**Module 7: Portfolio Project – Final Paper**

Drew A. Saunders

Colorado State University Global

MIS581: Capstone - Business Intelligence and Data Analytics

Dr. Osama Morad

November 24, 2024

Table of Contents

[Abstract 3](#_Toc183385453)

[Introduction 4](#_Toc183385454)

[Objectives 6](#_Toc183385455)

[Overview of Study 6](#_Toc183385456)

[Research Questions and Hypotheses 7](#_Toc183385457)

[Hypotheses 8](#_Toc183385458)

[Literature Review 9](#_Toc183385459)

[Research Design 14](#_Toc183385460)

[Findings 17](#_Toc183385461)

[Conclusion 22](#_Toc183385462)

[Recommendations 23](#_Toc183385463)

[Resources 24](#_Toc183385464)

# Abstract

Natural disasters can be severe, costly, and deadly. FEMA, the EPA, and other agencies work tirelessly to educate, prepare, provide aid, and lead recovery efforts for individuals that are likely to be impacted or have recently been impacted by disasters. The objective of this project is to advance the knowledge of declared disaster events in the US, specifically in their frequency and predictability. By finding significant patterns in the data, humanitarian organizations can improve upon previous procedures and infrastructure to better prepare themselves and others for future disasters. This research project reviews a FEMA dataset on declared disasters from 1953-2024 in the United States, focusing on the correlation and predictive modeling of the specific variables of incidentType (type of disaster), placeCode (zip code), and incidentBeginDate (date of disaster occurrence). Also, multiple studies around the topic of disasters were reviewed for research context, including studies on subtopics in general research project management, social media usage in disasters, and artificial intelligence and machine learning with disaster data. SAS Studio and Jupyter Notebook were used to view and analyze the data using correlation analysis and predictive regression modeling. The results of testing showed that Severe Storms show to be the most frequently occurring disaster type, and that the combination of disaster type, location, and date are not correlated at a statistically significant level of p<0.05. The implications of these findings are that additional factors must be considered to appropriately predict disasters at specific dates and locations. Humanitarian organizations must continue to learn and grow in this field to continually reduce the impacts of disasters in dollars and in lives.

*Keywords:* disaster, predictive modeling, FEMA, declared disasters

# Introduction

Prior to headlines being dominated by Election Day in the United States, a common topic in the news was of the devastation from disastrous weather events in the southern parts of North America. Reports of Hurricanes Helene and Milton, along with other disasters impacting communities in the southern United States rippled through the media, shaping the idea for a research project. This study is an attempt to learn more about weather disasters, reduce their ability to cause destruction, and limit negative effects on impacted areas and inhabitants.

Based on data from the National Oceanic and Atmospheric Administration’s (NOAA) National Centers for Environmental Information (NCEI), the year of 2023 was extremely costly and deadly compared to average years. In 2023, disasters in the United States totaled just under $93 billion, with 28 separate weather or climate events that each cost at least $1 billion (Smith, 2024). The disasters in the same year caused at least 492 direct or indirect fatalities, which is the 8th most since 1980 (Smith, 2024). While not all weather and climate events generate major costs or fatalities, serious events demand rapid responses and intense emergency management solutions to immediately contribute and jumpstart the recovery effort. Within hours, entire cities can be flooded, miles of property can be burned or destroyed, and huge chunks of the population can be displaced. Due to the speed and severity at which communities can be altered, being prepared prior to the disaster can be a key factor in providing necessary aid to devastated communities.

