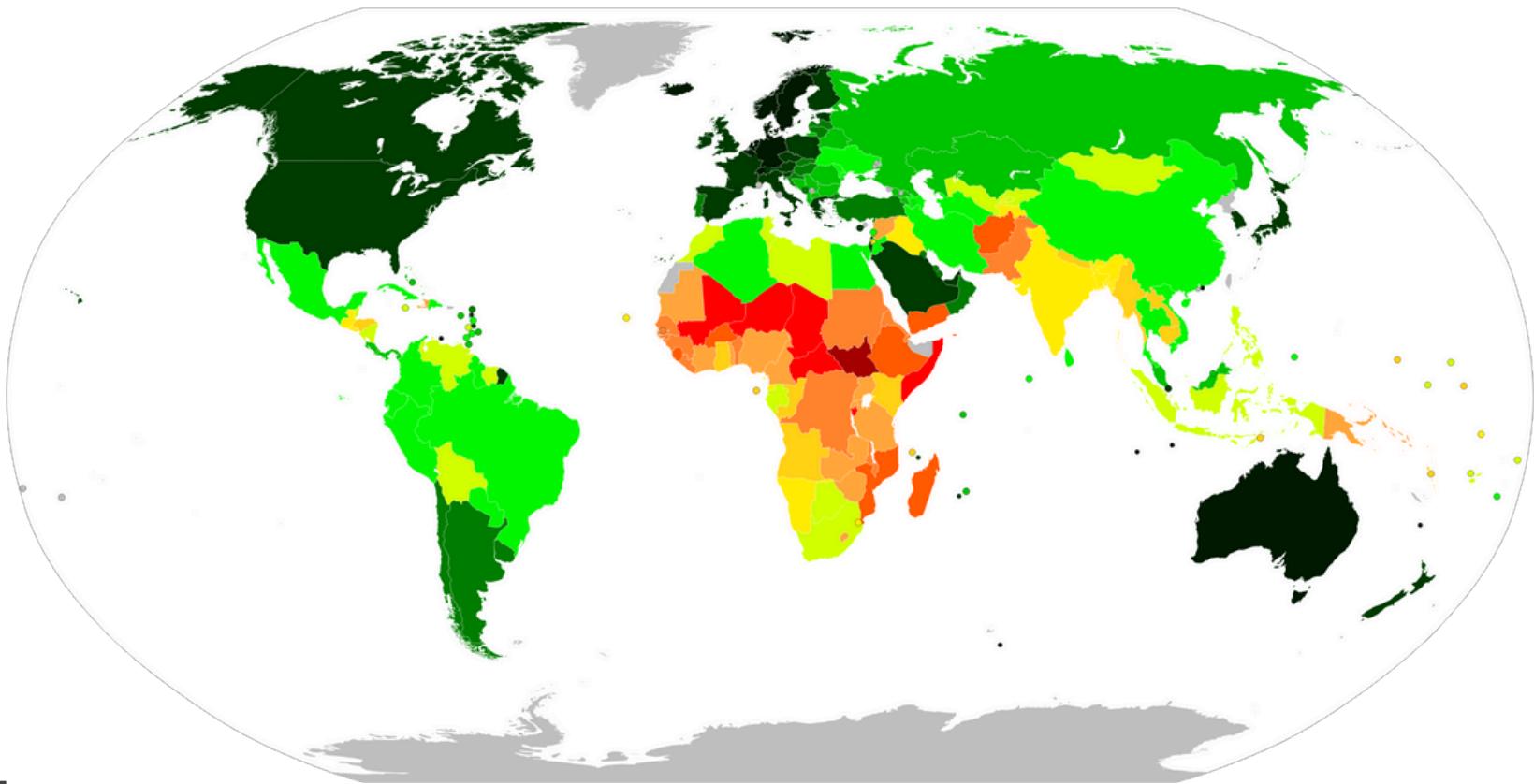


Possible Effects of Terrorist Actions on HDI(Human Development Index)



DSA210 Term Project Final
Presentation

Presented by :
Mehmet Rafet Yıldırım-34226

Table Of Contents

- 01 INTRODUCTION
- 02 DATA SOURCES
- 03 DATASET OVERVIEW
- 04 DATA PREPARATION
- 05 EXPLORATORY DATA ANALYSIS
- 06 HYPOTHESIS TESTING
- 07 MACHINE LEARNING IMPLEMENTATIONS
- 08 FINDINGS AND INSIGHTS
- 09 CONCLUSION

