

程式設計概論 Programming 101 —matplotlib繪製圖表

授課老師：邱淑怡

DATE: 5/24/2021



More on pandas

Dataframe 多重條件的篩選資料

```
import pandas as pd
per_df = pd.DataFrame() # 產生一個空的dataframe
col = ['class','name','Birthday','salary','height','weight']
data = [['class0', 'John', '1993-10-01',36000, 177, 76], ['class0', 'Bob', '1992-10-02',52000, 173, 68], ['class1', 'Helen', '1990-10-01',43000, 167, 55], ['class2', 'Alice', '1983-10-03', 27000, 169, 56], ['class1', 'Justin', '1991-10-02',22000, 180, 78], ['class0', 'David', '2001-10-03', 15000, 170, 69]]
per_df = pd.DataFrame(data,columns=col, index=['1','2','3','4','5','6'])
display(per_df)
```

篩選身高及體重兩個條件

```
# height >=170 & weight>=60
```

```
display(per_df[(per_df.height>=170) & (per_df.weight >=60)])
```

```
# or
```

```
display(per_df[(per_df['height']>=170) &(per_df['weight']>=60 )])
```

Matplotlib

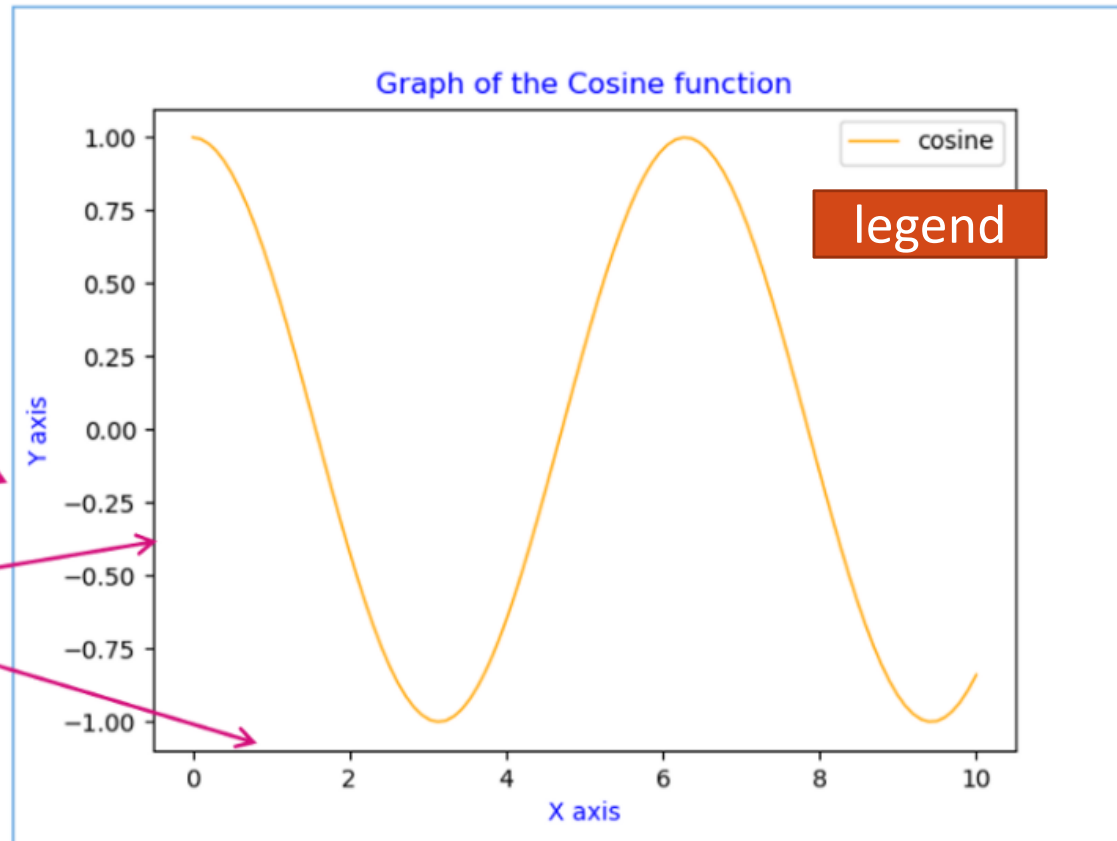
- Matplotlib (<https://matplotlib.org/>) 完全仿照Matlib 函數形式的繪圖介面
- Matplotlib最常運用的模組pyplot
- Matplotlib figure 架構
- Workflow
- 折線圖 (line plot): 使用函式 `plt.plot(x,y)`
- 長條圖 (bar plot): 使用函式 `plt.bar(x,y)`
- 直方圖 (histogram): 使用函式 `plt.hist(scores, bins, ...)`
- 散佈圖 (scatter plot): 使用函式 `plt.scatter(x,y)`
- 箱形圖 (boxplot): 使用函式 `dataframe.boxplot()`
- 圓形圖 (pie plot): 使用函式 `plt.pie(hours...)`