This project is limited to the United States and its territories, but even with this constraint, the large nation still has a variety of disaster types to manage on an annual basis. Hurricanes, tornadoes, earthquakes, and forest fires are several examples of disasters that affect different regions of the country. To effectively provide aid to in response to individuals and communities impacted by devastation, strategic emergency relief efforts must be organized and managed. Organizations such as FEMA, the EPA, and the Red Cross are examples of large agencies that specialize in these areas. These agencies have helpful websites with educational resources for developing a plan for emergencies, how to recover “safe drinking water”, to “avoid contact with flood water”, how to deal with mold, etc. (EPA, 2024). They also spend heavily in the risk mitigation area being discussed, with data showing over $20 billion spend on over 35,000 projects on over 150,000 properties going back to 1990 (FEMA, 2024). These resources emphasize the importance of preparation for emergencies prior to disaster striking. But what resources for after a disaster occurs? While these primary agencies quickly react to manage the situation, other organizations also seek to be involved in helping those individuals and communities being impacted. Walmart, for example, is a large corporate retailer that has taken it upon itself to use Walmart stores in effected areas as bases for the local communities to bathe, get essential supplies, receive food, and even do laundry (Walmart, 2024). While different agencies provide resources for these endeavors already, disasters still bring a heavy toll that may be able to be reduced through additional focus and data analysis. The more information that is gathered about disasters, the more prepared all organizations can be in organizing efforts to support effected communities.

The purpose of this project is to find meaningful and predictable patterns in the data to improve planning and preparation for potential and likely future disasters. Where this research project can improve the current state of events would be to show the correlation between time of year, physical location, and type of disaster event taking place to highlight areas that are likely to be impacted by certain disaster types during certain times of the year. If entities focused on humanitarian issues had better data for these predictions, they would be able to appropriately build additional and improved structures for safe zones or to provide educational or other resources to help prepare inhabitants for the upcoming disaster season. Destroyed assets can be reconstructed in similar fashions to their builds before the impact, or perhaps the reconstruction can also be done in ways that take recurring disasters into account to reduce future damages, future rebuilding costs, and future casualties. For improvements of that nature to be a common and recurring theme, the entities managing reconstruction of these areas must become aware of the specific patterns and predictability of natural or human-assisted disasters. Overall, these results could help to improve safety and reduce natural disaster casualties and damage costs.

# Objectives

This project is focused on finding patterns in the dataset of recorded large-scale disasters to help the populations of likely soon-to-be impacted areas increase awareness and preparedness for these critical events. The main objective is to find statistically significant results that will properly equip humanitarian organizations with knowledge to further improve their methods and tools for disaster preparation and relief. Finding correlations and better understanding disaster patterns in the combinations of time of year, physical location, and disaster type will help those groups and inhabitants of impact zones to better predict and prepare for disasters that are likely to materialize in the near future. Based on that information, time and money can be saved when initiating disaster relief efforts, which could limit the high costs and reduce casualties from weather and climate disasters.

# Overview of Study

This research paper first discusses the research questions and hypotheses that will send the research in the proper direction for finding results that will get closer to resolving the study’s problem and purpose statements. After discussing the hypotheses that will be tested, the paper will provide a literature review for the study’s topic area for five different peer-reviewed bodies of work. The initial work is a textbook for research project design and management. Three of the works are comprehensive studies in subsections of weather disasters including social media, artificial intelligence, and machine learning. The final work further delves into machine learning and AI while focusing on the theme of disaster prediction. After reviewing the relevant literature, the paper discusses the project’s research design, which focuses on the methodology, the methods used, the project limitations, and important ethical considerations. Upon reviewing the project’s research design, the paper transitions to its findings and concludes with its results and recommendations.

# Research Questions and Hypotheses

**Research Questions**

For this assignment, the below research questions were selected to guide the search for knowledge on disaster declarations while also providing an opportunity for testable hypotheses.

1. Do FEMA-declared disasters develop and strike in a geographically predictable pattern by disaster type?
2. Which disaster type is the most common?
   1. In the FEMA disaster declarations summaries dataset that has been identified for use in this Capstone project, the full list of disaster type options are: Biological, Chemical, Coastal Storm, Dam/Levee Break, Drought, Earthquake, Fire, Fishing Losses, Flood, Freezing, Human Cause, Hurricane, Mud/Landslide, Other, Severe Ice Storm, Severe Storm, Snow Storm, Straight-Line Winds, Terrorist, Tornado, Toxic Substances, Tropical Storm Tsunami, Typhoon, Volcanic Eruption, and Winter Storm (FEMA Dataset, 2024).

These research questions have been crafted in a manner that will cause the overall research to provide a direct yes/no answer to a question on predictability and a direct selection of most common disaster type. Both of those answers will help identify strengths or weaknesses in the dataset, which will progress the overall purpose of the study by better informing and preparing individuals and organizations to limit causalities and damages caused by disasters.

