Spatial density of vacant properties: A model for estimating the magnitude of background lead (Pb) exposure in children.

Ivan E. Castro

Syracuse University
Falk College, Department of Public Health

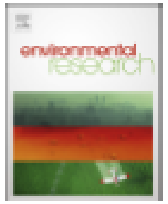
April 16, 2019

Environmental Research 170 (2019) 463–471

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres



Variability in the spatial density of vacant properties contributes to background lead (Pb) exposure in children

Ivan E. Castro[a], David A. Larsen[a], Bryce Hruska[a], Patrick J. Parsons[b,c], Christopher D. Palmer[b,c], Brooks B. Gump[a,*]

[a] Department of Public Health, Food Studies, and Nutrition, Syracuse University, Syracuse, NY 13204, USA
[b] Laboratory of Inorganic and Nuclear Chemistry, Wadsworth Center, New York State Department of Health, Albany, NY 12201, USA
[c] Department of Environmental Health Sciences, School of Public Health, University at Albany, Rensselaer, NY 12144, USA

ARTICLE INFO

Keywords:
Spatial modelling
Lead exposure
Environmental health
Blood lead
Spatial epidemiology

ABSTRACT

Background: Heightened blood lead levels (BLL) are associated with cognitive deficiencies and adverse behavioral outcomes. Lead-contaminated house dust is the primary source of exposure in U.S. children, and evidence suggests that even background (low-level) exposure has negative consequences. Identifying sources of background exposure is of great public health significance because of the larger number of children that can be affected.

Methods: Blood lead was assessed in a bi-racial sample of children from Syracuse, NY, aged 9–11, using established biomonitoring methods. The spatial density of vacant properties was modelled from publicly available georeferenced datasets. Further, regression models were used to measure the impact of this spatial density variable on children's BLL.

Results: In a sample of 221 children, with a mean BLL of 1.06 µg/dL (SD = 0.68), results showed increases in spatial density of vacant properties predict increases in median blood-PB levels, b = 0.14 (0.06–0.21), $p < .001$. This association held true even after accounting for demographic covariates, and age of individual housing. Further analysis showed spatial autocorrelation of the residuals changed from a clustered pattern to a random pattern once the spatial density variable was introduced to the model.

Discussion: This study is the first to identify a background-lead exposure source using spatial density modelling. As vacant properties deteriorate, lead-contaminated dust likely disperses into the surrounding environment. High-density areas have an accumulation of lead hazards in environmental media, namely soil and dust, putting more children at risk of exposure.

Background

- Lead (Pb) is the most common environmental toxicant among children
- Cognitive impairments
  - Decreased IQ
  - Impaired executive functioning
- Adverse behavioral outcomes
  - Increased anxiety, impulsivity, hostility
- Increased vascular responses to stress

Even at low levels…

Trends in blood lead

In Syracuse

- Average blood-lead in children under 6:
  - 8.77 mcg/dL in 1992
  - 3.94 mcg/dL in 2011
- Primary interventions
- Known mechanisms of exposure
- Children from deprived, urban neighborhoods still have the highest burden

  Why…?

Background exposure

- Less clear mechanism of low-level exposure
  - Airborne lead from resuspended soil
  - Differences in dietary exposure
  - Trace metals in street dust

- Little research on identifying mechanisms of chronic background exposure

Chronic background exposures unassumingly reach a larger number of children

Identifying sources of background exposure

- Current research provides initial understanding of housing age and condition as risk factors for exposure
- We consider older vacant properties (built pre-1978) as a source of exposure
    - 1978 - federal ban introduced on use of lead-based paints in housing
- We hypothesized that where and how these properties are located throughout space, create a mechanism of background exposure

Supporting evidence

- Lead dispersion from built environment
- Lead contaminated dust disperses outwards from demolition site
- Lead dust concentrations in residential entryways are a source of lead in street dust

Growing use of Geographic Information Systems (GIS) in lead exposure research

- Risk factors cluster at census-tract level
  - Age of housing stock
  - ***Proportion vacant ****
  - Median income
- Blood lead levels are spatially autocorrelated

GIS-based lead exposure research

Limitations in the literature

- Most studies do not provide input parameters or reference system
- Aggregated measurements of risk factors and/or outcomes
  - Median values / Percentages / Counts
  - Levels: census-tract, census-block, zip code
  - Modifiable Aerial Unit Problem (MAUP)
- Low blood lead consistently overlooked
  - Most studies dichotomize high/low
- Individual characteristics rarely accounted for
- Soil lead research
  - Kriging of sample sites

The Present Study

Hypothesis:

The spatial density distribution of vacant properties predict blood lead levels in children at their point of residence.

Methods - participants

- Syracuse Lead Study participants
- Blood lead data available for 270 children
- Data used for analysis - 221 children

