DEATHS DUE TO AIR POLLUTION

By Laura Zanetti
Email: laura.zanetti@studenti.univr.it
University of Verona - Master Degree in Data Science - Data Visualization

ABSTRACT

Air Pollution continues to be the reason for the killing of Innocent lives in the world. The data proposed in this paper once analysed will allow the user to get some insight into the situation.

INTRODUCTION

Air Pollution has always been at the center of many debates and issues during the years as it continues to be the reason for the killing of innocent lives in the world. What would be really interesting to analyse trough data and statistics, is how air pollution has change over the last decades by analyzing deaths per 100k (caused by pollution). Moreover let's see the countries that suffered from it the most and the differences in how outdoor, indoor and ozone pollution is affecting the population.

SCENARIO OF USE

The user can visualize the trend of Air pollution (deaths per 100k) during the last decades and compare the effects of the different types of pollution. An interactive interface allows the user to filter by year and country/ region. Visualization of maps provides an overview of where the highest numbers of death by pollution are concentrated.

DATA AND TASK ABSTRACTION

Dataset Type: Table
Data Types: 2 categorical attributes [**Entity, Code**], 5 quantitative attributes [**Year**, **Air pollution (total) (deaths per 100,000)**, **Indoor air pollution (deaths per 100,000)**, **Outdoor particulate matter (deaths per 100,000)**, **Outdoor ozone pollution (deaths per 100,000)**]

The Dataset (available at: https://www.kaggle.com/datasets/akshat0giri/death-due-to-air-pollution-19902017) has been divided in two major parts since it contained:

- items about each country by year and
- items about regions of the world (containing multiple countries) by year

Data Description

- **Entity**: It contains the name of the country or the region.
- **Code**: It contain Code of the country.
- **Year**: Years range from 1990 to 2017
- **Air pollution (total) (deaths per 100,000)** : Contains total deaths.
- **Indoor air pollution (deaths per 100,000)** : Contains deaths due to indoor air pollution.
- **Outdoor particulate matter (deaths per 100,000)** : Contains deaths due to outdoor pollution.
- **Outdoor ozone pollution (deaths per 100,000)** : Deaths due to ozone pollution.

Task abstraction:

- derive two group of samples (one with regions and the other with countries per year)
- Discover trend of deaths due to tot air pollution
- Compare trends (of different types of pollution)
- Discover distribution of deaths per country due to air pollution

answering to questions such as:

- What is the trend of deaths caused by air pollution? How it changed over the years?
- Which air pollution is more dangerous?
- In which countries there is a higher number of deaths (per 100k) caused by air pollution?

SOLUTION PROPOSED AND IMPLEMENTATION PLAN

Produce a visualization application and vertical 'storytelling' using **Streamlit**.

The application will provide its user with visual solutions and an interactive interface for the user to enjoy.