# **Hypotheses**

To properly answer these questions, they have been transformed into testable hypotheses. For each research question, there is a corresponding pair of alternate and null hypotheses listed below.

H1o – Fire is not the most common type of disaster (incidentType) in the United States.

H1a – Fire is the most common type of reported/declared disaster (incidentType) in the United States.

This group of hypotheses was selected to gain an understanding of what the overall most prevalent type of disaster is in the United States. Understanding the most common type of disaster will determine a good place to further delve into the data. If the alternate hypothesis is unable to be rejected, then further analysis can be done on fires to determine the locations and times of year where a fire disaster is most likely to occur. Fire was chosen as the hypothesized most common type here as it can occur in multiple regions during different times of year for different reasons. Most cities and towns in the United States also have a fire department, but not all cities and towns have a hurricane department, so it could be said that fire was seen to be chosen due to its general familiarity in American culture and its unfortunately opportunistic nature. While that selection and the description of the disaster type are subjective, it will be seen through analyzing the data if those opinions are found to be based around tested and measured facts.

H2o – There is no significant correlation between disaster types, geographic location, and time of year in the FEMA ‘DisasterDeclarationSummaries’ dataset.

H2a – FEMA-declared disasters are significantly correlated by disaster type (incidentType) and geographic area variables (placeCode) on certain dates (incidentBeginDate).

These hypotheses were selected to test some of the core variables in predicting disasters, which include disaster type, location, and time of year. Some specific examples of this hypothesis could be ‘Hurricanes are strongly correlated with disaster declarations in counties on the southern coast of Louisiana in October,’ or ‘Fires are strongly correlated in counties on the west coast and in northern California in June and July’. If these variables are significantly correlated, then the knowledge gained from testing this hypothesis can lead to increased preparation for US residents. This means it would directly achieve its goal of being “testable” in analysis and “focused” on the research problem and purpose statement (Polonsky, 2018).

# Literature Review

To bring contextually important information to the forefront of a study on predictions of data on FEMA declared disasters, the review of academic literature in the fields of disasters, data analysis, machine learning and artificial intelligence, and research project management is imperative. This review brings forth current findings in this specific academic landscape while highlighting holes in the current research in this area. The scope of the literature being reviewed is limited to published works in English with a focus on the prediction of and the recovery from disasters (with the exclusion of the course textbook on research project management). Using the thematic approach to reviewing literature in this area, this paper will first review a study focused on research methodology and research project management. Next, it reviews studies focused on using social media in disaster recovery, studies tracking disaster recovery using machine learning and artificial intelligence, and finally, studies focused on data analysis of disaster preparedness.

The availability of sources on the topic is healthy and growing in recent years, which is a great improvement compared to historical data. While weather prediction is a topic that has been around for hundreds of years, the availability of high-quantity data that enables expansive studies involving machine learning and artificial intelligence is newer and continuing to emerge as technological advancements in those areas continue to develop. From the available selection of studies, the consensus is that the accuracy in predicting specifics about storm and disaster developments in specific locations can continue to improve compared to the current state of the ability to predict these events. The way that humans leverage social media, artificial intelligence, and machine learning will be major opportunities for how we can improve predictions for these events and how we prepare for and respond to them.

A critical part of research in any is to follow outlined processes and follow critical steps that help to effectively design and manage a research project. By locating and following a “step-by-step guide” like Polonsky’s “Designing and Managing a Research Project,” it increases the likelihood the researcher can fix small mistakes upon reviews along the way in the research process instead of missing crucial steps that cause the entire project to crumble and start again from square one (Polonsky, 2018). Polonsky’s textbook helps to appropriately select “research topics”, “define problems” and “research objectives”, form “research questions and hypotheses”, gather and interpret data, and “present results” (Polonsky, 2018). The strength in this guide is that it is extremely helpful for individuals to form plans and become organized, and it offers examples to the reader better understand the text by seeing the suggestions in action. A weakness in this work could be that the examples are often small or vague, rather than a showcase of a full example of different areas to zoom out and show the big picture for how all the content aligns to form one cohesive document, study, or project.