Introduction

- Terrorism remains one of the most persistent global security challenges, with significant social, economic, and political consequences. Knowing that there are plenty of drivers for terrorism, such as political instability, economic factors, and history, we will be focusing on human development. One commonly used indicator of development is the Human Development Index (HDI), which captures health, education, and income dimensions of human well-being.
-
- This study investigates whether there exists a systematic relationship between human development and terrorist activity across countries over time. Specifically, it examines whether countries with lower HDI levels tend to experience more frequent terrorist attacks and greater attack severity

Data Sources

This project uses publicly available datasets from Kaggle and START(National Consortium for the Study of Terrorism and Responses to Terrorism)'s Global Terrorism Database.

Links for the datasets used:

- <https://www.kaggle.com/datasets/iamsouravbanerjee/human-development-index-dataset>
- <https://www.start.umd.edu/download-global-terrorism-database>

Dataset Overview

Used Variables in Terrorism Dataset

- Year
- Country
- Region
- Number of deaths
- Number of wounded

Used Variables in HDI Dataset

- HDI values for each year
- Country

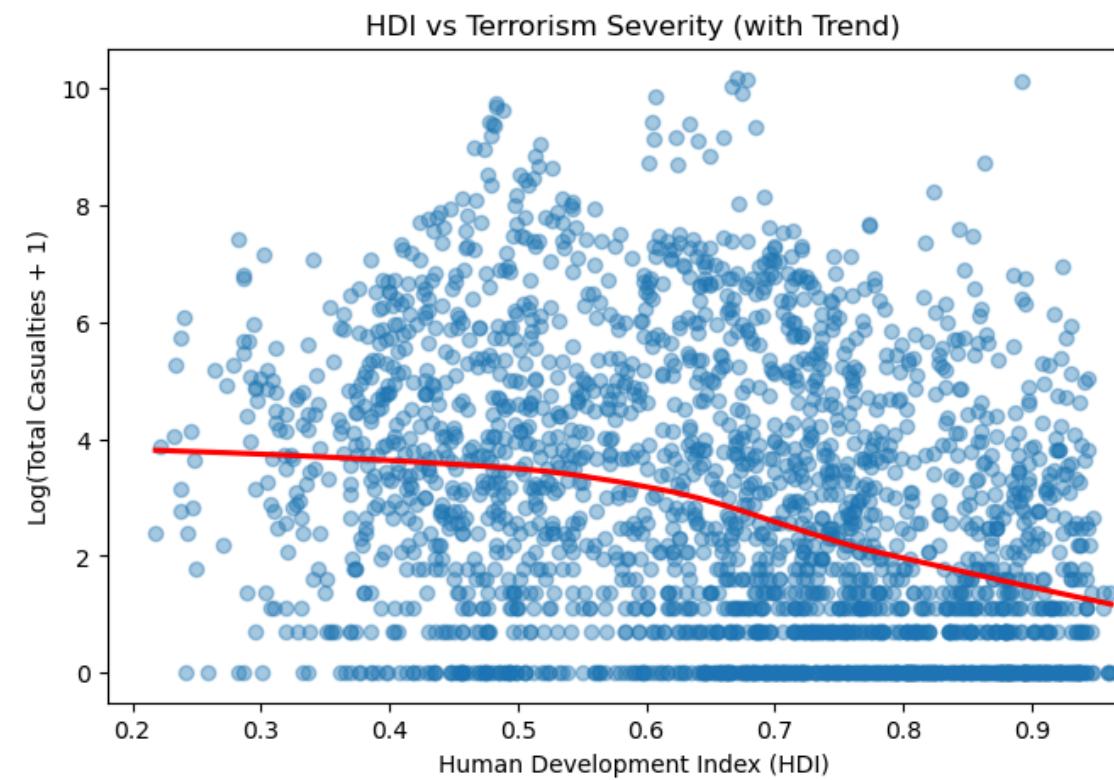
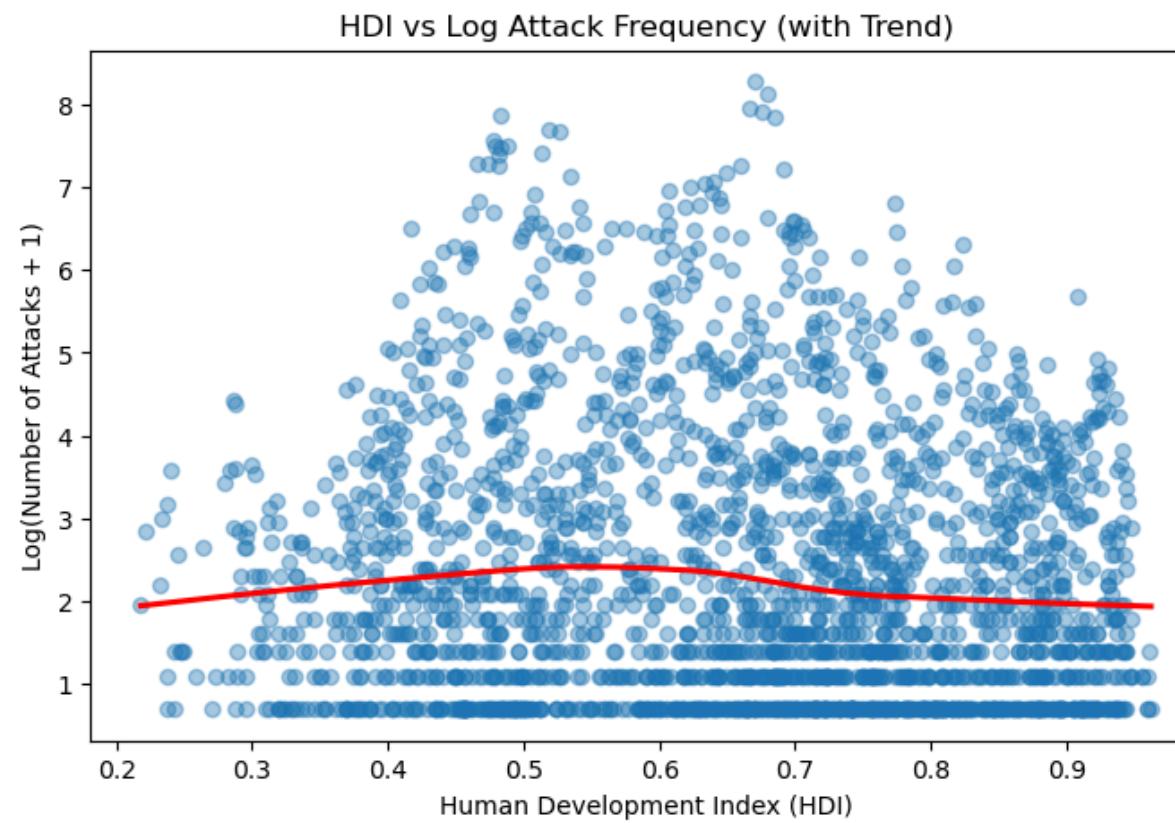
Data Preparation

