Case Study: San Francisco Auto-Burglaries Before and After the Pandemic

David Harper, PhD*
(Dated: August 29, 2023)

This case study delves into the intricate dynamics of auto-burglaries in various neighborhoods of San Francisco, both before and after the imposition of a Stay-at-Home Order (SAHO) during the Covid-19 pandemic. Using comprehensive datasets spanning the periods 'Pre-SAHO' and 'Post-SAHO', this research employs two logistic regression models to discern patterns and risk factors that influence such incidents. While the first model focuses on categorical variables such as the Neighborhood, Time-of-Day, Day-of-the-Week, and Month, the second model enriches this perspective by adding quantifiable variables related to neighborhoods characteristics standardized using z-scores. Additionally, this study adopts a bootstrap method to compute 99% confidence intervals for the daily average incidents in each neighborhoods during the two distinct periods. Findings suggest notable variations in the likelihood of auto-burglaries contingent on specific conditions and neighborhoods features. This case study offers insights and highlights the dynamic nature of a critical issue in the San Francisco area.

I. INTRODUCTION

San Francisco is often cited as a prime example of cities grappling with high rates of auto-burglaries. However, the true extent of this issue can sometimes be overshadowed by sensationalist reporting. In light of this, establishing a fact-driven perspective becomes crucial to accurately understand the auto burglary problem in San Francisco.

This project aims to assess the impact of specific risk factors on the prevalence of auto-burglaries in the city. These factors include location, time, and select census data believed to correlate with the number of automobiles in a given neighborhood.

II. METHODOLOGY

A. Data Collection

San Francisco is a city which offers a comprehensive collection of data available to the public which is relevant to various aspects and functions of the city. This data can be found on `https://data.sfgov.org`. We make use of this data for this analysis. In particular, we will be incorporating the following data sources as part of this project:

Data Source 1. Crime Data

- **Description**: Reported crime incidents in San Francisco beginning in January 2018 to the present.
- **Direct Link**: `https://data.sfgov.org/Public-Safety/Police-Department-Incident-Reports-2018-to-Present/wg3w-h783`

Data Source 2. On-Street Parking Census

- **Description**: Counts of publicly available, on-street parking for each street segment throughout the city.
- **Direct Link**: `https://data.sfgov.org/Transportation/On-Street-Parking-Census/9ivs-nf5y`

Data Source 3. Population Data

- **Description**: This dataset contains population and demographic estimates and associated margins of error obtained and derived from the US Census.
- **Direct Link**: `https://data.sfgov.org/Economy-and-Community/San-Francisco-Population-and-Demographic-Census-da/4qbq-hvtt`

Data Source 4. Analysis Neighborhood Limits

- **Description**: 41 neighborhoods by grouping 2010 Census tracts, using common real estate and residents' definitions.
- **Direct Link**: `https://data.sfgov.org/Geographic-Locations-and-Boundaries/Analysis-Neighborhoods/p5b7-5n3h`

Data Source 5. Parking Meter Locations

- **Description**: Location, cap color and other key attributes of parking meters in San Francisco.
- **Direct Link**: `https://data.sfgov.org/Transportation/Parking-Meters/8vzz-qzz9`

* Author email: dmichaelhii@gmail.com

B. Data Preparation

In ensuring the robustness of our analysis, a series of data preparation steps were implemented:

- **Filtering Relevant Incidents**: Our primary dataset encompassed a variety of crime incidents. To hone in on our research interest, we selectively extracted entries where the *Incident Subcategory* was marked as either *Larceny - From Vehicle* or *Theft From Vehicle.*

- **Temporal Segmentation**: A Stay-At-Home Order (SAHO) persisted from March 16, 2020 to January 25, 2021. For this reason, we decided to split the data into two distinct 2-year periods: Pre-SAHO (01/01/2018 - 12/31/2019) and Post-SAHO (06/06/21 - 05/31/23). This allowed for a comparative assessment of incident rates, factoring in the potential behavioral changes influenced by the SAHO.

- **Geospatial Analysis**: The 'Analysis Neighborhood Limits' data source provided shape files delineating Neighborhood boundaries. Employing QGIS, a spatial aggregation was performed to count the number of auto-burglary incidents, parking meters, and on-street parking spots within each defined Neighborhood. This geospatial clustering was instrumental in ensuring Neighborhood-centric insights were grounded in accurate spatial computations.

III. METHODOLOGY AND ANALYSIS

A. Features

In our analysis, we aim to discern the relationships between incident rates and the following risk factors, as well as gauge their relative impacts:

1. **Neighborhood**: Given the potential differences between areas, each Neighborhood serves as a categorical predictor, helping account for risks potentially unique to each locale.

2. **Time-of-Day**: Divided into morning (6am-12pm), afternoon (12pm-6pm), evening (6pm-12am), and night (12am-6am) as categorical intervals, this helps ascertain if incidents are likelier during particular periods.

3. **Day-of-the-Week**: Days are categorized to identify if certain days are more prone to incidents.

4. **Month**: Represented as a categorical variable, this aims to detect any seasonal influences or trends over the year.

5. **Neighborhood Population (z-score)**: The standardized population of a Neighborhood, where larger values indicate higher-than-average populations.