Figure

Figure

Axes

Axis



legend

Workflow

用matplotlib串列建新圖，基本步驟：

1. Step1: Prepare Data
2. Step2: Create Plot
3. Step3: Plot
4. Step4: Customize Plot
5. Step5: Save Plot
6. Step6: Show Plot

Workflow對應程式

折線圖 (line plot)

用matplotlib串列建新圖：

1. Step1: Prepare Data
2. Step2: Create Plot
3. Step3: Plot
4. Step4: Customize Plot
5. Step5: Save Plot
6. Step6: Show Plot



```
import matplotlib.pyplot as plt
x = [1, 2, 3, 4, 5] # Step 1
y = [5, 10, 20, 35, 45] # Step 1
plt.figure(figsize=(4,6)) # Step 2
plt.plot(x, y, color='blue', linewidth=2, marker='o') # Step 3
plt.xlabel("x value") # Step 4
plt.savefig('wk1.png') # Step 5
plt.show() # Step 6
```

折線圖 (line plot)

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

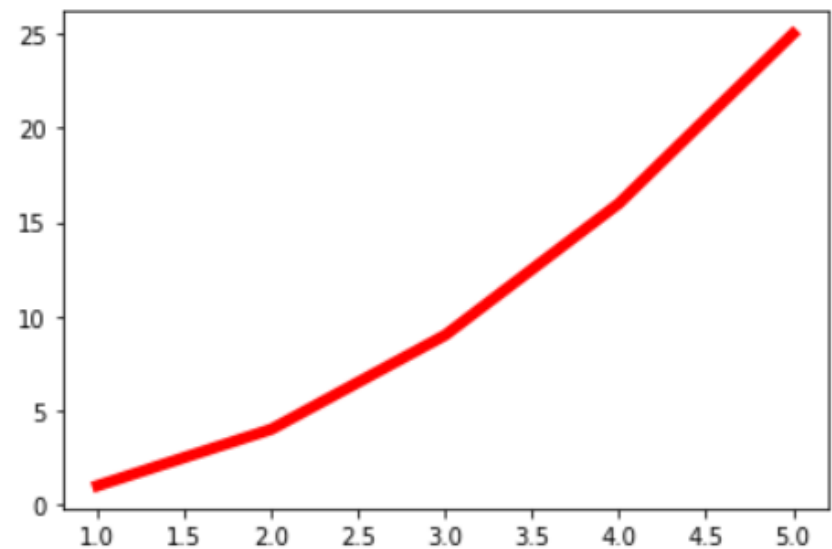
```
x=[1,2,3,4,5]
```

```
y=[1,4,9,16,25]
```

```
plt.plot(x,y, color="red",linewidth=5.0)
```

```
plt.show()
```

```
▶ plt.plot(x,y, color="red", linewidth=5.0)  
plt.show()
```

