

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

Xử lý, làm sạch dữ liệu																																																																															
Bước 1: Import thư viện cần thiết	#1. Import Necessary libraries <pre>import pandas as pd import numpy as np import matplotlib.pyplot as plt</pre>																																																																														
Bước 2: Tạo handle kết nối Jupyter Notebook vs ExcelFile	#2. Create handle to connect Python with Excel <pre>handle_excel = pd.ExcelFile('TestPython.xlsx') data = handle_excel.parse('InputData')</pre> Output: <pre>#2. Create handle to connect Python with Excel handle_excel = pd.ExcelFile('TestPython.xlsx') data = handle_excel.parse('InputData') data</pre> <table><thead><tr><th></th><th>Year_month</th><th>Y6M</th><th>Manager</th><th>Director</th><th>Division</th><th>Revenue</th><th>Asset</th><th>Level_number</th><th>Staff_quantity</th></tr></thead><tbody><tr><td>0</td><td>202301</td><td>T1&2</td><td>Trưởng phòng 1</td><td>Giám đốc A</td><td>Corporate</td><td>10</td><td>12</td><td>27</td><td>6</td></tr><tr><td>1</td><td>202301</td><td>T1&2</td><td>Trưởng phòng 1</td><td>Giám đốc A</td><td>Corporate</td><td>3</td><td>10</td><td>24</td><td>2</td></tr><tr><td>2</td><td>202301</td><td>T1&2</td><td>Trưởng phòng 1</td><td>Giám đốc A</td><td>Corporate</td><td>9</td><td>11</td><td>25</td><td>5</td></tr><tr><td>3</td><td>202301</td><td>T1&2</td><td>Trưởng phòng 2</td><td>Giám đốc A</td><td>Corporate</td><td>10</td><td>17</td><td>23</td><td>5</td></tr><tr><td>4</td><td>202301</td><td>T1&2</td><td>Trưởng phòng 2</td><td>Giám đốc A</td><td>Corporate</td><td>7</td><td>20</td><td>25</td><td>2</td></tr></tbody></table>										Year_month	Y6M	Manager	Director	Division	Revenue	Asset	Level_number	Staff_quantity	0	202301	T1&2	Trưởng phòng 1	Giám đốc A	Corporate	10	12	27	6	1	202301	T1&2	Trưởng phòng 1	Giám đốc A	Corporate	3	10	24	2	2	202301	T1&2	Trưởng phòng 1	Giám đốc A	Corporate	9	11	25	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	2										
	Year_month	Y6M	Manager	Director	Division	Revenue	Asset	Level_number	Staff_quantity																																																																						
0	202301	T1&2	Trưởng phòng 1	Giám đốc A	Corporate	10	12	27	6																																																																						
1	202301	T1&2	Trưởng phòng 1	Giám đốc A	Corporate	3	10	24	2																																																																						
2	202301	T1&2	Trưởng phòng 1	Giám đốc A	Corporate	9	11	25	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	2																																																																						
Bước 3: Chỉnh lại dữ liệu bị lỗi ở cột YearMont h	#2.1 Fix Errors in data <pre>data['Year_month'] = pd.Series(str(month).replace("204304","202304").replace("203303", "202303") for month in data['Year_month'])</pre> Output: <pre>4]: #2.1 Fix Errors in data data['Year_month'] = pd.Series(str(month).replace("204304","202304").replace("203303", "202303") data[data['Year_month'] == "202303"]</pre> <pre>4]:</pre> <table><thead><tr><th></th><th>Year_month</th><th>Y6M</th><th>Manager</th><th>Director</th><th>Division</th><th>Revenue</th><th>Asset</th><th>Level_number</th><th>Staff_quantity</th></tr></thead><tbody><tr><td>38</td><td>202303</td><td>T3&4</td><td>Trưởng phòng 1</td><td>Giám đốc A</td><td>Corporate</td><td>2</td><td>20</td><td>21</td><td>2</td></tr><tr><td>39</td><td>202303</td><td>T3&4</td><td>Trưởng phòng 1</td><td>Giám đốc A</td><td>Corporate</td><td>4</td><td>18</td><td>29</td><td>6</td></tr><tr><td>40</td><td>202303</td><td>T3&4</td><td>Trưởng phòng 1</td><td>Giám đốc A</td><td>Corporate</td><td>10</td><td>12</td><td>25</td><td>4</td></tr><tr><td>41</td><td>202303</td><td>T3&4</td><td>Trưởng phòng 2</td><td>Giám đốc A</td><td>Corporate</td><td>8</td><td>14</td><td>29</td><td>3</td></tr><tr><td>42</td><td>202303</td><td>T3&4</td><td>Trưởng phòng 2</td><td>Giám đốc A</td><td>Corporate</td><td>2</td><td>16</td><td>24</td><td>2</td></tr><tr><td>43</td><td>202303</td><td>T3&4</td><td>Trưởng phòng 2</td><td>Giám đốc A</td><td>Corporate</td><td>3</td><td>15</td><td>22</td><td>4</td></tr></tbody></table>										Year_month	Y6M	Manager	Director	Division	Revenue	Asset	Level_number	Staff_quantity	38	202303	T3&4	Trưởng phòng 1	Giám đốc A	Corporate	2	20	21	2	39	202303	T3&4	Trưởng phòng 1	Giám đốc A	Corporate	4	18	29	6	40	202303	T3&4	Trưởng phòng 1	Giám đốc A	Corporate	10	12	25	4	41	202303	T3&4	Trưởng phòng 2	Giám đốc A	Corporate	8	14	29	3	42	202303	T3&4	Trưởng phòng 2	Giám đốc A	Corporate	2	16	24	2	43	202303	T3&4	Trưởng phòng 2	Giám đốc A	Corporate	3	15	22	4
	Year_month	Y6M	Manager	Director	Division	Revenue	Asset	Level_number	Staff_quantity																																																																						
38	202303	T3&4	Trưởng phòng 1	Giám đốc A	Corporate	2	20	21	2																																																																						
39	202303	T3&4	Trưởng phòng 1	Giám đốc A	Corporate	4	18	29	6																																																																						
40	202303	T3&4	Trưởng phòng 1	Giám đốc A	Corporate	10	12	25	4																																																																						
41	202303	T3&4	Trưởng phòng 2	Giám đốc A	Corporate	8	14	29	3																																																																						
42	202303	T3&4	Trưởng phòng 2	Giám đốc A	Corporate	2	16	24	2																																																																						
43	202303	T3&4	Trưởng phòng 2	Giám đốc A	Corporate	3	15	22	4																																																																						

