Analysis Report: Assessing the impacts of extreme weather events on the Socioeconomics – A case study in Southeast Asian region

1 Introduction

In the last few decades, Southeast Asia has emerged as a region of significant economic development. The region however faces a daunting challenge as it is among those most vulnerable to the impacts of climate change and extreme weather events[1] . Different factors contribute to this vulnerability, including low-lying land that is susceptible to rising sea levels, frequent floods and droughts, large populations, heavy reliance on agriculture for economic stability, and limited community resilience to climate change. Some of the most extreme weather events have been experienced worldwide in emerging Southeast Asian nations such as the Philippines, Thailand, Indonesia, and Vietnam. This data project aims to provide insights into how extreme weather events have affected the socioeconomics of this region. Two main questions will be analyzed by examining underlying patterns in the datasets:

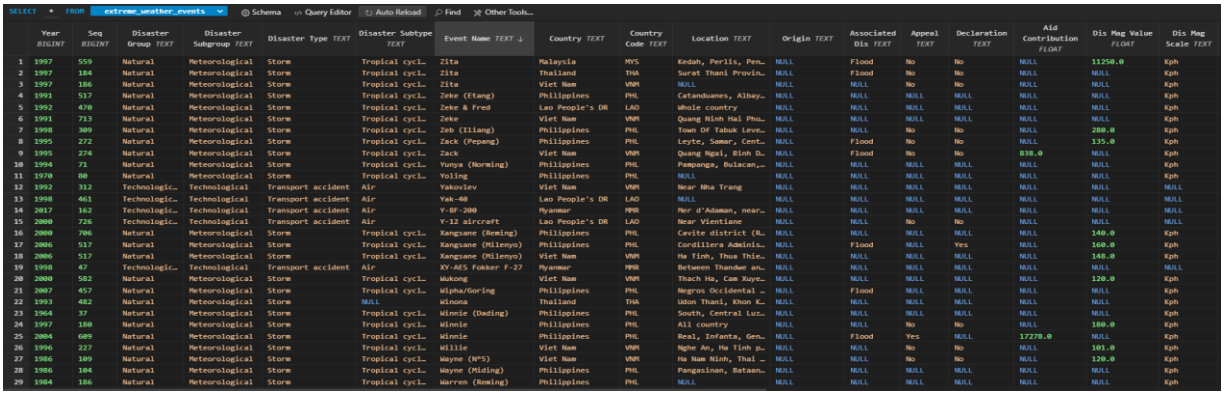
1. "What are the patterns in the occurrence of extreme weather events and the socioeconomic status of Southeast Asian countries?"
2. "How do disaster risks impact the socioeconomic landscape of Southeast Asian countries?"

2 Pipeline output

The ETL pipeline produces a structured SQLite database containing three tables *extreme_weather_events, socioeconomics, disaster_risk* . The tabular data can be retrieved through SQL execution that selects all rows from the respective tables.

2.1 Table extreme_weather_events:

```
import pandas as pd
import sqlite3
db_path = '../data/database.sqlite'
conn = sqlite3.connect(db_path)
cursor = conn.cursor()
df_events = pd.read_sql_query("SELECT * FROM extreme_weather_events", conn)
df_events.head()
```



