Data Visualization

ISM6419 Fall 2023

Prof. Johannes Reichgelt

Project Report

The Impact of Air Quality on Respiratory Health

BY

Sri Kumar Dundigalla

U48489930

MS in Business Analytics and Information Systems

Introduction

In my project, "The Impact of Air Quality on Respiratory Health," I delve into the pressing issue of how air pollution is affecting health outcomes around the globe. I've embarked on a journey to understand the correlation between ambient particulate matter pollution levels and all-cause mortality rates in different countries. My analysis spans from densely populated nations like China and India to the clearer skies of Sweden and New Zealand, revealing stark contrasts that call for a deeper investigation into the role of economic development, environmental policies, and public health initiatives in shaping these outcomes.

I'm driven by a series of research questions:

1. How is air pollution linked to the number of deaths on different continents?
2. Why is there such a big difference in air pollution's impact from one place to another?
3. Are some age groups more affected by air pollution than others, and how has this changed over the years?
4. How does the wealth of a region affect its air quality?
5. How common are lung diseases like asthma and COPD in different regions, and what might this show us about why they occur?

My project has a goal to find the stories hidden in the data about air pollution and health. I ask simple but important questions and use clear charts and graphs to help everyone understand this big issue that affects people all over the world.

Methodology

In examining the impact of air pollution on health and longevity, I have carefully curated data from diverse, authoritative sources. My methodical approach to data collection is designed to ensure both reliability and breadth of perspective. Here is a detailed overview of the data sources that I have incorporated into my research:

1. **WHO Air Quality Database**: This rich dataset from the World Health Organization offers an extensive assessment of particulate matter exposure, including PM10 and PM2.5, covering a vast array of countries and providing a foundation for my air quality analysis.

   Source: https://www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database
2. **HealthData.org**: From this resource, I have acquired critical statistics on deaths by risk factor, which is central to understanding the health implications of air pollution.

   Source: https://www.healthdata.org/
3. **Our World in Data (Indoor and Outdoor Air Pollution)**: This platform has supplied comprehensive data on death rates due to indoor and outdoor air pollution, along with age-specific mortality rates. It also facilitates historical comparisons by providing data over different time frames.

   Source: https://ourworldindata.org/indoor-air-pollution

4. **World Health Organization (Life Expectancy)**: This data allows for an analysis of life expectancy trends and their potential correlations with air quality over time.
   Source: Who Life Expectancy

5. **The World Bank (Population Data)**: Population figures are essential for contextualizing health statistics within the scope of global demographic changes.
   Source: https://data.worldbank.org/indicator/SP.POP.TOTL

6. **Centers for Disease Control and Prevention (CDC)**: The CDC's estimates on the prevalence of chronic obstructive pulmonary disease (COPD) inform the segment of my research focusing on respiratory illnesses.
   Source: https://www.cdc.gov/copd/data-and-statistics/county-estimates.html

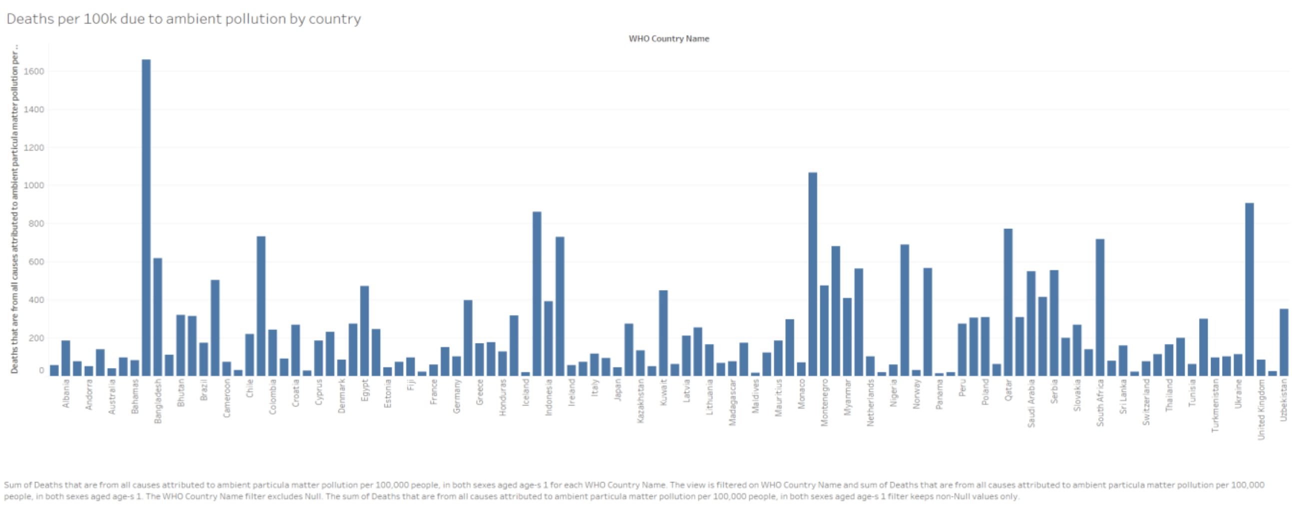
7. **Our World in Data (Asthma Prevalence)**: Data on asthma prevalence is crucial for assessing the respiratory health landscape and its relationship with air pollution levels.
   Source: https://ourworldindata.org/grapher/asthma-prevalence

8. **Science Magazine**: I reference greenhouse gas emissions data in relation to food production to connect environmental factors with health outcomes.
   Source: https://science.sciencemag.org/content/360/6392/987

9. **Our World in Data (Continents and Countries)**: This geographical categorization aids in analyzing data across different regions, ensuring a comprehensive global analysis.
   Source: https://ourworldindata.org/grapher/continents-according-to-our-world-in-data

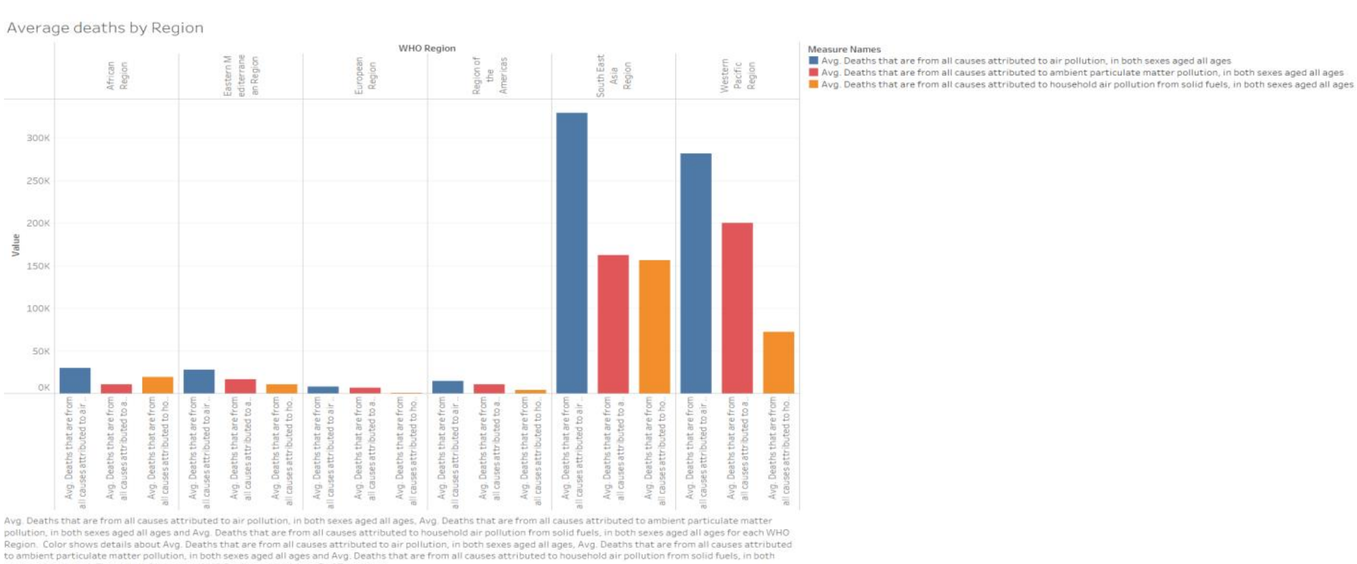
Analysis

Visualization 1: Global Respiratory Infection Death Rates



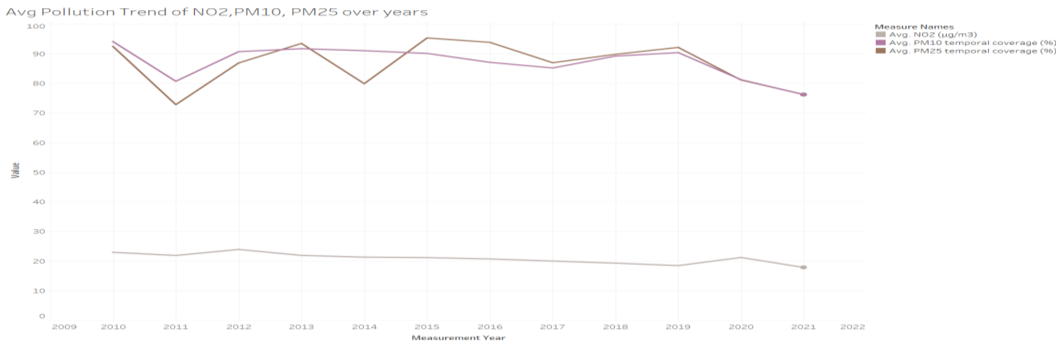
In this graph, the x-axis represents the different countries around the world, each marked with its respective country name. The y-axis quantifies the deaths per 100,000 people attributed to ambient particulate matter pollution for both sexes of all ages. The data shows significant variance among countries, with Bahrain, China, India, and Mongolia exhibiting alarmingly high rates, whereas nations like Sweden, New Zealand, and the Maldives report much lower figures. This disparity invites investigation into the factors contributing to the differences in pollution-related mortality rates, such as economic development, environmental policies, and public health initiatives.

Visualization 2: Temporal Trends in Respiratory Deaths



The graph I'm envisioning would have the regions along the x-axis, capturing diverse geographic areas. The y-axis would represent the average deaths, measured in thousands, for various types of air pollution. Looking at the data, it's immediately clear that the Southeast Asia and Western Pacific Regions experience a heavy toll, with average deaths due to air pollution far exceeding those in regions like Europe. This stark contrast highlights an urgent public health issue that demands further investigation into underlying causes, which could range from economic policies to energy consumption practices.

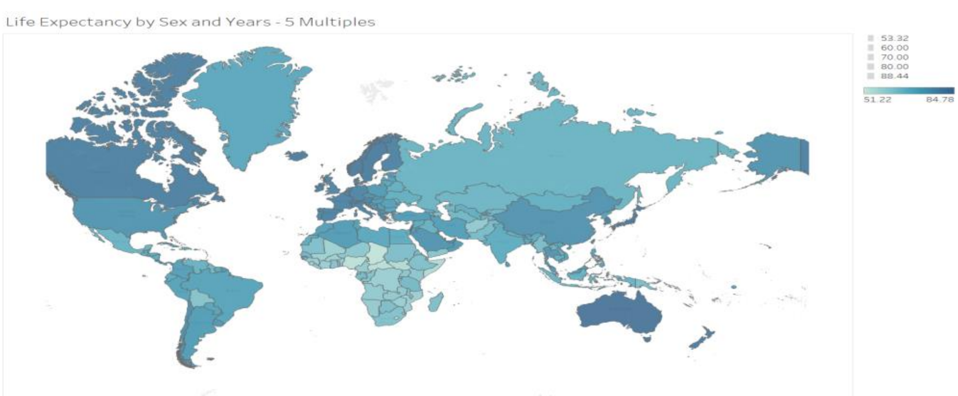
Visualization 3: Comparative Age-Specific Mortality



The plots of Avg. NO2 (µg/m3), Avg. PM10 temporal coverage (%) and Avg. PM25 temporal coverage (%) for Measurement Year. Color shows details about Avg. NO2 (µg/m3), Avg. PM10 temporal coverage (%) and Avg. PM25 temporal coverage (%). The data is filtered on sum of NO2 (µg/m3), which keeps non-Null values only.

In this tableau animation graph, the x-axis showcases the Measurement Year, ranging from 2010 to 2021, which allows us to track changes over time. The y-axis represents the Measure Values, capturing both the average NO2 concentration in µg/m3 and the temporal coverage percentages for PM10 and PM25. This setup reveals the trends in air pollution, specifically nitrogen dioxide and particulate matter, over a period of 12 years. From the data provided, we can discern that while NO2 concentrations have a specific value, PM10 and PM25 are measured by the completeness of their data records over time. The color coding on the graph delineates the three different measures, aiding in visual differentiation. With non-null NO2 values being a focus, the graph likely emphasizes changes in NO2 levels, while PM10 and PM25 temporal coverage indicates the reliability and extent of particulate matter data collection.

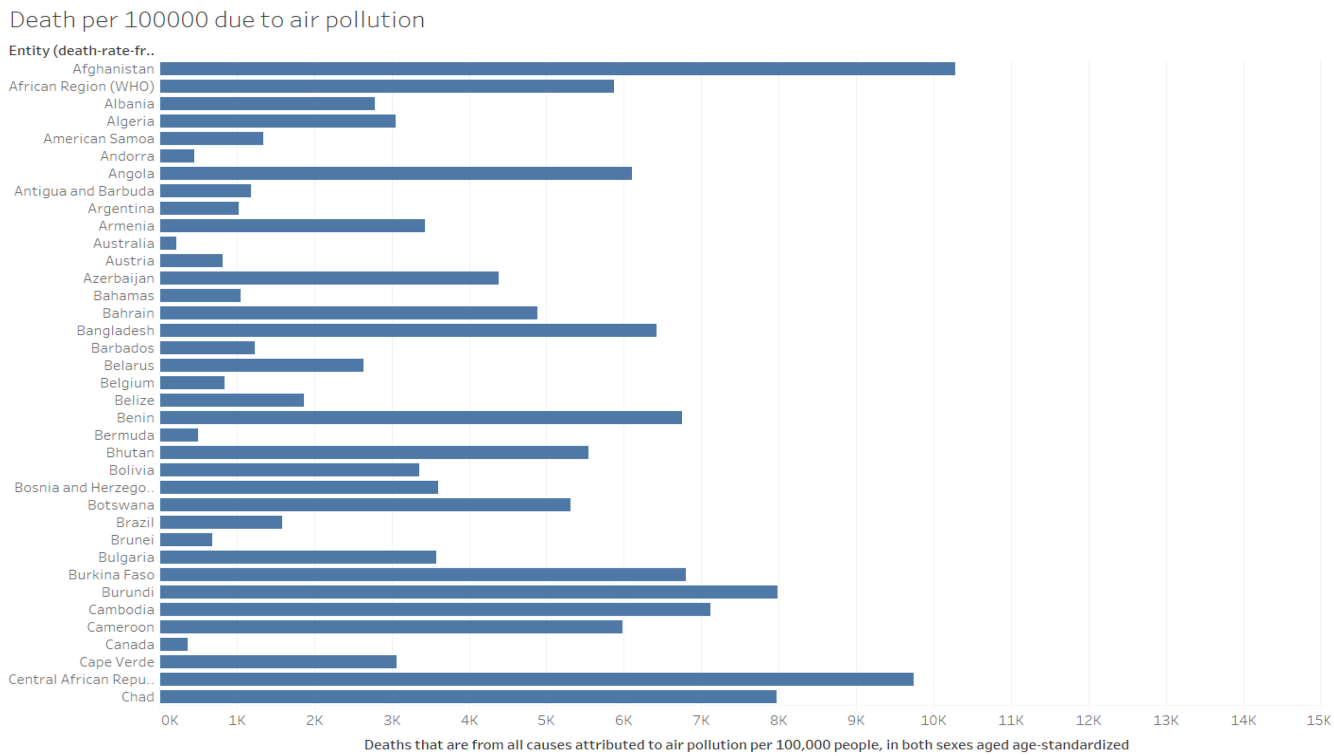
Visualization 4: Gender Disparity in Respiratory Infection Fatalities



Map based on Longitude (generated) and Latitude (generated). Color shows sum of Period life expectancy - Sex: male - Age: 0 (life-expectation-at-birth-1950-onwards-multiples-of-5.csv). Size shows sum of Period life expectancy - Sex: female - Age: 0 (life-expectation-at-birth-1950-onwards-multiples-of-5.csv). Details are shown for Entity (life-expectation-at-birth-1950-onwards-multiples-of-5.csv). The data is filtered on Year (life-expectation-at-birth-1950-onwards-multiples-of-5.csv), which keeps 2020. The view is filtered on Latitude (generated) and Longitude (generated). The Latitude (generated) filter keeps non-Null values only. The Longitude (generated) filter keeps non-Null values only.

