Project Title: The Impact of Housing Environment on Children's Blood Lead Levels

Team Members: Nicole Sosa, Seoeun Hong

Project Description:

Our project aims to investigate the relationship between children's elevated blood lead levels and their housing environment, as children can have elevated blood lead levels if they are exposed to old paint, old water pipelines, and lead dust. Unfortunately, children who have elevated blood lead levels can experience detrimental health effects such as delayed growth and development, damage to the brain and nervous system, and other serious health effects. We hope to provide insights that highlight the impact of housing conditions on children's health and development and to raise awareness of the critical importance of safe and healthy housing for all children. Through our analysis, we seek to find any trends that can influence practices or policies to improve housing conditions and reduce children's blood lead level incidence.

Who will benefit from our analytics?

- NYC Parents
- NYC Children
- NYC Housing Providers
- Policymakers

What insight will you derive from the data?

The insights that we want to derive from this data are to identify the housing conditions that are most strongly associated with blood lead levels in children. In addition, we want to understand the geographic distribution of lead exposure in children to better understand any geographical influences that may contribute to children's lead exposure.

Describe how you will check the goodness of the analytics. i.e., why do you believe that the results are accurate and can be trusted?

References to research papers:

- https://www.sciencedirect.com/science/article/abs/pii/S0013935103000367
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1447158/
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5904740/

The paper, "The effect of lead-based paint hazard remediation on blood lead level of lead-poisoned children in New York City," clearly shows that housing environments, such as dust, housing painting, and housing materials, have a strong correlation with

children's blood lead level. Moreover, the paper "Relation Between Housing Age, Housing Value, and Childhood Blood Lead Levels in Children in Jefferson County, Ky " concluded that elevated blood lead levels were associated with living in older housing. Therefore, we can assume that our approach is on the right track. Additionally, our two data sources are from Open Data NY - DOH and each of them is provided by the New York State Department of Health, which are government departments. Since all of our data is trustworthy, we can say that our results would be accurate and can be trusted.

Data sources (one per each team member). Include the source, link and schema of the datasource.

- Childhood Blood Lead Testing and Elevated Incidence by Zip Code: Beginning 2000
    - Link: https://health.data.ny.gov/Health/Childhood-Blood-Lead-Testing-and-Elevated-Incidenc/d54z-enu8
    - Schema of the datasource:

Field Name	Data Type	Brief Description
County	Plain Text	County of residence
County Code	Number	County FIPS code
Year	Number	Year the lead test was collected
Zip	Plain Text	Zip Code of residence
Tests	Number	Number of children who resided in the zip code and had a blood lead test within the specified year.
Less than 5 mcg/dL	Number	Number of children who resided in the zip code that were identified to have a unconfirmed and

		confirmed concentration of less than 5 micrograms per deciliter (mcg/dL) lead in their within the specified year.
5-10 mcg/dL	Number	Number of children who resided in the zip code that were identified to have a unconfirmed and confirmed concentration of lead in their blood between 5 – 10 micrograms per deciliter (mcg/dL) within the specified year.
10 – 15 mcg/dL	Number	Number of children who resided in the zip code that were identified for the first time to have a confirmed concentration of lead in their blood between 10 – 15 micrograms per deciliter (mcg/dL) within the specified year.
15 + mcg/dL	Number	Number of children who resided in the zip code that were identified for the first time to have a confirmed concentration of lead in blood greater than 15 micrograms per deciliter

		(mcg/dL) within the specified year.
Total Elevated Blood Levels	Number	Total number of children who resided in the zip code that were identified for the first time to have a confirmed elevated lead blood level within the specified year. Blood lead concentrations ≥ 10 mcg/dL are considered elevated.
Percent	Number	Total number of children who resided in the zip code that were identified for the first time with confirmed elevated lead blood levels within the specified year, divided by the total number of children who resided in the zip code that had a blood lead screening test in that same year multiplied by 100.
Rate per 1,000	Number	Total number of children who resided in the zip code that were identified for the first time with confirmed elevated lead blood levels within the specified year, divided by the total

		number of children who resided in the zip code that had a blood lead screening test in that same year multiplied by 1,000.
Zip Code Location	Location	Centroid location for zip code where child lives.
County Location	Location	Latitude and Longitude of New York State County's centroid

- Housing Maintenance Code Violations
  - Link: https://health.data.ny.gov/Health/Healthy-Neighborhoods-Program-Housing-Demographics/jbwf-vnai
  - Schema of the datasource:

Field Name	Data Type	Brief Description
Funding Cycle	Plain Text	County health departments submit applications to compete for funding to implement the Healthy Neighborhoods Program. Only full-service health departments are eligible to apply for funding. (A full-service health department has a Division of Nursing and a Division of Environmental Health.) Funding runs on cycles of 4 or 5 years. Both federal and state monies have funded the program; the current cycle is funded by the state budget.
County Name	Plain Text	The name of the county with a Healthy Neighborhoods Program

		(HNP). “All counties” is the combined data for all counties funded during a funding cycle.
Visit Type	Number	Types are Initial visits and Revisits. At the initial visit an HNP surveyor gathers information about the building, primary respondent and residents and conducts an assessment of the conditions of the home. At this visit, the surveyor provides the HNP interventions and guidance to assist residents in addressing potential health and safety hazards. About 25% of these homes are revisited 3-6 months after the initial visit. Homes needing a revisit are prioritized on the overall or severity of the conditions identified at the initial visit or homes with residents with asthma. The home is reassessed for ongoing or new conditions and additional interventions are provided as needed.
Item	Plain Text	Variables are grouped into item types: housing information, respondent demographics, and housing conditions.
Variable	Plain Text	Variables are the individual building characteristics, respondent demographics, or housing conditions assessed at the initial visit or revisit. Information about the building, primary respondent and residents are only collected at the initial visit.
Variable Detail	Plain Text	Variable detail are the responses possible for each variables. Housing conditions details are Yes, No, Unknown, and Not Applicable. The details for each building characteristic or respondent demographics depends on the variable. For example, the

		details for Age of Building are Pre-1950, 1950-1978, and Post-1978.
Frequency Count	Number	This is the number of dwellings that have a specific detail (e.g., the number of dwellings built before 1950, 1950-1978, after 1978, and Unknown). These sum to the number of dwellings visited.
Percent of Total Frequency	Number	
Date Closed	Date & Time	
Location	Location	The latitude and longitude of the county centroid.

Draw initial design diagram(s) using PowerPoint, Keynote, Google Drawing, etc. to describe your project work flow.