

**ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

❖ ❖ ❖



**ÁP DỤNG QUY TRÌNH OSEMN  
PHÂN TÍCH THĂM DÒ BỘ DỮ LIỆU  
“MASS DISASTERS IN VIETNAM 1953-2024”**

Nhóm 6			
Sinh viên thực hiện:			
STT	Họ tên	MSSV	Ngành
1	Bùi Quốc Bảo	23520091	KHDL
2	Lê Minh Khôi	23520767	KHDL
3	Huỳnh Phát Đạt	24520270	KHDL

**TP. HỒ CHÍ MINH – 12/2025**

## 1. GIỚI THIỆU

Đồ án này tập trung thực hiện phân tích dữ liệu thăm dò trên bộ dữ liệu “Mass Disasters in Vietnam 1953 – 2024” nhằm nhận diện các đặc điểm cốt lõi và đánh giá thiệt hại của thảm họa tại Việt Nam. Quy trình phân tích được triển khai theo hai trực chính là đặc điểm hình thái thảm họa và các chỉ số về thiệt hại kinh tế - xã hội. Bằng cách sử dụng ngôn ngữ Python cùng các thư viện chuyên dụng như Pandas, Matplotlib và Seaborn, chúng tôi tiến hành trực quan hóa dữ liệu để trích xuất những thông tin giá trị (insight). Kết quả đạt được không chỉ là những hiểu biết sâu sắc về quy luật biến động thiên tai qua các thời kỳ mà còn là hệ thống Dashboard trực quan, giúp truyền tải các phát hiện quan trọng đến mọi người một cách sinh động và hiệu quả.

Trong đề tài này, bộ dữ liệu phân tích được chúng tôi lấy tại link: [SOURCE](#) và tham khảo quy trình phân tích tại: [URL](#), song chúng tôi cam kết toàn bộ nội dung phân tích và thiết kế hệ thống đều do nhóm tự thực hiện. Khác với phân tích trước đây vốn chỉ dừng lại ở việc thống kê hình thái và thiệt hại cơ bản, chúng tôi đã đi sâu khai thác các insight ẩn còn tồn tại nhằm khai thác tất cả mọi khía cạnh bộ dữ liệu. Nhờ việc khai thác sâu này mà chúng tôi còn khám phá ra được khía cạnh kinh tế bằng cách phân tích sự tương quan giữa thiên tai và biến động lạm phát. Việc khám phá ra thêm những insight này giúp mang lại một cái nhìn toàn diện, hữu ích và có ý nghĩa thực tiễn cao hơn cho toàn bộ đề tài phân tích này của chúng tôi.

## 2. MÔ TẢ BỘ DỮ LIỆU

“Mass disasters in Vietnam from 1953 to 2024” là một bộ dữ liệu chứa thông tin về các thiên tai ở Việt Nam trong khoảng từ năm 1953 đến 2024 được trích xuất từ Cơ sở dữ liệu Thảm họa Quốc tế (EM-DAT), được duy trì bởi Trung tâm Nghiên cứu Dịch tễ học Thảm họa (CRED), do Open Development Việt Nam xuất bản.

Bộ dữ liệu ban đầu có tổng cộng 335 dòng và 46 cột nhưng qua xử lý chỉ còn 21 cột để phục vụ phân tích, trong đó có 7 biến phân loại và 14 biến số.

STT	Tên cột	Kiểu dữ liệu	Giải thích
1	Disaster Group	object	Nhóm lớn loại thảm họa
2	Disaster Subgroup	object	Nhóm phụ của loại thảm họa
3	Disaster Type	object	Loại thảm họa
4	Disaster Subtype	object	Loại thảm họa cụ thể
5	Location	object	Tên địa phương

6	Origin	object	Các yếu tố bối cảnh trước khi xảy ra thảm họa
7	Associated Types	object	Loại thảm họa kèm theo hoặc phát sinh
8	Magnitude	float	Cường độ thảm họa
9	Total Deaths	float	Tổng số người chết/mất tích
10	No. Injured	float	Số người bị thương
11	No. Affected	float	Số người bị ảnh hưởng trực tiếp
12	Total Affected	float	Tổng số người bị ảnh hưởng
13	Total Damage ('000 US\$)	float	Tổng thiệt hại về kinh tế
14	Total Damage, Adjusted ('000 US\$)	float	Tổng thiệt hại về kinh tế (điều chỉnh theo lạm phát)
15	CPI	float	Chỉ số giá tiêu dùng
16	Start/End + Day/month/Year	float/int	Ngày/Tháng/Năm xảy ra và kết thúc thảm họa

