UNIVERSITY OF PENNSYLVANIA

SCHOOL OF SOCIAL POLICY AND PRACTICE

MASTER OF SCIENCE IN SOCIAL POLICY AND DATA SCIENCE

MSSP 6070-001 202330 PRACTICAL PROGRAMMING FOR DATA SCIENCE (PROF. RICHARD HARTWELL)

CASE STUDY REPORT

# **FROM DATA TO INSIGHTS:** ANALYZING CHRONIC RESPIRATORY DISEASES WITH RESPECT TO AIR POLLUTION

## USING A DATA SCIENCE APPROACH

Rose Barragan Barranco

December 22, 2023

# CONTENTS


PAGE

PART 1

Foundations of the project 04

PART 2 10

Air Quality analysis: Global and US Approach

PART 3 25

Chronic Respiratory Disease Analysis: US Approach

PART 4 30

Testing the Hypothesis: explore the Link between Air Pollution and Chronic Respiratory Diseases and make predictions

PART 5 33

Recommendations and Conclusions

Appendix 36

References 37

# ABSTRACT

According to the World Health Organization (WHO), nearly 100% of the population breathes air that is harmful to their health due to air that exceeds established air quality limits. The most significant worldwide environmental threat to public health is air pollution, causing more than 7 million premature deaths yearly [4]. Like this, air pollution and respiratory diseases are closely linked, as all major pollutants have a dangerous impact on human health.

Therefore, this report examines the presence of three major air pollutants – PM2.5, PM10, and NO2 - in over 6000 cities in 117 countries around the world, as well as the reported fatalities due to pulmonary diseases.

The ultimate goal is to establish a direct relationship between the presence of these harmful particles in the air and the increase in deaths from chronic respiratory diseases. In addition, it is intended to offer concrete suggestions and strategies to reduce air pollution levels to protect public health.

# PART I
# FOUNDATIONS OF THE PROJECT


1. INTRODUCTION TO THE CONTEXT AND RELEVANCE OF THE STUDY

    1.2. ALIGNMENT WITH THE SUSTAINABLE DEVELOPMENT GOALS (SDGS)

2. GOALS OF THIS PROJECT AND METHODS

3. SCOPE AND STRUCTURE

4. STAKEHOLDERS (INTENDED AUDIENCE)

# 1. INTRODUCTION TO THE CONTEXT AND RELEVANCE OF THE STUDY

The urgency of addressing air pollution is evident in the current context, where the direct impact on public health has become a global health crisis. The alarming figures of premature deaths attributable to poor air quality, reaching nearly 7 million deaths yearly [4], highlight the imperative need for concrete and effective actions.

That is why the World Health Organization (WHO) [1] has responded to this growing concern by updating its Air Quality Guidelines [2], further reinforcing them to help countries better assess the safety of their air. In particular, standards for pollutants such as fine particles (PM2.5 and PM10) and nitrogen dioxide (NO2) have been made stricter.

- On the one hand, fine particles (PM10 and PM2.5) can reach a size smaller than 10 micrometers in diameter and smaller than 2.5 micrometers in diameter, respectively, posing the greatest problems, as they can penetrate deep into the lungs and enter the bloodstream, causing cardiovascular, cerebrovascular, and respiratory diseases [3].

- On the other hand, nitrogen dioxide is associated with respiratory diseases, especially asthma, and can lead to respiratory symptoms such as coughing or shortness of breath.

For that, WHO established average annual limits for these pollutants: **5 µg/m3 for PM2.5, 15 µg/m3 for PM10, and 10 µg/m3 for NO2** [2]. These limits seek to minimize risks to human health.

*Size of MP2.5 & PM10 in comparison with human hair and beach sand*


*Environmental Protection Agency. (2023)*

Air pollution has several sources of origin, coming mainly from human activity such as the combustion of fossil fuels for transportation and the incineration of waste materials, the burning of raw coal for cooking, heating, and lighting, as well as the use of kerosene to generate electricity, thus increasing air pollution [4].

It is, therefore, important to emphasize that this is a global problem that must be addressed from an international perspective, as set out in the United Nations (UN) Sustainable Development Goals (SDGs).

Global collaboration and the commitment of states are fundamental to achieving effective strategies and long-lasting solutions in the field of public health, as well as from the point of view of sustainability and the environment.

# 1.2. ALIGNMENT WITH THE SUSTAINABLE DEVELOPMENT GOALS

As mentioned in the previous section, it is essential to acknowledge that the problem of air quality and its harmful consequences for humans and the planet involves a joint international perspective through the Sustainable Development Goals (SDGs) of the United Nations [5].

The SDGs set out a universal agenda structured in 17 interconnected global goals established in 2015 to jointly address social and environmental challenges to achieve a more equitable, prosperous, and sustainable future.

Tackling air pollution aligns closely with several of the SDGs, in particular, Goal #3, *"Ensure healthy lives and promote well-being for all at all ages"* [6]. To accomplish this goal, there are several milestones or targets, such as #3.4 and #3.9 [7], which are especially relevant in the context of air quality and its impact on human health.


*Open Development Myanmar . (2020,)*

3.4

## Target 3.4

Air pollution is a significant factor in non-communicable (such as cardiovascular illnesses, cancer, diabetes, or chronic respiratory diseases). Thus, target 3.4 looks to decrease premature mortality from non-communicable diseases by a third through prevention and treatment to decrease the mortality rate associated with these conditions.

3.9

## Target 3.9

By 2030, decrease the number of fatalities and sicknesses caused by hazardous chemicals and air, water, and soil pollution. Air pollution, both household and ambient, is a key factor in this target. Thus, policies and actions must be created to improve air quality, having a direct impact on reducing the mortality rate attributed to air pollution

In addition, improving air quality has positive implications for other SDGs. For example, since air pollution and climate change are caused primarily by burning fossil fuels, clean air actions are also climate actions (SDG #13). Thus, improving air quality is not only fundamental to human health and well-being (SDG #3) but is also a critical component in the quest to combat inequality, climate change, and environmental degradation, thus contributing to a more holistic and practical approach to achieving the Sustainable Development Goals.

# 2. GOALS OF THIS PROJECT AND DS METHOD

This study is conducted in a global context where air pollution is one of our time's most pressing environmental and public health challenges. In this sense, the relevance of this study lies in the need to understand better the relationship between exposure to different types of pollutants (PM2.5, PM10, and NO2) and their particular effects on chronic respiratory diseases (CRD).

The research seeks to provide a sound scientific basis to guide public policies and individual actions to improve air quality. This knowledge is crucial for designing effective strategies to mitigate air pollution and, in turn, reduce the incidence of Chronic Respiratory Diseases. Thereby improving the quality of life for millions of people and addressing this global health and environmental crisis.

To achieve this objective, this report will analyze two datasets - the Air Quality Dataset by WHO and the Chronic Respiratory Diseases Dataset by the Institute for Health Metric and Evaluation - to:

- Identify patterns and monitor the current state of air quality around the world, highlighting the most affected regions
- Determine which air pollutant is most detrimental to CRDs and assess the impact of other pollutants
- Identify the most vulnerable age and gender groups of CRDs
- To draw forecasting trends among different regions

These valuable insights will be the basis for formulating evidence-based recommendations to improve air quality standards, reduce pollutant emissions, and ultimately mitigate the risk of CRDs. So, this project seeks to provide a detailed and up-to-date view of global air quality and its implications for public health from a data science perspective.

Using **Python and the Google Colab Platform**, a comprehensive analysis is developed that not only evidences the magnitude of the problem but also provides valuable tools for monitoring and informed decision-making. This technological and analytical approach is indispensable for understanding the complex dynamics of air pollution and for designing more effective strategies to mitigate its health impacts on a global scale.

# 3. SCOPE AND STRUCTURE

Following the last section's idea. This report showcases a comprehensive study that analyzes the relationship between air quality and the prevalence of CRD and defines a set of policy recommendations aligned with the SDGs. Given the comprehensive scope of the report, it is essential to establish a clear structure that facilitates more engaging storytelling and showcases the valuable insights obtained from the data.

**PART II**

focuses on Global Air Quality analysis using data [8] collected by WHO on the three air pollutants (PM2.5, PM10, and NO2, measured in µg/m3), covering over **6000 cities in 117 countries** grouped in 6 different regions [9]. This section will provide a global overview of air quality, identifying significant patterns and trends at the international level. It should be noted that this section will also include a detailed analysis of the understanding of pollutants in the United States of America.

**PART III**

focuses on Chronic Respiratory Diseases occurring in the **United States**. For this purpose, we will use data from the Institute for Health Metric and Evaluation and its study - Global Burden of Disease (GBD) [10] - that collects the number of deaths attributed to chronic respiratory diseases in five age groups (<20 years; 20-54 years; 55-89 years; 90-94 years; >95 years) and for both genders (female and male) across the 50 states.

**PART IV**

consists of the integration and comparative analysis. This section of the project aims to integrate the findings of the first two parts to test the hypothesis of whether there is a correlation between the number of deaths from chronic respiratory diseases and the levels of the three air pollutants based on insights from the data. In addition, we will seek to identify patterns and correlations that may suggest relationships between air quality and respiratory health, as well as identify future trends.

**PART V**

The final part of this project will present the conclusions derived from the integrated analysis, highlighting the key relationships identified and proposing recommendations based on the findings.

# 4. MAIN STAKEHOLDERS

## 1 Global Air Pollution and Health Technical Advisory Group (GAPH-TAG)

With 75 members [11], serves as a technical advisor to WHO, providing insight and guidance on assessing the health impacts of air pollution, as well as advising on methodologies and data entry for assessing population exposure to air pollutants. In connection with this project, GAPH-TAG is identified as a relevant stakeholder as it evaluates interventions and policies to address the health effects of air pollution.

## 2. Scientific Advisory Group on Air Pollution and Health (SAG)

Its eight members [12] advise WHO on issues related to ambient and domestic air pollution and its impact on health, as well as implementation guidelines and strategies. In addition, the SAG identifies critical issues and emerging problems for discussion and addressing by GAPH-TAG.

## 3. Air Quality and Health Unit (AQH)

The third identified stakeholder is the Air Quality and Health Unit (AQH) [13], which promotes healthy sectoral interventions and policies, addresses key health risks related to indoor and outdoor air pollution and contributes to the additional benefits of climate change mitigation policies. AQH leads the implementation of World Health Assembly Resolution 68.8 and focuses on monitoring the achievement of targets such as the mortality rate attributed to household and ambient air pollution.

## 4. US Office of Science and Technology Policy

From the national point of view, the US Office of Science and Technology Policy (OSTP) is responsible for directly advising the US President regarding science and technology [14]. The OSTP is made up of six teams, the most relevant for the present project being the following:

- The **Health team** uses science, technology, and innovation to improve health outcomes for all Americans, including health promotion, disease prevention and treatment, and improving access to and quality of health care [15].
- The **Climate and Environment team** works to advance government priorities related to climate, environmental justice, and nature. They seek to provide sound scientific information to inform government policies and actions, promote equity and inclusion, and collaborate with partners on climate and environmental issues [16].

# PART II

# AIR QUALITY ANALYSIS: GLOBAL AND US APPROACH

1. DATASETS OVERVIEW

1.1. DESCRIPTION OF THE DATASET: SOURCES

1.2. DESCRIPTION OF THE DATASET: CLARIFICATIONS AND LIMITATIONS

1.3. DESCRIPTION OF THE DATASET: VARIABLES

2. DESCRIPTIVE ANALYTICS FOR AIR QUALITY DATASET: GLOBAL APPROACH.

2.1. MEAN AND STANDARD DEVIATIONS

2.2. NAS, OUTLIERS AND DISTRIBUTIONS.

2.3. CORRELATION MATRIX AND HEAT MAP.

3. DESCRIPTIVE ANALYTICS FOR AIR QUALITY DATASET: US APPROACH.

3.1. NAS, OUTLIERS AND DISTRIBUTIONS.

4. EVOLUTION OF THE AIR POLLUTANTS.

4.1. PM2.5 (MG/M3)

4.2. PM10 (MG/M3)

4.3. NO2 (MG/M3)

4.4. REGION OF THE AMERICAS VS. USA

# 1. DATASET OVERVIEW

## 1.1. DESCRIPTION OF THE DATASET: SOURCES

The primary sources for this dataset are official reports submitted by countries to WHO, national and subnational air quality reports, and websites containing PM10, PM2.5, and NO2 measurements. In addition, measurements are provided by regional networks, such as Clean Air for Asia, the European Environment Agency, and the AirNow Program of US embassies and consulates. In cases where official data were not available, values provided by United Nations agencies, development agencies, and peer-reviewed journals were used [8].

## 1.2. DESCRIPTION OF THE DATASET: CLARIFICATIONS AND LIMITATIONS

The dataset was preprocessed before starting the study to adapt the dataset's variables to the focus of the project. This was achieved by eliminating other variables that were already compiled in the dataset (such as the number and type of monitoring stations, references, or IDs) without negatively affecting the quality and analysis of the present project. That said, it is important to note that the original datasets have certain limitations [18]:

Critical hotspot locations: locations designated as 'hotspots' (such as roads with a high traffic volume) were omitted, which often represent areas with exceptionally high pollution levels but were not considered representative of the average city population exposure.

Industrial areas: Data from sites located exclusively in industrial areas were omitted, as they present higher levels of contamination compared to areas where people reside. The exclusion of these areas is intended to focus on pollution levels that more accurately reflect typical human exposure.

Globalization: The comparability of data from different countries is limited due to different factors such as different locations of measurement stations, time coverage, heterogeneous quality of measurements, language barriers limiting access to data, etc.

