Analysis of the Police Department of San Francisco

College of Professional Studies

Northeastern University

Abhinav Jain
Department of Analytics
College of Professional Studies,
Northeastern University, Boston
jain.abhin@northeastern.edu

Chaitang Shah
Department of Analytics
College of Professional Studies,
Northeastern University, Boston
shah.chai@northeastern.edu

Kelvina Pethani
Department of Analytics
College of Professional Studies,
Northeastern University, Boston
pethani.k@northeastern.edu

Siddhartha Manda
Department of Analytics
College of Professional Studies,
Northeastern University, Boston
manda.s@northeastern.edu

Uma Sravanti Kala
Department of Analytics
College of Professional Studies,
Northeastern University, Boston
kala.u@northeastern.edu

Abstract – We will visualize the 'SFPD Incident Report from January 2018 to April 2022 dataset including information about properties in the United States in this paper. This dataset contains incident reports reported on or after January 1, 2018. Officers make these reports, or members of the public can self-report utilizing the SFPD's online reporting system. The data was provided to us for the purpose of understanding incident reports across neighborhoods, the police district, the supervisor district, and the number of investments made within the neighborhood. To visualize these crime data, we are using the tool, which is Tableau. Whenever we need to share the visualization, we must either publish the visualization on Tableau's website or the visualizer can simply insert the data file into Tableau.

Keywords: Police Department, Incident, Report, Resolution.

INTRODUCTION

To better understand crime incidents in San Francisco, we have created visualizations that help us explain what exactly happened. Based on how they were received and what type of incidents they describe, reports are organized into different categories. All occurrences are presented at the intersection level only, and incident reports can contain one or more related incident codes. To protect the privacy of individuals, the data has been anonymized. Datasets like this are limited

and do not capture all the information about crimes and police enforcement, just like incident reports do not. The dataset contains sensitive and personal information about the crime, victim, and witness; therefore, it has been limited due to the privacy settings. The theme behind the report analysis is to present the data in a way that the first dashboard assists about the incidents whereas, the second dashboard provides the information about the reports filed after the incident took place. Coming onto the final dashboard which shows the comparison of different cases that occurred in various districts.

Research Questions:

- What is the percentage of incidents that happen each year?
- What is the total number of incident resolutions?
- What are the number of incidents resolved per category?
- What are the number of incidents reported per month?
- What is the yearly percentage of incidents per category?
- What is the percentage of human trafficking cases each year?