Bảng 1. Mô tả bộ dữ liệu

### 3. QUY TRÌNH PHÂN TÍCH



Hình 1. Quy trình phân tích dữ liệu

#### 3.1. Chuẩn bị bộ dữ liệu (Data Source):

Thông qua nhiều nguồn khác nhau, chúng tôi tìm được bộ dữ liệu về các thảm họa lớn được ghi nhận tại Việt Nam từ năm 1953 đến năm 2024. Đối với chúng tôi,

bộ dữ liệu này rất gần gũi vì nó ghi nhận những sự kiện thảm họa gây ra nhiều thiệt hại cho đất nước. Chính vì vậy, việc khai thác và phân tích bộ dữ liệu này sẽ mang lại nhiều góc nhìn đáng giá và thiết thực đối với không chỉ riêng chúng tôi mà còn cho toàn thể mọi người đang sinh sống tại Việt Nam.

### **3.2. Xác định vấn đề (Define Problem):**

Thông qua quan sát sơ lược về bộ dữ liệu, chúng tôi xác định hai nội dung chính cần thực hiện phân tích chính là đặc điểm hình thái của các thảm họa và thiệt hại do chúng gây ra. Dựa trên hai nội dung này, chúng tôi sẽ rút ra các insights hàm chứa trong dữ liệu. Bên cạnh đó, trong quá trình thực hiện phân tích chúng tôi còn khám phá được thêm các khía cạnh về kinh tế, đặc biệt là trong lĩnh vực tiền tệ.

### **3.3. Làm sạch bộ dữ liệu (Data cleaning):**

Đầu tiên, chúng tôi tiến hành kiểm tra và loại bỏ các đặc trưng không phù hợp để đảm bảo tính hiệu quả cho quá trình phân tích dữ liệu. Những yếu tố bị loại bỏ bao gồm các cột bị khuyết dữ liệu tuyệt đối (100%), các đặc trưng chỉ chứa một giá trị duy nhất, hoặc các biến phân loại dư thừa do được kết hợp từ các đặc trưng khác. Sau quá trình thảo luận và chọn lọc kỹ lưỡng, bộ dữ liệu đã được tinh gọn từ 46 cột ban đầu xuống còn 30 cột mang giá trị sử dụng cao nhất để phục vụ cho việc nghiên cứu chuyên sâu.

### **3.4. Xử lý đặc trưng (Feature Engineering):**

Đầu tiên, do trong dữ liệu về thời gian như “Ngày-Tháng-Năm” trong bộ dữ liệu ban đầu được tách thành các đặc trưng riêng biệt nhau nên sẽ tiến hành tạo đặc trưng các đặc trưng mới bằng cách tổng ghép các cột dữ liệu riêng biệt nhau về “Ngày-Tháng-Năm” lại với nhau thành một đặc trưng có kiểu dữ liệu là datetime.

Tiếp theo, chúng tôi tiến hành điền khuyết cho dữ liệu, các đặc trưng được điền khuyết bao gồm “Total Affected”, “Total Deaths”, “Total Damage, Adjusted ('000 US\$)”, “Total Damage ('000 US\$)”. Các giá trị bị khuyết sẽ được điền bằng giá trị 0 do đây là các đặc trưng dạng số, và mang ý nghĩa định lượng. Ngoài ra, một thảm họa xảy ra có thể không gây ra thiệt hại về thương vong, tổn thất tài sản hoặc thiệt hại quá nhỏ để được thống kê nên việc điền cố định bằng giá trị 0 sẽ giúp đảm bảo không làm ảnh hưởng đến các dữ liệu khác và không phóng đại mức độ nghiêm trọng của thảm họa.

Cuối cùng, chúng tôi sử dụng sự trợ giúp của LLMs (Gemini) để tạo đặc trưng mới là “economic\_region” (vị trí của các vùng kinh tế-xã hội trọng điểm của Việt Nam xảy ra các thảm họa) dựa vào đặc trưng “Location”.

### **3.5. Khám phá và phân tích chuyên sâu (EDA):**

Khám phá và phân tích chuyên sâu là bước thực hiện then chốt giúp chúng tôi bóc tách được những thông tin giá trị từ bộ dữ liệu. Trực quan hóa dữ liệu giúp chúng tôi tiếp cận thông tin một cách trực diện thông qua các biểu đồ xu hướng và

phân bố, từ đó những khía cạnh ý nghĩa được xác định. Ngoài ra, các phương pháp kiểm định thống kê cũng được áp dụng với một số đặc trưng nhằm kiểm tra ý nghĩa hàm chứa trong dữ liệu.

