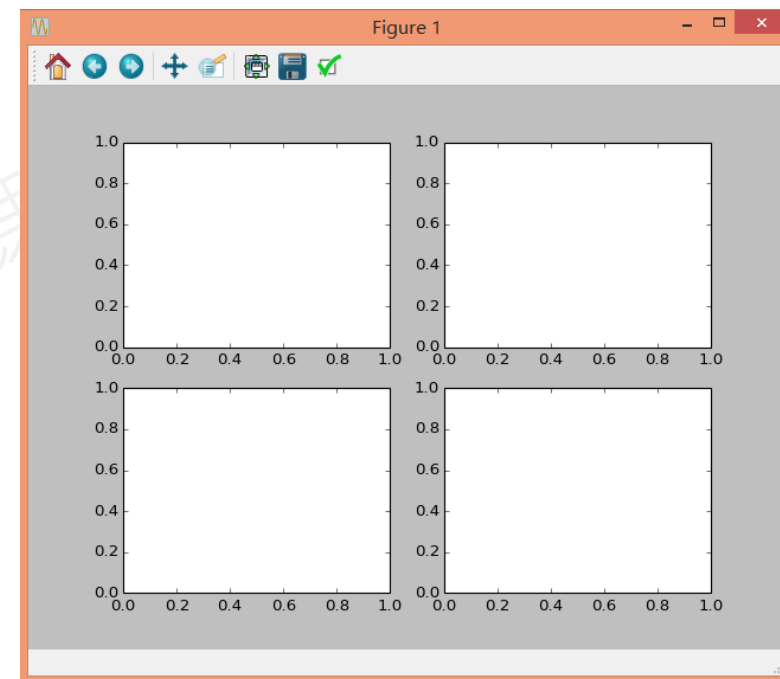
子图-subplots



```
plt.subplot(211)  
plt.subplot(212)
```



```
plt.subplot(121)  
plt.subplot(122)
```



```
plt.subplot(221)  
plt.subplot(222)  
plt.subplot(223)  
plt.subplot(224)
```

子图-subplots

File

Filename: 3.py

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

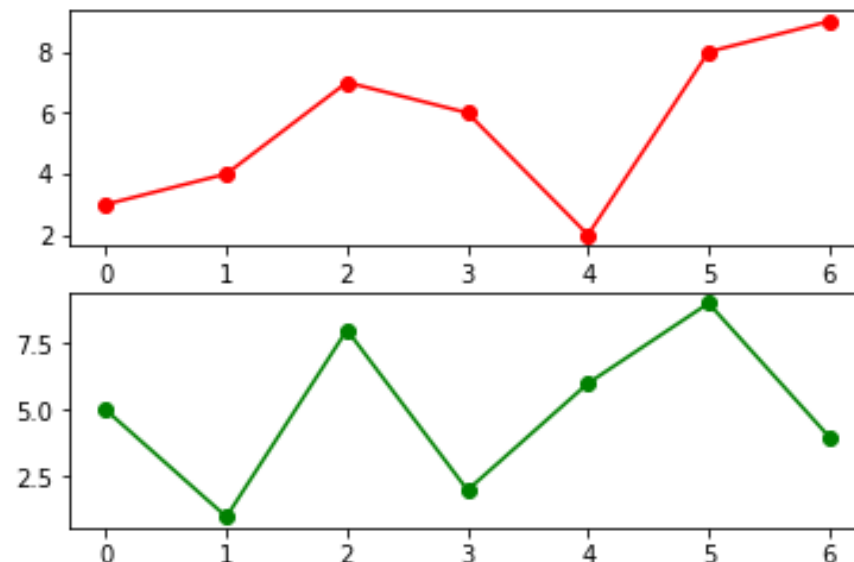
```
plt.figure(1) # 默认创建, 缺省
```

```
plt.subplot(211) # 第一个子图
```

```
plt.plot(range(7), [3, 4, 7, 6, 2, 8, 9], color = 'r', marker = 'o')
```

```
plt.subplot(212) # 第二个子图
```

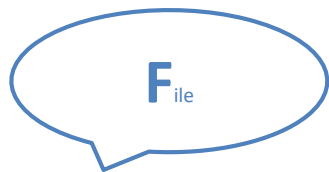
```
plt.plot(range(7), [5, 1, 8, 2, 6, 9, 4], color = 'green', marker = 'o')
```



子图-subplots

ax0 为子图1

ax1 为子图2



Filename: 4.py

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

```
fig, (ax0, ax1) = plt.subplots(2, 1)
```

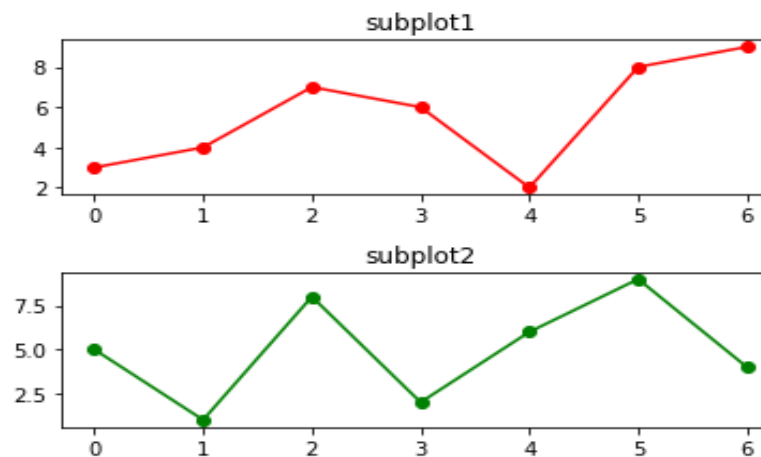
```
ax0.plot(range(7), [3, 4, 7, 6, 2, 8, 9], color = 'r', marker = 'o')
```

```
ax0.set_title('subplot1')
```

```
plt.subplots_adjust(hspace = 0.5)
```

```
ax1.plot(range(7), [5, 1, 8, 2, 6, 9, 4], color = 'green', marker = 'o')
```

```
ax1.set_title('subplot2')
```



