# Chic Clique

## Report

**By**


| | | |
|---|---|---|
| **Miss Nattarat** | **Panjachaipornpol** | **6488047** |
| **Miss Kanita** | **Karunkittikun** | **6488049** |
| **Miss Saranya** | **Suttisan** | **6488094** |
| **Miss Prakaikaew** | **Ratchawongsa** | **6488194** |
| **Miss Jidapa** | **Punrod** | **6488197** |
| **Miss Ravikarn** | **Jarungjitvittawas** | **6488210** |


**A Report Submitted in Partial Fulfillment of**
**the Requirements for**

**ITCS495 Special Topics in Database and Intelligent Systems**


**Faculty of Information and Communication Technology**
**Mahidol University**

**2024**

# Table of Contents

# Business Domain

## Business Domain Explanation

The United States faced numerous natural disasters each year, including hurricanes, tornadoes, and wildfires, which affected and resulted in billions of dollars of damage. The Federal Emergency Management Agency (FEMA) is in charge of "helping people before, during, and after disasters" by coordinating disaster response and providing relief funds (see also wikipedia). We will update this dataset regularly.

The dataset on U.S. Natural Disaster Declarations belong to the business domain of disaster management and risk mitigation. This domain focuses on the analysis, preparation, response, and recovery from natural disasters such as hurricanes, floods, wildfires, and earthquakes. The goal of disaster management is to reduce the impact of disasters on human lives, infrastructure, and the economy, while also improving the resilience of communities and industries to future events.

## Important Factors for Choosing this Domain

1. **Frequency of natural disasters**: The increasing frequency and intensity of natural disasters due to climate change make disaster management an urgent field of study.
2. **Economic and social impact**: Natural disasters have significant impacts on businesses, governments, and individuals. The ability to predict or mitigate these events can save lives, reduce costs, and minimize disruptions to the economy.
3. **Geographical variability**: Different regions experience different types of disasters e.g., earthquakes on the West Coast, hurricanes in the Southeast. Analyzing location-specific data is crucial for tailored disaster response strategies.
4. **Government and business preparedness**: Governments and businesses alike are responsible for preparing for and responding to natural disasters. Studying patterns in past disaster declarations can help improve preparedness plans and reduce the recovery time after disasters.

## Benefits of Analyzing this Domain

1. **Predictive Analysis**: By analyzing patterns in the occurrence of natural disasters, we can predict future events with greater accuracy. This can help communities prepare more effectively, reducing the overall damage.
2. **Improving Risk Management**: Businesses can use disaster data to develop risk mitigation strategies, ensuring that supply chains and critical infrastructure are protected during natural disasters.
3. **Resource Allocation**: Governments and organizations can use disaster data to allocate resources more effectively. This includes directing funds for disaster relief, emergency services, and infrastructure development in high-risk areas.
4. **Insurance and Financial Planning**: The insurance industry relies heavily on natural disaster data for setting premiums and assessing risks. Businesses and individuals can also use this data for financial planning to cover potential disaster related losses.
5. **Policy and Strategy Development**: The insights gained from analyzing natural disaster data can be used to inform policies related to urban planning, construction codes, and climate change adaptation strategies. It can also guide government investment in disaster-resistant infrastructure.
6. **Community Awareness and Preparedness**: Analyzing this data can help raise awareness among communities about the risks they face and encourage them to take preventive measures, thus reducing the overall impact of future disasters.

# Target Users

The Federal Emergency Management Agency (FEMA) is the primary federal organization responsible for coordinating disaster response and recovery efforts across the United States. The provided dataset offers crucial insights to help strengthen FEMA's strategies for disaster preparedness and response.

## Characteristics

FEMA requires comprehensive data to track and manage disaster events efficiently. This dataset provides essential geographic and program-specific insights, which are key for prioritizing resource distribution and managing recovery programs. Additionally, information on various assistance programs allows FEMA to oversee and manage the support provided to individuals and communities affected by disasters, ensuring a more efficient response and recovery process.