Once a research project is designed, it is time to execute the research. After determining a topic, it is critical to establish a base of knowledge on the topic formed by finding, reading, and categorizing information gathered by others on this topic to understand what information is available and where holes in existing research may lie. A quick and effective way to do this is to find comprehensive research done on the selected topic, such as literature reviews or systematic reviews. Several examples of literature reviews and systematic reviews follow on the topics of social media use in disaster recovery, big data and machine learning in natural disaster application, and trustworthy artificial intelligence applications in natural disasters.

One of the largest movements of the 21st century has been the transition from print media and traditional news media into social media. Ogie et al. (2022) provides an overview of the studies in the social media space as it relates to natural disasters for the primary purpose of reviewing the current landscape of information in that area and identifying gaps. The data for the research analysis was gathered through the selection and collection of many studies located by using key words, meeting screening conditions, and removing duplicates. The study points that there has been an existing “research gap” in some areas such as the post-disaster phases like disaster recovery and hopes to help fill that gap with their research (Ogie, 2022). The comprehensive study also highlights a strong usage in social media usage around directly before a disaster strikes as information-seeking individuals retrieve information about their area near the time of impact. This study’s second objective of filling the post-disaster research gap shows promising ideas about how individuals are beginning to use social media for resources to get them started on their recovery journey for financial, emotional, and mental support. A large portion of the data in the previously conducted studies used data that was not validated, and while it may not necessarily be inaccurate, it is alarming to understand that the data both researchers and disaster-stricken internet users were interpreting were not proven to be accurate. A strength of the study is the comprehensive analysis of so many different resources, but a weakness of the study is seeing that so many of the results were based on data that could be invalid or received through unvalidated sites and sources like Twitter and Facebook. When the study was conducted, some of the current fact-checking tools on social media were not in place, making the nature of social media data being shared and viewed by internet users more likely to be invalid, due to the speed at which information and misinformation can travel before being corrected on the internet. While using social media and other internet resources, it is important for users to understand that some of the information they are acting on may not always be accurate or proven to be accurate.

Other systematic reviews were performed on natural disasters with a different subsection of analysis: artificial intelligence and machine learning. Albahri et al. review the trustworthiness of artificial intelligence being applied to natural disasters as the technology can be used to “analyze datasets”, “identify patterns”, and “forecast potential disasters” to “mitigate the impact of disasters and alleviate human suffering” (Albahri, 2024). A strength of this study is that it breaks down and defines disasters by type for individual and grouped analysis. This helps to get a better understanding of what factors may be influencing certain patterns and behaviors. A weakness of the study is that the data is often not collected or defined in a way that is best for analysis. Going forward, an improvement could be made in technology for recording, collecting, and storing disaster and weather data on a broader scale so that future analyses can have cleaner, more impactful data to study and analyze for more actionable results. Another strength of the study is that it has a glossary for words and acronyms it uses in the study and results that may not be common in everyday language. The additional context helps to properly prepare readers. Another potential weakness of the study is that it is so large that it can be hard to ascertain the main points without likely missing some important information.

A third systematic study that was used to prepare for this study topic is “Natural Disaster Application on Big Data and Machine Learning: A Review” by authors R. Arinta and A. Emanuel (Arinta, 2019). This study was a large collection of information condensed into a concise document on the categorization of current information on applying big data analytics and machine learning to natural disaster data. Similar to the other comprehensive studies, the data for the research analysis was gathered through selecting many studies by searching for key words, meeting screening conditions, and removing duplicates. A strength of this study is that it classifies and organizes the data from previous studies into simple groups for analysis. It uses 63 studies to highlight the various types of data sources, analysis models/techniques, and category of study to see the available information at researchers’ fingertips. A weakness is that unlike the Albahri et al. review, this study does not add to any existing research or fill in gaps with its own analysis, but only provides descriptive statistics on previous studies. This collection of studies will help future researchers that focus on natural disaster data to have success in their findings, rather than be an addition to current information in its own findings.