### 3.6. Rút ra thông tin giá trị (Insights):

Dựa trên những kết quả đã được khám phá, phân tích, trực quan và đánh giá, chúng tôi đã rút ra được các thông tin có giá trị về tình hình xảy ra của các thảm họa tại Việt Nam trong giai đoạn 1953 đến 2024.

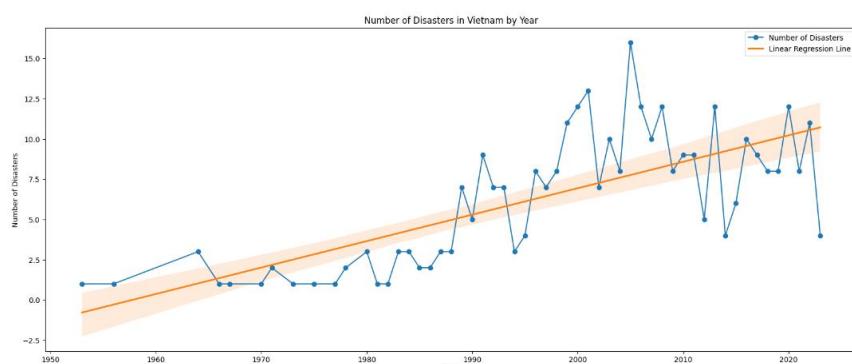
### 3.7. Truyền tải kết quả (DashBoard):

Để truyền tải các thông tin quan trọng (insight) đã được phát hiện trong quá trình khám phá và phân tích dữ liệu, nhóm chúng tôi đã xây dựng một Dashboard trên nền tảng Power BI. Đây là bước chuyển đổi các kết quả phân tích thành một giao diện trực quan, có thể tương tác.

## 4. PHÂN TÍCH ĐẶC ĐIỂM HÌNH THÁI CỦA THẨM HỌA

### 4.1. Xu hướng của thảm họa

Để có cái nhìn tổng quan về diễn biến thảm họa tại Việt Nam, *biểu đồ 1* dưới đây mô tả số lượng các vụ thảm họa xảy ra hàng năm từ năm 1953 đến 2024.

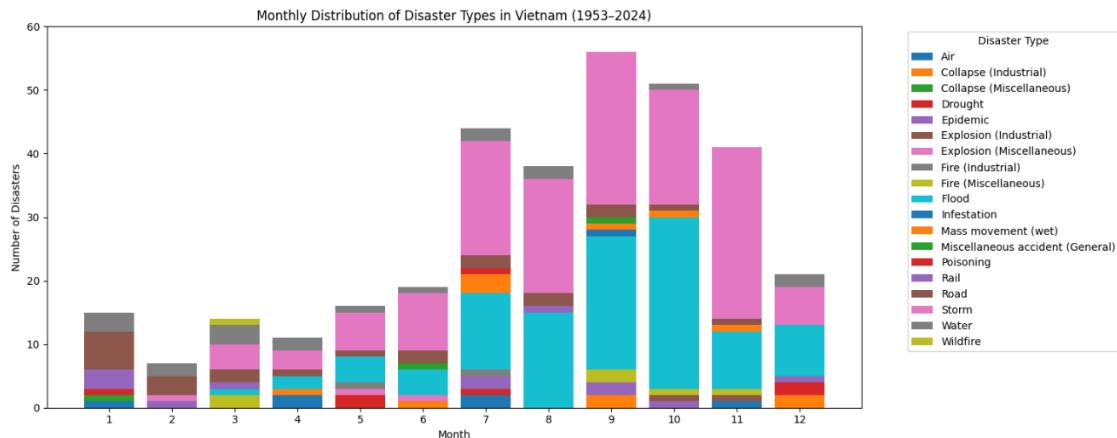


*Biểu đồ 1. Số lượng thảm họa tại Việt Nam qua từng năm từ 1953 - 2024*

Số lượng thảm họa tại Việt Nam từ năm 1953 đến 2024 có xu hướng tăng trưởng dài hạn rõ rệt với đường hồi quy tuyến tính có độ dốc đi lên. Dữ liệu cho thấy một sự phân hóa mạnh mẽ về tần suất theo thời gian, giai đoạn 1953 - 1989 số vụ thảm họa duy trì ở mức rất thấp và ổn định (thường từ 1 đến 3 vụ/năm), nhưng kể từ sau năm 1990, con số này đã bùng nổ và thường xuyên dao động ở ngưỡng cao từ 7 đến 13 vụ, đạt đỉnh điểm là khoảng 16 vụ vào năm 2005.Thêm vào đó, biên độ dao động của biểu đồ đường ngày càng lớn trong những năm gần đây, phản ánh tính chất bất ổn và khó dự đoán của thảm họa.