6. **Number of Parking Spots (z-score)**: The standardized count of street parking spots in a Neighborhood, with larger values indicating an above-average number of spots.

7. **Number of Parking Meters (z-score)**: The standardized count of parking meters in a Neighborhood, with larger values showing an above-average number of meters.

The z-score for a data set $x = \{x_i\}$ *is define as* $z_i = (x_i - \overline{x})/SD(x)$*, where* $SD(x)$ *is the standard deviation of* x*.*

Our analysis unfolds through two distinct angles, each contributing unique insights into the pattern and propensity of auto-burglary incidents in the region.

- **Descriptive Statistics**: In section IV, we present a series of descriptive statistics that provide an overview of the auto burglary incidents' distribution and characteristics.

- **Logistic Regression Analysis**: Using logistic regression models, we assess the likelihood of an auto-burglary incident based on selected variables. This method helps determine how specific factors, such as Neighborhood or Time-of-Day, impact the probability of an event.

- **Neighborhood-Centric Comparative Analysis**: This analysis compares the daily average incidents in each Neighborhood, both before and after the SAHO. The aim is to identify localized patterns and differences in incident rates across neighborhoods.

Together, these methods provide a broad and detailed perspective on auto-burglary incidents in the area.

B. Logistic Regression Models

Our analysis leverages two distinct logistic regression models, each incorporating various features. Importantly, both models are fit to the two separate datasets: Pre-SAHO and Post-SAHO.

1. *Model A*

Model A is designed to uncover the relationships and gauge the relative effects of risk factors such as Neighborhood, Time-of-Day, Day-of-the-Week, and Month on the probability of an auto burglary occurrence.

The training set for this model is constructed from combinations of the above features for each date in our study period. Each entry in this dataset represents whether an auto burglary took place under a particular set of conditions (for instance, on a Sunday afternoon in the Mission neighborhood during July). A crucial aspect of our dataset is its structure that ensures every Neighborhood has been equally represented across all day and Month combinations, thereby eliminating potential bias.

Using the logistic regression model, we aim to delineate the relationships between these factors and the probability of an incident occurring. Such insights are invaluable in pinpointing conditions that might be particularly prone to auto-burglaries.

It's also worth highlighting that, owing to the nature of our data, it's possible for multiple data points to share identical input features but differ in outcomes. This stems from the observational reality that even under identical conditions, incidents may or may not materialize.

2. *Model B*

Model B further advances our inquiry into the factors influencing auto-burglaries by incorporating a mix of categorical features and standardized numerical attributes. Retaining the Time-of-Day, Day-of-the-Week, and Month variables from Model A, this model integrates Neighborhood-centric metrics, diving deeper into potential localized risk factors.

The Time-of-Day, represented as morning, afternoon, evening, and night, explores the nuances of incident frequencies during distinct periods of the day. Similarly, the categorical variables Day-of-the-Week and Month are incorporated to tease out any patterns on specific days or any overarching seasonal influences.

Enhancing the granularity of our model, we introduce the Neighborhood Population represented as a z-score. This standardized value provides insights into whether neighborhoods with populations diverging from the mean — either above or below — exhibit different risks. Furthermore, the standardized count of street parking spots and parking meters in a Neighborhood offers a lens to gauge if the availability of parking and the presence of regulated parking zones, respectively, correlate with the occurrence of auto-burglaries.

C. Comparative Analysis of Daily Average Incidents by Neighborhood

Moving beyond multifaceted logistic models, this section hones in on neighborhoods as singular risk factors. The goal is to delve deeper into the locational influences on auto-burglaries. Specifically, the analysis is designed to capture the potential variations in the daily average number of incidents across neighborhoods and discern how these patterns might have shifted after the SAHO.

- **Average Incidents and Confidence Intervals**

  For both the Pre-SAHO and Post-SAHO datasets, the daily average number of incidents per Neighborhood was computed. Given the inherent variability in such data, it's vital to understand the uncertainty associated with these averages. To this end, we employed a bootstrap methodology to derive 99% confidence intervals around these averages. This non-parametric method offers a robust way to gauge the reliability of our estimates, even in the absence of assumptions about the underlying data distribution.

- **Change in Neighborhood Dynamics: Before and After SAHO**

  Armed with the daily averages and their confidence intervals for both periods, we embarked on a comparative analysis. The core of this assessment lies in understanding how the average number of incidents in each Neighborhood evolved as a result of the SAHO.

  To quantify these changes, we calculated the difference in means before and after the SAHO for each Neighborhood. Using the previously obtained confidence intervals from the bootstrap analysis, we were able to construct 98% confidence intervals for these differences. These intervals serve as indicators, helping us ascertain whether the observed differences are statistically significant or if they might have arisen due to random fluctuations.

  This comparative analysis sheds light on which neighborhoods experienced marked shifts in auto burglary rates Post-SAHO, guiding stakeholders in potential targeted interventions and policy decisions.

