

# 作品集

プログラム言語：python

撰寫人：翁雋閑

## 作品集

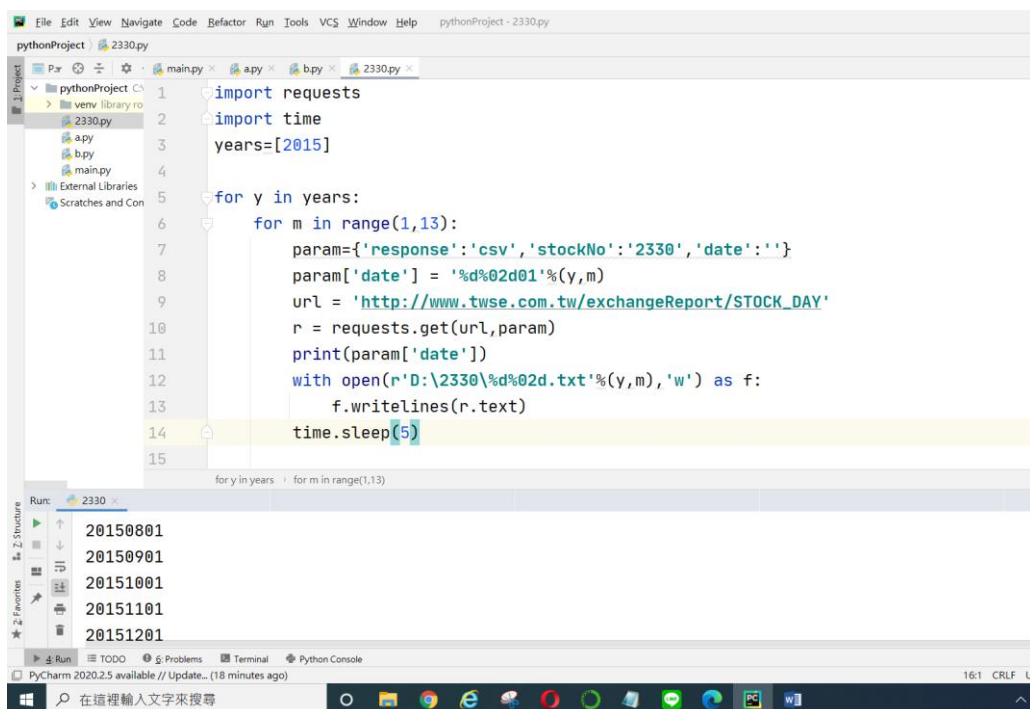
### 主題 1 : web crawler

目的：プログラム言語を使用して web 情報を収集すること

開発環境：Anaconda / PyCharm

開発例：ウェブサイト - 台湾証券取引所に株価を収集すること

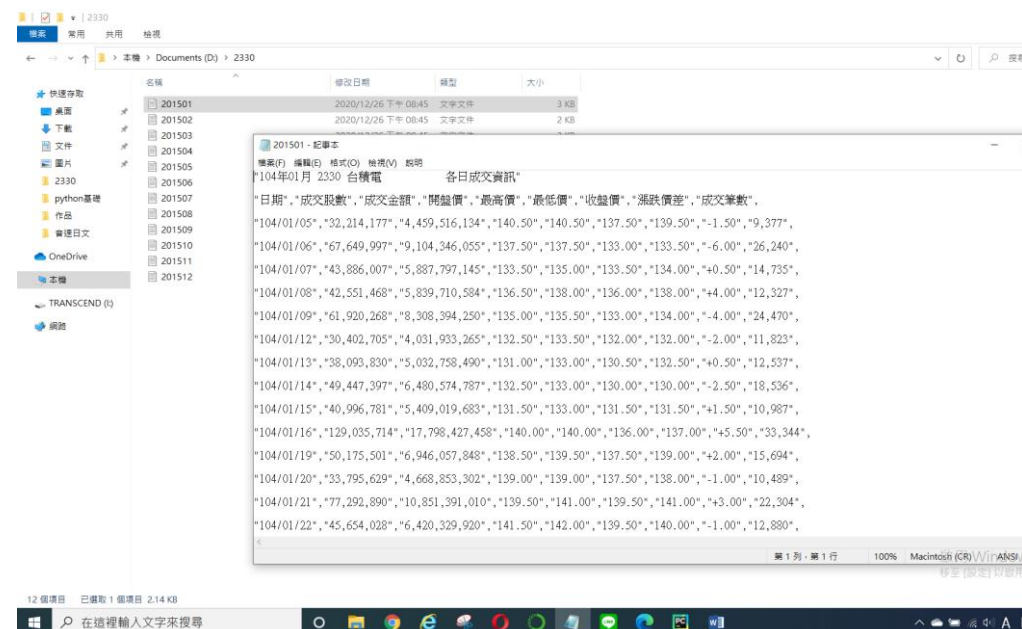
(会社名：TSMC、証券コード：2330、2015 年の株価)



```
1 import requests
2 import time
3 years=[2015]
4
5 for y in years:
6     for m in range(1,13):
7         param={'response':'csv','stockNo':'2330','date':''}
8         param['date'] = '%d%02d%01'%(y,m)
9         url = 'http://www.twse.com.tw/exchangeReport/STOCK_DAY'
10        r = requests.get(url,param)
11        print(param['date'])
12        with open(r'D:\2330\%d%02d.txt'%(y,m),'w') as f:
13            f.writelines(r.text)
14        time.sleep(5)
15
```

