Impacts of ShotSpotter on Housing Prices in Cambridge, MA

Sungwoo Noh, Suraj Godithi, Yun Wai Wai Oo(Melody)
CDS DS 539:Spark! Data Science X-Lab Practicum
December 20, 2024

Contents

Executive Summary

This study investigates the socioeconomic impact of the ShotSpotter gunshot detection system in Cambridge, MA, with a particular emphasis on its relationship to housing prices and demographic changes. Using data from the BridgeStat crime reports, the Cambridge Property dataset, and additional context from city council reports, this project conducts a detailed analysis of correlations between ShotSpotter deployments, and changes in property values. By examining these relationships, this study aims to provide data-driven insights to inform advocacy efforts led by The Black Response, with the goal of either challenging the continued use of ShotSpotter in Cambridge or offering critical guidance to other municipalities deploying or considering this technology.

Introduction

According to ShotSpotter, “ShotSpotter is gunshot detection technology that uses sophisticated acoustic sensors to detect, locate and alert law enforcement agencies and security personnel about illegal gunfire incidents in real-time. The digital alerts include a precise location on a map (latitude/longitude) with corresponding data such as the address, number of rounds fired, type of gunfire, etc. delivered to any browser-enabled smartphone or mobile laptop device as well as police vehicle MDC or desktop.” #StopShotSpotter is a campaign deeply concerned about its harmful impacts on the Black, brown, and poor people that the company surveys and wants to build community-led responses to address the root causes of gun violence.

Questions & Hypotheses

We seek to analyze the socioeconomic and public safety impacts of ShotSpotter deployments in Cambridge, MA, focusing on their relationship with housing prices, demographic shifts, and perceptions of safety.

Our primary hypothesis is that neighborhoods with increased ShotSpotter deployments exhibit correlations with rising housing prices and demographic changes, which may be driven by heightened surveillance and perceived safety improvements. However, we also hypothesize that these benefits are counterbalanced by concerns about over-surveillance, misidentification, and the disproportionate targeting of marginalized communities, which could accelerate displacement and gentrification.

To explore these hypotheses, we propose the following research questions:

Housing and Demographics

1. **Housing Impact**: What is the correlation between ShotSpotter deployments and changes in housing prices in Cambridge neighborhoods?
2. **Demographic Shifts**: How do ShotSpotter deployments affect neighborhood demographics, particularly in historically low-income or public housing areas?

Public Safety Effectiveness

3. **Gun Violence Reduction**: How effective is ShotSpotter in reducing gun violence?
4. **Accuracy and Reliability**: What is the accuracy of ShotSpotter in detecting gun-related incidents?

City Policy and Community Advocacy

5. **Advocacy for Termination**: What evidence-based arguments and data-driven insights are required to inform advocacy efforts aimed at ending the ShotSpotter contract?

Methodology

We will utilize data from the **BridgeStat crime reports** and the **Cambridge Property Dataset** covering fiscal years 2016 to 2024. This analysis will enable us to evaluate the relationship between ShotSpotter deployments and shifts in socioeconomic metrics, public safety performance, and policy developments. The results of this study will offer data-driven insights to guide informed decision-making and support advocacy efforts.

Methodology & Workflow

In order to evaluate this hypothesis, we started by converting Bridgestat data from PDF to CSV format. This process involved using text extraction tools to parse the structured and unstructured data within the PDFs, ensuring that the data was correctly formatted and organized for analysis. Each dataset was carefully reviewed to resolve inconsistencies and errors arising from the extraction process, such as mismatched columns, missing values, or erroneous entries.

To prepare the **Cambridge Property Dataset** for analysis, data from fiscal years 2016 through 2024 was merged into a single comprehensive dataset. The process began with data cleaning, which involved removing records with null values to ensure data integrity and standardizing the address format across all years for consistency. Once cleaned, the yearly datasets were integrated and sorted chronologically by the "Year of Assessment" to facilitate temporal analysis. Key columns such as property addresses, assessed values, building values, land values, and coordinates were retained, along with classifications like property type and zoning, to provide context for property trends. This prepared dataset serves as a robust foundation for exploring correlations between housing prices, property assessments, and ShotSpotter deployments, supporting both statistical and spatial analyses.