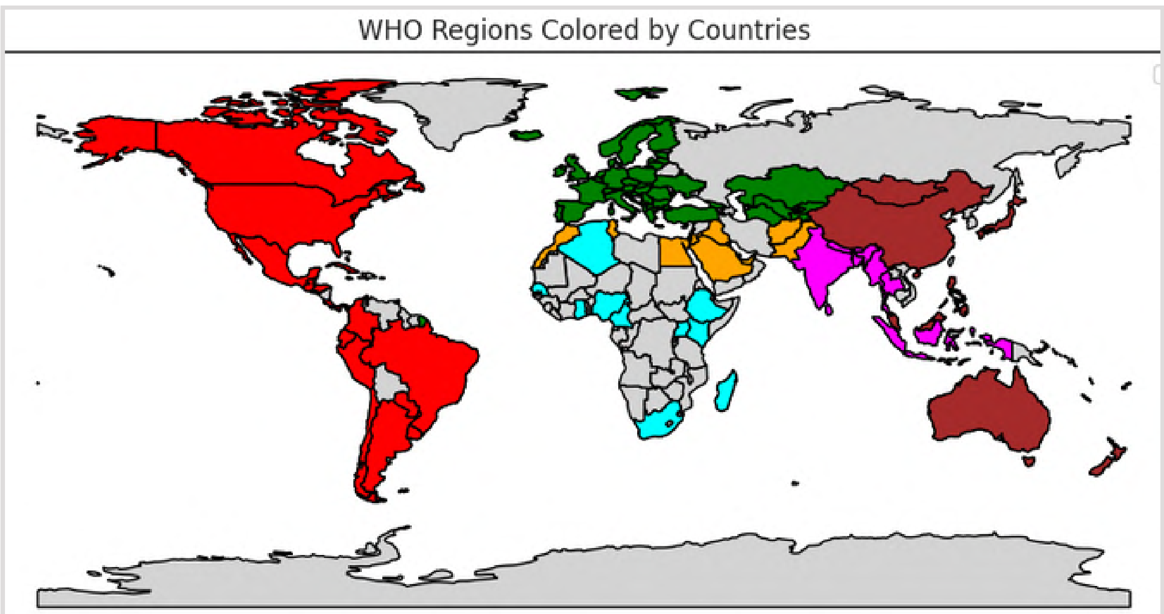
## 1.3. DESCRIPTION OF THE DATASET: VARIABLES

| ***Column name*** | ***Description*** |
|---|---|
| Global Air Quality Dataset | |
| **WHO Region** | A categorical variable that groups the countries according to the assigned region, identifying a total of six regions: African region, Region of the Americas, Southeast Asian region, European region, Eastern Mediterranean region, and Western Pacific region. |
| **ISO3** | A categorical variable that indicates the country code as an abbreviation. |
| **WHO Country Name** | A categorical variable that indicates the name of the country where the air pollutant record is obtained. There is a total of 117 countries in the dataset.<br>- African Region: 12 countries<br>- Eastern Mediterranean Region: 14 countries<br>- European Region: 48 countries<br>- Region of the Americas: 22 countries<br>- South East Asia Region: 9 countries<br>- Western Pacific Region: 12 countries |
| **City or Location** | Name of the city where the air pollutant record is obtained. |
| **Measurement Year** | Year of the annual mean concentration.<br>The dataset collects data from 2010 to 2020. |
| **PM10 (µg/m3)** | The annual mean concentration of particulate matter with a diameter of 10 µm or less. |
| **PM25 (µg/m3)** | The annual mean concentration of particulate matter with a diameter of 2.5µm or less. |
| **NO2 (µg/m3)** | The annual mean concentration of nitrogen dioxide |
| **State** | A categorical variable that indicates the abbreviation of the 50 US states |

# 2. DESCRIPTIVE ANALYTICS FOR AIR QUALITY DATASET: GLOBAL APPROACH.

## 2.1. MEAN AND STANDARD DEVIATIONS

The data is collected from more than 6,000 cities in 117 countries, organized in six different regions, ranging from 2010 to 2020, comprises a total of 32,152 records, through which the following statistical values are obtained:




**REGION OF THE AMERICAS:**

1. Brazil
2. Chile
3. Colombia
4. Ecuador
5. Mexico
6. Peru
7. Paraguay
8. El Salvador
9. United States of America
10. Guatemala
11. Venezuela
12. Jamaica
13. Bolivia
14. Canada
15. Costa Rica
16. Honduras
17. Panama
18. Argentina
19. Uruguay
20. Cuba
21. Trinidad and Tobago
22. Bahamas

**EUROPEAN REGION:**

1. Austria
2. Belgium
3. Bulgaria
4. Bosnia and Herzegovina
5. Switzerland
6. Cyprus
7. Czechia
8. Germany
9. Denmark
10. Spain
11. Estonia
12. Finland
13. France
14. United Kingdom
15. Greece
16. Croatia
17. Hungary
18. Ireland
19. Iceland
20. Italy
21. Lithuania
22. Luxembourg
23. Latvia
24. North Macedonia

1. Malta
2. Montenegro
3. Netherlands
4. Norway
5. Poland
6. Portugal
7. Romania
8. Serbia
9. Slovakia
10. Slovenia
11. Sweden
12. Turkey
13. Andorra
14. Israel
15. Monaco
16. Albania
17. Georgia
18. Russian Federation
19. Kazakhstan
20. Ukraine
21. Uzbekistan
22. Kyrgyzstan
23. Tajikistan
24. Turkmenistan

**EASTERN MEDITERRANEAN REGION**

1. Bahrain
2. Iran
3. Pakistan
4. Tunisia
5. United Arab Emirates
6. Morocco
7. Lebanon
8. Egypt
9. Kuwait
10. Saudi Arabia
11. Iraq
12. Jordan
13. Qatar
14. Afghanistan

**SOUTH EAST ASIA REGION:**

1. Bangladesh
2. Bhutan
3. India
4. Indonesia
5. Sri Lanka
6. Maldives
7. Myanmar
8. Nepal
9. Thailand

**AFRICAN REGION:**

1. Mauritius
2. Nigeria
3. Senegal
4. South Africa
5. United Republic of Tanzania
6. Cameroon
7. Ghana
8. Uganda
9. Madagascar
10. Ethiopia
11. Kenya
12. Algeria

**WESTERN PACIFIC REGION:**

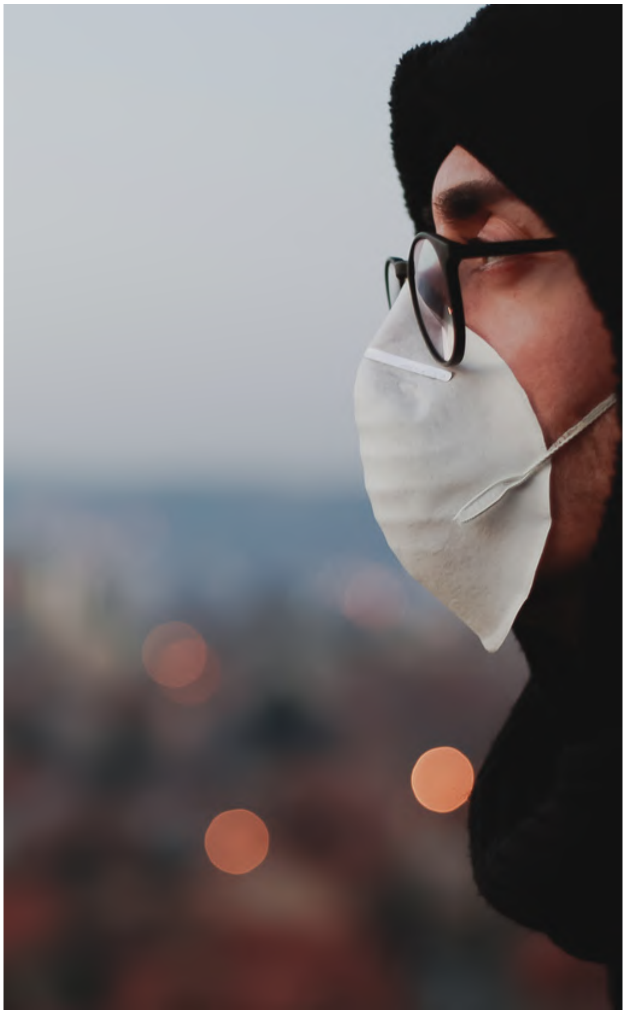
1. Australia
2. Japan
3. Republic of Korea
4. Mongolia
5. Malaysia
6. Singapore
7. China
8. Viet Nam
9. New Zealand
10. Philippines
11. Fiji
12. Lao People's Democratic Republic

*Source: Barragan,R., (December, 2023) (own).*

| *index* | PM2.5 (µg/m3) | PM10 (µg/m3) | NO2 (µg/m3) |
|---|---|---|---|
| *count* | 15032.0 | 21071.0 | 22186.0 |
| *mean* | 22.9199 | 30.3867 | 20.6211 |
| *std* | 17.9305 | 28.9775 | 12.1353 |
| *min* | 0.01 | 1.04 | 0.0 |
| *25%* | 10.34 | 16.97 | 12.0 |
| *50%* | 16.0 | 22.0 | 18.805 |
| *75%* | 31.0 | 31.215 | 27.1675 |
| *max* | 191.9 | 540.0 | 210.68 |

*Source: Barragan,R., (December, 2023) (own).*

The mean PM2.5 indicates that the concentration of fine particles in the air is approximately 22.92 micrograms per cubic meter (µg/m³); while the standard deviation equals 17.93 µg/m³. On the other hand, the mean fine particle PM10 equals 30.39 (µg/m³); while the standard deviation is 28.98 µg/m³. This great variability in both particulates could indicate that pollutant levels vary widely over time, which could be a concern regarding air quality stability.



For NO2, the mean value is 20.62 µg/m³, and the standard deviation equals 12.14 µg/m³, suggesting that the concentration of nitrogen dioxide tends to have fewer fluctuations over time than the other two pollutants, which may be preferable from the point of view of air quality stability.

On the other hand, as introduced in chapter 1.1. “Context and Relevance of the Study,” the average annual limits suggested by the WHO for these pollutants are 5 µg/m3 for PM2.5, 15 µg/m3 for PM10, and 10 µg/m3 for NO2. So the mean obtained for each pollutant - PM2.5 (22.92 µg/m³), PM10 (30.39 µg/m³), and NO2 (20.62 µg/m³) - indicates that they are well above the WHO air health limits.

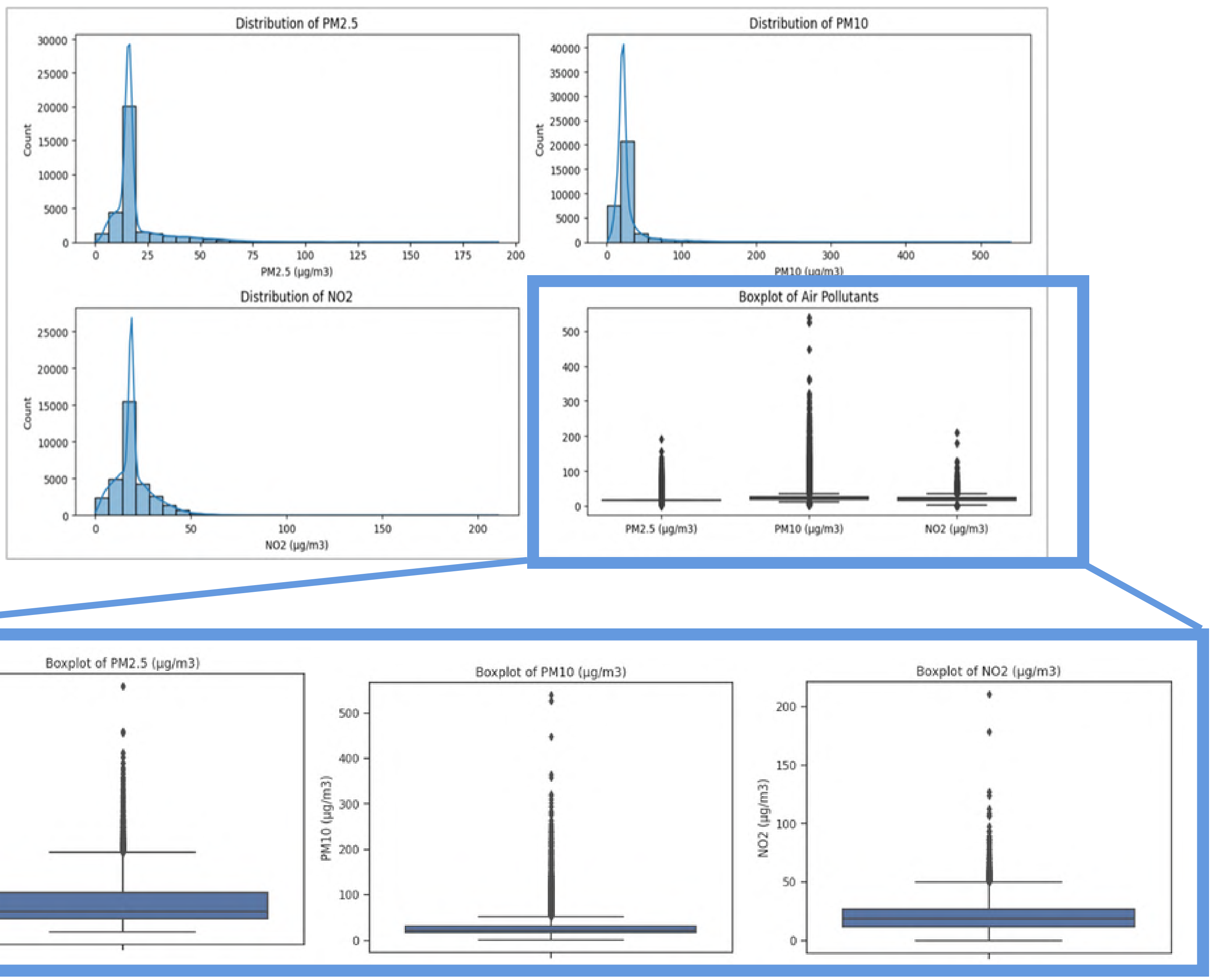
## 2.2. NAS, OUTLIERS AND DISTRIBUTIONS

Our dataset has a significant amount of missing values (NAs) in several columns. Specifically, the 'State' column has 30419 NAs. This is because the dataset includes cities from different locations, but the 'State' column is designed exclusively to associate US cities with their respective states.