Run: 2330

20150801  
20150901  
20151001  
20151101  
20151201



名前	修改日期	類型	大小
201501	2020/12/26 下午 06:45	文字文件	3 KB
201502	2020/12/26 下午 06:45	文字文件	2 KB
201503			
201504			
201505			
201506			
201507			
201508			
201509			
201510			
201511			
201512			

201501 - 記事本

檔案(F) 編輯(E) 格式(O) 檢視(V) 說明

104年01月 2330 台股電 各日成交資訊

"日期","成交股數","成交金額","開盤價","最高價","最低價","收盤價","漲跌價差","成交筆數",

"104/01/05","32,214,177","4,459,516,134","140.50","140.50","137.50","139.50",-1.50,"9,377",

"104/01/06","67,649,997","9,104,346,055","137.50","137.50","133.00","133.50",-6.00,"26,240",

"104/01/07","43,886,007","5,887,797,145","133.50","135.00","133.00","134.00",+0.50,"14,735",

"104/01/08","42,551,468","5,839,710,584","136.50","138.00","136.00","138.00",+4.00,"12,327",

"104/01/09","61,920,268","8,308,394,250","135.00","135.50","133.00","134.00",-4.00,"24,470",

"104/01/12","30,402,705","4,031,933,265","132.50","133.50","132.00","132.00",-2.00,"11,823",

"104/01/13","38,093,830","5,032,758,490","131.00","133.00","130.50","132.50",+0.50,"12,537",

"104/01/14","49,447,397","6,480,574,787","132.50","133.00","130.00","130.00",-2.50,"18,536",

"104/01/15","40,996,781","5,409,019,683","131.50","133.00","131.50","131.50",+1.50,"10,987",

"104/01/16","129,035,714","17,798,427,458","140.00","140.00","136.00","137.00",+5.50,"33,344",

"104/01/19","50,175,501","6,946,057,848","138.50","139.50","137.50","139.00",+2.00,"15,694",

"104/01/20","33,795,629","4,668,853,302","139.00","139.00","137.50","138.00",-1.00,"10,489",

"104/01/21","77,292,890","10,851,391,010","139.50","141.00","139.50","141.00",+3.00,"22,304",

