**Abstract**

Floods are the most prevalent natural catastrophe in both developed and developing countries, accounting for around 40% of all natural disasters. Flooding has serious consequences for human health before, during, and after the flood. Southeast Asia is highly susceptible to frequent and severe environmental disasters. The recent floods in Southeast Asia have been caused by a variety of events, including typhoons, severe rainfall, and tropical storms. Four tropical cyclones, as well as strong and extended monsoon rains, caused extensive damage in Thailand, Cambodia, the Philippines, Vietnam, and Indonesia. On a massive scale, these tremendous monsoon rains, typhoons, and tropical storms are the outcome of climate change, a complex system marked by dynamic relationships between land, bodies of water, and inhabitants. Southeast Asia is vulnerable to the adverse effects of climate change due to its rapidly rising population, the majority of whom are impoverished, insufficient food security, and decreasing natural resources. Small–scale mitigation has been undertaken by both governmental and non-governmental entities across Southeast Asia. These initiatives, however, are frequently unsustainable due to a lack of community connection and engagement. As a result, it caused tremendous and widespread damage across Southeast Asia regions livelihood and economy. The aim of this analysis is to construct an Exploratory Data Analysis of floods from the year 2000-2022 that will calculate the total damages to each country recorded in the data sets. Furthermore, it will be ranked up from each country based on the numbers of floods recorded and it will show the difference between the most damaged country based on the total damage to the least damaged country. This information will show precise and intact statistics of data that can help and prevent the damage caused by flood to every country in Southeast Asia.

**Reliability of Organizations Where the Data Sets Originated**

The researchers used the Emergency Events Database also known as “EM-DAT” launched by the Centre for research on the epidemiology of disasters which is a research unit that collaborates with the World Health Organization(WHO) as a centre since 1980 with this facts the records or datasets derived from the EM-DAT is authenticated for use and research even by the World Health Organization there backing up its reliability and credibility.

Project Proponents Evaluation of the X variables inside the Data sets References

**Objectives**

1.From 2001 to 2020, determine the top 5 countries in Southeast Asia in terms of the most damaged country based on overall damage down to the least damaged country.

2. Obtain information with relation to the data about the countries of Southeast Asia, which will be ranked from worst to best in terms of overall property damage.

3. Determine which of the top five countries have the most damage in terms of total deaths between 2001 and 2020.

4. Determine which flood subtype is the primary cause of overall damage in the top five countries.

**Data Sets Normalization**

Data normalization substantially aids in the reorganization and use of data acquired from multiple sources. It also increases data simplicity for group members, allowing for a more efficient method to generate data visualizations to focus on the main sections of a data collection by effectively suspending the unnecessary entries from the presentation to avoid data anomalies. The data set that has to be normalized is the South East Asian Countries 2000-2022 data collection since it contains rows that will not be useful for this project such as the Dis No, Seq, Glide, Disaster Group, Disaster Subgroup, Disaster Subsubtype, Event Name, Region, Continent, Origin, Location, Associated Dis, Associated Dis2, OFDA Response, Appeal, Declaration, Aid Contribution, Dis Mag Value, Dis Mag Scale, Local Time, Latitude, Longitude, Local Time River Basin, Start Year, Start Month, Start Day, End Year, End Month, End Day, Reconstruction Costs ('000 US$), Reconstruction Costs, Adjusted ('000 US$), Insured Damages ('000 US$), Insured Damages, Adjusted ('000 US$), Adm Level, Admin1 Code, Admin2 Code and Geo Locations. These filter rows will be deleted.

The analysts used Microsoft Excel to normalize the data set. Simply select the filter option from the home tab, then the sort and filter options from the menu. The filters were then applied to the selected data range using an arrow in the higher rows that were chosen to be filtered.

**Expected Output**

This research aims to bring about an exploratory data analysis of floods from the year 2000 to 2022 from the countries within the South Eastern region of Asia that resulted in the highest damage to property as well as their respective local governments that experienced the least amount of damage to property. The worldwide flood datasets derived from the Centre for Research on the Epidemiology of Disasters' Emergency Events Database (EM-DAT) will be used to rank up the number of damages to property to each country from the highest to lowest and then determine what location in the South Eastern Asia had the best countermeasures for flood damage. This data will be utilized to be able to develop a flood countermeasure strategy that can potentially help the countries most affected by floods be more capable of saving as much as they can. By collaborating with other countries in developing new strategies, next time floods arrived not much would be carried away.

**Sustainable Development Goals (SDGs) of the Project**

The SDGs that this project aims to achieve are Sustainable Cities and Communities and Climate Action. These SDGs are further described as follows;

1. **Goal 11: Sustainable Cities and Communities** - This SDG aims to make cities and human settlements inclusive, safe, resilient and sustainable.
2. **Goal 13: Climate Action** - This goal strives to take urgent action to combat climate change and its impacts.

These SDGs are in line with the Proposal Topic: Data Based Outlook to Reduce Flood to Property Damage EDA: Flood Damage to Property Reduction as it will provide a mitigation and quick response plan based from multiple frameworks from different countries within the South Eastern region of Asia that are tried and tested. These frameworks will help the governments of the countries that are take more damage caused by flood by assimilating the said frameworks from other countries and applying it to their own allowing them to mitigate damage to property caused by floods.

**Evaluation of the x variables inside the Data sets**

* Year - shows the year where the flood happens.
* Disaster Type - shows what kind of disaster occurred.
* Disaster Subtype – shows what type of flood is if it is flash flood, riverine flood and coastal flood.
* Country – shows the different country in Southeast Asia.
* ISO – shows the code assign for each country
* Number of Injured – shows the total number of injured people during the flood
* Number of Affected - shows the number of affected people during the flood
* Number of Homeless - shows the total number of people that became homeless during the flood
* Total Affected - shows the total number of injured people during the flood
* Total Damages - shows the initial cost in dollars.
* Total Damages, adjusted – shows the final cost in dollars
* CPI (Consumer Price Index) - measures the overall change in consumer prices over time based on a representative basket of goods and services