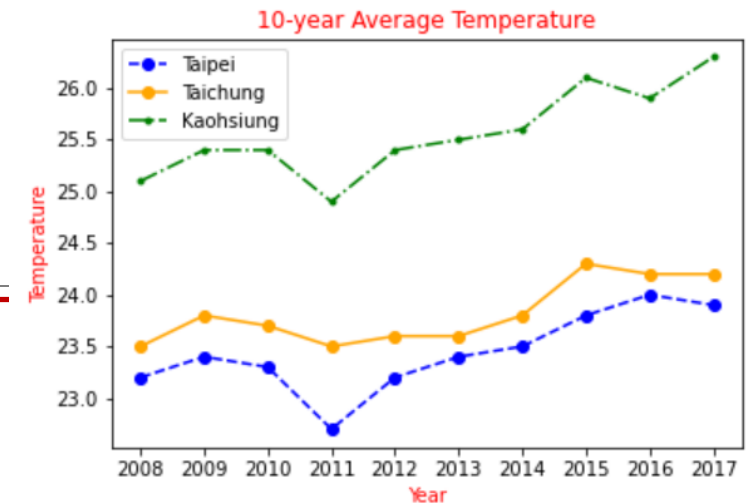


多條折線圖_程式例子

```
import matplotlib.pyplot as plt
Taipei_HTemp = [16.1, 16.5, 18.5, 21.9, 25.2, 27.7, 29.6, 29.2, 27.4, 24.5, 21.5, 17.9]
Taipei_LTemp = [13.9, 14.2, 15.8, 19.0, 22.3, 24.6, 26.3, 26.1, 24.8, 22.3, 19.3, 15.6]
month = range(1, 13)
plt.plot(month, Taipei_HTemp, 'red')
# 讀出month及Taipei_HTemp的資料，產生紅色線
plt.plot(month, Taipei_LTemp, 'blue')
# 讀出month及Taipei_LTemp的資料，產生藍色線
plt.xlabel('Month') # x軸的Label
plt.title('Taipei High and Low Temperature') # 此圖的title
plt.show()
```

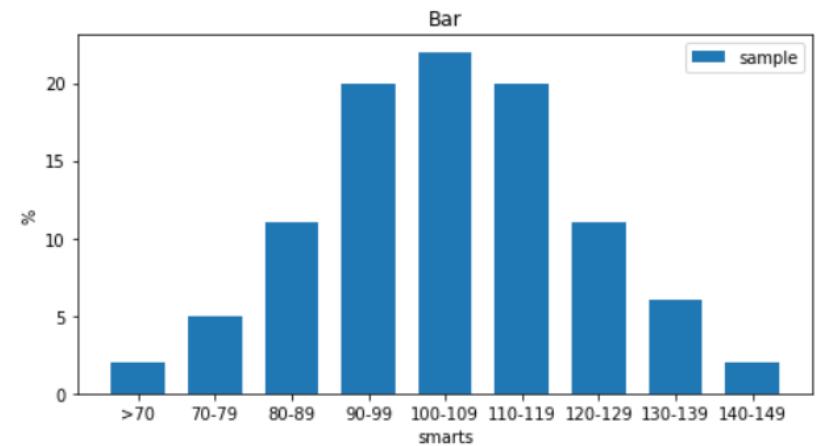
Add legend and save graph

```
from matplotlib import pyplot as plt
Taipei_temp = [23.2, 23.4, 23.3, 22.7, 23.2, 23.4, 23.5, 23.8, 24, 23.9]
Taichung_temp = [23.5, 23.8, 23.7, 23.5, 23.6, 23.6, 23.8, 24.3, 24.2, 24.2]
Kaohsiung_temp = [25.1, 25.4, 25.4, 24.9, 25.4, 25.5, 25.6, 26.1, 25.9, 26.3]
year = range(2008, 2018)
plt.plot(year, Taipei_temp, color = 'blue', marker='o', linestyle = '--', label='Taipei')
plt.plot(year, Taichung_temp, color = 'orange', marker='o', linestyle = '-', label='Taichung')
plt.plot(year, Kaohsiung_temp, color = 'green', marker='.', linestyle = '-.', label='Kaohsiung')
plt.legend(loc = 'upper left')
plt.xlabel('Year', color = 'red')
plt.ylabel('Temperature', color = 'red')
plt.title('10-year Average Temperature', color = 'red')
plt.xticks([2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017])
plt.savefig('wn2.png')
plt.show()
```