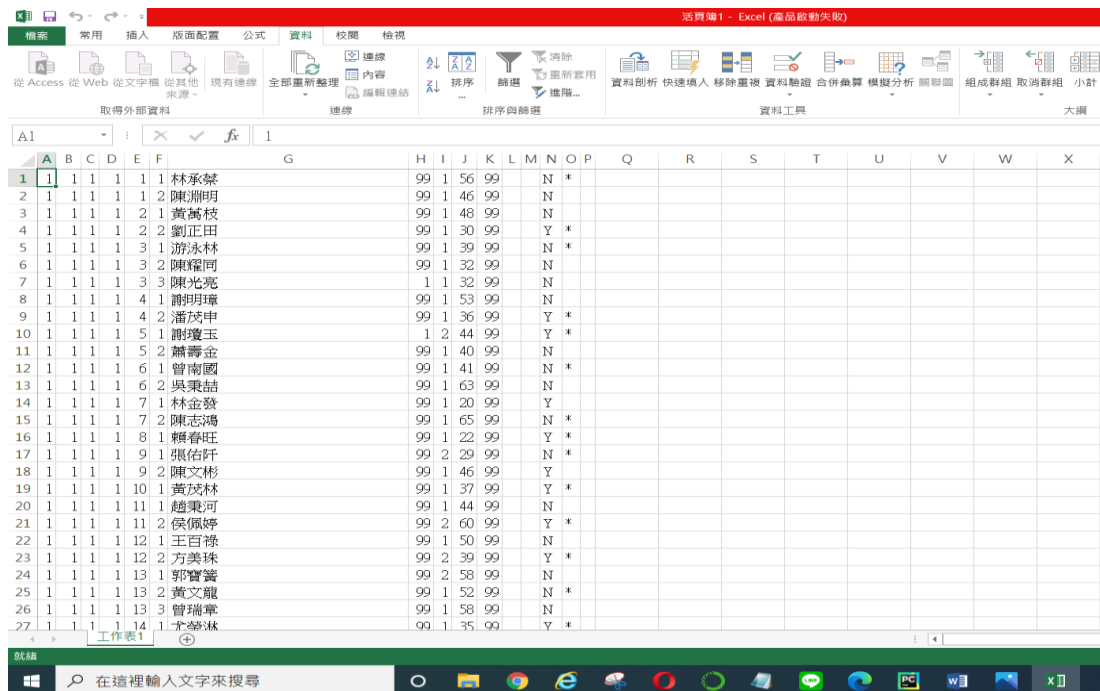
"104/01/22","45,654,028","6,420,329,920","141.50","142.00","139.50","140.00",-1.00,"12,880",

