

Người thực hiện: Đoàn Trọng Tân

Email: tandoannrc@gmail.com

Vị trí ứng tuyển: Business Intelligence

Câu 1: Vui lòng tham khảo file Jupyter NoteBook đính kèm

Bước 1: Import thư viện cần thiết Bước 2: Tạo handle kết nối Jupyter Notebook	import import import #2. Crea handle	numpy a matplotl	as pd	lihraries												
viện cần thiết Bước 2: Tạo handle kết nối Jupyter	import import #2. Crea handle_	numpy a matplotl	•	ibraries												
thiết Bước 2: Tạo handle kết nối Jupyter	import #2. Crea handle_	matplotl	import pandas as pd import numpy as np													
Bước 2: Tạo handle kết nối Jupyter	#2. Crea		import numpy as np import matplotlib.pyplot as plt													
Tạo handle kết nối Jupyter	handle_	ate hand														
kết nối Jupyter	_	#2. Create handle to connect Python with Excel handle_excel = pd.ExcelFile('TestPython.xlsx')														
Jupyter	data = handle_excel.parse('InputData')															
	Output:															
	#2. Create handle to connect Python with Excel															
vs ExelFile	handle_excel = pd.ExcelFile('TestPython.xlsx')															
	<pre>data = handle_excel.parse('InputData')</pre>															
	data															
	Yea	r_month	Y6M	Manager	Director	Division	Revenue	Asset	Level_number	r Staff_quantity						
	0	202301	T1&2	Trưởng phòng 1	Giám đốc A	Corporate	10	12	27	7 6						
	1	202301	T1&2	Trưởng phòng 1	Giám đốc A	Corporate	3	10	24	1 2						
	2	202301	T1&2	Trưởng phòng 1	Giám đốc A	Corporate	9	11	25	5 5						
	3	202301	T1&2	Trưởng phòng 2	Giám đốc A	Corporate	10	17	23	5						
	4	202301	T1&2	Trưởng phòng 2	Giám đốc A	Corporate	7	20	25	5 2						
Chỉnh lại dữ liệu bị lỗi ở cột YearMont h	<pre>data['Year_month'] = pd.Series(str(month).replace("204304","202304").replace("203303", "202303") month in data['Year_month']) Output:  #2.1 Fix Errors in data data['Year_month'] = pd.Series(str(month).replace("204304","202304").replace("203303", "202 data[data['Year_month'] == "202303"]</pre>															
YearMont	Output 4]: #2.:	1 Fix Er a['Year_	rors i	] = pd.Series(		.replace(										
YearMont h	Output 4]: #2.: data data	1 Fix Er a['Year_	rors i	] = pd.Series(		.replace(										
YearMont h	Output 4]: #2.:	1 Fix Er a['Year_	rors i month' Year_m	] = pd.Series( onth'] == "202	303"]		"204304"	,"2023(		"203303", "202						
YearMont h	Output 4]: #2.: data data	1 Fix Er a['Year_ a[data[' Year_mon	rors i month' Year_m	] = pd.Series( onth'] == "202	r Director	Division	"204304"	,"2023(	04").replace(	"203303", "202						
YearMont h	Output  4]: #2. data data data 4]:	1 Fix Er a['Year_a a[data[' Year_mor	rors i. Month' Year_m  onth Y6	] = pd.Series( onth'] == "202 M Manage	r Director  Giám đốc A	Division Corporate	"204304" Revenue	,"20230 Asset	04").replace( Level_number	"203303", "202 Staff_quantity						
YearMont h	Output 4]: #2. data data 4]: 4]:	1 Fix Er a['Year_a[data[' Year_mor	rors imonth' Year_m  th Y6 03 T38	] = pd.Series( onth'] == "202  M	r Director 1 Giám đốc A 1 Giám đốc A	Division Corporate Corporate	"204304".  Revenue 2	Asset 20	04").replace( Level_number 21	Staff_quantity 2						
YearMont h	Output  4]: #2. data data data 4]:  38 39	1 Fix Er a ['Year_a [data[' Year_mor 2023 2023	rors i month' Year_m  onth Y6  onth Y6	] = pd.Series( onth'] == "202  M	r Director 1 Giám đốc A 1 Giám đốc A	Division Corporate Corporate Corporate	"204304" Revenue 2 4	Asset 20	04").replace( Level_number 21 29	Staff_quantity  2 6						
YearMont h	Output 4]: #2. dat. dat. dat. 4]: 38 39 40	1 Fix Er a ['Year_a [data [' Year_mor 2023 2023 2023	rors i. month' Year_m  th Y6 03 T38 03 T38 03 T38	] = pd.Series( onth'] == "202  M Manage 4 Trưởng phòng 6 4 Trưởng phòng 6	r Director 1 Giám đốc A 1 Giám đốc A 1 Giám đốc A 2 Giám đốc A	Division Corporate Corporate Corporate Corporate	"204304".  Revenue 2 4 10	Asset 20 18	Devel_number 21 29 25	Staff_quantity  2 6 4						