## FEMA's Pain Points

Our team is focused on the significant pain points FEMA encounters in its disaster response and recovery efforts. These include prolonged delays in resource deployment, and logistical challenges that hinder timely assistance. By analyzing these critical issues, we aim to develop a comprehensive dataset that can provide valuable insights, enabling FEMA to improve its response speed and enhance overall operational efficiency in disaster management.

# Objectives of this work

FEMA struggles to anticipate future disasters and proactively prepare for necessary resources and funding [4].

Therefore, this project aims to develop predictive insights that help FEMA anticipate future disasters and proactively prepare by analyzing historical disaster trends, regional risk patterns, program demand, funding amount of each disaster, and obligation month for grants. The analysis aims to enhance FEMA's capacity to predict, mitigate, and assist people from the impact of future disasters through better preparedness, resources and funding allocation and grants, emergency management, and informed decision-making.

## Research Questions

1. What times of year are certain disasters most likely to occur?
2. Which states are most at risk for certain types of disasters?
3. Which assistance programs are most likely to be needed for future disasters?
4. What types of disaster incidents receive the most funding amount from FEMA?
5. What category of emergency management receives the most funding amount from FEMA?
6. Which states are the slowest to receive funding from FEMA?

# Data Description

## Title 1: US Natural Disaster Declarations (1953 - 2023)

**Description**: This dataset provides an overview of all reported disasters in the United States from 1953 to 2023. It was sourced from the Federal Emergency Management Agency (FEMA) website, and some basic data cleaning and formatting were applied. The dataset is detailed down to the county level, using FIPS codes to represent counties. It includes information on the type of disaster, timing, and specific FEMA aid programs.

**Link**: https://www.kaggle.com/datasets/headsortails/us-natural-disaster-declarations?select=us_disaster_declarations.csv

**Column Description:**

- **fema_declaration_string**:Unique identifier combining declaration type, disaster number, and state.
- **disaster_number**: Sequential number assigned to each disaster event.
- **State**: Identify US state, district, or territory.
- **declaration_type**:Type of declaration including major disaster(DR), emergency management(EM), or fire management(FM).
- **declaration_date**: Date when the disaster was reported.
- **fy_declared**: Fiscal year of the declaration.
- **incident_type**: Type of incident (e.g., Fire, Flood, Hurricane).
- **declaration_title**: Title for the disaster (e.g., "Hurricane Katrina", "Covid-19 Pandemic").
- **ih_program_declared:** Binary flag indicating whether the "Individuals and Households program" was declared for this disaster (IH provides financial and direct services to eligible households affected by a disaster[1]).
- **ia_program_declared:** Binary flag indicating whether the "Individual Assistance program" was declared for this disaster (IA provides financial and direct services to eligible individuals affected by a disaster[1]).
- **pa_program_declared:** Binary flag indicating whether the "Public Assistance program" was declared for this disaster (PA provides grants to State Agencies, Municipalities and Private Non Profit (PNPs) so that communities can quickly respond to and recover from major disasters[2]).

- **hm_program_declared**: Binary flag indicating whether the "Hazard Mitigation program" was declared for this disaster(HM provides funding for eligible mitigation measures that reduce disaster losses[3]).
- **incident_begin_date**: Beginning date of the incident.
- **incident_end_date:** Ending date of the incident.
- **disaster_closeout_date**: Date all financial transactions of all programs are completed.
- **fips**: 5-digit code for identifying counties.
- **place_code**: Unique FEMA location code.
- **designated_area**: Description of the affected area (e.g., Los Angeles, Statewide).
- **declaration_request_number**: Unique ID assigned to request for a disaster declaration.
- **hash**: MD5 hash generated from the fields and values within the record.
- **last_refresh**: The date when FEMA last updated the record.
- **id**: A unique identifier assigned to the record.
- **Last_ia_filing_date:** The deadline for individuals to apply for such assistance.

