"There is no post-hurricane world"

Laura Brown

labr1398@colorado.edu

Alexander Gonzalez-Torres

algo9142@colorado.edu

Joo Young Han

joha4925@colorado.edu

PROBLEM STATEMENT

Natural disasters have profound effects on both the economic and social development of nations. Research on disaster impacts, such as those caused by Hurricane Maria in Puerto Rico, demonstrates how these disruptions can create long-term challenges in governance, human development, and economic recovery. The idea that "there is no post-hurricane world," as a 2022 exhibit at the Whitney Museum of American art put it, suggests that disasters like hurricanes, floods, and earthquakes fundamentally alter a society's trajectory, leaving lasting scars on economic performance and social well-being. However, whether these disruptions produce temporary setbacks or long-term changes requires further exploration, particularly as it concerns social development metrics that are well-documented yet often omitted from disaster research. For example, access to clean cooking fuels, regulatory quality, access to financial institutions, etc.

This project aims to fill this research gap by applying data mining and statistical techniques to a combined dataset that includes: 1. information on natural disasters between 1960 and 2023 2. Economic and human development metrics for this time period. The analysis will address questions such as:

- Do different types of natural disasters produce distinct economic and social impacts?
- Are these impacts temporary, or do they result in long-term structural changes?
- How do disasters affect key economic indicators such as GDP growth, public debt, and trade, as well as social outcomes like education, healthcare access, and gender equality?
- What attributes are shared among countries that recover more effectively from disasters?

Answering these questions will provide insight into how countries can mitigate the impact of natural disasters and areas where preparedness and disaster management can improve.

LITERATURE SURVEY

By the late 20th century, disaster studies emerged as an interdisciplinary field, integrating sociology, geography, economics, and public policy. Recent research has emphasized the global interconnectedness of economies and the increasing frequency and intensity of climate-related disasters. A review of the literature for Hurricane Maria, which made landfall in Puerto Rico on September 20, 2017, serves as a clear example and inspiration for this project. In the age of intensifying climate change and natural disasters increasing in frequency and magnitude, we hope to build on the foundation set by prior studies to distill from the data whether climate change has intensified any previously known effects of natural disasters.

In "The Economy of Disasters? Puerto Rico Before and After Hurricane Maria" Carabello-Cueto 2021, it is discussed that while Puerto Rico had been trending into rising debt beforehand - Hurricane Maria had immediate effect on fiscal planning and economic factors thereafter. Statements within the research primarily relied on short-term before and after

period trends to provide insights: such as debt payments doubled, unemployment insurance claims rose, poverty increased, increase in population migration, home reconstruction claims represented a majority of Puerto Rican households and the planned budget to education had been further reduced[1].

Within our dataset, we can explore this moment in Puerto Rico's history and compare against the current literature, as well as extend any insights from 2021 to 2023. One could even explore the community's resilience and assess upward trends in the economy (especially since Puerto Rico had been undergoing public debt before the hurricane). Although this is one regional example, literature in this field has been shifted towards a global scale [2][3][4].. Our project will follow that path with the EM-DAT and World Development Index international datasets. We plan to explore global patterns around short and long-term economic effects (grouped by disaster) and possibly confirm climate change disaster research.

PROPOSED WORK

The dataset for this project combines the complete World Development Index–a World Bank Group database–with entries between 1960 and 2023 in the The International Disaster Database–a project of the Center for Research on the Epidemiology of Disasters. Due to the size of the dataset, collection will require downloading it in subsets. These subsets must then be combined into a singular DataFrame using Pandas.

To eliminate redundancy and improve data integrity, database normalization techniques will be applied to the new DataFrame. In particular, new primary keys must be developed. Additionally, NumPy and SciPy statistical methods will be used to eliminate outlying/null values and reduce dimensionality where appropriate.

Furthermore, attributes that appear in both the World Development Index and International Disaster Database—such as dates—must be standardized in format.