#### Bảng 1: Vẽ thống kê dạng bảng theo Manager, Director với Division là Corporate **#3. CORPORATE DIVISION** Bước 1: Loc ra #3.1 Filter Corporate Division corperate data = data[data['Division'] == 'Corporate'] Division Corporate Output: **#3. CORPORATE DIVISION** #3.1 Filter Corporate Division corperate\_data = data[data['Division'] == 'Corporate'] corperate\_data.head(5) Year\_month Y6M Manager Director Division Revenue Asset Level number Staff quantity 202301 T1&2 Trưởng phòng 1 Giám đốc A Corporate 0 10 12 27 6 202301 T1&2 Trưởng phòng 1 Giám đốc A Corporate 1 3 10 24 2 202301 T1&2 Trưởng phòng 1 Giám đốc A Corporate 2 9 11 25 5 202301 T1&2 Trưởng phòng 2 Giám đốc A Corporate 3 10 17 23 5 202301 T1&2 Trưởng phòng 2 Giám đốc A Corporate 7 20 25 2 Bước 2: #3.2 Create Pivot Table Pivot theo corporate table =pd.pivot table(corperate data, values =['Revenue', 'Asset', 'Level number', 'Staff quantity'], index =['Director', 'Manager'], Corporate Division columns =['Y6M'], aggfunc = np.sum) **Output:** #3.2 Create Pivot Table corporate\_table =pd.pivot\_table(corperate\_data, values =['Revenue', 'Ass columns =['Y6M'], aggfunc = np.sum) corporate table Asset Level\_number Revenue Staff\_quantity Y6M T1&2 T3&4 T1&2 T3&4 T1&2 T3&4 T1&2 T3&4 Director Manager Giám đốc A Trưởng phòng 1 78 97 145 154 46 35 26 25 Trưởng phòng 2 109 91 150 146 43 25 20 18 Giám đốc B Trưởng phòng 3 65 68 97 44 18 19 13 11 Trưởng phòng 4 43 43 83 73 15 17 12 13 Bước 3: #3.3 Create Subtotal for Director and Corporate Tính data 1 = pd.concat([ Subtotal y.append(y.sum().rename((x, ('Total of ' +x )))) for x, y in corporate table.groupby(level=0) (Giám đốc) và Total ]).append(corporate table.sum().rename(('Corporate', 'Total'))) của Output: Corporate

