***Mapping Eviction Trends in the City of Paterson* Data Discovery Project Blog**

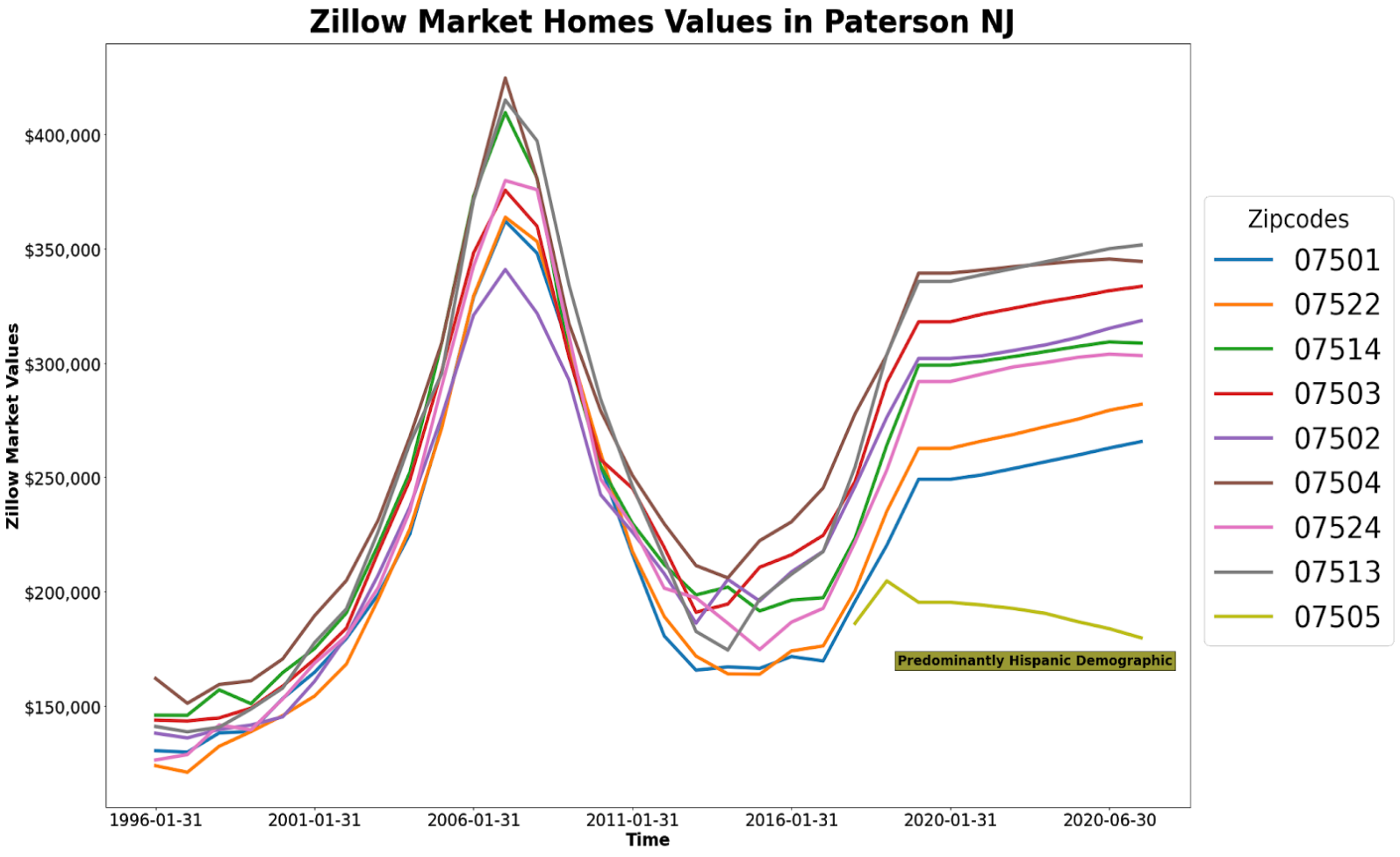
Research Mentor: Harsha Mallajosyula, Chief Data Officer, City of Paterson & Jigyasa Sharma, MPP, Goldman School of Public Policy

Contributors: Eva Sidlo, Rene Gamino, Nancy Xu

The City of Paterson, New Jersey wants to develop a dashboard that would map/visualize eviction trends in the City's neighbourhoods. The City aims to provide a spatial analysis of eviction trends to help organizations prioritize allocation of resources & address increased community needs. The goal of this study is to develop community and market insights into the issue of housing security in the City of Paterson and surrounding areas.