IV. RESULTS

A. Descriptive Statistics

Our dataset exploration began with Tables I and II, detailing the reported auto burglary incidents for both

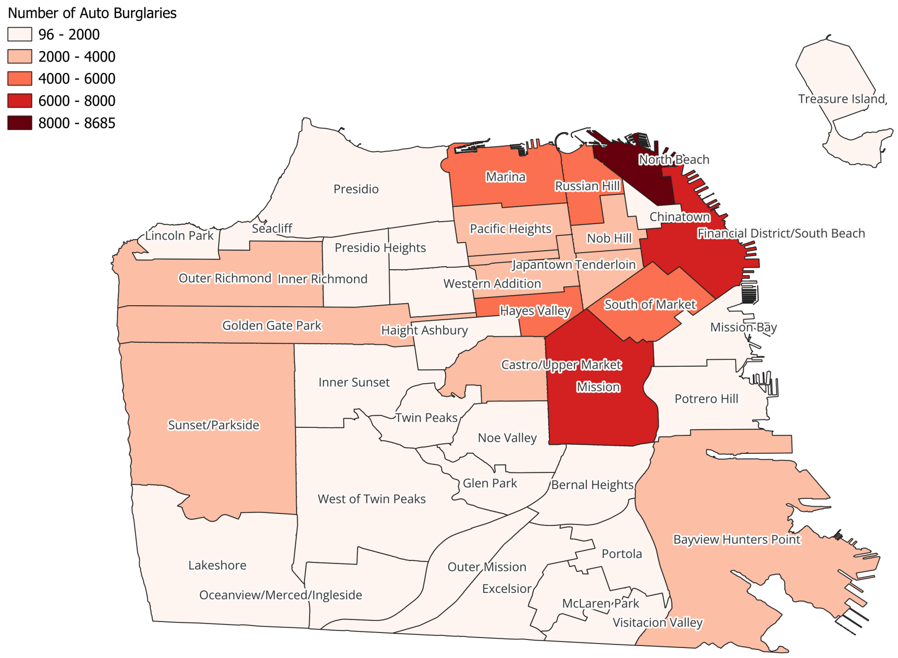

Figure 1. Total number of incidents in each Neighborhood included in the span of our analysis.

Table I. Pre-SAH

Incident Description	Count	%
Burglary, Vehicle (Arrest made)	637	1.2%
Burglary, Vehicle, Att. (Arrest made)	48	0.1%
Theft, From Locked Vehicle, \$200-\$950	4725	9.2%
Theft, From Locked Vehicle, \$50-\$200	1602	3.1%
Theft, From Locked Vehicle, <\$50	831	1.6%
Theft, From Locked Vehicle, >\$950	37173	72.5%
Theft, From Locked Vehicle, Att.	526	1.0%
Theft, From Unlocked Vehicle, \$200-\$950	524	1.0%
Theft, From Unlocked Vehicle, \$50-\$200	296	0.6%
Theft, From Unlocked Vehicle, <\$50	251	0.5%
Theft, From Unlocked Vehicle, >\$950	4639	9.0%
Theft, From Unlocked Vehicle, Att.	39	0.1%
Total Locked[a]	44857	87.5%
Total Unlocked[b]	5749	11.2%
Other[c]	685	1.3%

[a] Based on incident descriptions indicating locked vehicles.
[b] Based on incident descriptions indicating unlocked vehicles.
[c] Derived from incident descriptions that do not specify the status of vehicle locking.

the Pre-SAHO and Post-SAHO periods. These tables show aggregate counts and further classification by incident description, such as the burglarized vehicle's lock status and the value range of stolen items.

In Figure 1, a heat map, visualizes the distribution of incidents across neighborhoods, using data from both Pre-SAHO and Post-SAHO periods. This visualization highlights areas with higher incident counts and provides a geographical context to the neighborhoods.

In Figure 2, a time-series plot of reported incidents, supplemented with a 31-day moving average, captures the trends over time. Important time intervals, including the Pre and Post-SAHO periods, are marked for clarity.

B. Logistic Regression Outcomes

1. Model A

For both Pre-SAHO and Post-SAHO datasets, the logistic regression models utilized the same reference categories: Golden Gate Park (for the Neighborhood), Morning (for Time-of-Day), Monday (for Day-of-the-Week), and January (for Month).

- **Neighborhood Significance**: The most notable finding was the overwhelming significance of the Neighborhood category. When benchmarked against our reference (Golden Gate Park), multiple neighborhoods exhibited significantly different odds ratios. This suggests that the particular

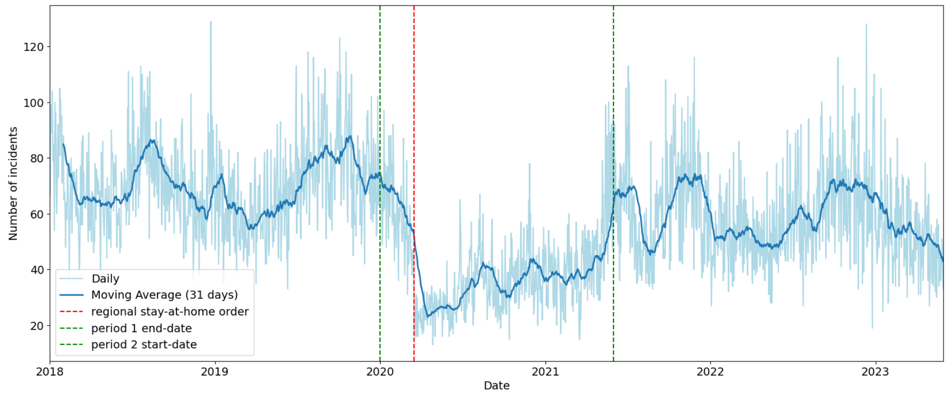

Figure 2. Reported auto-burglaries. The red vertical line indicates the start of a regional stay-at-home-order which went into effect March 16, 2020 until January 25, 2021. The green vertical lines indicate the end and start dates for the two period constituting our analysis.

