

NGÂN HÀNG NHÀ NƯỚC VIỆT NAM

BỘ GIÁO DỤC VÀ ĐÀO TẠO

TRƯỜNG ĐẠI HỌC NGÂN HÀNG TP. HỒ CHÍ MINH

-----o0o-----



BÁO CÁO THỰC TẬP

**ỨNG DỤNG XỬ LÝ VÀ TRỰC QUAN HÓA DỮ LIỆU TRONG
QUÁ TRÌNH XÂY DỰNG KẾ HOẠCH NĂM 2025 TẠI CÔNG
TY CỔ PHẦN ĐẦU TƯ PHAN VŨ**

Sinh viên thực hiện : NGUYỄN THANH VY

MSSV : 030237210219

Lớp : DH37KH01

Khóa : 37 (2021 – 2025)

Giảng viên hướng dẫn : Th.S VƯƠNG TRỌNG NHÂN

Tp. Hồ Chí Minh, tháng 1 năm 2025

MỤC LỤC

CHƯƠNG 1 : GIỚI THIỆU ĐƠN VỊ THỰC TẬP	1
1.1. Tổng quan Công ty	1
1.1.1. Thông tin công ty	1
1.1.2. Lịch sử hình thành và phát triển	1
1.1.3. Tầm nhìn, sứ mệnh và giá trị cốt lõi	2
1.1.4. Sơ đồ cơ cấu tổ chức	3
1.2. Quá trình thực tập tại đơn vị	3
1.2.1. Tổng quan bộ phận thực tập : Phòng Kế Hoạch Chuỗi Cung Ứng	3
1.2.2. Mô tả nghiệp vụ thực tập	4
1.2.3. Mục tiêu thực tập	4
1.2.4. Kết quả đạt được	4
CHƯƠNG 2 : ỨNG DỤNG XỬ LÝ VÀ TRỰC QUAN HÓA DỮ LIỆU TRONG QUÁ TRÌNH XÂY DỰNG KẾ HOẠCH NĂM 2025	5
2.1. Dữ liệu được sử dụng	5
2.2. Quá trình xử lý và trực quan dữ liệu	8
2.2.1. Microsoft Excel	8
2.2.2. Python	12
2.2.3. Power BI	14
2.3. Nhận xét và đề xuất giải pháp cải tiến nền tảng quản lý dữ liệu	20
CHƯƠNG 3 : KẾT LUẬN	23
3.1. Kết quả đạt được	23
3.2. Hướng phát triển	24
3.3. Kết luận	25
TÀI LIỆU THAM KHẢO	26

DANH MỤC HÌNH ẢNH, SƠ ĐỒ

Hình 1.1. Logo của Công ty Cổ Phần Đầu Tư Phan Vũ

Hình 1.2. Sơ đồ cơ cấu tổ chức của Công Ty Phan Vũ

Hình 2.1. Bảng dữ liệu thông tin các phân khúc thuộc ngành xây dựng

Hình 2.2. Bảng dữ liệu các chỉ số thị trường

Hình 2.3. Bảng thông tin số liệu về đối thủ cạnh tranh

Hình 2.4. Biểu mẫu 1 - Thông tin các cơ hội dự án năm 2025 của Văn Phòng Kinh Doanh

Hình 2.5. Biểu đồ dữ liệu dự báo theo quý

Hình 2.6. Biểu đồ dữ liệu dự báo theo năm

Hình 2.7. Biểu đồ thể hiện các chỉ số ROA, ROE, ROS, Growth của đối thủ

Hình 2.8. Pivot Table thể hiện dữ liệu của các đối thủ cạnh tranh

Hình 2.9. Biểu đồ các chỉ số từ Pivot Table

Hình 2.10. Code Python để nhập dữ liệu từ các tập Excel vào DataFrame

Hình 2.11. Code Python để gộp các tập dữ liệu thành một DataFrame

Hình 2.12. Code Python để làm sạch dữ liệu trong các DataFrame

Hình 2.13. Code Python để chỉnh sửa các giá trị trong các DataFrame

Hình 2.14. Code Python để xuất Dataframe ra file Excel mới

Hình 2.15. Thực hiện chuyển đổi dữ liệu trên Power Query Editor

Hình 2.16. Tổng quan Doanh số - Trang 1 của Dashboard

Hình 2.17. Tổng quan Doanh số theo chiều dọc - Trang 2 của Dashboard

Hình 2.18. Tổng quan Doanh số theo tháng năm 2025 - Trang 3 của Dashboard

CHƯƠNG 1 : GIỚI THIỆU ĐƠN VỊ THỰC TẬP

1.1. Tổng quan Công ty



Hình 1.1. Logo của Công ty Cổ Phần Đầu Tư Phan Vũ

Nguồn : Công ty Cổ Phần Đầu Tư Phan Vũ (2024)

1.1.1. Thông tin công ty

Tên công ty : Công ty Cổ Phần Đầu Tư Phan Vũ

Tên Quốc tế : Phan Vu Investment Corporation

Địa chỉ : A2 Trường Sơn, Phường 2, Quận Tân Bình, thành phố Hồ Chí Minh

Website : <https://www.phanvu.vn/>

Email : phanvu@phanvu.com

Điện thoại : +84 (028) 22200884-6

Fax : +84 (028) 22200887

1.1.2. Lịch sử hình thành và phát triển

Phan Vũ là một trong những doanh nghiệp tiên phong và hàng đầu trong lĩnh vực sản xuất, thiết kế, thi công xây dựng tại Việt Nam. Được thành lập từ năm 1996, Phan Vũ đã khẳng định vị thế tiên phong khi trở thành nhà sản xuất cọc bê tông ly tâm dự ứng lực đầu tiên tại Việt Nam. [1]

Khởi nguồn từ một công ty nhỏ tại Bình Dương mang tên Công ty Trách nhiệm hữu hạn Sản xuất và Xây dựng Phan Vũ, đến năm 2001 Công ty Cổ Phần Đầu Tư Phan Vũ ra đời. Liên tiếp trong các năm sau đó, Phan Vũ đã không ngừng mở rộng phạm vi hoạt động tại miền Trung và miền Nam Việt Nam. Năm 2007, các công ty con đã