## Title 2: Public Assistance Funded Projects Details

**Description**: This dataset contains the financial obligations details of the grant of FEMA in Public Assistance(PA) projects for debris elimination, damaged public restoration, and hazard mitigation and protective measures for future events. Furthermore, it includes PA recipients' or applicants' data, funded lists, and individual project worksheets.

**Link**: https://www.fema.gov/openfema-data-page/public-assistance-funded-projects-details-v1

**Column Description:**

- **disasterNumber:** The assigned number to identify the declared incident disaster.
- **declarationDate:** The declared date of the disaster.
- **incidentType:** The type of incidents.
- **pwNumber:** The assigned number to determine the distinctive project.
- **applicationTitle:** The non-unique title of the application.
- **applicantId:** The unique identification number of the Public Assistance applicant.
- **damageCategoryCode:** The code to categorize the types of work granted by Public Assistance(PA).

- **projectSize:** The project size includes Large and Small, which depends on the permitted amount in the damage survey.
- **county:** The country, independent city, territory name of the US.
- **countyCode:** The unique number to identify the country in the US comes from five digits of the Federal Information Processing Standard (FIPS) Code and uses only the last three digits.
- **state:** The state or territory name of the US.
- **stateCode:** The unique two characters to identify the state or territory name of the US.
- **stateNumberCode:** The unique number to identify the country in the US which comes from five digits of the Federal Information Processing Standard (FIPS) Code and uses only the first two digits.
- **projectAmount:** The estimation of the total cost without costs of administration in dollars of the project granted from Public Assistance(PA).
- **federalShareObligated:** The available funding for State grantees in dollars granted by the Public Assistance(PA).
- **totalObligated:** The total federal share amount of the Public Assistance(PA) to grant the amount to the eligible project in dollars.
- **obligatedDate:** The obligation date of the grant.
- **dcc:** The code to identify the damage category.
- **damageCategory:** The code to categorize the damage location.
- **lastRefresh:** The last updated date of the record.
- **hash:** The MD5 hash fields and the record values.
- **id:** The unique ID of the record.

# Alteryx

## Data Preparation Process

- **US Natural Disaster Declarations Dataset**

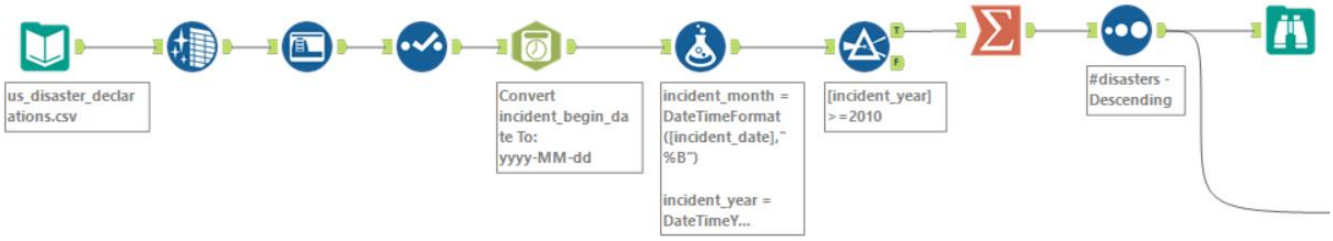


*Figure 1: Data Preparation Process on US Natural Disaster Declarations Dataset*

In the data preparation process for the US Natural Disaster Declarations dataset, we use the following tools:

- *Input Data* to import the 'us_disaster_declaration.csv' dataset.
- *Data Cleansing* to remove the null value and lead and trail whitespace.
- *Auto field* to automatically change the suitable field type, exclude 'ih_program_declared', 'ia_program_declared', 'pa_program_declared', and 'hm_program_declared' fields.
- *Select* to deselect unnecessary columns.
- *DateTime* and *Formula* on 'incident_begin_date' to convert the format of Date/Time to string into a new column named 'incident_date'.
- *Formula* on 'incident_date' to extract it into a month or 'incident_month' and year or 'incident_year'. Moreover, use it on 'ih_program_declared', 'ia_program_declared', 'pa_program_declared', and 'hm_program_declared' to replace the binary value with the understandable meaning, including '1' to 'Yes' and '0' to 'No'.
- *Filter* on 'incident_year' to show the data of incidents that happened from 2010 to 2023.
- *Summarize* to count the number of the incident types or 'incident_type' and change the field name as '#disaster' of the cleaned US Natural Disaster Declarations dataset, and then group by necessary fields, including 'state', 'incident_type', 'ih_program_declared', 'ia_program_declared','pa_program_declared','hm_program_declared','incident_month' , and 'incident_year'.
- *Sort* to order the data in descending based on the number of disasters or '#disaster'.

- **Public Assistance Funded Projects Details Dataset**

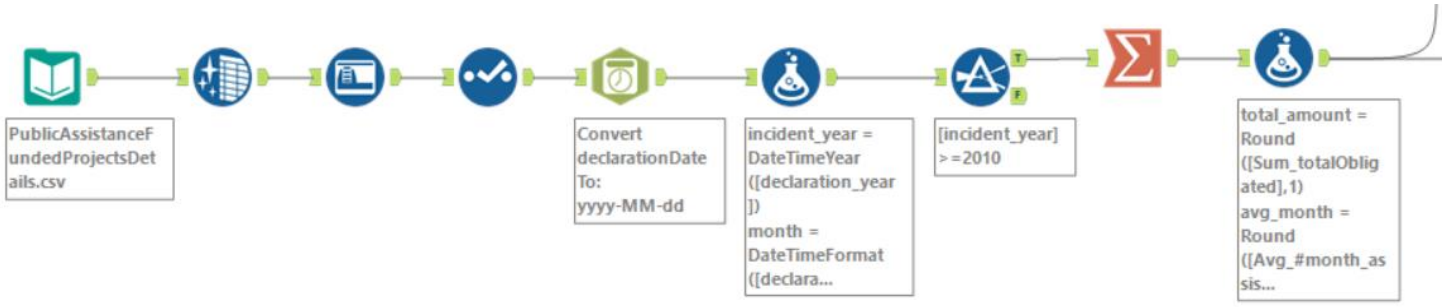


*Figure 2: Data Preparation Process on Public Assistance Funded Projects Details Dataset*

In the data preparation process for the Public Assistance Funded Projects Details dataset, we use the following tools:

- *Input Data* to import the 'PublicAssistanceFundedProjectsDetails.csv' dataset.
- *Data Cleansing* to remove the null value and lead and trail whitespace.
- *Auto field* to automatically change the suitable field type.
- *Select* to deselect unnecessary columns and rename the column of 'state' into 'state_name'.
- *DateTime* and *Formula* on 'declarationDate' to convert the format of Date/Time to string, and separate it into a year or 'declaration_year'.
- *Formula* on 'declaration_year' to extract it into a month or 'month' and a year or 'incident_year', Moreover, use it to calculate the period of the month or '#month_assistance' between the obligated grant date or 'obligatedDate' and the declared disaster date or 'declarationDate', and then remove the strings which are behind left parenthesis from incident type or 'incident_type'.
- *Filter* on 'incident_year' to show the data of incidents that happened from 2010 to 2023.
- *Summarize* to find the number of applicant projects by counting applicant ID or 'applicantID' and rename the field as '#projects', and calculate the average period of the month of obligation or '#month_assistance', and the sum of total obligated from Public Assistance(PA) or 'totalObligated', and change their field names as 'Avg_#month_assistance' and 'Sum_totalObligated'. Moreover, group by necessary fields, including 'state_name', 'stateCode', 'incidentType', 'incident_year', 'month', and 'damageCategory'.

- *Formula* to round the sum of total obligated from Public Assistance(PA) or ‘Sum_totalObligated’, and name the output column as ‘total_amount’, and the average period of the month of obligation or ‘Avg_#month_assistance’ and name the output as ‘avg_month’.

