

Article

Not peer-reviewed version

Predicting Household Disaster Preparedness: An Integrated Machine Learning and Simulation Framework

[Zhenlong Jiang](#) , Yudi Chen , Ting-Yeh Yang , Wenying Ji , [Sasha \(Zhijie\) Dong](#) , [Ran Ji](#) *

Posted Date: 3 March 2023

doi: [10.20944/preprints202303.0063.v1](https://doi.org/10.20944/preprints202303.0063.v1)

Keywords: disaster management; household preparedness prediction; machine learning; Monte Carlo Simulation



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Predicting Household Disaster Preparedness: An Integrated Machine Learning and Simulation Framework

Zhenlong Jiang ¹, Yudi Chen ², Ting-Yeh Yang ³, Wenyi Ji ⁴, Zhijie Sasha Dong ⁵ and Ran Ji ^{1,*}

¹ Department of Systems Engineering and Operations Research, George Mason University

² Department of Management Science and Engineering, College of Economics and Management, Nanjing University of Aeronautics and Astronautics

³ College of Engineering and Computing, George Mason University

⁴ Department of Civil, Environmental, and Infrastructure Engineering, George Mason University

⁵ Department of Construction Management, University of Houston

* Correspondence: rji2@gmu.edu; Tel.: +1-703-993-5993

Abstract: Effective household and individual disaster preparedness can minimize physical harm and property damage during catastrophic events. To assess the risk and vulnerability of affected areas, it is crucial for relief agencies to understand the level of public preparedness. Traditionally, government agencies have employed nationwide random telephone surveys to gauge the public's attitudes and actions towards disaster preparedness. However, these surveys may lack generalizability in certain affected locations due to low response rates or areas not covered by the survey. To address this issue and enhance the comprehensiveness of disaster preparedness assessments, we develop a framework that seamlessly integrates machine learning and simulation. Our approach leverages machine learning algorithms to establish relationships between public attitudes towards disaster preparedness and demographic characteristics. Using Monte Carlo simulation, we generate datasets that incorporate demographic information of the affected location based on government-provided demographic distribution data. The generated dataset is then input into the machine learning model to predict the disaster preparedness attitudes of the affected population. We demonstrate the effectiveness of our framework by applying it to Miami-Dade County, where it accurately predicts the level of disaster preparedness. With this innovative approach, relief agencies can have a clearer and more comprehensive understanding of public disaster preparedness.

Keywords: Disaster management; Household preparedness Prediction; Machine learning; Monte Carlo Simulation

1. Introduction

Natural disasters pose a significant threat to the lives and property of affected people, often resulting in shortages of essential resources (Rawls and Turnquist 2010, Elçi and Noyan 2018, Jiang *et al.* 2021). However, it is important to note that most fatalities and damages caused by disasters are preventable (Falkiner 2003). Adequate household disaster preparedness can significantly reduce the negative consequences of disasters and ensure that individuals can care for themselves and their families during the first 72 hours following a disaster (Diekman *et al.* 2007).

The US government estimates that investing one dollar in pre-disaster mitigation and preparedness efforts can prevent up to six dollars in losses from potential disasters (Multihazard Mitigation Council 2018). This is why most governmental relief agencies encourage the public to conduct disaster preparedness to hedge against the risk of natural disasters. For instance, the Federal Emergency Management Agency (FEMA) launched a campaign called *Get Ready* in 2003 to encourage people to prepare for emergencies and disasters. The campaign promotes the message to *get a kit, make a plan, and be informed* about different types of emergencies and their appropriate responses.

For disaster relief agencies, gaining knowledge about public disaster preparedness is essential for disaster risk and vulnerability assessments (Howe 2018). The survey is the most effective method to discover public attitudes and actions about disaster preparedness. Many relief agencies conduct surveys to learn public attitudes and actions about disaster preparedness. For example, FEMA conducts National Household Survey (NHS) annually to track the progress of household disaster preparedness via phone interviews since 2013 (FEMA 2021). Moreover, in academia, many researchers conduct surveys to analyze public attitudes about disaster preparedness in many countries and cities such as Japan (Onuma *et al.* 2017), Serbia (Cvetković *et al.* 2018), Istanbul (Tekeli-Yeşil *et al.* 2010), and Tehran (Najafi *et al.* 2015).

Conventional methods of surveying public disaster preparedness rely on either telephone surveys or convenience samples, both of which suffer from limitations. Telephone surveys are notorious for low response rates, while convenience samples have restricted generalizability. A case in point is the 2019 National Household Survey (NHS) conducted by FEMA (2021) in the United States, which sampled 5,025 households, including an oversampling of 510 interviewees from high-risk hurricane areas. Despite the focus on high-risk areas by the Federal Emergency Management Agency (FEMA), the 2019 NHS data show only 78 interviewees residing in Georgia, a state particularly vulnerable to hurricane damage. The limited sample size for Georgia renders the survey results unrepresentative of the actual state of household disaster preparedness, potentially leading decision-makers to miscalculate the population's vulnerability and overlook the needs of those most at risk. Therefore, more comprehensive and reliable methods of gauging public disaster preparedness are needed to enable better-informed relief efforts.

In academia, researchers often utilize statistical analysis techniques such as regression and hypothesis testing on historical survey data to address the generalizability issue of traditional surveys. Identifying the features that impact public attitudes towards disaster preparedness can guide disaster relief agencies in quickly identifying vulnerable groups and implementing targeted relief efforts when disasters occur (Baker 2011). For example, a logistic regression model was developed by Basolo *et al.* (2009) in 2009 based on surveys conducted in the New Orleans metro area and Los Angeles county. This model showed that a high level of confidence in government disaster management abilities and greater access to disaster preparedness information could improve residents' level of disaster preparedness.

Several studies have identified factors that impact disaster preparedness attitudes among the public. In Istanbul, higher education levels, earthquake experience, living in higher earthquake risk areas, home-ownership, and being between 35 and 54 years old were found to be positively correlated with disaster preparedness Tekeli-Yeşil *et al.* (2010). Similarly, in Florida, age, home ownership, house type, income, and race were found to have a significant impact on attitudes towards hurricane preparedness (Baker 2011). In Iran, income level, disaster experience, living area, and occupation were found to significantly affect preparedness scores (Najafi *et al.* 2015). Education level and disaster experience were found to strongly impact disaster preparedness in Thailand and the Philippines (Hoffmann and Muttarak 2017). In the US, informal support, income level, health status, and disaster experience were found to be correlated with disaster preparedness among elders (Kim and Zakour 2017). Similarly, it was discovered that disaster experience has positive impact on disaster preparedness of residents live in Japan (Onuma *et al.* 2017).