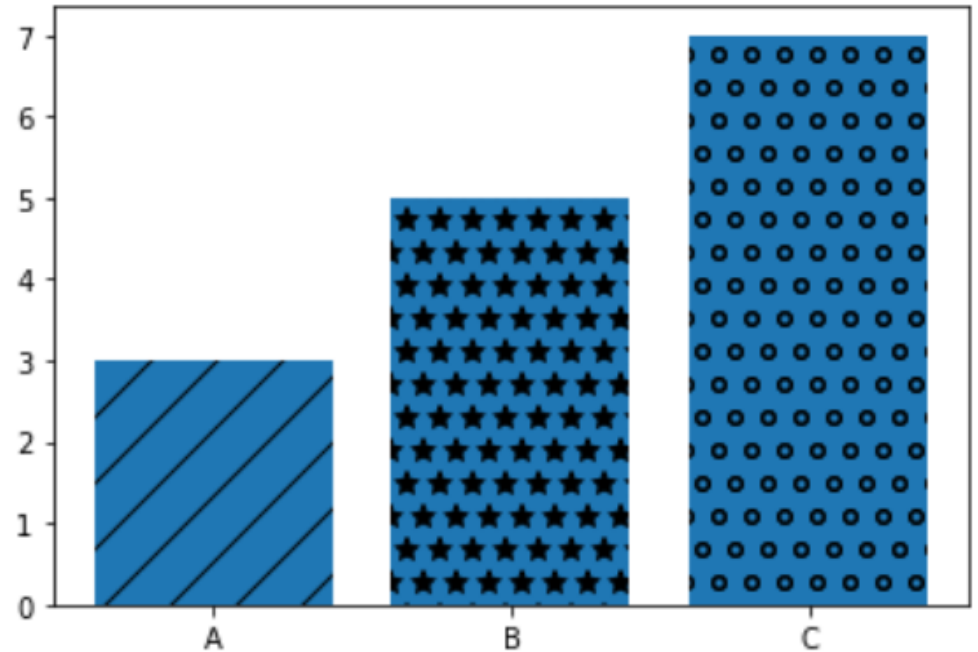
長條圖

```
import matplotlib.pyplot as plt
x=[70,80,90,100,110,120,130,140,150]
y=[2,5,11,20,22,20,11,6,2]
tit=['>70','70-79','80-89','90-99','100-109','110-119','120-129','130-139','140-149']
plt.figure(figsize=(8,4)) #8 inch, 4 inch
plt.bar(x,y, width=7, tick_label=tit,label='sample')
plt.legend() # 放置圖例
plt.xlabel('smarts')
plt.ylabel('%') # set label of y asix
plt.title('Bar') # set title
plt.show()
```



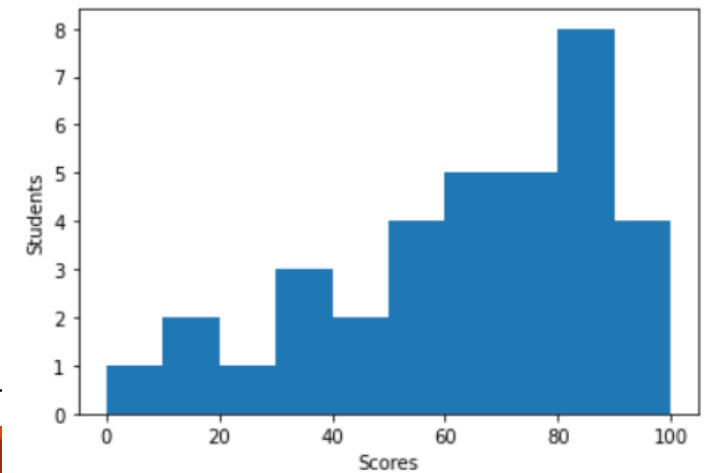
長條圖_改變圖案

```
x1=['A','B','C']  
y1=[3,5,7]  
bar1 = plt.bar(x1,y1)  
patterns=['/','*','o']  
for bar in bar1:  
    bar.set_hatch(patterns.pop(0))  
#bar1[0].set_hatch('/')  
#bar1[1].set_hatch('*')  
plt.show()
```



直方圖

```
import matplotlib.pyplot as plt
scores = [10, 15, 80, 22, 93, 55, 88, 62, 45, 75, 81, 34, 99, 84, 85, 55,
58, 63, 68, 82, 84, \
77, 69, 90, 100, 75, 65, 54, 34, 38, 48, 88, 71, 72, 5]
bins = [0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100]
plt.hist(scores, bins, histtype = "bar")
plt.xlabel("Scores")
plt.ylabel("Students")
plt.show()
```



圓形圖

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

```
activities = ["work", "sleep", "Internet", "others"]
```

```
hours = [8, 7, 2, 7]
```

```
colors = ["lightgreen", "lightblue", "yellow", "pink"]
```

```
plt.pie(hours, labels = activities, colors = colors, shadow = True, explode = (0, 0, 0.1, 0),  
autopct = "%1.1f%%") # 將數值百分比並留到小數點一位; 設定分隔的區塊位置
```

```
plt.axis("equal") # 使圓餅圖比例相等
```