	Year BIGINT	Seq BIGINT	Disaster Group TEXT	Disaster Subgroup TEXT	Disaster Type TEXT	Disaster Subtype TEXT	Event Name TEXT	Country TEXT	Country Code TEXT	Location TEXT	Origin TEXT	Associated Dis TEXT	Appeal TEXT	Declaration TEXT	Aid Contribution FLOAT	Dis Mag Value FLOAT	Dis Mag Scale TEXT
1	1997	559	Natural	Meteorological	Storm	Tropical cycl…	Zita	Malaysia	MYS	Kedah, Perlis, Pen…	NULL	Flood	No	No	NULL	11250.0	Kph
2	1997	184	Natural	Meteorological	Storm	Tropical cycl…	Zita	Thailand	THA	Surat Thani Provin…	NULL	Flood	No	No	NULL	NULL	Kph
3	1997	186	Natural	Meteorological	Storm	Tropical cycl…	Zita	Viet Nam	VNM	NULL	NULL	NULL	No	No	NULL	NULL	Kph
4	1991	517	Natural	Meteorological	Storm	Tropical cycl…	Zeke (Etang)	Philippines	PHL	Catanduanes, Albay…	NULL	NULL	NULL	NULL	NULL	NULL	Kph
5	1992	470	Natural	Meteorological	Storm	Tropical cycl…	Zeke & Fred	Lao People's DR	LAO	Whole country	NULL	NULL	NULL	NULL	NULL	NULL	Kph
6	1991	713	Natural	Meteorological	Storm	Tropical cycl…	Zeke	Viet Nam	VNM	Quang Ninh Hai Pho…	NULL	NULL	NULL	NULL	NULL	NULL	Kph
7	1998	389	Natural	Meteorological	Storm	Tropical cycl…	Zeb (Tiang)	Philippines	PHL	Town Of Tabuk Leve…	NULL	NULL	No	No	NULL	280.0	Kph
8	1995	272	Natural	Meteorological	Storm	Tropical cycl…	Zack (Pepang)	Philippines	PHL	Leyte, Samar, Cent…	NULL	Flood	No	No	NULL	135.0	Kph
9	1995	274	Natural	Meteorological	Storm	Tropical cycl…	Zack	Viet Nam	VNM	Quang Ngai, Binh D…	NULL	Flood	No	No	838.0	NULL	Kph
10	1994	71	Natural	Meteorological	Storm	Tropical cycl…	Yunya (Norming)	Philippines	PHL	Pampanga, Bulacan,…	NULL	NULL	NULL	NULL	NULL	NULL	Kph
11	1970	80	Natural	Meteorological	Storm	Tropical cycl…	Yoling	Philippines	PHL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	Kph
12	1992	312	Technologic…	Technological	Transport accident	Air	Yakovlev	Viet Nam	VNM	Near Nha Trang	NULL	NULL	NULL	NULL	NULL	NULL	NULL
13	1998	461	Technologic…	Technological	Transport accident	Air	Yak-40	Lao People's DR	LAO	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
14	2017	162	Technologic…	Technological	Transport accident	Air	Y-8F-200	Myanmar	MMR	Mer d'Adaman, near…	NULL	NULL	NULL	NULL	NULL	NULL	NULL
15	2000	726	Technologic…	Technological	Transport accident	Air	Y-12 aircraft	Lao People's DR	LAO	Near Vientiane	NULL	NULL	No	No	NULL	NULL	NULL
16	2000	706	Natural	Meteorological	Storm	Tropical cycl…	Xangsane (Reming)	Philippines	PHL	Cavite district (R…	NULL	NULL	NULL	NULL	NULL	140.0	Kph
17	2006	517	Natural	Meteorological	Storm	Tropical cycl…	Xangsane (Milenyo)	Philippines	PHL	Cordillera Adminis…	NULL	Flood	NULL	Yes	NULL	160.0	Kph
18	2006	517	Natural	Meteorological	Storm	Tropical cycl…	Xangsane (Milenyo)	Viet Nam	VNM	Ha Tinh, Thua Thie…	NULL	NULL	NULL	NULL	NULL	148.0	Kph
19	1998	47	Technologic…	Technological	Transport accident	Air	XY-AES Fokker F-27	Myanmar	MMR	Between Thandwe an…	NULL	NULL	NULL	NULL	NULL	NULL	NULL
20	2000	582	Natural	Meteorological	Storm	Tropical cycl…	Wukong	Viet Nam	VNM	Thach Ha, Cam Xuye…	NULL	NULL	NULL	NULL	NULL	120.0	Kph
21	2007	457	Natural	Meteorological	Storm	Tropical cycl…	Wipha/Goring	Philippines	PHL	Negros Occidental …	NULL	Flood	NULL	NULL	NULL	NULL	Kph
22	1993	482	Natural	Meteorological	Storm	NULL	Winona	Thailand	THA	Udon Thani, Khon K…	NULL	NULL	NULL	NULL	NULL	NULL	Kph
23	1964	37	Natural	Meteorological	Storm	Tropical cycl…	Winnie (Dading)	Philippines	PHL	South, Central Luz…	NULL	NULL	NULL	NULL	NULL	NULL	Kph
24	1997	180	Natural	Meteorological	Storm	Tropical cycl…	Winnie	Philippines	PHL	All country	NULL	NULL	No	No	NULL	180.0	Kph
25	2004	609	Natural	Meteorological	Storm	Tropical cycl…	Winnie	Philippines	PHL	Real, Infanta, Gen…	NULL	Flood	Yes	NULL	17278.0	NULL	Kph
26	1996	227	Natural	Meteorological	Storm	Tropical cycl…	Willie	Viet Nam	VNM	Nghe An, Ha Tinh p…	NULL	NULL	No	No	NULL	101.0	Kph
27	1986	189	Natural	Meteorological	Storm	Tropical cycl…	Wayne (HP5)	Viet Nam	VNM	Ha Nam Ninh, Thai …	NULL	NULL	No	No	NULL	120.0	Kph
28	1986	184	Natural	Meteorological	Storm	Tropical cycl…	Wayne (Miding)	Philippines	PHL	Pangasinan, Bataan…	NULL	NULL	NULL	NULL	NULL	NULL	Kph
29	1984	186	Natural	Meteorological	Storm	Tropical cycl…	Warren (Reming)	Philippines	PHL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	Kph