The impact of ethnicity on disaster preparedness has also been studied. Hispanics in the US were found to be less likely to conduct disaster preparedness, while income level, education, age, and disaster experience were positively related to preparedness (Donner and Lavariega-Montforti 2018). Gender, ethnic groups, age, medical conditions, healthcare access, number of children, income level, and evacuation experience were also found to significantly impact disaster preparedness attitudes among interviewees (Kyne *et al.* 2020). Lastly, gaining preparedness information or having disaster experience were found to increase the likelihood of disaster preparedness, while living in a rental

house or being Latino or Asian were negatively correlated with disaster preparedness (Rivera 2020). The summarized information of studies from the literature review is presented in Table 1.

While many factors have been found to be significant in shaping public attitudes toward disaster preparedness, this information alone is insufficient for relief agencies to address the diversity and complexity of the affected population. For instance, ethnicity and income level have been found to be positively related to disaster preparedness for those living in the Rio Grande Valley (Donner and Lavariega-Montforti 2018), but it is difficult to determine which community is most vulnerable among multiple communities with different ethnicities and income levels. Additionally, previous studies have focused on testing significance without evaluating model performance, such as prediction accuracy, leading to limited predictive capabilities. To address this gap, we apply multiple machine learning algorithms to FEMA NHS data to develop a more accurate and comprehensive estimation of residents' attitudes towards disaster preparedness. To address the generalizability challenge, we utilize the Monte Carlo Simulation approach to simulate residents' demographic features and use this data to train ML models to predict disaster preparedness in affected locations.

Our paper makes contributions to the literature from the following three perspectives:

- Our study presents a novel approach to disaster preparedness analysis by developing machine learning models that accurately predict household attitudes towards disaster preparedness. To our knowledge, this is the first time that machine learning algorithms such as XGBoost and artificial neural networks have been applied to this type of analysis.
- We have integrated the power of machine learning with a simulation approach to provide a systemic and accurate estimation of disaster preparedness levels for affected locations. This approach can be particularly useful in situations where the survey has not covered the affected areas, or the data in affected area is very limited.
- We further conducted sensitivity analysis to evaluate the impact of information awareness on disaster preparedness. The results show that higher information awareness improves disaster preparedness levels, which can help the relief agencies evaluate the effectiveness of their education and communication efforts in increasing public information awareness about disaster preparedness.

The remainder of this paper is organized as follows. Section 2 presents the proposed methodology for predicting the level of disaster preparedness. In Section 3, we present the results of extensive numerical tests to evaluate the efficiency of our proposed solution framework and assess the model's performance. We conclude the paper in Section 4 and suggest future research directions.

Table 1. Summary of the reviewed literature.

Paper	Regions	Year	Obser.	Meth.	Independent variable														
					DE	IA	A	I	G	SL	R	MS	C	EL	HL	HO	HT	JT	
Basolo <i>et al.</i> (2009)	NO, LA	1999	404	Log.	✓	✓ ⁺	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	
Tekeli-Yeşil <i>et al.</i> (2010)	Istanbul	2007	1123	Log.	✓	✓ ⁺	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	
Baker (2011)	Florida	2006	1200	Chi		✓ ⁺	✓ ⁺	✓	✓		✓ ⁺								
Najafi <i>et al.</i> (2015)	Iran	2014	1250	Lin.	✓ ⁺		✓ ⁺	✓	✓					✓	✓	✓	✓	✓ ⁺	
Hoffmann and Muttarak (2017)	SAC	2014	2199	Log.	✓ ⁺		✓	✓	✓			✓	✓ ⁺	✓ ⁺	✓	✓	✓	✓ ⁺	
Kim and Zakour (2017)	US	2014	719	Log.			✓	✓	✓		✓								
Onuma <i>et al.</i> (2017)	Japan	2013	20726	Lin	✓ ⁺		✓ ⁺	✓ ⁺	✓ ⁺				✓ ⁺						
Cox and Kim (2018)	US	2010	1711	Lin		✓ ⁺	✓ ⁺	✓ ⁺	✓ ⁺			✓ ⁺		✓	✓	✓	✓		
Donner and Lavariega-Montforti (2018)	RGV	2014	740	Lin	✓ ⁺		✓ ⁺	✓ ⁺	✓			✓							
Cvetković <i>et al.</i> (2018)	Serbia	2015	2500	Lin		✓ ⁺	✓ ⁺	✓ ⁺	✓ ⁺			✓ ⁺		✓ ⁺	✓ ⁺	✓ ⁺	✓ ⁺		
Nukpeah and Souja (2018)	US	2008	1137	Log.		✓ ⁺	✓ ⁺	✓ ⁺	✓ ⁺			✓ ⁺							
Kyne <i>et al.</i> (2020)	RGV	2017	590	Chi	✓ ⁺		✓ ⁺	✓ ⁺	✓ ⁺			✓ ⁺							
Rivera (2020)	US	2018	5003	Log.	✓ ⁺	✓ ⁺	✓	✓	✓		✓	✓ ⁺							
This paper	GC	17,18	1604	ML	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓ ⁺	✓ ⁺	

Meth. - Methodology; Lin. - Linear regression; Log. - Logistic regression; Chi. - Chi-Square test. NO, LA - New Orleans and Los Angeles; SD - San Diego, CA; SAC: Southeast Asia Countries, Thailand and Philippines; GC - Gulf Coast, US; RGV - Rio Grande Valley, Texas. DE - Disaster experience; IA - Information accessibility; I - Income; A - Age; G - Gender; R - Race; SL - Speaking language; MS - Marital status; C - Children in home. EL - Education level; HL - Home location; HO - Home ownership; HT - House type; JT - Job type. ✓ - Analyzed independent variable; ✓⁺ - Significant independent variable.

2. Methodology

The proposed framework is summarized in the Figure 1. There are four modules in the proposed framework including (1) Data Collection and Cleaning, (2) Model Training and Selection, (3) Residents' Feature Simulation, and (4) Preparedness Level Prediction, which will be introduced in details in the following subsections.

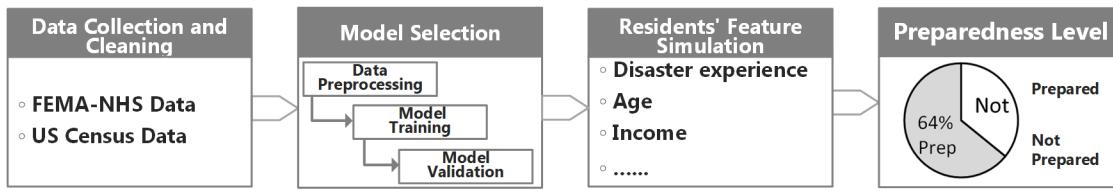


Figure 1. Flowchart of Proposed Framework

2.1. Data Collection and Cleaning

FEMA administers the National Household Survey (NHS) to investigate the American public's progress in personal disaster preparedness through telephone interviews. The survey consists of three major components: attitudes toward hazard preparedness, hazard experience, and demographic information. The interviewee's attitude about disaster preparedness behavior is classified into five levels: Pre-contemplation, Contemplation, Preparation, Action, and Maintenance. In our study, we consider individuals in the Action and Maintenance levels to have prepared for disasters, while those in the Pre-contemplation, Contemplation, and Preparation levels are deemed unprepared. We thus convert the multi-level categorical variable into a binary one (i.e., prepared vs. unprepared), and use it as the dependent variable in our predictive modeling analysis.