## 主題 2：データの視覚化

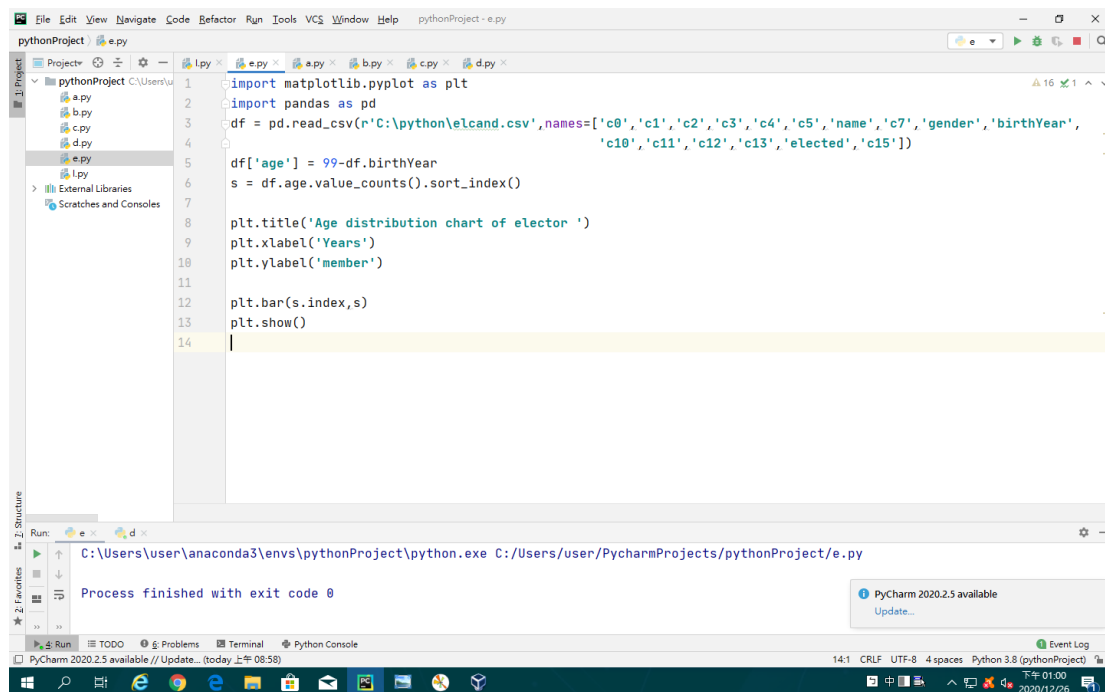
目的：収集したデータを使用して視覚化すること

開発環境：Anaconda / PyCharm

開発例：選挙人データ（elcand.csv）を読んで年齢に対して人数を統計して視覚化すること

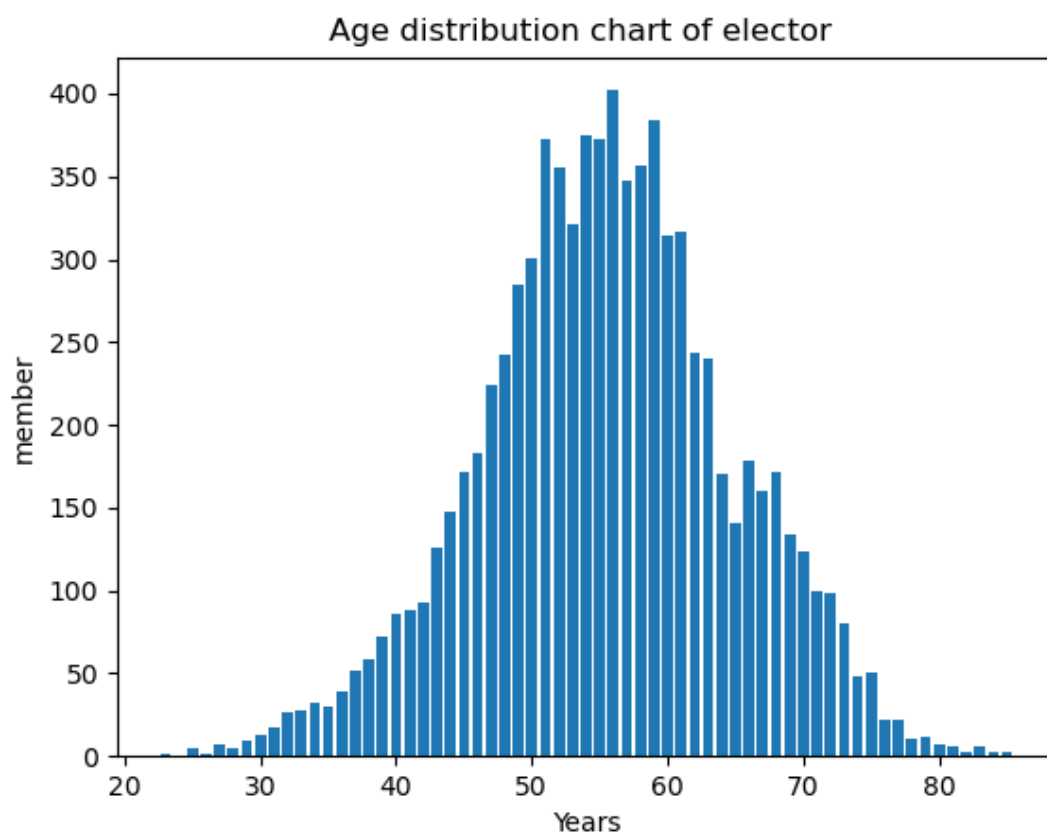


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	1	1	1	1	1	1	林承榮	99	1	56	99			N	*									
2	1	1	1	1	1	2	陳淵明	99	1	46	99			N										
3	1	1	1	1	2	1	黃萬枝	99	1	48	99			N										
4	1	1	1	1	2	2	劉正田	99	1	30	99			Y	*									
5	1	1	1	1	3	1	游泳林	99	1	39	99			N	*									
6	1	1	1	1	3	2	陳耀同	99	1	32	99			N										
7	1	1	1	1	3	3	陳光亮	1	1	32	99			N										
8	1	1	1	1	4	1	謝明璋	99	1	53	99			N										
9	1	1	1	1	4	2	潘茂申	99	1	36	99			Y	*									
10	1	1	1	1	5	1	謝瓊玉	1	2	44	99			Y	*									
11	1	1	1	1	5	2	蕭壽金	99	1	40	99			N										
12	1	1	1	1	6	1	曾南園	99	1	41	99			N	*									
13	1	1	1	1	6	2	吳秉結	99	1	63	99			N										
14	1	1	1	1	7	1	林金發	99	1	20	99			Y										
15	1	1	1	1	7	2	陳志鴻	99	1	65	99			N	*									
16	1	1	1	1	8	1	賴春旺	99	1	22	99			Y	*									
17	1	1	1	1	9	1	張佑軒	99	2	29	99			N	*									
18	1	1	1	1	9	2	陳文彬	99	1	46	99			Y										