CHƯƠNG 2 : ỨNG DỤNG XỬ LÝ VÀ TRỰC QUAN HÓA DỮ LIỆU TRONG QUÁ TRÌNH XÂY DỰNG KẾ HOẠCH NĂM 2025

2.1. Dữ liệu được sử dụng

Quá trình thu thập dữ liệu tại doanh nghiệp được triển khai một cách bài bản, có hệ thống, tập trung vào hai mảng dữ liệu chính : dữ liệu thị trường và dữ liệu doanh nghiệp [2]. Nhằm cung cấp cơ sở dữ liệu đầy đủ và chính xác để hỗ trợ doanh nghiệp phân tích xu hướng thị trường, đánh giá tình hình cạnh tranh và xây dựng các chiến lược kinh doanh hiệu quả. Việc này hướng đến mục tiêu dự đoán chính xác các chỉ số tài chính, xác định các yếu tố ảnh hưởng đến ngành và xây dựng KPI phù hợp cho năm 2025.

2.1.1. Dữ liệu thị trường

Dữ liệu thị trường bao gồm các thông tin định tính và định lượng như chỉ số tài chính, thông tin ngành nghề kinh doanh và đặc điểm kinh tế liên quan. Dữ liệu này được thu thập từ các nguồn chính thống và đáng tin cậy như các báo cáo của Chính phủ và Bộ Xây dựng cung cấp thông tin về :

- Chính sách đầu tư và ưu đãi thuế.
- Các dự án công lớn và phân khúc đầu tư.
- Các yếu tố ảnh hưởng đến hoạt động của ngành xây dựng và hoạt động.

Bất Động Sản	Tốc độ tăng trưởng của hoạt động kinh doanh bất động sản tăng 2,45% so với cùng kỳ năm 2023. Mức độ quan tâm nhà đất của nhà đầu tư trong quý III/2024 tăng khoảng 49% so với cùng kỳ năm 2023. Trong đó, phân khúc nhà riêng tăng 25%, chung cư tăng 24%, biệt thự tăng 22%.	Thị trường bất động sản Việt Nam có nhiều tín hiệu tích cực , dự kiến có giá trị 4,41 nghìn tỷ USD trong năm 2024.	- Quý 1/2025: Giai đoạn củng cố sẽ diễn ra khi điều kiện tiền tệ thuận lợi, nhà riêng và nhà phố sẽ có lợi thế. - Quý 2/2025: Thị trường dự kiến khởi sắc với đất nền và biệt thự được quan tâm nhiều hơn.
	Thị trường có khoảng 253.000 giao dịch thành công, bằng 110,26% so với 6 tháng cuối năm 2023, với lượng giao dịch tập trung vào đất nền. Nguồn cung nhà ở thương mại có 18 dự án hoàn thành, 23 dự án được cấp phép mới và 984 dự án đang triển khai. Nhóm BĐS khu công nghiệp tiếp tục đã tăng nhờ vốn FDI tích cực từ đầu năm	Chính sách giảm thuế bảo vệ môi trường đối với nhiên liệu và thuế giá trị gia tăng đối với một số hàng hóa, dịch vụ tiếp tục được áp dụng đến hết năm 2024 trong khi Luật Đất đai sửa đổi có hiệu lực từ tháng 8 sẽ tác động tích cực lên triển vọng ngành bất động sản.	
Công Nghiệp Nhẹ	- Tăng trưởng ổn định, nhu cầu cao từ thị trường. - Các sản phẩm thời trang và thực phẩm chế biến được ưa chuộng.	Tập trung vào đổi mới sáng tạo và nâng cao chất lượng sản phẩm.	Tăng trưởng tiếp tục do nhu cầu tiêu dùng tăng và thị trường xuất khẩu mở rộng.
	Quyết định số 1018/QĐ-TTg ngày 21/9/2024 ban hành Chiến lược phát triển công nghiệp bán dẫn Việt Nam đến năm 2030 và tầm nhìn 2050.	Mục tiêu thu hút FDI, phát triển doanh nghiệp thiết kế và nhà máy chế tạo chip, cùng việc đào tạo nguồn nhân lực kỹ sư.	Doanh thu dự kiến đạt 100 tỷ USD/năm vào năm 2050.

Hình 2.1. Bảng dữ liệu thông tin các phân khúc thuộc ngành xây dựng

Nguồn : File dữ liệu thông tin thị trường (2024)

Tập dữ liệu đóng vai trò là công cụ giúp doanh nghiệp đánh giá và phân tích toàn diện xu hướng và tình hình thị trường, từ đó xác định các yếu tố cơ bản tác động, cơ hội và thách thức để tối ưu hóa chiến lược kinh doanh hỗ trợ doanh nghiệp đưa ra quyết định chiến lược chính xác.

Ngoài ra, dữ liệu từ Tổng cục Thống kê, Báo Chính phủ, Báo Xây dựng và World Bank cũng được thu thập. Tập dữ liệu này bao gồm các chỉ số kinh tế vĩ mô quan trọng như :

- Tăng trưởng GDP và GDP bình quân đầu người (VND, USD).
- Chỉ số giá tiêu dùng (CPI) và tỷ lệ lạm phát.

Dữ liệu được thu thập từ năm 2015 đến năm 2024, với tần suất cả hàng quý và hàng năm với mục đích theo dõi chi tiết xu hướng, phân tích sự biến động qua từng giai đoạn và hỗ trợ doanh nghiệp dự báo các chỉ số kinh tế vào năm 2025 nhằm đưa ra các quyết định dựa trên xu hướng dự báo.

