EDA of the Global Terrorism Database: Hotspots, Groups & Attack Patterns

Introduction

The **Global Terrorism EDA Project** analyzes the **Global Terrorism Database (GTD)**, containing over **181,000 records** of terrorist attacks worldwide (1970–2017). The project aimed to uncover patterns, trends, and key insights to better understand global terrorism and its impact.

By acting as a **security/defense analyst**, the project focused on identifying **terrorism hot zones**, **lethal attack types**, and the **most active perpetrators**.

Tech Stack

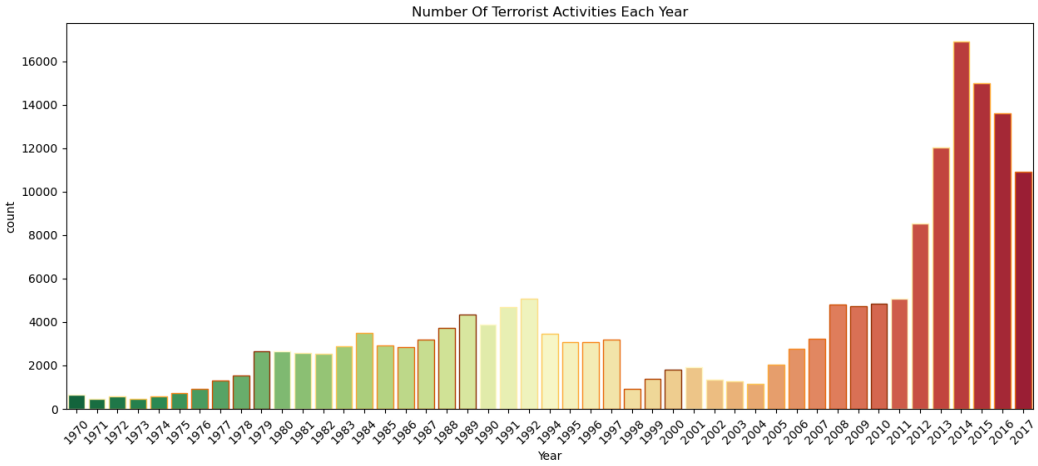
- Python (Jupyter Notebook)
- Pandas & NumPy (data cleaning & manipulation)
- Matplotlib & Seaborn (statistical visualizations)
- Folium (interactive maps)
- Kaggle Dataset (Global Terrorism Database)

Project Workflow

1. **Data Cleaning & Preprocessing**
    - Renamed obscure columns for readability
    - Filled missing values in Killed/Wounded columns with 0
    - Selected relevant features (Year, Country, Region, AttackType, etc.)
    - Created new **casualties** feature (Killed + Wounded)
2. **Exploratory Data Analysis**
    - **Temporal Analysis:** Attack frequency by year
    - **Geographical Analysis:** Hotspot countries and cities + heatmaps
    - **Attack Type Analysis:** Fatalities by attack type
    - **Casualty Analysis:** Human impact across countries
    - **Group Analysis:** Most active terrorist groups and their operational areas