19	1	1	1	1	10	1	黃茂林	99	1	37	99			Y	*									
20	1	1	1	1	11	1	趙秉河	99	1	44	99			N										
21	1	1	1	1	11	2	侯佩婷	99	2	60	99			Y	*									
22	1	1	1	1	12	1	王百祿	99	1	50	99			N										
23	1	1	1	1	12	2	方美珠	99	2	39	99			Y	*									
24	1	1	1	1	13	1	郭寶義	99	2	58	99			N										
25	1	1	1	1	13	2	黃文龍	99	1	52	99			N	*									
26	1	1	1	1	13	3	曾瑞寧	99	1	58	99			N										
27	1	1	1	1	14	1	李榮林	99	1	35	99			Y	*									



```
pythonProject - e.py
1 import matplotlib.pyplot as plt
2 import pandas as pd
3 df = pd.read_csv(r'C:\python\elcand.csv', names=['c0', 'c1', 'c2', 'c3', 'c4', 'c5', 'name', 'c7', 'gender', 'birthYear',
4          'c10', 'c11', 'c12', 'c13', 'elected', 'c15'])
5 df['age'] = 99 - df.birthYear
6 s = df.age.value_counts().sort_index()
7
8 plt.title('Age distribution chart of elector ')
9 plt.xlabel('Years')
10 plt.ylabel('member')
11
12 plt.bar(s.index, s)
13 plt.show()
14
```

Run: C:\Users\user\anaconda3\envs\pythonProject\python.exe C:/Users/user/PycharmProjects/pythonProject/e.py  
Process finished with exit code 0



