

p8105_final_project

Leila Yan, Tong Su, Shike Zhang, Shumei Liu, Alice Mao

2024-11-06

Tentative project title: The impact of air quality on respiratory illnesses in NYC

Motivation for this project: Since arriving in New York for school, I have experienced recurring symptoms of rhinitis. This persistent discomfort has led me to suspect that the city's air quality may be a contributing factor. I see this as an opportunity to conduct research on New York City's air quality, aiming to understand how it may impact respiratory health and to provide insights that could contribute to better public health outcomes.

Intended final products: We intended to create a website to report our findings.

Anticipated data sources:

- NYC Open Data
- DOHMH

Planned analyses / visualizations / coding challenges: Planned analyses

- Seasonal trends: analyze seasonal variations in pollutant levels and respiratory symptoms
- Geospatial analysis: create maps to visualize air quality levels and respiratory health outcomes by area

Visualization

- PM 2.5 mapping: display PM2.5 concentrations across different areas in NYC, highlighting pollution hotspots and spatial patterns that impact respiratory health.
- Line graph: Show trends in PM2.5 levels and respiratory illness cases over time, allowing for easy comparison of seasonal and daily fluctuations in both variables.

Challenges

- Missing data/incomplete data entries for a specific borough
- Multiple data sources: Air quality and health data come from multiple sources like government agencies, monitoring stations, and hospitals. Integrating these requires careful matching of location and time variables, adding complexity.

Planned timeline:

- Nov. 11-15: Data collection and preprocessing cleaning
- Nov. 16-30: Exploratory Data Analysis (EDA), perform R coding and analyses
- Dec. 1-7: Report + website + screencast + peer assessment
- Dec. 8-12: Finalize + wrap up