We retrieved the 2017, 2018, and 2019 FEMA NHS records from the official website ([FEMA 2021, 2019, 2020](#)) and manually selected seven states (Florida, Texas, Louisiana, South Carolina, Alabama, Georgia, and Mississippi) that were affected by hurricanes between 2015 and 2020 ([Heil 2019](#)) for analysis. To construct the machine learning models, we selected demographic features, including age, income, gender, education level, ethnicity, disaster experience, and awareness of information, based on previous studies that established a significant relationship between these variables and disaster preparedness ([Onuma et al. 2017, Basolo et al. 2009, Heller et al. 2005, Bourque 2015, Rivera 2021](#)), and the availability of NHS data. In addition, we kept the state name and zip codes from the raw data sets to identify the interviewee's address and evaluate the simulation results. The training data set comprised NHS 2017 and 2018, and NHS 2019 served as the testing data set to assess the model's performance and the predicted level of disaster preparedness accuracy. We removed observations with invalid information caused by the answer "Refused" or "Do not know" and matched the interviewees' zip code location with their state name; any unmatched observations were removed. The variables used in the study and their data types are summarized in Table 2, with the dependent variable referred to as Preparedness for brevity. Disaster experience, awareness of information, and education level are denoted as Experience, Information, and Education, respectively.

Table 2. Summary of Variables for Model Training

Variable	Definition of variable	Data type
Preparedness	Whether the household prepare for disaster	Binary
Experience	Whether the household experience a disaster	Binary
Information	Whether the household received disaster prevention information	Binary
Age	Demographic feature: interviewee's age	Integer
Income	Demographic feature: interviewee's income per year	Factor
Gender	Demographic feature: interviewee's gender	Binary
Education	Demographic feature: interviewee's levels of educational attainment	Factor
Ethnicity	Demographic feature: interviewee's ethnicity	Factor

2.2. Model Training and Selection

In this study, we utilized six different machine learning models, namely Logistic Regression (LR), Random Forest (RF), eXtreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Artificial Neural Network (ANN) to examine the relationship between public attitudes about disaster preparedness and the introduced variables. Logistic Regression is a linear model that is commonly used to predict binary outcomes, while Random Forest is an ensemble model that combines multiple decision trees to improve prediction performance. XGBoost is a gradient boosting model that utilizes decision trees and is particularly effective for large datasets. SVM is a binary classification model that seeks to find the best hyperplane that separates the different classes of data, while KNN is a non-parametric model that makes predictions based on the nearest neighbors of a given data point. Finally, Artificial Neural Network is a set of interconnected nodes organized into layers that can be used to make complex predictions. The choice of which machine learning model to employ depends on a variety of factors, including prediction accuracy and interpretability. Interested readers can refer to the classical machine learning book by [Hastie *et al.* \(2009\)](#) for more detailed descriptions of each of these models.

The module selection process begins with data pre-processing to transform predictor variables, such as Income, Education, and Ethnicity into binary dummy variables. The variable Age is re-scaled using min-max normalization to convert it between 0 and 1 with the following equation:

$$\text{Age}^* = \frac{\text{Age} - \min(\text{Age})}{\max(\text{Age}) - \min(\text{Age})}$$

When training the ML model, where a grid search approach and K-Fold Cross Validation are used to optimize hyperparameters. In the CV process, the training data is randomly and evenly divided into K folds, and the model is trained on $K - 1$ folds and tested on the remaining fold. This process is repeated K times to obtain an average accuracy. In our study, we use ten folds. The optimal parameter setting is obtained by comparing the CV accuracy obtained on each parameter grid.

The final step is model validation, where a hold-out test is used to evaluate model performance. We obtain the predicted preparedness attitude of each household based on the 2019 NHS data and use Accuracy, AUC, F-1 score, and Specificity to evaluate model performance. More technical details about model training will be illustrated in Section 3.

2.3. Residents' Feature Simulation

Monte Carlo simulation is a widely used technique in many fields, including engineering, physics, economics, and finance, to estimate the distribution of an uncertain variable ([Mooney 1997](#)). In essence, Monte Carlo simulation involves generating a large number of data points, usually following a probabilistic distribution, to simulate a real-world scenario. The generated data can then be used to assess the likelihood of certain outcomes or to inform decision-making processes. In the context of disaster preparedness, Monte Carlo simulation can be used to generate demographic profiles of residents living in an affected location. This is particularly useful when survey data (e.g., NHS data) on the demographics of the affected population is limited or not available.

In this module, we utilize the Monte Carlo simulation approach to simulate data points for selected demographic feature of residents in the affected county. To accomplish this, we first collect distribution information about residents' demographic features, such as age, income, education level, ethnicity, and gender, from the United States Census Bureau (USCB). To perform Monte Carlo simulation, a probabilistic distribution is selected based on available data or expert knowledge. For example, the distribution of age can be estimated from the USCB data and represented as a normal distribution. The simulation then generates a large number of data points, typically using a random number generator, that follow the selected distribution.

Additionally, we denote the “Experience” and “Information” as two variables following binary probability distributions, whose expected probability is respectively calculated the percentage of residents who have disaster experience and those who are aware of disaster preparedness information based on the 2017 and 2018 NHS data. Once a sufficient number of data points have been generated, they can be combined to create a demographic profile of a resident living in the affected location. This process is repeated multiple times, resulting in various simulated data of demographic profiles of residents with disaster-related experience and information awareness, which can be used as a good representation of the residents in affected areas with limited data. This approach helps ensure the generalizability of the proposed framework by accounting for the variability and uncertainty in the demographic features of residents in the affected areas.

2.4. Preparedness Level Prediction

In the Preparedness Level Prediction module, we utilize the top-performing model from the second module to predict the level of disaster preparedness for the impacted area. To do so, we construct data about the residents’ information for the affected location using Monte Carlo simulation in the third module. By feeding the simulated data set as the input data for the selected ML model, we can systematically and accurately determine whether a household in the affected location is engaged in disaster preparedness. By taking an average of each resident’s preparedness status, we can calculate the disaster preparedness level (as a percentage of total number of households) for the affected location.