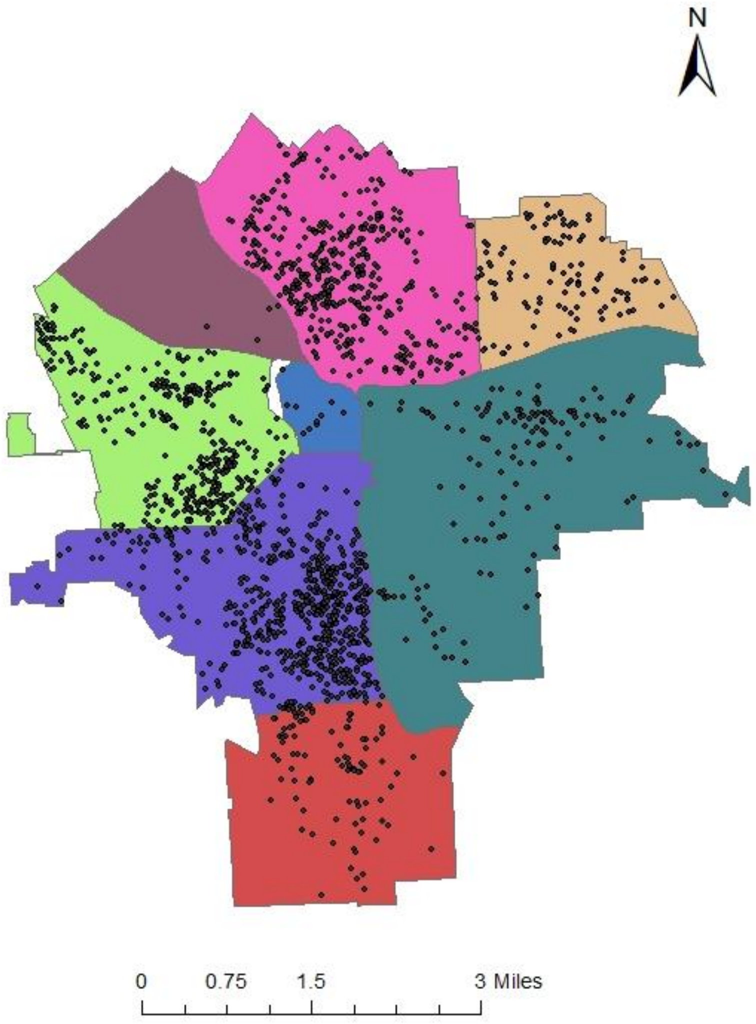
Linking measurements to a spatial point

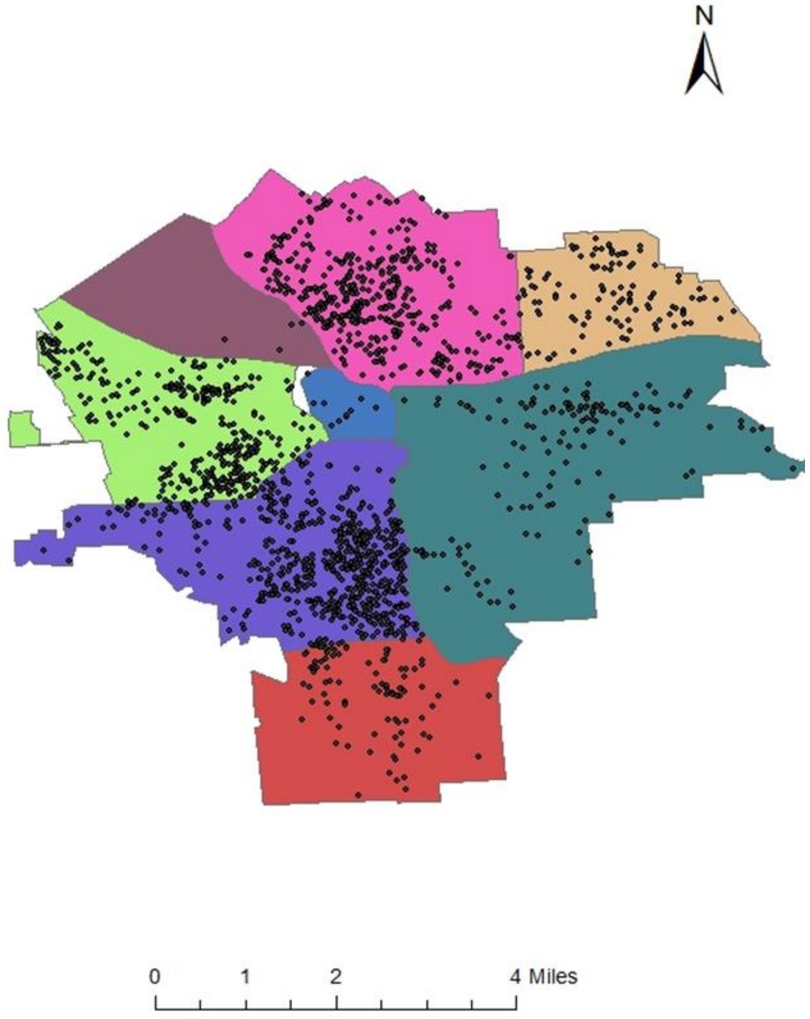
- Home addresses geocoded with NYS GIS Program server
- Projection: *NAD83/UTM Zone 18N*

Methods – parcel data

- Publicly available, downloaded from http://data.syrgov.net/
- Data in polygon shape
  - Converted to point shape: coordinates of the polygon centroid
- Total 1,828 points labelled *vacant (33* built post-1978)
  - 1,795 points used for analysis
- *Projection: NAD83/UTM Zone 18N*
  - Projection transforms coordinate system for accurate measurement of area and distance.

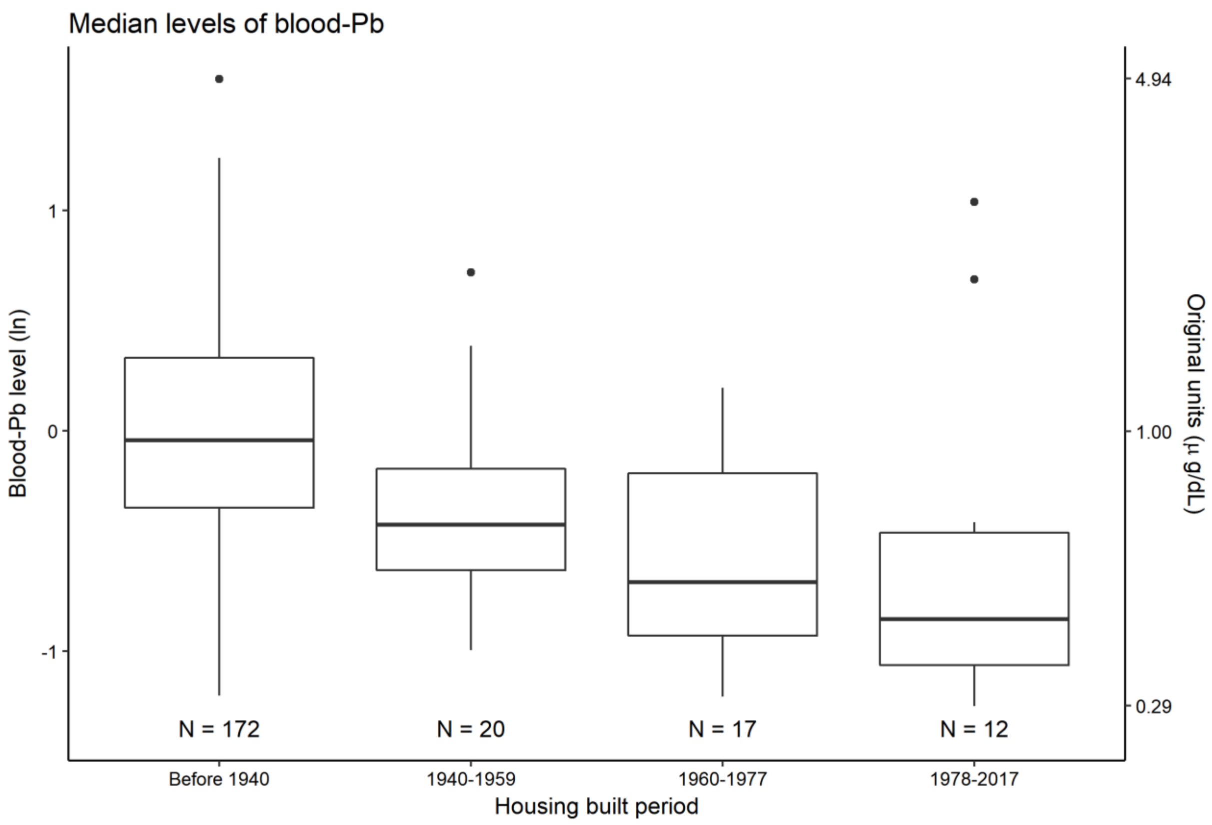
Projected v. Geographic coordinate systems





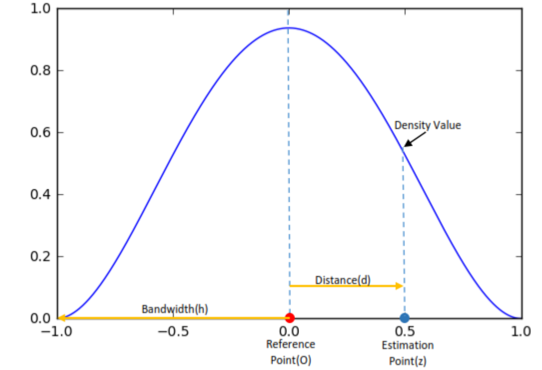
Methods – period built

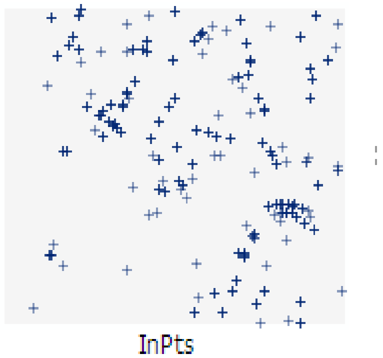
- Obtained from original parcel shapefile
- Categorized per previous research on prevalence of lead-based paints
- Significantly related to outcome
- Older housing is the primary source of exposure

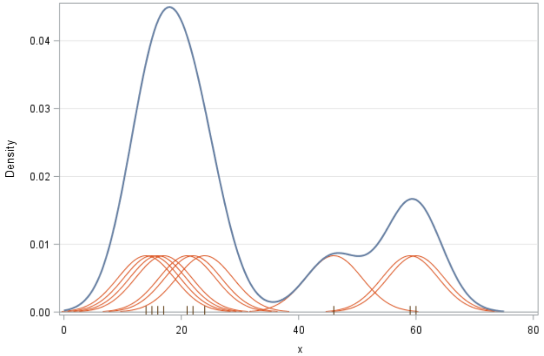


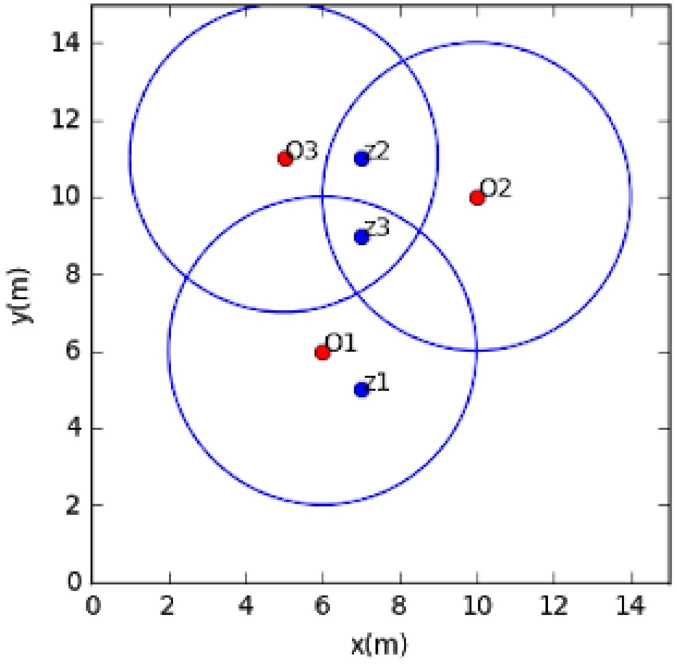
Methods – kernel density estimation

- Quartic kernel function estimates density of point features within a set window (bandwidth)
- A Gaussian distribution function
- Value is highest at point location, decreases to zero at the set bandwidth
- Kernel surfaces are added to calculate density at each cell.



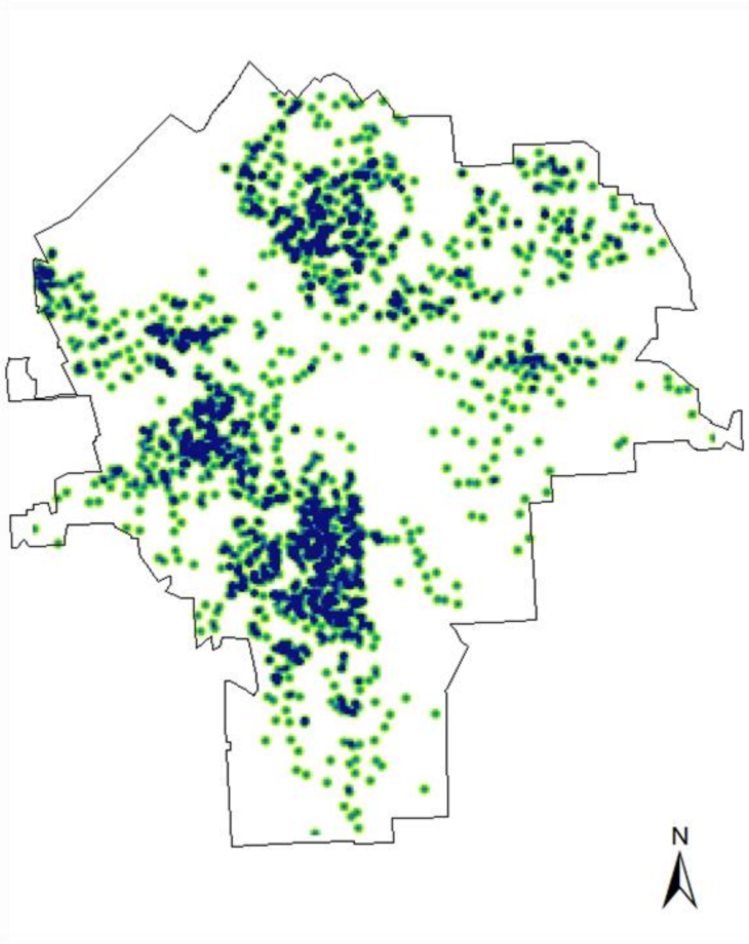


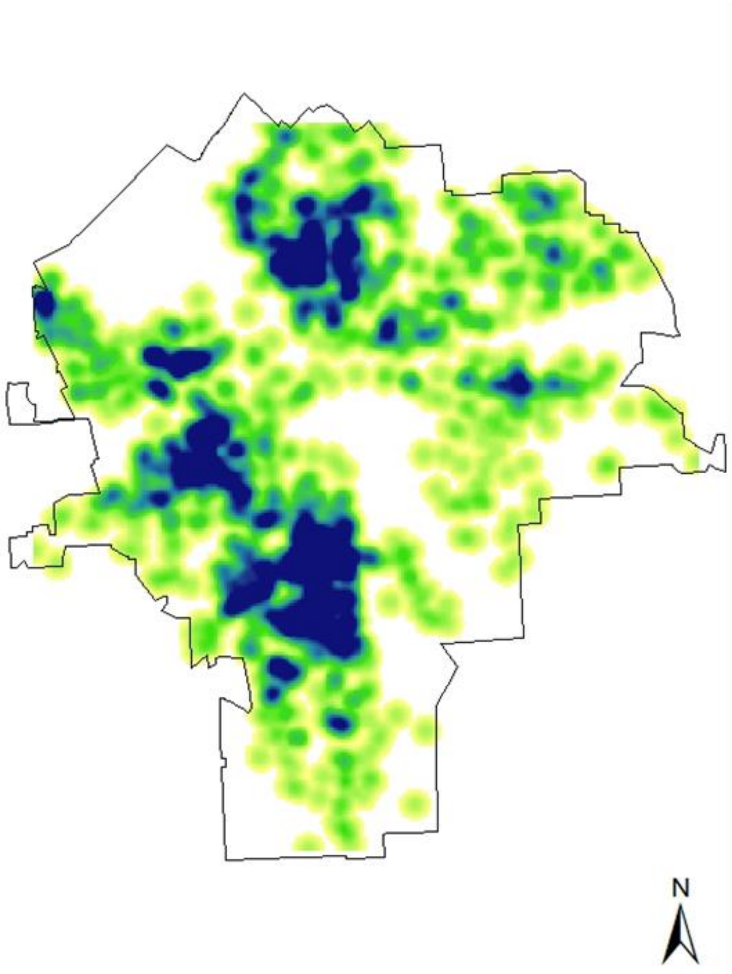




Methods - bandwidth

- Bandwidth defines the radius of the window
- Calculated bandwidths in 30 meter increments from 90-240 meters
- Each bandwidth was a significant predictor
- Larger bandwidths create smoother surfaces





Statistical Analysis

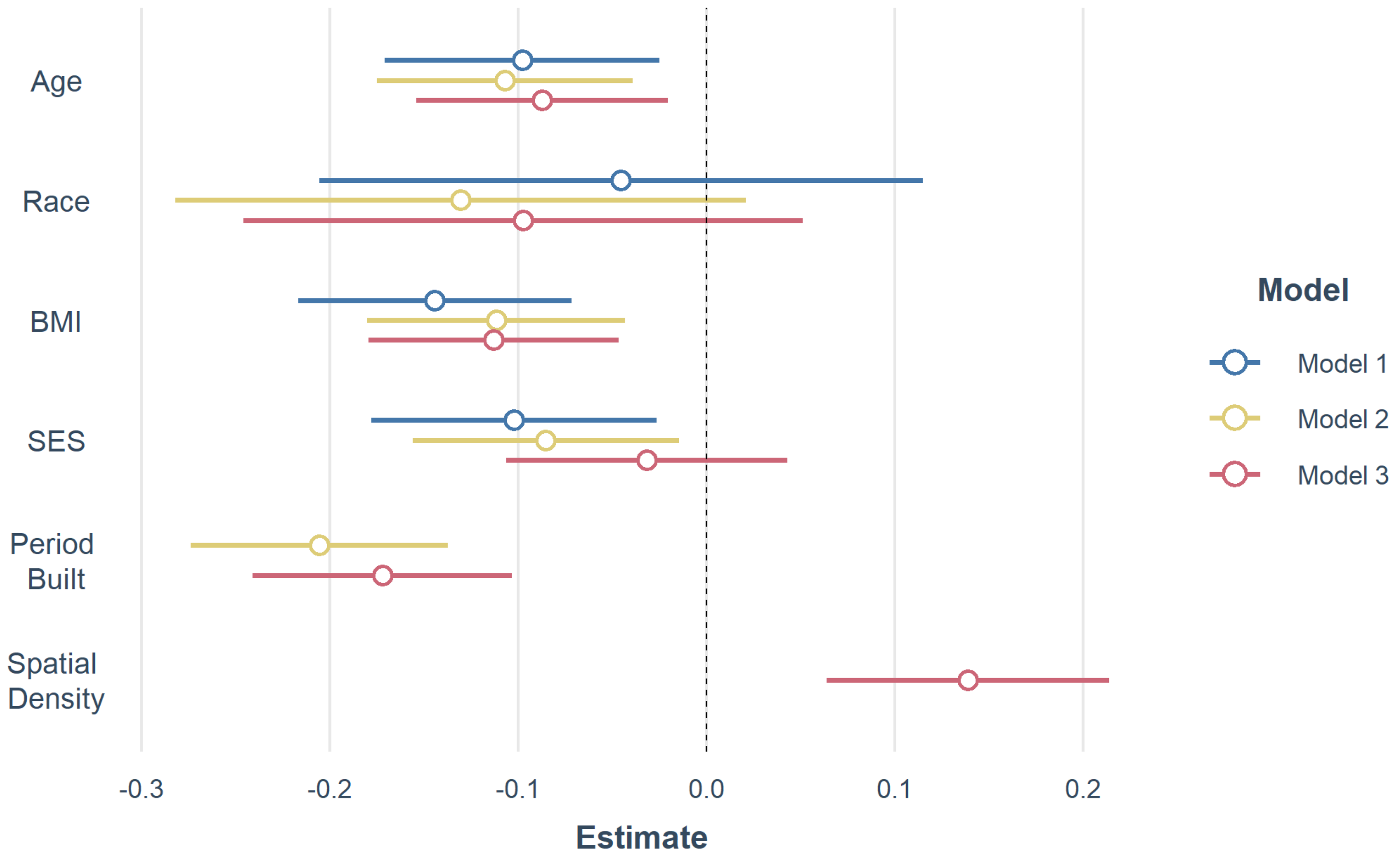
Regression Modelling

- Conducted in R version 3.4.1
- Blood lead measurement was log-transformed
- Density measurement was standardized
- Successive regression models
  - Model 1: Y ~ Age + Race + BMI + SES
  - Model 2: Y ~ Age + Race + BMI + SES + Period Built
  - Model 3: Y ~ Age + Race + BMI + SES + Period Built + Density

