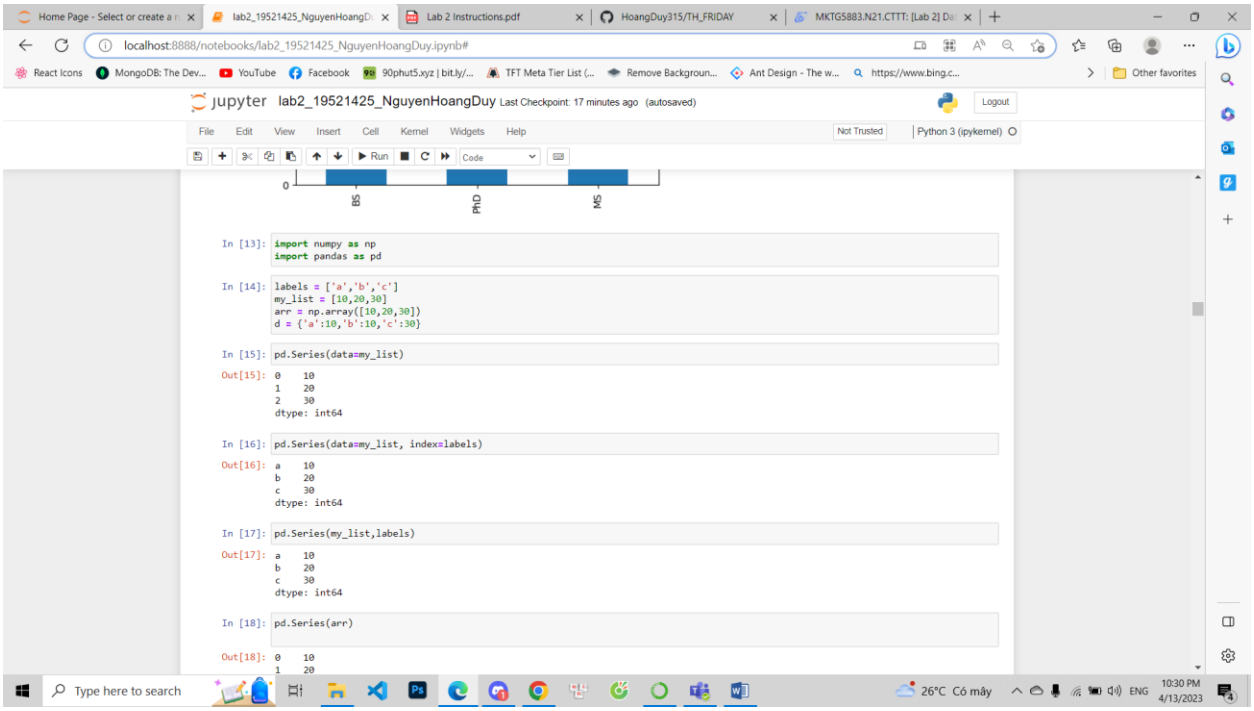
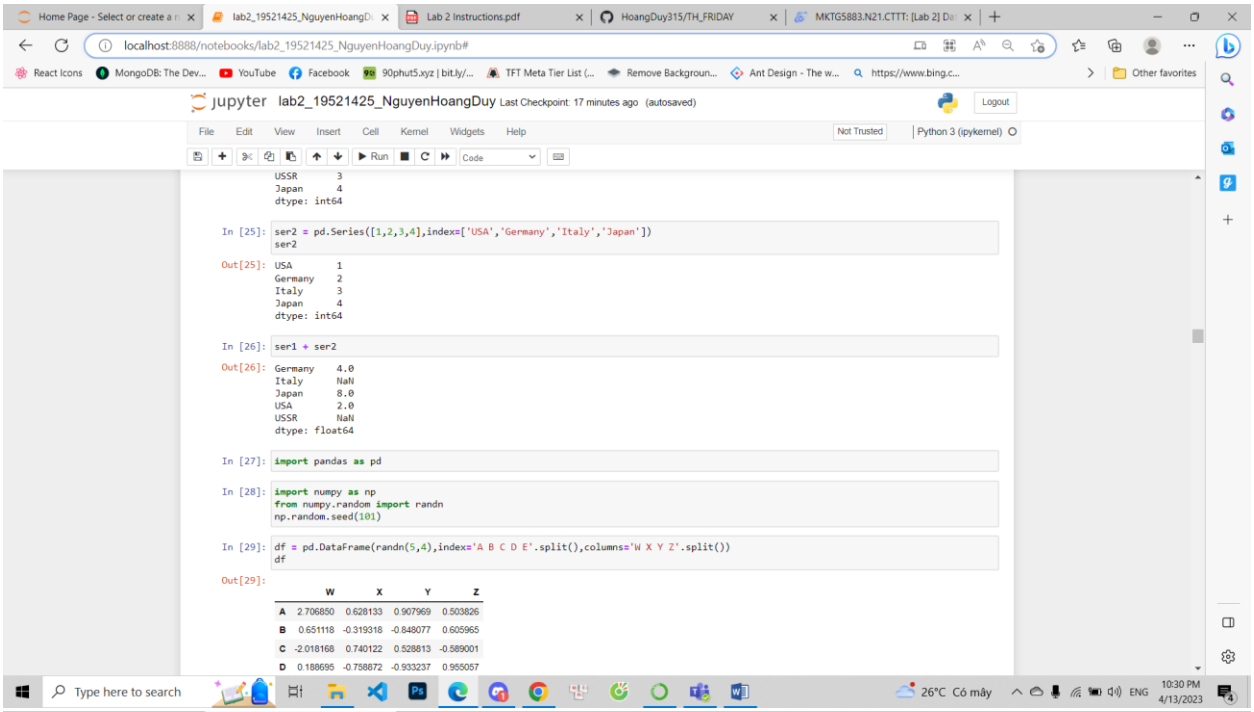
LAB2