To ensure robustness, we generate multiple simulated datasets in the third module, enabling us to obtain a disaster preparedness level for each data set by repeating the above process, and eventually obtain a distribution of preparedness level across all simulated datasets. Finally, we can determine the most preferred disaster preparedness level for the affected location by analyzing a histogram of the obtained disaster preparedness level.

3. Results and Discussion

3.1. Training Data Set Exploration

Following our data collection schemes introduced in Section 2.1, after performing data cleaning, we obtain 1604 observations in the training dataset (e.g., from NHS 2017 and 2018 dataset) and 864 observations in the testing dataset (e.g., from NHS 2019 dataset). Table 3 summarizes some descriptive statistics of variables in the training dataset. From Table 3, we can see that the majority of households (61.16%) have taken some action to prepare for potential disasters. Additionally, a significant proportion of families (72.63%) have experienced a disaster in the past, while only 60.29% have received information about disaster preparedness.

The age range of the interviewees is between 18 and 95 years old, with a mean age of 50.52 and a median age of 51. The income of the households varies widely, with 58% of interviewees earning more than \$4000 in household income, while 7.23% earn less than \$1000. Furthermore, most interviewees (72%) hold a college or higher degree. In terms of race, the majority of interviewees are white people (75%), followed by African-Americans (20.26%), and other races make up the remaining 5%. These demographic factors can help us gain preliminary insights into the disaster preparedness levels of households in the affected location.

Table 3. Training data set summary.

Preparedness		Income	
No	38.84%	Below 999	7.23%
Yes	61.16%	1000-3999	34.29%
Experience		4000-9999	
No	27.37%	Over 10000	40.03%
Yes	72.63%		18.45%
Information		Education	
No	39.71%	Less High School	5.67%
Yes	60.29%	High School	17.96%
Age		Vocational School	
Min	18	College	5.42%
Mean	50.52	College Graduate	24.31%
Median	51	Post Graduate	28.74%
Max	95		17.89%
Gender		Ethnicity	
Men	50.12%	White	75.00%
Women	49.88%	African-American	20.26%
		Asian	2.12%
		American Indian	1.68%
		Hawaiian	0.94%

3.2. Model Performance Evaluation

In this study, we utilized 10-fold cross-validation and accuracy as the evaluation metric to optimize the hyperparameters of our machine learning models. Table A1 in the Appendix lists the candidate and resulting tuned model parameters. We further evaluated the performance of the model on the testing dataset using additional metrics, including AUC (Area Under the ROC Curve), F-1 score, and Specificity. To be self-contained, we briefly describe each of these metrics below.

1. Accuracy: Accuracy is a widely used evaluation metric for classification models. It is defined as the proportion of correctly classified instances to the total number of instances. In other words, it measures the overall correctness of a model's predictions. The value of Accuracy is computed as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

where TP, TN, FP, FN represent True Positive, True Negative, False Positive, and False Negative, respectively.

2. AUC: The Area Under the ROC Curve (AUC) is a performance metric used to evaluate binary classification models at different classification thresholds. The AUC is computed by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) at different threshold values and then calculating the area under the resulting curve. The AUC ranges from 0 to 1, with higher values indicating better performance. A model with an AUC of 0.5 is equivalent to random guessing, while a model with an AUC of 1 is perfect. In summary, the AUC is a useful metric for evaluating the overall performance of a binary classification model.
3. F-1 score: The F-1 score is a popular evaluation metric for binary classification models that combines both precision and recall into a single score. Precision measures the proportion of correctly predicted positive instances out of all predicted positive instances, while recall measures the proportion of correctly predicted positive instances out of all actual positive instances. The harmonic mean of precision and recall is used to calculate the F-1 score, which ranges from 0 to 1, with higher values indicating better performance. A perfect classifier would have an F-1 score of 1, while a completely random classifier would have an F-1 score of 0. The formulation to compute the F-1 score is given below.

$$\text{F-1 score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

where precision = TP/(TP + FP) and recall = TP/(TP + FN).

4. Specificity: Specificity measures the proportion of true negative instances that are correctly identified by the model. It is calculated by dividing the number of true negative instances by the sum of true negative and false positive instances. In other words, specificity measures the model's ability to correctly identify negative instances as negative. A high specificity score indicates that the model is very good at avoiding false positives (i.e., it has a low rate of false alarms) and is useful in applications where the cost of a false positive is high, such as the classification of household preparedness in the disaster context as in our problem.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

Table 4 provides a summary of the training CV accuracy and hold-out testing performance of each of the six proposed ML models, which are evaluated using the above-mentioned metrics. The threshold used to calculate the CV and hold-out test accuracy is 0.5. The bold numbers in the table represent the top three highest values of each evaluation metric.

Table 4. Model Performance

	Training		Hold Out Test		
	Accuracy	Accuracy	AUC	F-1 Score	Specificity
LR	0.716	0.704	0.741	0.769	0.770
RF	0.708	0.700	0.727	0.770	0.800
XGBoost	0.719	0.704	0.745	0.772	0.786
SVM	0.702	0.686	0.728	0.761	0.783
KNN	0.645	0.652	0.665	0.731	0.759
ANN	0.720	0.686	0.734	0.749	0.732

From the results presented in Table 4, it is evident that LR, XGBoost, and ANN exhibit higher CV accuracy compared to the other three ML models. Among the six models, ANN achieves the highest CV accuracy of 72%, which is 1% and 2% higher than XGBoost and LR, respectively. The CV accuracy of RF and SVM does not exceed 71%, but it is over 70%. In contrast, KNN performs the worst among the six models, with a CV accuracy that does not exceed 65%.

In terms of hold-out testing performance, XGBoost outperforms all other models in all evaluation metrics except specificity. For instance, XGBoost has an AUC value of 0.745, which is 0.004, 0.018, 0.017, 0.08, and 0.011 higher than the AUC value of LR, RF, SVM, KNN, and ANN, respectively. Moreover, XGBoost has the second-highest specificity value, which is 0.014 lower than that of RF. The LR model also performs well in the disaster preparedness prediction hold-out test compared to RF, SVM, KNN, and ANN. Although the specificity value of the LR model is lower than that of RF, XGBoost, and SVM, the prediction accuracy (0.716) and AUC (0.741) of LR are the second-highest among the proposed models.

The RF and ANN models also perform well in some evaluation metrics but do not dominate over LR and XGBoost in all evaluation measures. Overall, XGBoost presents the best performance on all evaluation metrics for the hold-out test, making it the best model among the six. As such, we will employ LR and XGBoost for further analysis in the following sections, given that LR can present a more intuitive relationship between the dependent and independent variables.