I. Market Values Analysis

Due to Zillow having insufficient data on Paterson NJ rentals we utilized Zillow's data on home prices. Our goal was to identify changes in home prices in light of Covid-19. As of January 2020 home market values in Paterson began to plateau in value except for one zip code. The '07505' zip code is a newly established zip code located in the inner city of Paterson that has no data previous to the year 2018. The line graph below shows the only zip code in Paterson NJ that has experienced decreased home market values is one that is largely composed of Hispanic populations [City of Paterson].



To get an idea of market rent values in the city of Paterson, NJ we webscraped Craigslist because Zillow rental data was unavailable in New Jersey. Web scraping Craigslist was proven to be difficult due to unreliable user inputs that had to be manually removed after scraping. To do this listings that were duplicated, not actually located within Paterson NJ, or that were listed substantially below the market value of a region were dropped. Through this process square footage, URL, name, price, number of bedrooms, Craigslist ID, creation date, and latitude/longitude for each Craigslist listing were collected. Web scraping Craigslist was also limited by the inability to gather listings that were taken down/deleted. To rectify these issues the processes were repeated over a period of 5 months (Sept. 2020 to Jan. 2021) to create a dataset of ~750 listings total. Once the listings were aggregated and cleaned the census Geographical ID and zip codes were deciphered using the 'censusgeocode' package in Python. This was done to join the dataset with the shape file that outlines the neighborhood boundaries by census tract. By grouping the data points by the Geographical ID and averaging the price listings we were then able to make a choropleth of Craigslist listing prices.

Contracted rents prevent displacement of communities. Web scraping Craigslist rentals can identify areas where rent is being inflated above median contracted prices. A choropleth of contracted median rents was created by the City of Paterson (Fig. A). The darker shaded blue regions were oriented with higher priced contracted rents. Grey areas indicate where data was not available. Areas with low contract median rents (light shaded regions Fig. A) that are correlated with high median Craigslist listing prices (dark shaded regions Fig. B) are key areas where gentrification may be occurring. Areas with low contracted rent are sensitive to gentrification. The red arrow on the maps below indicate an example of where gentrification could be occurring in Paterson. Pinpointing areas of potential gentrification will help the city allocate the appropriate resources to prevent displacement of low-income communities.

Fig. A: *[Credited to City of Paterson, NJ]*
Choropleth Map of Paterson NJ Median Rent Contracts:

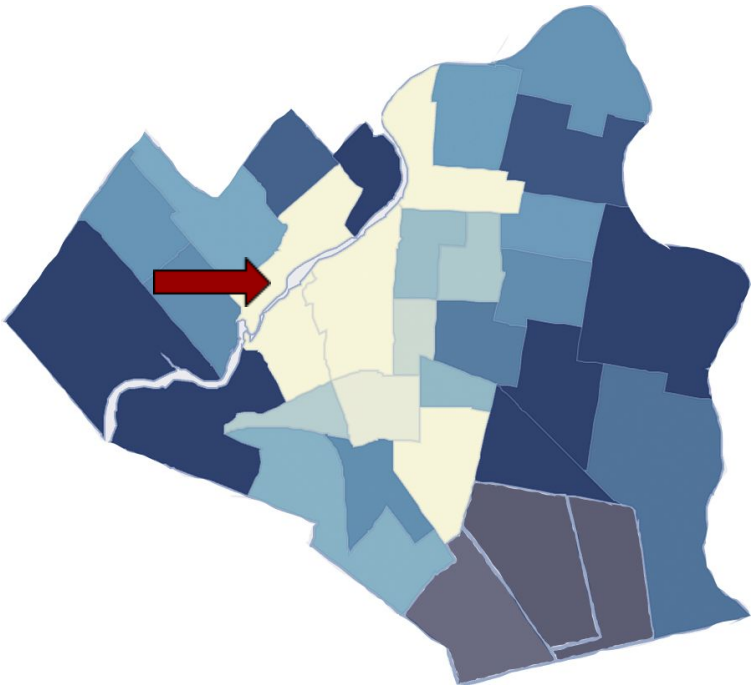
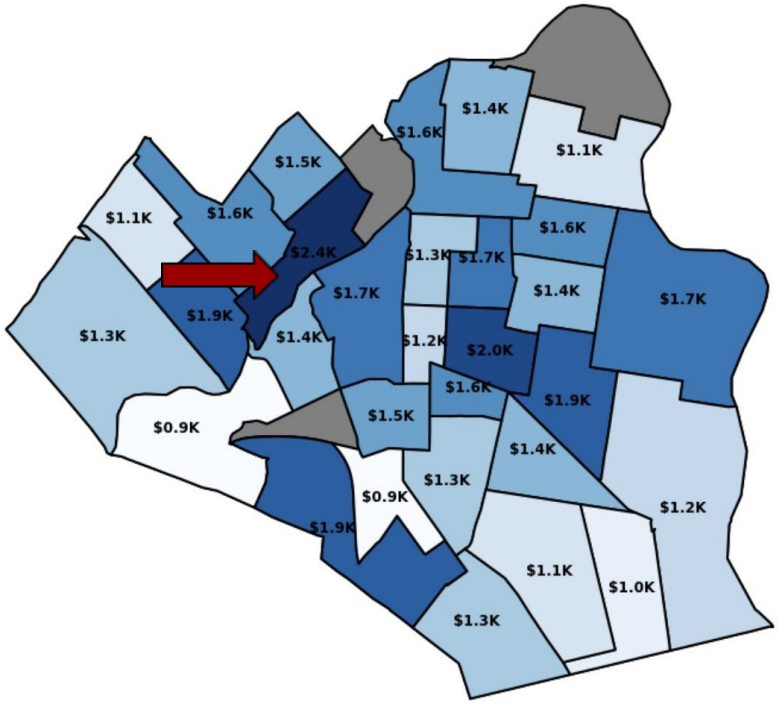


Fig. B: [From Sept. Nov. 2020]
Choropleth Map of Paterson NJ Median Craigslist Rental Prices:



II. Tweet Sentiment Analysis

We performed text and sentiment analysis of Tweets from Paterson and neighboring cities to explore and compare the change in public sentiment as a result of Covid-19 induced evictions. However, because we cannot personally encode each Tweet with a variable (i.e. a negative weight of -1 for overall negative sentiment), we navigated through the literature surrounding polarity analysis, which quantifies these sentiments as numeric values that are then referred as "polarity." Because of the compatibility with our aim and this toolbox, our priority for our sentiment analysis was to collect Twitter activity from users in the City of Paterson, render this activity usable for polarity analysis, and then prepare for visualization to finally explore and compare trends in public sentiment.