## The First Result Process


*Figure 3: The First Result Process*

After the preparation processes, we use the following tools to finalize the first result from the US Natural Disaster Declarations and the Public Assistance Funded Projects Details dataset before further utilization in creating a dashboard on Power BI:

- *Find Replace* to append the full name of the state or ‘state_name’ field from the Public Assistance Funded Projects Details dataset by finding values of ‘state’ from the US Natural Disaster Declarations dataset with ‘stateCode’ of another dataset.
- *Output Data* to save an output file as ‘us_disasters_overview.csv’.

## The Second Result Process

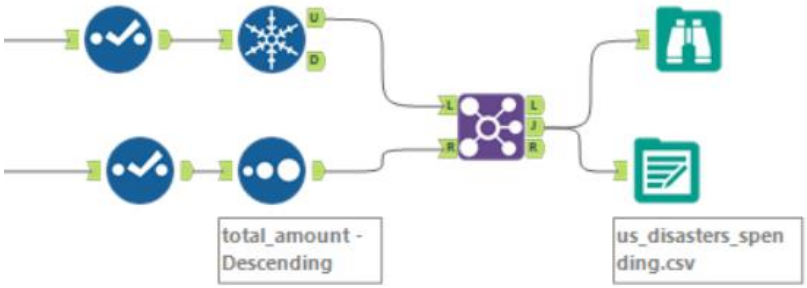


*Figure 4: The Second Result Process*

After the preparation processes, we use the following tools to finalize the second result from the US Natural Disaster Declarations and the Public Assistance Funded Projects Details dataset before further utilization in creating a dashboard on Power BI:

- *Select* to deselect unnecessary columns.
- *Unique* to remove the duplicated data in the incident type column or 'incident_type'.
- *Sort* to order the data in descending order based on the sum of total obligated from Public Assistance(PA) or 'total_amount'.
- *Join* to connect the fields of both datasets by their specific fields on 'incident_type' from the US Natural Disaster Declarations dataset with 'incidentType' from the Public Assistance Funded Projects Details dataset, and unselect the duplicated fields.
- *Output Data* to save an output file as 'us_disasters_spending.csv'.