子图-subplots

将可口可乐公司和IBM公司近一年来股票收盘价的月平均价绘制在一张图中

File

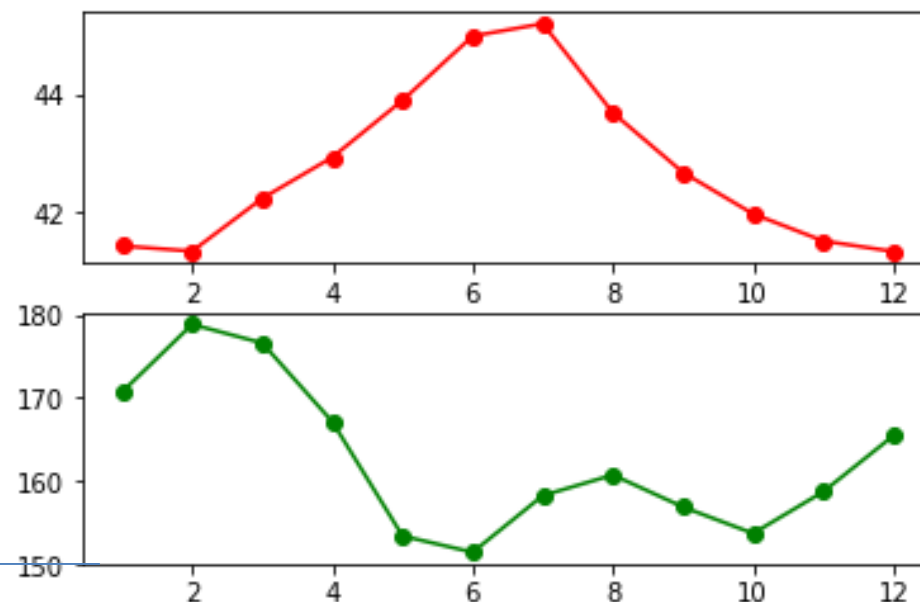
#The data of Coca-Cola and IBM is ready

```
plt.subplot(211)
```

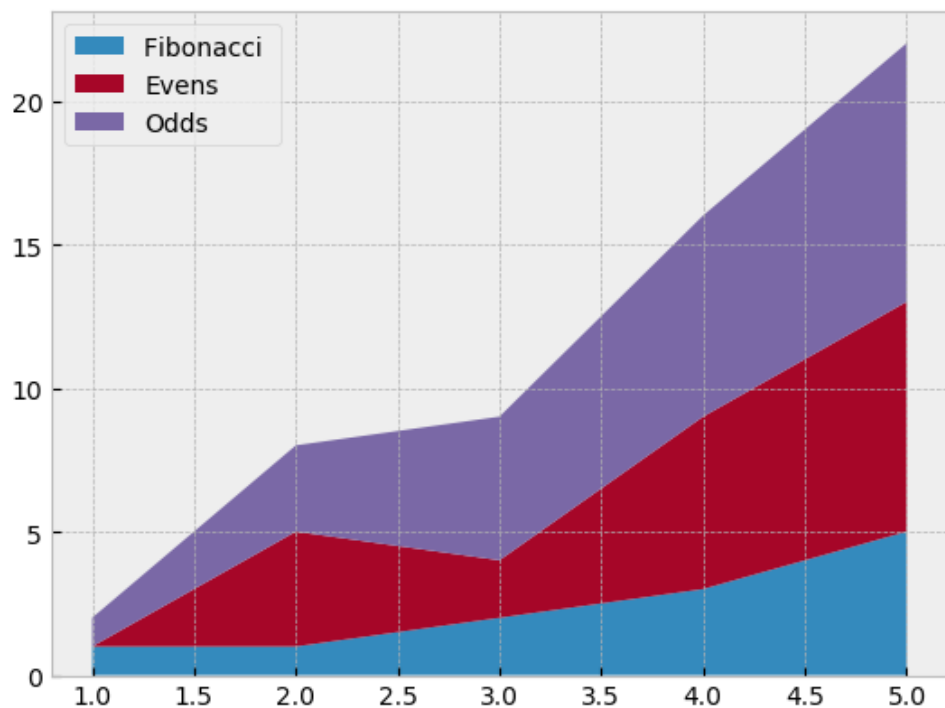
```
plt.plot(x, y, color = 'r', marker = 'o')
```

```
plt.subplot(212)
```

```
plt.plot(xi, yi, color = 'green', marker = 'o')
```



官网example — subplots()的理解



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

```
x = [1, 2, 3, 4, 5]
y1 = [1, 1, 2, 3, 5]
y2 = [0, 4, 2, 6, 8]
y3 = [1, 3, 5, 7, 9]
```

```
labels = ["Fibonacci ", "Evens", "Odds"]
```

```
fig, ax = plt.subplots()
ax.stackplot(x, y1, y2, y3, labels=labels)
ax.legend(loc='upper left')
plt.show()
```

子图-axes

`axes([left,bottom,width,height])`
参数范围为(0,1)

File

Filename: 5.py

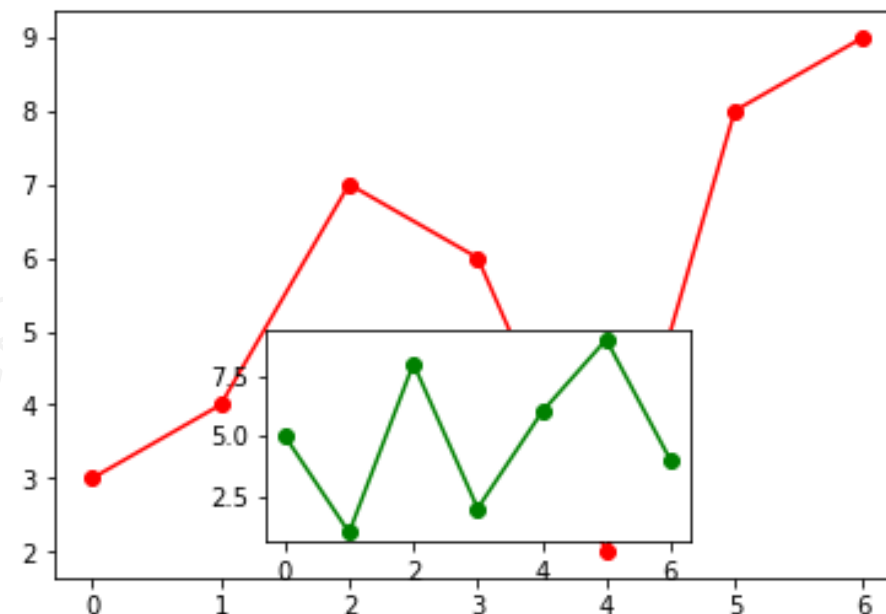
```
import matplotlib.pyplot as plt
```

```
plt.axes([.1, .1, 0.8, 0.8])
```

```
plt.plot(range(7), [3, 4, 7, 6, 2, 8, 9], color = 'r', marker = 'o')
```

```
plt.axes([.3, .15, 0.4, 0.3])
```

```
plt.plot(range(7), [5, 1, 8, 2, 6, 9, 4], color = 'green', marker = 'o')
```