B	C	D	AS	AT
GDP THỰC (tỷ)	01/03/2015	01/06/2015	01/09/2025	01/12/2025
Tổng sản phẩm trong nước theo giá so sánh	535.045	690.267	1.633.517	1.842.691
Nông nghiệp	49.790	114.829	118.423	146.061
Lâm nghiệp	3.617	4.163	7.943	9.060
Thủy sản	13.604	23.648	40.500	38.593
Công nghiệp	159.521	192.341	493.048	560.346
Ngành công nghiệp khác	46.353	59.687	39.171	41.696
Xây dựng	25.791	36.363	109.216	133.485
GDP DANH NGHĨA (tỷ)	01/03/2015	01/06/2015	01/09/2025	01/12/2025
GDP theo giá hiện hành	809.613	970.388	2.899.825	3.352.311
Nông nghiệp	71.249	156.728	232.456	279.884
Lâm nghiệp	5.674	6.731	14.004	17.296
Thủy sản	23.055	33.856	79.198	77.299
Công nghiệp	245.643	264.473	924.081	1.030.561
Công nghiệp chế biến, chế tạo	119.580	143.027	736.847	812.147
Xây dựng	36.905	49.579	182.510	223.415

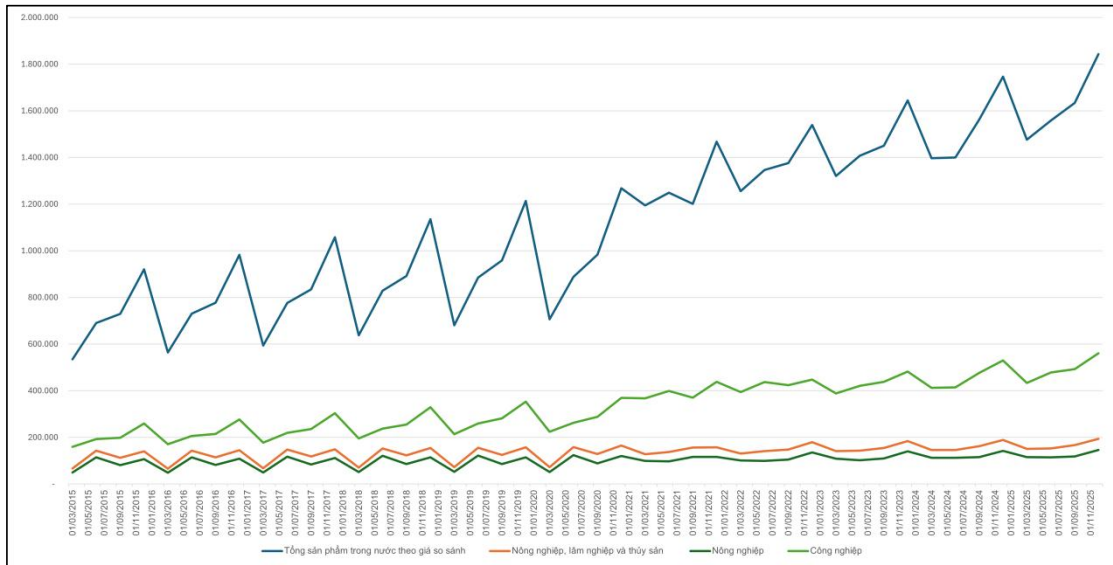
Hình 2.2. Bảng dữ liệu các chỉ số thị trường

Nguồn : File dữ liệu thông tin thị trường (2024)

Các dữ liệu này là bước đầu tiên trong việc hiểu rõ hơn về thị trường kinh tế, thị trường ngành xây dựng và phân khúc kinh tế mà doanh nghiệp đang hoạt động, qua

Microsoft Excel được sử dụng để tiến hành dự báo các chỉ số trong tập dữ liệu của thị trường qua từng giai đoạn với hàm FORECAST.ETS trên hai loại dữ liệu khác nhau [5] :

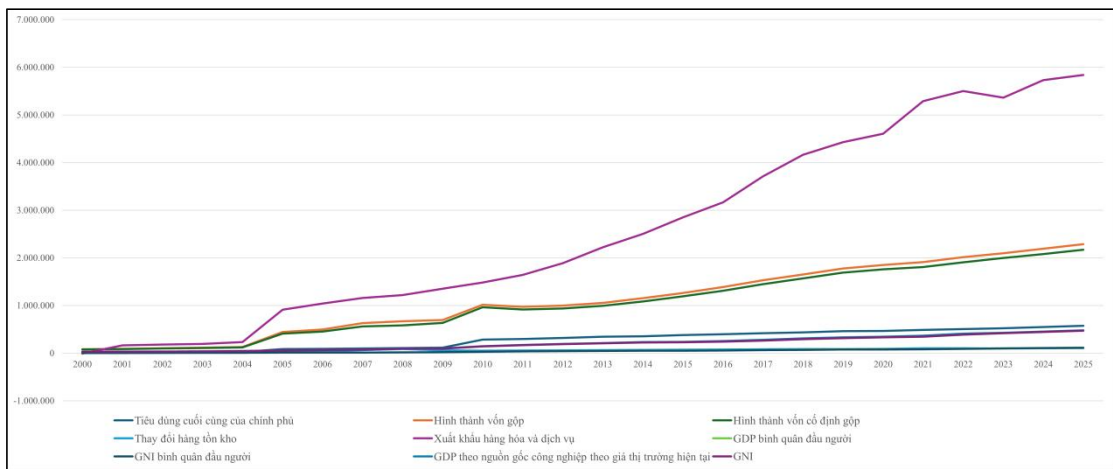
- Dữ liệu quý : Dựa trên các chỉ số theo từng quý bắt đầu từ năm 2017 đến tháng 9 năm 2024 để dự báo các chỉ số GDP, CPI, FDI,... đến hết năm 2025 với chu kỳ 4 quý.



Hình 2.5. Biểu đồ dữ liệu dự báo theo quý

Nguồn : File dữ liệu thông tin thị trường (2024)

- Dữ liệu năm : Dựa trên các chỉ số theo từng năm bắt đầu từ năm 2000 đến năm 2023 để dự báo các chỉ số GDP, CPI, FDI,... đến năm 2025 với chu kỳ 1 năm.