# Power BI Custom Visual

## Overview

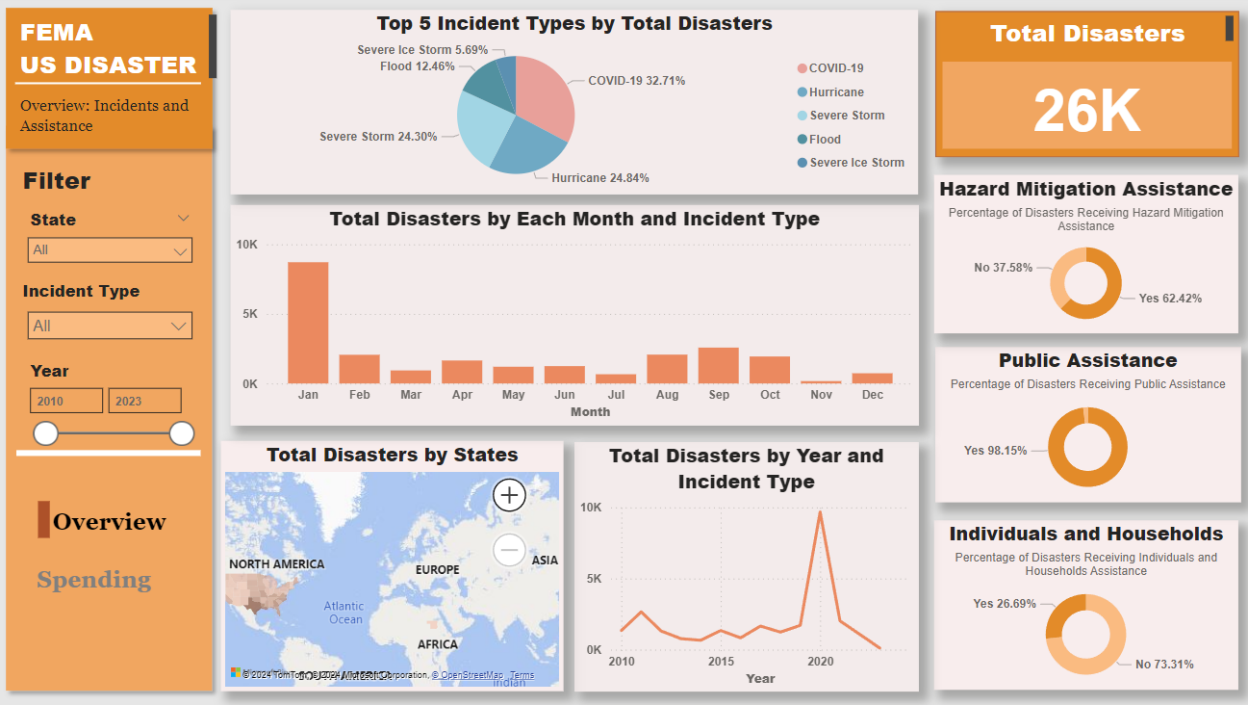


*Figure 5: The Overall of the Overview Page*

This dashboard is the “Overview” page. It shows the number of incidents in each state in the U.S. since 2010 categorized by several factors, including month and year. Furthermore, it illustrates the percentage of assistance programs which are Individuals and Households, Hazard Mitigation, and Public Assistance that have supported the state during disasters.