In a different assessment on leveraging machine learning to improve natural disaster prediction, Jain et al. looks to add research to the conversation instead of just organizing previous studies (Jain, 2023). Some of the main points of the study are that machine learning has seen success in this area in the past, there are great advantages like increased evacuation planning and response rate when using machine learning for predictions, and there are some challenges with machine learning in the area of natural disaster studies like small or incomplete data sources that do not allow for a full analysis with accurate findings. With incomplete data sources, models can be biased by what is not present. Like other studies, a call out was made to the need for increased use of advanced technology to help measure and analyze this type of information. This study and others mentioned that natural disaster recordings have more recently been increasingly active and cite climate change as a potential driver. As this type of variable has worldwide impacts, the need to continue to improve measurements and results in these areas remains strong. While these are not climate change studies, if that variable continues to produce increased counts of disasters at higher strengths of devastation, then the need to improve our understanding of specifics in prediction and forecasting will only continue to rise alongside it.

# Research Design

**Methodology**

While defining research methods, decisions must be made for how to design the research and engage in the study. The data collection step forces a determination for “qualitative” or “quantitative” data, meaning that the data can either be numeric or word-based (Scribbr Methodology, 2024). In this project, the data is quantitative and meant to be used in statistical analysis, as any measurable column that is not numeric is categorical and divides the items into groups. Another important decision for a research project is determining on how the data will be obtained, with options like “primary” or “secondary,” with primary referring to data that is collected by the researcher and secondary referring to data that is used by the researcher after it is obtained from a 3rd party collected that originally collected it (Scribbr Methodology, 2024). In this research project, the dataset was originally formed by FEMA, the United States Department of Homeland Security - Federal Emergency Management Agency, so its use is secondary. A third consideration for the data is whether the information will be “descriptive” or “experimental,” and in this project the data will be descriptive, as it will be measured “as is” instead of being used in an experiment (Scribbr Methodology, 2024).

**Methods**

Statistical analysis will be used to test this dataset. Jupyter Notebook and SAS Studio will be used to test for correlations between variables, run descriptive statistics for counts of different segments of the disaster data, and seek to prepare a predictive model based on correlated variables to help predict disaster patterns. The predictive model will test the initial hypothesis for predictable pattern recognition, and the descriptive statistics will address the second hypothesis for the most common type of disaster.

**Limitations**

The information in the dataset spans from 1953 to the present, as it is continually updated by FEMA as disasters are declared. Part of the limitation of the study deal with the nature of the timeframe the data was collected. In the dataset, there are obvious differences in quantities between observations recorded in 1953 and observations recorded in 2023. This could be due to multiple reasons. One could be that the processes or technological infrastructure for recording data that are in place today were not implemented in at as effective of a level in 1953. Another reason could be that changes in incentive for governing bodies to declare a disaster have changed over time but have not been well documented. These possibilities could point to why some areas or types of disasters do not repeat as often in the dataset. For example, if a weather event could have the characteristics of a ‘Severe Storm’ and hit the same location during the same time of year on an annual basis, that measurement would not be included in this dataset unless a governing body decided to declare the weather event as a disaster. These examples or others could explain why there are stark differences in the lower quantities of old records and the higher quantity of recent records. Regardless, a limitation to the study is that these differences are present and visible in the dataset, which cannot be ignored.

Another limitation is that the dataset is lacking additional variables which could be helpful in reaching the purpose and objective of the project. These missing variables include counts for casualties or dollars of damages recorded by declared disaster. These data points could help move the needle for conversations on where government agencies or other organizations could spend money to better prepare the population with knowledge and infrastructure.

**Ethical Considerations**

This dataset was obtained from the FEMA website and was not originally recorded during this study. However, as the data was previously recorded by another party, the ethics of the original collection and publication will be dissected. The dataset is anonymous and simple. It does not include casualty or death figures, nor does it include monetary damage totals or other ways of measuring the severity of devastation or impact per disaster. It strictly measures whether the disaster was declared, by which governing body/group it was declared, and the location information for occurrences. However, it could be assumed that this type of information could have been learned, at least in partiality, over the course of the data collection. As a commitment to ethics, the information excluded from any published version of the data. If that information was available, the sensitive information involving names, addresses, casualty details, or other information would have needed to be masked or altogether abandoned if it were to be used.

As the data type, statistical methods for analysis, limitations of the dataset and the study, and the considerations for ethics and technology have been described, the next step is to review the results of the study.

