

Problem Statement:

Can we accurately estimate financial property losses due to a disaster within a given zip code?

Data Acquisition:

We focused on acquiring our property data from Zillow's API. We used a Zillow API key to generate queries based on zip code. Obtained address data from openaddress.io and wrote code that will return a random sample of addresses within a given zip code. Run a query with the sample addresses to get property values using the Zillow API to obtain Zestimates. We then downloaded the occupancy of residential properties from census.gov.

WebApp Tool:

We designed a web application that will take in a given zip code and return a summary of financial estimates along with a map of the affected area. The application utilizes the Flask framework running on an AWS linux server. The application also uses Bokeh to generate a map with the zip code boundaries from Google maps API. The financial information is displayed in a data frame showing the mean property value, median property value, minimum property value, maximum property value, and estimated total damages. A histogram shows the distribution of Zillow Zestimate values within the affected area.

Limitations:

- The Zillow API only allows individual addresses to be queried and limits queries to 1000 per day.
- Because we can only get limited individual address data from Zillow, we have to sample addresses in the provided zip code, and this is computationally intensive and requires imputation of the remaining properties.
- We only have access to public address lists which do not provide 100% coverage for all zip codes (private lists may significantly improved coverage). Some areas have no address coverage at all.
- We rely on census data to estimate the remaining number of residential units. This data is also an estimate and is not timely.
- Our estimates are based on property prices being log-normally distributed. While this is relatively accurate, there are discussions in academic literature of more

accurate models for prices and scholars note that during periods of economic crisis prices become more difficult to model (see [T. Ohnishi et al., 2010](#))

Conclusions:

Our sampling method with only 100 samples, matches Zillow's home value index with an R2 score of 92%. We can effectively estimate in real time the financial losses of a zip code given an upcoming disaster.

Sources:

Slide Images:

Photo by Stephen Monroe on Unsplash

Photo by Ryan Wilson on Unsplash

Research:

<https://www.iii.org/fact-statistic/facts-statistics-us-catastrophes>