Through data transformation, extreme_weather_events table comprises 2738 entries with total 26 columns. Several columns contain a high number of missing values at random (MAR), including Dis Mag Value, Origin, Declaration, Appeal, Local Time, Aid Contribution, each with more than 2,000 missing values (NULL value). Imputing values in such cases can lead to a significant amount of speculation and introduce bias or noise, diminishing the reliability of the analysis. For instance, Aid Contribution involves financial figures which are difficult to predict without concrete data. However, the table is included due to its continued relevance in addressing the analysis question.

[1] According to UN Women "Southeast Asia belongs to one of the most disaster-prone regions in the world". https://wrd.unwomen.org/explore/regions/southeast-asia-asean

2.2 Table socioeconomics:

SELECT * FROM socioeconomics Schema Query Editor Auto Reload Find Other Tools...

	Country TEXT	Country Code TEXT ↑	IncomeGroup TEXT	Year BIGINT	Life Expectancy World Bank FLOAT	Prevelance of Undernourishment FLOAT	CO2 FLOAT	Health Expenditure % FLOAT	Education Expenditure % FLOAT	Unemployment FLOAT	Sanitation FLOAT	Injuries FLOAT	Communicable FLOAT	NonCommunicable FLOAT
1	Indonesia	IDN	Lower middle income	2001	66.037	19.2	302060.0	2.11850…	2.46003007…	6.07999992370…	58.92317706	6996569.3	30881627.35	41520804.67
2	Indonesia	IDN	Lower middle income	2002	66.321	19.1	305640.01464…	1.98062…	2.64568996…	6.59999990463…	39.648177265	6843320.92	30124966.45	42387112.22
3	Indonesia	IDN	Lower middle income	2003	66.631	18.9	333890.01464…	2.24837…	3.21799993…	6.65999984741…	39.871798780…	6701750.23	29355201.16	43214107.72
4	Indonesia	IDN	Lower middle income	2004	66.969	19.1	341239.99023…	2.12002…	2.74847006…	7.30000019073…	40.10086501	15958395.23	28468267.38	44013193.46
5	Indonesia	IDN	Lower middle income	2005	67.334	19.2	342149.99389…	2.53189…	2.87281990…	7.94000005722…	40.335014905	6392259.59	27418345.71	44810321.32
6	Indonesia	IDN	Lower middle income	2006	67.717	19.0	364470.00122…	2.61946…	4.15745811…	7.55000019073…	40.651187945 0…	6618365.34	26541945.32	45590359.5
7	Indonesia	IDN	Lower middle income	2007	68.105	18.5	379959.99145…	2.82385…	3.04425001…	8.06000041961…	40.967715045 0…	6134806.18	25622182.43	46353201.2
8	Indonesia	IDN	Lower middle income	2008	68.485	17.4	376140.01464…	2.54842…	2.90190005…	7.21000003814…	41.284514220 0…	5937786.39	24674868.93	47228732.36
9	Indonesia	IDN	Lower middle income	2009	68.853	15.4	391079.98657…	2.62324…	3.52513003…	6.11000013351…	41.453857255	5922454.1	23753521.06	48051280.1
10	Indonesia	IDN	Lower middle income	2010	69.205	13.0	415519.98901…	2.79087…	2.81227993…	5.61000013351…	41.62426946	5728815.31	22901985.11	48708035.12

The table provides comprehensive indicators on life expectancy and socioeconomics, containing 152 entries across 14 columns, with no missing values due to the utilization of built-in python function *interpolate* and *fillna* for data imputation.

