**Comprehensive Analysis of Natural Disasters from EM-DAT**

**1. Executive Summary**

This report analyzes disaster event records from the **EM-DAT (Emergency Events Database)**, focusing primarily on **flood-related disasters in African countries**. The analysis includes descriptive statistics, identification of trends in human and economic impacts, and the geographic distribution of disasters. The purpose is to support policymakers, humanitarian agencies, and disaster management professionals in understanding disaster frequency, severity, and areas of concern.

**2. Data Source and Structure**

**2.1 Source**

The data originates from **EM-DAT**, maintained by the Centre for Research on the Epidemiology of Disasters (CRED). It contains detailed entries on disasters globally, capturing various dimensions such as type, location, casualties, and economic losses.

**2.2 Structure**

* **Sheet Analyzed:** EM-DAT Data
* **Total Records:** 206 disaster entries
* **Coverage Period:** Implied to be from the year 2000 (based on DisNo. values)
* **Disaster Group:** All entries are under the group “Natural”
* **Subgroup:** Hydrological
* **Type:** Flood

**3. Data Cleaning and Integrity**

**3.1 Observations**

* **Missing Data:** Fields related to economic costs (reconstruction, insured damage) have extensive missing values.
* **Event Dates:** Dates are not explicit, but year can be parsed from DisNo.
* **Admin Units:** Locations are stored as JSON-like strings and may require transformation for GIS applications.
* **Currency Adjustment:** Economic loss figures are inflation-adjusted in one column (Total Damage, Adjusted ('000 US$)), enhancing comparability over time.

**3.**2 **Data cleaning process**

* Reading the excel file
* Checking for duplicates
* Checking for null values
* Dropping of columns
* Filling some columns with data
* Checking /changing of data typr
* Exporting cleaned data as csv file

**3.3 Tools used for cleaning and analysis**

* Numpy
* Pandas
* Matplotlib.pyplot
* Plotly.express

**4. Descriptive Statistics**

**4.1 Disaster Type Frequency**

All 206 records are classified as **Floods**, indicating a thematic focus in the dataset.

**4.2 Country-wise Distribution**

| **Country** | **No. of Events** |
| --- | --- |
| Nigeria | 18 |
| Burundi | 11 |
| Tanzania (United Republic of) | 10 |
| Mozambique | 10 |
| Democratic Republic of the Congo | 10 |

These countries appear most vulnerable to flood-related events.

**4.3 Human Impact**

| **Metric** | **Total** |
| --- | --- |
| Total Deaths | 3,286 |
| Total Affected | 17,817,661 people |

This includes direct deaths and individuals affected through displacement, injury, or livelihood loss.

**4.4 Economic Impact**

| **Metric** | **Value** |
| --- | --- |
| Total Damage (Adjusted) | $1.81 billion USD |

Due to missing values, this figure may underestimate the true economic loss.

**5. Exploratory Insights**

**5.1 Temporal Pattern (Suggested)**

The dataset spans multiple years, with event years embedded in the DisNo. column (e.g., 2000-0260-AGO). A parsed analysis would allow:

* Year-over-year trends in disaster frequency
* Correlation of death tolls with climate phenomena (e.g., El Niño years)

**5.2 Geographic Spread**

Admin Units provide administrative divisions down to district levels (adm2). With geocoding, this can be mapped to highlight:

* Flood hotspots
* Repeatedly affected regions
* Overlap with population density or poverty zones

**6. Limitations**

* **Incomplete Economic Data:** Most events lack reconstruction or insured loss estimates.
* **Temporal Gaps:** Explicit date/time fields are missing.
* **Single Disaster Type:** No other natural hazards (e.g., droughts, storms) are included for comparative analysis.

**7. Recommendations**

1. **Enrich with Time Series:** Extract and analyze year-wise trends for disaster frequency and impact.
2. **Geo-Mapping:** Convert Admin Units to latitude/longitude and create interactive maps.
3. **Impact Per Capita:** Normalize damage and death tolls by population for more actionable insights.
4. **Integrate Climate Data:** Overlay with rainfall patterns or river basin data to explore environmental drivers.
5. **Enhance Data Completeness:** Cross-validate with national disaster databases or insurance records.

**8. Conclusion**

This dataset provides a valuable snapshot of flood-related disasters in Africa, especially in high-risk countries like Nigeria and Mozambique. Despite data gaps, significant human and economic losses are evident. With further enhancement and integration, this data can strongly support targeted disaster risk reduction (DRR) and climate resilience planning.