To conduct further analysis, we will mainly focus on the results of logistic regression and XGBoost models considering the trade-off between prediction accuracy and interpretability. Logistic regression offers a simple and transparent equation, which can be used to interpret the impact of each predictor on individual attitudes toward disaster preparedness. Although it may not provide the most accurate predictions, it is an effective method to understand the relationship between the predictors and the outcome. On the other hand, XGBoost is a black-box model that provides high prediction accuracy when estimating individual attitudes about disaster preparedness. However, due to the ensemble

nature of XGBoost, it is challenging to obtain a transparent interpretation of the relationship between predictors and the outcome variable.

3.3. Feature Importance and Significance Analysis

This subsection focuses on the feature significance of the LR model and the importance of XGBoost. The coefficient and feature significance of the LR model are summarized in Table 5. The Disaster Experience (Yes) and Awareness of Information (Yes) predictors are both significant and positively correlated with disaster preparedness. The results also show that the elderly are more likely to prepare for disasters. Income is another essential factor that affects disaster preparedness, with those earning a monthly income of over \$4000 being more likely to prepare than those earning under \$999. Gender and education backgrounds do not significantly affect people's perceptions of disaster preparedness based on the LR results. In terms of race, Asian people are less likely to be prepared for disasters compared to white people, while the preparedness of African-American, American Indian, or Hawaiian people does not significantly differ from that of white people.

Table 5. Coefficients and Significance of Logistic Predictors

	Coefficient	Std. Error	Z Value	P Value
Intercept	-2.327	0.343	-6.790	1.09E-11
Experience (Yes)	0.941	0.130	7.242	4.44E-13
Information (Yes)	1.227	0.121	10.126	< 2E-16
Age	2.100	0.271	7.746	9.45E-15
Income (1000-3999)	0.302	0.234	1.292	0.196
Income (4000-9999)	1.145	0.243	4.710	2.47E-06
Income (Over 10000)	1.549	0.277	5.594	2.21E-08
Gender (Men)	0.169	0.119	1.415	0.157
Education (High School)	-0.311	0.275	-1.131	0.258
Education (Vocational School)	-0.048	0.347	-0.137	0.891
Education (College)	-0.184	0.271	-0.679	0.497
Education (College Graduate)	-0.337	0.276	-1.222	0.222
Education (Post Graduate)	-0.526	0.294	-1.787	0.074
Ethnicity (African-American)	-0.222	0.145	-1.535	0.125
Ethnicity (Asian)	-0.940	0.399	-2.355	0.0185
Ethnicity (American Indian)	0.031	0.452	0.068	0.946
Ethnicity (Hawaiian)	0.121	0.613	0.197	0.844

The XGBoost model has been used to compute the importance of each predictor, and the results are depicted in Figure 2. The importance values of the predictors reveal that Awareness of Information has the highest Gain importance value, which is over 0.3, making it the most significant predictor. The predictors Age and Disaster Experience (Yes) have the second and third highest importance values, with 0.237 and 0.14, respectively. The importance of Income (4000-9999) and Income (Over 10000) is also high, being close to 0.1. These predictors also have low P-Values in the LR results, which further support their significance in predicting disaster preparedness. On the other hand, the importance of Ethnicity (African-American), Gender (Men), Ethnicity (Asian), Education Level (High School), and Income (1000-3999) is lower, with values close to 0.02. The importance of the remaining predictors does not exceed 0.001. It is worth noting that these predictors are still considered significant, but their contribution to predicting disaster preparedness is relatively lower compared to the other predictors.

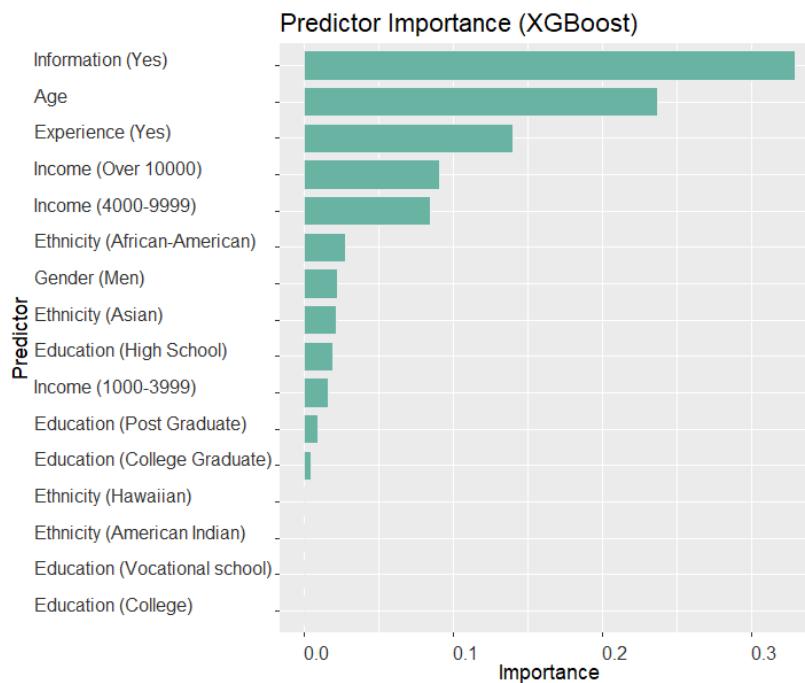


Figure 2. Predictors Importance.

The results from the XGBoost model and the LR model exhibit some similar insights in terms of feature importance. Out of the six predictors that show significance in the LR model, five of them (Awareness of Information (Yes), Age, Disaster Experience (Yes), Income (4000-9999), and Income (Over 10000)) also demonstrate relatively higher importance values in the XGBoost model.

The LR model indicates that if a household member has gained information about disaster preparedness, it is positively associated with preparedness. This finding is consistent with the XGBoost model, which shows that the Awareness of Information predictor has the highest importance value among all the predictors. These results suggest that FEMA's public awareness campaign, called "Ready," which provides disaster preparedness information through various channels such as television and website, is an effective strategy to increase preparedness. In addition, the LR model also shows that disaster experience is positively associated with preparedness, which is consistent with the XGBoost model, where the Disaster Experience (Yes) predictor is the third important feature. This indicates that individuals who have previously experienced disasters are more likely to be prepared for future disasters.

According to the LR results, age is positively associated with disaster preparedness, and it is the second most important feature in the XGBoost model. This suggests that older individuals are more likely to engage in disaster preparedness activities. One possible explanation for this trend is that the survey data show that older people are more likely to have received information about disaster preparedness than younger people. Specifically, the survey results indicate that 70% of people aged 65 years or older have learned about disaster preparedness, whereas only 43% of individuals surveyed by FEMA-NHS (FEMA 2020) reported having gained knowledge on the subject. This finding suggests that targeted campaigns to provide information and resources for disaster preparedness to younger age groups could be beneficial in improving overall disaster preparedness rates.