Hình 2.6. Biểu đồ dữ liệu dự báo theo năm

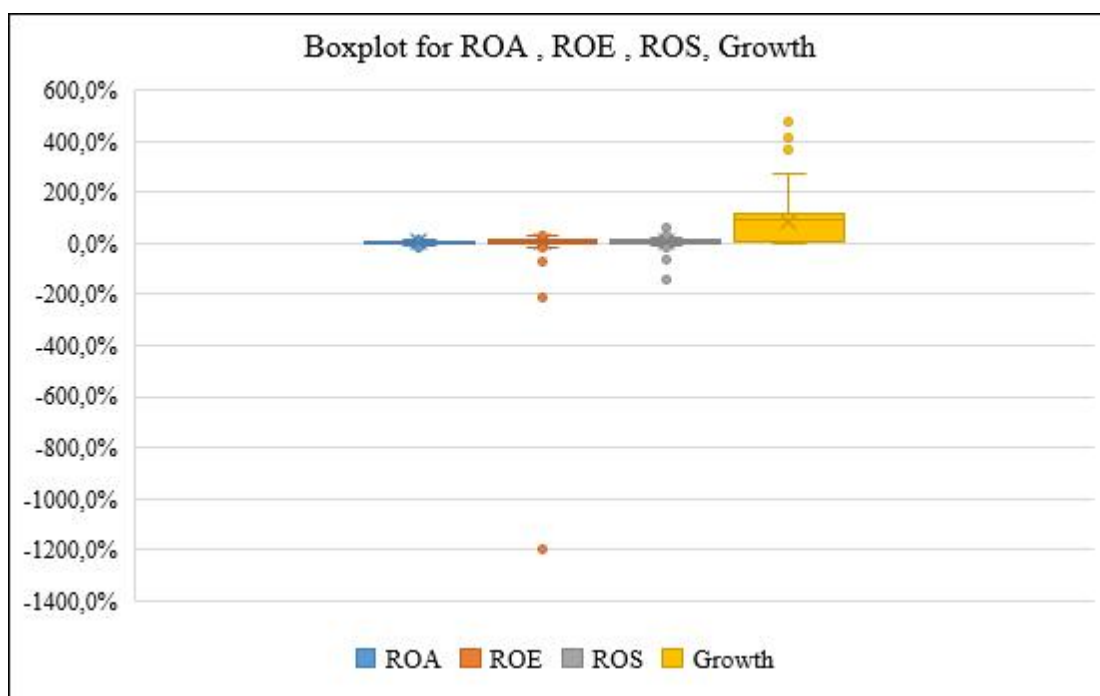
Nguồn : File dữ liệu thông tin thị trường (2024)

Mục đích của việc dự báo là để nhận diện xu hướng của thị trường qua từng chu kỳ khác nhau và cung cấp cái nhìn tổng quan về tình hình thị trường qua các năm. Kết quả cho thấy :

- Các giá trị dữ liệu có xu hướng lặp lại theo chu kỳ mỗi 4 quý
- Các chỉ số như GDP, CPI, và FDI đều có xu hướng đi lên theo từng năm, chứng minh thị trường đang phát triển tích cực và dự kiến sẽ tiếp tục mở rộng mạnh mẽ trong những năm tiếp theo.

Ngoài ra, Microsoft Excel còn được sử dụng để hỗ trợ phát hiện và tinh chỉnh các lỗi dữ liệu trong tập tin liên quan đến các công ty đối thủ cạnh tranh, đặc biệt là các trường hợp dữ liệu bị thiếu hoặc sai số :

- Phát hiện và tinh chỉnh dữ liệu bị thiếu hoặc sai số : sử dụng lineplot và boxplot trong tab Insert để xác định dữ liệu ngoại lệ (outliers). Cụ thể khi sử dụng boxplot, có thể phát hiện giá trị ROE của một số công ty đối thủ cạnh tranh có các giá trị bất thường dưới mức - 200%, giá trị Growth trên 300%. Những giá trị này sẽ được xác định và liên hệ với văn phòng kinh doanh để cập nhật lại.



Hình 2.7. Biểu đồ thể hiện các chỉ số ROA, ROE, ROS, Growth của đối thủ

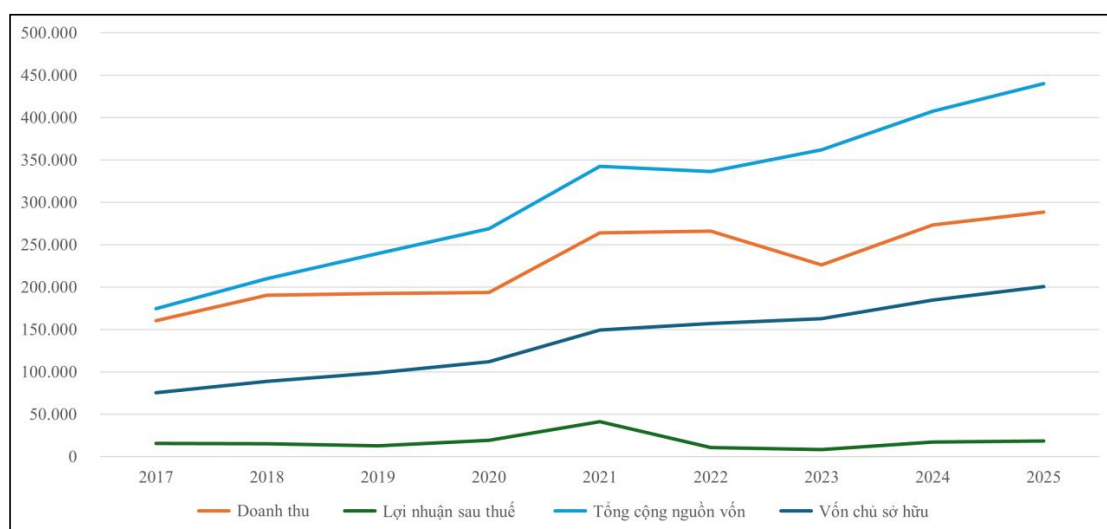
Nguồn : File dữ liệu thông tin về các đối thủ cạnh tranh (2024)

- Pivot Table : công cụ này giúp tóm tắt, phân tích, khám phá và trình bày dữ liệu một cách có tổ chức và đa chiều. Khi sử dụng Pivot Table, có thể dễ dàng lọc dữ liệu theo các giá trị định tính trong tập dữ liệu, nhóm dữ liệu theo từng mảng, từng lĩnh vực và có được các cột giá trị tính theo tổng và trung bình từng năm. Từ đó, có thể vẽ ra các biểu đồ thể hiện xu hướng khách quan của các chỉ số qua từng năm. Kết quả cho thấy các chỉ số doanh thu, lợi nhuận sau thuế và tổng nguồn vốn đều có xu hướng tăng trưởng trong tập dữ liệu.