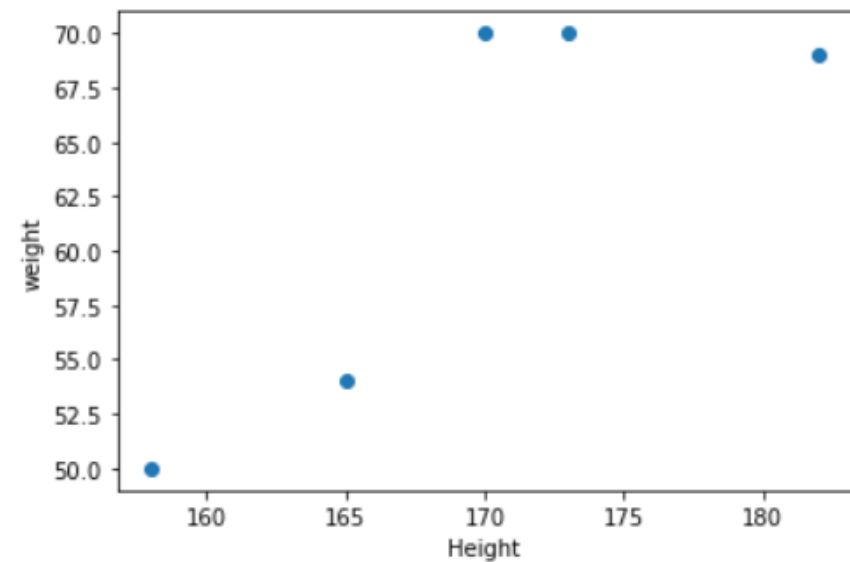
```
plt.title("Pie chart", {"fontsize" : 18})
```

```
plt.show()
```



散佈圖

```
x=[170,165,158,182,173]
y=[70,54,50,69,70]
plt.scatter(x,y)
plt.xlabel("Height")
plt.ylabel("weight")
plt.show()
```



範例說明

建立DataFrame

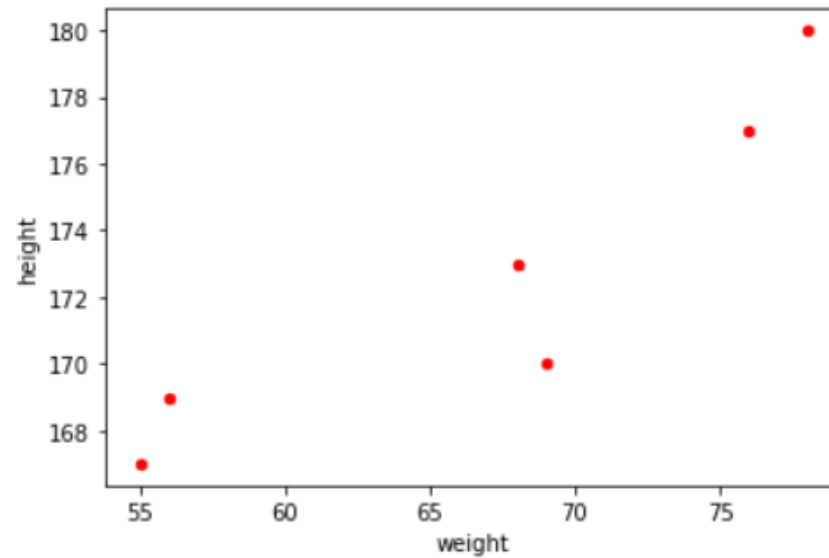
```
import matplotlib.pyplot as plt  
import pandas as pd
```

```
per_df = pd.DataFrame() # 產生一個空的dataframe  
col = ['class','name','Birthdate','salary','height','weight']  
data = [['class0', 'John', '1993-10-01', 36000, 177, 76], ['class0', 'Bob', '1992-10-02', 52000, 173, 68], ['class1', 'Helen', '1990-10-01', 43000, 167, 55], ['class2', 'Alice', '1983-10-03', 27000, 169, 56], ['class1', 'Justin', '1991-10-02', 22000, 180, 78], ['class0', 'David', '2001-10-03', 15000, 170, 69]]  
per_df = pd.DataFrame(data, columns=col, index=['1','2','3','4','5','6'])  
print(per_df)
```