### 4.2. Thời gian thảm họa xảy ra trong năm

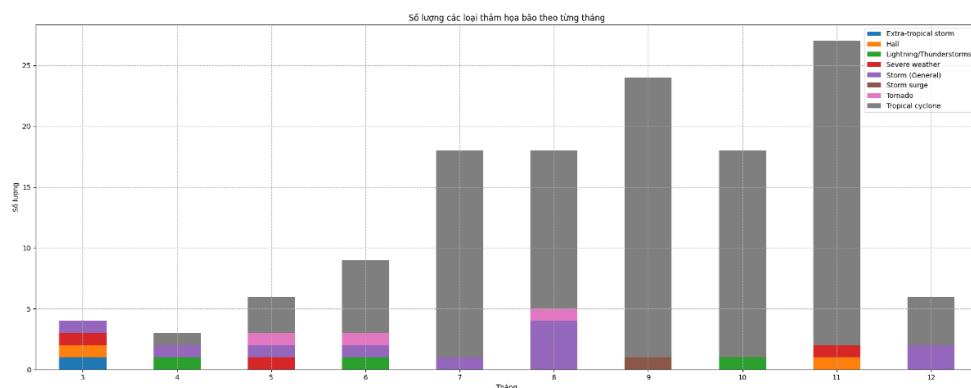
Sự phân bố của các loại hình thảm họa có sự khác biệt rõ rệt giữa các tháng trong năm, chi tiết về số lượng và thành phần các loại thảm họa được thể hiện cụ thể tại *biểu đồ 2*.



Biểu đồ 2. Số lượng thảm họa tại Việt Nam theo tháng từ 1953 - 2024

Biểu đồ cột chồng về sự phân bố thảm họa theo tháng tại Việt Nam cho thấy một quy luật mùa vụ cực kỳ rõ rệt, với sự tập trung cao độ của các biến cố vào nửa cuối năm, đặc biệt là giai đoạn từ tháng 7 đến tháng 11. Số lượng thảm họa đạt đỉnh điểm vào tháng 9 với hơn 55 vụ được ghi nhận, trong khi tháng 2 là thời điểm có tần suất thấp nhất với chỉ khoảng 7 vụ. Hai loại thảm họa chiếm tỷ trọng áp đảo và chi phối hoàn toàn xu hướng biểu đồ là "Storm" (Bão) và "Flood" (Lũ lụt), phản ánh trực tiếp tác động của mùa bão lũ tại Việt Nam. Các loại thảm họa khác xuất hiện rải rác và ít có sự biến động lớn theo mùa hơn.

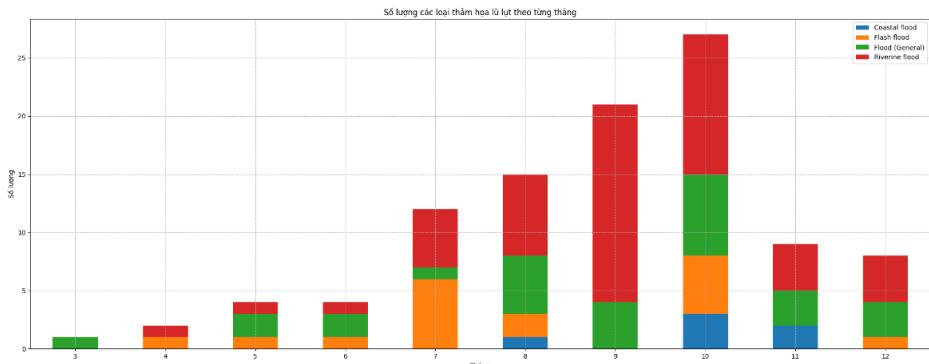
Do nhận thấy loại hình thiên tai bão (Storm) và lũ lụt (Flood) xảy rất phổ biến ở Việt Nam nên chúng tôi sẽ tiến hành phân tích sâu hơn về 2 loại hình thiên tai này. Đầu tiên và về bão (Storm), chúng tôi tiến hành quan phân bố của chúng theo tiến trình thời gian và các loại hình bão.



Biểu đồ 3: Phân bố các loại bão theo thời gian và mùa vụ

Biểu đồ cho thấy tần suất gia tăng dần đều của bão theo thời gian từ đầu năm đến khoảng giữa cuối năm. Các cơn bão thường xuất hiện vào khoảng tháng 6 trong năm, giai đoạn cao điểm thường kéo dài từ tháng 7 đến tháng 11 và đạt đỉnh điểm vào tháng 11. Bão nhiệt đới (Tropical storm) chiếm tỉ trọng áp đảo trong năm và đã xuất hiện ở hầu hết các tháng trong năm cho thấy mức độ ảnh hưởng rộng rãi của loại bão này.