# Findings

**Hypothesis 1 – Incident Type Frequency**

To test the first hypothesis on disaster occurrence frequency, the dataset was analyzed to find the count of each incident type. The proposed alternate hypothesis stated that the most common type of disaster would be fire, with the null hypothesis stating that fire would not be the most common incident type. Based on the descriptive and summary statistics from this dataset, the most frequently occurring incident type is the Severe Storm event, with 18,371 of the declared disasters of the total 67,111 records at the time of the data pull. The top four incident types make up over 75% of the declared disasters recorded in the US since 1953.

**Incident Type Frequency Results Table**

A table with text on it

Description automatically generated

**Pie Chart for Count of Incident Type**



Based on the analysis of the data, Fire is not the most common incidentType. Severe Storm is the most common incidentType, at a frequency in the dataset of being 27% of all declared disasters. Thus, the null hypothesis is accepted for the first hypothesis test on disaster type frequency.

**Hypothesis 2 – Disaster Prediction**

Next is to test and review the second hypothesis, which discusses the significant correlation of the dataset variables involving disasters by location data, time of year, and disaster type. These variables translate to placeCode, incidentBeginDate, and incidentType. The hypothesis suggests the necessary testing of all three concepts (type of event, location, and date/time) as being correlated together. Therefore, the variables were used in a correlation analysis and a predictive regression model.

**Correlation Analysis Results**

***Groupings Captures***

A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated A screenshot of a computer screen

Description automatically generated

***Analysis of Correlation Analysis Test***

Of the 25 types of disasters, only 3 show a mildly strong correlation, above 0.70, which includes Dam/Levee Break, Other, and Toxic Substances. However, Biological, Chemical, Straight-Line Winds, and Tsunami all resulted as blank for various reasons. The Biological event records in the dataset all came from the same day, 1/20/2020, referring to the beginning of the COVID-19 pandemic and had occurrences across many states in the US. The Chemical records are all from one disaster on 1/09/14 in WV across multiple counties. The Straight-Line Winds event only has two records in the dataset, both occurring in the same event on the same day in two different counties in MT. The Tsunami event records all arise from one disaster on 3/11/2011 across locations in CA, HI, and OR. These show that there is not enough data from different events to show a true correlation. Based on these results, it cannot be shown that the combination of placeCode, incidentBeginDate, and incidentType are strongly correlated. To further test this hypothesis, an additional test was done for predictive regression modeling.

**Predictive Regression Model**

***Results Capture***

A screenshot of a computer

Description automatically generated

***Analysis of Predictive Regression Model Test***

The predictive regression model was set up by using incidentBeginDate as the dependent variable, incidentType and placeCode as the classification variables, and the data being partitioned so 80% is training data and 20% is testing data. Based on the results of the model, it was confirmed that the variables are not correlated at a statistically significant level. The best model using these variables only had an Adjusted R-Square of 0.3913. These results also show that the combination of placeCode, incidentBeginDate, and incidentType are not strongly correlated at a statistically significant level of p<0.05.

Thus, the null hypothesis is accepted for the second hypothesis test on correlation between variables and predicting disasters solely based on event type, location, and date.

# Conclusion

In conclusion, the null hypotheses were accepted in both cases of testing. The most common type of disaster that has been declared in the US since 1953 is the Severe Storm, at 27% of all disasters, not Fire, at 6% of all disasters. The variables representing disaster type, geographic location, and date were not found to be correlated at a statistically significant level. For the literature on the topic, the available resources provided a well-rounded view of disaster recovery and preparation, predictive data analysis, artificial intelligence and machine learning, and research project management. A collection of resources involving both comprehensive studies and individual studies provide an array of specific examples and broad findings that help to illustrate a more complete view of the big picture for current events in predicting disasters and managing relief efforts.

Overall, this research did not unlock a new knowledge for individuals or organizations to advance in understanding on this topic. Research must continue in this area for improvement.

# Recommendations

For future studies on this topic, a suggestion would be to find a dataset that includes all weather events, not just those that are declared ‘disasters’ by FEMA or other groups. This will not only improve the quantity of the dataset, but also provide a dataset that does not include government or human bias on how events are declared to be ‘disasters’.