Table II. Post-SAH

Incident Description	Count	%
Burglary, Vehicle (Arrest made)	297	0.7%
Burglary, Vehicle, Att. (Arrest made)	13	< 0.1 %
Theft, From Locked Vehicle, \$200-\$950	3377	7.9%
Theft, From Locked Vehicle, \$50-\$200	1166	2.7%
Theft, From Locked Vehicle, <\$50	814	1.9%
Theft, From Locked Vehicle, >\$950	30562	71.6%
Theft, From Locked Vehicle, Att.	637	1.5%
Theft, From Unlocked Vehicle, \$200-\$950	378	0.9%
Theft, From Unlocked Vehicle, \$50-\$200	201	0.5%
Theft, From Unlocked Vehicle, <\$50	185	0.4%
Theft, From Unlocked Vehicle, >\$950	5042	11.8%
Theft, From Unlocked Vehicle, Att.	32	0.1%
Total Locked[a]	36556	85.6%
Total Unlocked[b]	5838	13.7%
Other[c]	310	0.7%

[a] Based on incident descriptions indicating locked vehicles.
[b] Based on incident descriptions indicating unlocked vehicles.
[c] Derived from incident descriptions that do not specify the status of vehicle locking.

Neighborhood plays a pivotal role in predicting the likelihood of an auto-burglary incident, even more so than time-based categories.

- **Time-of-Day Influence**: In both the Pre-SAHO and Post-SAHO models, every Time-of-Day category demonstrated strong statistical significance when compared to the reference Morning category. Not only were these categories significant, but their effect sizes (as gauged by the odds ratios) were also notably large, especially when juxtaposed with other factors like Day-of-the-Week and Month. This highlights the crucial role that specific times of the day play in the likelihood of an incident.

- **Day-of-the-Week Influence**: The significance of specific days as predictors exhibits variation between the Pre-SAHO and Post-SAHO models. In the Pre-SAHO period, multiple days demonstrate statistical significance in predicting auto-burglary likelihood. However, in the Post-SAHO scenario, only Friday stands out as a statistically significant predictor.

- **Month Influence**: The influence of individual months remains evident in both Pre-SAHO and Post-SAHO models. However, the specific months that are statistically significant, along with the direction of their influence (whether increasing or decreasing odds), show variability. One potential explanation for these shifts might be the disruption of seasonal tourism caused by the SAHO, affecting the typical patterns associated with certain months.

For a detailed breakdown of the results obtained by this part of the analysis please refer to tables III and IV.

2. Model B

Model B, which integrated the effects of Neighborhood population, number of parking meters, and parking spots, offered valuable insights, even if its overall fit might not have been as strong as that of Model A. Model B largely mirrors the findings from Model A for common predictors:

- **Consistent Predictors**: The influence of common predictors, such as Time-of-Day, Day-of-the-Week, and Month, remain consistent with Model A across both Pre-SAHO and Post-SAHO periods. This consistency underscores the robustness of these relationships in predicting auto-burglary incidents.
- **Parking Infrastructure's Role**: Among the numerical features, perhaps surprisingly the number of parking meters in a Neighborhood proved to be the most influential even above population. This suggests areas with more parking meters could witness a heightened likelihood of auto-burglaries. This aligns with the intuitive notion that areas with metered parking might have more transient traffic, possibly increasing the risk profile.

  However, an unexpected trend was the inverse relationship between the number of parking spots and the incidence of auto-burglaries. Holding population and parking meters constant, an increase in the number of parking spots predicted a decline in auto-burglary likelihood. This could hint at a correlation between higher parking availability (and thus potentially lower vehicle density) and decreased incidents, a factor that could be examined in more detail in future investigations.

For a detailed breakdown of the results obtained by this part of the analysis please refer to tables V and VI.

C. Comparative Analysis of Daily Average Incidents by Neighborhood

The differences in auto-burglary incidents before and after the SAHO became especially evident when assessing the mean number of incidents at the Neighborhood level. Across the city, there was a notable decline in the average daily incidents from approximately 70 to 58 between the Pre-SAHO and Post-SAHO periods. This translates to an approximate 17% decrease.

Delving deeper into the Neighborhood-specific data:

- **Consistent Rankings**: Despite the overall reduction, the hierarchy of neighborhoods with respect to their auto-burglary frequencies remained relatively consistent.
- **Significant Variances**: Among the neighborhoods, Mission stood out with a sharp 44% reduction in auto-burglaries. In contrast, North Beach exhibited an unexpected trend, seeing an increase of roughly 46% in incidents. This uptick in North Beach juxtaposed with the citywide decline underscores the heterogeneity in Neighborhood responses to the broader citywide dynamics.

This Neighborhood-centric analysis offers a granular perspective on how city-wide trends manifest at the local level. While the general decline in incidents following the SAHO is reassuring, the stark differences between neighborhoods like Mission and North Beach reiterate the importance of localized strategies and interventions in crime prevention.