NAME: NGUYEN HOANG DUY

ID:19521425

LINK GITHUB: [HoangDuy315/TH_FRIDAY (github.com)]





```
USSR      3
Japan     4
dtype: int64
```

In [25]:
```
ser2 = pd.Series([1,2,3,4],index=['USA','Germany','Italy','Japan'])
ser2
```

Out[25]:
```
USA       1
Germany   2
Italy     3
Japan     4
dtype: int64
```

In [26]:
```
ser1 + ser2
```

Out[26]:
```
Germany    4.0
Italy      NaN
Japan      8.0
USA        2.0
USSR       NaN
dtype: float64
```

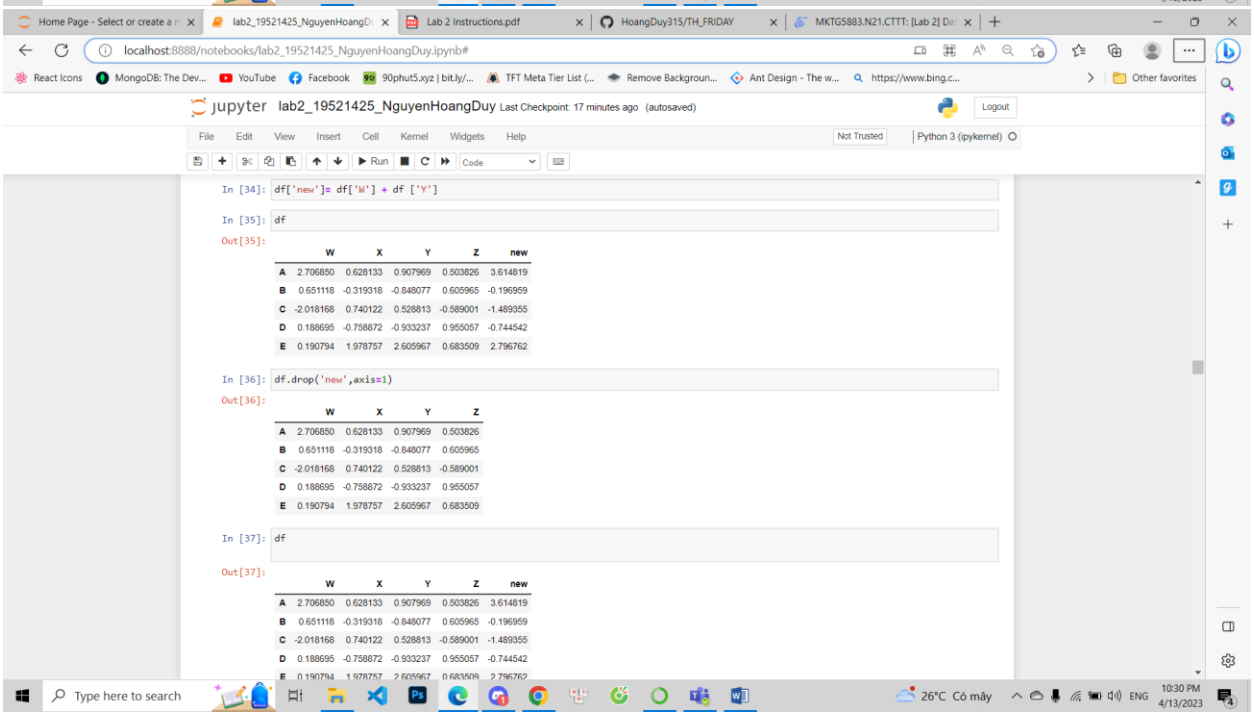
In [27]:
```
import pandas as pd
```

In [28]:
```
import numpy as np
from numpy.random import randn
np.random.seed(101)
```

In [29]:
```
