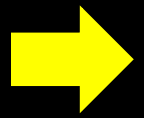
Python 製作樞紐分析表

樞紐分析表 Pivot Table

raw_data

NAME	PRODUCT	AMOUNT	PRICE
BIG	A	10	100
SMALL	B		200
BIG	C	10	
SMALL	A	20	300
BIG	B	20	400
SMALL	C	30	500
BIG	A	30	600
SMALL	B		700
BIG	C	40	
SMALL	A	40	800
BIG	B	50	900
SMALL	C	50	1,000



pt

NAME	PRODUCT	AMOUNT	PRICE
BIG	A	40	350
BIG	B	70	650
BIG	C	50	0
SMALL	A	60	550
SMALL	B	0	450
SMALL	C	80	750

依據 NAME 及 PRODUCT 欄位分組，
計算 AMOUNT 欄位的總和及
PRICE 欄位的平均。

Step 1 : 讀取 Excel 檔

pandas.read_excel( )

Step 2 : 製作樞紐分析表

pandas.pivot_table( )

Step 3 : 還原整數索引欄位

pandas.DataFrame.reset_index( )

Step 4 : DataFrame 寫入 Excel 檔

pandas.DataFrame.to_excel( )

Step 1 : pandas.read_excel( )

```
In [2]: xlsx_path = 'D:\RPA_UiPath\Python x RPA\Pivot Table\Input\SampleData.xlsx'

        raw_data = pd.read_excel( xlsx_path )
        raw_data
```

Out[2]:

	NAME	PRODUCT	AMOUNT	PRICE
0	BIG	A	10.0	100.0
1	SMALL	B	NaN	200.0
2	BIG	C	10.0	NaN
3	SMALL	A	20.0	300.0
4	BIG	B	20.0	400.0
5	SMALL	C	30.0	500.0
6	BIG	A	30.0	600.0
7	SMALL	B	NaN	700.0
8	BIG	C	40.0	NaN
9	SMALL	A	40.0	800.0
10	BIG	B	50.0	900.0
11	SMALL	C	50.0	1000.0

Step 2 : pandas.pivot_table( )

In [3]:
```
pt = pd.pivot_table( raw_data,
                     index = ['NAME','PRODUCT'],
                     values = ['AMOUNT','PRICE'],
                     aggfunc = { 'AMOUNT':np.sum,'PRICE':np.mean },
                     dropna = False, fill_value = 0 )

pt
```

Out[3]:

		AMOUNT	PRICE
NAME	PRODUCT		
BIG	A	40	350
	B	70	650
	C	50	0
SMALL	A	60	550
	B	0	450
	C	80	750

Step 3 : pandas.DataFrame.reset_index( )

```
In [4]: pt.reset_index( inplace =True )
        pt
```

Out[4]:

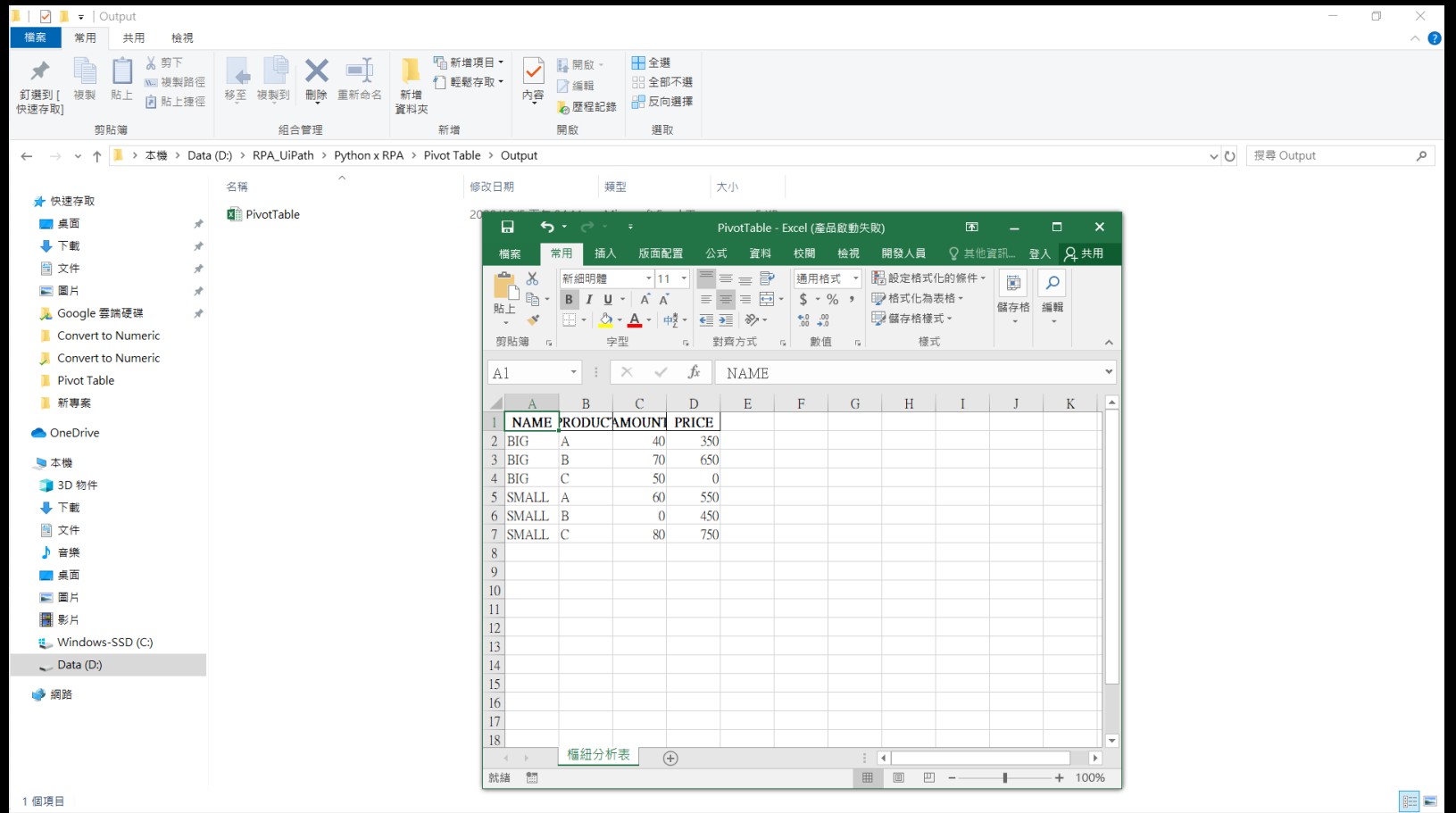
	NAME	PRODUCT	AMOUNT	PRICE
0	BIG	A	40	350
1	BIG	B	70	650
2	BIG	C	50	0
3	SMALL	A	60	550
4	SMALL	B	0	450
5	SMALL	C	80	750

Step 4 : pandas.DataFrame.to_excel( )

```
In [5]: output_path = 'D:\RPA_UiPath\Python x RPA\Pivot Table\Output\PivotTable.xlsx'

        pt.to_excel( output_path, sheet_name = '樞紐分析表', index = False )
```

Output



Another Way of Pivot Table

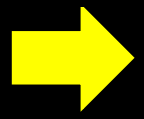
```
In [6]: total_amount = raw_data.groupby(['NAME','PRODUCT'])[['AMOUNT']].sum()
        mean_price = raw_data.groupby(['NAME','PRODUCT'])[['PRICE']].mean()
        pivot_table = pd.merge( total_amount, mean_price, on=['NAME','PRODUCT'], how='left' )

        display( total_amount )
        display( mean_price )
        display( pivot_table )
```

		AMOUNT
NAME	PRODUCT	
BIG	A	40.0
	B	70.0
	C	50.0
SMALL	A	60.0
	B	0.0
	C	80.0

		PRICE
NAME	PRODUCT	
BIG	A	350.0
	B	650.0
	C	NaN
SMALL	A	550.0
	B	450.0
	C	750.0

merge



		AMOUNT	PRICE
NAME	PRODUCT		
BIG	A	40.0	350.0
	B	70.0	650.0
	C	50.0	NaN
SMALL	A	60.0	550.0
	B	0.0	450.0
	C	80.0	750.0