SAFER TOGETHER: WASHINGTON D.C

2014-2017

Cross sectional analysis of gunshots predictors

Examination of gunshot prevalence across the seven police districts of Washington D.C using open source data.

By Brian Collins, Harsha Goonewardana, Priya Kakkar and Kihoon Sohn

SAFER TOGETHER: WASHINGTON D.C

Cross sectional analysis of gunshots predictors

INTRODUCTION

ShotSpotter is an integrated acoustic surveillance system which uses sensors to isolate and triangulate the sound of a gunshot and alerts the authorities to the location.[1] Numerous research studies have postulated that the violent crime, for which gunshots can be a proxy, is selectively underreported and ShotSpotter data provide an important alternative outcome measure that is not affected by underreporting.[2]Other studies have shown a causal link between amount of violent crime and property values. Across United States of America, reduction in gun crimes have shown to increase property values by close to two thousand dollars.[3] A correlation between calls for service and crime have been shown by the literature and researchers have suggest that 311 calls for service are a leading indicator of municipal neglect.[4]

PROBLEM STATEMENT

The project team examined a feature set including gunshot locations, calls for service, and property information. The aim was to identify and describe expected relationships with data supplied by the ShotSpotter system in the seven police districts in the City of Washington, D.C. Two independent time series analysis models using gunshot data from 2014 to 2018 was used to predict the incidents of gunshots in each police district for the subsequent 12 months. A logistic regression model trained on a combination of spatial and temporal data was utilized for inference analysis and was able to classify the presence of gunshots improved the prediction accuracy by 20% when compared to the baseline measure.

DATASET

The data spanned 51 months with daily observations of type and location of gunshots, description and volume of 311 calls to service and monthly real estate data[5]. The geospatial data allocated to match the boundaries of the seven police districts; please note the Second Police District is not part of the ShotSpotter program.

The feature engineered dataset contained three numerical and seven categorical features.

The ShotSpotter data was riddled with false positives in 2014-2015 with fireworks often being mistaken for gunshots around the holiday periods. In the subsequent time periods, the parameters for identification

[1] Williams, Clarence. "How ShotSpotter locates gunfire, helps police catch shooters and works to 'denormalize' gun violence" [Washington D.C]10 May 2017

[2] The geography, incidence, and underreporting of gun violence: new evidence using ShotSpotter data (with Jillian B. Carr). 2016. Brookings Research Paper.

[3] https://repository.upenn.edu/cgi/viewcontent.cgi?article=1352&context=fnce_papers

[4] Wheeler, Andrew "The effects of 311 Calls for Service on Crime in D.C at Microplaces" Crime & Delinquency, 23 June 2017

[5] Gunshot and 311 data was sourced from http://opendata.dc.gov/ and real estate information was retrieved from https://www.redfin.com/blog/data-center

were updated to mitigate this misclassification, but the gunshot data still contains outliers around the holiday periods.

METHODOLOGY

Macro level analysis and prediction was conducted using the entire dataset for Washington D.C. Subsequently the dataset was disaggregated to the police district strata for further examination.

A logistic regression model trained on the feature set successfully classified gunshots with an 80% accuracy. The binary classification model was also able to recognize indicators correlated with incidence of gunshots.

An Augmented Dickey-Fuller test was conducted to confirm the stationarity of the time series data. Two independent time series models were trained on 1461 daily gunshot incidents from 2014 to 2018. Autoregression Integrated Moving Average (ARIMA) and FB PROPHET models provided similar predictions on the gunshot incidences in the subsequent 12 months.

First order autocorrelation is prevalent among the data indicating presence of gunshots can lead to other firearm incidents with one to two days of the original incident

RESULTS

The analysis showed Real Estate data correlation

On aggregate, the length of time to resolve a 311 call and the closure of a 311 call both showed the most correlation followed by garbage collection, graffiti removal and street light repairs.

Police districts 6 & 7 showed the most amounts of gunshots. The incidents were also clustered around the weekend. Both the gunshot prediction models anticipated a reduction in overall gunshots in most areas with the city limits in the period to 2014-2018 except police areas 6 & 7. These two areas show that the gun violence is continuing along the present trendline.

RECOMMENDATIONS

The findings recommend continued focus on the quick resolution of 311 service calls to signal the attention of local government, preempting civic negligence which can lead to violent crime.

Preventive police presence should be increased during the weekends and in areas of gun violence in the subsequent days.

Police districts 6 & 7 can be targeted with extra policing resources to reduce the incidence of firearm violence, especially if resource become available due to the expected reduction in gun crime in other areas.

AREAS FOR FURTHER RESEARCH

Better harmonization of real estate data with police districts using GPS data points is an important next step to disaggregate the seven police districts to the 52 Police Service Areas for more granular examination of gun violence at the neighborhood level.