The dataset is organized in datasets listed in the World Bank Data Catalog, also provided under the Creative Commons Attribution 4.0 International License (CC BY-SA 4.0). When downloading the dataset, users are also agreeing to comply with the terms of CC BY-SA 40 license as detailed in section 2.4. Furthermore, it is necessary to provide attribution to The World Bank and its data providers in the format "The World Bank: Dataset name: Data source (if known)", including in any sublicenses granted to others.

2.3 Table disaster_risk:

SELECT * FROM disaster_risk Schema Query Editor Auto Reload Find Other Tools...

	Country TEXT	Country Code TEXT	Year BIGINT	World Risk Index FLOAT	Exposure FLOAT	Vulnerability FLOAT	Susceptibility FLOAT	Lack of Coping Capabilities FLOAT	Lack of Adaptive Capabilities FLOAT
1	Brunei Darussalam	BRN	2000	1.35	0.33	5.55	5.68	2.45	12.26
2	Cambodia	KHM	2000	10.63	2.49	45.34	25.0	65.02	57.33
3	Indonesia	IDN	2000	41.63	39.31	44.09	29.26	56.61	51.73
4	Lao People's Democr…	LAO	2000	3.56	0.38	33.4	39.91	14.49	64.42
5	Malaysia	MYS	2000	13.05	8.56	19.89	22.63	9.95	34.94
6	Myanmar	MMR	2000	36.94	23.58	57.87	55.54	57.85	60.32
7	Philippines	PHL	2000	44.97	40.62	49.79	38.95	56.21	56.37
8	Singapore	SGP	2000	0.6	0.15	2.39	2.61	0.71	7.34
9	Thailand	THA	2000	22.84	13.79	37.84	43.74	47.48	26.08
10	Timor-Leste	TLS	2000	7.89	2.8	22.22	57.39	3.2	59.75

The World Risk Index provides comprehensive indicators to evaluate disaster risk worldwide. For the specific research in this project, it is restricted to Southeast Asian countries. The output table disaster_risk contains 264 entries across 9 columns, featuring 6 standard indicators with no missing values thanks to data preprocessing.

2.4 Data license:

Three datasets are licensed under the Creative Commons Attribution 4.0 International License (CC BY-SA 4.0). It allows for the free use, sharing, adaptation of the licensed. Users can modify, remix, transform, and build upon the dataset for any purpose, even commercially, as long as they credit the original creator(s) and indicate if changes were made. To comply with the license, I would provide clear and proper attribution and, if applicable, include the dataset's title. For the sake of the creator's rights, it is also advisable to provide a clear link to the license itself, allowing others to understand their rights and obligations when using the dataset.

3 Analysis

3.1 What are the patterns in the occurrence of extreme weather events and the socioeconomic status of Southeast Asian countries?

The report utilizes python libraries *matplot.lib, seaborn*, and open-source data mining tool *RapidMiner* to visualize the patterns across three datasets. Figure 1 highlights that Indonesia and the Philippines experienced the highest overall number of disaster occurrences. Some countries seem particularly vulnerable to specific disaster types. For instance, Indonesia stands out with the highest number of total fatalities, primarily due to earthquakes and volcanic eruptions, reflecting its location along the Pacific Ring of Fire. Myanmar is the second highest in total fatalities, mainly caused by storms, followed by the Philippines in the third place. Floods are the most common disaster type across most countries, implying the region's vulnerability to heavy rainfall and potential infrastructure challenges in water management. This suggests the need for targeted disaster preparation and mitigation measures across these countries, which could save many lives.

The impact of extreme weather events varies significantly across countries. Singapore, Malaysia, Timor-Leste stand out as countries with consistently low numbers of fatalities from extreme weather events over the years. This could be attributed to several factors, including geographical location, effective disaster preparedness measures, and robust infrastructure.