Scatter plot

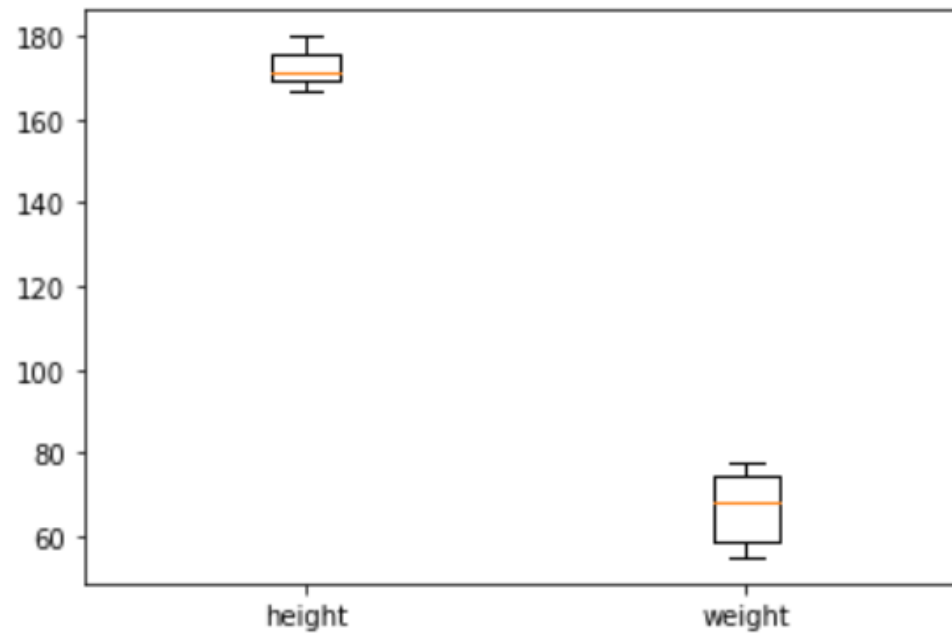
```
per_df.plot(kind='scatter',x='weight',y='height',color='red')  
plt.show()
```

```
per_df.plot(kind='scatter',x='weight',y='height',color='red')  
plt.show()
```



Boxplot

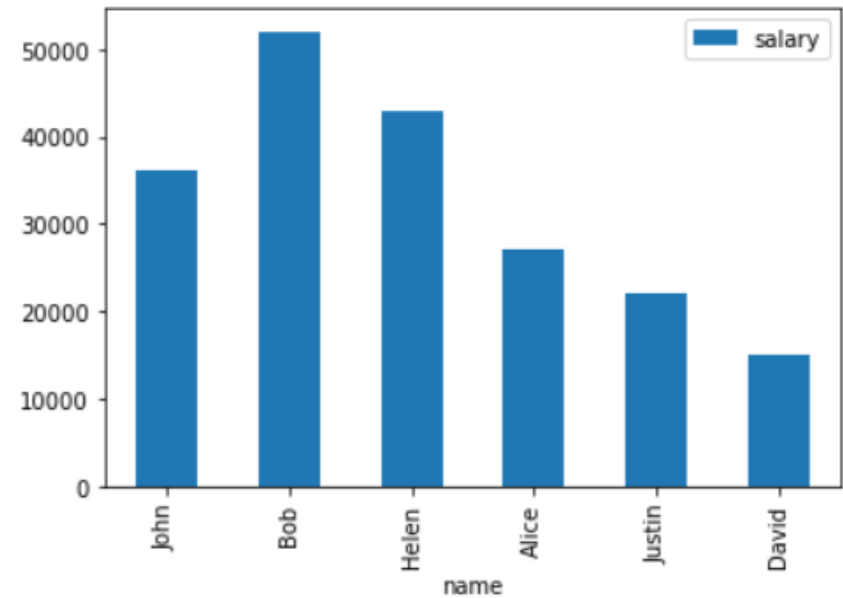
```
plt.boxplot([per_df.height,per_df.weight],labels=['height ','weight'])
```



Bar plot

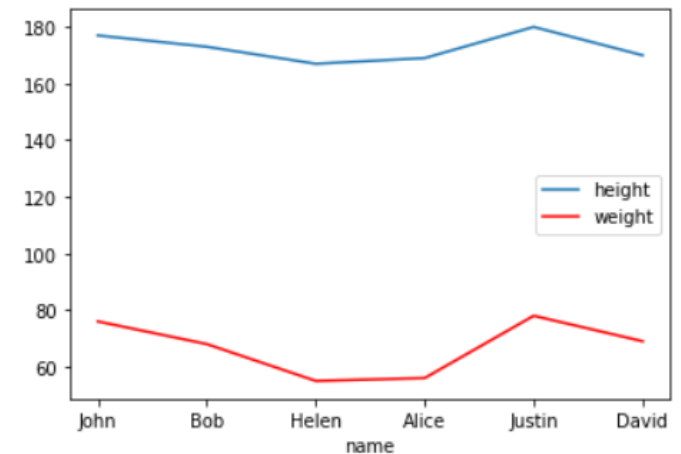
a simple line plot

```
per_df.plot(kind='bar',x='name',y='salary')
```