According to the XGBoost model, Income (4000-9999) and Income (Over 10000) are the fourth and fifth most important variables, respectively. The LR model also indicates that both of these income brackets have a positive relationship with disaster preparedness. These findings align with our previous analysis, which showed that middle and high-income households are more likely to engage in disaster preparedness activities compared to those with lower incomes. This may be because higher-income households have more resources and are better able to afford emergency supplies,

insurance, and other preparations. Additionally, they may have more education and awareness of the importance of disaster preparedness.

Although ethnicity (Asian) is a significant predictor in the LR model, it is not a top-important variable in the XGBoost model. This may be because the surveyed Asian people are generally younger and have less disaster experience. Specifically, the training data set shows that the average age of the Asian interviewees is 36.53, which is much younger than the average age of the other ethnic groups. Additionally, the percentage of surveyed Asian people who have disaster experience is 58.8%, which is lower than the percentage of white people who have experienced a disaster, which is 73.2%. These factors could contribute to the lower importance value of ethnicity (Asian) in the XGBoost model, indicating that it may have a weaker relationship with disaster preparedness compared to other predictors.

3.4. Preparedness Level Prediction

In this section, we use Miami-Dade County (FL) as an example to predict its preparedness level based on 1000 simulated datasets, each containing 1000 simulated residents with demographic features and disaster-related features. The demographic features include Age, Income, Gender, Education Level, and Ethnicity, which are simulated based on the distribution obtained from the *2019 American Community Survey Single-Year Estimates* (US Census Bureau 2020). The probability of whether a person has a disaster experience and awareness of information are set to 0.72 and 0.43, respectively, which are observed from the FEMA NHS 2018 (FEMA 2020).

The density plot of the predicted disaster preparedness level by the LR model is presented in Figure 3. The blue dashed line in the figure represents the overall preparedness level (63.5%) of Miami-Dade County, as observed from the FEMA NHS 2019 (FEMA 2021). The red dashed line interval indicates the 95% confidence interval (CI) of the predicted preparedness level. The histogram is symmetric and has a peak at 0.645, which is the percentage of preparedness. The 95% CI ranges from 0.63 to 0.66, covering the Miami-Dade preparedness level from the real survey data. However, the observed preparedness level is not close to the axis of symmetry and is on the left-hand side. This indicates that the predicted preparedness level is higher than the surveyed value in most scenarios. Therefore, the results of the LR model are more optimistic than the surveyed preparedness level in Miami-Dade County. Figure 4 displays the preparedness level density obtained by the XGBoost model, as a comparison benchmark. The histogram of XGBoost is also symmetric, with a peak near 0.63, which is smaller than that of the LR model. The 95% CI ranges from 0.615 to 0.65, containing the surveyed Miami-Dade preparedness level. Unlike the LR model, the surveyed preparedness level is very close to the axis of symmetry in XGBoost. Based on the plot, the proposed framework using XGBoost shows better ability to obtain predicted results that are closer to the surveyed value in Miami-Dade County. This result indicates that the XGBoost model is more accurate in predicting the preparedness level than the LR model.

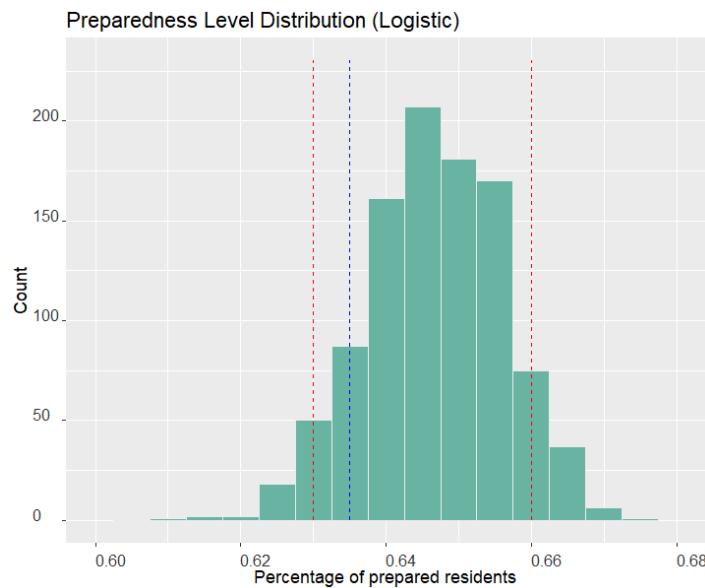


Figure 3. Predicted preparedness level (LR).

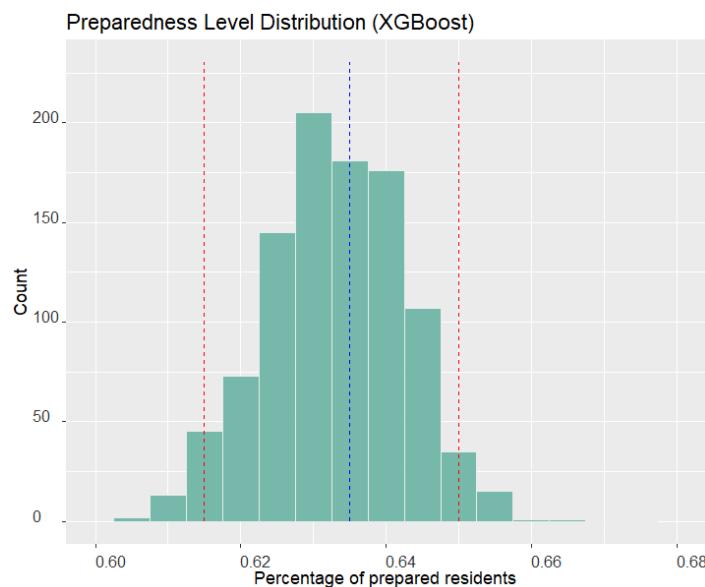


Figure 4. Predicted preparedness level (XGBoost).

The density plots obtained by LR and XGBoost for Miami-Dade County show that the results of XGBoost are closer to the surveyed preparedness level. One reason for this is that the distribution of some demographic factors in the training dataset may differ from the actual distribution in the Miami-Dade County population. For instance, Table 5 shows that all other ethnicities, except white people, have negative coefficients, indicating that they are less likely to conduct disaster preparedness. However, around 75% of Miami-Dade residents are white people, according to the 2019 American Community Survey Single-Year Estimates ([US Census Bureau 2020](#)). This difference in the distribution of the ethnicity factor may lead to the predicted preparedness level by the LR model being higher than the actual level. In contrast, XGBoost has identified the top five predictors that have the most significant power to affect the preparedness level. The summation of the importance values of these five predictors is about 0.88, which means that these factors are more critical for the prediction of the preparedness level than the other factors. The ethnicity factor is not among the top five predictors in

the XGBoost model, which may explain why the preparedness level predicted by XGBoost is closer to the surveyed value than LR.

