An analysis of residential home values after the Flint water crisis

Ilse Paniagua & Zeping Tao

Background & Motivation

- The **Flint water crisis** began in 2014, after the drinking water source for the city of Flint, Michigan was changed from Lake Huron and the Detroit River to the Flint River.
- Due to insufficient water treatment, **lead leached from water pipes into the drinking water**, exposing over 100,000 residents to elevated lead levels.
- **Previous studies** have explored the effects of the water crisis on children's lead levels (Hanna-Attisha, 2016) and depopulation in the city due to perceptions of unsafe water (Morckel and Greg Rybarczyk, 2018)

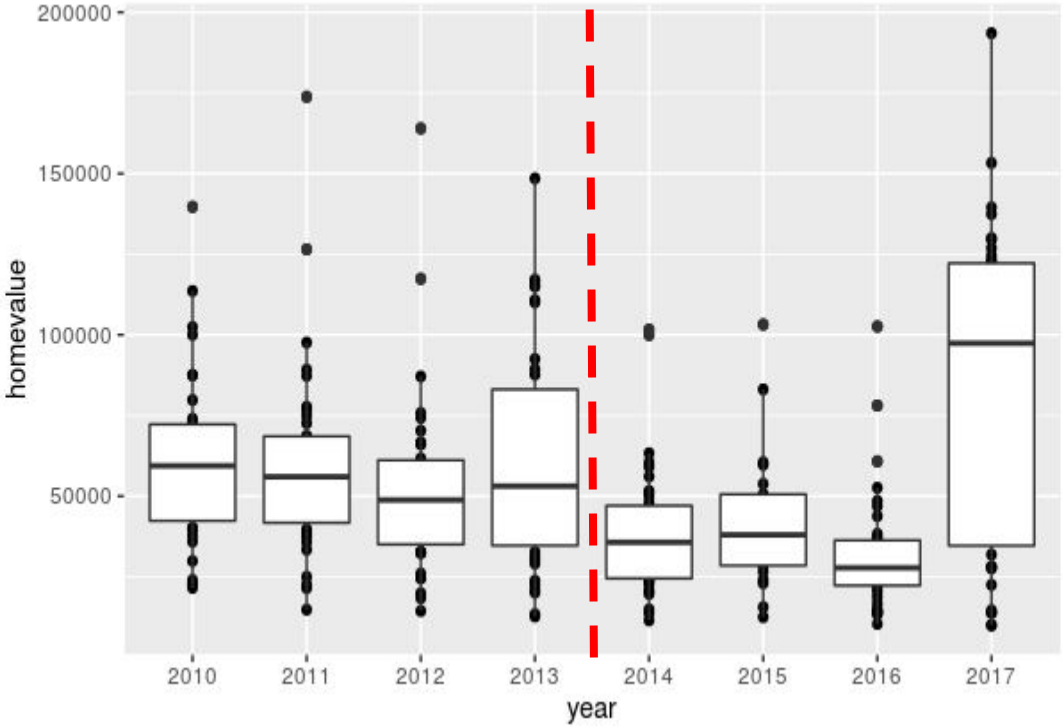
Credit: https://www.idcide.com/citydata/mi/flint.htm

Why alternative sources of data?

Limitations of self-reported administrative data

- **Measurement error:** Lack of knowledge, social desirability bias
- **Frequency:** Data on home values available on a yearly basis
- **Aggregation:** Average value across city (unable to study neighborhoods most affected)

Figure 1. Self-report home values in Flint, MI from ACS 2010-2017



Start of Flint Water Crisis

Research Questions

1. How did home values change in Flint, MI between 2010 and 2017?
2. Do self-report and Zillow estimates of home values agree?
3. Was the change in home values higher in areas with lead pipes?

Data gathering steps

- American Community Survey Data API (2010-2017)
  - API call for demographic variables
    - Racial composition
    - Age of housing structure
    - Number of children under 5
    - Percent under the poverty line
- Zillow API (www.zillow.com)
  - Gather list of random streets throughout Flint
  - API call for all home values in that street using regular expression (goal of ~100 Zillow prices per census tract)
  - Link addresses to census tract
- Location of lead pipes
  - GIS shapefile from University of Michigan-Flint

Appendix

Data Analysis Methods

Research Question 1:

- Wrangle Zillow data from xml format into tidy data format (variables: address, tax assessment, tax assessment date, lat and long, Zestimate)
- Geocode Zillow data to census tract
- For each census tract, calculate median sale price by year
- Analysis of variance (ANOVA) tests will be performed to test the differences between years
- Median sale price by tract will be mapped using ArcGIS Pro, combined with a shapefile of known lead pipes as of November 2016