Blockers

Our largest barrier has been data collection and data cleaning BridgeStats data from pdf to csv, which has been the main focus of this semester. ……..

From the **2013–2016 BridgeStat data**, there is no specific information on gunshot incidents; the data only includes records of murders. In contrast, the 2017 dataset contains gunshot incident data but is incomplete, with some records missing key details such as location, time, and date. Keywords such as "gunshot," "shots fired," "shots on," "homicide," "firearm," and "gun" were used to search and extract relevant data, which was subsequently compiled into a CSV file and uploaded to GitHub. Due to these limitations, it is expected to be challenging to conduct an in-depth comparative analysis between the 2013–2017 BridgeStat data and the more comprehensive 2018–2024 dataset.

Results

Data Visualization

As part of our preliminary analysis, we created the visualization **"Average Total Assessed Value of Assets Over Time"** to examine trends in property assessments in Cambridge from 2016 to 2024. The x-axis shows the years of assessment, while the y-axis represents average assessed values. The blue line illustrates the mean values, and the red dashed line represents the median, both showing a consistent upward trend, reflecting rising property values. A slight fluctuation in median values is observed around 2020–2021 before resuming the overall growth trajectory. Shaded areas around the lines indicate variability in the data. The visualization highlights a steady increase in assessed values over the years, with mean and median trends closely aligned. Based on these insights, we plan to explore factors driving the rise in housing prices.

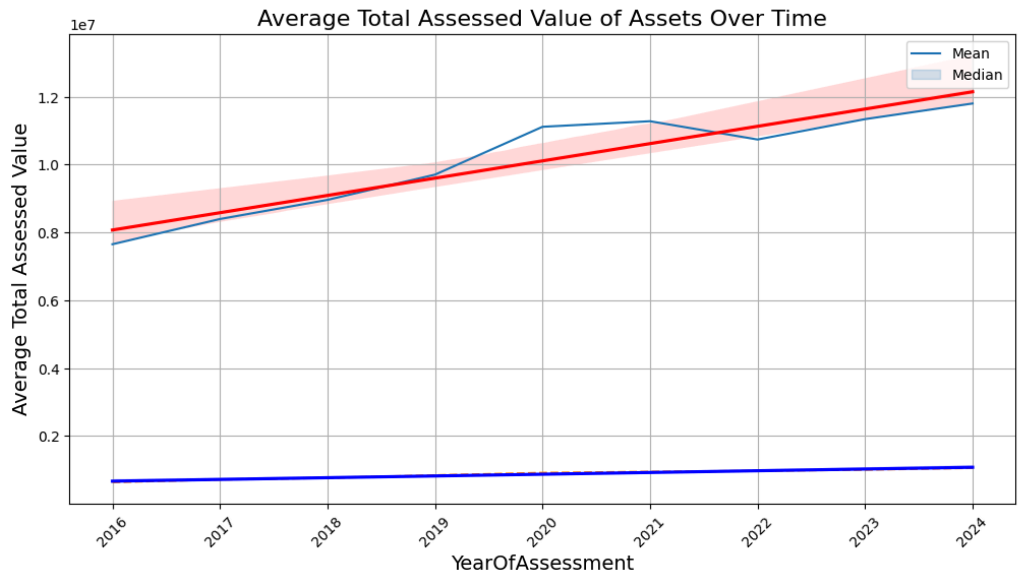

Figure 1. Average Total Assessed Value of Assets Over Time (2016–2024)

We also visualized the top 10 locations in Cambridge with the highest frequency of gunshot incidents based on the ShotSpotter data. The x-axis lists the specific locations, while the y-axis represents the count of incidents. Locations such as "Cherry St & School St" and the "100 block of Harvard St" show the highest counts, each with three incidents. Other notable locations include "Pine St & Washington St," "Market St & Windsor St," and "Clement Morgan Park & 600 block Main St," each with slightly fewer incidents. This visualization provides a clear view of hotspots for gunshot activity within the analyzed dataset.