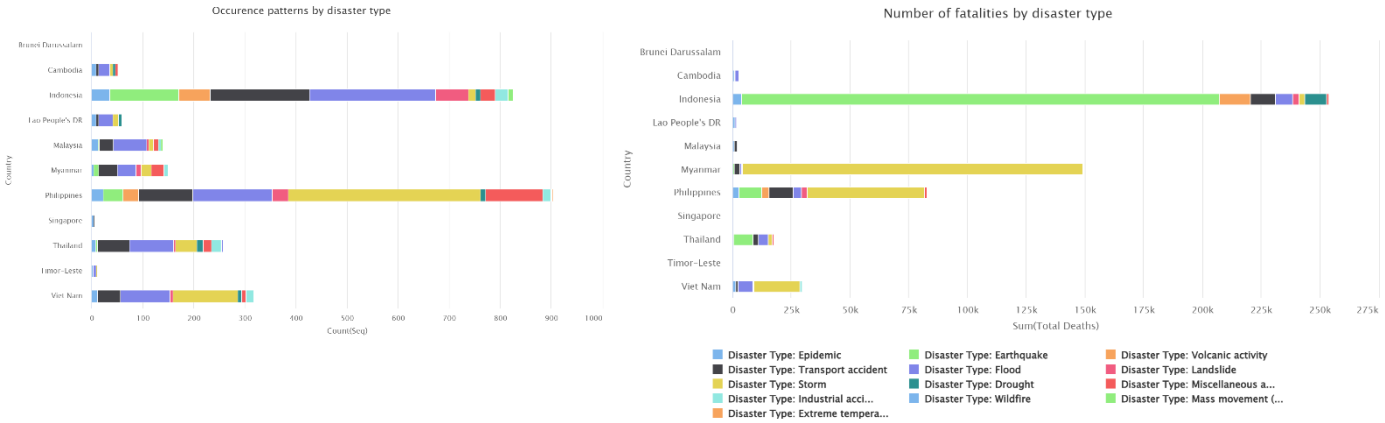


Figure 1. Disaster occurrences and fatalities by disaster type in Southeast Asian countries

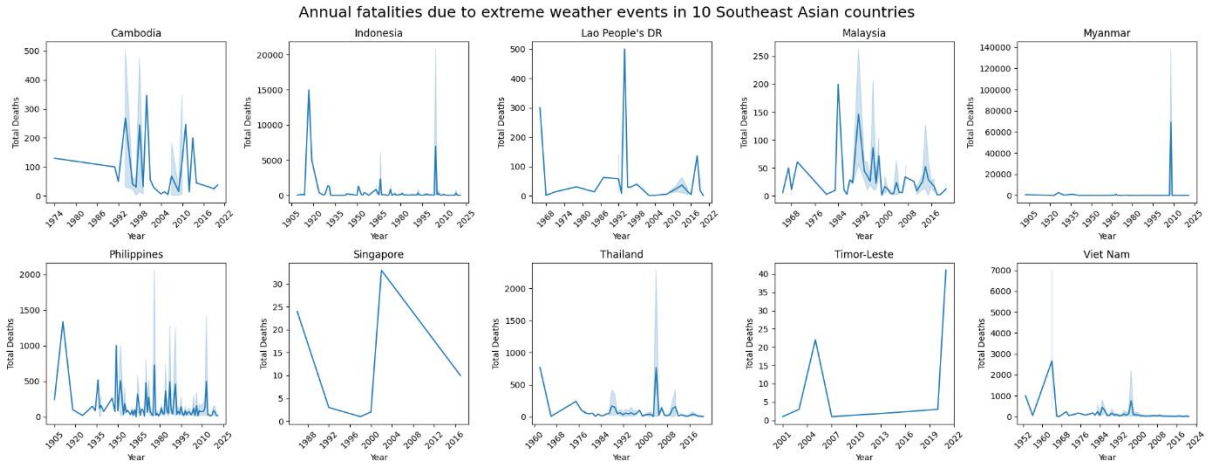


Figure 2. Annual fatalities across 10 Southeast Asian countries since 1902

Meanwhile, Myanmar witnessed devastating events, possibly due to a severe storm, resulting in a high death toll in 2010, as depicted in Figure 2. Several countries lack data for certain years or have significantly shorter data series compared to others. This could be variations in data collection methods. Brunei Darussalam was therefore excluded from this figure due to insufficient recorded data.

