

**Preventing Childhood
Lead Poisoning**
Chicago, IL

Station #2

The Bottom Line	Data Science for Social Good (DSSG) developed a predictive model to determine the risk of a child being poisoned by lead. This model encouraged city inspectors to prioritize potential hazards and identify children most at risk.
Problem	Lead poisoning is a significant public health issue in the United States. Exposure to lead has been linked to premature birth, neurological development issues, and lowered IQs. Like many resource-constrained agencies, the Chicago Department of Public Health (CDPH) does not have the funding to inspect and remediate all of the older buildings in Chicago. When interventions do occur, it is often after a child is poisoned.
Leadership	The project was lead by CDPH. CDPH's leaders played a central role in supporting the initiative and guaranteeing that the analysts had the support they needed to complete the project.
Staffing	DSSG worked directly with the Lead Poisoning Prevention Program in CDPH. The team included the Commissioner of CDPH, an epidemiologist, administrative manager of the lead program, and a representative from the local Women Infant and Children (WIC) clinics.
Data	Two datasets key to the researchers' analysis included blood lead level tests and home lead inspection records. DSSG also leveraged open data such as building footprint data, shapefiles of census tract and ward boundaries, assessed property value, building classification, and census tract variables from the American Community Five-Year survey. (Examples of variables include socio-demographics, education, health insurance, and home ownership rates.)
Technology	To clean and aggregate the data, the team from DSSG used PostgreSQL with the geospatial extension PostGIS. Python was utilized to assemble datasets and clean them by removing duplicates. All source code is available on GitHub .
Methodology	The team from DSSG built a classifier to predict a child's risk of being poisoned by lead. This model was then evaluated by using a cross-validation strategy.
Results	CDPH plans on making the following changes to its lead poison prevention strategies: <ul style="list-style-type: none">- Encouraging pregnant women and parents of young children to request home inspections via billboards and other outreach efforts;- Partnering with doctors and healthcare providers to facilitate inspections for patients living in high-risk areas;

- Piloting a program to include risk scores in a child's medical records; and
- Ratcheting up outreach and education efforts for landowners and housing providers.

Replication

To replicate the project, cities need guaranteed access to the data sources leveraged in the project, including but not limited to the age of housing stock and building inspection history. Some of the data sources are readily available to the public like Census data, while other datasets are somewhat proprietary (e.g., information related to the historical blood levels for children). To add value to the project, cities must have inspection and remediation resources to address their lead problems.

Learn More

Data Science for Social Good | Center for Data Science and Public Policy, University of Chicago | www.dssg.uchicago.edu | datascifellows@gmail.com