		Asset		Level_i	number	Reven	ue	Staff_q	uantity		
	Y6M	T1&2	T3&4	T1&2	T3&4	T1&2	T3&4	T1&2	T3&4		
Director	Manager										
Giám đốc A	Trưởng phòng 1	78	97	145	154	46	35	26	25		
	Trưởng phòng 2	109	91	150	146	43	25	20	18		
	Total of Giám đốc A	187	188	295	300	89	60	46	43		
Giám đốc B	Trưởng phòng 3	44	65	68	97	18	19	13	11		
	Trưởng phòng 4	43	43	83	73	15	17	12	13		
	Total of Giám đốc B	87	108	151	170	33	36	25	24		
Corporate	Total	274	296	446	470	122	96	71	67		
ble_2 = pd. dex =['Dired	DIVISION = data[data['Divi pivot_table(priva ctor', 'Manager'], columns =['Y6M	ate_d	ata, v	alues =	=['Reve	enue',	'Asse	t', 'Lev	el_num	ber', 'St	aff_quant
rivate_data able_2 = pd. dex =['Direc  utput: #4. PRIVA private_d	= data[data['Divi pivot_table(priva ctor', 'Manager'],	ate_d , ]'], ag ata[ ble(	ata, v gfunc 'Divi priva	alues = = np.s ision ate_d	=['Reve sum) '] == ata,	'Pr	ivato	e'] ['Rev		'Asse	
rivate_data able_2 = pd. dex =['Direc  utput: #4. PRIVA private_d	= data[data['Divi pivot_table(priva ctor', 'Manager'], columns =['Y6M TE DIVISION ata = data[da	ate_d , ]'], ag ata[ ble(	ata, v gfunc 'Divi priva	alues = = np.s ision ate_d	=['Reve sum) '] == ata,	'Pr	ivato	e'] ['Rev	enue',	'Asse	
rivate_data able_2 = pd. dex =['Direc  utput: #4. PRIVA  private_d table_2 =	= data[data['Divi pivot_table(priva ctor', 'Manager'], columns =['Y6M TE DIVISION ata = data[da	ate_d , ]'], ag ata[ ble(	ata, v gfunc 'Divi priva	alues = = np.s ision ate_d	=['Reve sum) '] == ata,	'Pr	ivato	e'] ['Rev	enue',	'Asse	
rivate_data able_2 = pd. dex =['Direc  utput: #4. PRIVA  private_d table_2 =	= data[data['Divi pivot_table(priva ctor', 'Manager'], columns =['Y6M TE DIVISION ata = data[da	ate_d , l'], ag ata[ ble(	ata, v gfunc 'Divi priva	alues = = np.s ision ate_d	=['Reve sum) '] == ata, ['Y6M	'Pr valu	ivat es = aggf	e'] ['Rev	enue', np.su	'Asse um)	
rivate_data able_2 = pd. dex =['Direc  utput: #4. PRIVA  private_d table_2 =	= data[data['Divi pivot_table(priva ctor', 'Manager'], columns =['Y6M TE DIVISION ata = data[da	ate_d , n'], ag ata[ ble(	ata, v gfunc 'Divi priva colur	= np.s ision ate_d nns =	=['Reve sum) '] == ata, ['Y6M	'Pr valu '],	ivat es = aggf	e'] ['Rev unc =	enue', np.su	'Asse um) Staff_q	et', 'L
rivate_data able_2 = pd. dex =['Direc  utput: #4. PRIVA  private_d table_2 =	= data[data['Divi pivot_table(priva ctor', 'Manager'], columns =['Y6M TE DIVISION ata = data[da pd.pivot_tal	ate_d , n'], ag ata[ ble(	ata, v gfunc 'Divi priva colur	= np.s ision ate_d nns =	=['Reve sum) '] == ata, ['Y6M	'Pr valu '],	ivate es = aggf	e'] ['Rev unc =	enue', np.su	'Asse um) Staff_q	et', 'L uantity
rivate_data able_2 = pd. dex =['Direc  utput: #4. PRIVA  private_d table_2 = table_2	= data[data['Divi pivot_table(priva ctor', 'Manager'], columns =['Y6M TE DIVISION ata = data[da pd.pivot_tal	ate_d , n'], ag ata[ ble(	ata, v gfunc 'Divi priva colur	= np.s ision ate_d nns =	=['Reve sum)  '] == ata, ['Y6M  Leve	'Pr valu '],	ivate es = aggf	e'] ['Rev unc =	enue', np.su	'Asse um) Staff_q	et', 'L uantity
rivate_data able_2 = pd. dex =['Direct  utput: #4. PRIVA  private_d table_2 = table_2  Directo	= data[data['Divi pivot_table(priva ctor', 'Manager'], columns =['Y6M TE DIVISION ata = data[da pd.pivot_tal	ate_d , n'], ag ata[ ble( ble(	ata, v gfunc 'Divi priva colur	alues = np.s ision ate_d nns =	=['Reve sum)  '] == ata, ['Y6M  Leve T1&2	'Pr valu '], ' I_num 2 T3	ivate es = aggfo nber &4	e'] ['Rev unc = Rever T1&2	enue', np.su uue T3&4	'Asse im) Staff_q T1&2	uantity

	Y6M	T1&2	T3&4	T1&2	T3&4	T1&2	T3&4	T1&2	T3&4
Director	Manager								
Giám đốc B	Trưởng phòng 3	46	33	73	51	11	2	11	7
	Trưởng phòng 4	42	48	73	75	27	7	5	14
Giám đốc C	Trưởng phòng 5	84	80	151	161	34	31	21	23
	Trưởng phòng 6	107	121	192	205	36	53	31	28

Bước 5: Tính Subtotal và lấy dòng cuối Private Total

Bước 4: Lọc ra Private Division và Pivot kết quả

> #4.1 Create Subtotal for Director and Private, then take the final row as Private Total data\_2 = pd.concat([

y.append(y.sum().rename((x, ('Total of ' +x ))))

for x, y in table\_2.groupby(level=0)