3 基于PANDAS的绘图

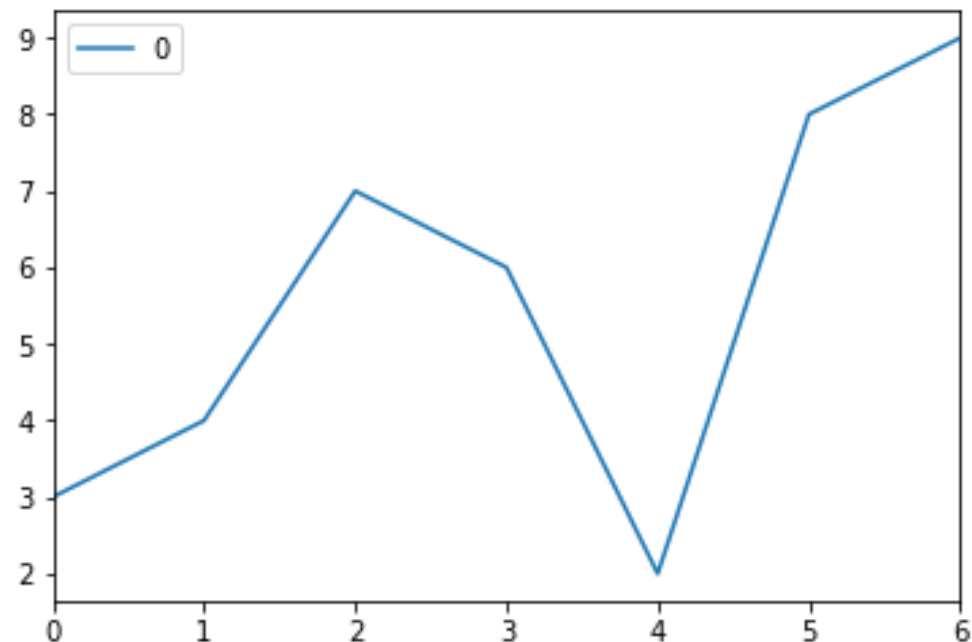
人工智能程序设计

南京大学本科《人工智能程序设计》课程专用

pandas绘图

Source

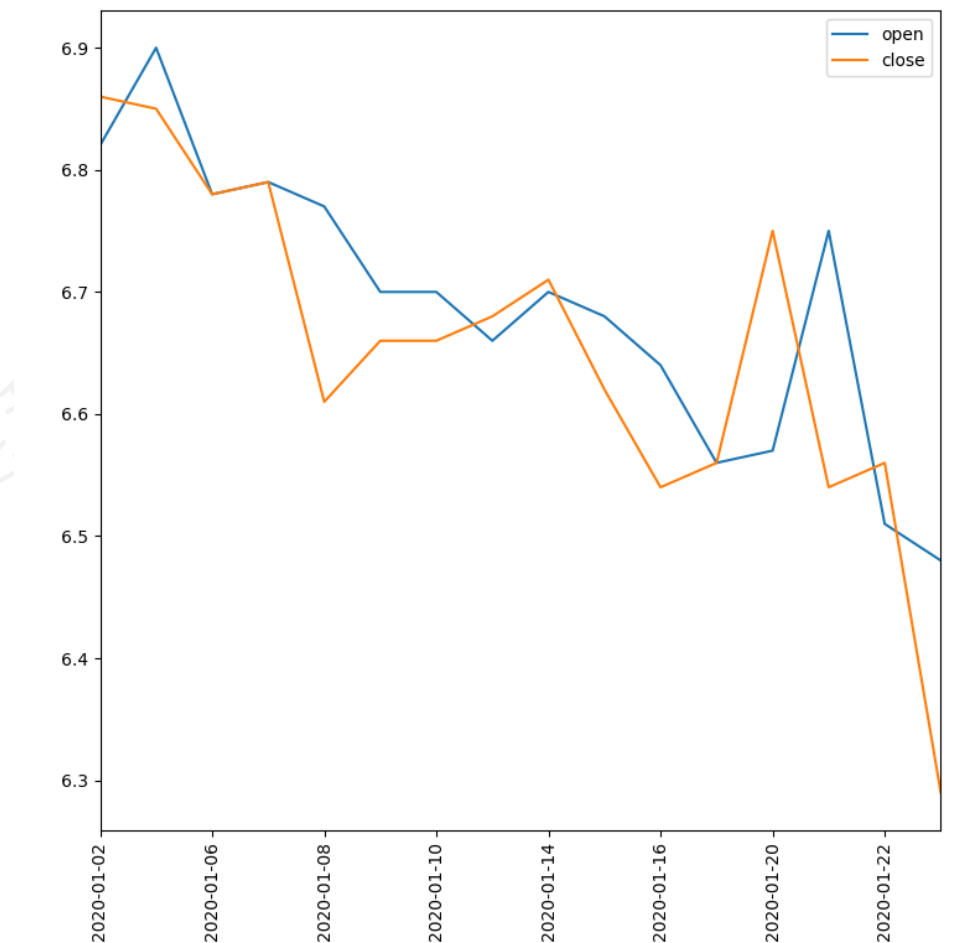
```
>>> import pandas as pd
>>> data = [3, 4, 7, 6, 2, 8, 9]
>>> pDF = pd.DataFrame(data)
>>> pDF.plot()
```



