**San Francisco Police Department Calls and Incidents Using Data from Kaggle.com**

Authors: Jesus Diaz, Hye Lim Kim, Zehao Chen

Department of Information Systems, California State University, Los Angeles

CIS 3200-02: DATA PROCESSNG & ANALYTICS

e-mail: [jdiaz294@casltatela.edu](mailto:jdiaz294@casltatela.edu) , [hkim192@calstatela.edu](mailto:hkim192@calstatela.edu) , [zchen49@calstatela.edu](mailto:zchen49@calstatela.edu)

**Abstract:** This document presents a comprehensive analysis of the calls and incidents reported in the vibrant city of San Francisco, California, as provided by the San Francisco Police Department. The study aims to provide valuable insights into the patterns, trends, and dynamics of law enforcement activities within the city, thereby contributing to a better understanding of public safety, crime prevention, and overall community well-being.

**1. Introduction**

The analysis in this paper is focused on the dataset focused on an in-depth examination of the data sourced from the San Francisco Police Department, this analysis encompasses various aspects of policing, including but not limited to, the types of incidents reported, geographical distribution of incidents, temporal patterns, and the demographic characteristics associated with different types of calls. The study leverages advanced statistical methods and data visualization techniques to distill meaningful information from the vast dataset, offering a nuanced perspective on the city's law enforcement landscape.

Moreover, this analysis seeks to identify potential hotspots of criminal activity, enabling law enforcement agencies, policymakers, and community leaders to allocate resources strategically and implement targeted intervention strategies. By understanding the temporal trends in incident reports, this study also endeavors to highlight any seasonality or periodicity in criminal activities, facilitating the development of proactive measures to address specific challenges during peak periods.

Furthermore, an examination of the socio-economic and demographic factors associated with reported incidents aims to shed light on the underlying determinants of criminal behavior in San Francisco. This multifaceted approach allows for a more holistic understanding of the intricate interplay between social, economic, and environmental factors that contribute to crime in urban settings.

In conclusion, the findings of this analysis are intended to inform evidence-based decision-making processes for law enforcement agencies, city officials, and community stakeholders. By providing a detailed and nuanced examination of calls and incidents in San Francisco, this document serves as a valuable resource for those seeking to enhance public safety, optimize resource allocation, and foster a safer and more resilient community.

**2. Related Work**

Previous works that have analyzed police calls and incidents are Hot spots and burning times: A spatiotemporal analysis of calls for service to establish police demand, a study, conducted in Antwerp, Belgium, which explores the spatiotemporal patterns of calls for police service, emphasizing the importance of prioritization through the use of dispatch-assigned priority codes. The analysis challenges traditional assumptions, revealing that urgent incidents, contrary to expectations, exhibit a spatially less clustered distribution. Additionally, a temporal concentration of urgent calls on Friday and Saturday evenings/nights is identified. Moreover, the study highlights the intersection of spatial and temporal dynamics, indicating that more urgent calls not only have distinct spatial and temporal characteristics but also require a higher allocation of police resources. This study provides valuable insights that may be extrapolated to enhance the efficiency of law enforcement operations. Understanding the nuanced spatiotemporal patterns of incidents, particularly in relation to their urgency, could inform the allocation of SFPD resources. By incorporating priority codes into patrol routing solutions, the SFPD might develop more realistic and adaptive patrol strategies.

      The analysis in this paper distinguishes itself from previous works mentioned in that the analysis was produced using Elastic Cloud. For example, one CSV file we analyzed contains more than 100k records, and without deploying this technology and it would have been impossible to analyze the dataset on our computer.

**3. Platform Specification**

We used Kibana and ElasticSearch and we set the size of 1 GB RAM.

*Table 1: Platform Specifications*

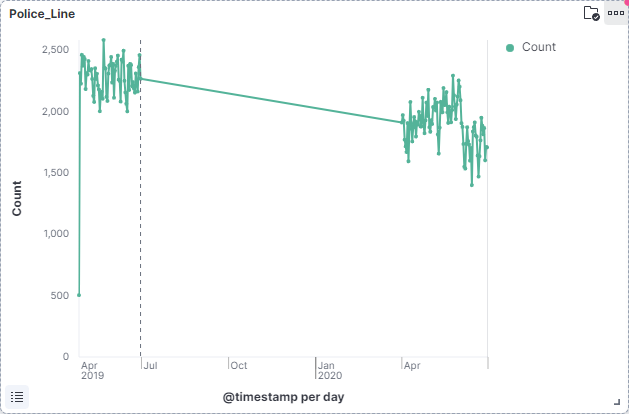
|  |  |
| --- | --- |
| **Number of nodes  per zone** | 1 |
| **Version** | ElasticSearch & Kibana |
| **Set Size** | 1 GB RAM |