The statistical methods of this project will focus on time-series analysis for all attributes to determine 1. their state before a natural disaster 2. their state for years after. We acknowledge that, in some cases, we must control for global financial events like oil price crises, the Great Recession, etc. Additionally, we must control for political instability unrelated to natural disasters. Development metrics with noticeable post-disaster trends, if any, will receive further statistical analysis to draw insights not produced during time-series analysis.

DATA SET
1. [World Development Index](#)
2. [The International Disaster Dataset](#)

Both databases listed above produce CSV files. The World Development Index has 1488 attributes ("series"), including economic metrics such as GDP growth and GINI coefficient, and social development metrics such as HIV infection rates and access to clean cooking fuels. Changes in these attributes are described year over year ("time"), with data spanning between 1960 and 2023. Accordingly, we extract entries between 1960 and 2023 for The International Disaster Dataset. Of note, the Center for Research on the Epidemiology of Disasters indicates that entries before 1960 in their database are subject to reporting biases.

In total–without excluding null values in the World Development Index, which are common in older data–the complete dataset contains 26,480,562 data points spread between 1,533 attributes.

The data is quite granular as it concerns human development metrics and we hope to make novel

findings as we analyze these attributes following natural disasters.

EVALUATION METHODS

To ensure the rigor and reliability of this data mining project, four key evaluation methods will be employed: accuracy of insights, comparative analysis, reproducibility, and clarity of visualizations.

The accuracy of insights will be assessed by comparing our findings to existing literature on the economic and social impacts of natural disasters. This evaluation will determine whether the patterns we uncover—such as GDP trends, shifts in social metrics, or changes in governance—align with previously documented evidence. Discrepancies between our results and the literature will be addressed through comparative analysis to identify potential causes, such as differences in data sources, sample periods, or methodological approaches.

Reproducibility will be a cornerstone of the project, as it ensures that the findings can be validated by other projects. This involves transparent documentation of data sources, transformations, and methodologies. We will provide clear descriptions of the data mining and statistical techniques we applied to ensure that others can replicate our analysis in the dataset.

Finally, the clarity of visualizations will be evaluated to guarantee that our results are digestible by a diverse audience. We aim to create visual tools–like graphs and dashboards– that effectively communicate our findings.

TOOLS

Programming Language: Python
Data Analytics: Pandas, NumPy, SciPy, Matplotlib, Seaborn
Collaboration: Github, Kanban, Jupyter

MILESTONES

1. **October 11:** Collect complete dataset
2. **October 18:** Data preprocessing and storage in a normalized and standardized DataFrame
3. **October 25:** Exploratory Data Analysis with a focus on corroborating completeness and accuracy of dataset, including time-series analysis of economic metrics
4. **November 1:** Time-series analysis of human development attributes, reduce dimensionality of dataset to focus on extracting further insights from metrics
5. **November 15:** Complete further analysis of human development metrics, corroborate resources against any existing literature
6. **November 22:** Employ Evaluation Methods outlined on this paper, produce final report

REFERENCES

[1] Jose Caraballo-Cueto. The Economy of Disasters? Puerto Rico Before and After Hurricane Maria. Centro Journal. 33, 1. (Spring 2021), 66+.
link.gale.com/apps/doc/A668270706/IFME?u=anon~d81b8026&sid=googleScholar&xid=e24c277b

[2] W. J. Wouter Botzen, Olivier Deschenes, and Mark Sanders. 2019. The Economic Impacts of Natural Disasters: A review of Models and Empirical studies. *Review of Environmental Economics and Policy* 13, 2 (May 2019), 167–188. https://doi.org/10.1093/reep/rez004

[3] Evi Susanti Tasri, Kasman Karimi, and Irwan Muslim. 2021. The effect of economic variables on natural disasters and the impact of disasters on economic variables. *Heliyon* 8, 1 (December 2021), e08678.
https://doi.org/10.1016/j.heliyon.2021.e08678

[4] Jones, R.L., Guha-Sapir, D. & Tubeuf, S. Human and economic impacts of natural disasters: can we trust the global data?. *Sci Data* 9, 572 (2022).
https://doi.org/10.1038/s41597-022-01667-x