Line plot, multiple columns

```
# gca stands for 'get current axis'  
ax = plt.gca()  
per_df.plot(kind='line',x='name',y='height',ax=ax)  
per_df.plot(kind='line',x='name',y='weight', color='red', ax=ax)  
plt.show()
```



Save plot to file

plt.show() 改成 plt.savefig('outputfile.png')

```
per_df.plot(kind='bar',x='name',y='height')  
# the plot gets saved to 'output.png'  
plt.savefig('output.png')  
plt.show()
```

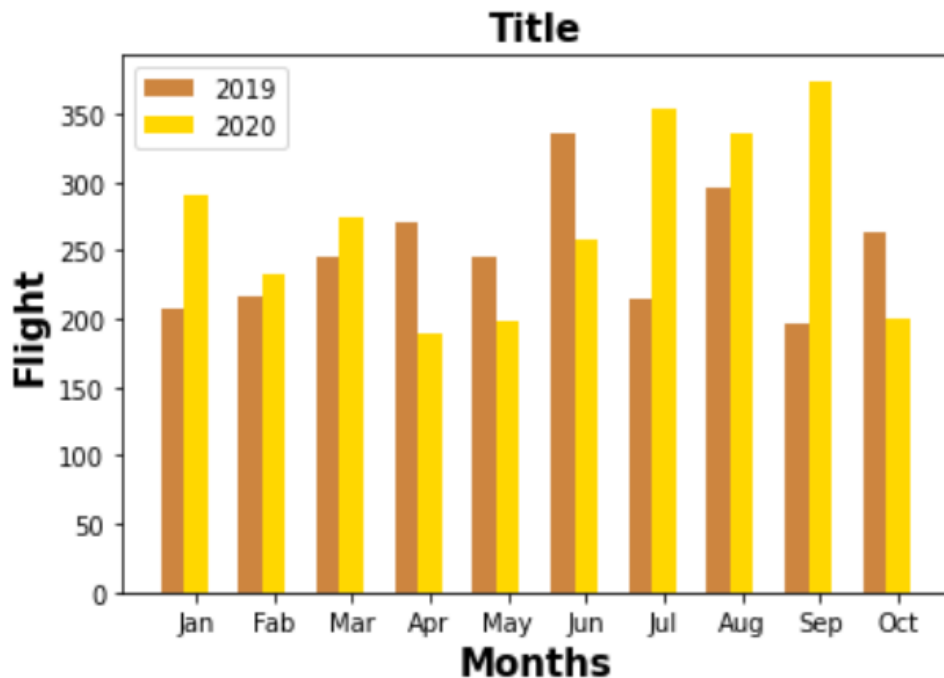
多個圖形的呈現

```
names = ['group_a', 'group_b', 'group_c']  
values = [1, 10, 100]  
  
plt.figure(figsize=(9, 3))  
  
plt.subplot(131)  
plt.bar(names, values)  
plt.subplot(132)  
plt.scatter(names, values)  
plt.subplot(133)  
plt.plot(names, values)  
plt.suptitle('Categorical Plotting')  
plt.show()
```

匯入csv file 後產生折線圖

```
import pandas as pd
from matplotlib import pyplot as plt
plt.figure(figsize=(8,5))
gas_df = pd.read_csv("gas_prices.csv")
for country in gas_df:
    print(country)
    if country != 'Year' and country!='group':
        plt.plot(gas_df.Year, gas_df[country], marker='.',label=country)
plt.legend()
x_list= list(range(1990,2009,2))
plt.xticks(x_list)
plt.show()
```

Grouped barplot by column



```
import numpy as np
import matplotlib.pyplot as plt
# set width of bar
barWidth = 0.3
# set height of bar
bars1=[208,216,246,270,246,336,214,296,196,264]
bars2=[290,232,274,190,198,258,354,336,374,200]
# Set position of bar on X axis
r1 = np.arange(len(bars1))
r2 = [x + barWidth for x in r1]
# Make the plot
plt.bar(r1, bars1, color='peru', width=barWidth, label='2019')
plt.bar(r2, bars2, color='gold', width=barWidth, label='2020')
# Add xticks on the middle of the group bars
plt.xlabel('Months',fontweight='bold',fontsize = 15)
plt.ylabel('Flight',fontweight='bold',fontsize = 15)
plt.xticks([r + barWidth for r in range(len(bars1))],
['Jan','Feb','Mar','Apr', 'May','Jun','Jul','Aug','Sep','Oct'])
# Create legend & Show graphic
plt.title(' Title',fontsize = 15,fontweight='bold')
plt.legend()
plt.show()
```