For a detailed breakdown of the results obtained by this part of the analysis please refer to table VII, VIII and IX.

V. CONCLUSION

This case study embarked on a investigation into the factors influencing auto-burglaries in San Francisco, specifically examining the periods before and after the Stay-at-Home Order (SAHO) issued during the COVID-19 pandemic.

Two logistic regression models were employed:

- Model A focused on categorical predictors such as Neighborhood, Time-of-Day, Day-of-the-Week, and Month. The results indicated that the Neighborhood played a pivotal role in influencing the likelihood of incidents, with Time-of-Day following closely in terms of significance.
- Model B, a variant of the initial model, incorporated the Neighborhood's population, parking meters, and parking spots – all normalized using z-scores. Time-of-Day emerged as a significant predictor, followed by parking infrastructure features. Interestingly, an increase in parking spots correlated with a decrease in incidents, potentially suggesting that denser vehicle populations in certain areas might be linked to higher auto-burglary rates.

In a complementary analysis focusing on daily average incidents by Neighborhood, a noticeable reduction in citywide auto-burglaries was observed Post-SAHO. This trend was further evident at the Neighborhood level. Notably, while some neighborhoods like the Mission experienced a considerable decrease in burglaries, others like North Beach saw an increase.

Overall, the findings of this study offer valuable insights into the factors that can elevate or mitigate the risk of auto-burglaries in San Francisco. The distinct patterns observed in different neighborhoods both before and after the SAHO period underscore the dynamic nature of urban crime and highlight the importance of localized strategies for prevention and response.

Table III. Pre-SAHO Model A Estimates with odds ratios with 95%-confidence intervals

Feature	OR (CI)	Feature	OR (CI)
Neighborhood (Ref. Golden Gate Park)		Outer Mission	$0.71^{***}(0.61, 0.83)$
Mission	$17.86^{***}(15.93, 20.01)$	Presidio	$0.67^{***}(0.58, 0.79)$
Financial District/South Beach	$13.28^{***}(11.87, 14.85)$	Presidio Heights	$0.63^{***}(0.54, 0.74)$
South of Market	$11.30^{***}(10.10, 12.63)$	Oceanview/Merced/Ingleside	$0.62^{***}(0.53, 0.72)$
North Beach	$9.66^{***}(8.65, 10.79)$	Lincoln Park	$0.55^{***}(0.47, 0.65)$
Hayes Valley	$8.83^{***}(7.91, 9.87)$	Visitacion Valley	$0.52^{***}(0.44, 0.62)$
Western Addition	$7.89^{***}(7.06, 8.81)$	Treasure Island	$0.23^{***}(0.19, 0.29)$
Russian Hill	$7.85^{***}(7.03, 8.78)$	Seacliff	$0.17^{***}(0.13, 0.22)$
Tenderloin	$7.34^{***}(6.57, 8.20)$	McLaren Park	$0.10^{***}(0.07, 0.13)$
Marina	$5.90^{***}(5.28, 6.59)$	Month (Ref. January)	
Outer Richmond	$5.21^{***}(4.66, 5.82)$	February	$0.79^{***}(0.74, 0.85)$
Nob Hill	$4.35^{***}(3.88, 4.86)$	March	$0.76^{***}(0.71, 0.82)$
Bayview Hunters Point	$4.18^{***}(3.73, 4.68)$	April	$0.79^{***}(0.74, 0.85)$
Castro/Upper Market	$3.82^{***}(3.41, 4.28)$	May	$0.83^{***}(0.78, 0.89)$
Pacific Heights	$3.75^{***}(3.34, 4.20)$	June	$0.88^{***}(0.82, 0.95)$
Japantown	$3.68^{***}(3.28, 4.12)$	July	$1.07(1.00, 1.15)$
Sunset/Parkside	$3.52^{***}(3.14, 3.94)$	August	$1.10^{**}(1.03, 1.18)$
Potrero Hill	$2.88^{***}(2.56, 3.23)$	September	$0.98(0.91, 1.05)$
West of Twin Peaks	$2.37^{***}(2.11, 2.67)$	October	$1.06(0.99, 1.14)$
Inner Sunset	$1.99^{***}(1.76, 2.25)$	November	$0.86^{***}(0.80, 0.92)$
Inner Richmond	$1.99^{***}(1.76, 2.25)$	December	$0.95(0.89, 1.02)$
Haight Ashbury	$1.97^{***}(1.74, 2.22)$	Day-of-the-Week (Ref. Monday)	
Chinatown	$1.90^{***}(1.68, 2.14)$	Tuesday	$0.91^{***}(0.86, 0.96)$
Lakeshore	$1.80^{***}(1.59, 2.03)$	Wednesday	$0.97(0.92, 1.03)$
Mission Bay	$1.67^{***}(1.48, 1.90)$	Thursday	$0.94^{*}(0.89, 0.99)$
Bernal Heights	$1.64^{***}(1.45, 1.86)$	Friday	$1.06^{*}(1.00, 1.12)$
Lone Mountain/USF	$1.47^{***}(1.30, 1.67)$	Saturday	$1.09^{**}(1.03, 1.15)$
Twin Peaks	$1.39^{***}(1.22, 1.58)$	Sunday	$1.04(0.98, 1.10)$
Portola	$1.09(0.95, 1.25)$	Time-of-Day (Ref. Morning)	
Noe Valley	$1.05(0.92, 1.21)$	Afternoon	$1.96^{***}(1.89, 2.04)$
Excelsior	$0.79^{***}(0.68, 0.91)$	Evening	$2.81^{***}(2.70, 2.93)$
Glen Park	$0.73^{***}(0.63, 0.85)$	Night	$0.55^{***}(0.53, 0.58)$

