Extracting Building Values from Zillow

Katy Chow, Mason Childress,
Cesar Mendoza, Joey Ward

Executive Summary

During times of disaster it is fairly difficult to predict the amount of economic loss that will occur. Economic loss is not limited to physical infrastructure during times of disaster. It includes wages lost and other externalities that are challenging to measure. For the purpose of this project, we use a proxy equation to estimate economic lost given a disaster level parameter (ranging 0-10) and a list of zip codes. To make the information easily digestible, a tool was built to display demographic information on a map with a pop up over the specific area of interest. This allows for users such as government entities to input a zip code along with the disaster level parameter to understand the population within a geography and the possible losses if a certain type of disaster was to go through the area. We believe the value add in this tool is that insurance companies and other government entities can use our tool to prepare for the amounts of payout and grants that are used for rebuilding a community after a disaster. The users of this tool can also enrich the data with their own information and feed it into the display tool.

Problem Statement

Determine the amount of economic loss within residential areas due to natural disasters in specific zip codes

Collected Data

Data collected from:

- uszipcode python package
- Zillow Databases
- Openaddresses

Column Name	Data Type	Brief Description
lat	float	Latitude
lng	float	Longitude
median_home_value	float	Median home values from uszipcode python
median_household_income	float	Median household income
occupied_housing_units	int	Occupied housing units
population	int	Population size
Zip	str	Zip code
City	str	City Name
State	str	State
Median_Home_Value_Z	float	Median home values from Zillow Dataset
Top_Tier_Median_Home_Value	float	Top third median home values from Zillow Dataset
Bottom_Tier_Median_Home_Value	float	Bottom third median home values from Zillow Dataset

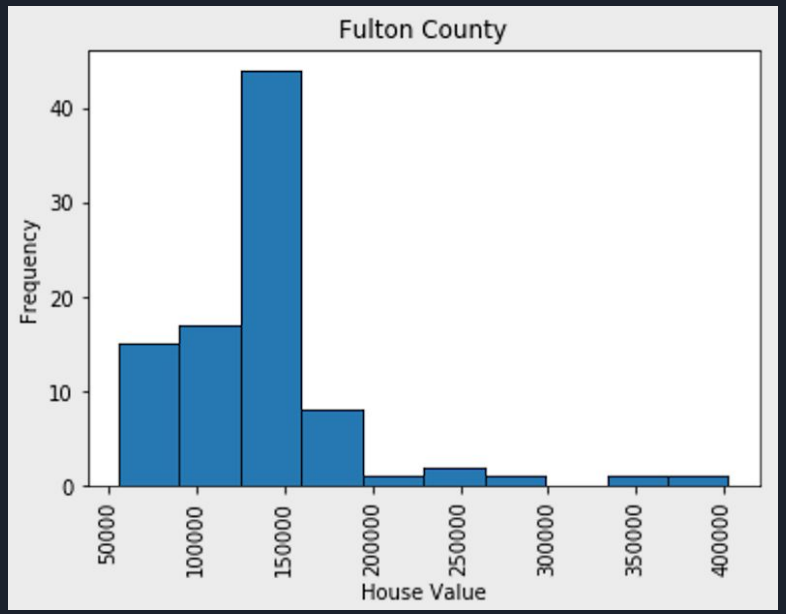
Zestimate API

Valuation based on list of addresses

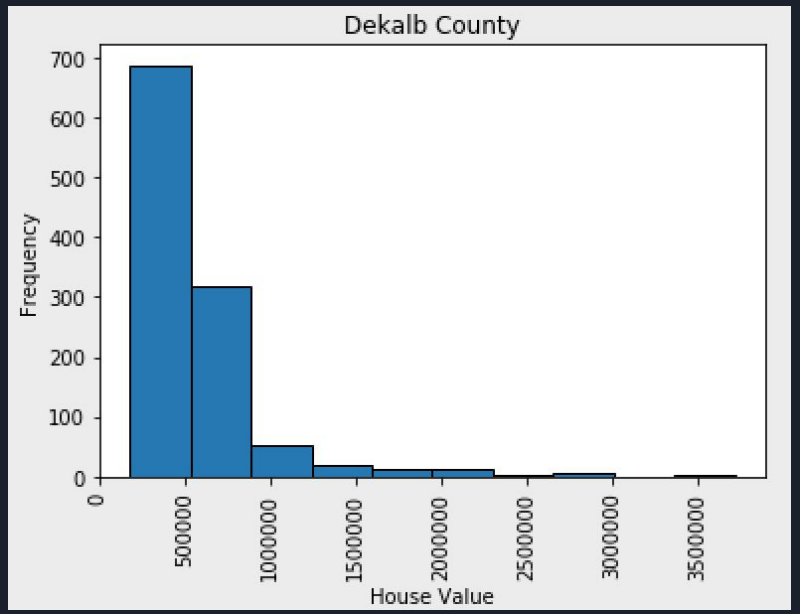
- The Zillow API allows for the user to upload a list of addresses which will be then matched onto the Zillow database to extrapolate housing values based on matches and estimate overall housing value within a zipcode
- From the API you can pull metrics such as Min, Max, Median, Skewness, Kurtosis, and other metrics while plotting them.
- Limitations to this API is the time it takes to fully estimate a given area is extremely long and not including data prep

EDA

Fulton county
(30310,30311,30314)



Dekalb county
(30307, 30306, 30317)



Estimating Economic Loss Function

Function : (takes disaster level 0 - 10)

- ***Economic Loss*** = *(% value of disaster damage)* x *Occupied_Housing_Units* x *Median_Household_Value_Z* + *(% value of disaster/2)* x *(1/12)*** x *Population* x *Median_Household_Income*
  - We are scaling disaster effects on income down by ½ due to the possibilities of transactions to occur via the web and workers that have remote options
  - Additionally, we make the assumption that at most 1 month of income would be loss in event of a disaster
  - We only compute economic loss due to occupied housing units because those effects are immediate. Unoccupied homes while effected, will not impact the population as much.

Walkthrough Economic Loss Function

Inputs:

Disaster Level = 5 (~ 50%)

Occupied Housing Units = 100

Median Household Values = 250000

Population = 10000

Median Household Income = 50000

Step 1: Calculate Infrastructure Losses:

- (0.5) x 100 x 250,000 = 12,500,000

Step 2: Calculate Income Losses:

- (0.5) x (0.083) x 10,000 x 50,000 = 20,833,333

Step 3: Sum the two losses:

- 12,500,000 + 20,833,333 = 33,333,333

Folium

Using a Python library called “Folium”, we were able to create an interactive map that allows the user to input a list of zip codes.

The function returns a map with an indicator where the zip codes are. When hovered over, the map will bring up a text box with statistics on the zip codes.

A demo is provided...

Future Improvements

1. Find a data source that has complete housing information base on zip codes
2. Enrich disaster economic loss function sliced by types of disasters.
3. Enrich display with Zestimate API data