Grouped barplot by column

```
import numpy as np
```

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

```
# set width of bar
```

```
barWidth = 0.25
```

```
# set height of bar
```

```
bars1 = [12, 30, 1, 8, 22]
```

```
bars2 = [28, 6, 16, 5, 10]
```

```
bars3 = [29, 3, 24, 25, 17]
```

```
# Set position of bar on X axis
```

```
r1 = np.arange(len(bars1))
```

```
r2 = [x + barWidth for x in r1]
```

```
r3 = [x + barWidth for x in r2]
```

```
# Make the plot
```

```
plt.bar(r1, bars1, color='gray', width=barWidth, edgecolor='white',  
label='var1')
```

```
plt.bar(r2, bars2, color='darkorange', width=barWidth,  
edgecolor='white', label='var2')
```

```
plt.bar(r3, bars3, color='navy', width=barWidth, edgecolor='white',  
label='var3')
```

```
# Add xticks on the middle of the group bars
```

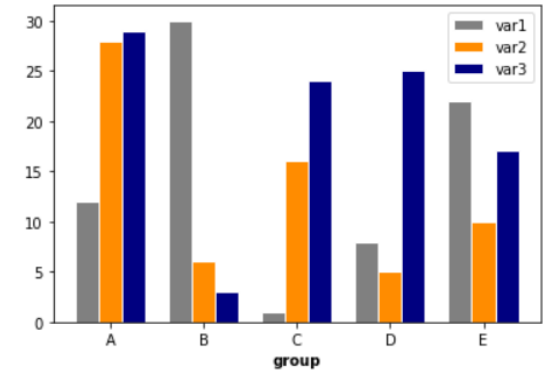
```
plt.xlabel('group', fontweight='bold')
```

```
plt.xticks([r + barWidth for r in range(len(bars1))], ['A', 'B', 'C', 'D', 'E'])
```

```
# Create legend & Show graphic
```

```
plt.legend()
```

```
plt.show()
```



Ref: set color

Google search: color picker

颜色可以参考下面:

	black		bisque		lightgreen		slategrey
	k		darkorange		forestgreen		lightsteelblue
	dimgray		burlywood		limegreen		cornflowerblue
	dimgrey		antiquewhite		darkgreen		royalblue
	grey		tan		green		ghostwhite
	gray		navajowhite		g		lavender
	darkgrey		blanchedalmond		lime		midnightblue
	darkgray		papayawhip		seagreen		navy
	silver		moccasin		mediumseagreen		darkblue
	lightgray		orange		springgreen		mediumblue
	lightgrey		wheat		mintcream		blue
	gainsboro		oldlace		mediumspringgreen		b
	whitesmoke		floralwhite		mediumaquamarine		slateblue
	white		darkgoldenrod		aquamarine		darkslateblue
	w		goldenrod		turquoise		mediumslateblue
	snow		cornsilk		lightseagreen		mediumpurple
	rosybrown		lemonchiffon		mediumturquoise		blueviolet
	lightcoral		khaki		azure		indigo
	indianred		palegoldenrod		lightcyan		darkorchid
	brown		darkkhaki		paleturquoise		darkviolet
	firebrick		ivory		darkslategray		mediumorchid
	maroon		beige		darkslategrey		thistle
	darkred		lightyellow		teal		plum
	red		lightgoldenrodyellow		darkcyan		violet
	r		olive		c		purple
	mistyrose		y		aqua		darkmagenta
	salmon		yellow		darkturquoise		m
	tomato		olivedrab		cadetblue		fuchsia
	darksalmon		yellowgreen		powderblue		magenta
	coral		darkolivegreen		lightblue		orchid
	orangered		greenyellow		deepskyblue		mediumvioletred
	lightsalmon		chartreuse		skyblue		deeppink
	sienna		lawngreen		lightskyblue		hotpink
	seashell		sage		steelblue		lavenderblush
	chocolate		lightsage		aliceblue		palevioletred
	saddlebrown		darksage		dodgerblue		crimson
	sandybrown		honeydew		lightslategrey		pink
	peachpuff		darkseagreen		lightslategray		lightpink
	peru		palegreen		slategray		
	linen						

專題討論進度
畫流程圖
整理資料、計算資料
繪製圖形