**4. Data Analysis**

**4.1 The data analysis of number of police calls in Q2 2019 and Q2 2020**

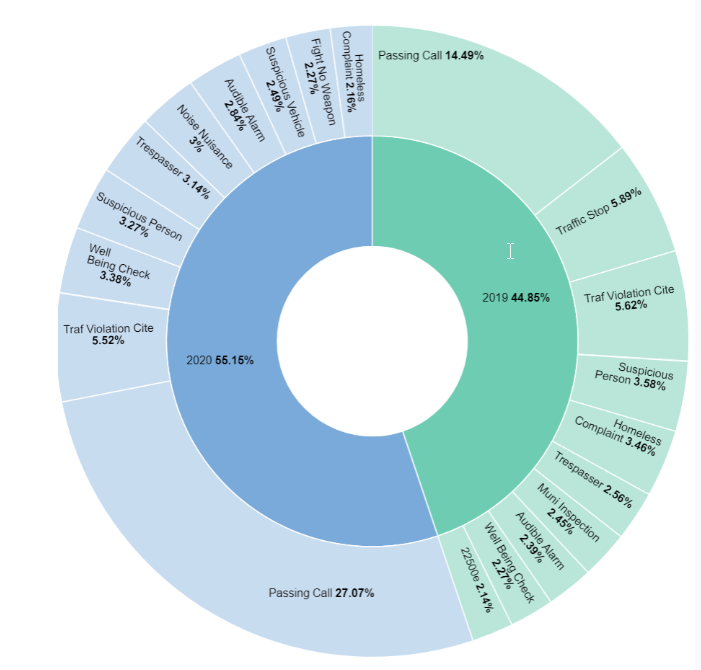
      Number count of police calls represent the numbers of police calls receive by the San Francisco Police Department.

           Calculating the average monthly call volume from daily data involves a straightforward statistical process. For each month, add up the daily call counts to get a total call volume for the month. This sum gives the total number of calls received by the police department over that month. Divide the total by the number of days in the respective month to get the average daily call volume for that month. Doing this provides a clearer picture of the daily workload handled by the police department. This method offers a standardized way to compare call volumes across different time periods, which can be particularly useful for identifying trends, assessing the impact of interventions or policies, such as those related to COVID-19, and for planning and resource allocation within the police department.



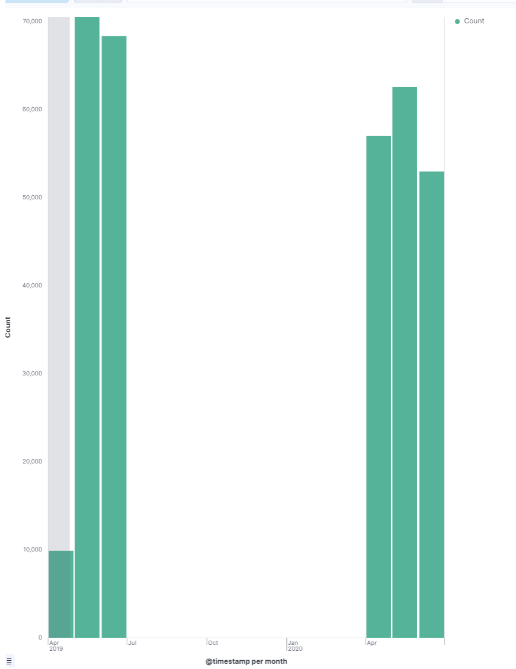
*Figure 1 – Line chart of number of police calls*

      Figure 1 the line graph shows a trend over time on a daily basis. If the downward trend corresponds to the onset of COVID-19 restrictions, it may suggest that there was a decrease in daily police calls. This could be due to lockdowns, curfews, or people being less likely to engage in activities that would typically result in police calls.



*Figure 2 – Types of police calls*

      Figure 3 shows the types of police calls made throughout Q2 of 2019 and 2020. There are many different types of police calls including passing calls, suspicious person, and trespassers. As the pie chart shown, in 2020 and 2019 a good percentage are Passing Calls.



*Figure 3 – Vertical chart of police calls*

Figure 3 shows the number of calls made throughout the Q2 of 2019 and 2020. As seen from the visualization above, in June 2019 there were about 70,000 calls which is the most out of the data we had. It started to decrease in Q2 of 2020, due to Covid-19.

**5. Conclusion**

 In conclusion, leveraging and running a dataset in the Elastic Cloud offers a myriad of benefits, particularly in the context of law enforcement and public safety analytics. The Elastic Cloud, powered by Elasticsearch, Kibana, Logstash, and Beats, provides a scalable and flexible platform that can greatly enhance the analysis of datasets, such as those related to police calls and incidents. Elastic Cloud allows for easy scalability, enabling organizations to handle datasets of varying sizes. This is particularly crucial in law enforcement, where datasets can be extensive and dynamic. The platform can seamlessly scale up or down based on the evolving demands of data analysis. The integration of Kibana with Elasticsearch also enables intuitive data visualization. Law enforcement officials can create dynamic dashboards and visual representations of the dataset, aiding in the interpretation of complex information.

**References**

[1] Dewinter, Maite, et al. “Hot Spots and Burning Times: A Spatiotemporal Analysis of Calls for Service to Establish Police Demand.” Login - EZ Proxy Server - Primary, www-sciencedirect-com.mimas.calstatela.edu/science/article/pii/S0143622822000832#sec3. Accessed 16 Dec. 2023.