股票数据绘制

绘制“葛洲坝
600068”2020年1月
份的股票数据开盘价
和收盘价的折线图

```
df.sort_index().plot()  
plt.xticks(rotation = 90)
```



pandas绘图

File

Filename: 6.py

```
import pandas as pd
```

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

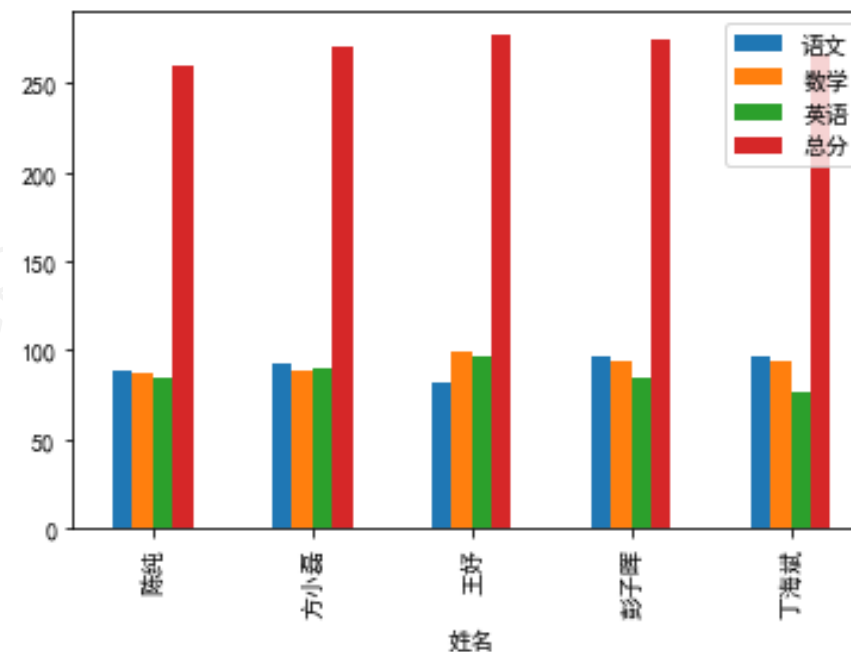
```
# 解决图中中文显示方块问题
```

```
import matplotlib as mpl
```

```
mpl.rcParams['font.sans-serif'] = ['SimHei']
```

```
df = pd.read_csv('score.csv', encoding = 'gb2312', index_col = '姓名')
```

```
df.plot(kind = 'bar')
```



pandas绘图

F_{ile}

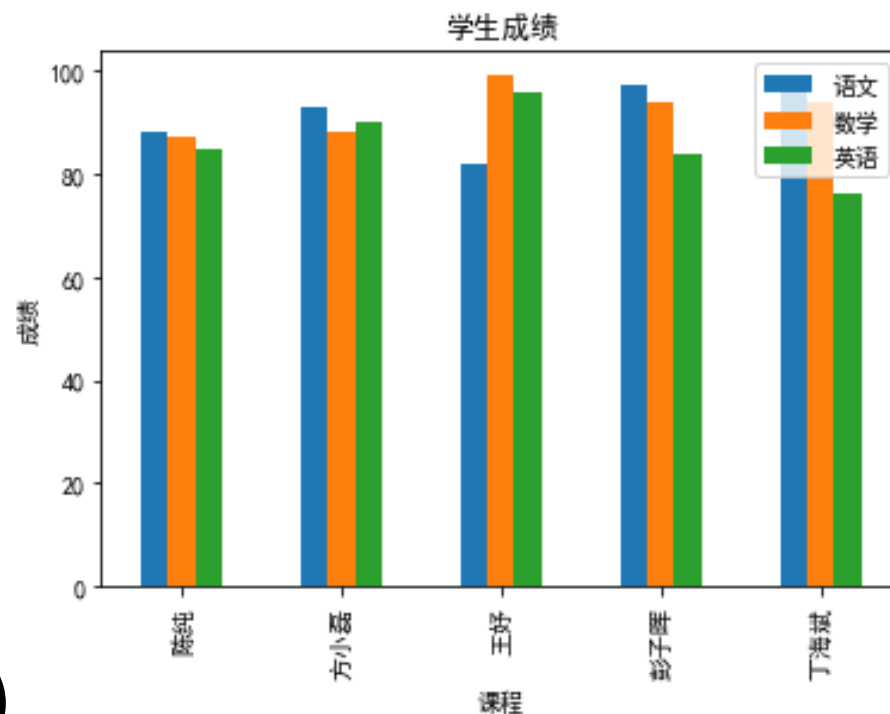
Filename: 6.py

...

```
df = pd.DataFrame(data)
df1 = df.iloc[:, :4]
Ax = df1.plot(kind='bar', title='学生成绩')
Ax.set(xlabel='课程', ylabel='成绩')
```

```
cht.set_xlabel('课程')
```

```
plt.title('学生成绩')
plt.xlabel('课程')
plt.ylabel('成绩')
```

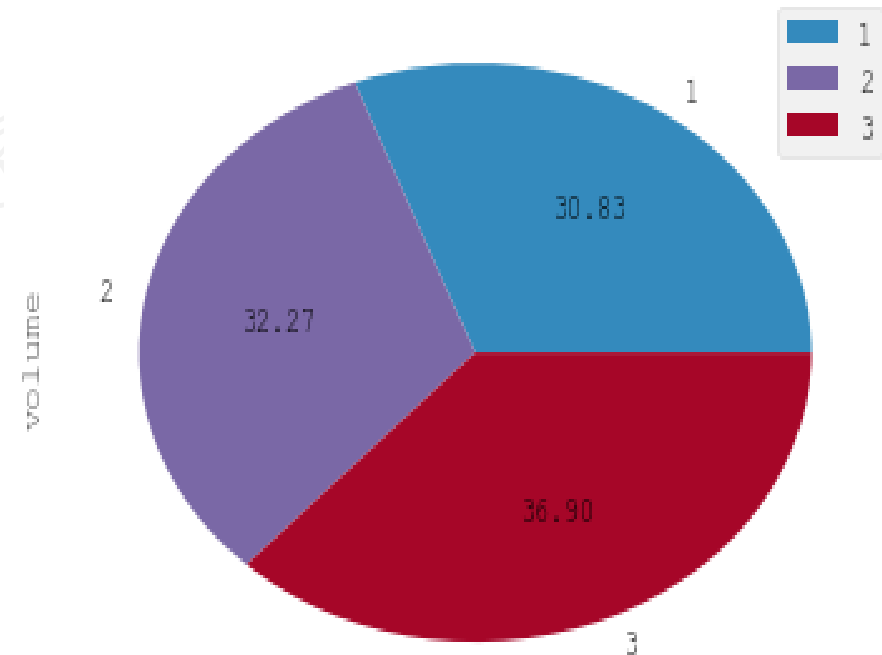


pandas控制图像形式

Intel公司本年度前3个月每个月股票收盘价的占比

```
quotesINTC.plot()
```

```
quotesINTC.plot(kind = 'pie',  
subplots = True, autopct = '%.2f')
```

