Final Project Abstract

Cara Kuuskvere, Jenn McNeill, and Emma Kaufman

2024-02-26

library(readr)

## 1. State your research question(s) and provide 1-2 paragraphs of background information needed to understand the question or issues.

We are interested in understanding the relationship between different air pollutant concentrations and air quality in North Carolina in 2018 and 2019. We are specifically interested in concentrations of ozone and PM2.5. We plan to examine the following questions in our report:

* Are Ozone and PM2.5 concentrations correlated?
* Is Mean Air Quality Index (AQI) correlated with Ozone and PM2.5?
* How correlated are PM2.5 and Ozone concentrations with longitude across North Carolina?

Ozone and PM2.5 are two of six major air pollutants with standards set by the EPA to protect human health and welfare [1]. When concentrations of these pollutants are too high in ambient air, they can cause serious health and environmental problems [1]. Some common sources of these pollutants are coal-fired power plants, cars, and industrial manufacturers. We are interested in examining these pollutant concentrations throughout North Carolina to see if a disproportionate region of the state is at higher risk for health conditions due to air pollutants PM2.5 and Ozone. We are also interested in whether these two pollutant concentrations are correlated.

## 2. Explain why this question is relevant and interesting to env. Science, management, or policy in 1 paragraph

High levels of air pollutants are associated with a range of health problems, including respiratory diseases like asthma, chronic obstructive pulmonary disease (COPD), and lung cancer, as well as cardiovascular issues like heart attacks and strokes [2]. By studying air pollutant concentrations through statistical analysis, public health officials and policymakers can develop targeted interventions to reduce exposure and mitigate health risks. Understanding the relationship between air quality and public health allows for the implementation of policies and strategies aimed at protecting vulnerable populations and promoting equitable public health management solutions. Our research question is relevant because it will provide clarity on which regions of North Carolina have been more affected than others and which air pollutants are the most threatening to the state.

## 3. Describe your data, including the dependent variable(s) and independent variables in the dataset. What is your unit of analysis (e.g., cities, nations, corporations, individuals, plots of land, rivers, etc.)? What type of data are your dependent variables (categorical, nominal, continuous)? List the data source (website, faculty project, etc.).

Our data comes from the EPA. It is a data set that provides daily data at 10 different monitoring sites in 9 different counties across North Carolina for a time period of two years. The latitude and longitude of each site is also provided. Each day, each monitoring site has a measurement for PM2.5, Ozone, and AQI. All data are continuous and independent of one another with the exception of AQI. The AQI is calculated by converting measured pollutant concentrations to a uniform index which is based on the health effects associated with a pollutant. The health benchmarks used for calculating the AQI are pollutant specific and are established by the EPA through the National Ambient Air Quality Standards.

## 4. With the understanding that we haven’t yet discussed the multivariate statistics and models that you will use to analyze your data, if you do have an idea of the statistical tools/tests you might use (e.g., ANOVA, multiple linear regression, and generalized linear models) briefly discuss them. If not, skip this part.

We will conduct hypothesis testing to examine whether our variables of interest are correlated with other variables in our data set. Hypothesis testing will be conducted using several potential tools such as ANOVA and TukeyHSD testing. This investigation aims to shed light on the behavior of air pollutants around the state of North Carolina. Beyond hypothesis testing for correlation between variables, we may also be able to learn more about potential correlations between the variables studied through linear modeling.

## 5. State any concerns you have about completing the project or specific information or skills you might need that are not listed on the syllabus.

Our group looks forward to learning more of the skills required throughout the course to successfully complete our analysis. We aim to continue to refine our R skills through both coursework and outside research and learning as needed.

## References

1. <https://www.deq.nc.gov/about/divisions/air-quality/air-quality-monitoring>
2. <https://www.who.int/teams/environment-climate-change-and-health/air-quality-energy-and-health/health-impacts>