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

AUC-ROC: 0.79

Table IV. Post-SAHO Model A Estimates with odds ratios with 95%-confidence intervals

Feature	OR (CI)	Feature	OR (CI)
Neighborhood (Ref. Golden Gate Park)		Outer Mission	$0.70^{***}(0.61, 0.81)$
North Beach	$11.94^{***}(10.71, 13.30)$	Portola	$0.61^{***}(0.53, 0.71)$
Financial District/South Beach	$9.72^{***}(8.73, 10.83)$	Presidio Heights	$0.47^{***}(0.39, 0.55)$
Russian Hill	$6.92^{***}(6.22, 7.70)$	Glen Park	$0.44^{***}(0.37, 0.52)$
Mission	$6.70^{***}(6.02, 7.45)$	Presidio	$0.43^{***}(0.36, 0.51)$
Hayes Valley	$5.41^{***}(4.86, 6.02)$	Seacliff	$0.32^{***}(0.26, 0.39)$
Marina	$4.90^{***}(4.41, 5.46)$	Visitacion Valley	$0.32^{***}(0.26, 0.38)$
South of Market	$4.72^{***}(4.24, 5.25)$	Treasure Island	$0.14^{***}(0.11, 0.19)$
Outer Richmond	$3.76^{***}(3.37, 4.19)$	McLaren Park	$0.12^{***}(0.09, 0.16)$
Western Addition	$3.67^{***}(3.29, 4.09)$	Month (Ref. January)	
Japantown	$3.30^{***}(2.96, 3.68)$	February	$0.96(0.89, 1.03)$
Bayview Hunters Point	$3.19^{***}(2.86, 3.56)$	March	$1.03(0.96, 1.11)$
Nob Hill	$2.77^{***}(2.47, 3.09)$	April	$0.98(0.91, 1.06)$
Sunset/Parkside	$2.72^{***}(2.43, 3.04)$	May	$0.90^{**}(0.84, 0.97)$
Tenderloin	$2.59^{***}(2.32, 2.90)$	June	$1.14^{***}(1.06, 1.23)$
Castro/Upper Market	$2.28^{***}(2.04, 2.56)$	July	$1.09^{*}(1.02, 1.17)$
Pacific Heights	$2.25^{***}(2.01, 2.53)$	August	$1.03(0.96, 1.11)$
Inner Sunset	$2.25^{***}(2.01, 2.52)$	September	$1.24^{***}(1.16, 1.34)$
Haight Ashbury	$2.19^{***}(1.95, 2.46)$	October	$1.33^{***}(1.23, 1.42)$
West of Twin Peaks	$1.64^{***}(1.46, 1.85)$	November	$1.33^{***}(1.24, 1.43)$
Inner Richmond	$1.63^{***}(1.45, 1.84)$	December	$1.17^{***}(1.09, 1.26)$
Lakeshore	$1.52^{***}(1.35, 1.72)$	Day-of-the-Week (Ref. Monday)	
Chinatown	$1.46^{***}(1.29, 1.65)$	Tuesday	$0.96(0.91, 1.02)$
Potrero Hill	$1.41^{***}(1.24, 1.59)$	Wednesday	$1.05(1.00, 1.11)$
Bernal Heights	$1.22^{**}(1.07, 1.38)$	Thursday	$1.00(0.95, 1.06)$
Noe Valley	$1.05(0.92, 1.19)$	Friday	$1.14^{***}(1.08, 1.20)$
Excelsior	$0.99(0.86, 1.13)$	Saturday	$1.02(0.96, 1.08)$
Mission Bay	$0.90(0.79, 1.03)$	Sunday	$0.98(0.93, 1.04)$
Twin Peaks	$0.84^{**}(0.73, 0.97)$	Time-of-Day (Ref. Morning)	
Lone Mountain/USF	$0.77^{***}(0.67, 0.89)$	Afternoon	$1.74^{***}(1.67, 1.81)$
Lincoln Park	$0.72^{***}(0.62, 0.83)$	Evening	$1.70^{***}(1.63, 1.77)$
Oceanview/Merced/Ingleside	$0.71^{***}(0.62, 0.82)$	Night	$0.63^{***}(0.60, 0.66)$

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

AUC-ROC: 0.76

Table V. Pre-SAHO Model B Estimates with odds ratios with 95%-confidence intervals