Bảng 1: Vẽ thống kê dạng bảng theo Manager, Director với Division là Corporate

Bước 1:
Lọc ra
Division
Corporate

#3. CORPORATE DIVISION
#3.1 Filter Corporate Division
corperate_data = data[data['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
0	202301	T1&2	Trưởng phòng 1	Giám đốc A	Corporate	10	12	27	6
1	202301	T1&2	Trưởng phòng 1	Giám đốc A	Corporate	3	10	24	2
2	202301	T1&2	Trưởng phòng 1	Giám đốc A	Corporate	9	11	25	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	2

Bước 2:
Pivot theo
Corporate
Division

#3.2 Create Pivot Table
corporate_table = pd.pivot_table(corperate_data, values =['Revenue', 'Asset', 'Level_number', 'Staff_quantity'], index =['Director', 'Manager'],
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	44	65	68	97	18	19	13	11
	Trưởng phòng 4	43	43	83	73	15	17	12	13

Bước 3:
Tính
Subtotal
(Giám đốc)
và Total
của
Corporate

#3.3 Create Subtotal for Director and Corporate
data_1 = pd.concat([
y.append(y.sum().rename((x, ('Total of ' + x))))
for x, y in corporate_table.groupby(level=0)
]).append(corporate_table.sum().rename(('Corporate', 'Total')))

Output:

</

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_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
	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:
Tính sự
tăng giảm
(dòng vòng
lặp for)

#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&2']) / df_corporate[column][['T1&2']]['T1&2'], 2) * 100
df_corporate.reindex(sorted(df_corporate.columns), axis=1) #reorde columns
Công thức: $[(\text{Value T3\&4} - \text{Value T1\&2}) / \text{Value T1\&2}] * 100$
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&2']) / df_corporate[column][['T1&2']]['T1&2'], 2) * 100

df_corporate.reindex(sorted(df_corporate.columns), axis=1) #reorder columns

		Asset		Asset	Level_number		Level_number	Revenue			Revenue	Staff_quantity		Staff_quantity	
				tăng/giảm (%)			tăng/giảm (%)				tăng/giảm (%)			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:

Y6M	Director	Manager	Asset		Asset tăng/giảm (%)		Level_number		Level_number tăng/giảm (%)		Revenue		Revenue tăng/giảm (%)		Staff_quantity		Staff_quantity tăng/giảm (%)	
			T1&2	T3&4	T1&2	T3&4	T1&2	T3&4	T1&2	T3&4	T1&2	T3&4	T1&2	T3&4	T1&2	T3&4	T1&2	T3&4
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.

Về line chart thể hiện Revenue Performance của Manager qua các tháng

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

#7 Pivot Chart

```
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:

Year_month	Revenue					
	Trưởng phòng 1	Trưởng phòng 2	Trưởng phòng 3	Trưởng phòng 4	Trưởng phòng 5	Trưởng phòng 6
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

Bước 2:
Vẽ chart

#8 Line chart visualization

```
plot_df = df_chart_pivot['Revenue']
```

```
ax = plot_df.plot(grid=True, figsize=(15, 10), marker='o', ylabel = "Revenue", title = "Sales Performance  
of Managers by Month")
```

```
#Show data value here
```

```
for col in plot_df.columns:
```

```
    for idx, val in enumerate(plot_df[col]):
```

```
        ax.text(idx, val, str(val))
```

Output:

