

Analysis of the effect of PM10 Concentration,
Temperature, Geographic Location,
Population, and distance to
Electricity-Generation Combustion Points on
the concentration of fine particulate matter
(PM2.5) in North Carolina during the year
2018.

https://github.com/fr55/DataAnalytics_FinalProject

Felipe Raby Amadori

Abstract

Experimental overview. This section should be no longer than 250 words.

Contents

1	Research Question and Rationale	5
2	Dataset Information	6
2.1	EPA PM2.5 Dataset	6
2.1.1	Data Content Information	6
2.2	EPA PM10 Dataset	6

List of Tables

1	Selections	6
2	Dataset content	7
3	Selections	7

List of Figures

1 Research Question and Rationale

Nowadays air pollution is one of the most relevant health issues in the world. It refers to the contamination of the air by chemicals, biological materials, and other types of pollutants that are harmful to human health. To solve the problem of air pollution, it's necessary to understand the problem, what are the causes, and search for solutions based on the findings.

Particulate matter with a diameter of less than 2.5 micrometers is called PM2.5, and it is a extremely harmful air pollutant because it consists of particles with diameters that are less than or equal to 2.5 microns in size, which can get deeply into the lung, and ultimately impair lung function.

This study focus on trying to understand how PM2.5 concentration in North Carolina vary with temperature, PM10 concentration, zoning (piedmont, coastal, mountain), population, elevation, and distance to combustion points for electricity generation. This last variable was included because according to the EPA combustion for electricity generation is the major point-source sector for PM2.5 in the USA (EPA, 2019).

The research question is: What are the effects of temperature, PM10 concentration, zoning (piedmont, coastal, mountain), population, elevation, distance to combustion points for electricity generation, in PM2.5 concentrations within North Carolina in the year 2018?

2 Dataset Information

For the analysis the following datasets were considered:

2.1 EPA PM2.5 Dataset

This dataset contains data from air quality monitoring of PM2.5 in North Carolina in 2018, and it was obtained using the Download Daily Data Tool in the United States Environmental Protection Agency (EPA) webpage <https://www.epa.gov/outdoor-air-quality-data/download-daily-data> where the options showed in Table 1 were selected:

Option	Selection
Pollutant	PM2.5
Year	2018
Geographic Area	North Carolina
Monitor Site	All Sites
Download	Download CSV (spreadsheet)

Table 1: Selections

The downloaded file was saved in the project folder path `./Data/Raw/` as `EPAair_PM25_NC2018_raw.csv` on 2019-03-31.

2.1.1 Data Content Information

The dataset contains daily mean PM2.5 concentration in $\mu\text{g}/\text{m}^3$ in 2018. Data from 24 stations in 21 different counties of North Carolina with their location in NAD83 lat/long coordinates.

The dataset contains 19 columns, which are shown in Table 2. Column names without description are self-explanatory.

2.2 EPA PM10 Dataset

This dataset contains data from air quality monitoring of PM10 in North Carolina in 2018, and it was obtained using the Download Daily Data Tool in the United States Environmental Protection Agency (EPA) webpage <https://www.epa.gov/outdoor-air-quality-data/download-daily-data> where the options showed in Table 3 were selected:

Column	Description
Date	mm/dd/YY
Source	AQS (Air Quality System)
Site ID	A unique number identifying the site.
POC	“Parameter Occurrence Code”, distinguishes different instruments that measure the same parameter at the same site.
Daily Mean PM2.5 Concentration	
Units	Concentration Units
DAILY_AQI_VALUE	AQI = Air quality index
Site Name	
DAILY_OBS_COUNT	
PERCENT_COMPLETE	
AQS_PARAMETER_CODE	
AQS_PARAMETER_DESC	
CBSA_CODE	
CBSA_NAME	
STATE_CODE	
COUNTY_CODE	A unique number identifying the County.
COUNTY	
SITE_LATITUDE	NAD83
SITE_LONGITUDE	NAD83

Table 2: Dataset content

Option	Selection
Pollutant	PM10
Year	2018
Geographic Area	North Carolina
Monitor Site	All Sites
Download	Download CSV (spreadsheet)

Table 3: Selections