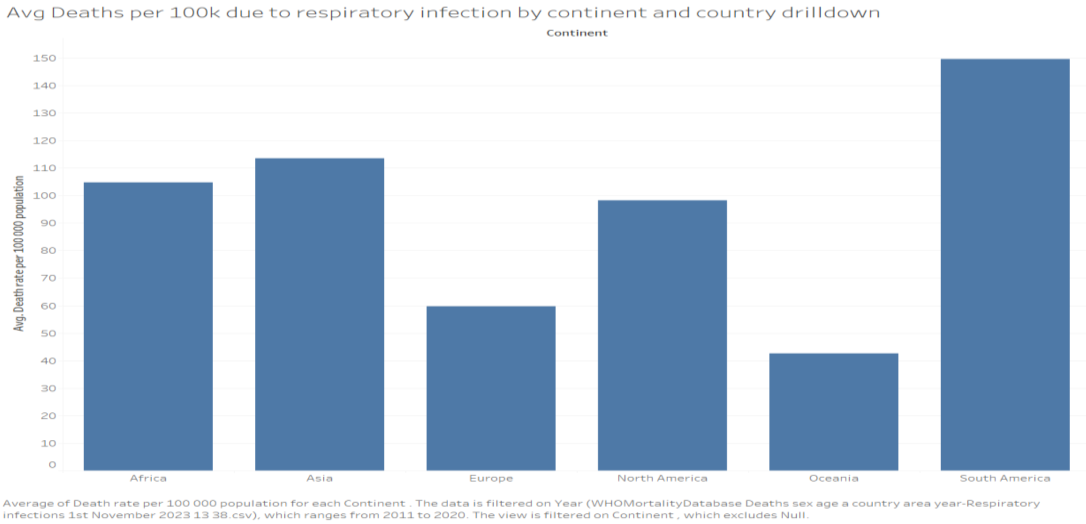
Up to year 2020 The country with the highest life expectancy for females at birth is Hong Kong, with a life expectancy of 88.2818 years, and for males, it is also Hong Kong, with a life expectancy of 82.2385 years. On the other end of the spectrum, the country with the lowest life expectancy for females at birth is the Central African Republic, with a life expectancy of 56.7567 years, and for males, it is also the Central African Republic, with a life expectancy of 52.5817 years. These figures illustrate the significant variation in life expectancy across different countries. The data points would represent the life expectancy for females and males at birth in each country, with latitude and longitude determining their position on the map. The color or size of these points could then indicate the magnitude of the life expectancy, allowing for a visual comparison across regions. The data description indicates life expectancy figures from a dataset spanning from 1950 to 2020, with measurements taken at intervals of every five years.

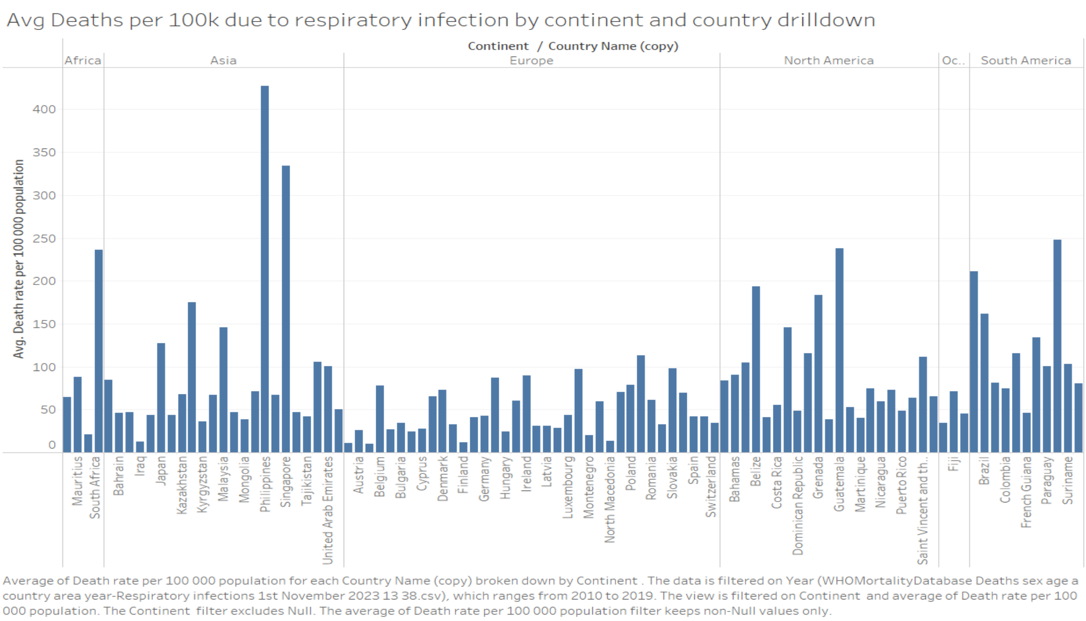
Visualization 5: Air Pollution and Respiratory Deaths Correlation



I plot a Horizontal Bar graph with the y-axis representing countries. The x-axis would quantify the death rate per 100,000 people. Afghanistan, with the highest death rate of 10,277.53, would be at the top, while Andorra with a rate of 451.71, one of the lowest, would be positioned towards the bottom. This visual representation would highlight the stark contrasts in air pollution's impact on public health across different regions. Wealthier countries generally have lower rates, possibly reflecting more robust air pollution control policies and healthcare systems.