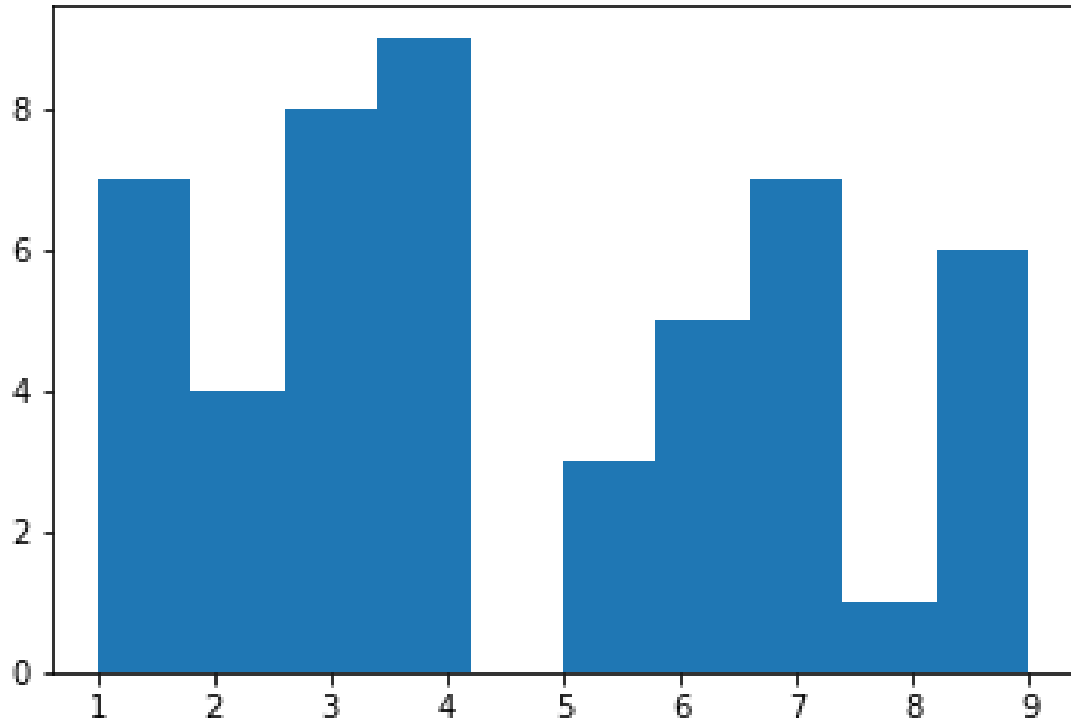


4 常见类型图举例

人工智能程序设计

南京大学本科《人工智能程序设计》课程专用

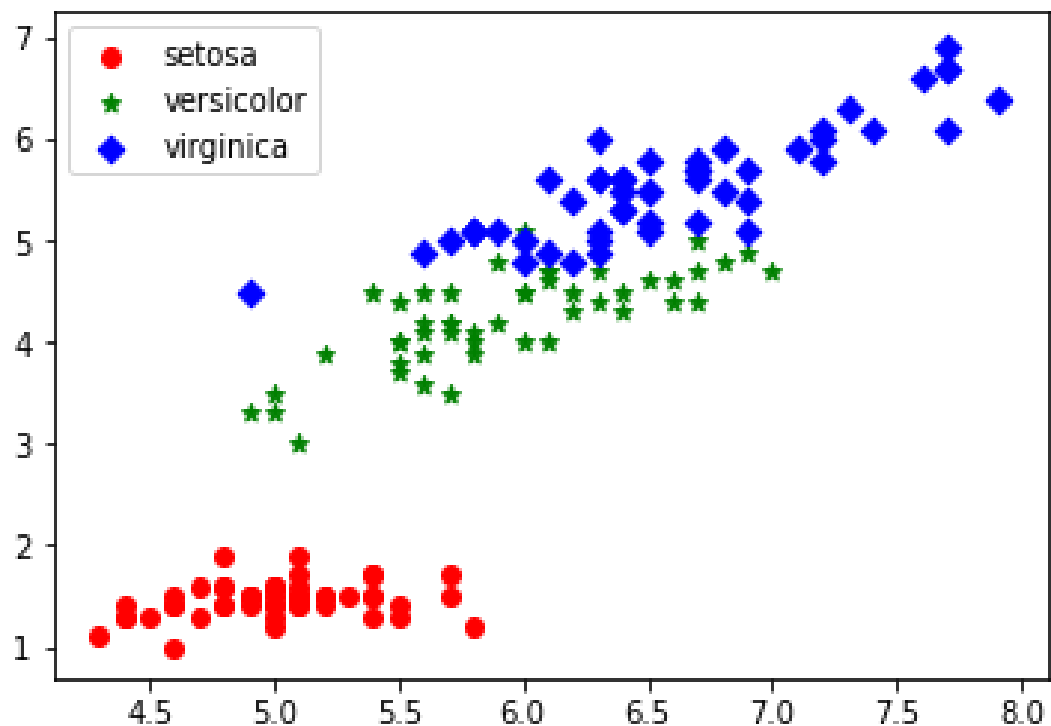
直方图与直方图中的信息



```
x = np.random.randint(1,10,50)  
plt.hist(data)
```

```
array([7, 2, 3, 4, 7, 1, 4, 2, 9, 9, 7, 3, 7, 1, 8, 4, 4, 4, 5, 9, 2, 1, 1, 3, 9,  
6, 1, 6, 3, 6, 1, 6, 5, 9, 4, 7, 1, 3, 3, 6, 2, 3, 3, 7, 4, 9, 4, 4, 5, 7])
```

散点图与散点图中的信息



```
from sklearn import datasets
import matplotlib.pyplot as plt
```

```
iris = datasets.load_iris()
print(iris.data, iris.target)
```

```
X = [item[0] for item in iris.data] # 获取萼片长度
```

```
Y = [item[2] for item in iris.data] # 获取花瓣长度
```

```
plt.scatter(X[:50], Y[:50], color='red', marker='o', label='setosa')
```

```
plt.scatter(X[50:100], Y[50:100], color='green', marker='*', label='versicolor')
```

```
plt.scatter(X[100:], Y[100:], color='blue', marker='D', label='virginica')
```