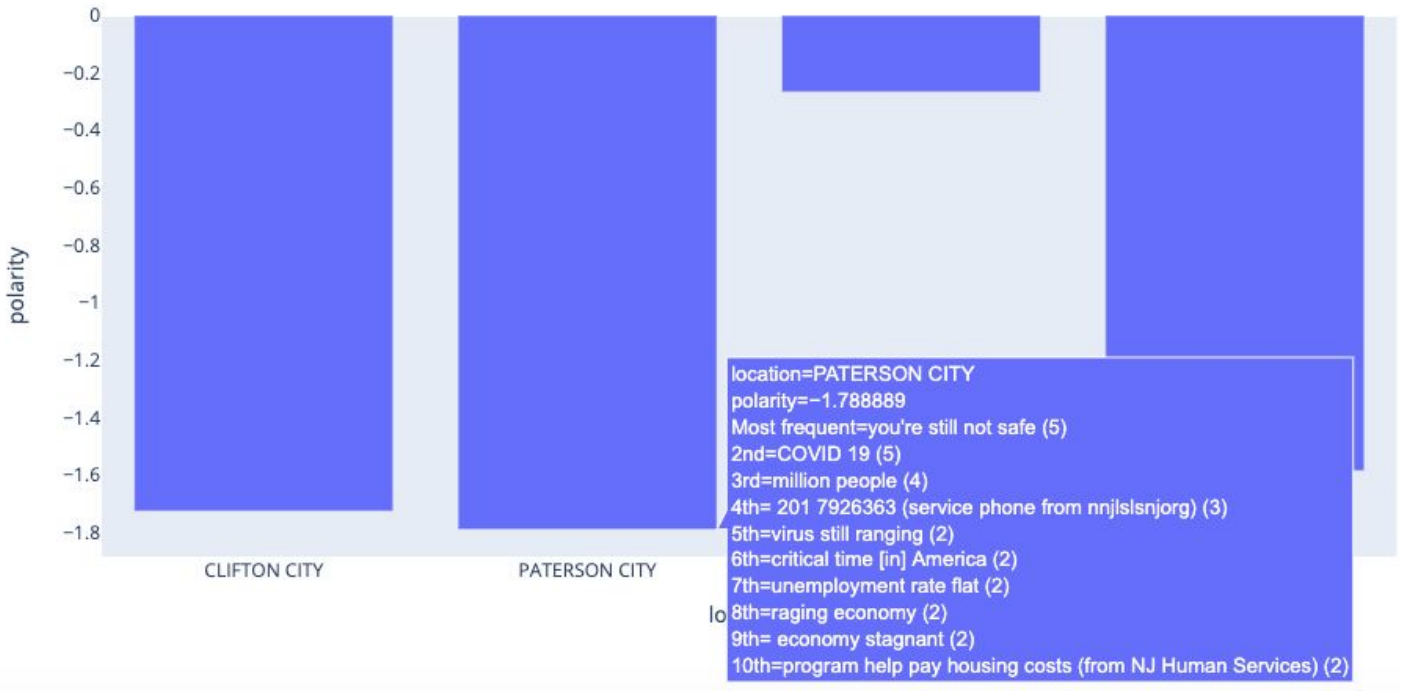
Data collection started by setting up target keywords/hashtags relating to evictions, in both English and Spanish and the topic of interests adjacent to housing security (Appendix A). These tags consist of commonly used phrases collected from performing searches on posts related to activism and housing security, both inside and outside of the City of Paterson. In addition, we gathered a list of the Twitter ID's for community leaders and non-profit organizations for the purpose of comparing the response of community leaders/population.

Our original goal was to collect content from the City of Paterson and then perform geospatial analysis using the assumed coordinates. However, we discovered through our exploration of the Twitter API that the use of coordinates as the location of Tweets were replaced with the nearest place or city. Twitter removed the search timeline function, which made it difficult to obtain Tweets from a few months ago needed for analysis on sentiment trends. Thus we shifted our focus to collect data for cities around Paterson, including Jersey City, Newark, and cities in Passaic County (Paterson and Clifton). Then, we used the gathered list of Twitter ID's for community leaders and nonprofits to extract their tweets using the Python library "tweepy". To collect tweets from residents, we used the "snscrape" library to first extract the twitter ID of eviction tweets based on the list of keywords/hashtags, and then collected the tweets using Twitter API and "tweepy".