Tiếp theo, chúng tôi tiến hành quan phân bố của thảm họa lũ lụt theo tiến trình thời gian và các loại hình lũ lụt.

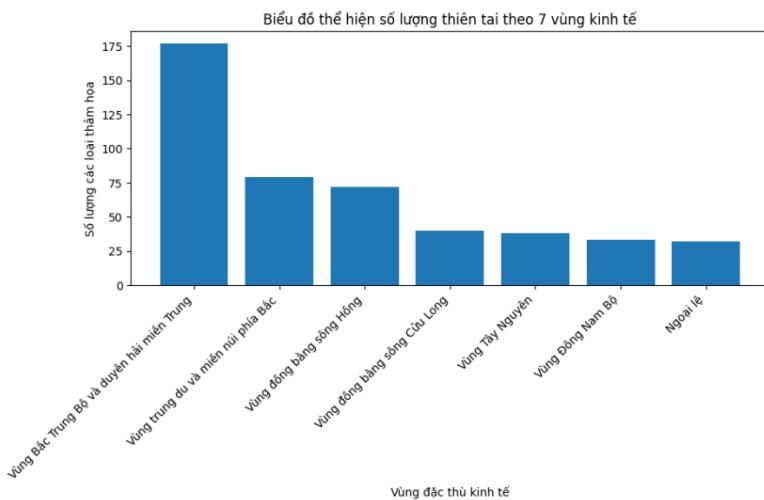


Biểu đồ 4: Phân bố các loại lũ theo thời gian và mùa vụ

Biểu đồ cho thấy các sự kiện lũ tại Việt Nam thường bắt đầu từ tháng 7 và kéo dài đến các tháng cuối năm, đạt đỉnh điểm vào hai tháng 9 và 10. Các cơn lũ xảy ra thường xuất hiện dưới dạng bão do nước sông dâng lên (Riverine Flood) và các cơn lũ bình thường (Flood general) thường do nước ở thượng nguồn chảy về hoặc là các cơn lũ chưa được phân loại. Cần chú ý sự xuất hiện của lũ quét vào các tháng 7, 8 và tháng 10 để đảm bảo công tác phòng chống và ứng cứu kịp thời.

### 4.3. Vùng thảm họa

Để đánh giá mức độ ảnh hưởng của thiên tai trên các vùng lãnh thổ khác nhau, *biểu đồ 5* thống kê tổng số lượng thảm họa xảy ra tại 7 vùng kinh tế trọng điểm của Việt Nam.



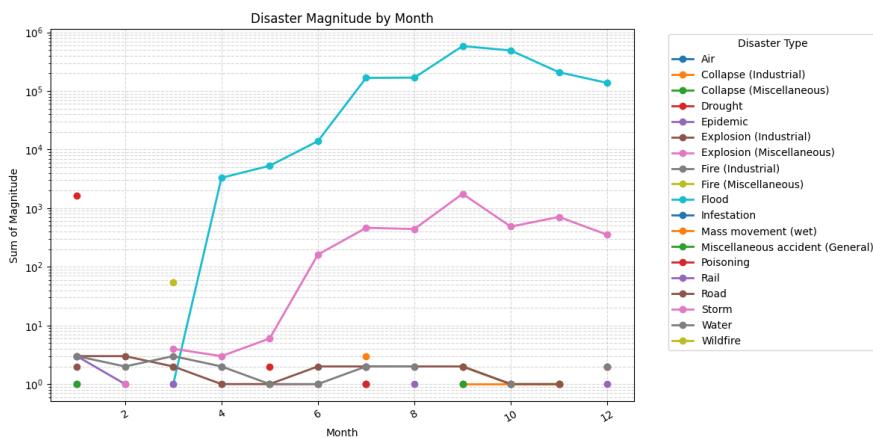
Biểu đồ 5. Số lượng thảm họa theo các vùng kinh tế Việt Nam từ 1953 - 2024

Biểu đồ cột minh họa sự chênh lệch lớn về mức độ rủi ro giữa các vùng kinh tế Việt Nam. Trong đó, Bắc Trung Bộ và duyên hải miền Trung chịu ảnh hưởng nặng nề nhất với hơn 175 vụ, cao gấp đôi vùng xếp thứ hai là Trung du và miền núi phía Bắc với khoảng 80 vụ. Ngược lại, Đông Nam Bộ và nhóm Ngoại lề ghi nhận

tần suất thấp nhất, chỉ dao động quanh mức 30 vụ. Sự tập trung thảm họa tại dải miền Trung xuất phát từ đặc thù địa lý giáp biển cùng địa hình hẹp và dốc, khiến nơi đây trở thành tâm điểm của bão lũ. Điều này cho thấy công tác quy hoạch hạ tầng và dự phòng nguồn lực quốc gia cần được ưu tiên đặc biệt cho khu vực miền Trung và miền núi phía Bắc để giảm thiểu tối đa thiệt hại.