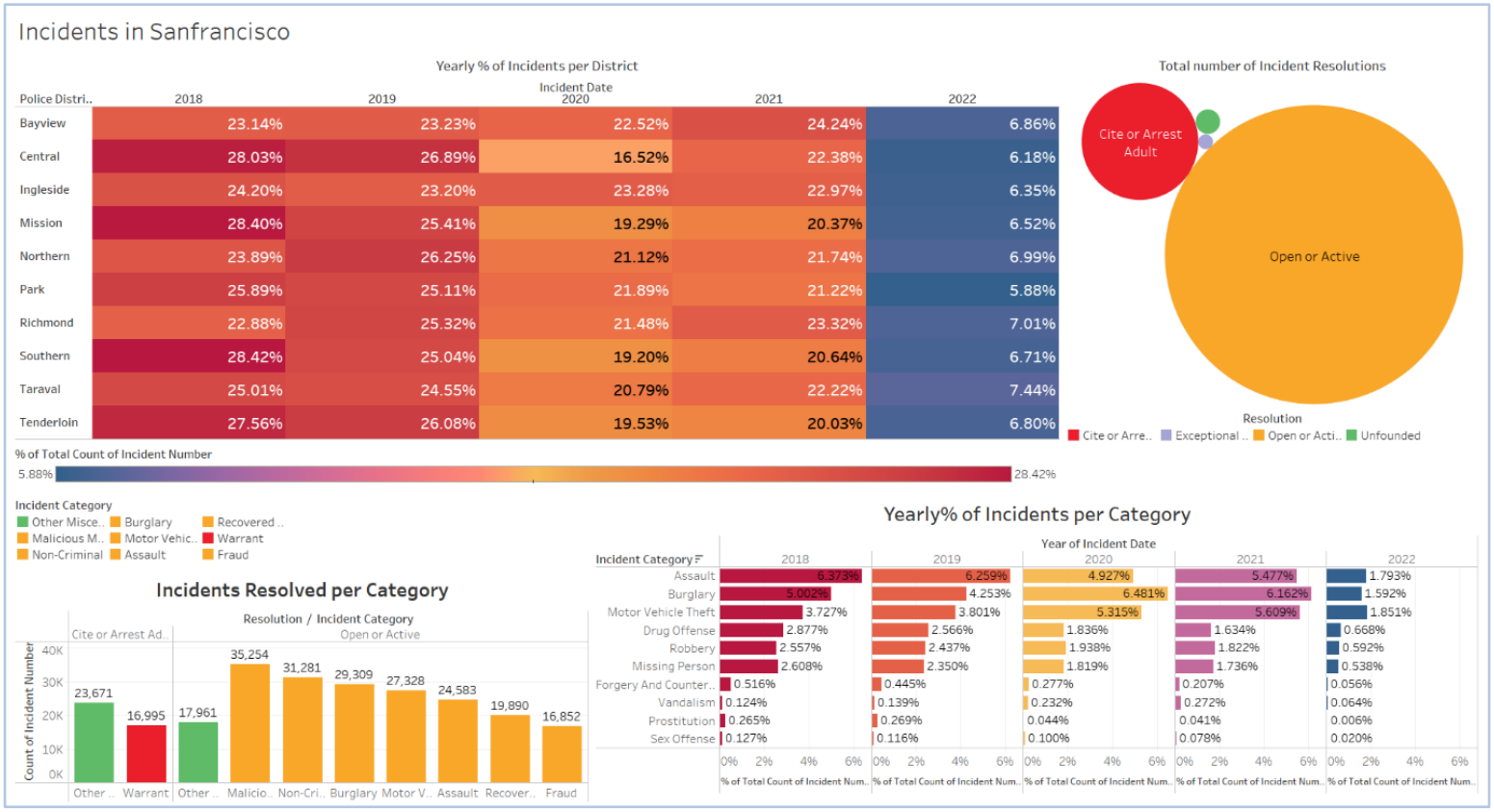
DATA ANALYSIS

1. Dashboard I

Starting with the first dashboard, we are primarily focusing on the incidents happening in the city of San Francisco. Different aspects have been considered to foresee the total number of incidents that occurred in every district in a period of five years. From determining the overall percentage of incidents happening annually to observing the outcome of the matters to spotting the resolved cases, we can easily judge the given dataset.

As shown in the below figure, we can see that the ratio of incidents has changed for every district five years down the line. It is strange to see that the three districts that recorded the highest number of cases in the year 2018 saw fewer cases for the recent year as compared to other districts. Despite the Richmond district having the least cases in the first year registered the highest number of cases in the year 2022. This indicates that the drop or increase in incidents is impacted due to the strictness or liberalness of the police.

For each category, there has been a decline in the incidents with an increase in the cases solved compared to the previous years. However, we can detect that the police department was not successful in solving the fraud cases unlike the mischiefs and non-criminal matters which have the highest resolved count. Also, the arrested count is way less than the total active cases along with a significant number of cases that are still unfounded. This means that the police department is surely not able to cope with every type of incident in all the given districts.