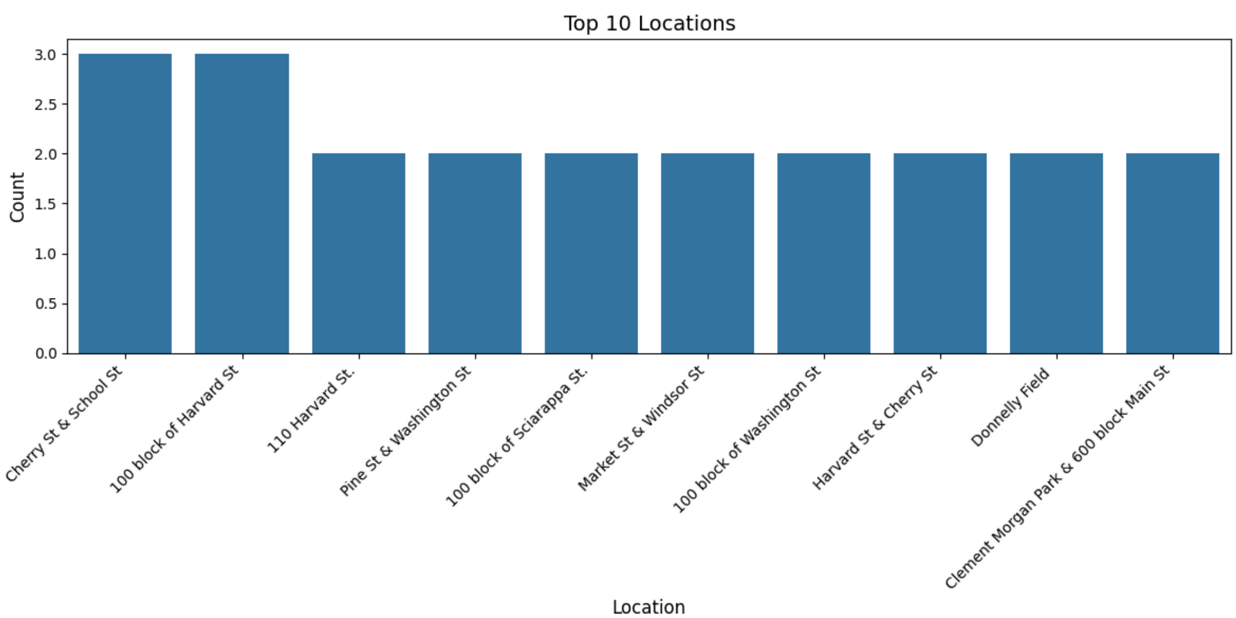

Figure 2: Top 10 Locations for Gunshot Incidents

In Figure 2, we compared average housing prices in neighborhoods with ShotSpotter locations between 2015 and 2023. Prices in these areas rose comparably to the broader city trends, suggesting no distinct pattern unique to ShotSpotter placements that would confirm our hypothesis. However, this appearance is misleading, as further analysis reveals.