]).append(table\_2.sum().rename(('Private', 'Total')))

data\_2.iloc[[-1]] #Note

Output:

		Asset		Level_	number	Reven	iue	Staff_quantity		
	Y6M	T1&2	T3&4	T1&2	T3&4	T1&2	T3&4	T1&2	T3&4	
Director	Manager									
Private	Total	279	282	489	492	108	93	68	72	

Bước 6: Append dòng cuối của Private vào Corporate để có bảng Pivot tổng #4.2 Get the Final Pivot Result of CORPORATE df\_corporate = data\_1.append(data\_2.iloc[[-1]])

**Output:** 

		Asset		Level_i	number	Reven	ue	Staff_quantity	
	Y6M	T1&2	T3&4	T1&2	T3&4	T1&2	T3&4	T1&2	T3&4
Director	Manager								
Giám đốc A	Trưởng phòng 1	78	97	145	154	46	35	26	25
	Trưởng phòng 2	109	91	150	146	43	25	20	18
	Total of Giám đốc A	187	188	295	300	89	60	46	43
Giám đốc B	Trưởng phòng 3	44	65	68	97	18	19	13	11
	Trưởng phòng 4	43	43	83	73	15	17	12	13
	Total of Giám đốc B	87	108	151	170	33	36	25	24
Corporate	Total	274	296	446	470	122	96	71	67
Private	Total	279	282	489	492	108	93	68	72

Bước 7: Lấy list cột cần tính toán và so sánh #4.3 Get the column names of the Final Pivot Table columns = list(set([item[0] for item in df\_corporate.columns]))

## **Output:**

#4.3 Get the column names of the Final Pivot Table
columns = list(set([item[0] for item in df\_corporate.columns]))
columns

['Staff\_quantity', 'Asset', 'Revenue', 'Level\_number']

Bước 8: Tinh sự tang giảm (dòng vòng lặp for) #5 FINAL RESULT OF CORPORATE DIVISION

for column in columns:

 $\label{lem:df_corporate} $$ df_corporate[column + 'tăng/giảm (%)'] = round((df_corporate[column][['T3&4']]['T3&4'] - df_corporate[column][['T1&2']]['T1&2']) / df_corporate[column][['T1&2']]['T1&2'], 2) * 100 df_corporate.reindex(sorted(df_corporate.columns), axis=1) #reorde columns$ 

Công thức: [(Value T3&4 – Value T1&2)/ Value T1&2] \*100 Output:

#5 FINAL RESULT OF CORPORATE DIVISION

for column in columns:

df\_corporate[column + ' tăng/giảm (%)'] = round((df\_corporate[column][['T3&4']]['T3&4'] - df\_corporate[column][['T1&2']]['T1&df\_corporate.reindex(sorted(df\_corporate.columns), axis=1) #reorder columns|

		Asset		Asset tăng/giảm (%)	Level_	number	Level_number tăng/giảm (%)		Reven	ue	Revenue tăng/giảm (%)	Staff_c	uantity	Staff_quantity tăng/giảm (%)
	Y6M	T1&2	T3&4		T1&2	T3&4			T1&2	T3&4		T1&2	T3&4	
Director	Manager													
Giám đốc A	Trưởng phòng 1	78	97	24.0	145	154		6.0	46	35	-24.0	26	25	-4.0
	Trưởng phòng 2	109	91	-17.0	150	146		-3.0	43	25	-42.0	20	18	-10.0
	Total of Giám đốc A	187	188	1.0	295	300		2.0	89	60	-33.0	46	43	-7.0
Giám đốc B	Trưởng phòng 3	44	65	48.0	68	97		43.0	18	19	6.0	13	11	-15.0
	Trưởng phòng 4	43	43	0.0	83	73		-12.0	15	17	13.0	12	13	8.0
	Total of Giám đốc B	87	108	24.0	151	170		13.0	33	36	9.0	25	24	-4.0
Corporate	Total	274	296	8.0	446	470		5.0	122	96	-21.0	71	67	-6.0
Private	Total	279	282	1.0	489	492		1.0	108	93	-14.0	68	72	6.0