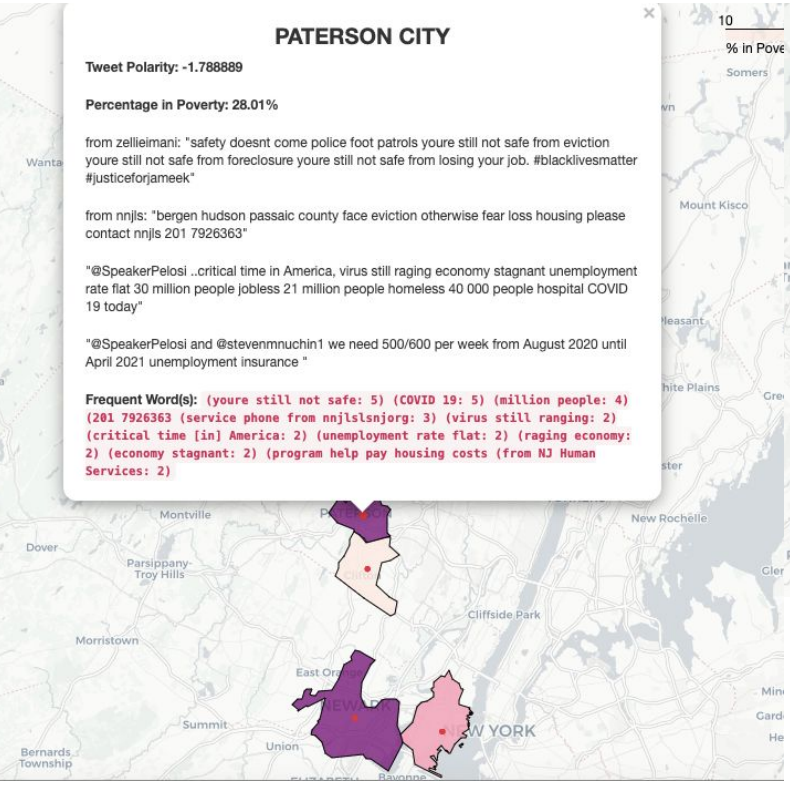
The data set has some sample bias due to the nature of social platforms: not every local non-profit organization or community leader advocating for housing protection is active on Twitter under a local chapter or small-scale organization, and not every resident facing eviction/having opinions about housing legislations is active on Twitter. Despite this, it's possible people who have strong opinions on evictions/eviction-related policies are more likely to post and engage with Twitter. Our sample underestimates those who have been evicted since the nation-wide response towards the COVID-19 pandemic, and overestimates the non-profit organizations due to the comparative ease in data collection.

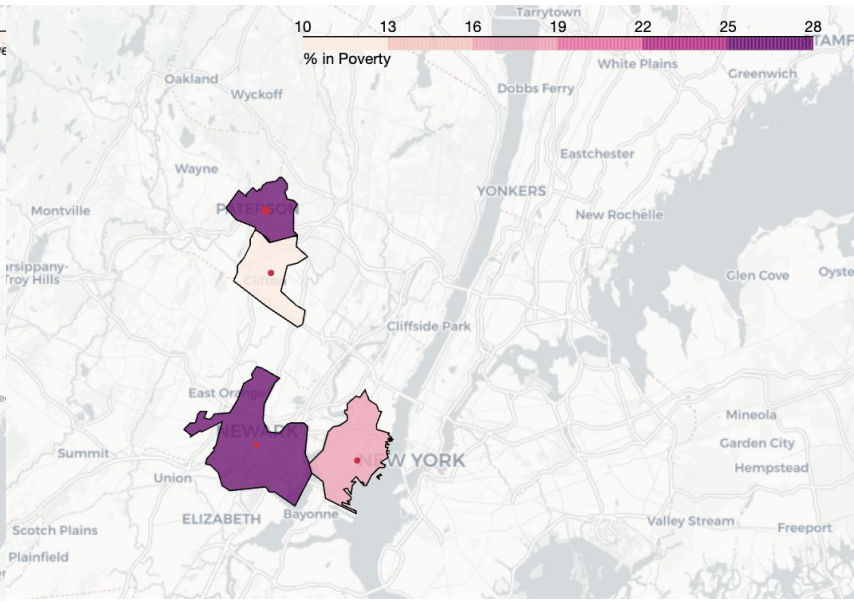
We analyzed the polarity of the tweets using "VADER" (Valence Aware Dictionary for Sentiment Reasoning) sentiment analysis, which allowed us to code sequences of stemmed words by their polarity scores. We used the Python library "wordcloud" to visualize the most common phrases in these datasets and based the size of these phrases on their frequency. We then manually removed phrases that were irrelevant/intelligible.