Dashboard I: Incidents in San Francisco

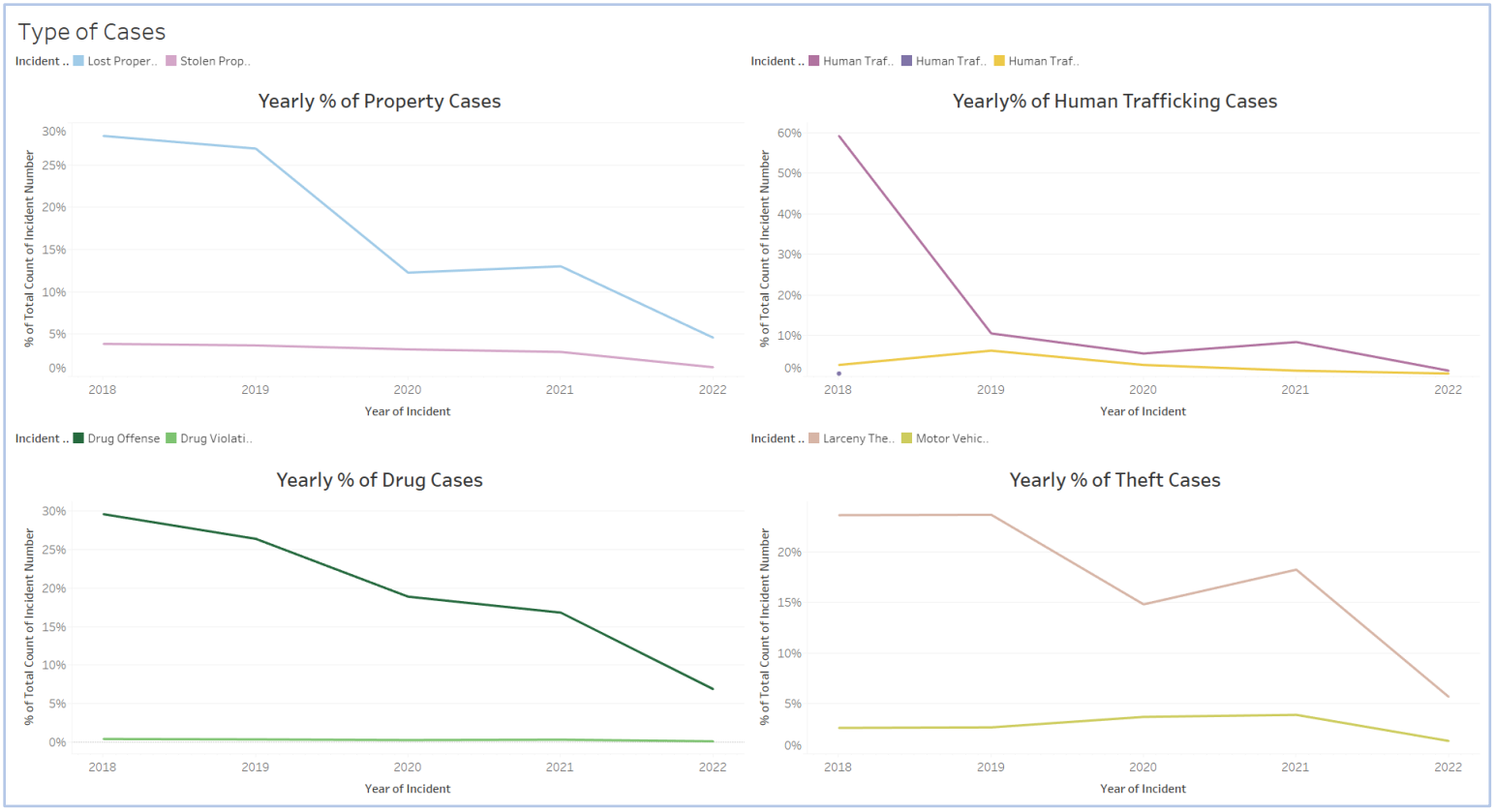
2. Dashboard II

In this dashboard we have used two bar charts and a line chart that displays the number of incidents reported. In the first graph we will be focusing on the Police Districts like Bayview, Central and Park and the total sum of incident numbers for five consecutive years from 2018 to 2022. Secondly, we have a bar graph that represents the year of the report date and the count of report time. Lastly, a line graph which has the total incident percentage per year with respect to reporting type considering the total count of the incident number and the year of the incident date.