3.5. Sensitivity Analysis for Public Information Awareness

In the context of disaster management, public information awareness about disaster preparedness refers to the extent to which individuals have knowledge and understanding of how to prepare for, respond to, and recover from natural disasters. The level of public information awareness about disaster preparedness can vary based on several factors, such as the frequency and type of disasters in a given region, the quality and accessibility of information provided by government agencies and other organizations, and individual attitudes and beliefs about the likelihood and impact of disasters. Effective communication and education initiatives can play a vital role in enhancing the level of public information awareness about disaster preparedness. By providing accessible and reliable information through multiple channels, including social media, television, and websites, individuals can become better equipped to handle the consequences of disasters. Moreover, individuals with higher levels of public information awareness about disaster preparedness are better able to make informed decisions and take appropriate actions during and after disasters. This, in turn, can help reduce the negative impact of disasters on individuals, families, and communities.

We next perform a sensitivity analysis on the awareness of information variable to investigate the impact of information awareness on disaster preparedness. The sensitivity analysis allows relief agencies to determine the effectiveness of their education and communication efforts in increasing public awareness about disaster preparedness. Specifically, it can help agencies identify which communication methods are most effective and which areas require more attention to improve public awareness. The results of the sensitivity analysis can also be used to justify the allocation of resources to disaster preparedness education and communication efforts.

Figures 5 and 6 illustrate how the disaster preparedness level changes when the percentage of residents who are aware of disaster preparedness information varies from 38% to 48%. As the percentage of information awareness decreases from 43% to 38%, both LR and XGBoost models show a leftward shift in the density curve, indicating a decrease in the disaster preparedness level. Conversely, when the percentage increases from 43% to 48%, the density curve shifts to the right, indicating an increase in the disaster preparedness level. These results suggest that both LR and XGBoost models are sensitive to changes in the percentage of residents who have awareness of disaster preparedness information. Small increases in this percentage can lead to significant improvements in disaster preparedness levels, which can be a compelling argument for investing in education and communication efforts.

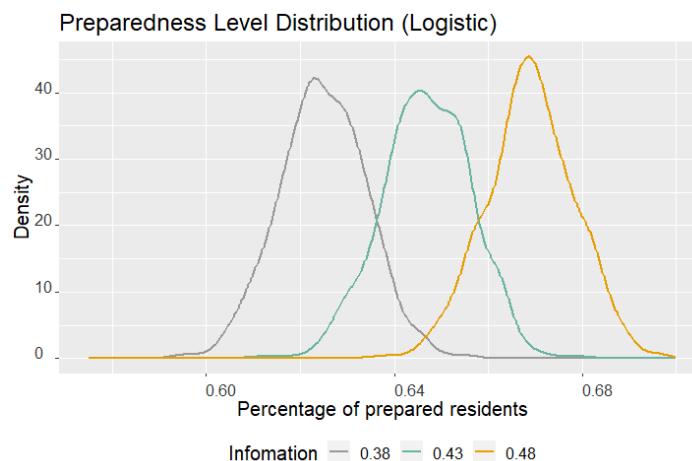


Figure 5. Predicted preparedness level comparison (LR).

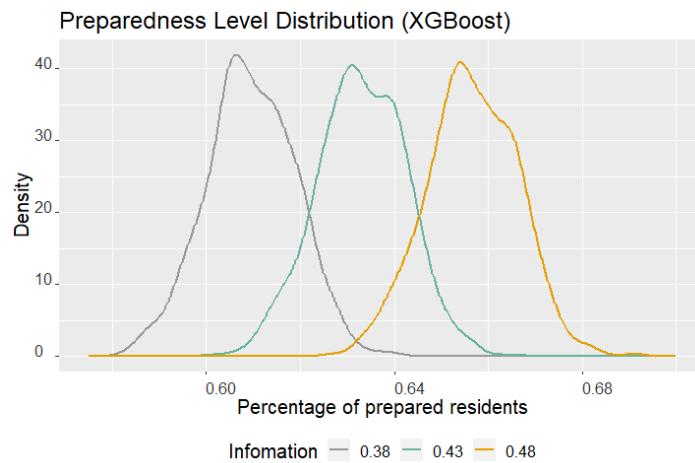


Figure 6. Predicted preparedness level comparison (XGBoost).

4. Conclusions

In this study, we propose a framework that combines machine learning models and simulation approaches to predict the level of disaster preparedness at the county level. Our framework uses 2017-2019 FEMA NHS data for model training and testing and simulates demographic features based on 2019 census data of Miami-Dade County. The results show that both the LR and XGBoost models can predict the level of disaster preparedness in Miami-Dade County, but XGBoost is more accurate than the LR model. The importance analysis of the predictors conducted in our study reveals that awareness of information, age, disaster experience, and income are significant predictors of disaster preparedness. These findings can inform public policymakers and emergency managers about the factors that influence the preparedness level of the population and can guide the development of targeted intervention programs.

Our framework provides a novel and effective approach for predicting disaster preparedness levels in a specific geographical area. It can help local authorities identify areas that are most vulnerable to disasters and design strategies to improve their level of preparedness. We believe that our framework can be useful for disaster management agencies and emergency responders in making more informed decisions and taking proactive measures to prevent or mitigate the impact of disasters. Future research can further refine the framework by incorporating additional demographic factors and expanding the geographic scope of the study to other regions.

Author Contributions: Conceptualization, Zhenlong Jiang, Yudi chen, Wenying Ji, and Ran Ji; methodology, Zhenlong Jiang, Yudi chen, Wenying Ji, and Ran Ji; software, Zhenlong Jiang, and Ting-Yeh Yang; validation, Zhenlong Jiang, and Ting-Yeh Yang; formal analysis, Zhenlong Jiang; data curation, Zhenlong Jiang; writing—original draft preparation, Zhenlong Jiang; writing—review and editing, Zhenlong Jiang, Yudi chen, Wenying Ji, Zhijie Dong, and Ran Ji; visualization, Zhenlong Jiang; supervision, Ran Ji; project administration, Ran Ji. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: <https://www.fema.gov/about/openfema/data-sets/national-household-survey>.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1 show the hyper parameters tuned for each ML models, respectively. In the table, the first and second columns list the model and parameter names, while the third column illustrates the candidate values for grid search. The last column shows the optimal value for selected model. The models are obtained in R 4.1.2 with package caret 6.0, xgboost 1.5.0.2, e1071 1.7, nnet 7.3.