Regression Results

Predictor	b	b 95% CI	β	β 95% CI	sr^2	sr^2 95% CI	r	Fit	Difference
Model 1 (Intercept)	1.37**	(0.57, 2.18)							
Age	-0.10**	(-0.18, -0.03)	-0.17	(-0.30, -0.04)	.03	(-.01, .07)	-.22**		
Race	-0.05	(-0.21, 0.11)	-0.04	(-0.17, 0.10)	.00	(-.01, .01)	-.09		
BMI	-0.00**	(-0.01, -0.00)	-0.25	(-0.38, -0.13)	.06	(.00, .12)	-.27**		
SES	-0.13**	(-0.22, -0.03)	-0.18	(-0.31, -0.05)	.03	(-.01, .07)	-.18**	AIC = 357.24	
								R^2 = .142** 95% CI(.06,.22)	
Model 2 (Intercept)	1.86**	(1.10, 2.62)							
Age	-0.11**	(-0.19, -0.04)	-0.19	(-0.31, -0.07)	.03	(-.01, .07)	-.22**		
Race	-0.13	(-0.28, 0.02)	-0.11	(-0.24, 0.02)	.01	(-.01, .03)	-.09		
BMI	-0.00**	(-0.01, -0.00)	-0.20	(-0.32, -0.08)	.04	(-.01, .08)	-.27**		
SES	-0.11*	(-0.20, -0.02)	-0.15	(-0.27, -0.03)	.02	(-.01, .05)	-.18**		
Period built	-0.24**	(-0.32, -0.16)	-0.36	(-0.48, -0.24)	.12	(.05, .20)	-.36**	AIC = 325.77	
								R^2 = .262** 95% CI(.15,.34)	ΔR^2 = .12** 95% CI(.05, .20)
Model 3 (Intercept)	1.55**	(0.79, 2.31)							
Age	-0.09*	(-0.16, -0.02)	-0.15	(-0.27, -0.04)	.02	(-.01, .05)	-.22**		
Race	-0.10	(-0.25, 0.05)	-0.08	(-0.21, 0.04)	.01	(-.01, .02)	-.09		
BMI	-0.00**	(-0.01, -0.00)	-0.20	(-0.31, -0.08)	.04	(-.01, .08)	-.27**		
SES	-0.04	(-0.13, 0.05)	-0.06	(-0.19, 0.08)	.00	(-.01, .01)	-.18**		
Period built	-0.20**	(-0.28, -0.12)	-0.30	(-0.42, -0.18)	.08	(.02, .14)	-.36**		
Kernel density	0.14**	(0.06, 0.21)	0.24	(0.11, 0.37)	.04	(-.00, .09)	.38**	AIC = 314.49	
								R^2 = .306** 95% CI(.19,.38)	ΔR^2 = .04** 95% CI(-.00, .09)

