Exploring the relationship between the Air Quality Index (AQI) and respiratory conditions based on age and geographic location.

Abstract:

Respiratory illness across America continues to increase year by year. This can be attributed to the poor air quality due to traffic pollution and global warming. We can help increase the awareness of respiratory health problems caused by poor air quality based on the candidate's age, geographic location, and the current season.

Motivation:

As climate changes, so do the major health issues faced by humanity. This is particularly prevalent in the case of air quality. Local air quality affects how we live and breathe. Moreover, it changes from day to day. A key tool in monitoring these fluctuations in the air quality is the Air Quality Index (AQI). It tells you the concentration of various particles or compounds that are present in the air at any given time. Particularly, it measures ground ozone, particle pollution, Carbon Monoxide and Sulfur Dioxide.

Global warming directly contributes to ground level ozone through decelerating the rate at which CO2 is recycled from our atmosphere. Our primary motivation behind pursuing this problem statement is to help increase the awareness of the impact that air quality has on our respiratory systems. Subsequently, we would like to influence policy making to help those who are most at risk.

Literature Survey:

1.    **The relationship between air pollution and asthma in Malaysian school children**

Summary: The authors have investigated asthma symptoms among Malaysian school due to exposure to traffic-related air pollution in the urban and semi-rural areas. They have performed multiple logistic regression analysis and found that asthma and dry cough are significantly higher among urban living school children compared to semi-rural school children.

2.    **PM2.5 air pollution and cause-specific cardiovascular disease mortality**

Summary: Chances of dying from a CVD or IHD increase with PM2.5 concentration. The key parameter that indicated chanced of death based on AQI was indicated as the Hazard Ratio (HR). The methods used in this paper included:

-      Modeling exposure response as a continuous variable (regression).

-      Cox regression modeling with time-dependent covariates was used to model the hazard ratio (HR) and 95% confidence intervals (CIs) for CVD mortality in relation to ambient PM2.5

Conclusion: Each increase of 10 milligrams/m3 PM2.5 was associated, in fully adjusted models, with a 16% increase in mortality from ischaemic heart disease [hazard ratio (HR) 1.16; 95% CI 1.09-1.22] and a 14% increase in mortality from stroke (HR 1.14; CI 1.02-1.27)

3.    **Research on the Topological properties of Air Quality Index Based on a Complex Network**

Summary: Topological AQI Properties in Beijing was measured from 1 November 2013 to 31 October 2017 by converting AQI time series into simplified symbol sequence. Then each sequence was used to construct AQI direct-weighted network. Complex network theory was utilized on time series for analyzing the centrality, cluster-ability and ranking of AQI network.. Power law distribution found between main components in the network.

4.    **Traffic-related pollution and asthma prevalence in children**

Summary: This study has investigated Asthma symptoms in children with the prevalence of Nitrogen dioxide in air due to traffic condition. The researchers have collected data for a period of 1 year - interviewing the parents in a 12 month period and monitoring No2 content in air. From the previous literature, the possible confounding factors are  (1) indoor—gas stoves, pets, damp etc.; (2) socioeconomic—occupation, education etc.; (3) to Peicbacco smoke—current parental, in utero; (4) demographic— age, sex, ethnicity etc.; and (5) other—breast feeding, parental allergies, past respiratory infections etc.

The techniques used are land use regression, interpolation from monitors and dispersion models. The conclusion is the evidence for associations with NO2 was mixed and that the causality of any associations was judged to fall somewhere between “sufficient” and “suggestive but not sufficient” to infer causality.

5.    **Outdoor air pollution and health effects in urban children with moderate to severe asthma.**

Summary: This study investigates particulate matter less than 2.5 μm in diameter association with Asthma. Data collects in NY city for 2 periods (Summer and winter) for outdoor particulate matter and lung function are collected. The odd ratio for pollutant concentration and Asthma seems no longer significant during summer but negatively associated with symptoms during winter.

6.    **Impact of Long-Term Exposures to Ambient PM2.5 and Ozone on ARDS Risk for Older Adults in the United States**

Summary: People who are of age that is greater than 65, who have reported ARDS due to poor air quality. They used a linear model to determine if someone would report that they had trouble breathing due to the poor air quality during that day.

Methodology:

1. Experiment design:
   1. Literature Survey
   2. Data Collection
   3. Data Signature verification
   4. Problem statement generation and hypothesis

We are continuously exploring the available data and have reached out to a few public organizations such as \_\_\_ to obtain anonymous patient ER visit data.

1. Algorithm/Model design:

We plan on using similar techniques to those that were done in the aforementioned literature. Our plan is to use regression, dispersion modeling/ clustering, and interpolation to reach a conclusion for our hypothesis.

The proposed regression model would be multi-dimensional. It would encompass the additional factors of age and geographic location into the present AQI. Then, it would provide a prediction regarding the likely profile of the people who are prone to visit the hospital for a respiratory illness on any given date.

1. Evaluation methods:
2. How well our data represents the real world
3. Other factors we missed out while looking at our problem

The efficacy of our model depends on how well the data represents the real world scenarios, and that is an important point which we will consider in our model evaluation

Deliverables:

Technical Paper

- Abstract

- Methods

- Results

- Formulas: Mathematical representation of air quality impact on respiratory health by age and location

- Discussion

-  Visualizations

**Presentation**

-      Condensed version of the paper

**Code**

**-** Github repository

-      Jupyter notebook

-      Datasets

Team Members

|  |  |
| --- | --- |
| Task | Member |
| Data Collection - Air Quality | Subarna, Manmeet |
| Data Collection - Health Data | Jeyasri, Pranav |
| Literature Survey | All |
| Data Cleaning | All |
| Experiment Design | All |
| Implementation | Jeyasri and Subarna |
| Visualization | Pranav and Manmeet |
| Evaluation | All |
| Paper | All |

Reference

1.    <https://academic.oup.com/ije/advance-article/doi/10.1093/ije/dyz114/5529269>

2.    <https://www.mdpi.com/2071-1050/10/4/1073/pdf>

3.    <https://www.sciencedirect.com/science/article/abs/pii/S0012369219307111>

4.    <https://ieeexplore.ieee.org/abstract/document/8679335>

5.    <http://www.baaqmd.gov/about-air-quality/current-air-quality/air-monitoring-data/#/aqi?id=316&date=2015-09-21&view=monthly>

6. <https://airnow.gov/index.cfm?action=airnow.local_state&stateid=5&mapcenter=0&tabs=0>

7.    <https://aqicn.org/api/>