Regarding geographical factor, mainland Southeast Asian countries, including Cambodia, Lao People's DR, Myanmar, Thailand, Vietnam tend to experience fewer disasters compared to the island nations such as Indonesia and the Philippines, especially in terms of the diversity of disaster types. In countries with greater economic development, such as Singapore and Brunei, disasters are less likely to occur, possibly because their sophisticated infrastructure and more effective disaster preparedness strategy. Figure 3 of the World Risk Index (WRI) also provides additional analysis on this. While Indonesia and the Philippines display the highest overall WRI, reflecting its vulnerability to natural hazards, others like Singapore and Brunei with fewer disaster occurrences, maintain the two lowest WRI.

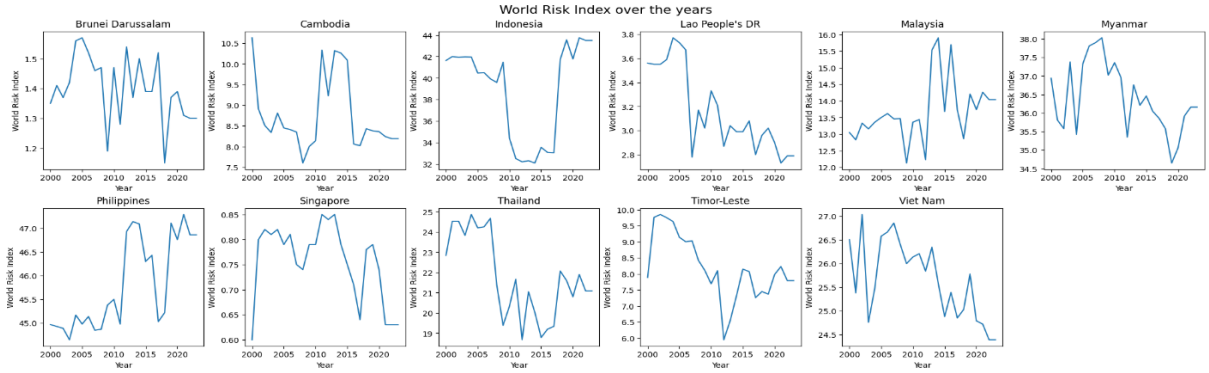


Figure 3. World Risk Index over the years in Southeast Asian countries

3.2 How do disaster risks impact the socioeconomic landscape of Southeast Asian countries?

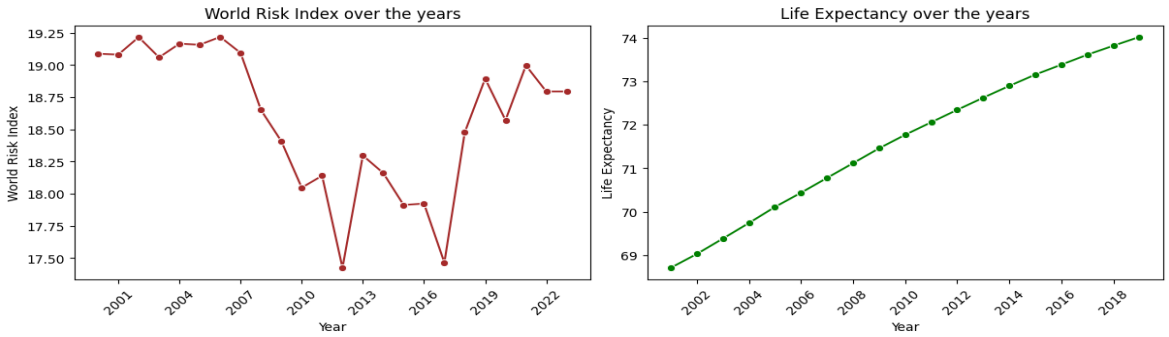


Figure 4. World Risk Index and Life Expectancy tendency over the years

The plots above do not reveal a direct correlation between the World Risk Index and life expectancy. Despite fluctuations in WRI, life expectancy continues to rise steadily over the years. However, as shown in Figure 5's correlation matrix, life expectancy exhibits a temperate negative correlation with the World Risk Index (-0.5) and its components: exposure (-0.3), vulnerability (-0.8), susceptibility (-0.7), lack of coping capabilities (-0.6), and lack of adaptive capacities (-0.7). This indicates that countries with higher disaster risk tend to have lower life expectancies. Countries with higher exposure, vulnerability and susceptibility to natural hazards, combined with limited coping mechanisms, are more likely to face a higher overall risk. In regions where the risk of exposure to hazards is higher, sanitation conditions are generally worse (negative correlation of -0.5). The strong positive correlations between Exposure and Injuries (0.7) and Communicable Diseases (0.8) indicate that a significant proportion of injuries can be directly attributed to hazard exposure. Disaster-prone

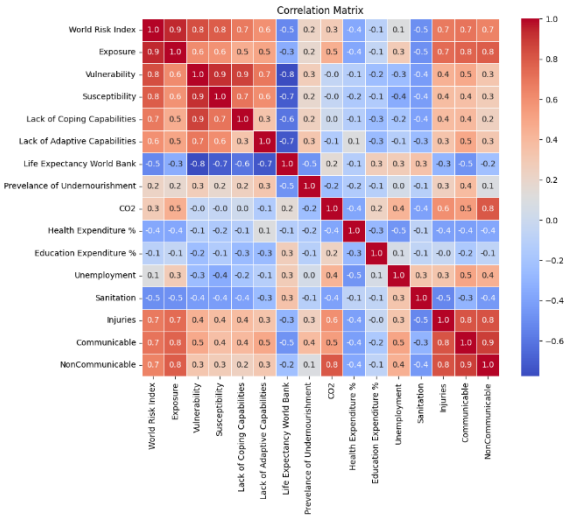


Figure 5. Correlation matrix of disaster risks and socioeconomic indicators