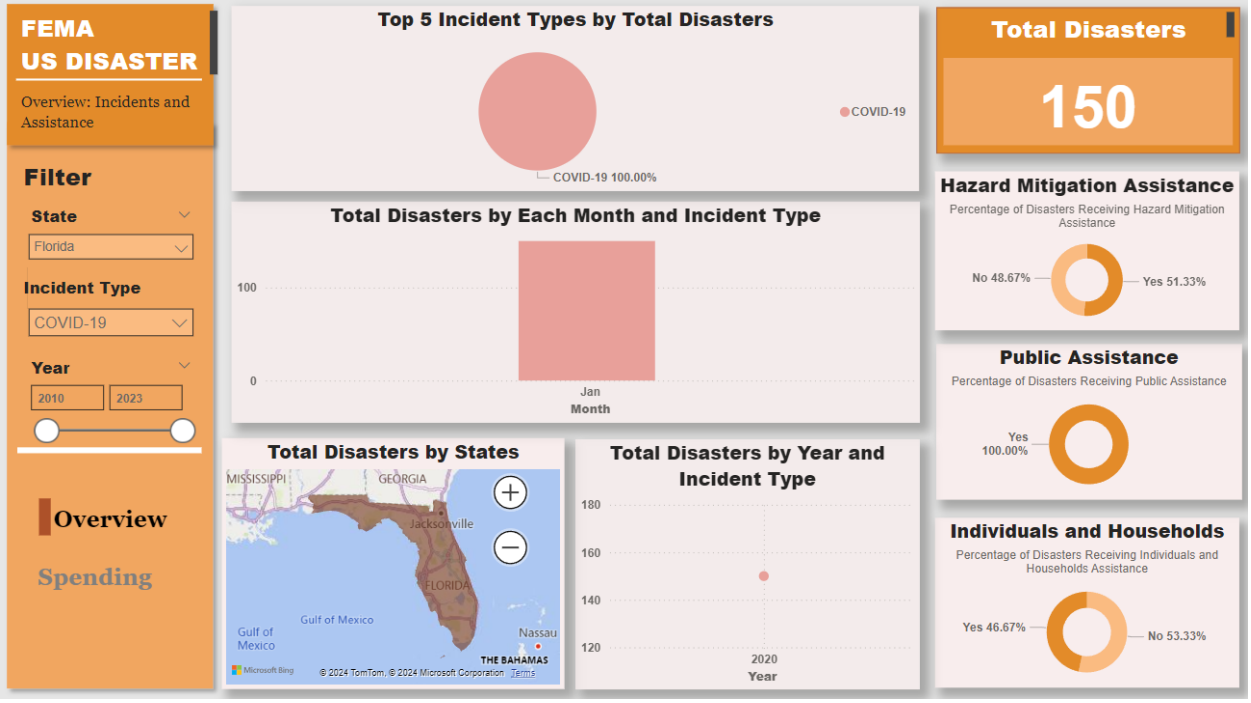


*Figure 6: The Example Usage of the Overview Page*

Consequently, this dashboard aims to predict the trends of each type of disaster and help FEMA better prepare resources and funds for assistance programs.

## Spending

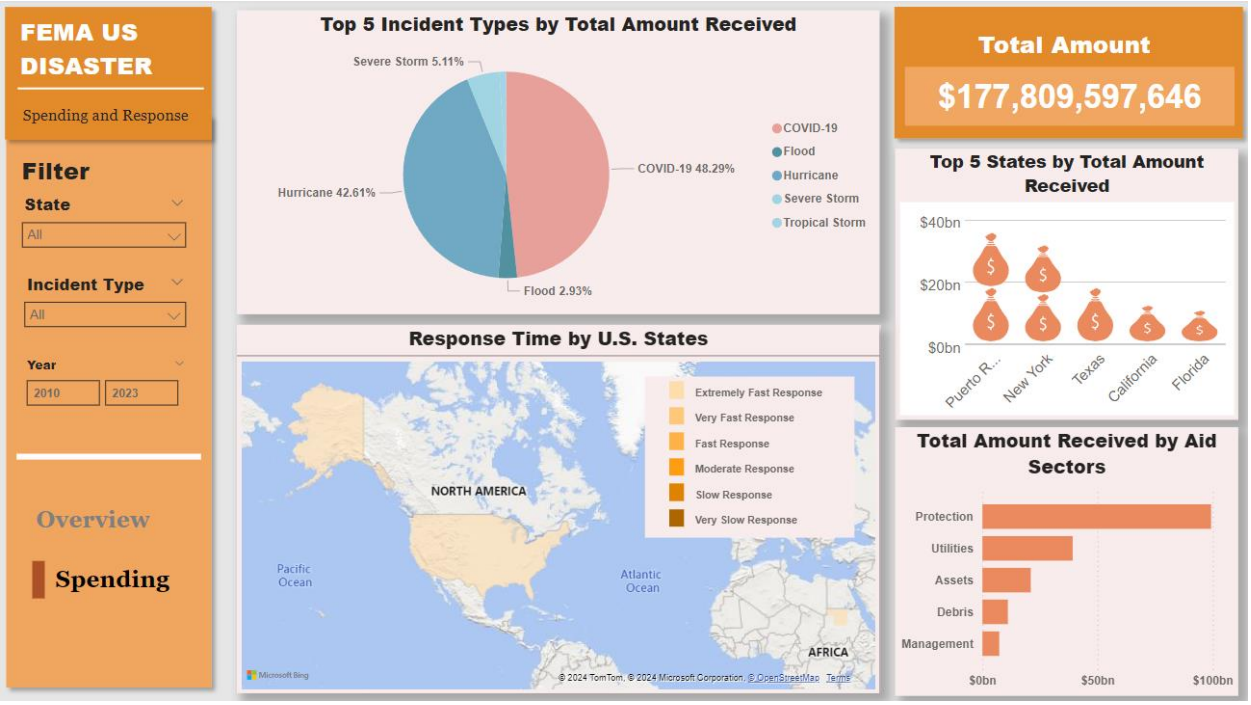


*Figure 7: The Overall of the Spending Page*

This dashboard is the "Spending" page. It presents the amount of FEMA funds allocated for assistance based on the types of incidents and emergency management categories. Additionally, it shows the response time of FEMA in providing assistance funding in each U.S. state.