#### 4.4. Cường độ thảm họa

Nhằm phân tích mức độ khốc liệt của từng loại hình thiên tai, *biểu đồ 6* thể hiện cường độ của các thảm họa theo từng tháng, giúp làm nổi bật xu hướng về cường độ của các loại thảm họa trong năm



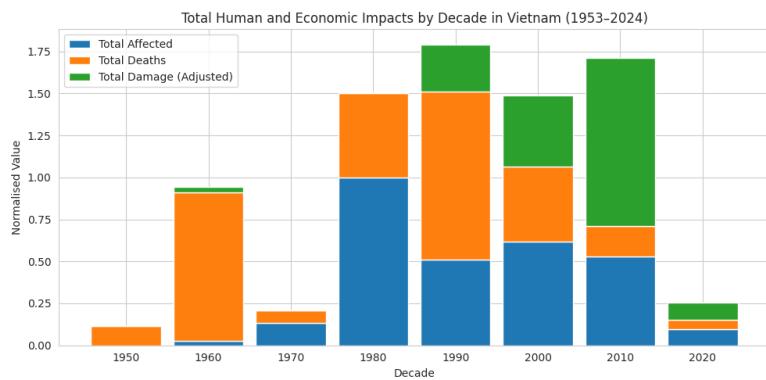
Biểu đồ 6. Cường độ của các loại thảm họa theo từng tháng

Biểu đồ đường về tổng cường độ của các loại thảm họa theo tháng sử dụng thang đo logarit cho thấy sự áp đảo hoàn toàn về cường độ hoạt động của lũ lụt (Flood) và bão (Storm) so với các loại hình thảm họa còn lại. Tổng cường độ thảm họa đạt đỉnh điểm vào tháng 9, trong đó lũ lụt chiếm vị thế độc tôn với giá khoảng  $10^6$ , trong khi tháng 2 và tháng 3 ghi nhận mức cường độ thấp nhất đối với các loại hình thảm họa diễn ra theo mùa. Một xu hướng chuyển biến mạnh mẽ được ghi nhận từ tháng 4 trở đi khi cường độ lũ lụt bắt đầu tăng mạnh và duy trì sức ảnh hưởng cực lớn cho đến hết năm. Qua đó có thể thấy là dù các vụ tai nạn công nghiệp hay hỏa hoạn có thể xảy ra rải rác, nhưng bão và lũ lụt mới chính là yếu tố gây ra tác động hủy diệt khủng khiếp nhất, gấp nhiều lần các loại hình khác. Điều này đòi hỏi các kế hoạch quản lý rủi ro phải ưu tiên tuyệt đối nguồn lực cho việc kiểm soát bão lũ và gia cố hạ tầng thoát nước, đặc biệt là trong giai đoạn cao điểm từ tháng 7 đến tháng 11.

### 5. THIỆT HẠI

#### 5.1. Thiệt hại theo năm

Để đánh giá hệ quả của thiên tai đối với xã hội và nền kinh tế, *biểu đồ 7* minh họa tổng mức độ tác động về người và tài sản tại Việt Nam qua từng thập kỷ, từ năm 1953 đến 2024.