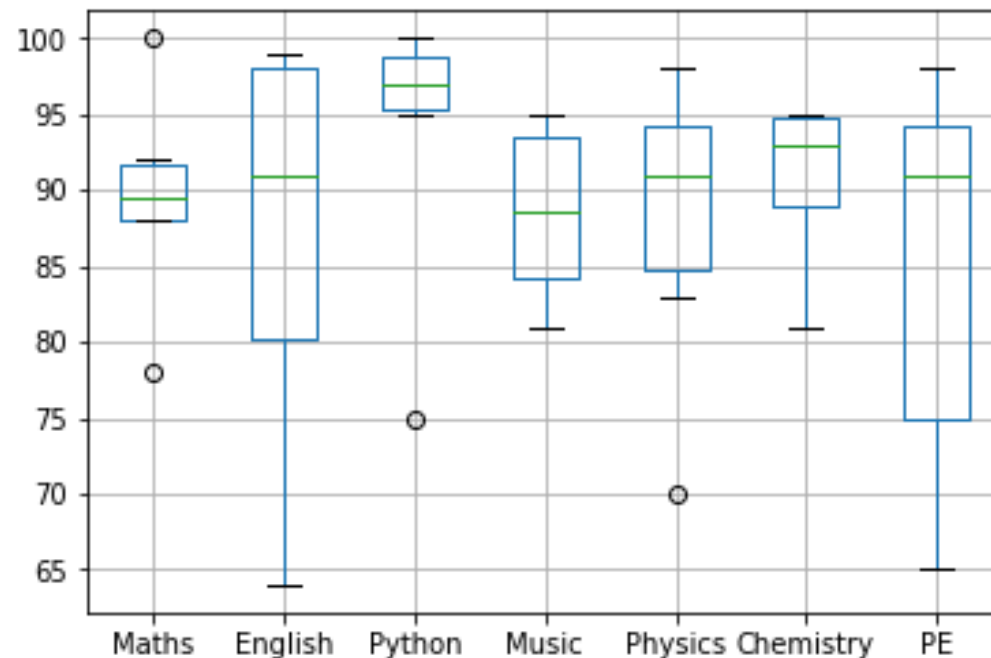
```
plt.legend(loc='best')
```

```
plt.show()
```


箱形图与箱形图中的信息

	Maths	English	Python	Music	Physics	Chemistry	PE
Wang	88	64	96	85	90	81	95
Ma	92	99	95	94	92	94	90
Liu	91	87	99	95	95	92	70
Qian	78	99	75	81	83	88	92
Meng	88	78	98	84	70	95	98
Song	100	95	100	92	98	95	65

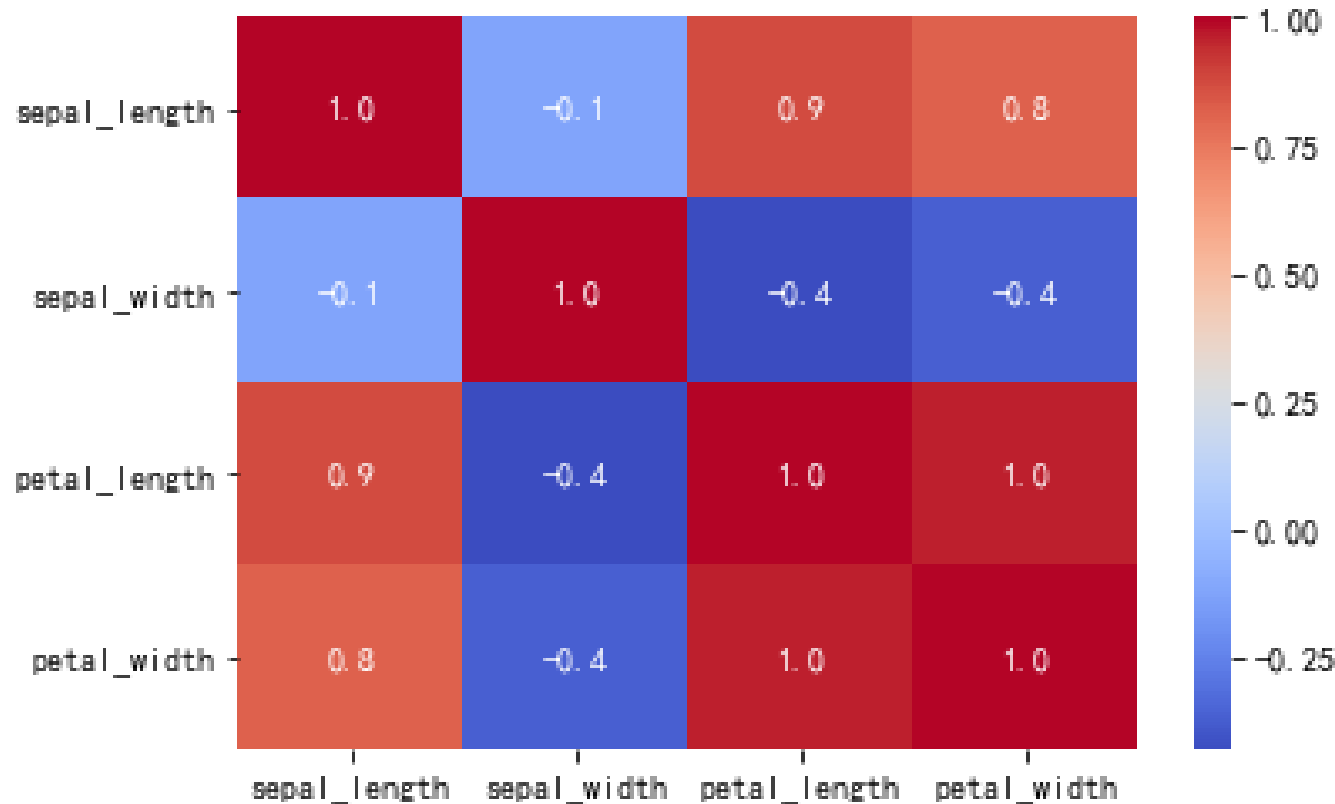
最大值、最小值、中位数、
下四分位数、上四分位



```
plt.boxplot(df)
```