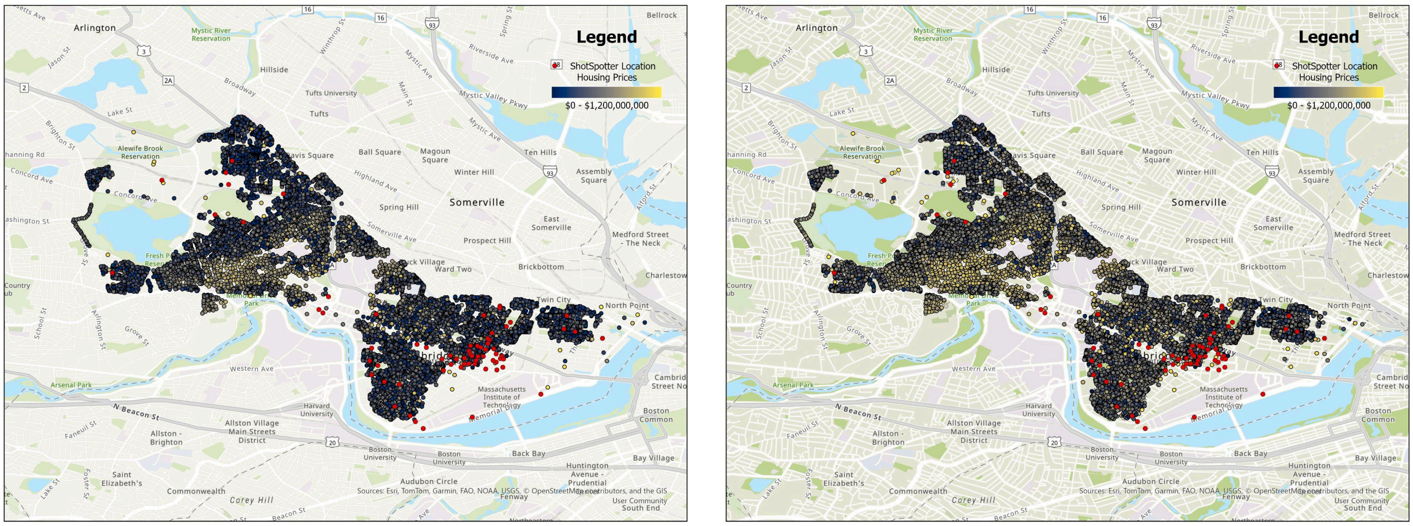

Figure 3. ArcGIS Pro screenshot of housing prices over time by Cambridge neighborhoods(Left: 2016, Right: 2024)

The above maps visualize the correlation between ShotSpotter deployment locations (marked in red) and housing price trends in Cambridge neighborhoods over time. The gradient shading represents housing prices, ranging from $0 (dark blue) to $1,200,000,000 or more (gold). The maps show how housing prices have evolved from 2016 (left) to 2024 (right) while highlighting the areas where ShotSpotter has been deployed. Between 2016 and 2024, areas shaded in gold, indicative of high housing prices, have expanded significantly. Neighborhoods that were previously in mid-range price tiers (yellow or light blue) are now dominated by high-value properties. The red dots show a concentration of ShotSpotter deployments in neighborhoods that initially had lower housing prices in 2016. Over time, these areas have seen an upward shift in housing prices, as evidenced by the spread of gold shading in 2024. The overlay of ShotSpotter deployments with housing price increases suggests a potential relationship. Deployment locations coincide with historically low-income neighborhoods, which may now be experiencing gentrification or increased investment, leading to higher housing prices.The most significant housing price increases appear to occur near major urban hubs and transportation corridors, where ShotSpotter deployment is dense. This could reflect broader economic trends exacerbated by increased surveillance and public safety investments.

