Air Pollution in Poland – Group 2

**Project Plan for GRS35306** - Dong Liang (1399853), Hudson Passos (1311182), Nikshep Trinetra Bangalore Suresh (1302590), Pamungkas Intan (1280937), Qin Xu (1274376), Sabrina Ramadwiriani (1340905)

Dong Liang, Hudson Passos, Intan Pamungkas, Qin Xu, Sabrina Ramadwiriani

**Introduction & Context:**

Air pollution in Europe has been a significant issue for several decades, but there have been considerable improvements in recent years. Despite these improvements, air pollution still poses a major health and economic burden in Europe. In 2019, it was estimated that 307,000 premature deaths were caused by exposure to fine particulate matter (Juginović et al., 2021).

Poland has been a major contributor to European dust, sulfur dioxide, and nitrogen oxides emissions (Sawicka-Kapusta & Zakrzewska, 1998). The detrimental health effects of air pollution in Poland include increased mortality rates, higher prevalence of respiratory diseases such as asthma and lung cancer, and a higher risk of COVID-19 infections (Nazar & Niedoszytko, 2022). In six Polish cities (Wroclaw, Lublin, Warsaw, Katowice, Olsztyn, Szczecin) long-term exposures to PM2.5, PM10, NO2, and O3, above the 2021 WHO guideline values, led to a total of about 6700 deaths for non-accidental causes, 2310 deaths for cardiovascular diseases, and 500 deaths for respiratory diseases over the period of 2015–2020 (Cakaj et al., 2023).

For this project, air pollution includes the presence of seven main pollutants identified by the European Union - ammonia (NH3), nitrogen oxides (NOx), carbon monoxide (CO), particulate matter (PM), sulfur dioxide (SO2), ozone, and non-methane volatile organic compounds (NMVOCs).

**Objective and Research Questions:**

* To investigate the causes of air pollution and how these causes differ across the various regions of Poland.
* To investigate the disproportionate health impacts of air pollution in Poland.
* To understand the perception of Polish residents on air pollution in Poland.

**Hypothesis:**

Air pollution is caused by:

* + Energy industry
  + Motor vehicles/traffic
  + Agricultural activities
  + Forest Fire

**Methodology**

* The study area chosen for this project is Poland.
* Data needed and strategy to acquire this data:
  + Strategy: Download the data or use API to call data from portals. For social media data, use web scrapers.
  + Air Quality – AQI, Pollutants with locations
  + Wind Direction
  + Power Plants – with location
  + Land use
  + Forest Fire
  + Traffic Movement
  + Respiratory illness data – medical data, mask sold data
  + Demographic data
  + Social media – Opinions
* Proposed approach and techniques used for analyses and visualization.
  + Geospatial Analysis to find geospatial correlation between the factors and the air quality– GeoDa, GeoPandas, QGIS, ArcGIS,
  + Statistical Analysis to find the correlation between air quality and health impacts – Pandas, Machine Learning
  + Sentiment Analysis to find perception of residents on air pollution – Pattern, BERT, TextBlob
  + Mapping to visualize the results – ArcGIS, QGIS
  + Graphs to visualize the results – Matplotlib
  + Story Map to present the results – ArcGIS Pro

**Societal and Ethical/Privacy Issues**

* What societal issues will your project address?
* Who or what can be affected by your project? (societal groups and/or environments, animals, etc.)
* Can you only identify potential benefits or also potential harms?

As our project focuses on understanding the caused of the dynamics of air pollution in Poland addresses several critical societal issues and has the potential to impact the entities. The primary areas of consideration encompass environmental, public health, and socio-economic dimensions. The research project delves into the relative contribution of each factor of air pollution across different regions. This regional perspective is crucial for tailoring interventions to the specific challenges faced by diverse communities.

In examining the trend of air pollutant concentrations in Poland serves as a step in recognizing the environmental challenge highlighted in Poland. Elevated levels of pollutants can have reached the consequences of ecosystem, wildlife, and the overall environmental quality. Identifying regions with high concentrations is essential for implementing targeted intervention to preserve the ecological balance.

~~Analyzing the respiratory illness occurrence in relation to socioeconomic status brings attention to social disparities in health outcomes. The project could highlight the vulnerable population that may be disproportionately affected by poor air quality, fostering awareness.~~ Understanding the source and contributor to air pollution provides policymakers with insights to develop effective strategies for reducing pollution levels. This, in turn, positively impacts the health and well-being of the Poland citizens.

Furthermore, investigating the trend of perception of air pollution in Poland adds a subjective dimension to the study. Public perception influences public engagement, policy advocacy, and community action. Understanding how people perceive air pollution can contribute to effective communication strategies, and the formulation of policies that resonate with the concerns of the population.

To recognize the potential harms in the substantial potential benefits of this research project, risk of misinterpretation of data or ethical concern related to privacy in health -related studies are occurred. During the project ensuring the responsible and the transparency is paramount to navigate these potential challenges and maximize the positive impact of the study on both societal and environmental well-being.

**Expected Results & Deliverables (in a Story Map)**

* Trend of concentration of air pollutants in Poland
* Correlation between potential factors/causes and air pollution
* Relative contribution of each factor of air pollution, differing by region
* Respiratory illness occurrence, differing by socioeconomic status
* Trend of perception of air pollution in Poland

**Expertise in the team**

* Dong: Data analysis and visualization by Python, R, GEE, ArcGIS, QGIS
* Nikshep: Systems Analysis, GIS Tools, Python
* Qin Xu: GIS background, Python, R, ArcGIS
* Intan: Python background, system analysis, visual design, data interpretation.
* Sabrina: Spreadsheet and pivot, visual design, system analysis
* Hudson: QGIS, ArcGIS, Python, R, Geostatistics.

**Feasibility/Risks**

* Incomplete data – not finding data in the required resolution or in the required geography or data with incomplete elements.
* Timeframe – potential lack of time to complete the analyses.
* Resource availability – potential lack of skills or software to complete the analyses.
* Methodological feasibility – potential improper use of analysis methods data to establish causation/correlation.

**Task Division**

**Project Plan Drafting**

Nik

Sabrina

**Data searching: 2013-2023 (Preferable), 2018-2023 (second option)**

* Social media – Opinions >> Nik

**Data preprocessing** (data cleaning, structured data)

Dong

Sabrina >> scripting in jupyter notebook

Nik

Olive >> wants to learn more about coding

Backup : Intan

**Data analysis**

* Trend of concentration of air pollutants (excel and phyton) >> Sabrina, Dong, Hudson
* Correlation between potential factors/causes and air pollution (GeoDa and python) >> Nik (any kind of factor), Sabrina (Fire forest), Olive, Dong, Hudson
* Relative contribution of each factor of air pollution, differing by region (GeoDa/Python) >> Intan, Olive, Dong, Sabrina, Hudson
* Respiratory illness occurrence, differing by socioeconomic status (Python)>> Dong, Nik, Intan, Olive
* Trend of perception of air pollution in Poland (Python)>> Nik, Dong, Olive

**Visualization (mapping, storymap)**

All of us!

**Presentation**