5 人工智能程序设计 其他绘图模块例

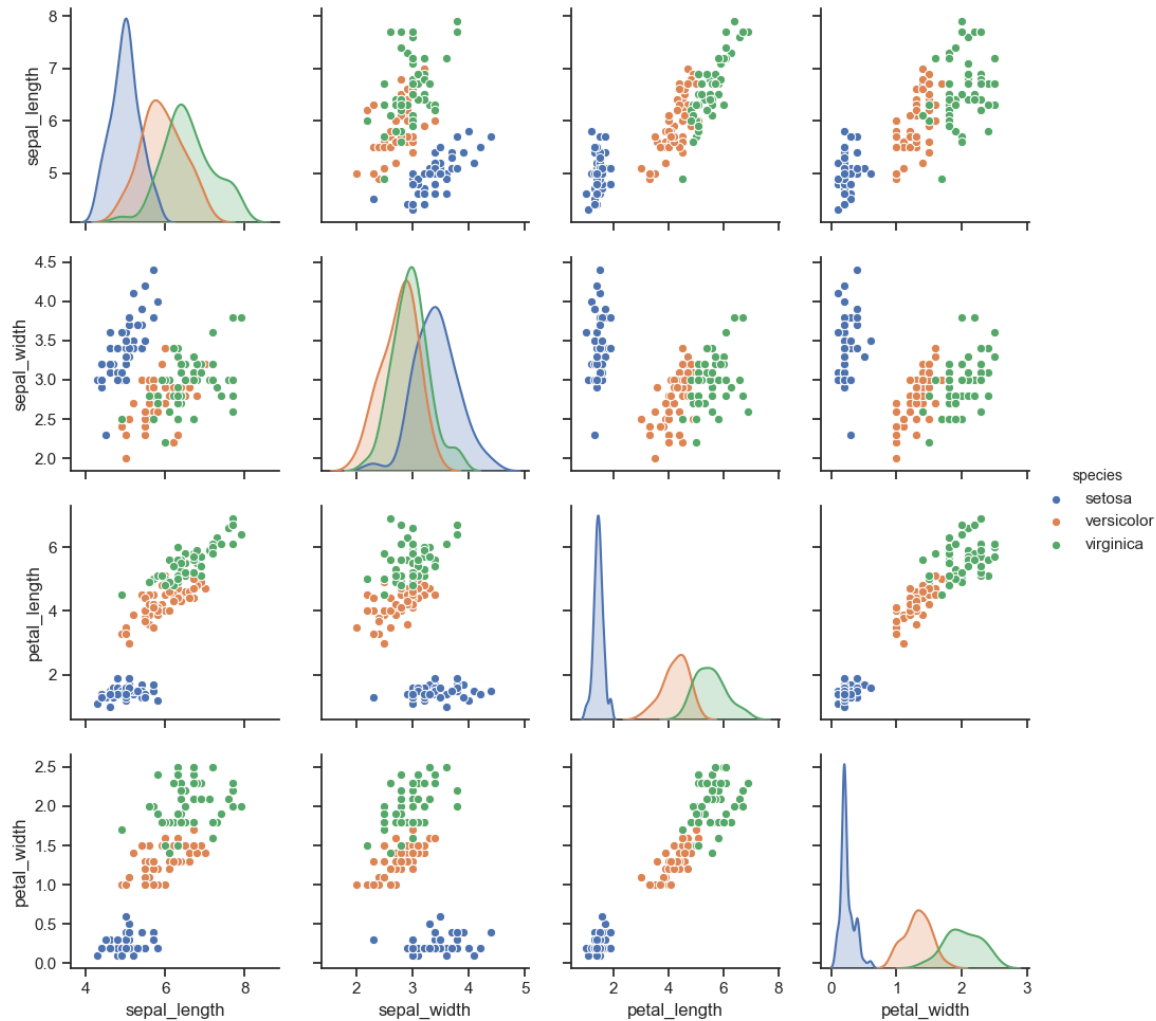
seaborn



<http://seaborn.pydata.org/>

```
iris = sns.load_dataset('iris')  
sns.heatmap(iris.corr(), annot = True, fmt = '.1f', cmap = 'coolwarm')
```

seaborn

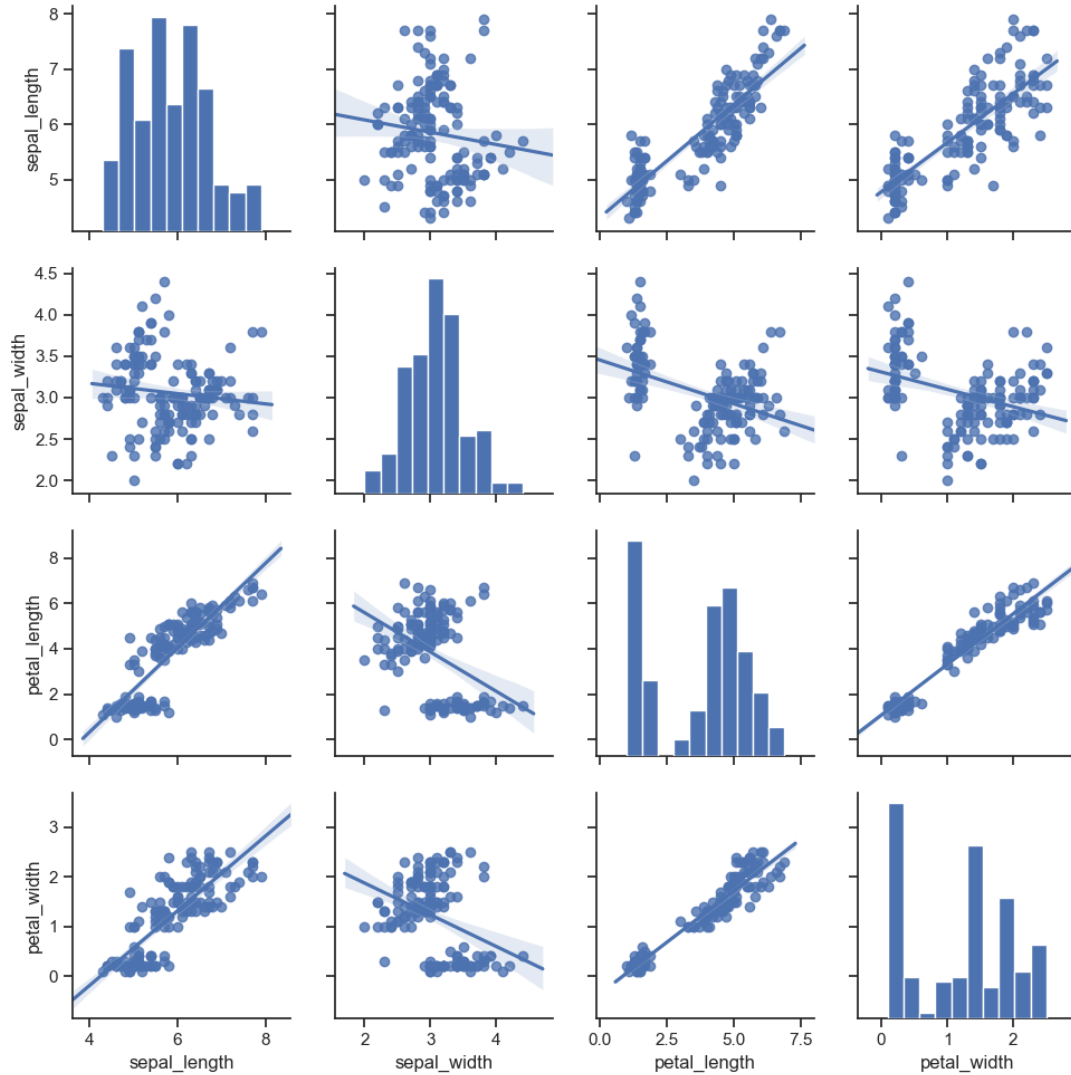


```
import seaborn as sns
```

```
sns.set(style="ticks")
```

```
df = sns.load_dataset("iris")  
sns.pairplot(df, hue="species")  
plt.show()
```