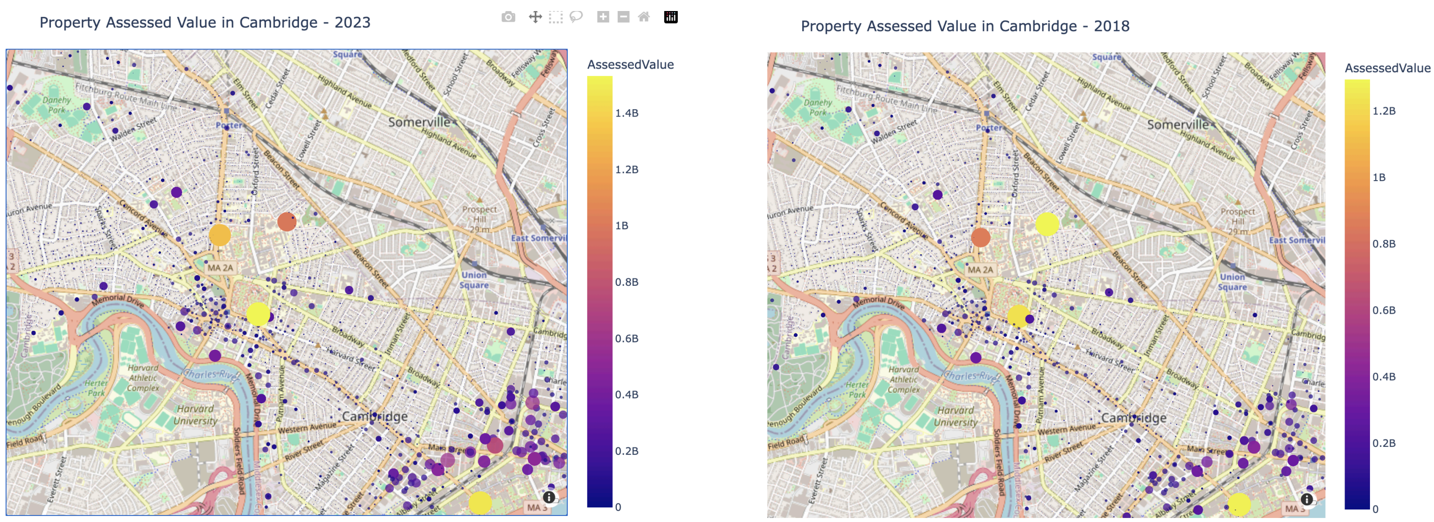

Figure 4. Housing Prices in Cambridge Over Time: 2018 to 2023

Looking at the map showing housing prices, where each dot represents a property’s value, the closer the color is to yellow, the higher the property value. Similarly, larger dots indicate higher values. From 2018 to 2024, we can observe that the yellow-colored dots, representing higher housing prices, have expanded outward, indicating an increase in property values over time. By 2023, clusters of high-value properties had extended further outward, encompassing more areas of central Cambridge and some surrounding neighborhoods.

Long-term residents in these areas may face the risk of displacement as housing costs continue to rise. The combination of increasing housing prices and surveillance tools like ShotSpotter could contribute to a cycle where marginalized groups are pushed out, reducing the city's diversity.

Examining the final map, where properties are categorized by their type and the size of dots represent their prices, it is evident that these areas are predominantly residential. This suggests that the people living in these neighborhoods are likely to experience potential displacement due to the rising cost of housing.