Table A1. Parameter setting

Models	parameter	Candidate Parameter	Selected Parameter
Random Forest	ntrees	300, 400, 500, 600, 700	500
	nodesize	40, 50, 60, 70	60
	mtry	1, 2, 3, 4, 5	3
	nrounds	30, 35, 40, 45, 50	45
XGBoost	max_depth	1, 2, 3	1
	eta	0.2, 0.3, 0.4, 0.5, 0.6	0.4
	gamma	0.1, 0.15, 0.2, 0.25, 0.3	0.2
	colsample_bytree	0.3, 0.5, 0.7, 1	0.5
	min_child_weight	0.5, 1, 3, 5	1
	subsample	0.3, 0.5, 0.7, 1	0.5
	C	3, 5, 7, 9	7
SVM	sigma	0.005, 0.01, 0.015, 0.02	0.01
	K	2, 3, 4, 5, 6, 7, 8	5
KNN	size	1, 2, 3, 4, 5	2
ANN	decay	0.05, 0.1, 0.15, 0.2, 0.25	0.1

References

Rawls, C.G.; Turnquist, M.A. Pre-positioning of emergency supplies for disaster response. *Transportation Research Part B: Methodological* **2010**, *44*, 521–534.

Elçi, Ö.; Noyan, N. A chance-constrained two-stage stochastic programming model for humanitarian relief network design. *Transportation Research Part B: Methodological* **2018**, *108*, 55–83.

Jiang, Z.; Ji, R.; Dong, S. A distributionally robust chance-constrained model for humanitarian relief network design. Available at SSRN 3929286 **2021**.

Falkiner, L. Impact analysis of the Canadian red cross expect the unexpected program. Retrieved August 2003, 15, 2011.

Diekman, S.T.; Kearney, S.P.; O'neil, M.E.; Mack, K.A. Qualitative study of homeowners' emergency preparedness: Experiences, perceptions, and practices. *Prehospital and Disaster Medicine* **2007**, *22*, 494–501.

Multihazard Mitigation Council. Natural hazard mitigation saves: 2017 interim report, 2018. https://www.fema.gov/sites/default/files/2020-07/fema_ms2_interim_report_2017.pdf, accessed 15.3.2022.

Howe, P.D. Modeling geographic variation in household disaster preparedness across US states and metropolitan areas. *The Professional Geographer* **2018**, *70*, 491–503.

FEMA. Preparedness research, 2021. <https://www.fema.gov/about/openfema/data-sets/national-household-survey>, accessed 15.3.2022.

Onuma, H.; Shin, K.J.; Managi, S. Household preparedness for natural disasters: Impact of disaster experience and implications for future disaster risks in Japan. *International Journal of Disaster Risk Reduction* **2017**, *21*, 148–158.

Cvetković, V.M.; Roder, G.; Öcal, A.; Tarolli, P.; Dragičević, S. The role of gender in preparedness and response behaviors towards flood risk in Serbia. *International Journal of Environmental Research and Public Health* **2018**, *15*, 2761.

Tekeli-Yeşil, S.; Dedeoğlu, N.; Braun-Fahrlaender, C.; Tanner, M. Factors motivating individuals to take precautionary action for an expected earthquake in Istanbul. *Risk Analysis: An International Journal* **2010**, *30*, 1181–1195.

Najafi, M.; Ardalani, A.; Akbarisari, A.; Noorbala, A.A.; Jabbari, H. Demographic determinants of disaster preparedness behaviors amongst Tehran inhabitants, Iran. *PLoS currents* **2015**, *7*.

FEMA. 2019 national household survey, 2021. <https://www.fema.gov/about/openfema/data-sets/national-household-survey>, accessed 15.3.2022.

Baker, E.J. Household preparedness for the aftermath of hurricanes in Florida. *Applied Geography* **2011**, *31*, 46–52.

Basolo, V.; Steinberg, L.J.; Burby, R.J.; Levine, J.; Cruz, A.M.; Huang, C. The effects of confidence in government and information on perceived and actual preparedness for disasters. *Environment and Behavior* **2009**, *41*, 338–364.

Hoffmann, R.; Muttarak, R. Learn from the past, prepare for the future: Impacts of education and experience on disaster preparedness in the Philippines and Thailand. *World Development* **2017**, *96*, 32–51.

Kim, H.; Zakour, M. Disaster preparedness among older adults: Social support, community participation, and demographic characteristics. *Journal of Social Service Research* **2017**, *43*, 498–509.

Donner, W.R.; Lavariega-Montforti, J. Ethnicity, income, and disaster preparedness in deep South Texas, United States. *Disasters* **2018**, *42*, 719–733.

Kyne, D.; Cisneros, L.; Delacruz, J.; Lopez, B.; Madrid, C.; Moran, R.; Provencio, A.; Ramos, F.; Silva, M.F. Empirical evaluation of disaster preparedness for hurricanes in the Rio Grande Valley. *Progress in Disaster Science* **2020**, *5*, 100061.

Rivera, J.D. The likelihood of having a household emergency plan: Understanding factors in the US context. *Natural Hazards* **2020**, *104*, 1331–1343.

Cox, K.; Kim, B. Race and income disparities in disaster preparedness in old age. *Journal of Gerontological Social Work* **2018**, *61*, 719–734.

Nukpezah, J.A.; Souja, I. Creating emergency prepared households—What really are the determinants of household emergency preparedness? *Risk, Hazards & Crisis in Public Policy* **2018**, *9*, 480–504.

FEMA. 2017 national household survey, 2019. <https://www.fema.gov/about/openfema/data-sets/national-household-survey>, accessed 15.3.2022.

FEMA. 2018 national household survey, 2020. <https://www.fema.gov/about/openfema/data-sets/national-household-survey>, accessed 15.3.2022.

Heil, J. 10 U.S. states where hurricanes hit most often, 2019. <https://universalproperty.com/united-states-where-hurricanes-hit-most>, accessed 15.3.2022.

Heller, K.; Alexander, D.B.; Gatz, M.; Knight, B.G.; Rose, T. Social and personal factors as predictors of earthquake preparation: The role of support provision, network discussion, negative affect, age, and education 1. *Journal of Applied Social Psychology* **2005**, *35*, 399–422.

Bourque, L.B. Demographic characteristics, sources of information, and preparedness for earthquakes in California. *Earthquake Spectra* **2015**, *31*, 1909–1930.

Rivera, J.D. Factors influencing individual disaster preparedness information seeking behavior: Analysis of US households. *Natural Hazards Review* **2021**, *22*, 04021042.

Hastie, T.; Tibshirani, R.; Friedman, J.H.; Friedman, J.H. *The elements of statistical learning: Data mining, inference, and prediction*; Vol. 2, Springer, 2009.

Mooney, C.Z. *Monte carlo simulation*; Number 116, Sage, 1997.

US Census Bureau. 2019 American community survey single-year estimates, 2020. <https://www.census.gov/newsroom/press-kits/2020/acs-1year.html>, accessed 15.3.2022.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.