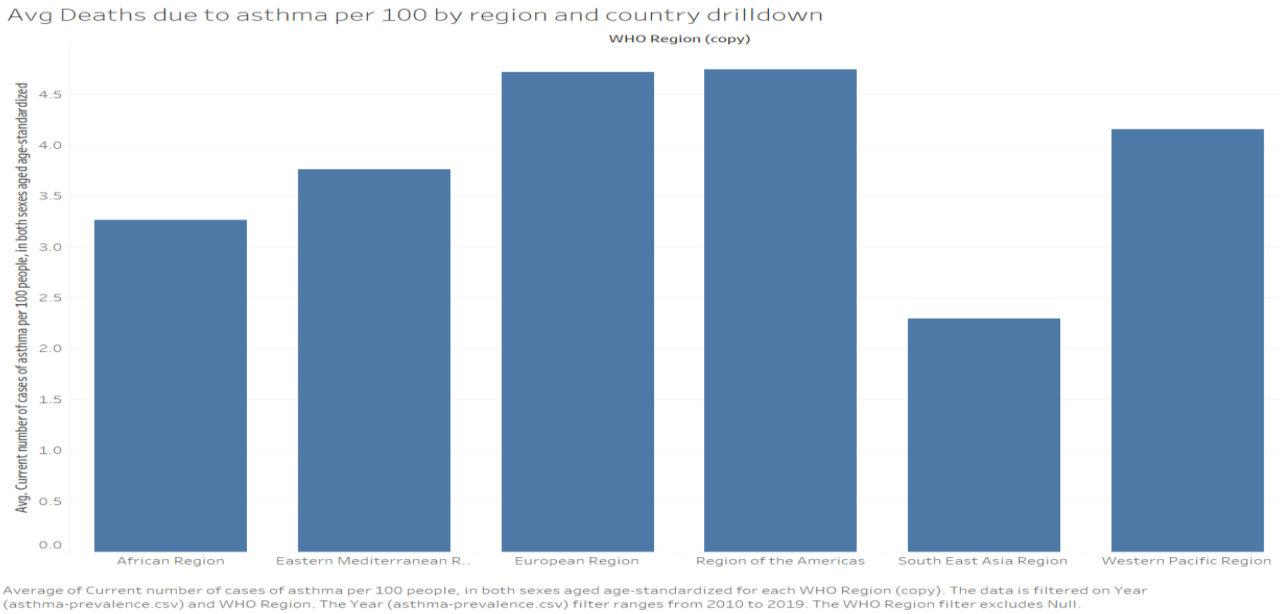
Visualization 6: Respiratory Infection Death Rates by Continent

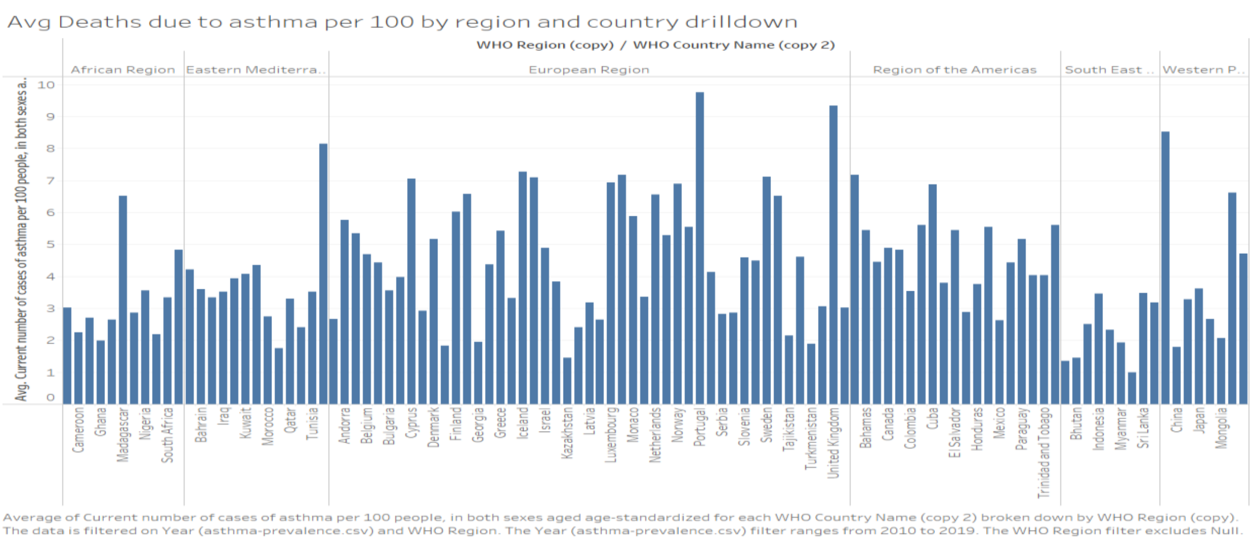




In my interpretation of the graph, the x-axis would represent the continents, with a further drill-down option to individual countries. The y-axis would quantify the average death rate per 100,000 population attributed to air pollution. The data highlights a gradient of pollution-related mortality across continents, with Oceania at the lower end and South America at the higher end of the spectrum. This continental view is nuanced by the country-specific data which shows significant within-continent variability, suggesting that continent-level trends do not always reflect the situation within individual countries. For countries, Belarus in Europe presents a much lower death rate compared to other European countries, and similarly, some Asian countries have lower rates than the continental average. This discrepancy raises questions about the interplay of environmental policies, healthcare systems, and pollution control measures across different regions.