Where most attention needs to be paid is, on the one hand, the null values for PM2.5 µg/m3 (17120); PM10µg/m3 (11081), and NO2 µg/m3 (9966), and on the other hand, the outliers. Any values more than three standard deviations away from the mean are considered outliers. Therefore, having identified the outliers, they are handled using z-scores and replaced with missing values (NaN).

Next, all missing values (NaN) in the data frame are imputed by replacing them with the median value of each respective column.

It is also observed that the data for each air pollutant (PM2.5 µg/m3, PM10µg/m3 y NO2 µg/m3) is not normally distributed. In each histogram, the data is positively skewed, meaning that the majority of the values are low, i.e., concentrated at the lower end.




*Source: Barragan,R., (December, 2023) (own).*

## 2.3. CORRELATION MATRIX AND HEAT MAP

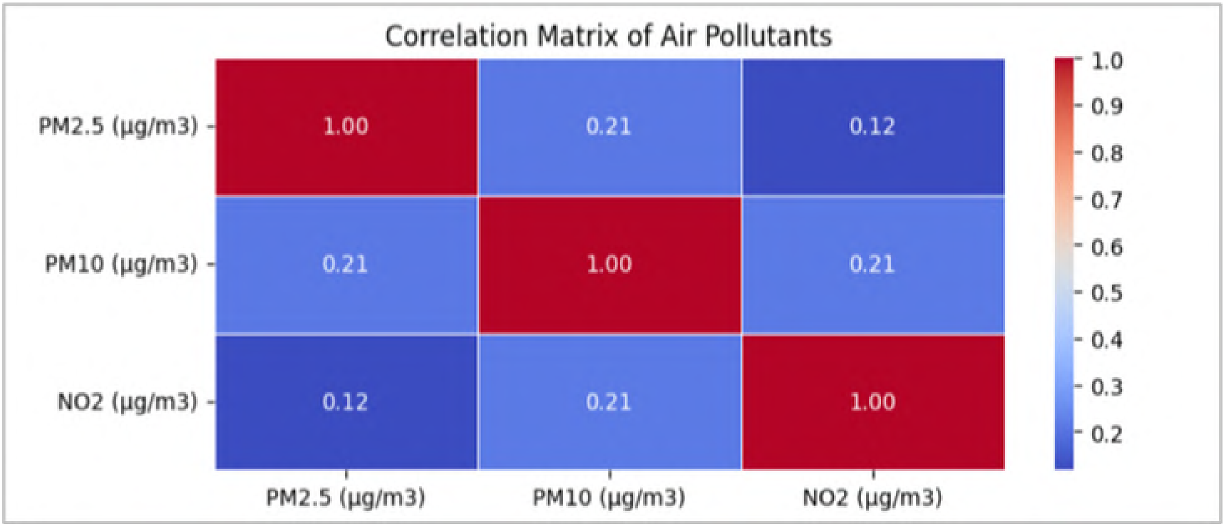
The correlation matrix shows the correlation coefficients between three air pollutants: PM2.5, PM10, and NO2:

- PM2.5 and PM10 reflect a positive but weak correlation coefficient of 0.21, indicating that the higher the levels of PM2.5, the higher the levels of PM10.

- PM2.5 and NO2 reflect a positive but even weaker correlation coefficient of 0.12, implying that changes in PM2.5 levels are just slightly associated with NO2 concentrations, although the relationship is not strong.

- PM10 and NO2 show a correlation coefficient similar to the correlation between PM2.5 and PM10, with a coefficient equal to 0.21, indicating a weak positive correlation. When the levels of PM10 increase, so do the levels of NO2.

All in all, It should be noted that the values for each correlation PM2.5 & PM10 (0.21), PM2.5 & NO2 (0.12), and PM10 & NO2 (0.21) are quite weak yet indicate a positive correlation.




*Source: Barragan,R., (December, 2023) (own).*

# 3. DESCRIPTIVE ANALYTICS FOR AIR QUALITY DATASET: US APPROACH.

## 3.1. MEAN AND STANDARD DEVIATIONS

Based on the previous analysis, this section will focus on the US data entries on the Air Quality Dataset, comprising a total of 1776 records, through which the following statistical values are obtained:

| *index* | PM2.5 (µg/m3) | PM10 (µg/m3) | NO2 (µg/m3) |
|---|---|---|---|
| *count* | 1776.0 | 1776.0 | 1776.0 |
| *mean* | 13.3327 | 20.7853 | 16.6251 |
| *std* | 4.0923 | 6.7184 | 6.2799 |
| *min* | 1.8 | 3.8 | 0.5 |
| *25%* | 10.2875 | 17.3725 | 13.4 |
| *50%* | 16.0 | 22.0 | 18.805 |
| *75%* | 16.0 | 22.0 | 18.805 |
| *max* | 20.4 | 65.3 | 41.100 |

*Source: Barragan,R., (December, 2023) (own).*

The mean PM2.5 indicates that the concentration of fine particles in the air is approximately 13.32 micrograms per cubic meter (µg/m³), while the standard deviation equals 4.092 µg/m³. On the other hand, the mean fine particle PM10 equals 20.78 (µg/m³), while the standard deviation is 6.718 µg/m³. For NO2, the mean value is 16.62 µg/m³, and the standard deviation equals 6.27 µg/m³, suggesting that the concentration of these air pollutants tends to have fewer fluctuations over time than in the global scenario, which may be preferable from the point of view of air quality stability.

As commented, the average annual limits recommended by the WHO for these pollutants are 5 µg/m3 for PM2.5, 15 µg/m3 for PM10, and 10 µg/m3 for NO2.

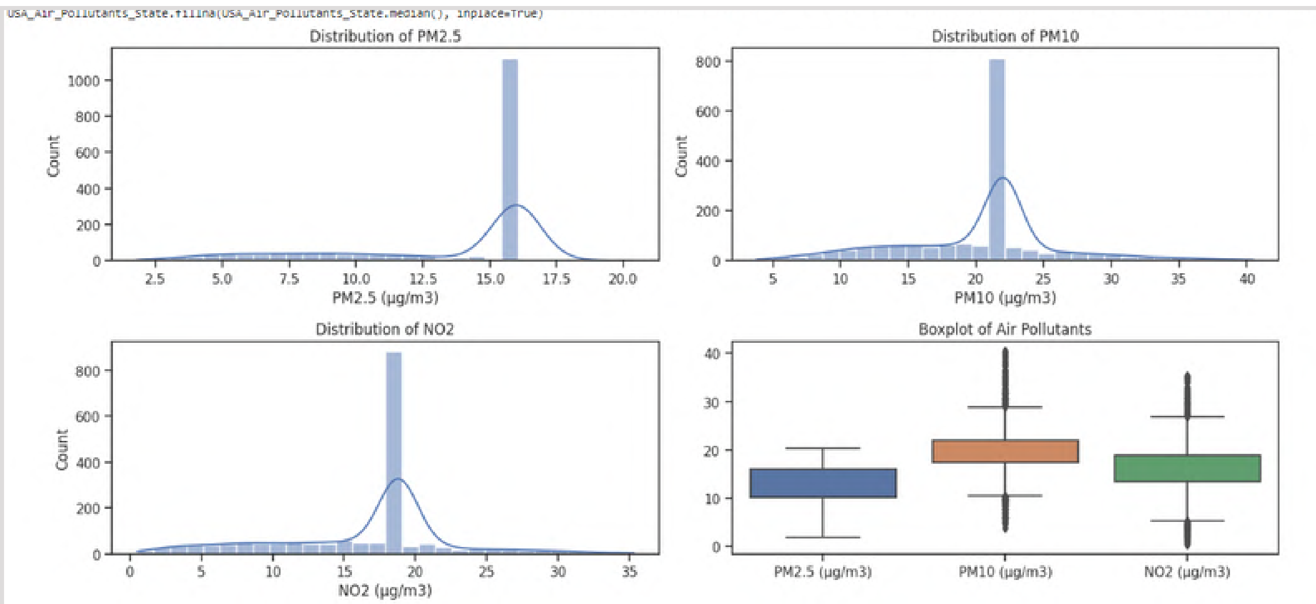
Therefore, although the United States is performing better, having lower values for each air pollutant, in comparison with the global approach - PM2.5 (22.92 µg/m³), PM10 (30.39 µg/m³), and NO2 (20.62 µg/m³), it is worth noting, that the US values are still above the WHO threshold for quality air.

## 3.2. NAS, OUTLIERS AND DISTRIBUTIONS

In the previous section, missing values (NAs) were handled for all values except for the "State" variable. Therefore, 43 NAs for "State" have been detected in the dataset's subsection.

To complete the dataset with the abbreviations of the states in the missing cells, the cities were matched with the state. In other words, a custom function was defined to impute the missing State abbreviations based on the city or locality names. Later, the function checked if a locality was in the dictionary (previously created) and assigned the corresponding name abbreviation for the state.

Hence, it was imputed that the state name was based on the city column, resolving the NAs problem and resulting in a more complete and accurate dataset.



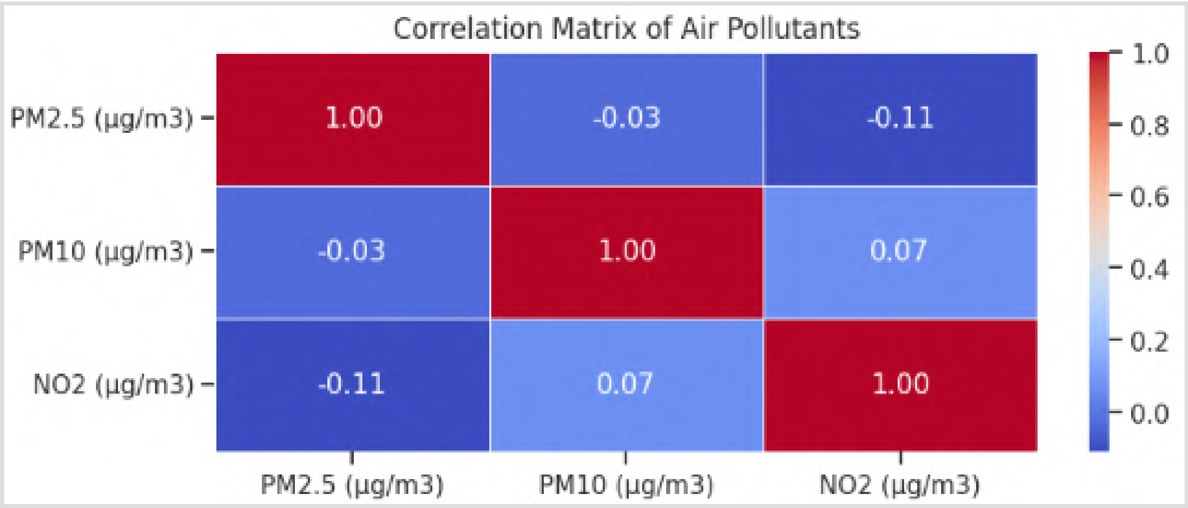

*Source: Barragan,R., (December, 2023) (own).*

## 3.3. CORRELATION MATRIX AND HEAT MAP

The correlation matrix shows the correlation coefficients between three air pollutants: PM2.5, PM10, and NO2, only this time they are following the US approach.

- PM2.5 and PM10 reflect an extremely weak correlation coefficient of -0.03, meaning that when the level of PM2.5 increases, the PM10 concentration levels decrease.
- PM2.5 and NO2 reflect an even more negative and weaker correlation coefficient of -0.11, implying that a rise in PM2.5 levels translates into a reduction in the concentration levels of NO2.
- PM10 and NO2 show a correlation coefficient equal to 0.07, indicating a very weak yet positive correlation. This suggests that when the levels of PM10 increase, so do the levels of NO2.

Overall, the correlation coefficient of Air Pollutants in the US is considerably lower and weaker than the coefficients of Air pollutants at a global scale. This may be because of the complex interplay of the other world regions with higher pollution levels, which can overshadow the local variations of pollutants in the United States.