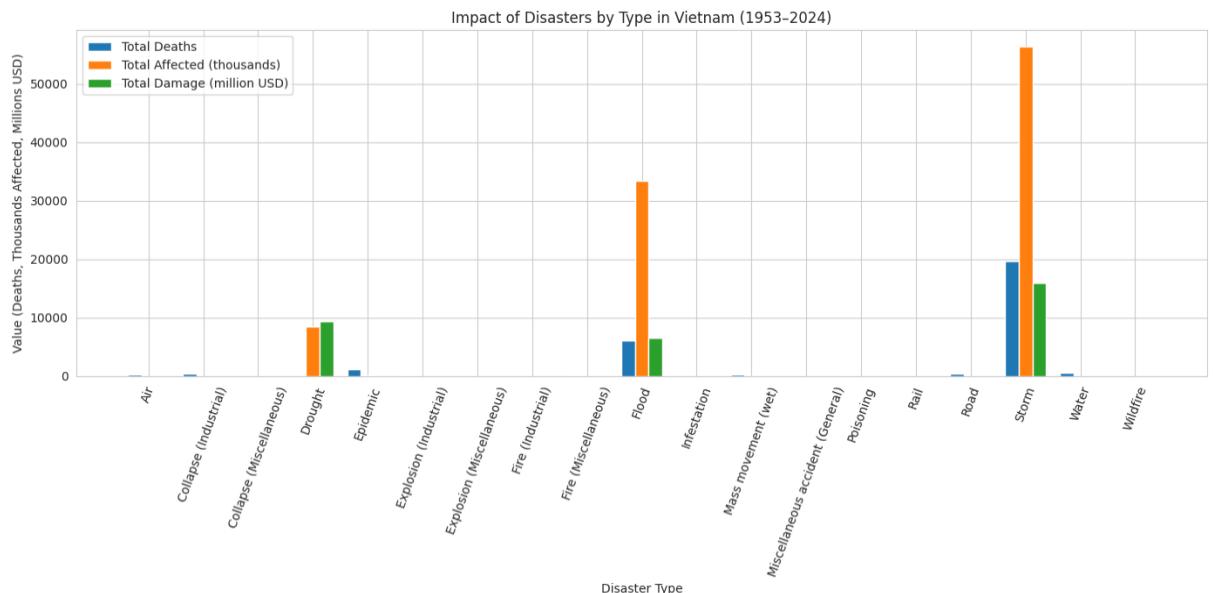


*Biểu đồ 7. Tổng thiệt hại về người và kinh tế theo thập kỷ từ 1953- 2024*

Biểu đồ cột chồng minh họa sự dịch chuyển rõ rệt về tác động của thảm họa qua từng thập kỷ. Trong khi thập niên 1950 ghi nhận chỉ số thấp nhất do hạn chế về dữ liệu và quy mô kinh tế, tổng mức độ thiệt hại đã đạt đỉnh điểm vào thập niên 1990. Đáng chú ý là sự thay đổi trong cấu trúc thiệt hại rằng là tỉ lệ tử vong chiếm ưu thế vào thập niên 1960, thì từ sau năm 1990, thiệt hại kinh tế bắt đầu tăng vọt và đạt mức cao nhất vào thập niên 2010. Bên cạnh đó, số người bị ảnh hưởng cũng ghi nhận mức bùng nổ vào thập niên 1980. Kết quả này phản ánh rằng dù công tác dự báo đã giúp giảm thiểu đáng kể thương vong, nhưng quá trình đô thị hóa và phát triển kinh tế lại khiến giá trị tài sản chịu rủi ro tăng mạnh, dẫn đến thiệt hại tài chính ngày càng nặng nề trong những năm gần đây.

## 5.2. Thiệt hại theo loại hình thảm họa

Để xác định những loại hình thiên tai gây thiệt hại nghiêm trọng nhất, *biểu đồ 8* so sánh tác động về người và kinh tế giữa các loại thảm họa khác nhau tại Việt Nam sau khi đã chuẩn hóa dữ liệu.



*Biểu đồ 8. Thiệt hại theo từng loại thảm họa*

Biểu đồ cột ghép minh họa sự chênh lệch không lồ về mức độ tàn phá giữa các loại hình thảm họa. Trong đó, bão (Storm) chiếm vị trí dẫn đầu tuyệt đối ở cả ba chỉ số là số người bị ảnh hưởng, số người thiệt mạng và thiệt hại kinh tế. Đứng thứ hai là lũ lụt (Flood) với tác động đáng kể trên mọi phương diện, còn hạn hán (Drought) bộc lộ đặc điểm riêng biệt khi gây thiệt hại kinh tế rất cao dù thương vong trực tiếp thấp hơn nhiều so với bão. Ngược lại, các loại hình thảm họa như tai nạn kỹ thuật hay giao thông ghi nhận mức tác động gần như không đáng kể khi so sánh với các thiên tai lớn. Qua đó, có thể thấy Việt Nam cần ưu tiên tập trung bảo vệ tính mạng trong bão lũ và xây dựng các phương pháp bảo vệ tài sản để đối phó hiệu quả với tính chất gây thiệt hại kinh tế của hạn hán.