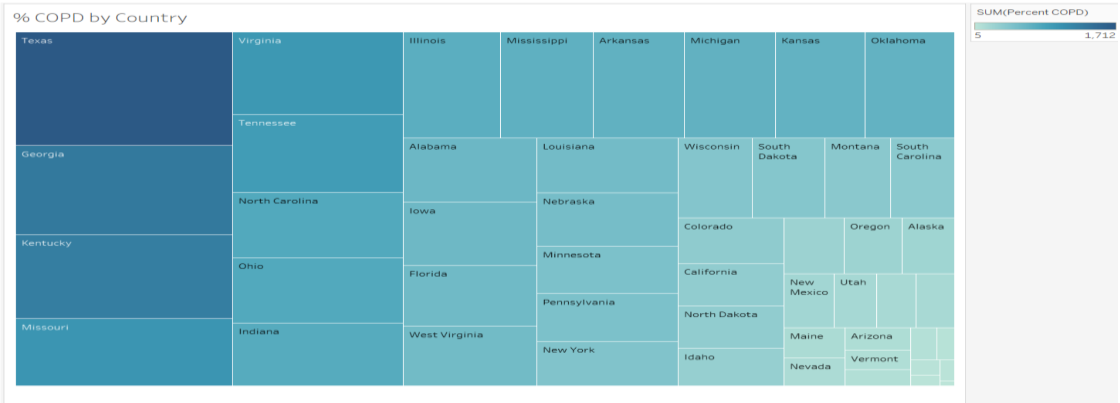
Visualization 7: Asthma Prevalence by WHO Region





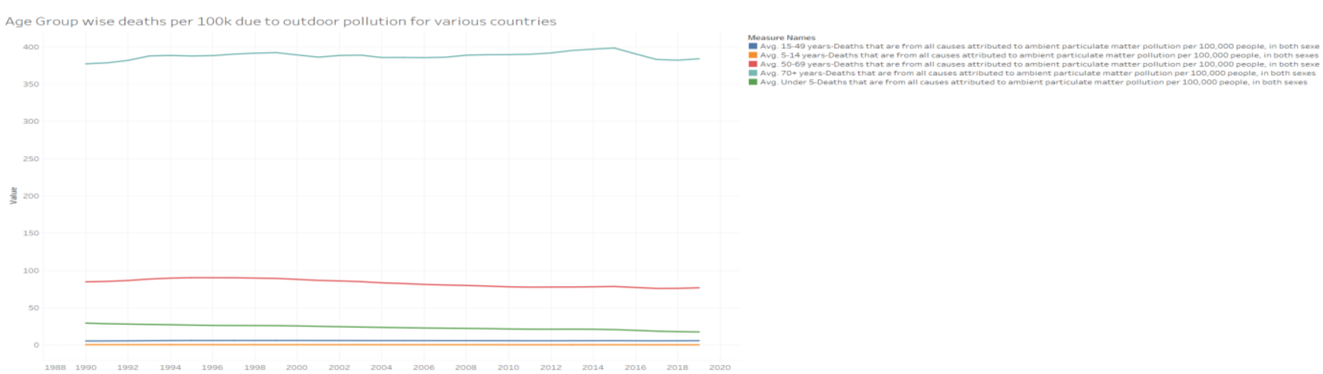
From my perspective, the x-axis on the graph represents the regions and the countries within them, and the y-axis indicates the average current number of cases of asthma per 100 people, age standardized. The data shows a range from 2.3 cases in the Southeast Asia Region up to 4.74 cases in the Region of the Americas at a regional level. Notably, there's substantial variation within regions; for instance, in the European Region, Kazakhstan has only 1.45 cases, while the United Kingdom has 9.35 cases. This suggests a complex interplay of genetic, environmental, and healthcare-related factors influencing asthma prevalence. Analyzing these differences could reveal important insights into the effectiveness of public health interventions and the impact of environmental conditions on respiratory health.

Visualization 8: Percentage COPD Prevalence by States in USA



The tree map graph displays the percentage of Chronic Obstructive Pulmonary Disease (COPD) by state in the USA. Texas appears to have the largest area on the tree map, indicating it has the highest rate of COPD among the states shown. In contrast, Hawaii, which is not visible on the provided segment of the graph, presumably has a much smaller area, denoting the lowest disease rate. Tree maps represent data hierarchically through nested rectangles, with area size corresponding to data value; hence, larger rectangles such as Texas signify higher values. The varying shades of blue may represent different intensities or ranges of COPD rates, with darker shades typically indicating higher values. This visual distribution provides a quick comparative view of COPD prevalence across states in the USA.