Công ty	(All)		
Lĩnh vực KD	(All)		
Mảng	(All)		
Row Labels	Sum of Doanh thu (01)	Sum of Lợi nhuận sau thuế (60)	Sum of Tổng cộng nguồn vốn (440)
2017			
2018			
2019			
2020			
2021			
2022			
2023			
6TD 2024			
Grand Total			

Hình 2.8. Pivot Table thể hiện dữ liệu của các đối thủ cạnh tranh

Nguồn : File dữ liệu thông tin về các đối thủ cạnh tranh (2024)



Hình 2.9. Biểu đồ các chỉ số từ Pivot Table

Nguồn : File dữ liệu thông tin về các đối thủ cạnh tranh (2024)

Kết luận : Microsoft Excel chứng minh khả năng mạnh mẽ trong việc quản lý các tập dữ liệu vừa và nhỏ, phát hiện các điểm bất thường, trực quan hóa và dự báo dữ liệu hiệu quả. Tuy nhiên, khi làm việc với các tập dữ liệu lớn, yêu cầu tích hợp

nhiều nguồn dữ liệu hoặc cập nhật dữ liệu theo thời gian thực, Excel gặp một số hạn chế. Để xử lý hiệu quả hơn trong những trường hợp này, các công cụ như Python trên Jupyter Notebook hoặc Power BI có thể được áp dụng để đạt hiệu suất cao hơn.

2.2.2. Python

Python được sử dụng để xử lý các tập dữ liệu liên quan đến thông tin dự án từ các văn phòng kinh doanh, nhằm mục đích xây dựng một tập dữ liệu về các dự án chính xác và đầy đủ nhất, phục vụ cho việc trực quan hóa thông tin một cách hiệu quả [3] [4]. Quy trình thực hiện bao gồm :

- Tích hợp dữ liệu : Tổng hợp các tập dữ liệu từ 6 văn phòng kinh doanh để tạo thành một tập dữ liệu hoàn chỉnh, phục vụ cho các mô hình phân tích. Mục đích là để thống nhất giá trị giữa các tập dữ liệu từ các văn phòng để đảm bảo dữ liệu đồng nhất, dễ dàng xử lý và phân tích. [6]
- Load dữ liệu vào dataframe

```
import seaborn as sns
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import unicode

# Đường dẫn đến thư mục chứa các file Excel
folder_path = r"D:\Thực tập\Thực Hiện\Data\BM_1"
# Danh sách tên file Excel (bao gồm đuôi .xlsx)
file_names = ["Miền Tây.xlsx", "Miền Trung.xlsx", "Tổng Hợp.xlsx",
               "Trung Tâm.xlsx", "Duyên Hải.xlsx", "Nam Hà Nội.xlsx"]
# Danh sách để chứa các DataFrame
df_list = []

# Đọc dữ liệu từ các file Excel và tạo các DataFrame độc lập
for file_name in file_names:
    # Tạo đường dẫn đầy đủ đến file
    file_path = f"{folder_path}\\{file_name}"
    # Tên DataFrame (bỏ phần .xlsx trong tên file và chuyển thành chữ thường không dấu)
    df_name = unicode.unidecode(file_name.replace(".xlsx", "").lower())
    # Đọc dữ liệu từ sheet
    try:
        df = pd.read_excel(file_path, sheet_name='BM_01.Thông tin cơ hội dự án')
        # Thêm DataFrame vào danh sách
        df_list.append(df)
    except Exception as e:
        print(f"Lỗi khi đọc file: {file_name} - {e}")
```

Hình 2.10. Code Python để nhập dữ liệu từ các tập Excel vào DataFrame

Nguồn : Jupyter Notebook về xử lý dữ liệu (2024)

- Làm sạch dữ liệu : Loại bỏ ký tự không hợp lệ, chuẩn hóa định dạng và sửa lỗi nhập liệu do người dùng gây ra. [7]
- Gộp các tập dữ liệu lại thành một tập dữ liệu tổng hợp đầy đủ

```
# Gộp tất cả các DataFrame thành một DataFrame duy nhất
df_combined = pd.concat(df_list, ignore_index=True)
```

Hình 2.11. Code Python để gộp các tập dữ liệu thành một DataFrame

Nguồn : Jupyter Notebook về xử lý dữ liệu (2024)

- Tinh chỉnh và làm sạch dữ liệu

```
# Xử lý DataFrame chung
# Loại bỏ hàng trống
df_combined = df_combined.dropna(how='all')
# Loại bỏ các giá trị trùng lặp
df_combined = df_combined.drop_duplicates()
# Loại bỏ khoảng trắng ở đầu và cuối tên các cột
df_combined.columns = df_combined.columns.str.strip()
# Loại bỏ khoảng trắng và dấu \n trong các giá trị chuỗi
df_combined[df_combined.columns] = df_combined[df_combined.columns]/
.apply(lambda x: x.str.strip().str.replace("\n", " ") if x.dtype == "object" else x)

print(f"Đã gộp và xử lý DataFrame chung với {df_combined.shape[0]} dòng /
và {df_combined.shape[1]} cột.")
```

Hình 2.12. Code Python để làm sạch dữ liệu trong các DataFrame

Nguồn : Jupyter Notebook về xử lý dữ liệu (2024)

```
# Loại bỏ dấu * trước khi đổi tên cột
df_combined.columns = [col.replace("*", "").strip() /
                        for col in df_combined.columns]
# Đổi tên cột thành dạng sentence
df_combined.columns = [col.capitalize() for col in df_combined.columns]

# Nhóm các cột theo định dạng
columns_title = ["Phân loại", "Nhân viên phụ trách", "Nguồn vốn"]
columns_upper = ["Dự án", "Khách hàng", "Vnd"]
columns_number = ["Tổng giá trị da ước tính", "Tỷ lệ thành công kỹ vọng", "Jan", "Feb", "Mar", "Apr", "May",
                  "Jun", "Jul", "Aug", "Sep", "Oct", "Nov", "Dec"]
columns_datetime = ["Tháng bắt đầu", "Tháng kết thúc"]

# Áp dụng định dạng cho từng nhóm cột
df_combined[columns_title] = df_combined[columns_title].apply(lambda col: col.str.title())
df_combined[columns_upper] = df_combined[columns_upper].apply(lambda col: col.str.upper())
df_combined[columns_number] = df_combined[columns_number].replace({' ': ' ', '%': '%'}, regex=True).astype(float)
df_combined[columns_datetime] = df_combined[columns_datetime].apply(pd.to_datetime, errors='coerce')

df_combined
```