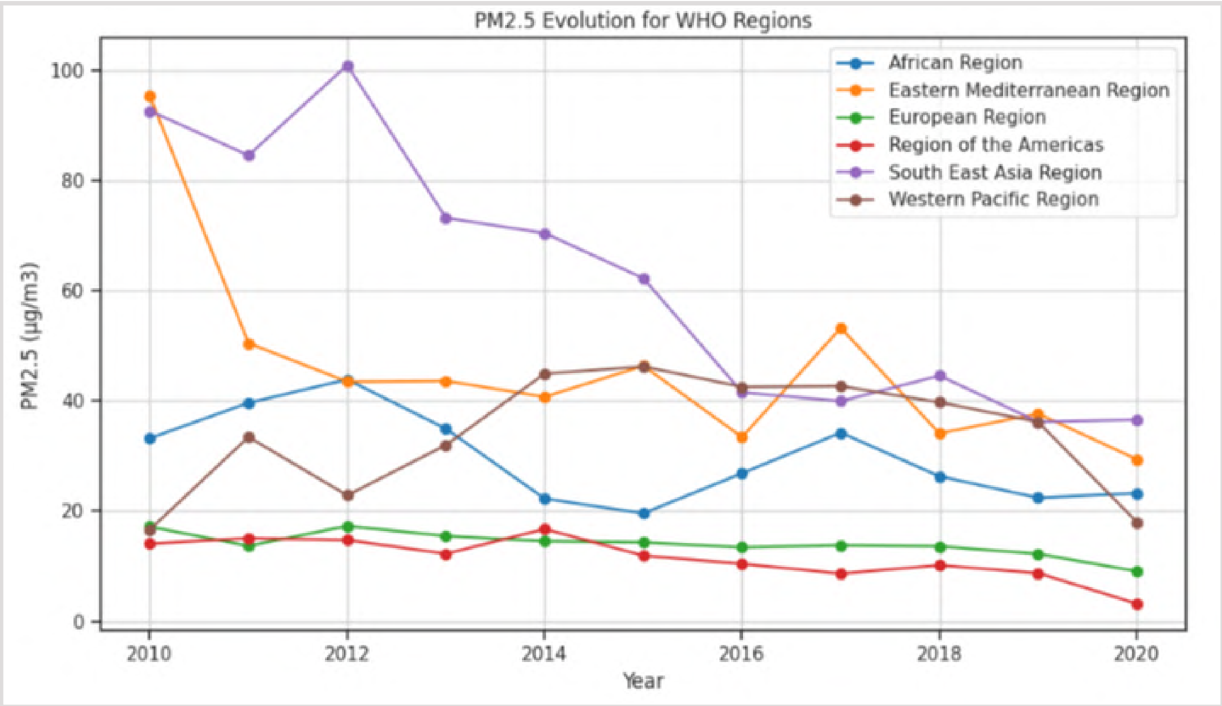



*Source: Barragan,R., (December, 2023) (own).*

# 4. EVOLUTION OF THE AIR POLLUTANTS.

Understanding the evolution and comparing the trends of each air pollutant over a decade is essential to assess its impact on public health and the environment. In this chapter, three graphs representing the concentrations of 'PM2.5 (µg/m3)', 'PM10 (µg/m3)' and 'NO2 (µg/m3)' in six different regions will be analyzed. These graphs will provide a detailed view of how these pollutants have evolved over time [See more visualization for each Region in Appendix #1].

## 4.1. PM2.5 (µg/m3)




*Source: Barragan,R., (December, 2023) (own).*

This graph represents the evolution of PM2.5 (µg/m³). The first thing that can be noticed is the wide range of values in each region in a comparative way. That being said, the following stands out:

1. South East Asia and Eastern Mediterranean Region had the highest PM2.5 levels, peaking around 2010. However, as we get closer to the present, these are the ones that show more fluctuation, going from a level close to 100(µg/m³) in 2010 to less than 40(µg/m³) in 2020. The primary sources of PM2.5 are energy, transport, industry, and windblown dust [4] due to the proximity of the desertic areas of the Arabian Peninsula and the Middle East.

2. The Africa and Western Pacific Region stand out at intermediate levels.

- For the African region, the major sources of PM2.5 are desert dust (due to the presence of large areas of arid land such as the Sahara Desert and the African savannas), landscape fires, and residential pollution (mainly caused by cooking and heating with biomass).
- Meanwhile, for the West Pacific Region, the major sources of PM2.5 are residential pollution, burning combustion, and industry.

3. Finally, the European Region and the Region of the Americas offer the lowest and most constant values, probably because they are the most developed regions and have the strictest agreements and implementation of environmental policies on the pollution levels to be complied with by each city.

- European Region: These values are caused by windblown dust, agriculture, energy production, and residential pollution.
- Region of the Americas: These values are caused by transport, landscape fires, and industry to a greater extent.

A common aspect all regions share is the violation of the WHO threshold of 5 µg/m3 for PM2.5 Another fact to consider is the decrease of PM2.5 (µg/m³) in all regions in the year 2020 with respect to the starting year (2010) due to the lockdown caused by the pandemic that reduced pollutant emissions by slowing down the economy, industry production and tourism.

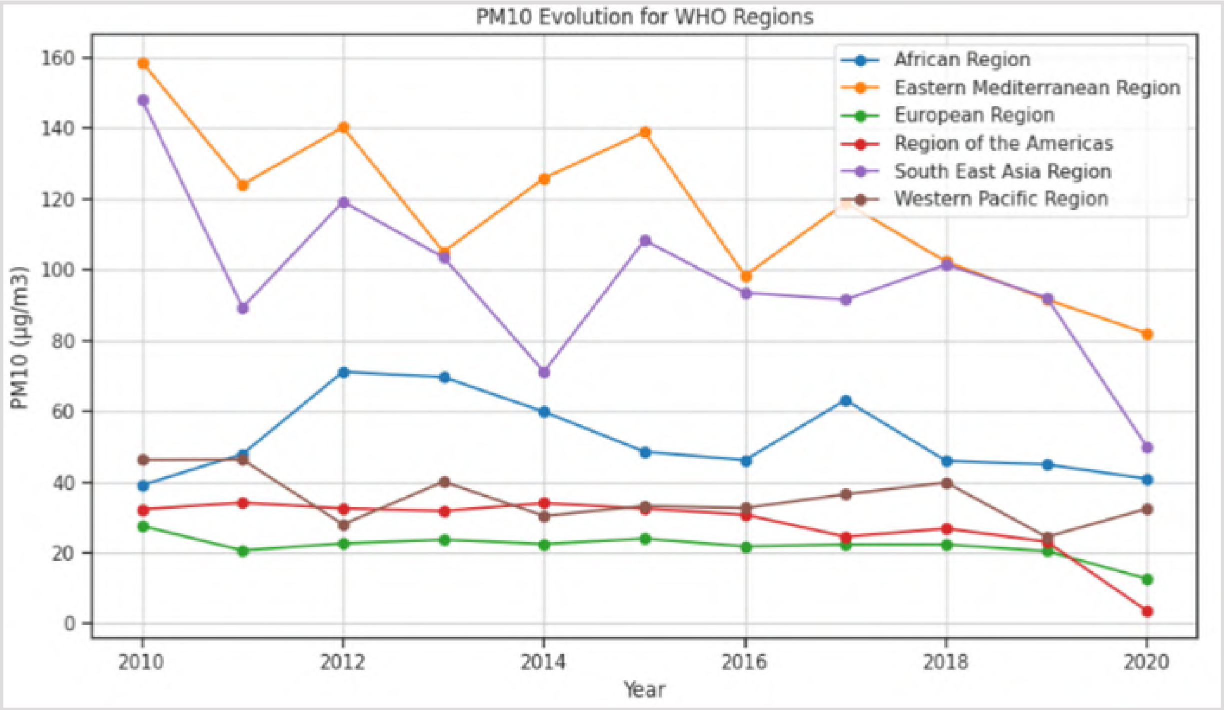
## 4.2. PM10 (µg/m3)

As mentioned in "Part I - 1. Introduction to the Context and Relevance of the Project", PM10 is another type of fine particulate matter (like PM2.5) suspended in the air with a diameter of less than 10 micrometers and which represents a health risk.

Since PM10 and PM2.5 share the same nature, PM10 will show a similar evolution in all regions. Although the values are higher than those of PM 2.5. This is because the WHO threshold is higher for PM10, being 15 µg/m³. Despite this, the levels still exceed this threshold.

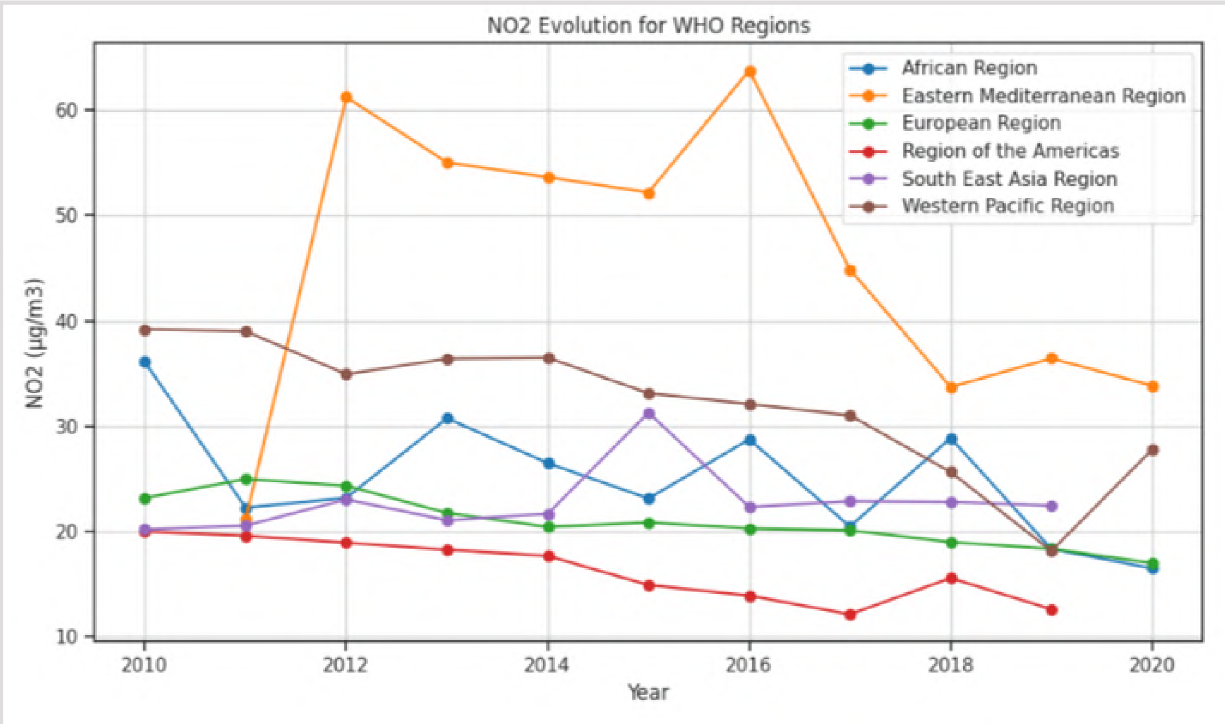
The African and Western Pacific regions continue to stand out for their high PM10 levels and a negative trend in their evolution (starting from initial values much higher than those of PM2.5), i.e., in 2010, levels were close to 160, and by 2020, the average was around 70.

However, there is a more obvious difference between this first group of regions and the remaining four. However, three of them stand out in particular: the Americas, Europe, and Western Pacific regions, as their levels remain much more stable over the 10 years at an average of less than 40 µg/m³.




*Source: Barragan,R., (December, 2023) (own).*

## 4.3. NO2 (µg/m3)

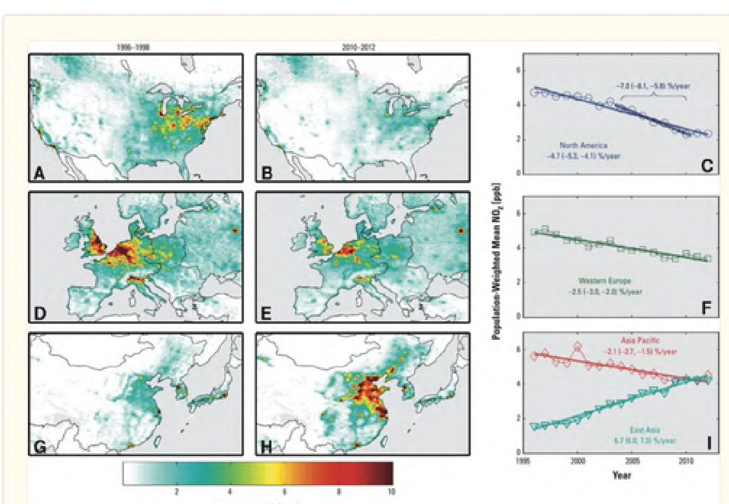


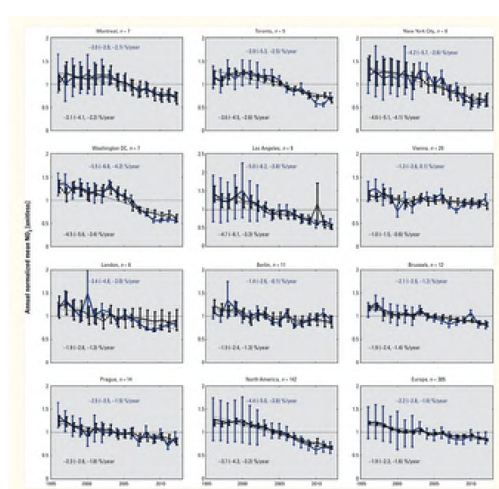

*Source: Barragan,R., (December, 2023) (own).*

Unlike the two previous cases, Nitrogen dioxide ($NO_2$) is not a fine particle, sino que belonging to the nitrogen oxides (NOx) group, comprising highly reactive gases. $NO_2$ primarily originates from the combustion of fuels in various sources, including motor vehicles and industrial activities.

Its impact on human and environmental health is significant, especially concerning cardiovascular and respiratory well-being, particularly asthma, and can trigger symptoms like coughing and shortness of breath. To safeguard human health, the World Health Organization (WHO) has set an annual average limit of 10 µg/m$^3$ for NO2 [2].

Globally, the concentration of $NO_2$ varies greatly, with many regions experiencing substantial declines in NO2 levels in recent years, while some countries still grapple with alarmingly high levels. To better understand this result, we will refer to the 2026 - research paper titled *"Long-Term Trends Worldwide in Ambient NO2 Concentrations Inferred from Satellite Observations"* [20].





*Sources: [20] Geddes, J. A., Martin, R. V., Boys, B. L., & van Donkelaar, A. (2016)*

Our findings reveal that $NO_2$ levels in the Region of the Americas show a decreasing pattern, starting around 21 µg/m³ in 2010 and ending around 11 µg/m³ in 2020, suggesting a decrease at a rate of approximately 10 µg/m³ over 10 years. Following this pattern, the levels of $NO_2$ in the Western Pacific Region decreased from around 39 µg/m³ in 2010 to about 29 µg/m³ in 2020, which would be a rate of 10 µg/m³ over 10 years.

Similarly, the European Eegion also witnessed a decrease in the levels of $NO_2$, starting at 24 µg/m³ in 2010 and dropping to about 17 µg/m³ in 2020. This indicates a decrease at a rate of about 7 µg/m³ over 10 years, similar to the Region of the Americas, likely influenced by stricter standards governing vehicle emissions.

The highest concentrations of $NO_2$ were observed in the Eastern Mediterranean Region, starting at 20 µg/m³ in 2010 and rising to about 65 µg/m³ in 2016, and finally decreasing to 35 µg/m³. Also, the Southeast Asia Region experienced a slight increase in nitrogen dioxide levels, possibly due to the region's increased urbanization over the same time period.

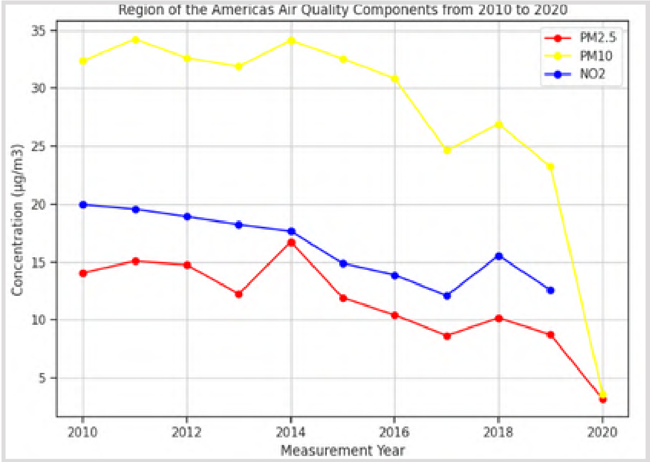
These fluctuations in $NO_2$ levels underscore the dynamic nature of this pollutant and its complex relationship with human activities and environmental factors.

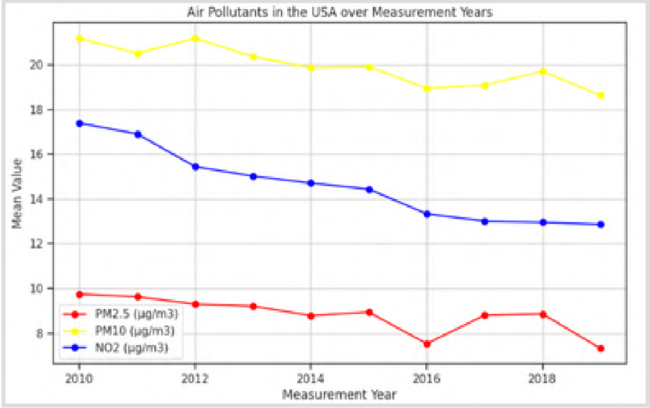
## 4.4. Region of the Americas vs. The United States of America

As observed, the United States is performing so much better than the average values in the Region of the Americas. In the second plot, the air pollutant levels in the United States have consistently declined from 2010 to 2020, following the WHO thresholds - especially for PM2.5 (5 µg/m³) and $NO_2$ (13 µg/m³).

These reductions can be attributed to environmental regulations to reduce emissions from transportation combustion, energy, and industry.

- U.S.-Canada Air Quality Agreement (1991) [21]: A pact between the U.S. and Canada to combat cross-border air pollution, addressing issues like acid rain and transboundary smog.
- UN ESCAP Regional Action Programme on Air Pollution (2022) [22]: Focuses on improving air quality, monitoring, data sharing, capacity building, and international cooperation.
- UNECE Convention on Long-Range Transboundary Air Pollution (CLRTAP) [23]: An international treaty reducing transboundary air pollution, with over 90% sulfur dioxide emission reduction since 1979.
- Arctic Council (1996) [24]: Promotes cooperation among Arctic States and aims to reduce black carbon emissions by 25-33% from 2013 levels by 2025.




*Source: Barragan,R., (December, 2023) (own).*

## 4..4.1 The U.S. states with the highest air pollutants levels

This reseatch takes a step further, and identifies some patterns and similarities regarding the demographic and geographic characteristics among the states with the highest levels of 'PM2.5 (µg/m3), 'PM10 (µg/m3), and 'NO2 (µg/m3):

**Demographic Similarities:**

Industrial Activity: It appears that some of the states with the highest pollutant levels (e.g., IL, OH, MI, IN) have a notable industrial presence. High industrial activity can contribute to air pollution due to emissions from factories and manufacturing processes.

Population Density: Some of the states (e.g., PA, IL, MD) have relatively high population densities, which can lead to increased vehicle traffic and urban pollution sources, contributing to higher pollutant levels.

Urban Areas: Many of the states with high pollution levels have major urban centers or cities known for industrial or transportation hubs. Urban areas tend to have higher pollution levels due to increased vehicular traffic and industrial activity.

**Geographic Similarities:**

Proximity to Water Bodies: States like PA and MD have coastlines along the Atlantic Ocean, and IL has access to the Great Lakes. Proximity to water bodies can influence air quality, as it can affect weather patterns and dispersion of pollutants.

Midwestern and Eastern U.S.: Several of the states (IL, OH, MI, IN, PA, MD) are located in the Midwest and eastern regions of the United States. These areas may share common weather patterns and sources of pollution.

### HIGHEST

| State | Measurement Year | PM2.5 (µg/m3) |
|---|---|---|
| PA | 2010 | 16.00 |
| | 2011 | 13.85 |
| | 2012 | 14.25 |
| OR | 2013 | 13.60 |
| PA | 2014 | 13.40 |
| OH | 2015 | 12.10 |
| PA | 2016 | 11.80 |
| OR | 2017 | 13.28 |
| | 2018 | 15.40 |
| MI | 2019 | 11.00 |

| State | Measurement Year | PM10 (µg/m3) |
|---|---|---|
| IL | 2010 | 42.90 |
| | 2011 | 41.80 |
| | 2012 | 45.15 |
| SD | 2013 | 37.70 |
| IL | 2014 | 42.90 |
| MD | 2015 | 37.40 |
| | 2016 | 39.00 |
| NM | 2017 | 32.00 |
| | 2018 | 34.20 |
| IN | 2019 | 33.15 |

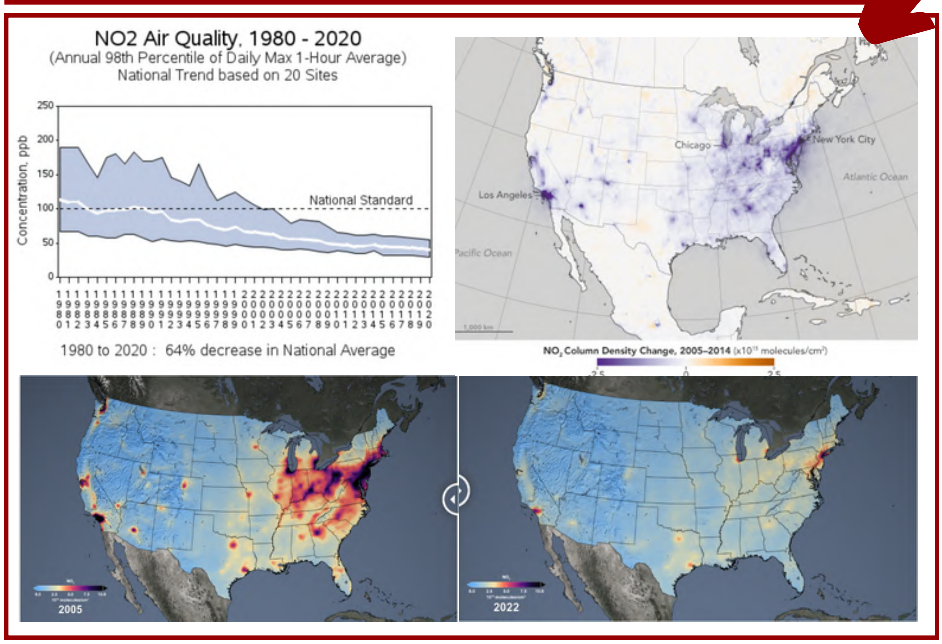
| State | Measurement Year | NO2 (µg/m3) |
|---|---|---|
| [illegible] | 2010 | 29.70 |
| NV | 2011 | 28.05 |
| AZ | 2012 | 27.56 |
| UT | 2013 | 25.87 |
| NH | 2014 | 30.95 |
| MI | 2015 | 27.65 |
| NV | 2016 | 30.20 |
| | 2017 | 31.70 |
| CT | 2018 | 23.10 |
| [illegible] | 2019 | 21.98 |

### LOWEST

| State | Measurement Year | PM2.5 (µg/m3) |
|---|---|---|
| NV | 2010 | 3.40 |
| MT | 2011 | 3.80 |
| MI | 2012 | 4.10 |
| SD | 2013 | 3.10 |
| | 2014 | 2.40 |
| VT | 2015 | 3.50 |
| SD | 2016 | 1.90 |
| HI | 2017 | 3.23 |
| | 2018 | 3.45 |
| NV | 2019 | 3.00 |

| State | Measurement Year | PM10 (µg/m3) |
|---|---|---|
| VT | 2010 | 7.40 |
| ND | 2011 | 7.60 |
| VT | 2012 | 6.80 |
| | 2013 | 6.60 |
| ND | 2014 | 5.90 |
| | 2015 | 5.90 |
| | 2016 | 3.80 |
| VT | 2017 | 5.90 |
| MA | 2018 | 5.10 |
| VT | 2019 | 4.90 |

| State | Measurement Year | NO2 (µg/m3) |
|---|---|---|
| ND | 2010 | 2.20 |
| | 2011 | 3.92 |
| MT | 2012 | 3.10 |
| | 2013 | 2.55 |
| | 2014 | 2.50 |
| | 2015 | 1.70 |
| WY | 2016 | 1.60 |
| MT | 2017 | 1.30 |
| | 2018 | 1.70 |
| WY | 2019 | 1.40 |




*Sources for Table Barragan,R., December, 2023)*

*For the four images: [25]Bishop, S. (2021) [26] NASA. (2014) [27] NASA. (2022)*

# PART III
# CHRONIC RESPIRATORY DISEASE ANALYSIS: US APPROACH

1. DATASETS OVERVIEW
    1.1. DESCRIPTION OF THE DATASET: SOURCES
    1.2. DESCRIPTION OF THE DATASET: VARIABLES
    1.3. NAS, OUTLIERS AND DISTRIBUTION

2. DESCRIPTIVE ANALYTICS
    2.1. GENDER
    2.2. AGE

# 1. DATASET OVERVIEW

## 1.1. DESCRIPTION OF THE DATASET: SOURCES

The Chronic Respiratory Diseases (CRD) dataset is provided by the Institute for Health Metric and Evaluation and its study - Global Burden of Disease (GBD) [10]. The GBD research offers a thorough overview of numerous diseases, injuries, and risks and their impact on mortality and disability rates across nations over time (considering aspects like age and gender) to improve the healthcare system.

In this project, the CRD dataset collects the number of deaths attributed to chronic respiratory diseases in five age groups (<20 years; 20-54 years; 55-89 years; 90-94 years; >95 years) and for both genders (female and male) across the 50 states.

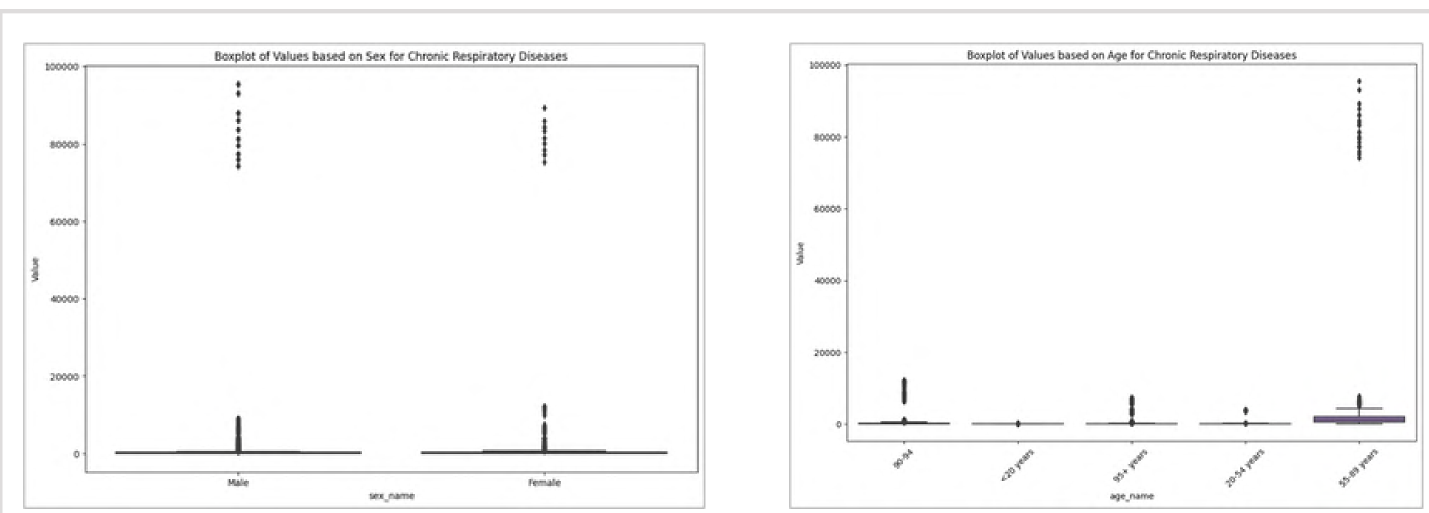
## 1.2. DESCRIPTION OF THE DATASET: VARIABLES

| *Column name* | *Description* |
|---|---|
| Chronic Respiratory Diseases Dataset | |
| **Measure Name** | The categorical variable indicates the type of medical incident. For the purpose of this project, the dataset collects the medical incidence collected as "deaths". |
| **Location_Name** | Full name of the US state where the death occurred, collecting records from all 50 states. |
| **Sex_Name** | Binary variable that segments the dataset into two groups: women and men. |
| **Age_Name** | A categorical variable that classifies chronic respiratory disease data according to five age groups (<20 years; 20-54 years; 55-89 years; 90-94 years; >95 years). |
| **Cause_Name** | Categorical variable indicating the type of illness. For the purpose of this project, the dataset collects the medical disease,s collected as "Chronic Respiratory Diseases". |
| **Measurement Year** | Year of the annual mean concentration. The dataset collects data from 2010 to 2019. |
| **Value** | Numerical variable that indicates the number of registered deaths caused by Chronic Respiratory Diseases |
| **State** | A new column was created in the dataset, indicating the abbreviation of the 50 US states. |

*Source: Barragan,R., (December, 2023) (own).*

## 1.3 NAS, OUTLIERS AND DISTRIBUTIONS

No imputation is needed since there are no null values in the dataset. Furthermore, box plots (sex_name and age_name) show that the data presents a great deal of outliers. However, instead of handling or eliminating them, we will keep them. The reason is that removing this considerable amount of data will reduce the data variability, potentially losing a lot of relevant information and accuracy in this way.




*Source: Barragan,R., (December, 2023) (own).*

# 2. DESCRIPTIVE ANALYTICS

The dataset collects a total of 5200 entries for chronic respiratory diseases in the United States from 2010 to 2019. This is a very complex yet comprehensive dataset, which is worth start exploring from a double perspective: gender and age.

## 2.1. GENDER

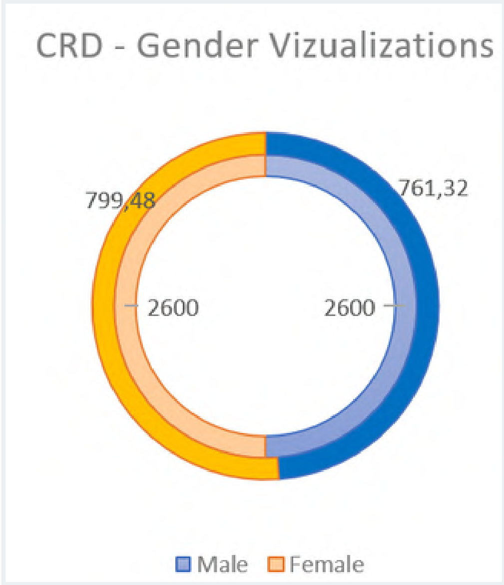
From the point of view of gender, the dataset is split 50-50 between women and men. In other words, it includes 2600 female entries (50%) and 2600 male entries (50%)

In order for the results of this study to apply to society, the sample must be representative of that population. For this, we will reference the Research called "Total population of the United States by gender 2010-2027," performed by the Statista Research Department in 2023 [19]. The study confirms that since 2013, the gender distribution in the United States has remained stable, with women comprising about 51.1 percent of the population.

Bias is a natural part of collecting and analyzing data. Even when data scientists try to ensure that the data represents the real world accurately, there can still be a small chance of bias.

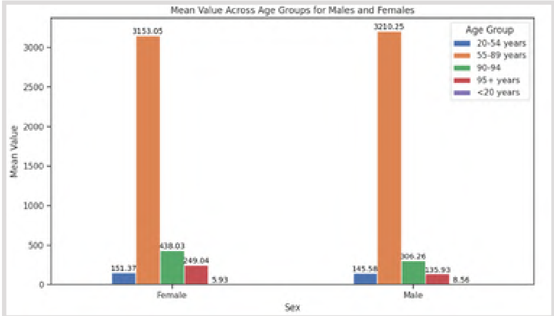
In our dataset, we did find a slight bias, but it doesn't compromise the trustworthiness of our study.

Having clarified this aspect, the dataset shows that women have a higher risk of suffering from respiratory diseases since the number of deaths averages 799.5, with respect to the average for men, which is equivalent to 761.3 deaths. This results in a total mean of 780.399976 and a standard deviation of 5230.36.



Source: Barragan,R., (December, 2023) (own).

Another similarity found in these two groups is the close average number of deaths for each age group, i.e., where there is almost no difference between the number of deaths for men and women.



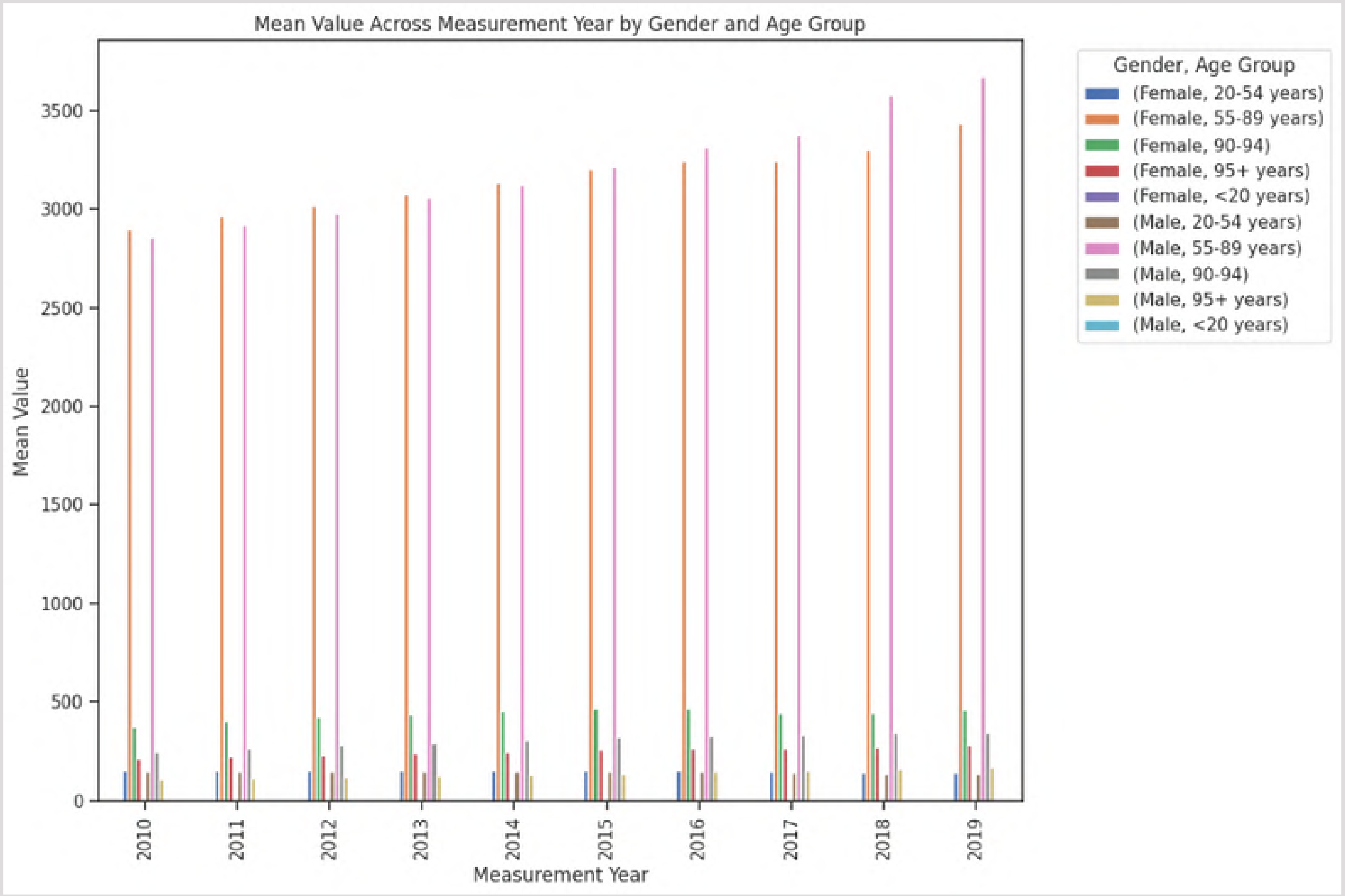
Source: Barragan,R., (December, 2023) (own).

## 2.2. AGE

When taking the Age variable, there are three main aspects to highlight from the analysis.

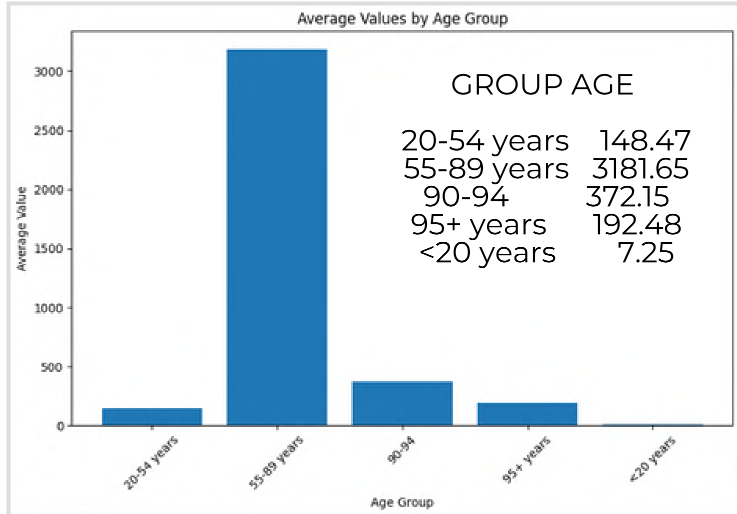
- Upward time trend: the first thing to be observed in the histogram is an upward trend in the mean value over the years for all groups. This could indicate a worsening of the measured health condition or an increase in the incidence of agents harmful to health (such as air pollution), causing an increase in mortality over the years, with the most worrisome year being 2019.

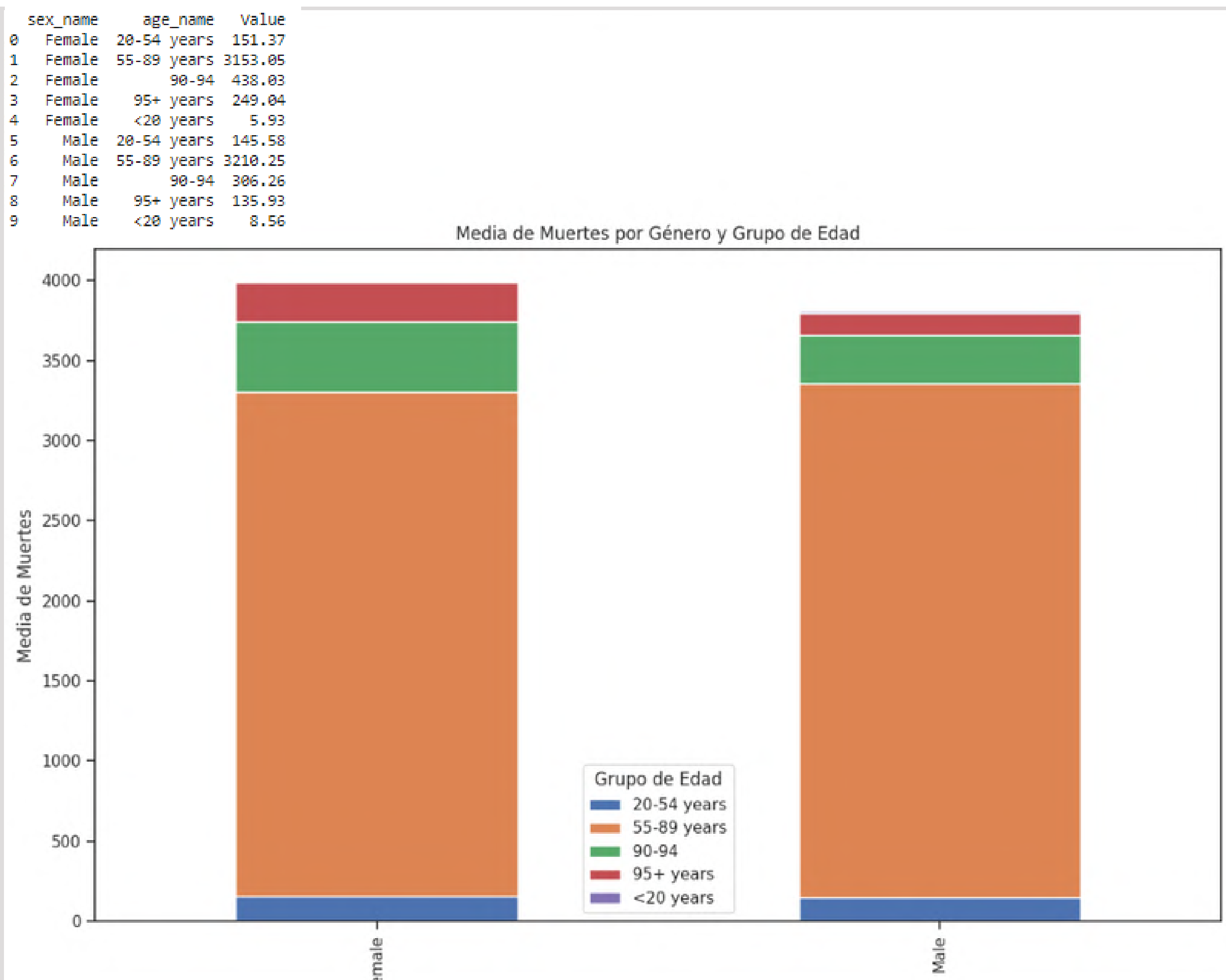
- The second aspect that stands out from the histogram is the group with the highest number of deaths, this being the age range of 55 - 89 years, both for females and males in all 10 years. This may be due to
    - an acceleration in the aging of the human body that occurs in middle age (e.g., menopause).
    - Increased susceptibility to developing severe or chronic diseases.
    - Worsening lifestyle and increased sedentary lifestyle.
    - Oxidation of cells due to increased exposure to harmful pollutants and ultraviolet rays that accelerate organ degradation.