Another recommendation for future versions of this study would be to isolate weather events or disasters by type, rather than looking at the entire dataset at once. Some disaster types may be more likely to repeat themselves or be more predictable based on date and location than others. By focusing on which disaster types are the most predictable, that would provide different organizations with the best possible starting point for implementing new procedures to help limit future disaster damage and devastation.

A final recommendation is based on the literature review. The reviewed studies suggest there is an opportunity for improvement in the areas of machine learning, artificial intelligence, and the use of social media. Those areas could be key in providing critical knowledge to individuals who are likely to be impacted by disasters in the future. If individuals focused in areas of software development or data science, rather than more strictly focusing on data analytics, there could be new discoveries in machine learning, artificial intelligence, and social media usage for interpreting, predicting, and communicating information prior to or during a critical weather event.

# Resources

Albahri, A. S., Khaleel, Y. L., Habeeb, M. A., Ismael, R. D., Hameed, Q. A., Deveci, M.,

Homod, R. Z., Albahri, O. S., Alamoodi, A. H., & Alzubaidi, L. (2024). A systematic

review of trustworthy artificial intelligence applications in natural disasters. *Computers*

*and Electrical Engineering*, 118, 109409.

**https://doi.org/10.1016/j.compeleceng.2024.109409**

Arinta, R. R. & Emanuel, A. W. R. (2019, November 22). *Natural Disaster Application on Big*

*Data and Machine Learning: A Review*. The 4th International Conference on Information

Technology, Information Systems and Electrical Engineering (ICITISEE 2019).

**https://doi.org/10.1109/ICITISEE48480.2019.9003984**

EPA. (2024). Preparing for and Responding to Natural Disasters. *United States Environmental*

*Protection Agency*. Retrieved October 13, 2024, from **https://www.epa.gov/natural-**

**disasters/preparing-and-responding-natural-disasters**

FEMA. (2024). Hazard Mitigation Assistance. *United States Department of Homeland Security –*

*Federal Emergency Management Agency*. Retrieved October 13, 2024, from

**https://www.fema.gov/data-visualizations/hazard-mitigation-assistance-obligations**

FEMA Dataset. (2024). OpenFEMA Dataset: Disaster Declarations Summaries - v2. FEMA

Reports & Data. Retrieved October 27, 2024, from https://www.fema.gov/openfema-

data-page/disaster-declarations-summaries-v2

Jain, H., Dhupper, R., Shrivastava, A., Kumar, D., & Kumari, M. (2023). Leveraging machine

learning algorithms for improved disaster preparedness and response through accurate

weather pattern and natural disaster prediction. *Frontiers in Environmental Science*, 11,

1194918. **https://doi.org/10.3389/fenvs.2023.1194918**

Ogie, R. I., James, S., Moore, A., Dilworth, T., Amirghasemi, M., & Whittaker, J. (2022). Social

media use in disaster recovery: A systematic literature review. *International Journal of*

*Disaster Risk Reduction*, 70, 102783. **https://doi.org/10.1016/j.ijdrr.2022.102783**

Polonsky, M. J. & Waller, D. S. (2018) *Designing and Managing a Research Project: A*

*Business Student’s Guide* (4th ed.). SAGE Publications, Inc.

Scibbir Methodology. (2024). Research Methods: Definitions, Types, Examples. *Scribbr*

*Methodology*. Retrieved October 20, 2024, from

**https://www.scribbr.com/category/methodology/**

Smith, A. B. (2023, January 10). 2023: A Historic Year of U.S. Billion-dollar Weather and

Climate disasters. *Climate.gov.* Retrieved October 13, 2024, from

**https://www.climate.gov/news-features/blogs/beyond-data/2023-historic-year-us-**

**billion-dollar-weather-and-climate-disasters**

Walmart, Inc. (2024). Hurricane Helene. *Walmart Disaster Response*. Retrieved October 13,

2024, from **https://corporate.walmart.com/purpose/community/disaster-**

**response/hurricane-helene?cid=linkedin:social:Hurricane-Helene-Support:p1583:-**

**:image:wmt-wmt:d5467**