# Bảng 2: Vẽ thống kê dạng bảng theo Manager, Director với Division là Private (option)

Cách thức giống như cách xử lý và tính toán Corporate, #6 FINAL RESULT OF PRIVATE DIVISION

for column in columns:

df\_private[column + ' tăng/giảm (%)'] = round((df\_private[column][['T3&4']]['T3&4'] - df\_private[column][['T1&2']]['T1&2']) / df\_private[column][['T1&2']]['T1&2'], 2) \* 100 df private.reindex(sorted(df corporate.columns), axis=1)

### **Output:**

		Asset		Asset tăng/giảm (%)	Level_	number	Level_number tăng/giảm (%)	Reve	nue	Revenue tăng/giảm (%)	Staff_c	quantity	Staff_quantity tăng/giảm (%)
	Y6M	T1&2	T3&4		T1&2	T3&4		T1&2	T3&4		T1&2	T3&4	
Director	Manager												
Giám đốc B	Trưởng phòng 3	46	33	-28.0	73	51	-30.0	11	2	-82.0	11	7	-36.0
	Trưởng phòng 4	42	48	14.0	73	75	3.0	27	7	-74.0	5	14	180.0
	Total of Giám đốc B	88	81	-8.0	146	126	-14.0	38	9	-76.0	16	21	31.0
Giám đốc C	Trưởng phòng 5	84	80	-5.0	151	161	7.0	34	31	-9.0	21	23	10.0
	Trưởng phòng 6	107	121	13.0	192	205	7.0	36	53	47.0	31	28	-10.0
	Total of Giám đốc C	191	201	5.0	343	366	7.0	70	84	20.0	52	51	-2.0
Private	Total	279	282	1.0	489	492	1.0	108	93	-14.0	68	72	6.0
Corporate	Total	274	296	8.0	446	470	5.0	122	96	-21.0	71	67	-6.0

## Nhận xét về Revenue

- Doanh thu của nhìn chung giảm 18% (Corporate giảm 21% và Private giảm 14%) trong giai đoạn tháng 3-4/2023 so với tháng 1-2/2023.
- Xét Corporate Division:
- Trong đội ngũ của Giám Đốc A, xét thấy Trưởng phòng 2 giảm đều ở Staff\_quantity, Level\_number, Asset và Revenue. Vậy có thể đặt giả định là 1 trong 2 nhân viên vừa nghỉ trong T3&4 của Trưởng phòng 2 đóng góp khá nhiều vào Performance của đội ngũ.
- So sánh làm tăng độ tin cậy của độ tin cậy trên như sau:
- Xét đội ngũ Giám Đốc B, trưởng phòng 3, có 2 nhân viên nghỉ trong giai đoạn T3&4, nhưng các chỉ số còn lại không giảm, thậm chí tang đáng kể ở Asset và Level\_number, vậy thì 1 trong 2 nhân viên đó nghỉ việc mang lại dấu hiệu tốt để tái cấu trúc đội ngũ.
- ⇒ Vậy Trưởng phòng 2 cần xem lại cách quản lý, để tìm hiểu lý do nghỉ việc của 1 trong 2 nhân viên đóng góp hiệu quả cho đội ngũ, để đề xuất Ban giám đốc có những chế độ, phương pháp tốt hơn để giữ nhân viên tốt.
- Xét Private Division:
- Đội ngũ Giám Đốc B, dù có tăng nhẹ ở phần doanh thu ở Corporate Division, tuy nhiên là ở Private giảm đáng kể, đơn cử Trưởng phòng 4 dù số lượng nhân viên tăng lên 9 người (180%), nhưng doanh thu giảm tận 74%), vậy thì Giám Đốc B cần xem xét xem trưởng phòng 4 có phù hợp cho phân khúc Private nữa hay không để có định hướng phù hợp.

ve line chart the hiện Revenue Performand	ce cua ivianager qua cac thang
#7 Pivot Chart	

Bước 1: Lấy ra các cột cần thiết để Pivot thông tin

df\_chart = data[['Year\_month', 'Manager', 'Revenue']]

df chart pivot = pd.pivot table(df chart, values =['Revenue'], index =['Year month'],

columns =['Manager'], aggfunc = np.sum)

**Output:** 

Revenue

Manager	Trưởng phòng 1	Trưởng phòng 2	Trưởng phòng 3	Trưởng phòng 4	Trưởng phòng 5	Trưởng phòng 6
Year_month						
202301	22	22	11	27	19	13
202302	24	21	18	15	15	23
202303	16	13	18	17	15	18
202304	19	12	3	7	16	35