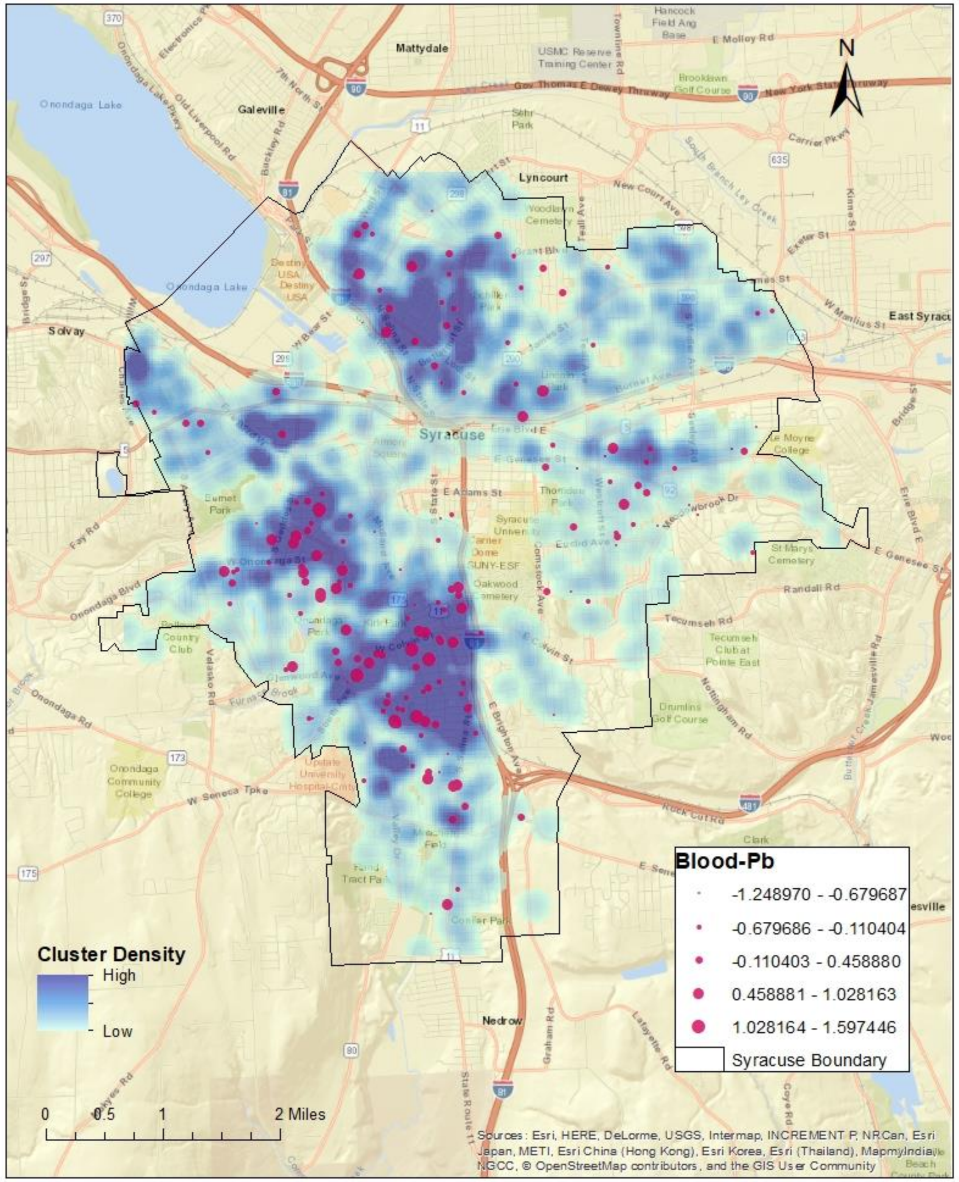
Regression Results



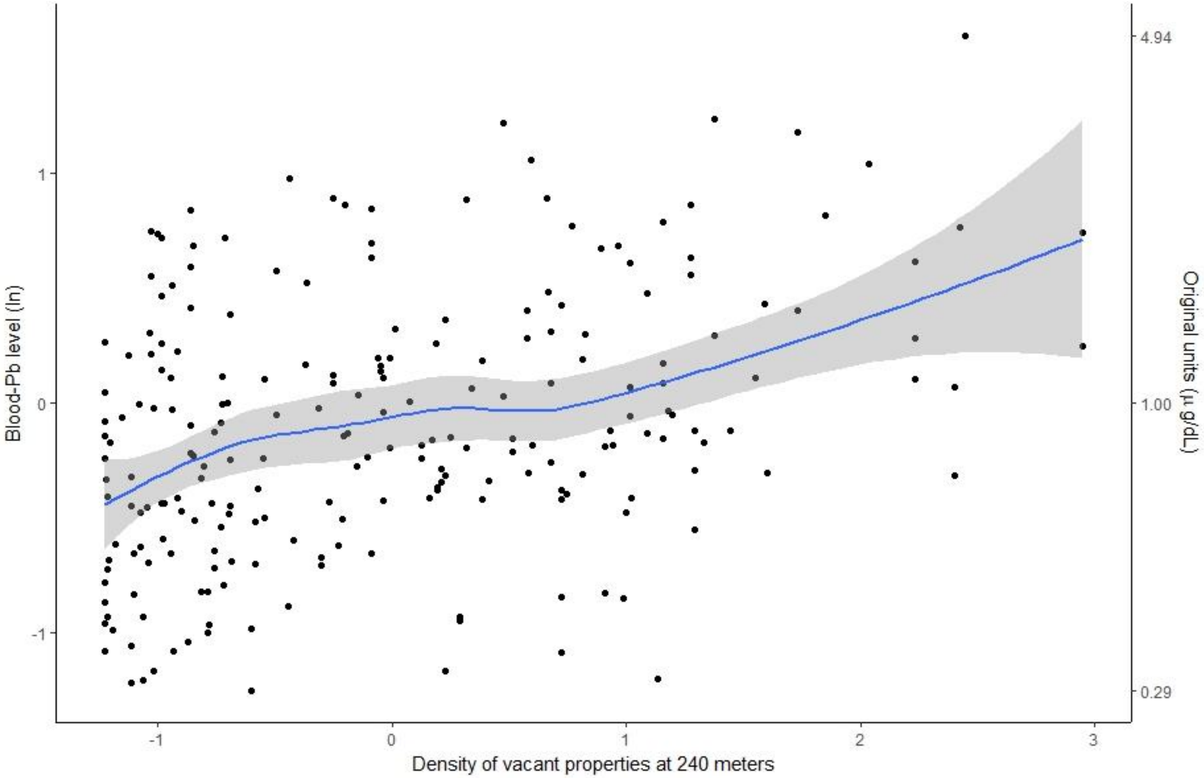
Interpretation

Regression

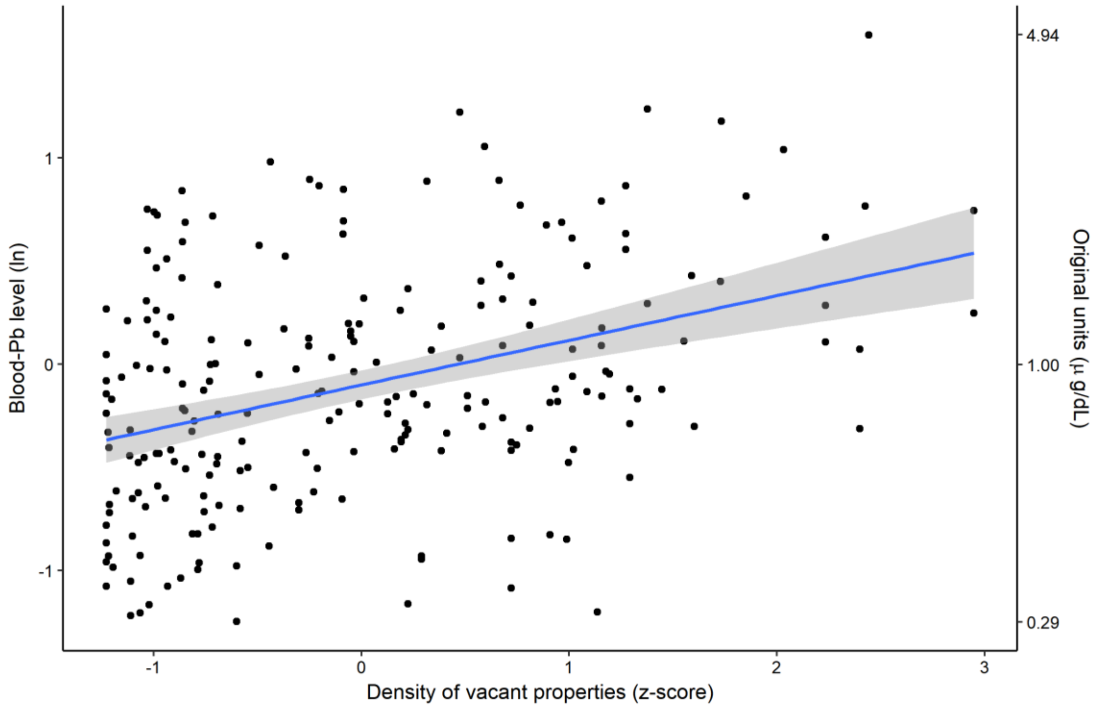
- Age, SES, and BMI have significant negative relationship with outcome
- Housing age also has significant negative relationship
- Increases in density unit show a 15% increase in median blood-Pb
- We interpret this due to the log function of the outcome
    - Exp(0.14) = 1.15
    - (1.15 – 1) x 100 = 15%

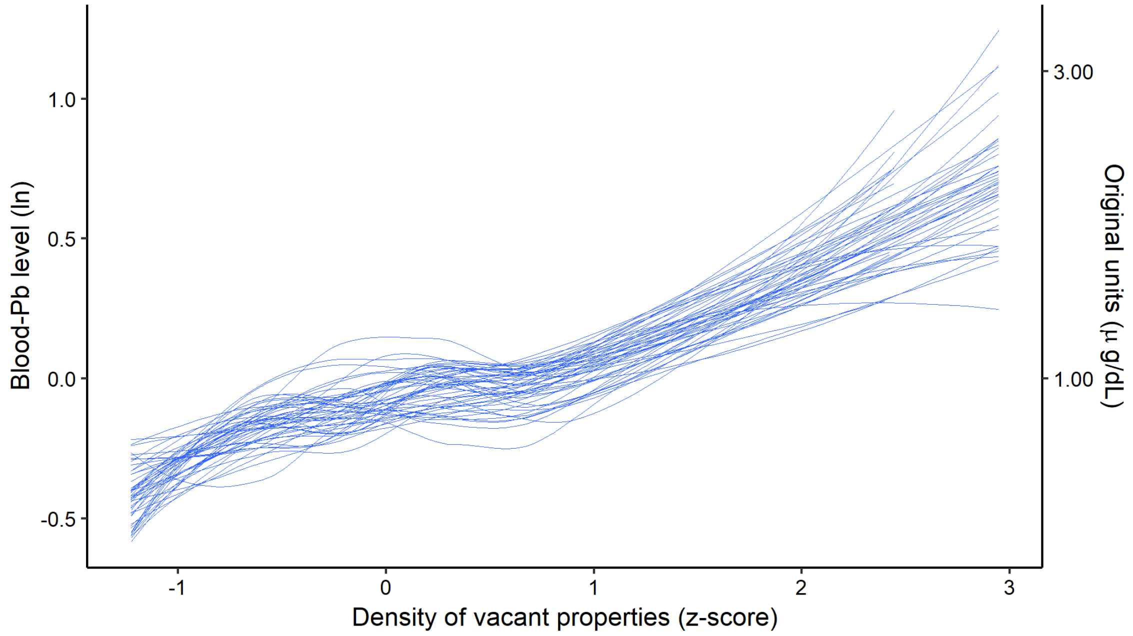


Relationship between density and blood-Pb



Another way to visualize uncertainty





Limitations

- Cross-sectional design does not allow for establishing a causal link
- We don't know how long any particular property had been vacant
- Lead measurements collected as far back as 2013
- We were unable to measure lead in dust or soil