Hình 2.13. Code Python để chỉnh sửa các giá trị trong các DataFrame

Nguồn : Jupyter Notebook về xử lý dữ liệu (2024)

- Xuất file dữ liệu cuối cùng : từ tập DataFrame tổng hợp sau khi được làm sạch sẽ đưa về tập file Excel để dễ dàng xử lý cho các bước xây dựng model trực quan hay trao đổi giữa các phòng ban khác nhau.

```
# Xuất DataFrame ra file Excel
output_file_path = r"D:\Thực tập\Thực Hiện\Data\Tất cả VPKD.xlsx"
df_combined.to_excel(output_file_path, sheet_name='BM_01.Thông tin cơ hội dự án'/
                    , index=False)

print(f"Đã xuất file Excel tại {output_file_path}")
```

Hình 2.14. Code Python để xuất Dataframe ra file Excel mới

Nguồn : Jupyter Notebook về xử lý dữ liệu (2024)

Kết quả : Có 6 tập dữ liệu được thu thập từ các văn phòng kinh doanh ban đầu mỗi tập có 35 cột. Sau khi làm sạch và hợp lại, tập dữ liệu tổng hợp có 315 dòng và 35 cột. Các dữ liệu thiếu và sai số đã được khắc phục thông qua việc liên hệ trực tiếp với các văn phòng kinh doanh để điều chỉnh. Các ký tự không hợp lệ như ký tự xuống dòng hoặc ký tự đặc biệt đã được loại bỏ hoàn toàn. Đồng thời, các danh từ địa điểm được thống nhất về mặt ngữ nghĩa, và các từ viết tắt được chỉnh sửa theo một định dạng đồng nhất. Tập dữ liệu sau khi xử lý đã sẵn sàng được sử dụng trong mô hình Power BI.

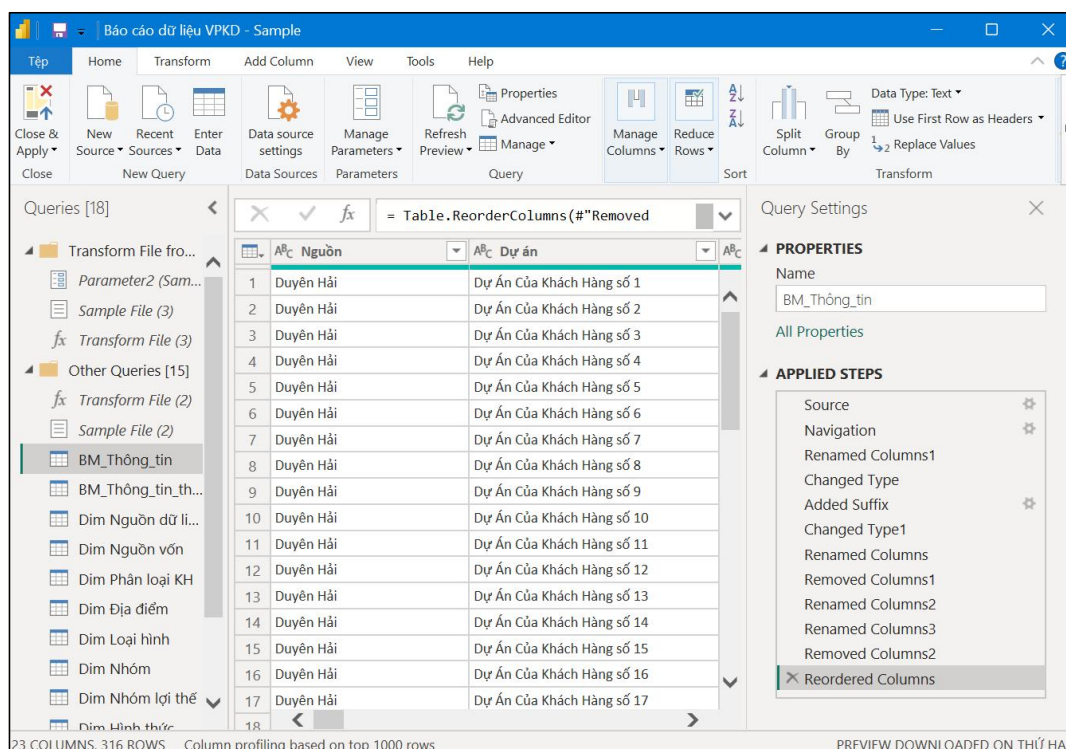
2.2.3. Power BI

2.2.3.1. Sử dụng Power BI để chuyển đổi dữ liệu

Doanh nghiệp sử dụng Power BI làm công cụ chính để hiển thị và phân tích dữ liệu từ hệ thống dữ liệu của công ty. Dữ liệu từ các máy chủ khác nhau, sau khi được tổng hợp và xử lý qua Python, sẽ được nhập vào Power BI để tạo các mô hình, biểu đồ và dashboard phù hợp với mục đích phân tích. Power BI không chỉ giúp trực quan hóa các phân tích dưới dạng biểu đồ và dashboard, mà còn tiếp tục xử lý dữ liệu đã được làm sạch qua Python, đảm bảo dữ liệu được chuyển đổi sao cho phù hợp với định dạng yêu cầu của công cụ này và hỗ trợ việc xây dựng các measure tính toán :

- Load và Transform dữ liệu vào model [8] :
 - Nhập dữ liệu từ file Excel vào Power BI.
 - Transform Data : Chỉnh sửa dữ liệu trong Power Query Editor.

- Dùng Transform để chỉnh sửa các cột cho phù hợp với định dạng của Power BI : Replace Values, Split Column, Merge Columns, Add Suffix, Change Type,....
- Sử dụng các bảng Dimension lưu trữ các giá trị duy nhất của từng cột dữ liệu hỗ trợ việc lọc và phân tích từng loại dữ liệu dễ dàng.