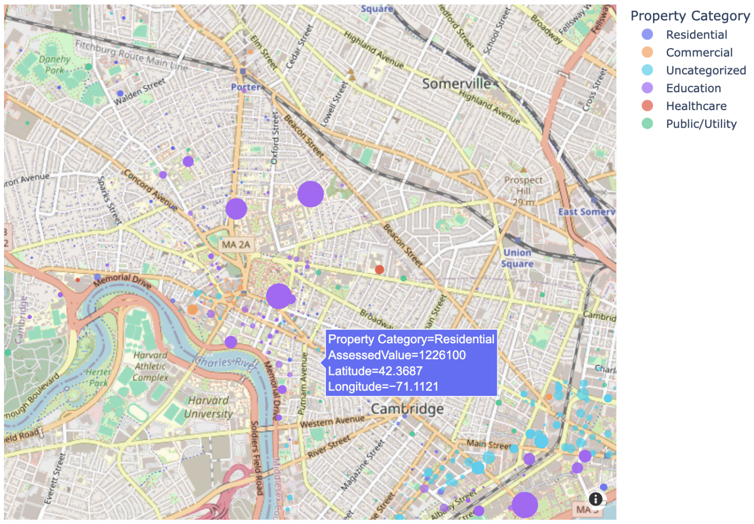

Figure 5. Property Locations in Cambridge by Categories

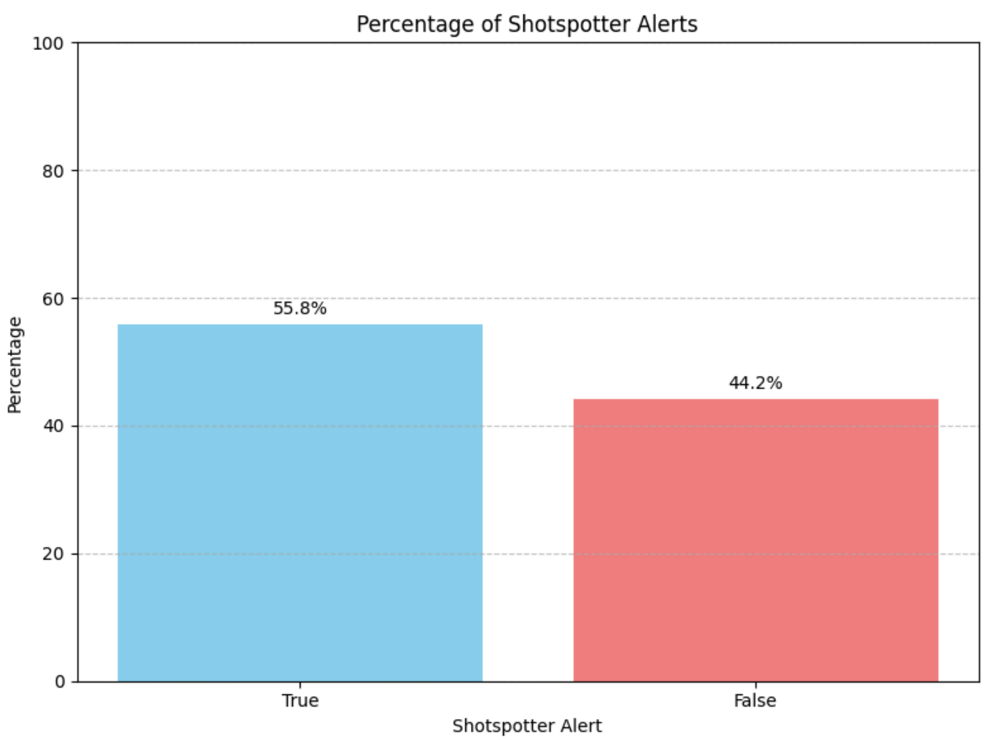

Figure 6. Accuracy of ShotSpotter in Detecting Gunshots

This graph illustrates the actual accuracy of ShotSpotter in detecting gunshots compared to its claimed performance. While ShotSpotter advertises a detection accuracy of 97%, the analysis reveals that it only detects gunshots 55.8% of the time, with 44.2% of incidents going undetected. This discrepancy highlights significant limitations in the system's reliability and raises concerns about its effectiveness as a public safety tool. The substantial gap between reported and actual accuracy underscores the need for independent validation of ShotSpotter's claims. Furthermore, false negatives could undermine public trust in the system and affect the timely response to gun-related incidents

Correlation Analysis

To explore the potential relationship between housing prices and ShotSpotter implementation, we conducted an OLS regression analysis. This method was chosen for its simplicity, which allows for straightforward inference.

Given the significant imbalance in size between the housing price data and the ShotSpotter data, we employed stratified sampling on the housing price dataset. This approach allowed us to match the size of the ShotSpotter data while maintaining the same proportional distribution of the housing price data. We believed this was the most suitable method given our limitations, as it helps address data imbalance and ensures a baseline level of validity for our data and analysis.

The independent variables in the analysis included land area, sales price, ShotSpotter alert status, and whether an arrest occurred. The dependent variable was assessed value, defined as the sum of land value and building value, which serves as a key indicator of housing price changes and a potential signal of gentrification. As a supplementary analysis, we conducted a two-tailed t-test to compare two groups: one where the ShotSpotter alert was 0 (indicating no detection of gunshot sounds) and another where the ShotSpotter alert was 1 (indicating successful detection).

Starting with the regression model, the p-value for the F-statistic is 0.000, indicating that at least one of the independent variables has a statistically significant effect on the dependent variable. The R-squared value is 0.6173, suggesting that approximately 62% of the variation in assessed value is explained by the independent variables, making the model reasonably reliable.

In our model, the ShotSpotter alert variable is dummy-coded, with the group where ShotSpotter = 0 serving as the reference category. For the group where ShotSpotter = 1, the coefficient is -1,174,053, suggesting a negative relationship with assessed value. However, its p-value of 0.870 indicates that this result is not statistically significant. Similarly, the arrest variable also has negative coefficients across the board but with p-values well above the common alpha level of 0.05, rendering these variables statistically insignificant.

While land area and sales price are included in the model, they were not analyzed further, as our project's primary focus is the relationship between ShotSpotter and housing prices rather than the broader factors influencing housing price changes.

Source	SS	df	MS
Model	1.2588e+17	6	2.0980e+16
Residual	7.8056e+16	79	9.8805e+14
Total	2.0394e+17	85	2.3992e+15

Number of obs	=	86
F(6, 79)	=	21.23
Prob > F	=	0.0000
R-squared	=	0.6173
Adj R-squared	=	0.5882
Root MSE	=	3.1e+07

assessedvalue	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
1.user_shotspotter_alert	-1174053	7138067	-0.16	0.870	-1.54e+07	1.30e+07
user_arrests						
1	-6751922	9935291	-0.68	0.499	-2.65e+07	1.30e+07
2	-5810392	1.35e+07	-0.43	0.668	-3.27e+07	2.11e+07
3	-7431897	2.30e+07	-0.32	0.748	-5.33e+07	3.84e+07
landarea	2472.06	227.9881	10.84	0.000	2018.261	2925.859
saleprice	-2.669639	3.22258	-0.83	0.410	-9.084025	3.744746
_cons	45766.26	6358092	0.01	0.994	-1.26e+07	1.27e+07