df = pd.DataFrame(randn(5,4),index='A B C D E'.split(),columns='W X Y Z'.split())
df
```

Out[29]:

	W	X	Y	Z
A	2.706850	0.628133	0.907969	0.503826
B	0.651118	-0.319318	-0.848077	0.605965
C	-2.018168	0.740122	0.528813	-0.589001
D	0.188695	-0.758872	-0.933237	0.955057



In [34]:
```
df['new']= df['W'] + df ['Y']
```

In [35]:
```
df
```

Out[35]:

	W	X	Y	Z	new
A	2.706850	0.628133	0.907969	0.503826	3.614819
B	0.651118	-0.319318	-0.848077	0.605965	-0.196959
C	-2.018168	0.740122	0.528813	-0.589001	-1.489355
D	0.188695	-0.758872	-0.933237	0.955057	-0.744542
E	0.190794	1.978757	2.605967	0.683509	2.796762

In [36]:
```
df.drop('new',axis=1)
```

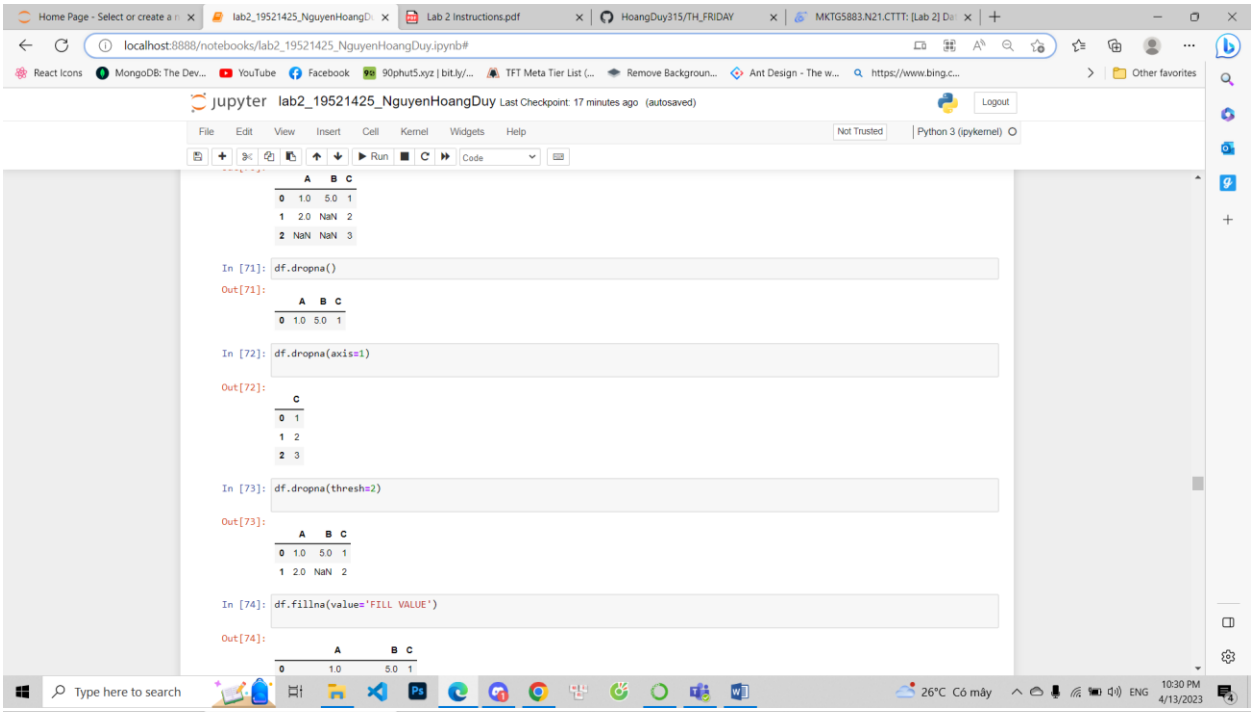
Out[36]:

	W	X	Y	Z
A	2.706850	0.628133	0.907969	0.503826
B	0.651118	-0.319318	-0.848077	0.605965
C	-2.018168	0.740122	0.528813	-0.589001
D	0.188695	-0.758872	-0.933237	0.955057
E	0.190794	1.978757	2.605967	0.683509

In [37]:
```
df
```

Out[37]:

	W	X	Y	Z	new
A	2.706850	0.628133	0.907969	0.503826	3.614819
B	0.651118	-0.319318	-0.848077	0.605965	-0.196959
C	-2.018168	0.740122	0.528813	-0.589001	-1.489355
D	0.188695	-0.758872	-0.933237	0.955057	-0.744542
E	0.190794	1.978757	2.605967	0.683509	2.796762



Home Page - Select or create a n | lab2_19521425_NguyenHoangDu | Lab 2 Instructions.pdf | HoangDuy315/TH_FRIDAY | MKTG5883.N21.CTTT: [Lab 2] Da

localhost:8888/notebooks/lab2_19521425_NguyenHoangDuy.ipynb#

React Icons | MongoDB: The Dev... | YouTube | Facebook | 90phut5.xyz | bit.ly/... | TFT Meta Tier List (... | Remove Backgroun... | Ant Design - The w... | https://www.bing.c... | Other favorites

jupyter lab2_19521425_NguyenHoangDuy Last Checkpoint: 17 minutes ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

Run Code

	A	B	C
0	1.0	5.0	1
1	2.0	NaN	2
2	NaN	NaN	3

In [71]:
```
df.dropna()
```

Out[71]:

	A	B	C
0	1.0	5.0	1

In [72]:
```
df.dropna(axis=1)
```

Out[72]:

	C
0	1
1	2
2	3

In [73]:
```
df.dropna(thresh=2)
```

Out[73]:

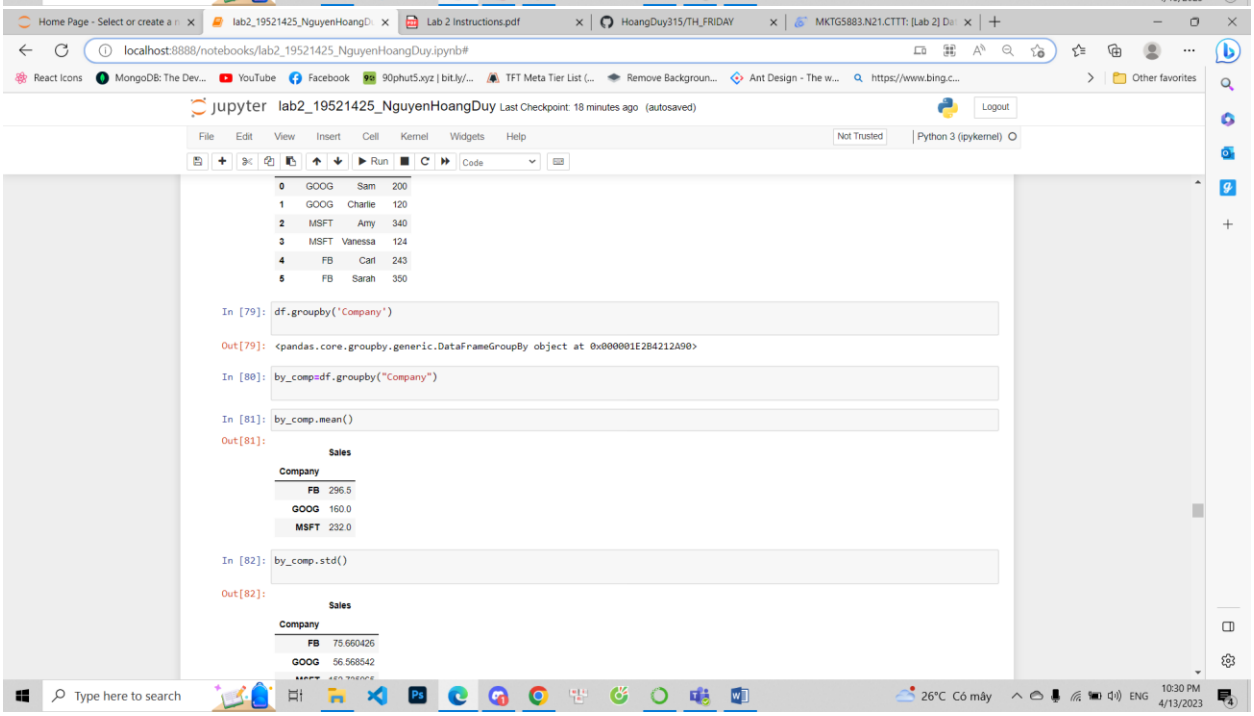
	A	B	C
0	1.0	5.0	1
1	2.0	NaN	2

In [74]:
```
df.fillna(value='FILL VALUE')
```

Out[74]:

	A	B	C
0	1.0	5.0	1

Type here to search 26°C Có mây ENG 10:30 PM 4/13/2023



Home Page - Select or create a n | lab2_19521425_NguyenHoangDu | Lab 2 Instructions.pdf | HoangDuy315/TH_FRIDAY | MKTG5883.N21.CTTT: [Lab 2] Da

localhost:8888/notebooks/lab2_19521425_NguyenHoangDuy.ipynb#

React Icons | MongoDB: The Dev... | YouTube | Facebook | 90phut5.xyz | bit.ly/... | TFT Meta Tier List (... | Remove Backgroun... | Ant Design - The w... | https://www.bing.c... | Other favorites

jupyter lab2_19521425_NguyenHoangDuy Last Checkpoint: 18 minutes ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

Run Code

0	GOOG	Sam	200
1	GOOG	Charlie	120
2	MSFT	Amy	340
3	MSFT	Vanessa	124
4	FB	Carl	243
5	FB	Sarah	350

In [79]:
```
df.groupby('Company')
```

Out[79]:
```
<pandas.core.groupby.generic.DataFrameGroupBy object at 0x000001E2B4212A90>
```

In [80]:
```
by_comp=df.groupby("Company")
```

In [81]:
```
by_comp.mean()
```

Out[81]:

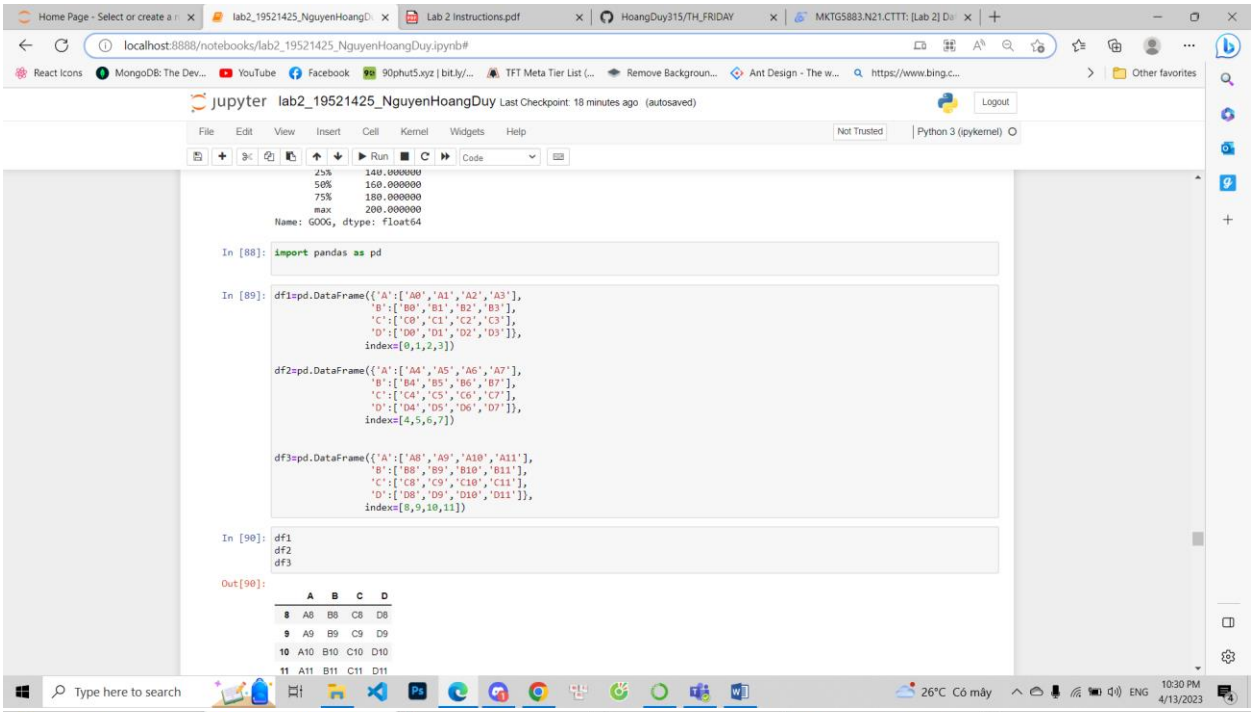
Company	Sales
FB	296.5
GOOG	160.0
MSFT	232.0

In [82]:
```
by_comp.std()
```

Out[82]:

Company	Sales
FB	75.660426
GOOG	56.568542

Type here to search 26°C Có mây ENG 10:30 PM 4/13/2023



```
25%     140.000000
50%     160.000000
75%     180.000000
max     200.000000
Name: GOOG, dtype: float64
```