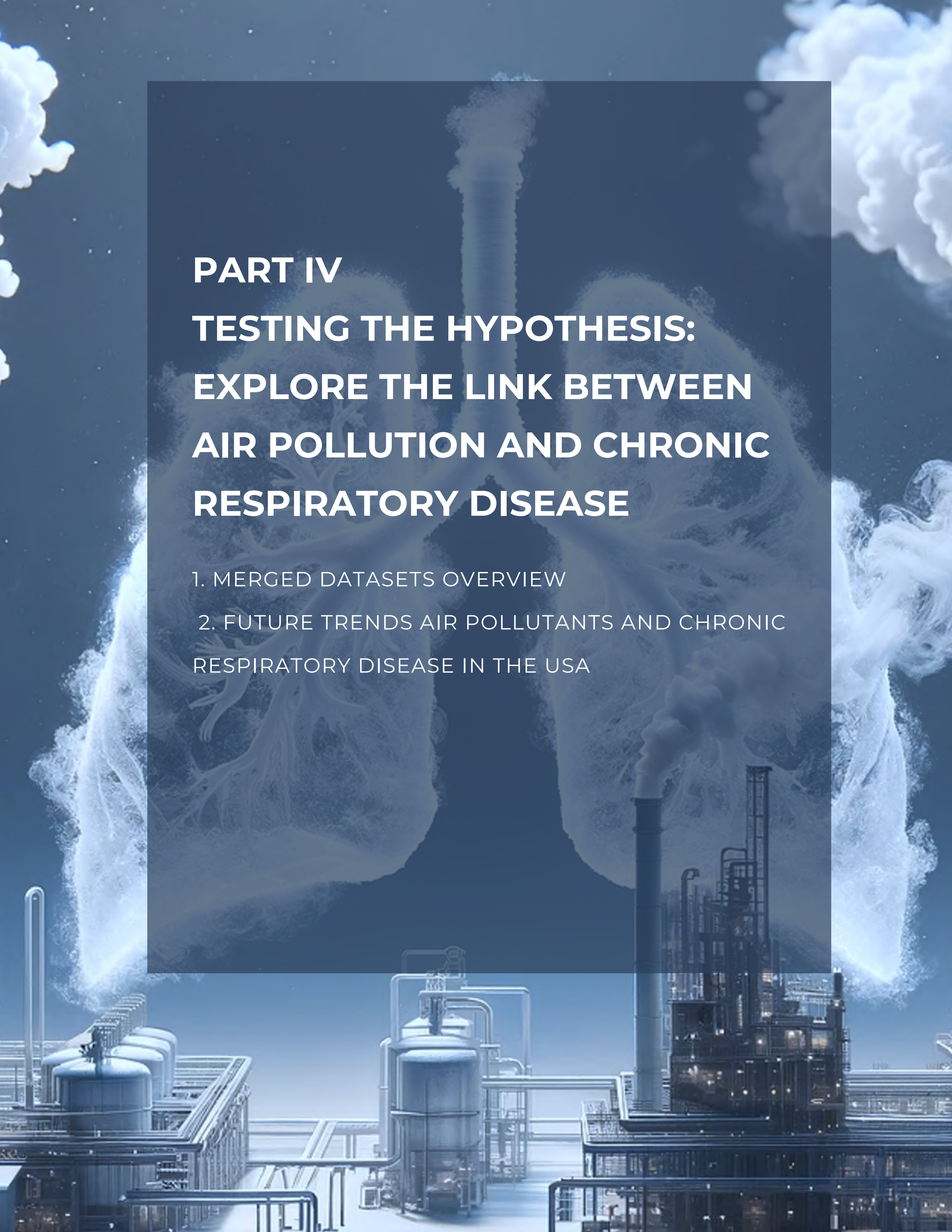

*Source: Barragan,R., (December, 2023) (own).*

- Finally, and linking it with the last idea, the existence of similarity in patterns for each age group: From 2010 to 2019, there is a pattern in the number of deaths in each age group that is also apparent in the gender ranking:
-
    - 1: 55-89 years (male and female)
    - 2: 90-94 years (male and female)
    - 3: >95 years (male and female)
    - 4: 20-54 years (male and female)
    - 5: < 20 (male and female)




*Source: Barragan,R., (December, 2023) (own).*



# PART IV
# TESTING THE HYPOTHESIS: EXPLORE THE LINK BETWEEN AIR POLLUTION AND CHRONIC RESPIRATORY DISEASE

1. MERGED DATASETS OVERVIEW
2. FUTURE TRENDS AIR POLLUTANTS AND CHRONIC RESPIRATORY DISEASE IN THE USA

# 1. MERGED DATASET OVERVIEW

After thoroughly analyzing each dataset, the Air Quality Dataset and Chronic Respiratory disease dataset, having solved the limitations and matched the dataset terminology, we can proceed to unify both datasets.

The new combined US-focused dataset has 17670 rows and 15 columns. It will now be possible to understand better the relevance of the effect of pollutants in the air on the number of deaths from respiratory diseases, as well as to visualize future trend patterns up to 2025.

## 1.2. CORRELATION MATRIX AND HEAT MAP

The five variables to be highlighted are Measurement Year, PM2.5, PM10, NO2, and Value to see the correlation between them.

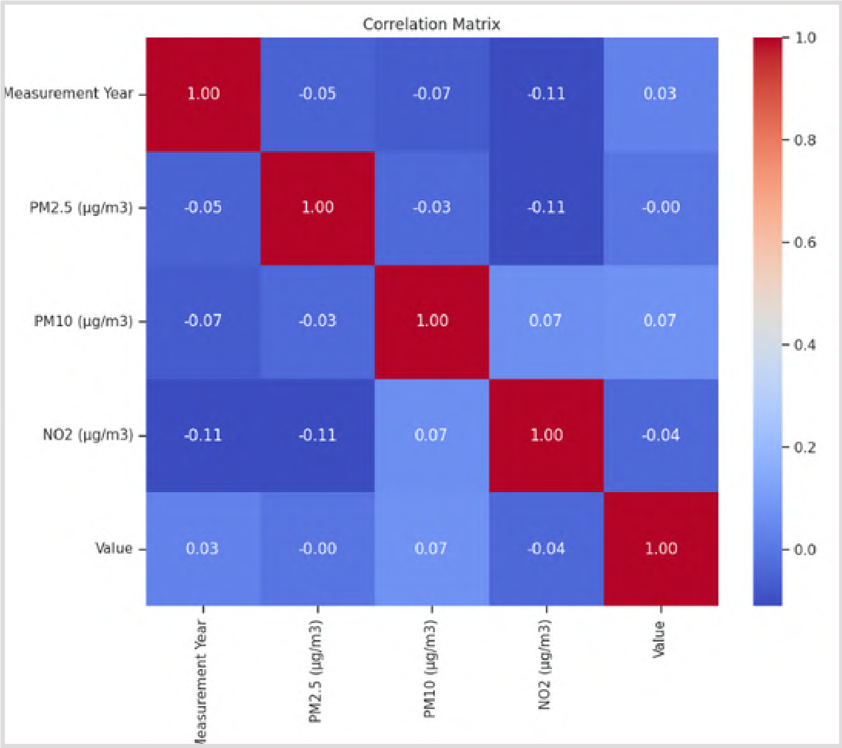
Measurement Year does not have a strong correlation with the other variables PM2.5 (-0.05), PM10 (-0-07), and NO2 (-0.11). This indicates that as the years go by, air pollution levels decrease, probably due to implementing new measures and policies and Agenda 2030 goals.

PM2.5 and PM10 have a negative correlation, indicated by the correlation coefficient equal to -0.03. However, this correlation is very weak, suggesting that when one fine particle decreases, so does the other.

PM10 and Value have a positive correlation coefficient (0.07); although the relationship is weak, it is positive, suggesting that when the levels of PM10 increase, so does the number of deaths.
The same happens with PM10 and NO2, with a correlation coefficient of 0.07. This means that there is a weak yet positive linear relationship.

NO2 also has a weak negative correlation with "Value" (-0.04), suggesting that there is not a strong linear relationship between NO2 concentration and the number of deaths from chronic respiratory diseases.

*Merged dataset (Air Quality & Chronic Diseases)*




*Source: Barragan,R., (December, 2023) (own).*

## 1.2. FUTURE TRENDS AIR POLLUTANTS AND CRD IN THE USA

The visualizations present the historical analysis as well as future projections on the evolution of air pollutants and the number of deaths caused by chronic respiratory diseases from 2010 to 2025.

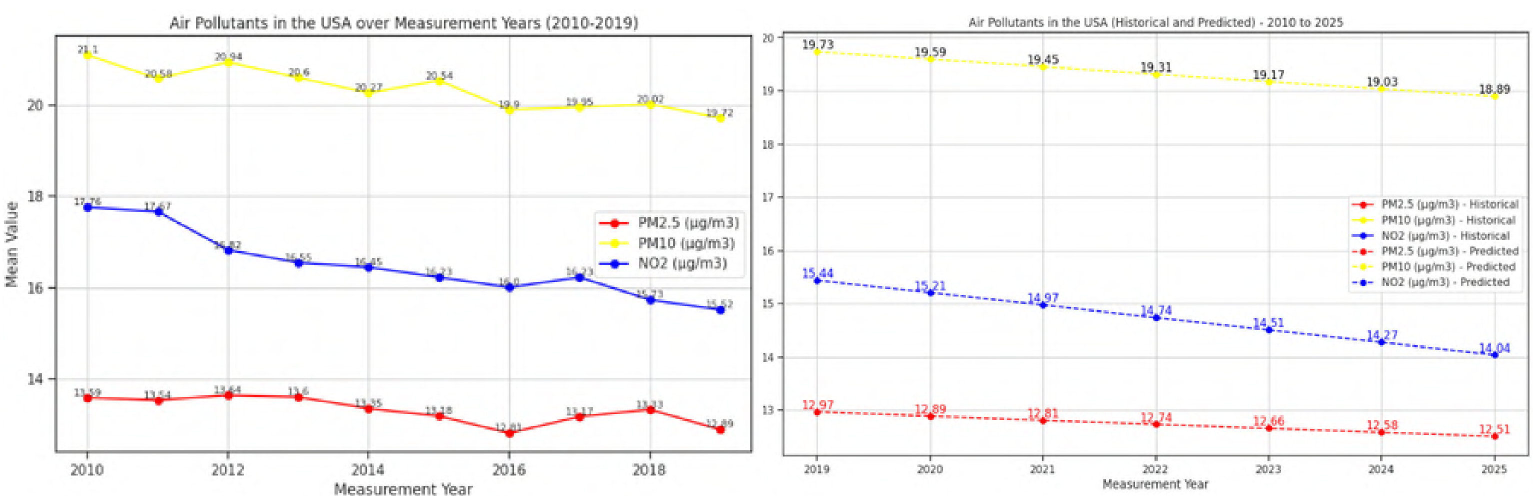
On the one hand, for the projection of pollutant gases, we see a progressive decline in the level of each of the three pollutants. PM10 and PM2.5 (two fine particles) show a prolonged and similar rate of decline.

Meanwhile, for PM10, the decline trend is more pronounced. This may be because as the main sources of PM10 are motor vehicles and industrial activities, in the future, more sustainable alternatives such as electric vehicles or recycling or upcycling of materials may become more democratized.

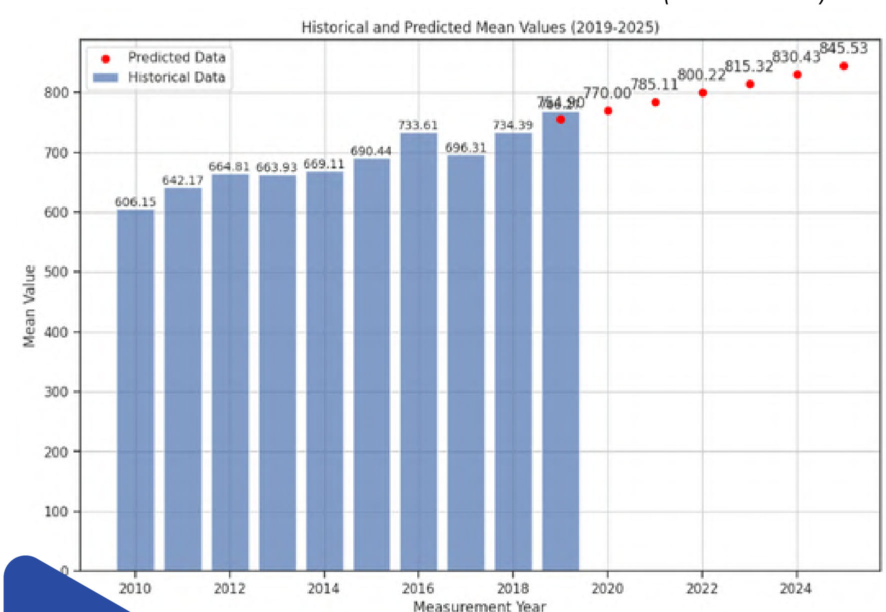
On the other hand, the third graph shows a clear increase in the number of deaths up to 2025. This could reflect several factors, such as an aging population, an increase in the prevalence of chronic respiratory diseases, possible changes in air quality, or even the impact of factors such as smoking and other environmental exposures.