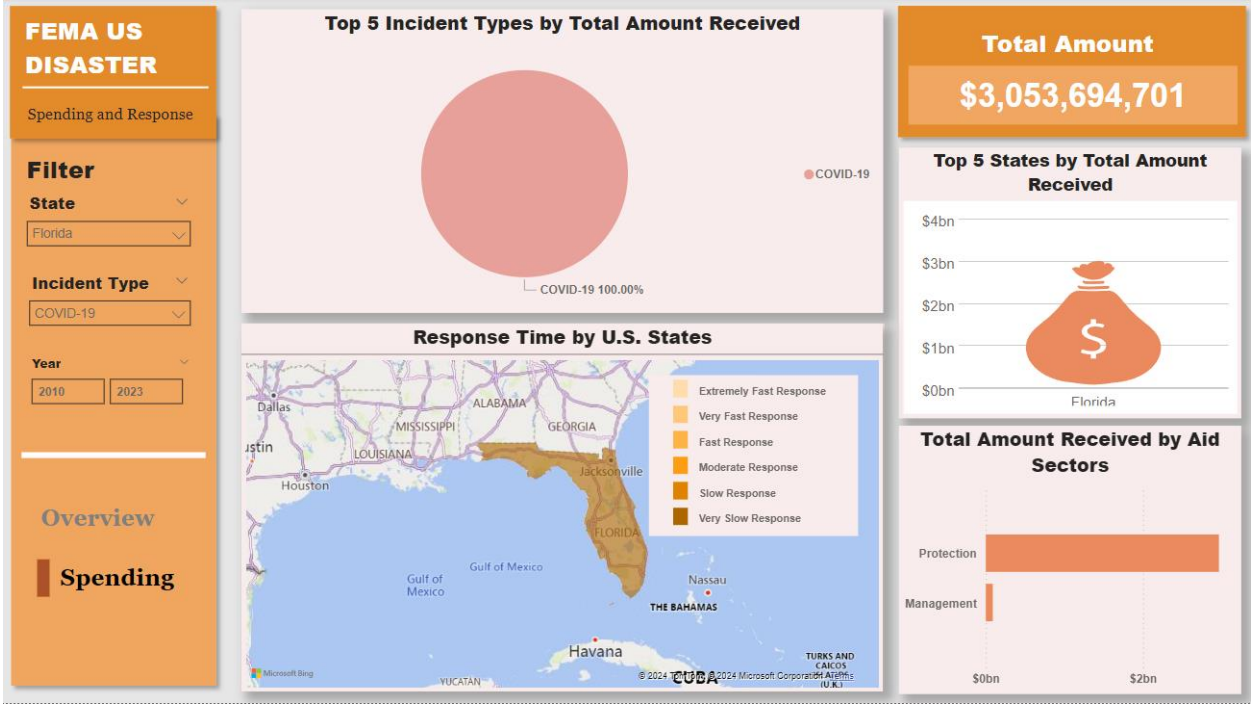


*Figure 8: The Example Usage of the Spending Page*

Hence, this dashboard aims to enhance the response time of FEMA for assistance funding and highlight the funding amount in different emergency management categories to help FEMA prepare and allocate appropriate funds for each type of emergency management.

# References

[1] *Individuals and households program* (no date) *FEMA.gov*. Available at: https://www.fema.gov/assistance/individual/program (Accessed: 04 September 2024).

[2] *Portal de transparencia de COR3* (no date) *Transparency Portal of COR3*. Available at: https://recovery.pr.gov/en/recovery-programs/public-assistance (Accessed: 04 September 2024).

[3] *Hazard mitigation assistance grants* (2020) *FEMA.gov*. Available at: https://www.fema.gov/grants/mitigation (Accessed: 04 September 2024).

[4] Freking, K. and Long, C. (2024) *Mayorkas warns FEMA doesn't have enough funding to last through hurricane season*, *AP News*. Available at: https://apnews.com/article/hurricane-helene-congress-fema-funding-5be4f18e00ce2b509d6830410cf2c1cb (Accessed: 03 October 2024).