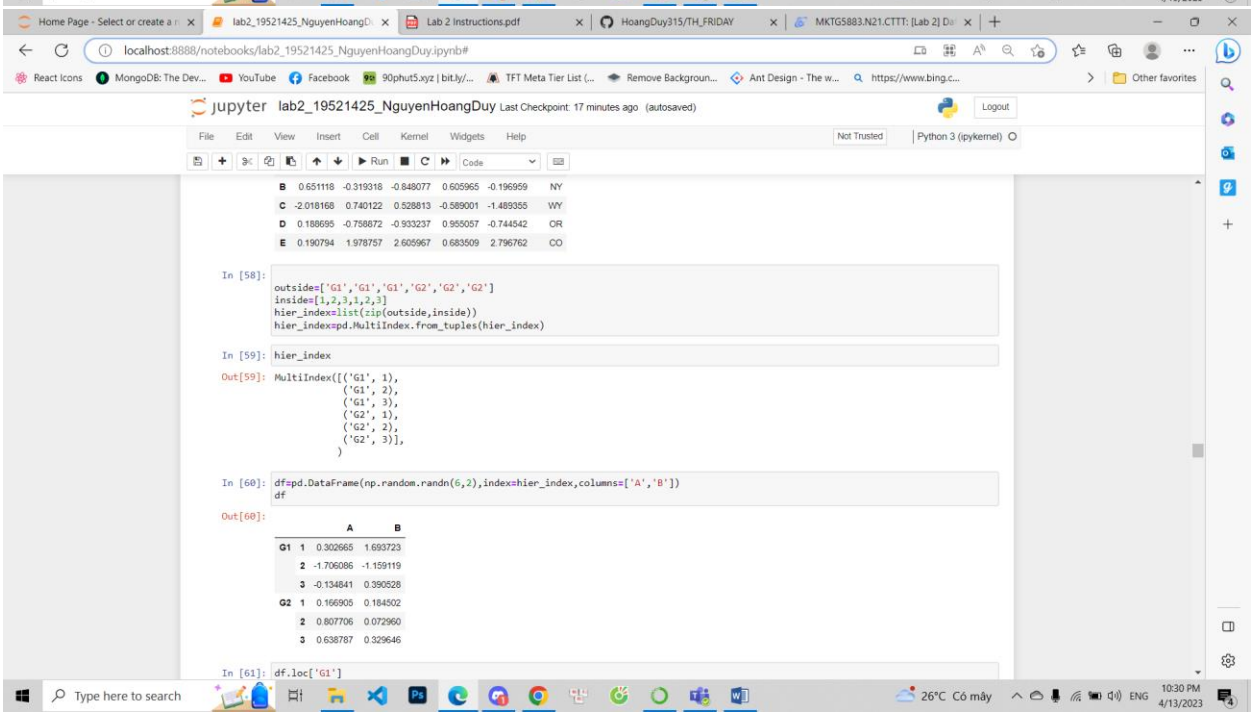
In [88]:
```
import pandas as pd
```

In [89]:
```
df1=pd.DataFrame({'A':['A0','A1','A2','A3'],
                  'B':['B0','B1','B2','B3'],
                  'C':['C0','C1','C2','C3'],
                  'D':['D0','D1','D2','D3']},
                 index=[0,1,2,3])

df2=pd.DataFrame({'A':['A4','A5','A6','A7'],
                  'B':['B4','B5','B6','B7'],
                  'C':['C4','C5','C6','C7'],
                  'D':['D4','D5','D6','D7']},
                 index=[4,5,6,7])


df3=pd.DataFrame({'A':['A8','A9','A10','A11'],
                  'B':['B8','B9','B10','B11'],
                  'C':['C8','C9','C10','C11'],
                  'D':['D8','D9','D10','D11']},
                 index=[8,9,10,11])
```

In [90]:
```
df1
df2
df3
```

Out[90]:

	A	B	C	D
8	A8	B8	C8	D8
9	A9	B9	C9	D9
10	A10	B10	C10	D10
11	A11	B11	C11	D11



B	0.651118	-0.319318	-0.848077	0.605965	-0.196959	NY
C	-2.018168	0.740122	0.528813	-0.589001	-1.489355	WY
D	0.188695	-0.758872	-0.933237	0.955057	-0.744542	OR
E	0.190794	1.978757	2.605967	0.683509	2.796762	CO

In [58]:
```
outside=['G1','G1','G1','G2','G2','G2']
inside=[1,2,3,1,2,3]
hier_index=list(zip(outside,inside))
hier_index=pd.MultiIndex.from_tuples(hier_index)
```

In [59]:
```
hier_index
```

Out[59]:
```
MultiIndex([('G1', 1),
            ('G1', 2),
            ('G1', 3),
            ('G2', 1),
            ('G2', 2),
            ('G2', 3)],
           )
```

In [60]:
```
df=pd.DataFrame(np.random.randn(6,2),index=hier_index,columns=['A','B'])
df
```

Out[60]:

		A	B
G1	1	0.302665	1.693723
	2	-1.706086	-1.159119
	3	-0.134841	0.390528
G2	1	0.166905	0.184502
	2	0.807706	0.072960
	3	0.638787	0.329646

In [61]:
```
df.loc['G1']
```