Table of Contents

DATASET

1.1 Dataset Description

1.1.1 Global Terrorism Database (GTD)

We make use of the Global Terrorism Database (GTD), which is kept up to date by START (University of Maryland) which records individual terrorist incidents worldwide. The dataset includes 135 variables and 181,691 attacks from 1970 to 2017 (1993 is missing). With fine spatial resolution down to the city/province level and more than 36,000 individual cities (including "Unknown"), it covers 205 nations in 12 regions worldwide. Timing (year, month, day), location, attack features (attack type, weapon, suicide, success), targets and victims, perpetrators and motivations, repercussions (dead, wounded, property damage, hostages, ransom), and international dimension (logistics and ideology) are all considered variables.

1.1.2 World Development Indicators (WDI-Based Dataset)

The World Bank's World Development Indicators (WDI) is a derived dataset related to the World Bank series. The WDI is structured as a country–year panel that covers 1960–2023 with annual data for about 274 countries and territories, comprising 17,272 observations and 50 variables. The indicators span governance and institutions (e.g. control of corruption, government effectiveness, political stability, rule of law, regulatory quality, voice and accountability), macroeconomic and fiscal conditions (GDP, inflation, debt, tax revenue, military and public expenditure, trade in services), development and human capital (human capital index, Doing Business–type measures, education and health spending, poverty and inequality), demography and health (population, life expectancy, birth/death rates, homicides), and environment and land use (agricultural and forest land, precipitation, CO_2 and other emissions, population density).

1.1.3 Corruption Perceptions Index (CPI) 2023

As an external governance metric, we employ Transparency International's Corruption Perceptions Index (CPI) 2023. With variables for Country, ISO3, Region, CPI_2023 (0–100, greater means less corruption), and CPI_Rank, the dataset is a country-level cross-section for 2023 that covers about 180 nations. It is used with the terrorism-development panel to offer a similar, current measure of perceived corruption, enhancing the empirical analysis's governance component.

1.2 Data Preprocessing and Cleaning

Global Terrorism Database (GTD)

A subset of GTD variables related to time, location, casualties and the governance link was retained and renamed using more descriptive labels (e.g. nkill → Killed). A valid date variable was constructed from iyear, imonth and iday, replacing unknown months and days (coded as 0) by 1, excluding the year 1993, and retaining only confirmed terrorist incidents

(doubtterr = 0). Text fields were standardised and exact duplicate rows removed. Casualty fields were converted to numeric, so that non-numeric entries (such as “Unknown”) became missing and were subsequently treated as zero in the computation of total casualties. A derived Casualties variable was defined as Killed + Wounded, and incomplete location fields (e.g. province, city) were preserved but explicitly labelled as “(Unknown)”.

World Development Indicators (WDI)

Keys were harmonised by standardising the Country variable and deriving a numeric Year from the date field. From the full WDI dataset, only a targeted subset of governance and structural indicators, together with population for per-capita calculations, was retained. For the indicators used in the analysis, within-country linear interpolation with forward and backward fill was applied to obtain complete annual series for 1997–2017, after which remaining missing values in the merged analysis dataset were removed.

CPI 2023

The CPI excel was reorganized by using the "Country / Territory" row as the header and keeping only Country, ISO3, Region, CPI_2023, and CPI_Rank. The remaining missing values were imputed (median for numbers, “(Unknown)" for text), ISO3 codes were standardized, and CPI variables were converted to numeric.

Construction of the analysis panel

Finally, we aggregated GTD incidents by Country–Year to obtain Attacks, Total_Casualties and Has_Casualties, and merged this with WDI population to compute Attacks_per_million and Casualties_per_million. Restricting to 1997–2017, we merged terrorism, WDI and CPI data into a single, complete country–year dataset used for the empirical analysis.

1.3 Exploratory Data Analysis (EDA)

1.3.1 Key Descriptive Statistics

After cleaning (excluding 1993 and doubtful cases), the GTD comprises 138,905 confirmed incidents for the period 1970–2017. Annual attack counts are highly volatile, with a median of roughly 1,700–1,800 incidents per year and peaks above 13,000 in the mid-2010s. Casualties per incident follow a strongly right-skewed distribution: most attacks cause only 1–2 casualties, while a small number of extreme events account for a disproportionate share of total deaths and injuries. At the country–year level, the majority of observations display zero or very low casualties per million inhabitants, and only a limited group of states shows very high per-capita risk.

1.3.2 Temporal and Spatial Patterns

Examining the time series reveals wave-like dynamics rather than a steady upward trend: levels are low and stable before 2000, rise in the early 2000s, reach a pronounced peak around 2009–2014, and subsequently decline, though remaining above pre-2000 values. In per-capita terms, global terrorism risk is low on average yet exhibits the same rise-and-fall

pattern over the last 25 years, providing an empirical basis for AQ1 on whether the world is “increasingly dangerous”. Spatially, incidents and casualties are highly concentrated in the Middle East & North Africa, South Asia and parts of Sub-Saharan Africa, whereas many high-income countries register almost no casualties per million over long periods.

1.3.3 Governance, Development and Early Findings

Combining the terrorism panel with WDI and CPI 2023 indicates a negative relationship between governance quality and terrorism risk: higher CPI and stronger governance scores are typically associated with near-zero casualties per million, while poorly governed, conflict-affected states show extreme per-capita values. A simple country clustering based on per-capita casualties, attack frequency and selected economic/governance variables separates low-risk stable countries from high-risk conflict states, directly motivating AQ2 and AQ3.