Hình 2.15. Thực hiện chuyển đổi dữ liệu trên Power Query Editor

Nguồn : Báo cáo cơ hội dự án của các Văn phòng kinh doanh (2024)

2.2.3.2. Trực quan hóa dữ liệu, tạo dashboard báo cáo cho các cơ hội dự án

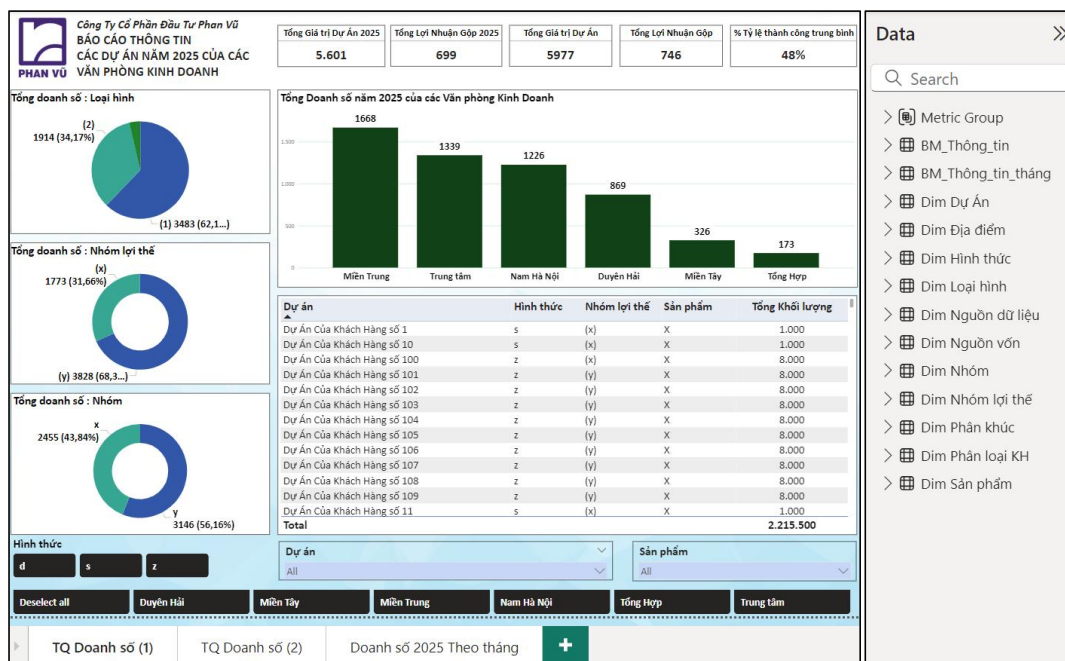
Dựa trên các file dữ liệu về cơ hội dự án được gửi từ các văn phòng kinh doanh, dữ liệu đã được xử lý và làm sạch bằng Python sẽ được sử dụng để xây dựng dashboard báo cáo trong Power BI. Dashboard này được thiết kế với cấu trúc linh hoạt, cho phép nhập dữ liệu mới từ file dữ liệu mà không cần chỉnh sửa phức tạp.

- Yêu cầu đối với dashboard :

- Đảm bảo dữ liệu từ 6 văn phòng kinh doanh đầy đủ và chính xác, chuẩn hóa kiểu dữ liệu để phù hợp với yêu cầu báo cáo doanh nghiệp. Sử dụng các bảng dimension để dễ dàng lọc và phân tích từng loại dữ liệu, tối ưu hóa trải nghiệm người dùng trong việc theo dõi thông tin. Áp dụng unpivot columns

vào dữ liệu sao chép để chuyển dữ liệu doanh số từ dạng ngang (theo tháng) sang dạng dọc, tạo điều kiện thuận lợi cho việc phân tích.

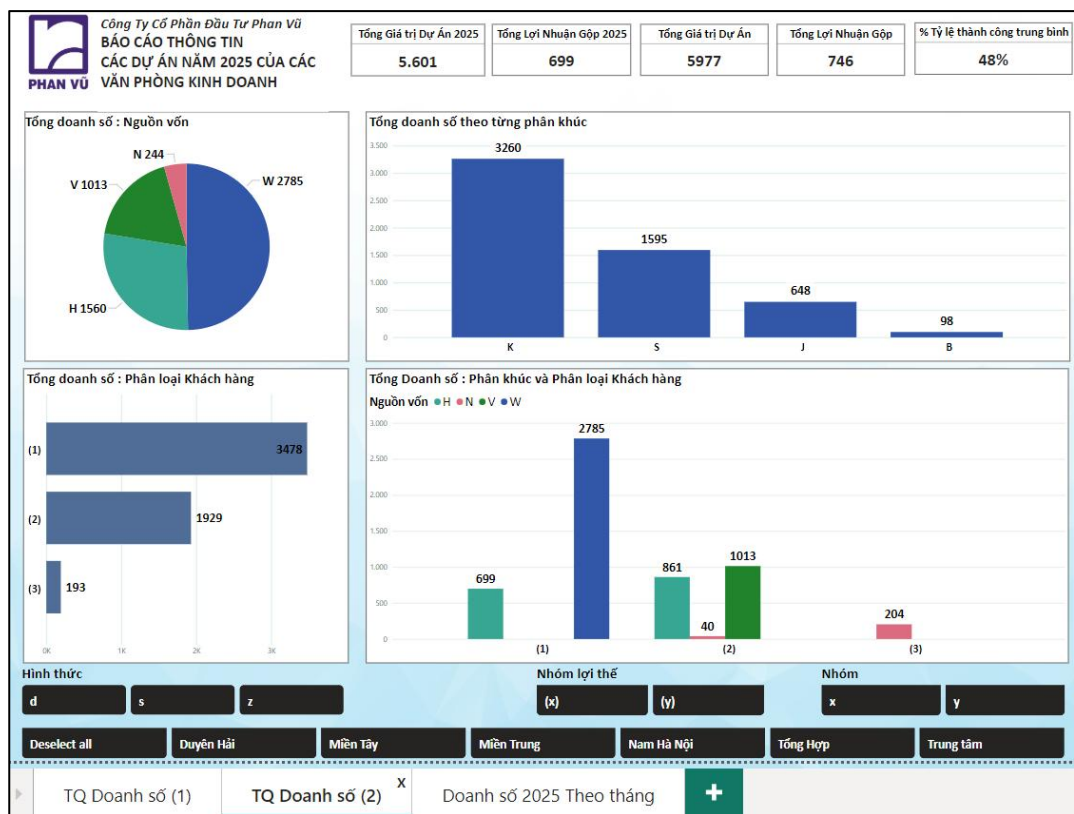
- Sử dụng calculation group để tổ chức các measure, giúp các chỉ số đo lường (KPIs) hiển thị một cách rõ ràng và tập trung.
- Đảm bảo dashboard có khả năng xử lý khối lượng dữ liệu lớn, giao diện dễ đọc, dễ hiểu và phù hợp với yêu cầu báo cáo doanh nghiệp.