Fig. C. *The interactive bar graph shows the level of polarity for cities of interest with sufficient data. The frequency of the phrases are indicated with the number in the parenthesis.*



Polarity and Poverty Mapping





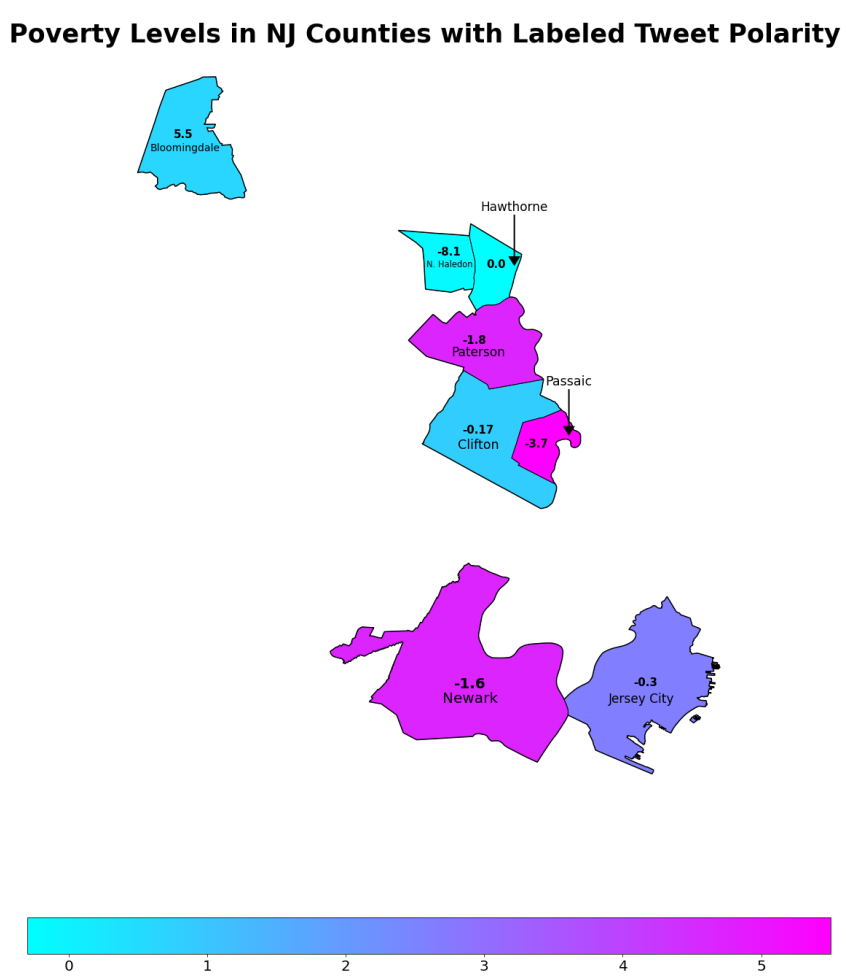


Fig. D: *The map above shows the relationship between poverty and tweet polarity in Paterson, Newark, Clifton, and Jersey City, and has pop ups that include segments of tweets and the frequency of different N-gram phrases of tweets in that city. The map to the right shows a static representation of the interactive map we made above, showing the relationship between polarity (labeled) and poverty (gradient).*