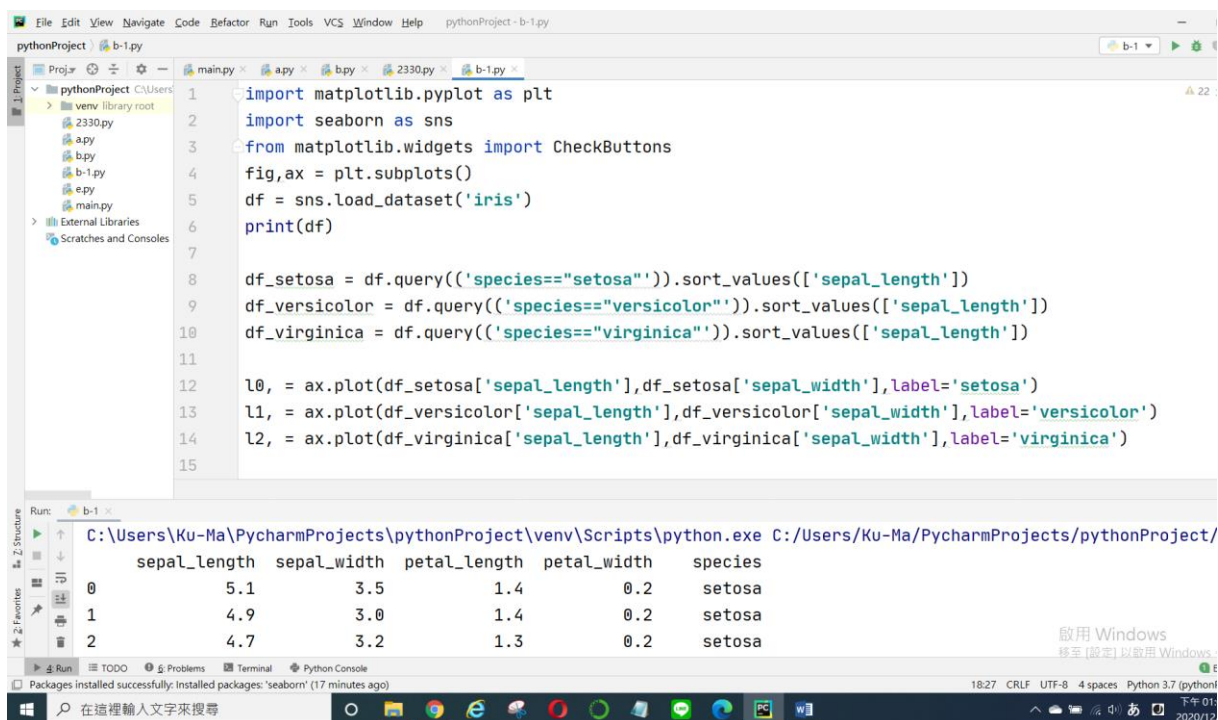
### 主題 3：データの視覚化と使用者の交流

目的：データを視覚化して使用者と交流しできるのこと。

開発環境：Anaconda / PyCharm

開発例：1.データ - 'iris' を読んで花卉の寸法を視覚化すること。

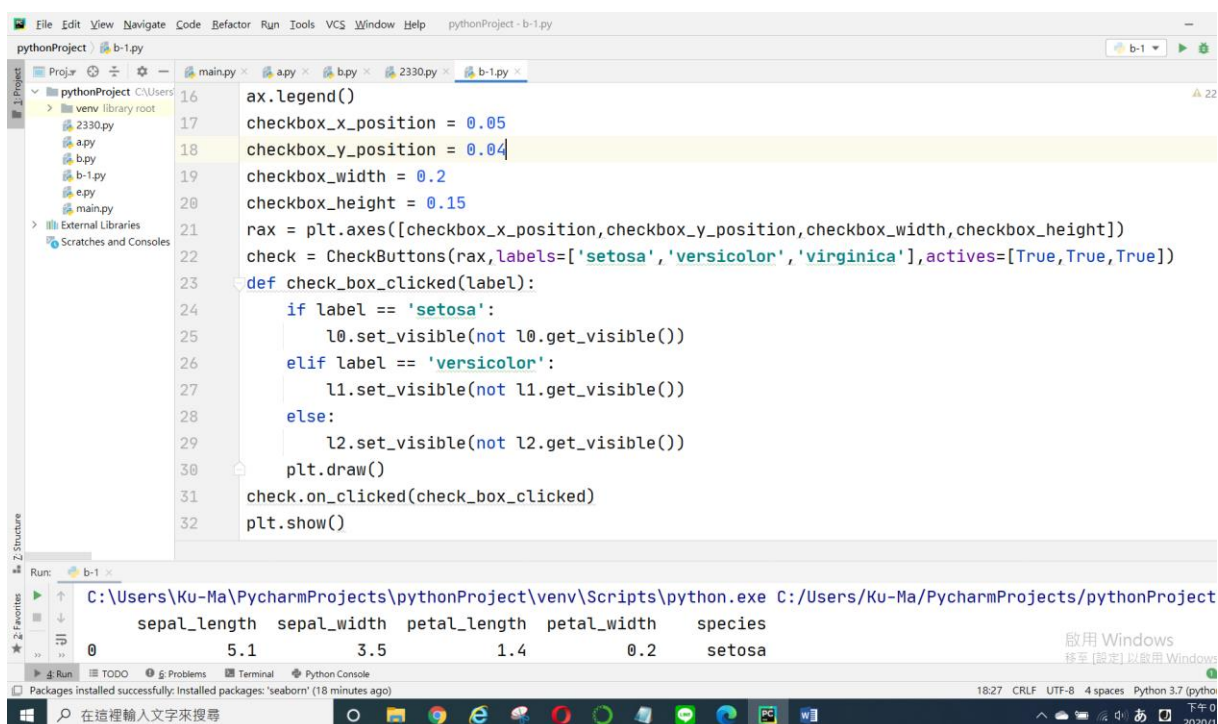
2.使用者が花の品種に対して選べるのこと。



```
1 import matplotlib.pyplot as plt
2 import seaborn as sns
3 from matplotlib.widgets import CheckButtons
4 fig, ax = plt.subplots()
5 df = sns.load_dataset('iris')
6 print(df)
7
8 df_setosa = df.query(('species=="setosa"')).sort_values(['sepal_length'])
9 df_versicolor = df.query(('species=="versicolor"')).sort_values(['sepal_length'])
10 df_virginica = df.query(('species=="virginica"')).sort_values(['sepal_length'])
11
12 l0, = ax.plot(df_setosa['sepal_length'], df_setosa['sepal_width'], label='setosa')
13 l1, = ax.plot(df_versicolor['sepal_length'], df_versicolor['sepal_width'], label='versicolor')
14 l2, = ax.plot(df_virginica['sepal_length'], df_virginica['sepal_width'], label='virginica')
15
```