Data preparation included:

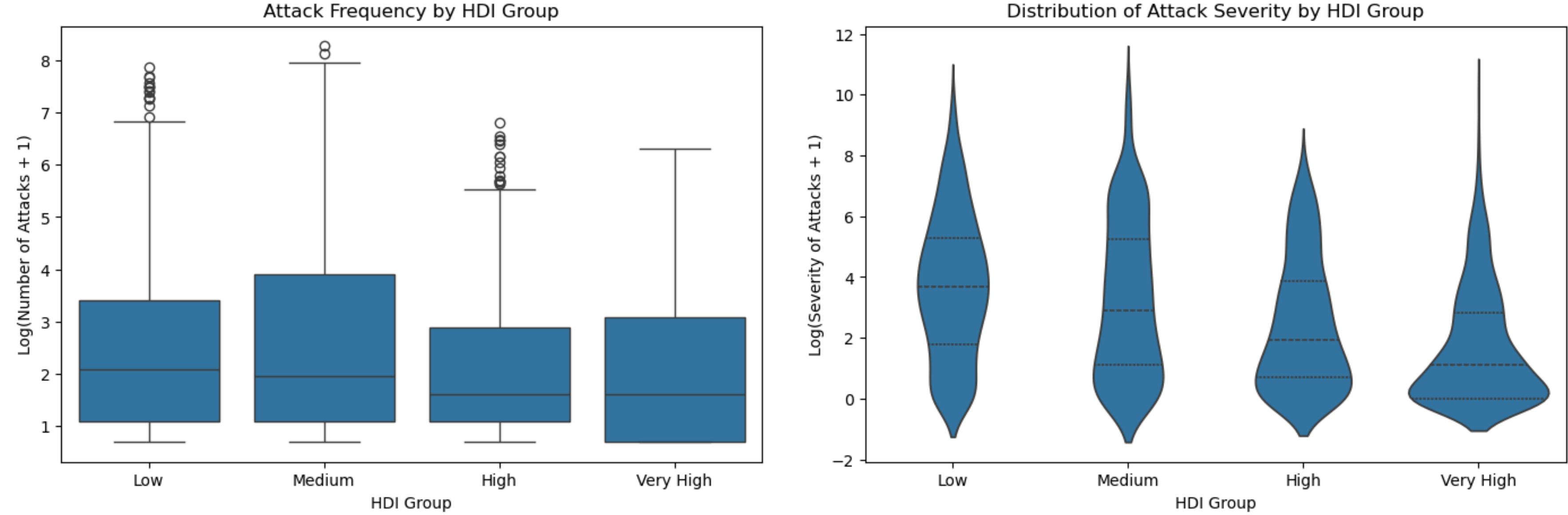
- Removal of identifier columns(e.g., resolution, coordinates, targets)
- Handling missing values via removal
- Log transformations since the terrorism data is highly skewed, logarithmic transformations were applied to the attack frequency and severity variables.
- Additional features such as HDI categories and per-attack severity measures were constructed.

Exploratory Data Analysis

Exploratory Data Analysis (EDA) reveals several notable patterns:



Scatter plots of HDI against log-transformed terrorism measures suggest a generally negative relationship, particularly at lower HDI levels.



Countries with low HDI values display higher variability in both attack frequency and severity.



Regional analysis shows that geography plays a crucial role, with certain regions consistently experiencing higher levels of terrorist activity regardless of HDI.

Hypothesis Testing

To formally assess the relationship between terrorism and development, the following hypotheses were tested:

Hypothesis 1 (Frequency)

- Null Hypothesis (H_{01}): There is no statistically significant relationship between the number of terrorist attacks and HDI.
- Alternative Hypothesis (H_{11}): There is a statistically significant negative relationship between the number of terrorist attacks and HDI.

Hypothesis 2 (Severity)

- Null Hypothesis (H_{02}): There is no statistically significant relationship between the severity of terrorist attacks and HDI.
- Alternative Hypothesis (H_{12}): There is a statistically significant negative relationship between attack severity and HDI.

Pearson and Spearman correlation tests were conducted. The results indicate a negative association between terrorism measures and HDI; however, in some cases, this relationship does not reach conventional levels of statistical significance. When p-values exceed 0.05, the null hypothesis cannot be rejected, implying insufficient evidence to confirm a statistically significant relationship.

ML Implementations

To complement statistical inference, several machine learning models were implemented:

- Linear Regression as a baseline model
- Decision Tree Regressor to capture non-linear patterns
- Random Forest Regressor to model complex interactions

Models were trained using a time-based split and evaluated using MAE, RMSE, and R² metrics. Across all target variables, Random Forest models consistently outperformed linear models, indicating that the relationship between HDI and terrorism outcomes is non-linear and context-dependent.

Findings and Insights

Key findings of the study include:

- HDI is more strongly associated with terrorism frequency than with attack severity.
- Geographic and country-specific factors explain a large portion of the variance in terrorism outcomes.
- HDI remains an important predictive feature even after controlling for region, country, and time.
- Machine learning models reveal patterns that traditional linear methods fail to capture.

These findings suggest that development plays a role in shaping terrorism risk, but it is not the sole determinant.

Limitations and Future Work

This study has some limitations. The analysis uses country-year level data, which may hide differences within countries where terrorist attacks are often localized. HDI is a broad development indicator and does not include political or security-related factors that may influence terrorism. In addition, the analysis focuses on correlation and prediction rather than causality, so the results do not prove that HDI directly affects terrorism outcomes. Terrorism may also negatively impact development, creating possible reverse causality. Future studies could use more detailed data, include political variables, and apply causal or time-series methods to better understand the relationship between development and terrorism. I personally want to be able to forecast which country is going on the way to have more frequent and violent terrorist attacks.