Feature	OR (CI)	Feature	OR (CI)
Numerical Feature (z-scores)		November	$0.87^{***}(0.81, 0.93)$
Population	$1.28^{***}(1.26, 1.30)$	December	$0.96(0.89, 1.02)$
Parking spots	$0.95^{***}(0.93, 0.96)$	Day-of-the-Week (Ref. Monday)	
Parking meters	$1.76^{***}(1.73, 1.78)$	Tuesday	$0.91^{***}(0.87, 0.96)$
Month (Ref. January)		Wednesday	$0.98(0.93, 1.03)$
February	$0.81^{***}(0.76, 0.87)$	Thursday	$0.94^{*}(0.89, 0.99)$
March	$0.78^{***}(0.73, 0.83)$	Friday	$1.05(1.00, 1.11)$
April	$0.81^{***}(0.75, 0.86)$	Saturday	$1.08^{**}(1.02, 1.13)$
May	$0.85^{***}(0.79, 0.91)$	Sunday	$1.03(0.98, 1.09)$
June	$0.89^{***}(0.83, 0.95)$	Time-of-Day (Ref. Morning)	
July	$1.06(1.00, 1.14)$	Afternoon	$1.86^{***}(1.79, 1.93)$
August	$1.09^{*}(1.02, 1.16)$	Evening	$2.56^{***}(2.46, 2.66)$
September	$0.98(0.91, 1.05)$	Night	$0.56^{***}(0.54, 0.59)$
October	$1.06(0.99, 1.13)$		

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$
AUC-ROC: 0.73

Table VI. Post-SAHO Model B with odds ratios with 95%-confidence intervals

Feature	OR (CI)	Feature	OR (CI)
Numerical Feature (z-scores)		November	$1.30^{***}(1.21, 1.39)$
Population	$1.21^{***}(1.19, 1.23)$	December	$1.16^{***}(1.08, 1.24)$
Parking spots	$0.97^{***}(0.95, 0.99)$	Day-of-the-Week (Ref. Monday)	
Parking meters	$1.54^{***}(1.52, 1.56)$	Tuesday	$0.97(0.92, 1.02)$
Month (Ref. January)		Wednesday	$1.05(1.00, 1.11)$
February	$0.96(0.89, 1.03)$	Thursday	$1.00(0.95, 1.05)$
March	$1.03(0.96, 1.10)$	Friday	$1.13^{***}(1.07, 1.19)$
April	$0.99(0.92, 1.06)$	Saturday	$1.02(0.96, 1.07)$
May	$0.91^{**}(0.85, 0.97)$	Sunday	$0.98(0.93, 1.04)$
June	$1.13^{***}(1.05, 1.21)$	Time-of-Day (Ref. Morning)	
July	$1.08^{*}(1.01, 1.16)$	Afternoon	$1.62^{***}(1.56, 1.69)$
August	$1.02(0.96, 1.10)$	Evening	$1.66^{***}(1.60, 1.73)$
September	$1.22^{***}(1.14, 1.31)$	Night	$0.65^{***}(0.62, 0.68)$
October	$1.29^{***}(1.21, 1.39)$		

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$
AUC-ROC: 0.68

Table VII. Pre-SAHO Daily Average number of incidents 99%-confidence intervals

Neighborhood	Avg (CI)	Neighborhood	Avg (CI)
Mission	5.99(5.60, 6.39)	Haight Ashbury	1.14(0.98, 1.31)
Financial District/South Beach	5.74(5.35, 6.12)	Lakeshore	1.12(0.95, 1.29)
North Beach	4.82(4.45, 5.21)	Inner Sunset	1.09(0.95, 1.24)
South of Market	4.38(4.09, 4.69)	Mission Bay	0.98(0.84, 1.14)
Russian Hill	3.88(3.53, 4.23)	Bernal Heights	0.88(0.75, 1.00)
Hayes Valley	3.73(3.42, 4.05)	Twin Peaks	0.86(0.70, 1.01)
Western Addition	3.41(3.14, 3.69)	Lone Mountain/USF	0.83(0.71, 0.96)
Tenderloin	3.17(2.92, 3.44)	Noe Valley	0.59(0.50, 0.70)
Marina	2.94(2.68, 3.21)	Portola	0.59(0.49, 0.70)
Outer Richmond	2.71(2.45, 2.98)	Excelsior	0.46(0.37, 0.56)
Japantown	2.03(1.83, 2.24)	Glen Park	0.43(0.34, 0.52)
Nob Hill	2.01(1.82, 2.21)	Presidio	0.42(0.33, 0.52)
Bayview Hunters Point	1.87(1.68, 2.05)	Outer Mission	0.42(0.33, 0.51)
Castro/Upper Market	1.85(1.65, 2.05)	Presidio Heights	0.37(0.30, 0.45)
Pacific Heights	1.76(1.59, 1.95)	Oceanview/Merced/Ingleside	0.36(0.28, 0.45)
Sunset/Parkside	1.70(1.52, 1.88)	Lincoln Park	0.34(0.27, 0.42)
Potrero Hill	1.61(1.41, 1.80)	Visitacion Valley	0.30(0.23, 0.38)
Golden Gate Park	1.43(1.24, 1.62)	Treasure Island	0.15(0.10, 0.20)
West of Twin Peaks	1.29(1.12, 1.47)	Seacliff	0.11(0.07, 0.15)
Chinatown	1.25(1.10, 1.43)	McLaren Park	0.06(0.03, 0.09)
Inner Richmond	1.19(1.03, 1.37)		