Run: b-1

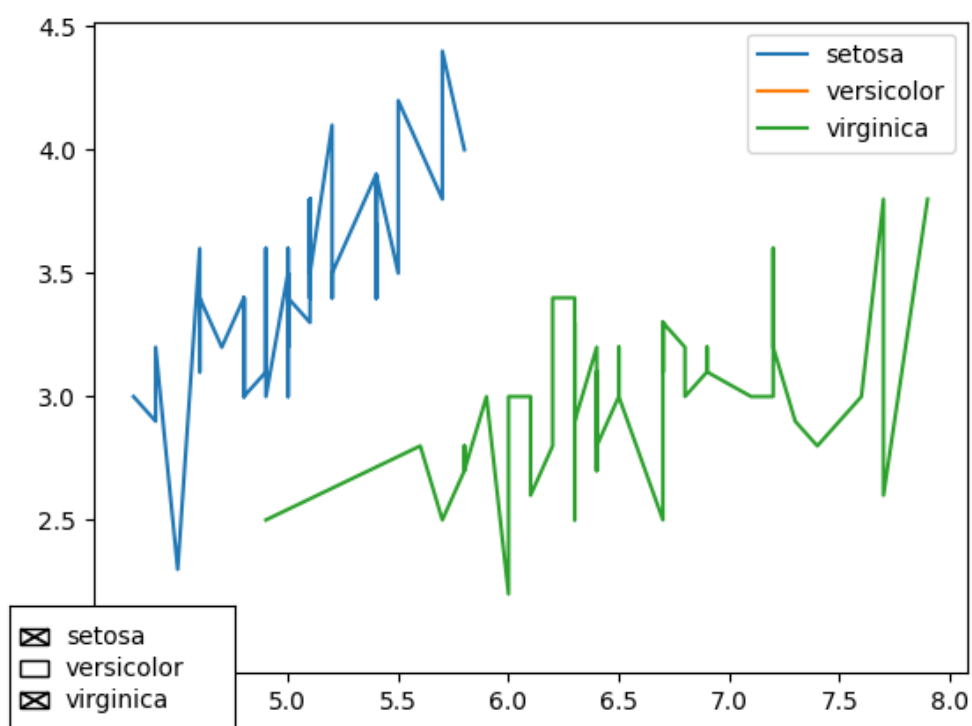
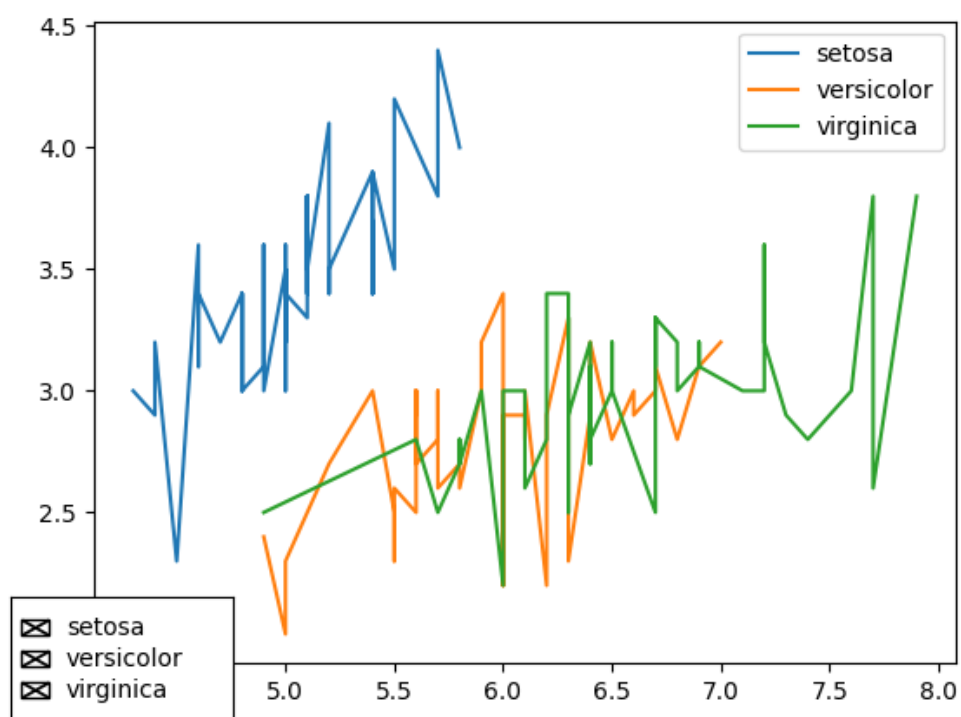
	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa



```
16 ax.legend()
17 checkbox_x_position = 0.05
18 checkbox_y_position = 0.04
19 checkbox_width = 0.2
20 checkbox_height = 0.15
21 rax = plt.axes([checkboxbox_x_position, checkbox_y_position, checkbox_width, checkbox_height])
22 check = CheckButtons(rax, labels=['setosa', 'versicolor', 'virginica'], active=[True, True, True])
23 def check_box_clicked(label):
24     if label == 'setosa':
25         l0.set_visible(not l0.get_visible())
26     elif label == 'versicolor':
27         l1.set_visible(not l1.get_visible())
28     else:
29         l2.set_visible(not l2.get_visible())
30     plt.draw()
31 check.on_clicked(check_box_clicked)
32 plt.show()
```