Figure 7. Regression Model to analyze relationship between Shotspotter Detection and Housing Prices

Two-sample t test with equal variances

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]
0	38	1.87e+07	1.19e+07	7.31e+07	-5351751	4.27e+07
1	48	1563556	341971	2369244	875599.4	2251513
Combined	86	9118210	5281872	4.90e+07	-1383566	1.96e+07
diff		1.71e+07	1.05e+07		-3852878	3.80e+07

```
    diff = mean(0) - mean(1)                                      t =   1.6229
H0: diff = 0                                     Degrees of freedom =       84

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9458         Pr(|T| > |t|) = 0.1084          Pr(T > t) = 0.0542
```

Figure 8. T-test Result: Group 0(No Detection) and Group 1 (Detection) for mean of Assessed Value of Properties

While our regression model proved to be ineffective in establishing a relationship between ShotSpotter and gentrification, the t-test yielded some meaningful findings that suggest a potential connection. In our t-test, the null hypothesis was that there is no difference in assessed value between properties in group 0 (no ShotSpotter detection) and group 1 (ShotSpotter detection). We tested three alternative hypotheses: (1) group 1 has a higher mean assessed value than group 0, (2) group 0 and group 1 have different assessed values, and (3) group 0 has a higher mean assessed value than group 1.

The first and second hypotheses failed to reject the null hypothesis, with p-values of 0.9458 and 0.1084, respectively, exceeding common significance levels of 0.1, 0.05, and 0.01. However, the third hypothesis successfully rejected the null hypothesis at the common alpha level 0.1, closer to 0.05 with a value of 0.0542. This suggests that properties where ShotSpotter is not detecting incidents have higher assessed values compared to properties where ShotSpotter is active.

These results provide several insights. First, ShotSpotter may not be effective in maintaining or increasing property values. The t-test results suggest a potential trend where properties in areas without ShotSpotter detection have higher assessed values than those with detection.

Second, the presence of active ShotSpotter detection could indicate gun-related incidents in the area, which may negatively impact housing prices. The presence of active ShotSpotter detections (Group 1) could imply frequent gun-related incidents in those areas, which might contribute to lower property values due to a perception of higher crime and reduced safety.

Lastly, properties in areas without ShotSpotter detection have higher assessed values, potentially suggesting a perception of safety or desirability in those areas. Properties in areas without ShotSpotter detection (Group 0) tend to have higher assessed values. This could suggest that these areas are perceived as safer or more desirable. However, this interpretation should be taken with caution, as the analysis does not directly account for other factors that may influence property values, such as socioeconomic status, neighborhood characteristics, or the presence of other crime-prevention measures.

However, these insights are counter intuitive to our initial expectation that shotspotter increases housing prices, ultimately leading to gentrification. However, This is likely due to data limitations and because we examine areas where ShotSpotter is working versus where it is not, rather than a direct comparison between areas with and without ShotSpotter deployment. So, we should not conclude here, rather treat it as one of the possibilities. Further research and the collection of more comprehensive and reliable data are necessary to validate or challenge these results.

Limitations

There are significant limitations in our data and analysis that must be addressed to draw more valid conclusions. First, we did not directly analyze the presence or absence of ShotSpotter deployment but instead focused on whether ShotSpotter detected gunshots. This creates a significant limitation, as the model outcomes are more reflective of whether non-functioning ShotSpotter systems correlate with higher housing prices rather than whether ShotSpotter itself impacts housing prices.

Moreover, there is a clear gap between ShotSpotter data and housing price data. We faced challenges in matching the deployment locations of ShotSpotter systems to nearby properties. Without precise spatial matching, it remains unclear which properties fall within the effective coverage area of ShotSpotter systems. As a workaround, we used detection status as a proxy, but this approach introduces ambiguity.

Additionally, there was a notable imbalance between the ShotSpotter data and the housing price data, with ShotSpotter data being disproportionately smaller. While stratified sampling was applied to balance the datasets by extracting observations from housing price data to match ShotSpotter data, we cannot ensure that this fully mitigated bias in the analysis.

