# Group 10 PDA Project Preliminary Questions

## Introduction

We chose the "Intentional Homicide Victims" crime dataset compiled by the United Nations Office on Drugs and Crime (UNODC) and disseminated via the UNODC Data Portal.

## Scope & structure

Observations are organized by country and year with attributes such as Iso3_code, Region, Subregion, Country, Source, Dimension (e.g., by citizenship), Category (e.g., National citizens, Foreign citizens, Total), Sex (Male/Female/Total), Age (often Total), Year, Unit of measurement, and value. The time span covers roughly 1990–2022, enabling long term trend analysis.

## Source

UNODC's series integrates criminal justice data (law enforcement and court records) and public health data (vital registration/causes of death).

## Questions for EDA

1. Which countries have the highest and lowest homicide rates (per 100,000 population) in the most recent year available?
2. How has the global homicide rate (or total counts) changed over time for the 'World' region?
3. How do homicide rates compare across the five major regions (Americas, Europe, Asia, Africa, Oceania)?
4. How concentrated are global counts—what share of total homicides occur in the top 10 countries each year?
5. What is the distribution of homicide by mechanisms (e.g., 'Firearms or explosives', 'Another weapon') globally or by a specific region?
6. Is there a significant difference in crime rates (either total or homicide) between male and female victims?
7. Which age group has the highest number of homicide victims in the dataset?
8. For a given country, how does the proportion of homicides committed by intimate partners or family members compare to those committed by unknown perpetrators?
9. How do the crime statistics differ between 'National citizens' and 'Foreign citizens' across a selected set of European countries?
10. What is the distribution of VALUE (counts) across all observations, and where are the outliers?