regions, such as Indonesia and the Philippines, which frequently experience earthquakes, floods, and storms, are particularly vulnerable to casualties and outbreaks of infectious diseases. Surprisingly, there is also a high positive correlation of 0.8 between Exposure and Non-communicable diseases (NCDs). Disasters do directly cause NCDs like heart disease and cancer, but the resulting stress, trauma, and healthcare disruptions can worsen these conditions and increase the risk of developing them. Overall, understanding these correlations helps policymakers and public health officials prioritize interventions for climate change and disaster exposure, leading to better preparedness and more resilient societies. However, keep in mind that correlation does not imply causation; further research is necessary to understand the underlying causal mechanisms revealed by the matrix.

4 Conclusion

To sum up, this data science project has uncovered several key takeaways, particularly the significant impact of climate change-induced extreme weather events on socioeconomics:

- **Unequal distribution in the region**: The impact of extreme weather events is not evenly distributed across the region. Countries like Indonesia and the Philippines experience significantly higher disaster occurrences compared to others like Singapore, Brunei or Laos. This highlights the need for tailored approaches to disaster risk reduction and policy management for these two vulnerable island countries. Government should incorporate disaster-resistant building codes and construction practices in all infrastructure projects.
- **Reduce exposure to hazards:** In developing Southeast Asian countries, where extreme weather events occur frequently, policymakers can implement strategic urban planning to minimize exposure to high-risk areas prone to floods, landslides, and other natural hazards. In Myanmar, the Philippines and Vietnam, a strong infrastructure as building storm-resistant housing, strengthening seawalls, improving drainage systems should be considered to reduce damage and casualties during severe storms.
- **Prioritizing health and resilience**: Correlation analysis reveals the potential for targeted policy interventions to minimize disaster risks and significantly enhance socioeconomic resilience. Prioritizing investments in health and education, while mitigating natural hazards, can create a positive feedback loop, reducing WRI, exposure, unemployment, injuries, and disease burden.

Limitation: The datasets cover limited timeframes and still contain many missing values. Some crucial aspects of disaster risks and socioeconomic aspects are therefore difficult to quantify. Further research on different policy interventions for disaster risks and their potential unintended consequences to the socioeconomic growth would be beneficial in the future.