seaborn

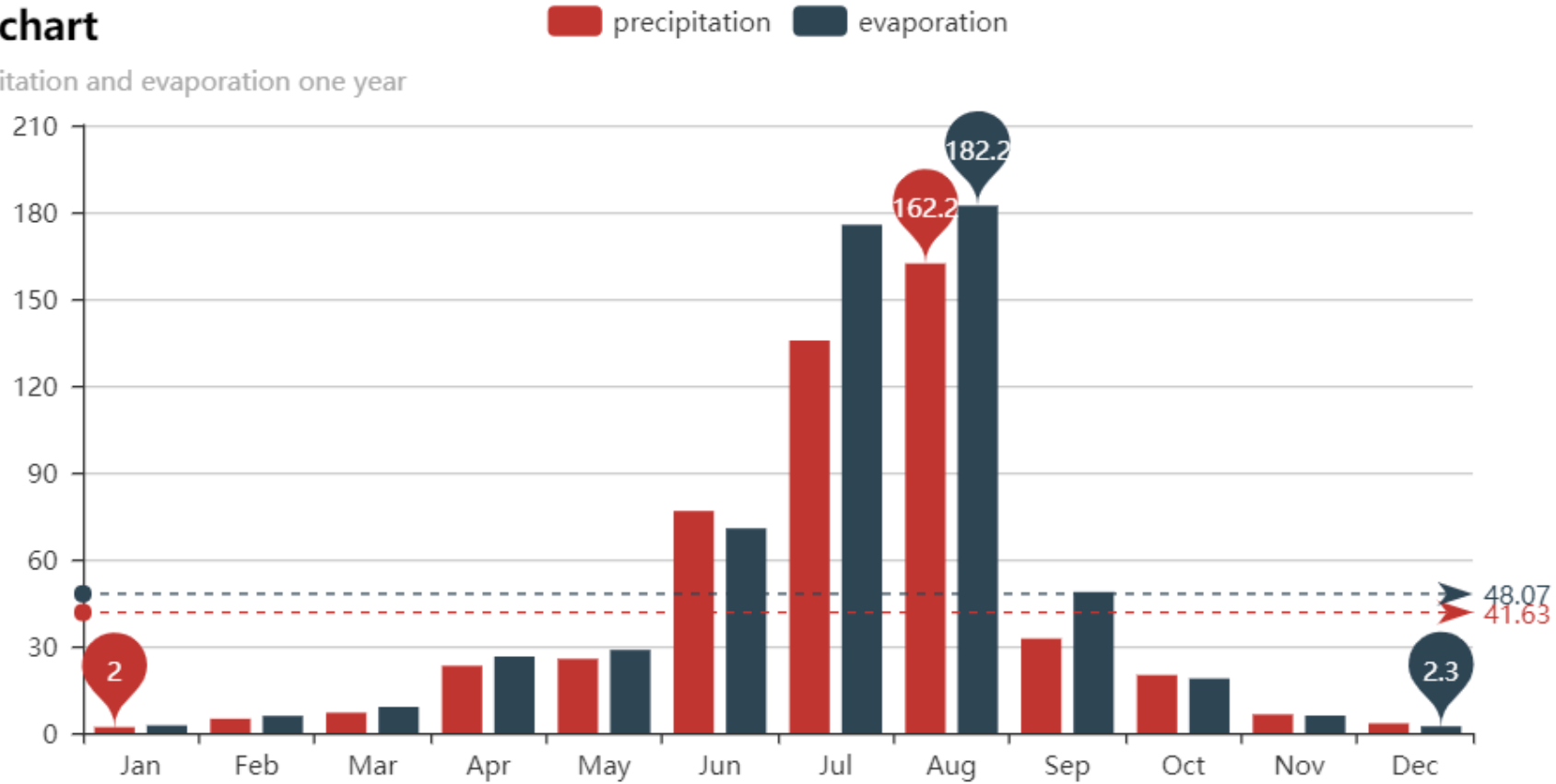


`sns.pairplot(df, kind='reg')`

pyecharts

Bar chart

precipitation and evaporation one year



<http://pyecharts.org/>

pyecharts

```
from pyecharts import Bar
```

```
attr = ["Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug", "Sep", "Oct", "Nov", "Dec"]
```

```
v1 = [2.0, 4.9, 7.0, 23.2, 25.6, 76.7, 135.6, 162.2, 32.6, 20.0, 6.4, 3.3]
```

```
v2 = [2.6, 5.9, 9.0, 26.4, 28.7, 70.7, 175.6, 182.2, 48.7, 18.8, 6.0, 2.3]
```

```
bar = Bar("Bar chart", "precipitation and evaporation one year")
```

```
bar.add("precipitation", attr, v1, mark_line=["average"], mark_point=["max", "min"])
```

```
bar.add("evaporation", attr, v2, mark_line=["average"], mark_point=["max", "min"])
```

```
bar.render('render.html')
```

M2 小结

00 SciPy生态系统

01 NumPy与科学计算

02 pandas与数据分析

03 Matplotlib与可视化