DATA INFO

Description

IS THERE AN INTENDED PURPOSE FOR THE DATASET? WHAT DOMAIN WAS IT DESIGNED FOR?

The dataset is designed to investigate potential correlations between numerous poverty indicators (all categorized into levels of low, medium, and high poverty) and casualties in the West Bank and Gaza region. Specifically, it aims to explore whether a direct correlation exists. The analysis will delve into the dataset to better understand the potential impact of socio-economic conditions through the levels of poverty and casualties in the West Bank and Gaza region.

WAS ANY PREPROCESSING/CLEANING/LABELING OF THE DATA DONE? IF SO, PLEASE PROVIDE A DESCRIPTION.

When joining two datasets, there was a need for removing unnecessary columns/variables that didn't contribute to the overall story and observations, such as **Palestinians and Israelis injuries** from the casualties dataset. Also, from the poverty dataset, we removed the **indicator codes** variable and also the **country codes** variable.

Composition

DOES THE DATASET HAVE A METADATA REPOSITORY OR DATA DICTIONARY? IF YES, PLEASE PROVIDE THE LINK AND IF NOT, PLEASE EXPLAIN WHAT EACH FIELD MEANS.

As far as repositories go. the dataset can be found here (https://github.com/INFO-201-Fall-2023-Final/final-projects-NahomeYo.git) . The fields in the dataset include Year, Country Name, Indicator Name, Value, Month, Palestinians Killed, Israelis Killed, Poverty_Level, Total_Killed.

- **Year:** The year of the data entry.
- **Country Name:** The region or area (West Bank and Gaza in this case).
- **Indicator Name:** The type of indicator (Population living in slums (% of urban population)).
- **Value:** The numerical value associated with the indicator.
- **Month**: The month of the data entry.
- **Palestinians Killed:** Number of Palestinians killed.
- **Israelis Killed:** Number of Israelis killed.
- **Poverty_Level:** The poverty level associated with the data entry.
- **Total_Killed:** The total number of people killed.

Levels of Poverty

Low Poverty

Indicator names:

- Survey mean consumption or income per capita, total population (2017 PPP $ per day)
- Annualized average growth rate in per capita real survey mean consumption or income, total population (%)

Medium Poverty

Indicator names:

- Population living in slums (% of urban population)
- Income share held by second 20%
- Income share held by third 20%
- Income share held by fourth 20%
- Income share held by highest 20%
- Income share held by highest 10%
- Proportion of people living below 50 percent of median income (%)
- Gini index
- Multidimensional poverty headcount ratio (% of total population)
- Multidimensional poverty headcount ratio, household (% of total households)
- Multidimensional poverty intensity (average share of deprivations experienced by the poor)
- Multidimensional poverty index (scale 0-1)
- Poverty headcount ratio at national poverty lines (% of population)

High Poverty

Indicator names:

- Poverty headcount ratio at $2.15 a day (2017 PPP) (% of population)
- Poverty gap at $2.15 a day (2017 PPP) (%)
- Poverty headcount ratio at $3.65 a day (2017 PPP) (% of population)
- Poverty gap at $3.65 a day (2017 PPP) (%)
- Poverty headcount ratio at $6.85 a day (2017 PPP) (% of population)
- Poverty gap at $6.85 a day (2017 PPP) (%)
- Poverty headcount ratio at $6.85 a day (2017 PPP) (% of population)
- Poverty gap at $6.85 a day (2017 PPP) (%)
- Income share held by lowest 10%
- Income share held by lowest 20%
- Poverty headcount ratio at $6.85 a day (2017 PPP) (% of population)
- Poverty gap at $6.85 a day (2017 PPP) (%)
- Income share held by lowest 10%
- Income share held by lowest 20%
- Poverty headcount ratio at $2.15 a day (2017 PPP) (% of population)
- Poverty gap at $2.15 a day (2017 PPP) (%)

OBJECTIVES & ALERTS

No Yes Maybe

Objectives: Correlation Analysis:

- Investigate potential correlations between various poverty indicators categorized into levels of low, medium, and high poverty and casualties in the West Bank and Gaza region.

ALERTS

- A sudden increase or decrease in casualties compared to previous years, particularly focusing on the periods 2012 to 2014 and 2018 to 2020, which may indicate a critical situation requiring attention.

Objectives: temporal trends

- Identify temporal trends in poverty levels and casualties to understand the evolution of these poverty indicators over the specified time period.

ALERTS

- Deviations in the correlation between poverty levels and casualties counts from the historical trend. Notably, if the year with the highest casualties, like 2014, shows a medium poverty level instead of high, it signals potential irregularities and requires further investigation.

Objectives: Comparative Analysis

- Compare the relationship between poverty levels and casualties for different categories of poverty (low, medium, and high) to determine if specific poverty indicators exhibit a stronger correlation with casualties.

ALERTS

- The geographical setting remains constant throughout the entire data creation year range, emphasizing the relevance and stability of our focused factors.