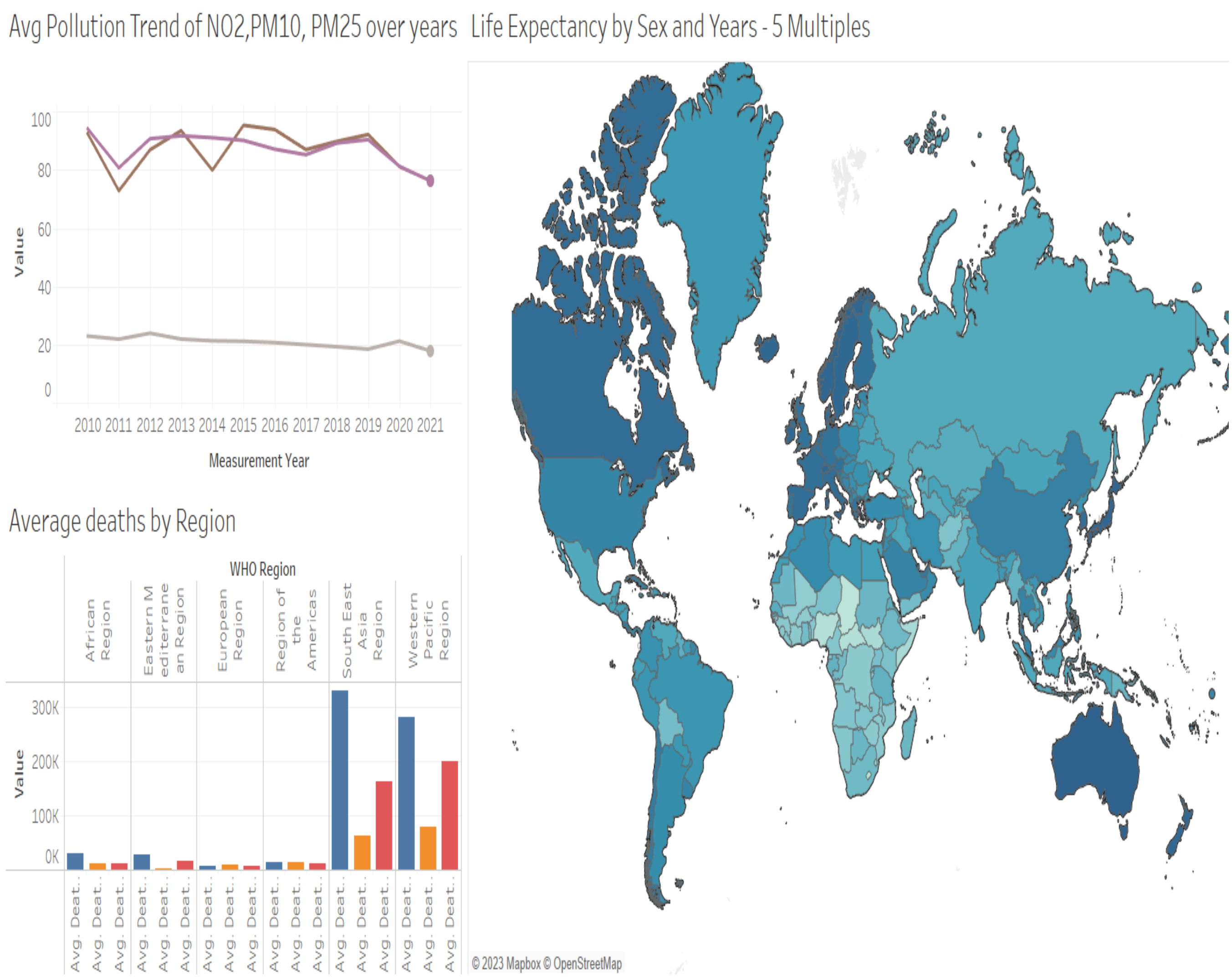
Visualization 9: Outdoor Pollution and Age-Specific Mortality



The trends of Avg. 15-49 years-Deaths that are from all causes attributed to ambient particulate matter pollution per 100,000 people, in both sexe, Avg. 5-14 years-Deaths that are from all causes attributed to ambient particulate matter pollution per 100,000 people, in both sexes, Avg. 50-69 years-Deaths that are from all causes attributed to ambient particulate matter pollution per 100,000 people, in both sexe, Avg. 70+ years-Deaths that are from all causes attributed to ambient particulate matter pollution per 100,000 people, in both sexes and Avg. Under 5-Deaths that are from all causes attributed to ambient particulate matter pollution per 100,000 people, in both sexes for Year (outdoor-pollution-rates-by-age.csv). Color shows details about Avg. 15-49 years-Deaths that are from all causes attributed to ambient particulate matter pollution per 100,000 people, in both sexe, Avg. 5-14 years-Deaths that are from all causes attributed to ambient particulate matter pollution per 100,000 people, in both sexes, Avg. 50-69 years-Deaths that are from all causes attributed to ambient particulate matter pollution per 100,000 people, in both sexe, Avg. 70+ years-Deaths that are from all causes attributed to ambient particulate matter pollution per 100,000 people, in both sexes and Avg. Under 5-Deaths that are from all causes attributed to ambient particulate matter pollution per 100,000 people, in both sexes.