Run: b-1

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa



↑ 使用者が花の品種に対して選びできます。

## 主題 4：データの 3D 視覚化と使用者の交流

目的：データを視覚化して使用者と交流しできるのこと。

開発環境：Anaconda / PyCharm

開発例：1.データ - 'iris' を読んで花卉の寸法を 3D 視覚化すること。

```
pythonProject - 3D.py
pythonProject
├── venv
│   ├── 3D.py
│   ├── 2330.py
│   ├── a.py
│   ├── b.py
│   ├── b-1.py
│   ├── e.py
│   └── main.py
└── External Libraries
    └── Scratches and Snippets

1 import seaborn as sns
2 import pandas as pd
3 import matplotlib.pyplot as plt
4 from mpl_toolkits.mplot3d import Axes3D
5
6 df = sns.load_dataset('iris')
7 df_setosa = df.query('species=="setosa"')
8 df_versicolor = df.query('species=="versicolor"')
9 df_virginica = df.query('species=="virginica"')
10
11 fig = plt.figure()
12 ax = fig.add_subplot(1,1,1,projection='3d')
13
14 ax.scatter(df_setosa['sepal_length'],df_setosa['sepal_width'],label='setosa')
15 ax.scatter(df_versicolor['sepal_length'],df_versicolor['sepal_width'],label='versicolor')
16 ax.scatter(df_virginica['sepal_length'],df_virginica['sepal_width'],label='virginica')
17
18 ax.legend()
19 plt.show()
20
```

Run: C:\Users\Ku-Ma\PycharmProjects\pythonProject\venv\Scripts\python.exe C:\Users\Ku-Ma\PycharmProjects\pythonProject\3D.py

7:42 CRLF