Data Analysis Methods

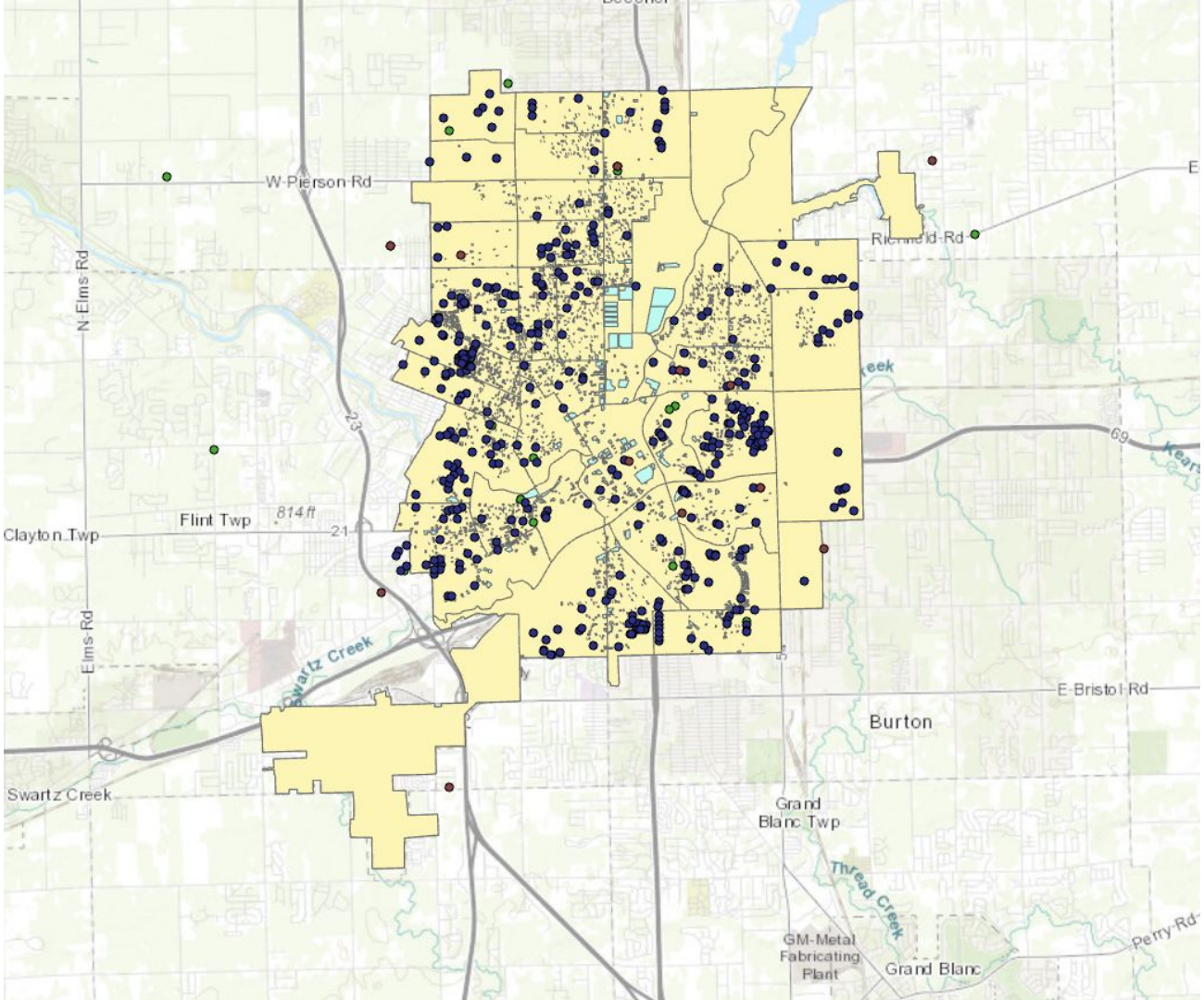
Research Question 2:

- Using the dataset described for research question 1
- Create categorical variables for median sale price based on the distribution of the ACS data
- By tract, create a dummy variable for whether Zillow data categories match the average ACS median sale price

Data Analysis Methods

Research Question 3:

- Geocode Zillow data using lat and lang
- Map Zillow data and lead pipes shapefile and using ArcGIS to create a map visualization


W Pierson Rd
N Elms Rd
Richfield Rd
Flint Twp
814 ft
Clayton Twp
Elms Rd
Swartz Creek
E Bristol Rd
Burton
Swartz Creek
Grand Blanc Twp
Thread Creek
GM-Metal Fabricating Plant
Grand Blanc
Perry Rd
23
69
21