Furthermore, while these insights reflect observable patterns in the data, it is crucial to acknowledge the limitations of the analysis. The data does not establish a causal relationship between ShotSpotter and property values. Instead, the observed trends could be influenced by other variables not accounted for in this analysis. Additionally, this is not a direct comparison of ShotSpotter deployment versus non-deployment but rather a comparison of areas with and without ShotSpotter detections.

Suggestion

Given these limitations, I strongly recommend ensuring accurate matching between ShotSpotter deployment locations and nearby properties. Also, clearly defining and incorporating the distance or coverage radius of ShotSpotter systems in relation to gunshot incidents and properties can be useful in validating analysis. Finally, resolving the data imbalance between ShotSpotter and housing price datasets is necessary to reduce potential bias. By addressing these limitations, future analyses can provide more reliable and actionable insights regarding the relationship between ShotSpotter and housing prices.

We need a direct comparison between areas with ShotSpotter deployment and those without it to address these issues and generate valid and meaningful insight. It is essential to consider this limitation and incorporate it into their project design, further research, and modeling to ensure more robust and accurate conclusions. To resolve this, as data imputation, we initially considered filling in missing ShotSpotter data with a default assumption of "no ShotSpotter." However, I realized that spatial information about ShotSpotter coverage—such as the range within which it detects gunshot incidents—is essential for this approach. If this spatial information can be collected and integrated into the dataset, this imputation strategy could become a viable option.

I recommend using panel analysis, which accounts for both temporal and cross-sectional dimensions in the data. This method allows you to isolate the effect of time-invariant factors, such as ShotSpotter deployment, while controlling for unobserved differences between cross-sectional units In this context, the temporal dimension could represent years, and the cross-sectional dimension could represent streets or neighborhoods in Cambridge.

Impact

The deployment of ShotSpotter technology in Cambridge has profound implications beyond its intended purpose of gunshot detection. By examining its socioeconomic and public safety impacts, we identified several critical areas of influence:

Socioeconomic Transformation:

Housing Prices: ShotSpotter deployments potentially correlate with increases in housing prices, particularly in historically low-income neighborhoods. This trend suggests a possible link between increased surveillance and gentrification, as higher property values often lead to the displacement of vulnerable populations.

Public Safety Concerns:

Inaccurate Detection: ShotSpotter's actual gunshot detection accuracy (55.8%) is far below its advertised claim of 97%, raising questions about its reliability. This gap undermines its effectiveness and casts doubt on its ability to improve public safety meaningfully.

Community Distrust: The technology's limitations, combined with concerns about over-surveillance, may erode trust between law enforcement and communities, particularly in neighborhoods with a history of systemic inequities.

Policy and Resource Allocation:

The financial investment in ShotSpotter contracts must be critically evaluated against its modest impact on public safety and its broader socioeconomic consequences. Alternative investments in community-driven safety programs or housing equity initiatives could yield more sustainable and equitable outcomes.

These impacts highlight the need for careful consideration of the unintended consequences of deploying surveillance technologies like ShotSpotter. Policymakers must weigh the benefits against the potential harm to ensure decisions align with the community's best interests.

Conclusion

This study sheds light on the complex dynamics between ShotSpotter technology, public safety, and socioeconomic shifts in Cambridge. While the system promises to enhance law enforcement capabilities, its limited accuracy and broader implications raise critical questions about its effectiveness and ethical use.

The data reveal that ShotSpotter deployments could be associated with rising housing prices and potential gentrification, which could disproportionately impact low-income communities. Additionally, the system's failure to deliver on its accuracy claims undermines its value as a public safety tool and highlights the imp ortance of independent validation of its performance.

Moving forward, policymakers and community leaders should:

- Reassess the cost-effectiveness of ShotSpotter and explore alternative public safety investments.
- Consider the socioeconomic impacts of surveillance technologies on vulnerable populations.
- Engage in transparent dialogue with residents to address concerns and ensure equitable decision-making.

Ultimately, this analysis serves as a foundation for informed advocacy and decision-making, equipping stakeholders with the evidence needed to challenge or refine the use of ShotSpotter in Cambridge. By prioritizing data-driven insights and community voices, cities can pursue safety strategies that genuinely benefit all residents.