Clouds of the Most Frequent Words in Tweets of Nonprofits vs. Residents

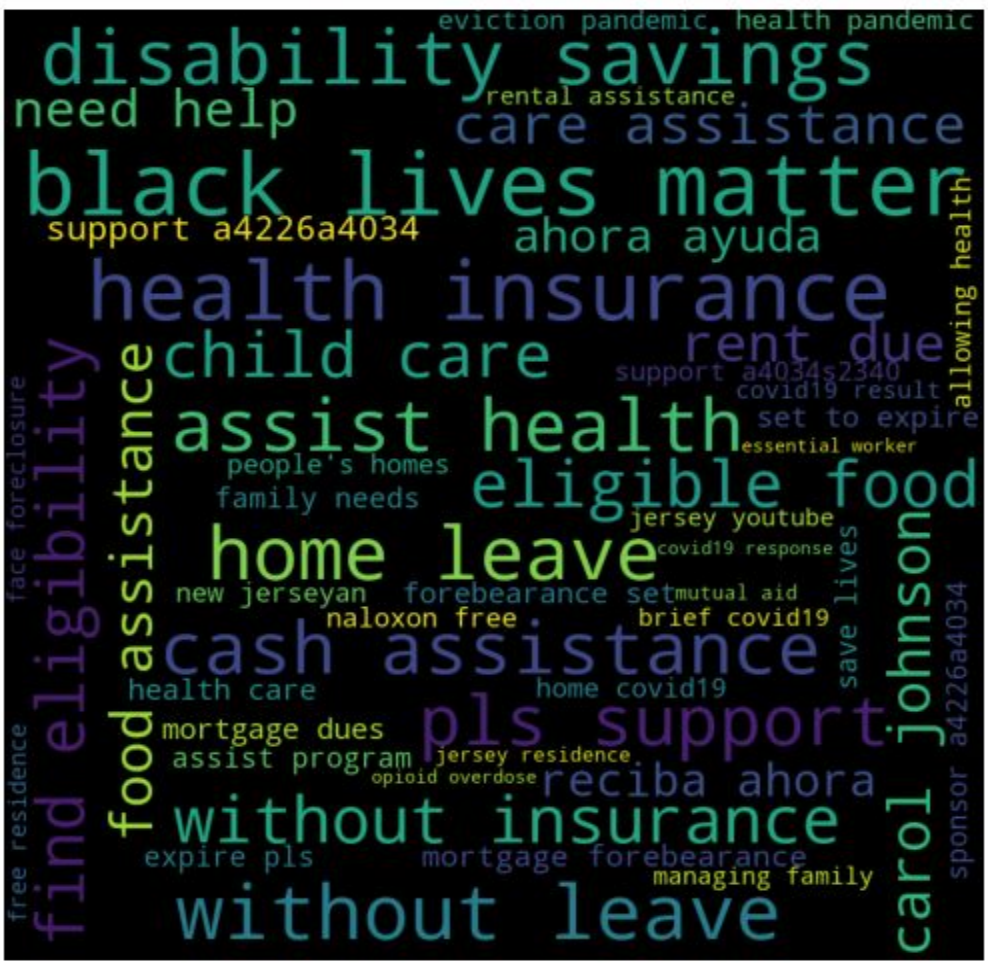


***Fig. E**: The infographic above displays the most frequent words that are present in the relevant Tweets corresponding to non-profit organizations and residents.*

Fig. F: *The infographic on the left displays the most frequent bigrams that are present in the relevant Tweets corresponding to the composite collection of Twitter activity, from both non-profit organizations and the residents.*