Hình 2.16. Tổng quan Doanh số - Trang 1 của Dashboard

Nguồn : Báo cáo cơ hội dự án của các Văn phòng kinh doanh (2024)

- Dashboard bao gồm các biểu đồ trực quan thể hiện doanh số của các dự án theo chiều dọc (dữ liệu tổng hợp) và theo chiều ngang (doanh số từng tháng) :
- Dữ liệu doanh số theo chiều dọc : Dữ liệu doanh số theo chiều dọc là kiểu dữ liệu trong đó mỗi hàng biểu thị một quan sát riêng biệt và các thông tin liên quan đến quan sát này được lưu trữ trong các cột, vì vậy có thể dễ dàng lọc, nhóm hoặc phân tích theo các yếu tố như thời gian, khu vực, hay danh mục sản phẩm.



Hình 2.17. Tổng quan Doanh số theo chiều dọc - Trang 2 của Dashboard

Nguồn : Báo cáo cơ hội dự án của các Văn phòng kinh doanh (2024)

- Dữ liệu doanh số theo chiều ngang : Dữ liệu doanh số theo chiều ngang tổ chức thông tin theo các cột, với mỗi cột đại diện cho một nhóm thông tin riêng biệt, ví dụ như doanh thu của từng sản phẩm hoặc từng thời kỳ. Mỗi dòng trong bảng sẽ biểu thị doanh thu của tất cả các sản phẩm cho một khoảng thời gian cụ thể.



Hình 2.18. Tổng quan Doanh số theo tháng năm 2025 - Trang 3 của Dashboard

Nguồn : Báo cáo cơ hội dự án của các Văn phòng kinh doanh (2024)

- Chi tiết cách áp dụng các kỹ thuật trong Dashboard :
 - Bảng số liệu (Table) được thiết kế để hiển thị thông tin kết hợp nhiều tiêu chí, hỗ trợ người dùng tra cứu dữ liệu cụ thể. Với bảng số liệu ở hình 2.16 có thể giúp người dùng xem thông tin như tên dự án, tổng khối lượng sản phẩm, hình thức dự án và số sản phẩm đã đặt hàng.
 - Biểu đồ tròn được sử dụng để hiển thị tỷ trọng của từng mục so với tổng thể. Hình 2.16 có tỷ lệ phân phối doanh số theo loại hình cũng được minh họa giúp người dùng nắm bắt rõ tỷ trọng của từng nhóm. Và hình 2.17 có tỷ lệ doanh thu theo nguồn vốn, trong đó mỗi phần của biểu đồ đại diện cho một loại nguồn vốn, cung cấp cái nhìn trực quan về mức độ đóng góp của từng loại vào tổng doanh số.
 - Biểu đồ Donut : là cải tiến của biểu đồ tròn, mang lại sự trực quan hơn nhờ khoảng trống trung tâm. Hình 2.17 có phân phối doanh số theo nhóm lợi thế,

trong đó tỷ lệ doanh số từ từng nhóm được thể hiện rõ ràng, giúp đánh giá mức độ đóng góp của mỗi nhóm vào tổng doanh số.

- Biểu đồ cột : được sử dụng để minh họa và so sánh dữ liệu theo các danh mục hoặc mốc thời gian, với hai cách thể hiện doanh số :
 - Doanh số theo chiều dọc : các cột được sắp xếp theo trục đứng, giúp người xem dễ dàng so sánh giá trị giữa các danh mục như biểu đồ doanh số theo nguồn vốn ở hình 2.17 với mỗi cột tương ứng với từng loại nguồn vốn. Ngoài ra còn có biểu đồ doanh số theo phân loại khách hàng hình 2.18 với mỗi cột đại diện cho một nhóm khách hàng, giúp so sánh hiệu quả doanh số giữa các nhóm với nhau.
 - Doanh số theo chiều ngang : các cột được sắp xếp theo trục ngang, thường được sử dụng để thể hiện dữ liệu thời gian hoặc phân nhóm cụ thể như doanh số từng tháng trong năm 2025 ở hình 2.18 với các cột nằm ngang biểu diễn doanh số từng tháng, được phân tách theo văn phòng và nhóm kinh doanh, giúp dễ dàng so sánh giữa các tháng và đánh giá xu hướng thời gian.
- Biểu đồ đường : thể hiện sự biến đổi theo thời gian, giúp nhận diện xu hướng một cách rõ ràng. Trong hình 2.18, biểu đồ doanh số qua từng tháng cho thấy xu hướng tổng thể, từ tăng trưởng đều đặn đến những biến động, giúp người dùng phân tích các thay đổi quan trọng.
- Biểu đồ đường kết hợp cột là kỹ thuật minh họa cả xu hướng và giá trị cụ thể tại từng mốc thời gian. Ở hình 2.18 có biểu đồ doanh số theo thời gian, với đường biểu đồ thể hiện sự tăng giảm theo từng tháng, trong khi cột biểu đồ biểu diễn giá trị cụ thể tại các điểm. Sự kết hợp này giúp người xem nắm rõ cả xu hướng và chi tiết từng giai đoạn.
- Slicer (Công cụ lọc) : hỗ trợ lọc dữ liệu nhanh chóng theo nhiều tiêu chí, mang lại sự tiện lợi cho người dùng. Ví dụ, lọc dữ liệu theo văn phòng kinh doanh, theo loại dự án, hoặc theo thời gian từng tháng hoặc giai đoạn cụ thể trong năm.