## 6. KẾT QUẢ PHÂN TÍCH

Kết quả phân tích bộ dữ liệu ghi nhận lại các thảm họa tại Việt Nam giai đoạn 1953 – 2024 cho thấy xu hướng gia tăng rõ rệt về tần suất xảy ra của các loại thảm họa, đặc biệt là trong giai đoạn sau năm 1990. Về mặt hình thức, Các loại thảm họa xảy ra dưới nhiều hình thức khác nhau từ thảm họa tự nhiên đến thảm họa do con người và phân bố rải rác ở khắp các khu vực trên Việt Nam, trong đó hai loại hình thảm họa bão và lũ chiếm số lượng áp đảo cả về tần suất xảy ra, cường độ tàn phá lẫn thiệt gây ra, chủ yếu tập trung vào khoảng thời gian nửa cuối năm (từ tháng 7 đến 11) và thường đạt đỉnh điểm vào hai tháng 9,10. Về mặt địa lý, khu vực Bắc Trung Bộ và duyên hải miền Trung được xác định là vùng trọng điểm thu hút nhiều thiên tai nhất và chịu ảnh hưởng nặng nề nhất. Về mặt thiệt hại, sự chuyển dịch quan trọng từ thiệt hại về người sang thiệt hại về kinh tế. Tuy nhiên, tỉ lệ tử vong giảm đáng kể cho thấy công tác dự báo và ứng phó đã được cải thiện hơn thời gian trước rất nhiều, nhưng vì quá trình phát triển kinh tế và đô thị hóa ngày càng gia tăng khiến dẫn đến thiệt hại kinh tế ngày càng trầm trọng trong khoảng thời gian từ sau 1990.

Nhằm truyền tải các insight quan trọng và tiêu biểu nhất về đặc điểm hình thái của thảm họa và thiệt hại do chúng gây ra cho kinh tế và xã hội Việt Nam, chúng tôi đã xây dựng hệ thống Dashboard trực quan trên nền tảng PowerBI như *hình 2*. Để xem và tương tác trực tiếp với Dashboard, có thể nhấn vào [link](#).



Hình 2. Dashboard

## 7. KẾT LUẬN

Đồ án phân tích thăm dò toàn diện trên bộ dữ liệu “Mass Disasters in Vietnam 1953 – 2024” đã giúp chúng tôi bóc tách và thu nhận được những thông tin vô cùng giá trị về các khía cạnh hình thái của thảm họa và thiệt hại do chúng gây ra tại Việt Nam trong khoảng hơn 70 năm qua. Thông qua quy trình thực hiện bài bản, từ khâu tiền xử lý dữ liệu, xử lý đặc trưng cho đến phân tích chuyên sâu, nhóm đã bóc tách thành công những quy luật ẩn chứa sâu trong dữ liệu. Việc kết hợp sức mạnh tính toán của ngôn ngữ Python với khả năng trực quan hóa sinh động của Power BI không chỉ giúp nhóm đạt được những hiểu biết sâu sắc về sự biến động của thiên tai mà còn tạo ra một hệ thống Dashboard hiệu quả, giúp truyền tải thông tin một cách trực diện và dễ tiếp cận đối với mọi người.

Nhìn lại quá trình thực hiện, những phát hiện về sự gia tăng tần suất thảm họa sau năm 1990, tính vụ khắc nghiệt của bão lũ tại miền Trung, hay sự chuyển dịch từ thiệt hại về người sang thiệt hại về kinh tế đã minh chứng cho tính thực tiễn cao của đề tài. Những kết quả này không chỉ đơn thuần là các con số thống kê, mà còn là cơ sở khoa học quan trọng hỗ trợ cho công tác quan sát, dự báo và tối ưu hóa việc phân bổ nguồn lực ứng phó thiên tai tại Việt Nam. Đồ án đã thực sự mang lại một cái nhìn toàn diện hơn để cộng đồng và các nhà quản lý có thêm góc nhìn khách quan để xây dựng các phương án phòng chống thiên tai bền vững hơn trong tương lai.

## TÀI LIỆU THAM KHẢO

- [1] Mass disasters in Vietnam from 1953 to 2024. Link:  
<https://www.kaggle.com/datasets/patricklford/mass-disasters-in-vietnam-from-1900-to-2024> (23/11/2025)

## PHỤ LỤC PHÂN CÔNG NHIỆM VỤ

STT	Thành viên	Nhiệm vụ	
1	Bùi Quốc Bảo	Viết báo cáo từ phần 1 đến 3 + Làm phần tiền xử lý dữ liệu + Code phân tích các đặc điểm hình thái thảm họa	Vẽ Dashboard
2	Lê Minh Khôi	Làm slide thuyết trình + tạo đặc trưng mới + Code phân tích thiệt hại của thảm họa	
3	Huỳnh Phát Đạt	Viết báo cáo từ phần 4 đến 7 + Code phân tích các đặc điểm kinh tế + Phân tích các đặc điểm hình thái của thảm họa	