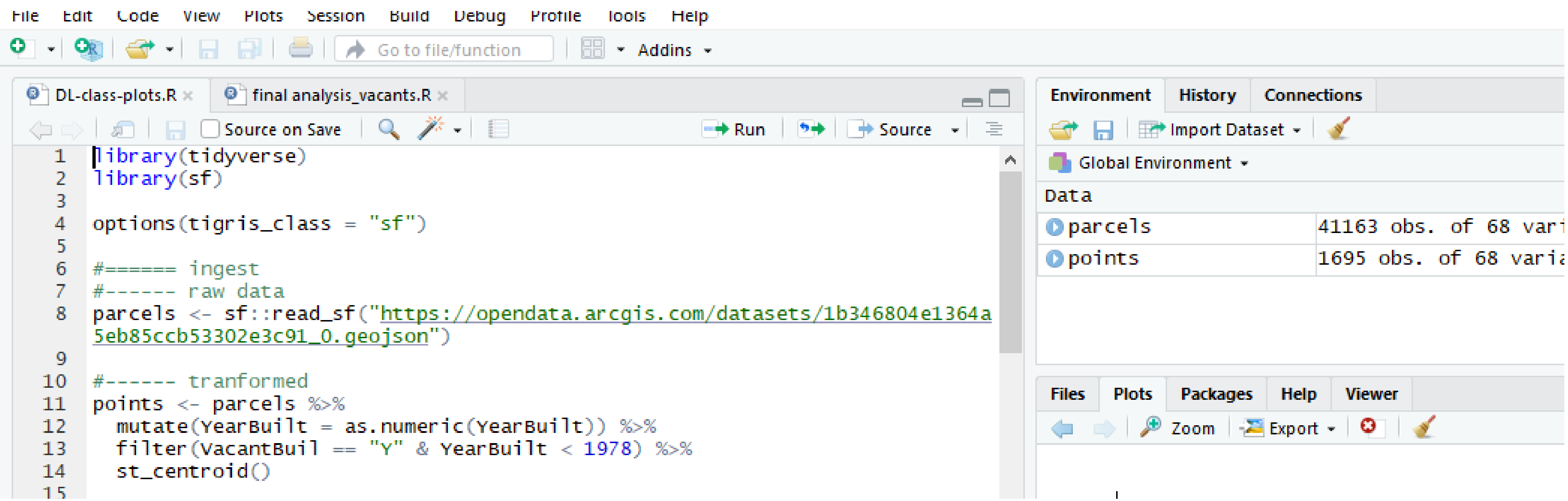
Still… we believe this can be replicated

What about reproducibility?

- Methods are detailed
    - And expanded after peer-review
- I can describe how we selected vacant buildings for analysis
    - but I cannot show you the steps I took

Learn to write code

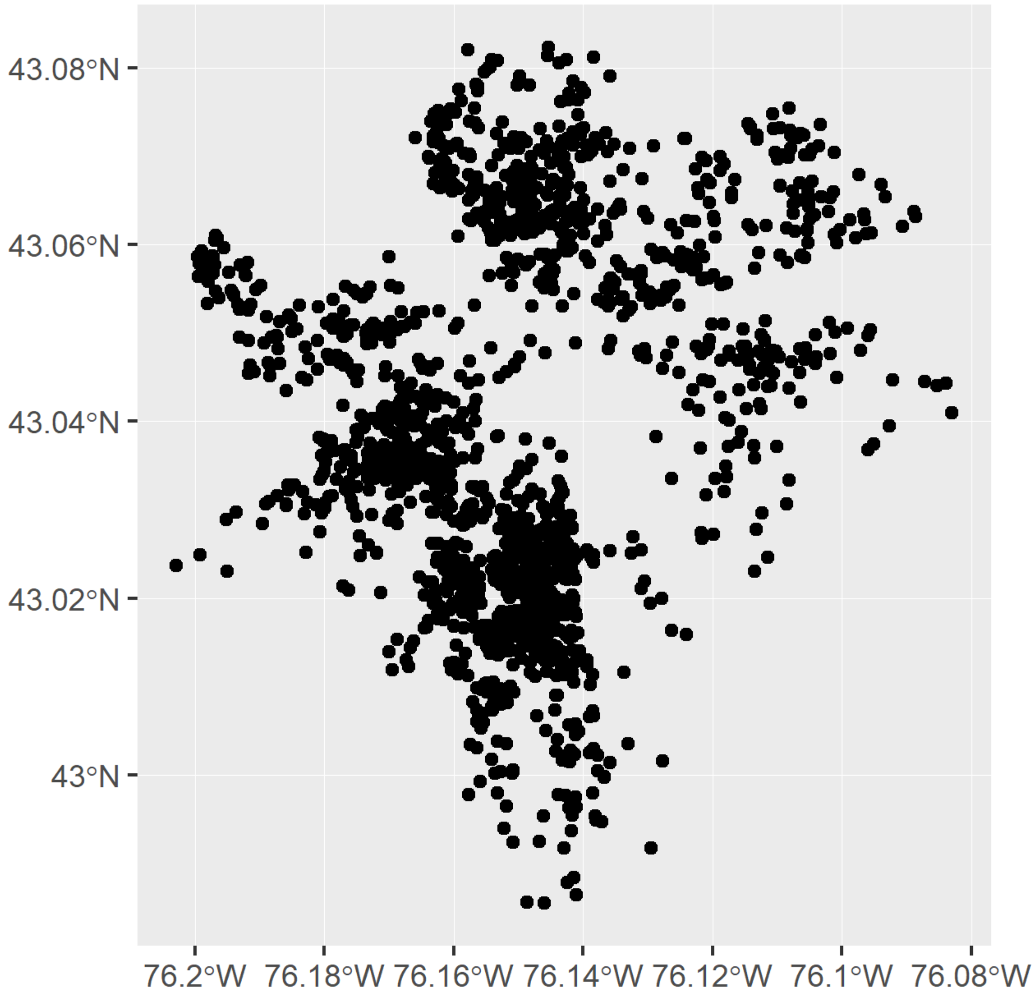
A reproducible example:



Simple Features (sf) for R

R as a GIS

```
ggplot() +
  geom_sf(data = points)
```

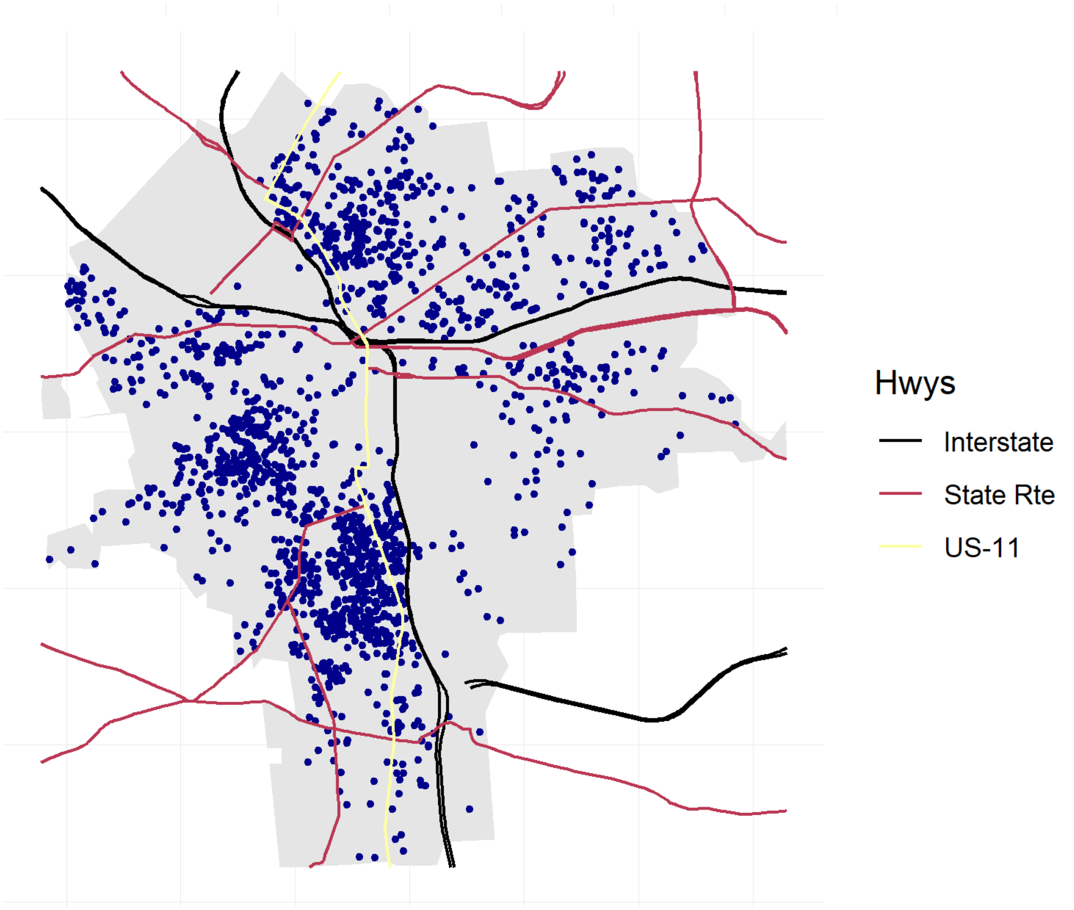


Spatial data collection

```r
options(tigris_class = "sf")


syr <- tigris::places("NY", cb = TRUE) %>%
  tigris::filter_place("Syracuse")


roads <- tigris::roads("NY","Onondaga")

isTRUE(st_crs(syr) == st_crs(roads))


sy_rds <- roads %>%
  filter(RTTYP %in% c("I", "U", "S")) %>%
  st_crop(syr) %>%
  mutate(Hwys = case_when(
    RTTYP == "I" ~ "Interstate",
    RTTYP == "U" ~ "US-11",
    RTTYP == "S" ~ "State Rte"))
```

Plotting simple features (sf)

```
ggplot() +
  theme_minimal() +
  geom_sf(data = syr, fill = "gray90",
          color = NA) +
 geom_sf(data = points, size = .75,
          color = "blue4", fill = NA) +
 geom_sf(data = sy_rds, aes(color =
          Hwys), show.legend = "line") +
 scale_color_viridis_d(option =
          "inferno") +
 theme(axis.text = element_blank())
```

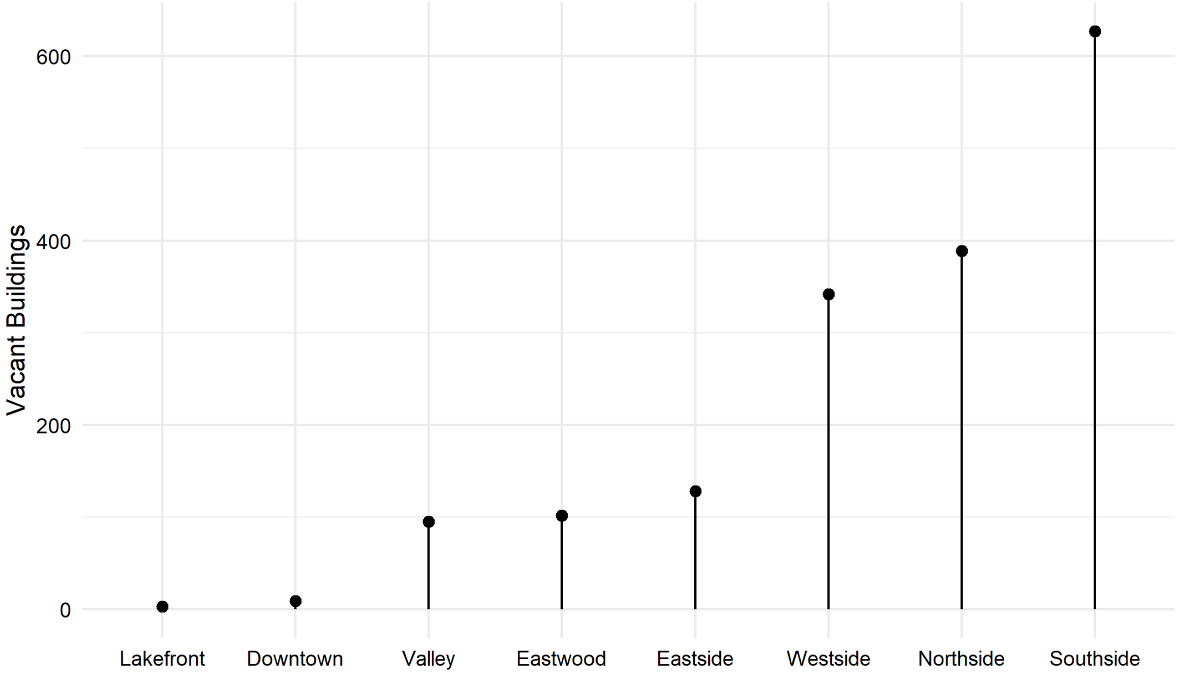


Tidy analysis

```r
points %>%
  st_set_geometry(NULL) %>%

  group_by(TNT_NAME) %>%
  summarise(n=n()) %>%
  arrange(desc(n)) %>%

  ggplot(aes(reorder(TNT_NAME, n), n)) +
  geom_segment(aes(xend = TNT_NAME,
                   y = 0, yend = n)) +
  geom_point(size = 2) +
  theme_minimal() +
  labs(x = NULL, y = "Vacant Buildings") +
  theme(axis.text = element_text(color = "black"))
```

Null Geometry

Aggregate & Summarize



Questions?

Conclusion

- Areas with densely grouped vacant properties create persistent background exposure
- Higher density likely increases the availability of dust-lead
  - accumulates as it disperses from multiple structures
- Density of vacant properties has more explanatory power than demographic variables
  - SES non-significant predictor
- Clear opportunity for improving exposure mitigation

What now?

- Utopian solution: rebuild the city
  - New Orleans - 10 years post-Katrina
- Pilot interventions
  - Meaningful changes in exposure won’t be noticeable for years
- Can we sell public health interventions as economic development?