Future projections indicate that the upward trend is expected to continue, increasing significantly year after year, reaching a peak in 2025.

*Historical and Predicted Air Pollutants in the USA (2010 - 2025)*




*Historical and Predicted CRD in the USA (2010 - 2025)*




*Source: Barragan,R., (December, 2023) (own).*

# PART V
# CONCLUSIONS AND RECOMMENDATIONS

1. CONCLUSIONS
2. RECOMMENDATIONS

# 1. CONCLUSIONS

The urgency of addressing air pollution is evident in the current context, which is becoming a global health crisis with millions of premature deaths annually attributable to poor air quality. This scenario underlines the imperative need for concrete and effective actions.

The present study has provided a detailed correlation analysis between air pollutants and mortality. In particular, fine particles, such as **PM2.5 and PM10**, capable of penetrating deep into the lungs and bloodstream, have shown a variable distribution globally.

The **Southeast Asia and Eastern Mediterranean Region** had the highest, though fluctuating, levels, while the **regions of the Americas and Europe** maintained the lowest and most constant values, possibly due to stricter environmental policies. This variability underscores the complexity of managing air quality in different geographic and development contexts.

Focusing on the United States, this world power is making noteworthy efforts to keep pollution levels under control, in line with the thresholds set by the World Health Organization.

Although each pollutant has been thoroughly studied, it is essential to mention that the **most relevant pollutant** that stands out for its positive correlation (0.07) with the number of deaths in the **United States is nitrogen dioxide** ($NO_2$).

This coefficient indicates that an **increase in $NO_2$ levels may be linked to an increase in the number of deaths**. This finding is particularly relevant for formulating public health policies, as it highlights the importance of **monitoring and controlling $NO_2$ levels (and their sources** - mainly motor vehicles and industrial activity) to protect the population's health.

Another **aspect to highlight** in the conclusions is age. Age emerges as a critical factor, especially with the **55-89 age group** registering the **highest number of deaths**. This pattern points to the need to pay special attention to older populations, which may be more vulnerable due to natural aging, the prevalence of chronic diseases, and cumulative exposure to environmental risk factors.

In conclusion, the results emphasize the importance of comprehensive strategies that encompass the reduction of pollutants - especially $NO_2$ - together with the implementation of health policies targeting the most at-risk populations. Only through a multifaceted approach will we be able to mitigate the growing public health crisis arising from air pollution and move towards meeting the Sustainable Development Goals related to environmental health and well-being.

# 2. RECOMMENDATIONS

Based on the insights and conclusions of this study, the key recommendations that further support addressing the problems related to air pollution and its impact on public health are as follows:

**Strengthen environmental legislation:**

Update and strengthen air quality regulations to align them with stricter WHO guidelines, setting new, lower emission limits especially for NO2.

**Improve air quality monitoring**:

Invest in a more extensive and technologically advanced monitoring network to obtain accurate, real-time data on air pollution levels and trace their source of origin.

**Provide more personalized health services**

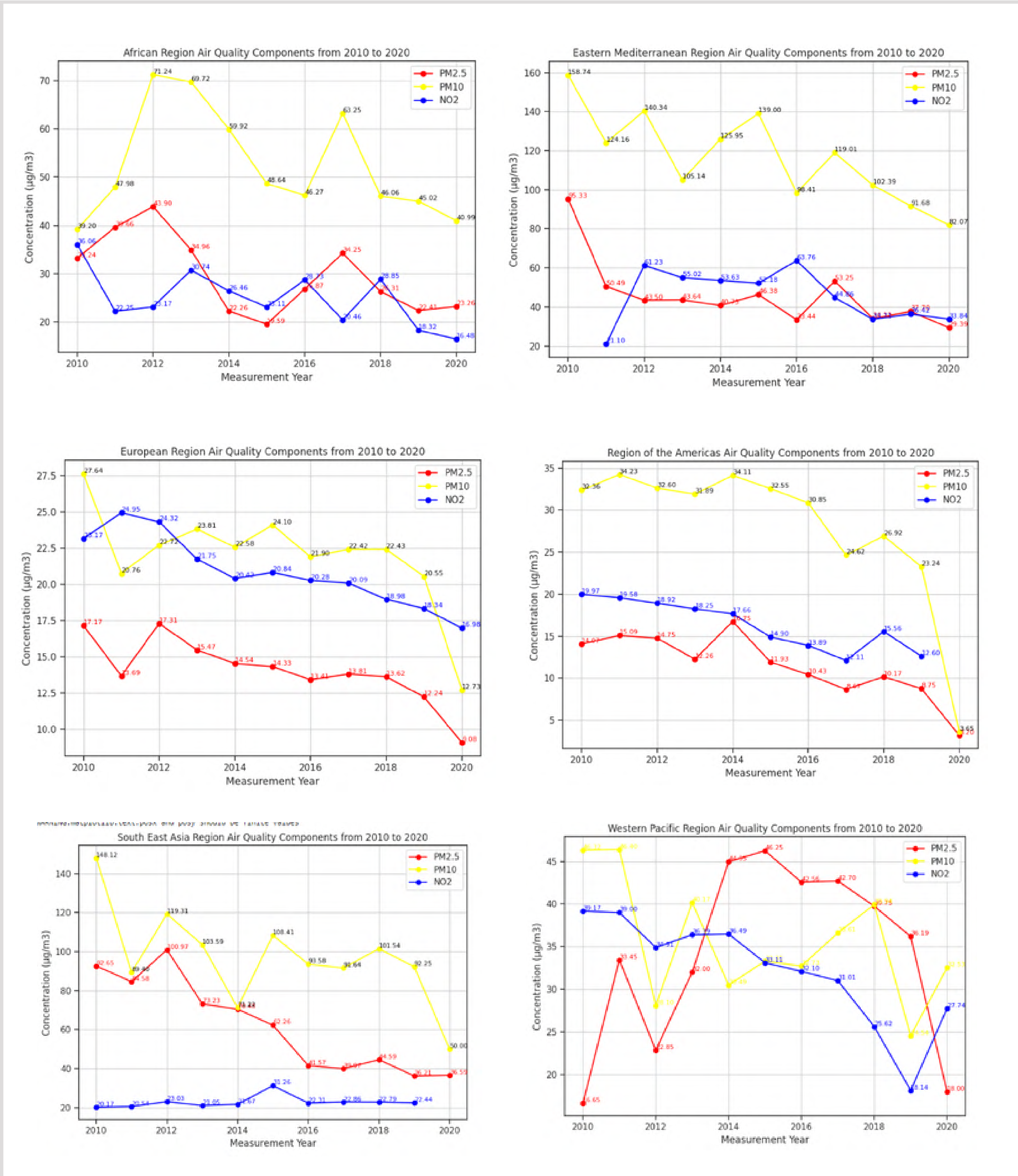
to the most vulnerable age groups (55 - 89 years), implementing greater follow-up and preventive and palliative treatments focused on improving respiratory well-being and preventing the contraction of respiratory diseases.

**Promote clean energy:**

Given that the most relevant pollutant gas is NO2, and its sources of origin are mainly motor vehicles and industrial activity; the use of renewable and clean energy sources should be promoted to reduce NO2 emissions derived from fossil fuel combustion and industrial activities. Linked to this point, encourage public transportation, the use of electric and non-motorized vehicles, and develop infrastructures that support their use in order to reduce pollution generated by vehicles.

# APPENDIX

**Appendix #1** (mentioned in Part II.4 - Evolution of the Air Pollutants. Page 20)




*Source: Barragan,R., (December, 2023) (own).*

# REFERENCES


[1] World Health Organization. (n.d.). Who we are. World Health Organization. https://www.who.int/about/who-we-are/

[2] WHO. (2021, September). WHO Air Quality Guidelines. C40 Knowledge Community. https://www.c40knowledgehub.org/s/article/WHO-Air-Quality-Guidelines?language=en_US#:~:text=By%20reducing%20air%20pollution%20levels,3%20%2D%204%20days%20per%20year.

[3] Environmental Protection Agency. (2023, June 26). Conceptos básicos sobre el material particulado . EPA. https://espanol.epa.gov/espanol/conceptos-basicos-sobre-el-material-particulado-pm-por-sus-siglas-en-ingles

[4] UNEP. (2021, September 7). Air pollution note – data you need to know. United Nations Environment Program. https://www.unep.org/interactives/air-pollution-note/

[5] Department of Economic and Social Affairs. (n.d.). The 17 goals | sustainable development. United Nations. https://sdgs.un.org/goals

[6] UNICEF. (2023, September 20). SDG goal 3: Good health and well-being. UNICEF DATA. https://data.unicef.org/sdgs/goal-3-good-health-wellbeing/

[7] Department of Economic and Social Affairs. (2022). Goal 3 | Department of Economic and Social Affairs. United Nations. https://sdgs.un.org/goals/goal3#targets_and_indicators

[8] World Health Organization. (2023, May 15). WHO Ambient Air quality database. World Health Organization. https://www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database

[9] WHO. (2022, April 4). Billions of people still breathe unhealthy air: New WHO data. World Health Organization. https://www.who.int/news/item/04-04-2022-billions-of-people-still-breathe-unhealthy-air-new-who-data#:~:text=A%20record%20number%20of%20over,countries%20suffering%20the%20highest%20exposures.

[10] Global Burden of Disease (GBD). (2019). GBD results. Institute for Health Metrics and Evaluation. https://vizhub.healthdata.org/gbd-results/

# REFERENCES


[11] World Health Organization. (n.d.). Global Air Pollution and health - technical advisory group. World Health Organization. https://www.who.int/groups/global-air-pollution-and-health---technical-advisory-group/about

[12] World Health Organization. (n.d.). Scientific Advisory Group on air pollution and health. World Health Organization. https://www.who.int/groups/sdg-11-6-2-working-group/about/scientific-advisory-group-on-air-pollution-and-health#:~:text=The%20SAG%20provides%20expert%20guidance,addressing%20air%20pollution%20and%20health.

[13]World Health Organization. (n.d.). Air quality and health - About Us. World Health Organization. https://www.who.int/teams/environment-climate-change-and-health/air-quality-and-health/about

[14] The United States Government. (2022). OSTP's teams. The White House. https://www.whitehouse.gov/ostp/ostps-teams/

[15] The United States Government. (2022). Health outcomes. The White House. https://www.whitehouse.gov/ostp/ostps-teams/health-outcomes/

[16] The United States Government. (2022). Climate and environment. The White House. https://www.whitehouse.gov/ostp/ostps-teams/climate-and-environment/

[17] Open Development Myanmar . (2020, January 6). SDG 3 Good health and well-being - open development myanmar. Open Development Myanmar . https://opendevelopmentmyanmar.net/topics/sdg-3-good-health-and-well-being/

[18] WHO. (2022). (rep.). WHO ambient air quality database, 2022 update: status report (Vol. ISBN: 978-92-4-004769-3, pp. 1–34). https://www.who.int/publications/i/item/9789240047693

[19] Statista Research Department. (2023, October 6). U.S. population by gender 2027. Statista. https://www.statista.com/statistics/737923/us-population-by-gender/#:~:text=Projection%20estimates%20calculated%20using%20the,US%20Census%20data%20for%202021

[20] Geddes, J. A., Martin, R. V., Boys, B. L., & van Donkelaar, A. (2016). Long-term trends worldwide in ambient no$_2$ concentrations inferred from satellite observations. Environmental Health Perspectives, 124(3), 281–289. https://doi.org/10.1289/ehp.1409567

# REFERENCES


[21]U.S. Environmental Protection Agency | US EPA. (2020). https://www.epa.gov/system/files/documents/2023-03/22119.02%20CAN-US%20Air%20Quality%20Agreement-EN_V08_508_3012023_508pass.pdf

[22]Operationalizing the Regional Action Programme on air pollution: Inception meetings of the open-ended working groups on air quality standards and Data. ESCAP. (2023). https://www.unescap.org/events/2023/operationalizing-regional-action-programme-air-pollution-inception-meetings-open-ended

[23]1979 Convention on Long Range Transboundary Air Pollution - UNECE. (1979). https://unece.org/sites/default/files/2021-05/1979%20CLRTAP.e.pdf

[24]About the Arctic Council. Arctic Council. (1996). https://arctic-council.org/about/#:~:text=The%20Arctic%20Council%20was%20established,States%20signed%20the%20Ottawa%20Declaration.

[25]Bishop, S. (2021, September 27). Air Quality Measurements Series: Nitrogen Dioxide. Clarity. https://www.clarity.io/blog/air-quality-measurements-series-nitrogen-dioxide

[26] NASA. (2014). A decade of change for Nitrogen Dioxide. Earthe Observatory. https://earthobservatory.nasa.gov/images/87392/a-decade-of-change-for-nitrogen-dioxide

[27] NASA. (2022). Nitrogen Dioxide. NASA. https://airquality.gsfc.nasa.gov/no2

Penn Social Policy & Practice
UNIVERSITY of PENNSYLVANIA

Rose Barragan Barranco

Master of Science in Social Policy and Data Science

rosabb@upenn.edu

https://www.linkedin.com/in/rose-barragan/