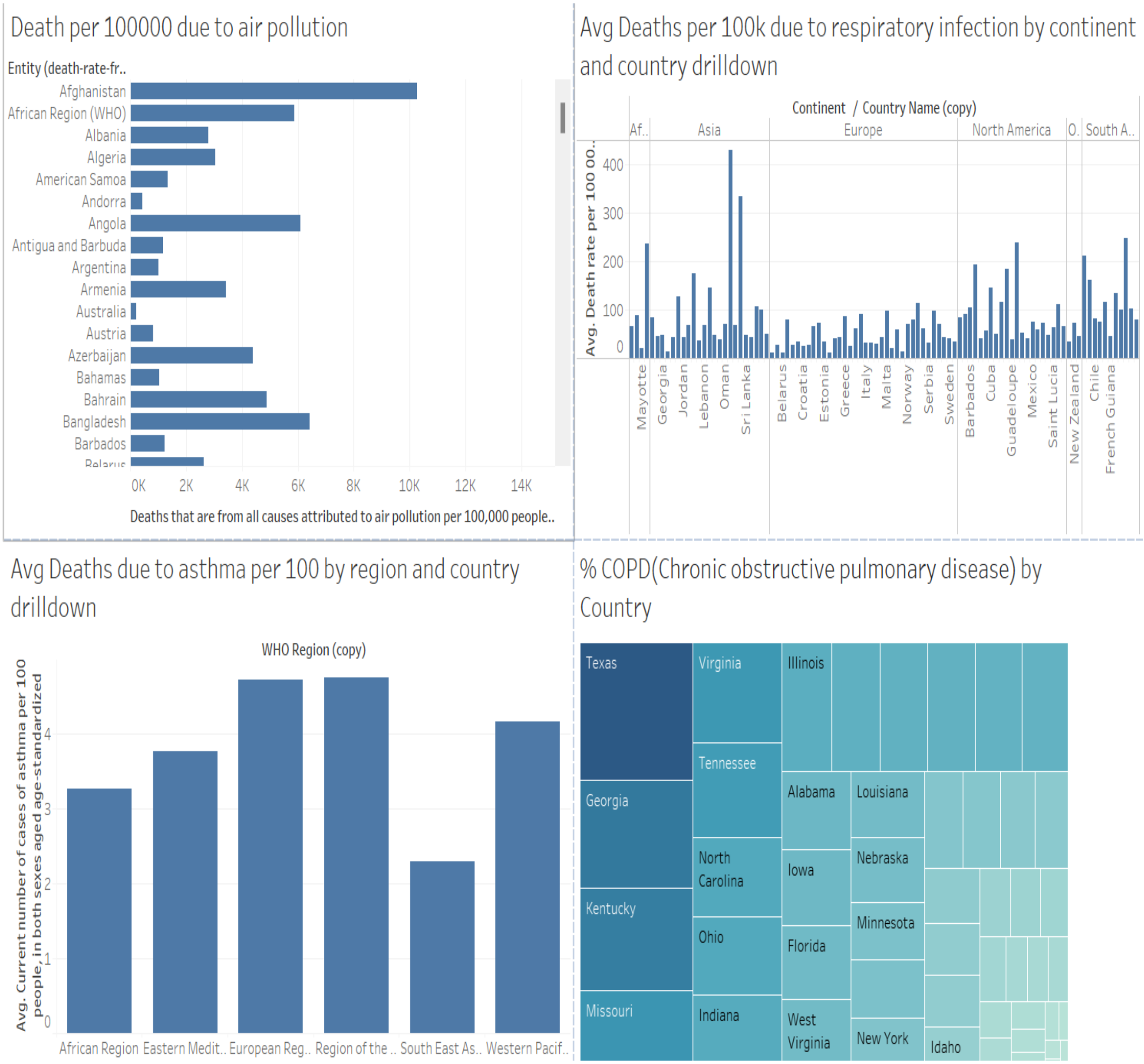
In the graph, the x-axis represents the years from 1990 to 2019, while the y-axis displays the death rate per 100,000 people. Each line or bar would correspond to a different age group, from under 5s to 70+ years. The data shows an overall trend where the youngest and oldest age groups have higher death rates due to ambient particulate matter pollution, with the 70+ age group consistently experiencing the highest rates. Interestingly, while other age groups show some fluctuation, the rates for the 70+ category appear to increase over time, indicating heightened vulnerability or changes in population demographics and pollution levels.

Dashboard 1:



The dashboard provides a comprehensive view of pollution and its impact on health globally, conveyed through several data visualizations. It includes a line graph that reveals the trends in nitrogen dioxide (NO2), PM10, and PM25 levels from 2010 to 2021, showing a consistent pattern with minor fluctuations. A bar chart further breaks down the average pollution-related deaths by WHO regions, with color-coded bars suggesting differentiation by pollutant or category. Additionally, there's a choropleth map that depicts life expectancy across the world, with varying shades possibly indicating life expectancy ranges, although the legend isn't shown.

Dashboard 2:



The dashboard analysis presents a detailed overview of the health impact of air pollution across various global regions. The first chart shows the death rate per 100,000 individuals attributable to air pollution for a range of countries, indicating a significant variance between them. Moving on, there's a bar chart detailing average deaths per 100,000 due to respiratory infections by continent, which allows for a comparison between different geographies. Additionally, I can see a focused bar chart that provides average deaths due to asthma per 100 individuals, categorized by WHO regions. This visual seems to highlight regional disparities in the prevalence of asthma-related fatalities. Lastly, there's a heat map representing the percentage of chronic obstructive pulmonary disease (COPD) by country, with a more granular breakdown by states within the United States.

Conclusions

For the research question on how ambient particulate matter pollution levels correlate with all-cause mortality rates across different countries, the conclusion drawn from the analysis is that there is a significant correlation between high levels of particulate matter pollution and increased mortality rates. Countries with higher pollution levels, such as Bahrain, China, India, and Mongolia, show alarmingly high mortality rates. This suggests that pollution is a critical factor in health outcomes, likely due to its impact on respiratory and cardiovascular diseases. The variance observed across countries highlights the need for targeted public health policies and pollution control measures, particularly in regions with economic challenges that may limit the capacity for environmental regulation.

Regarding the question on the prevalence of chronic respiratory diseases like asthma and COPD across different regions, the analysis concludes that these diseases are unevenly distributed across WHO regions and countries. Environmental factors, healthcare accessibility, and genetic predispositions appear to play significant roles in this variability. For example, the high number of asthma cases in the Region of the Americas could be influenced by a combination of urbanization, air quality, and healthcare practices. In contrast, the lower rates of asthma observed in Southeast Asia may relate to different lifestyle factors and environmental exposures. These findings point to the importance of considering regional characteristics when developing health interventions and underscore the complex interplay of factors contributing to respiratory health outcomes.

Additional Research Questions:

1. Is there a discernible impact of improved air quality on life expectancy within countries that have implemented stringent air pollution controls?
2. Can we isolate the effect of air pollution on COPD and asthma prevalence from other contributing factors, using longitudinal cohort studies?
3. How does exposure to indoor versus outdoor air pollution contribute to the overall risk of death from respiratory conditions, and how has this risk changed over the last decade?
4. What are the economic impacts of air pollution in terms of healthcare costs associated with pollution-related diseases across different WHO regions?
5. What role does air pollution play in the global burden of disease, and how does this burden shift with changing pollution levels?
6. How might climate change influence the levels of air pollutants and thus affect health outcomes related to air pollution in the future?