Table VIII. Post-SAHO Daily Average number of incidents 99%-confidence intervals

Neighborhood	Avg (CI)	Neighborhood	Avg (CI)
North Beach	7.08(6.52, 7.66)	Chinatown	0.99(0.86, 1.13)
Financial District/South Beach	5.03(4.66, 5.42)	West of Twin Peaks	0.95(0.83, 1.08)
Russian Hill	3.95(3.60, 4.32)	Potrero Hill	0.82(0.71, 0.94)
Mission	3.27(3.00, 3.56)	Bernal Heights	0.72(0.60, 0.83)
Hayes Valley	3.00(2.73, 3.28)	Noe Valley	0.63(0.52, 0.74)
Marina	2.73(2.47, 3.00)	Excelsior	0.61(0.51, 0.72)
South of Market	2.53(2.29, 2.78)	Mission Bay	0.56(0.47, 0.67)
Outer Richmond	2.16(1.94, 2.37)	Twin Peaks	0.54(0.44, 0.65)
Japantown	2.12(1.88, 2.36)	Lincoln Park	0.50(0.40, 0.61)
Western Addition	2.02(1.84, 2.21)	Lone Mountain/USF	0.47(0.38, 0.56)
Bayview Hunters Point	1.62(1.45, 1.78)	Oceanview/Merced/Ingleside	0.46(0.38, 0.56)
Golden Gate Park	1.54(1.37, 1.73)	Outer Mission	0.44(0.36, 0.53)
Nob Hill	1.51(1.35, 1.68)	Portola	0.38(0.30, 0.46)
Sunset/Parkside	1.50(1.33, 1.68)	Presidio Heights	0.30(0.22, 0.38)
Haight Ashbury	1.50(1.31, 1.71)	Presidio	0.29(0.22, 0.37)
Tenderloin	1.46(1.30, 1.63)	Glen Park	0.27(0.21, 0.34)
Inner Sunset	1.38(1.22, 1.55)	Seacliff	0.21(0.15, 0.26)
Castro/Upper Market	1.29(1.14, 1.45)	Visitacion Valley	0.20(0.14, 0.26)
Pacific Heights	1.22(1.08, 1.37)	Treasure Island	0.10(0.06, 0.15)
Inner Richmond	1.06(0.90, 1.23)	McLaren Park	0.07(0.04, 0.11)
Lakeshore	1.03(0.87, 1.19)		

Table IX. Difference between Post-SAHO and Post-SAHO Daily Average number of incidents 98%-confidence intervals

Neighborhood	Avg (CI)	Neighborhood	Avg (CI)
North Beach	$2.26^{*}(1.31, 3.22)$	Sunset/Parkside	$-0.20(-0.55, 0.17)$
Haight Ashbury	$0.36(0.00, 0.73)$	Portola	$-0.22^{*}(-0.40, -0.03)$
Inner Sunset	$0.29(-0.02, 0.60)$	Marina	$-0.22(-0.75, 0.32)$
Lincoln Park	$0.16(-0.02, 0.34)$	Bayview Hunters Point	$-0.25(-0.60, 0.10)$
Excelsior	$0.15(-0.05, 0.35)$	Chinatown	$-0.26(-0.57, 0.03)$
Golden Gate Park	$0.11(-0.26, 0.49)$	Twin Peaks	$-0.32^{*}(-0.57, -0.05)$
Seacliff	$0.10(0.00, 0.20)$	West of Twin Peaks	$-0.35^{*}(-0.65, -0.05)$
Oceanview/Merced/Ingleside	$0.10(-0.08, 0.28)$	Lone Mountain/USF	$-0.36^{*}(-0.58, -0.15)$
Japantown	$0.08(-0.36, 0.53)$	Mission Bay	$-0.42^{*}(-0.67, -0.18)$
Russian Hill	$0.08(-0.62, 0.78)$	Nob Hill	$-0.50^{*}(-0.86, -0.14)$
Noe Valley	$0.03(-0.18, 0.24)$	Pacific Heights	$-0.54^{*}(-0.87, -0.22)$
Outer Mission	$0.02(-0.15, 0.20)$	Outer Richmond	$-0.55^{*}(-1.05, -0.08)$
McLaren Park	$0.02(-0.05, 0.08)$	Castro/Upper Market	$-0.56^{*}(-0.91, -0.21)$
Treasure Island	$-0.04(-0.14, 0.05)$	Financial District/South Beach	$-0.70(-1.46, 0.07)$
Presidio Heights	$-0.07(-0.23, 0.09)$	Hayes Valley	$-0.73^{*}(-1.31, -0.13)$
Lakeshore	$-0.09(-0.42, 0.24)$	Potrero Hill	$-0.79^{*}(-1.10, -0.47)$
Visitacion Valley	$-0.11(-0.24, 0.03)$	Western Addition	$-1.40^{*}(-1.85, -0.93)$
Presidio	$-0.13(-0.31, 0.04)$	Tenderloin	$-1.71^{*}(-2.14, -1.29)$
Inner Richmond	$-0.14(-0.47, 0.20)$	South of Market	$-1.85^{*}(-2.40, -1.31)$
Glen Park	$-0.15(-0.32, 0.00)$	Mission	$-2.72^{*}(-3.40, -2.03)$
Bernal Heights	$-0.17(-0.40, 0.08)$		

Note: * indicates a significant ($p < 0.02$) change