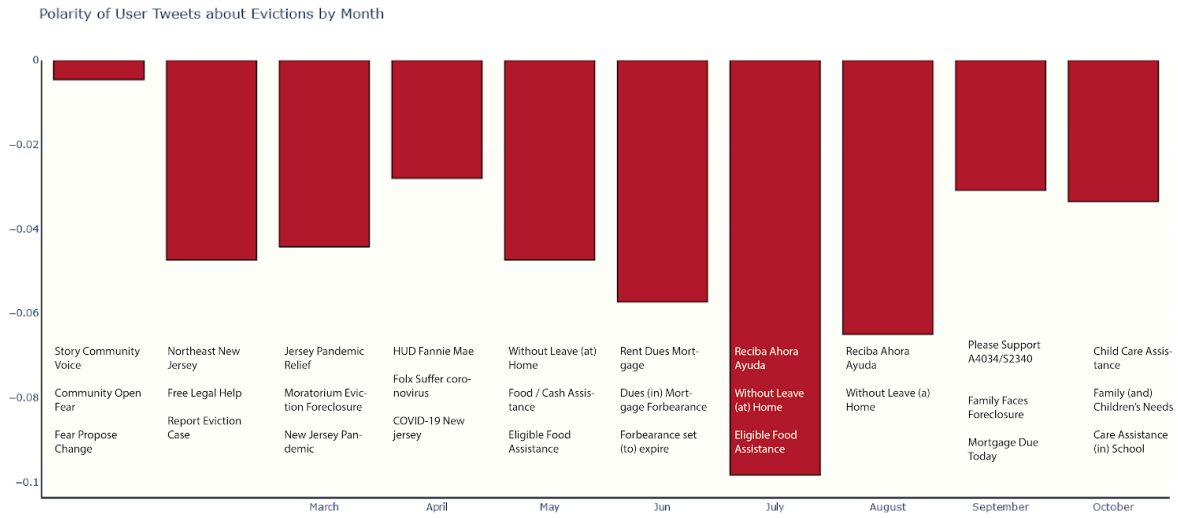


Fig. G *The bar graph above displays the relationship between the month for the year 2020 and the aggregate value (mean) for the composite polarity score assigned to each trigram. Each "bar" is modified to contain the most frequent trigrams for that month.*

We arrive at the following conclusions from our analysis:

- First, as shown by Fig. D, cities with higher poverty levels, such as Paterson and Newark, tend to have more negative sentiment towards evictions and topics related to evictions.
- Second, as shown by Fig. E, residents tend to tweet more frequently about reactions to local legislations (such as "moratorium" referring to the eviction moratorium, or the bill "a4034", which authorizes the governor to reduce rent payment responsibility and allow mortgage forbearance).
- Third, the change in polarity over the months (Fig. G) could reflect the effect of public policy. For instance, the eviction moratorium was issued by the New Jersey Governor in March, and this could be related to the slight increase of polarity in April. On the other hand, polarity decreases dramatically (more negative sentiments) in June and July, as June has the highest unemployment rate (16.8%) in New Jersey out of all the months.

Our visualizations can be useful to policymakers in gauging city-level sentiment regarding to COVID-19 pandemic and its effect on housing security. It is unsurprising that throughout our visualizations, the Twitter activity partially captured the social impact of the eviction moratorium executive order issued in March by the New Jersey Governor and the residents' support for the legislature NJ A4034, which authorizes the governor to reduce rent payment responsibility and allow mortgage forbearance.

Appendix A. Twitter Hashtags and Nonprofits

Twitter Accounts

English Topics/ Hashtags	Spanish Topics	Spanish Hashtags	Nonprofits Twitter Accounts
Evictions Homelessness Homeless Tenants AffordableHousing CancelRent RentStrike TaxTheRich EvictionMoratorium RentMoratorium EvictionFree NoEvictions Housing assistance Housingrelief Rentstrike Housingcrisis Rentburden HousingForAll Evictionfreeze HousingAffordability EquitableDevelopment Unemployment Foreclosure SuspendRent RentFreeze Displacement RentalAssistance feedthehungry patersonstrong HousingIsHealth localfoodbank foodbank fooddrive endhunger	desalojo desahucio casa domicilio sin hogar sin techo inquilin tenencia arrenda arriendo arrendamiento arrendatari residente vecino huésped persona cubierta alberg amparo asilo resguardar encubrir dueño casero propietario posadero mesonero asequible alquilar gentrificación asistencia moratoria congelar dipacement alojamiento propiedad huelga ayuda	CancelLaRenta SinDesalojos CancelLaRenta SinDesalojos ComidaSiRentaNo NoCobramosNoPagamos RentaEstabilizada HuelgadeAlquileres CancelenDeudas LaRentaNoSeCome CansadosDeLimosnas QueremosJusticia ImpactoLatino	St.Paul's Community Development Corporation https://twitter.com/stpaulscdcnj Table To Table https://twitter.com/TableToTableOrg CUMAC https://twitter.com/CUMACFeeds Oasis Paterson https://twitter.com/OasisPatersonNJ Eva's Village https://twitter.com/EvasVillageNJ Habitat for Humanity https://twitter.com/PatersonHabitat Zellie Thomas - activist in Paterson https://twitter.com/zellieimani NE NJ Legal Services https://twitter.com/NNJLS Fair Housing Council of North NJ https://twitter.com/nnjfairhousing Housing/community Development Network of NJ https://twitter.com/HCDNNJ NJHumanServices https://twitter.com/NJDHS NJ Relief Fund https://twitter.com/NJ_PRF Paterson Times https://twitter.com/PatersonTimes