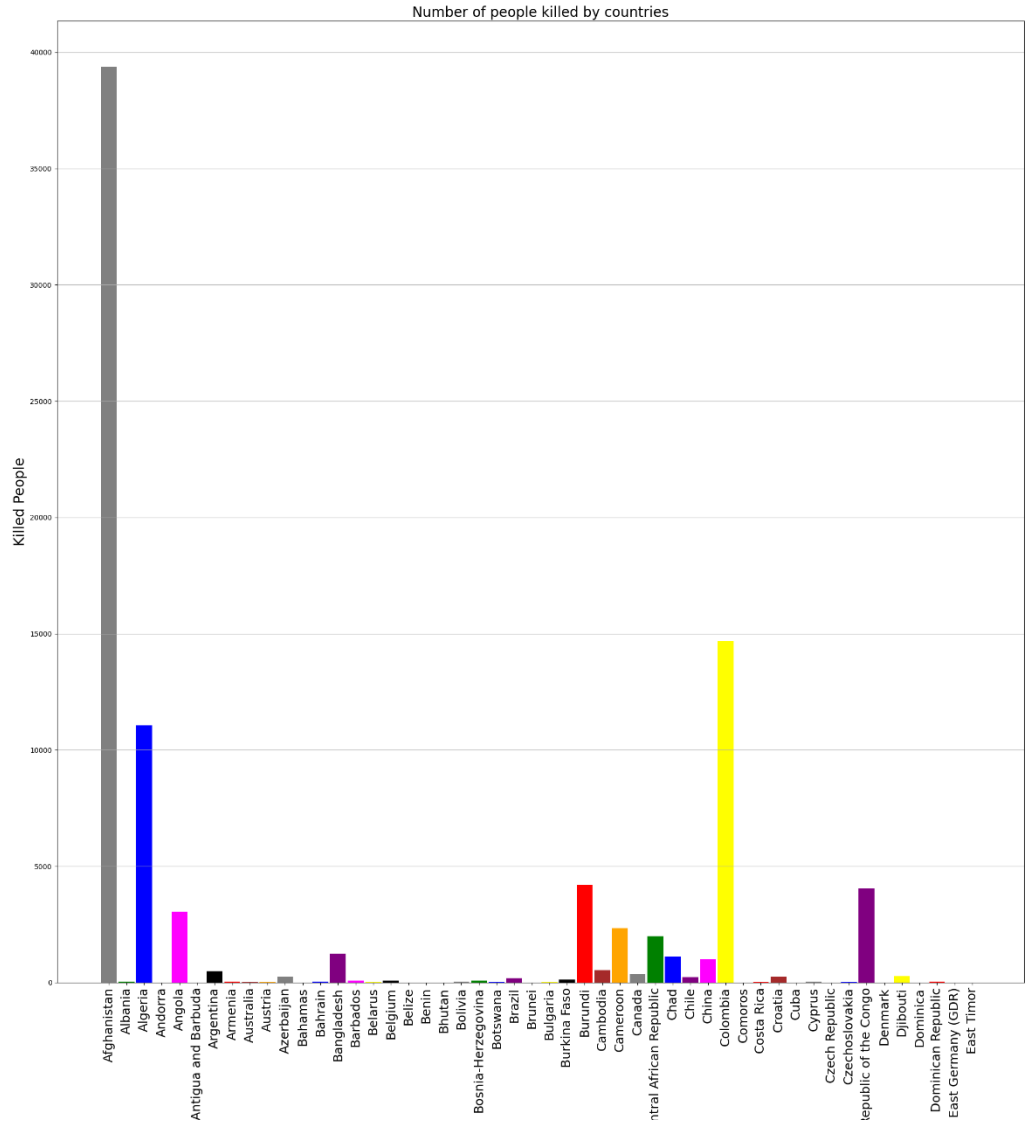
Key Analyses & Visuals

1. Yearly Attack Trends



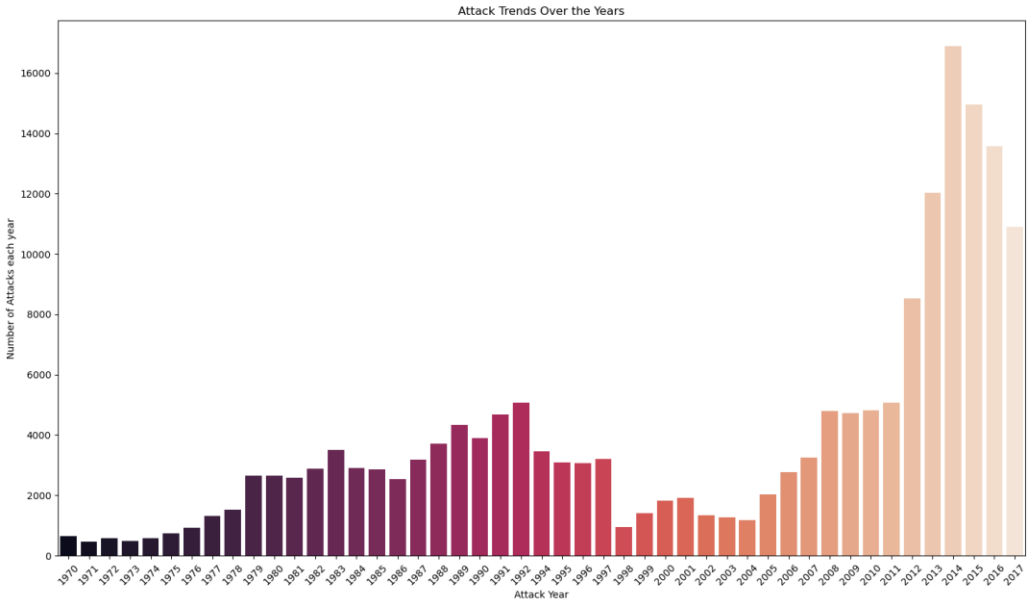
- **Finding:** Terrorist incidents increased dramatically over the decades, with a sharp rise after 2010.

2. Casualties by Country



- **Finding:** Iraq, Afghanistan, and Pakistan recorded the highest number of terrorism-related deaths, highlighting that terrorism disproportionately impacts these regions.

3. Attack Trends Over Time



- **Finding:** Bombing/Explosion and Armed Assault consistently emerge as the most frequent and lethal attack types worldwide.

Key Insights

- **Hot Zones:** MENA region (Iraq, Afghanistan, Pakistan) leads in attacks and casualties
- **Escalation:** Terrorism escalated dramatically in the 2010s
- **Attack Methods:** Explosives and armed assaults dominate globally
- **Groups:** Taliban and ISIL are among the deadliest organizations
- **Victims:** Muslim-majority countries suffer the highest casualties

Analytical Impact

- Showed how **data-driven methods can reveal global security patterns**
- Demonstrated **geopolitical insights** that challenge misconceptions (terrorism mainly impacts Western nations)
- Built reproducible workflows for **large-scale EDA** using Python

Learning Outcomes

- Mastered **large dataset handling (181K+ records, 135 features)**
- Enhanced **EDA, visualization, and feature engineering skills**
- Strengthened **analytical storytelling** with real-world impact

Future Enhancements

- Add **recent years' data (post-2017)** for updated insights
- Apply **predictive models** for terrorism risk forecasting
- Explore **NLP on textual fields** (attack motives, group statements)
- Deploy an **interactive dashboard (Streamlit/Dash)**