Dashboard II: Incidents Reported

3. Dashboard III

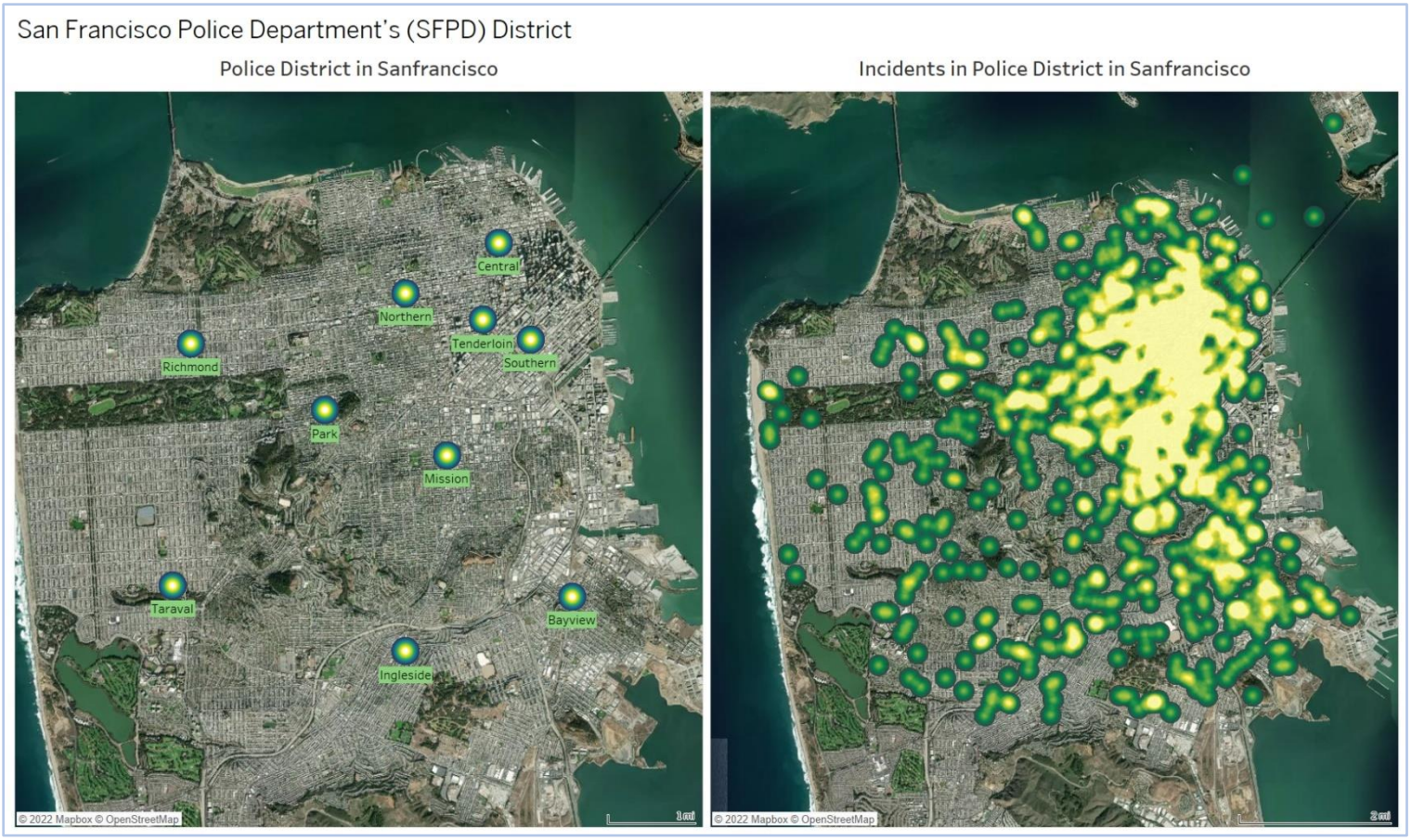
In the dashboard three its mainly about the types of cases which consists of property cases, drug cases, theft cases and trafficking cases. In all the graphs it is mainly about percentage of total count of the incident number and the year of the incident from the year 2018 to 2022. As the number of cases is trending low in each case in comparison to the previous years.



Dashboard III: Types of Cases

4. Dashboard IV

Overviewing the final dashboard, we are using two density maps for showing the number of police divisions in the San Francisco district and the number of incidents that happened in the San Francisco police district. We can clearly see that the number of incidents happening near Richmond and Taraval is less compared to the incidents happening near the Central, Northern, Southern, and Tenderloin. The San Francisco police department needs to improvise its system by increasing the patrolling and alerting the people about the same. The below map shows the divisions of the police department in San Francisco city. Considering the range of each division we can acknowledge that the Richmond neighborhood is the safest. Despite having a maximum range compared to other neighborhoods Richmond has the minimal number of cases which is 543,941 as compared to Bayview and Central having the highest number of cases.



Dashboard IV: SFPD District

CONCLUSION

Crime data can be processed and analyzed efficiently using visualization techniques.

My experience with Tableau has been a great one so far. Data can be displayed in more than 24 different graphical views with Tableau. It is a very effective tool for graphical representation. Tableau allows us to create dashboards more easily and, in less time, even if the dataset is very complex or very large. Additionally, I appreciated that the palette was color-blind friendly, which made the visualizations more accessible. It is crucial to understand that utilizing all the capabilities of a visualization tool can result in complicated and confusing visuals. Keeping things simple and following good design guidelines are also important. In this study, we will primarily focus on the information police departments need to investigate incidents that have occurred in the past, for them to reduce cases in the future.

1. Approximately half of all crimes are committed in the central, northern, and mission sides of the police district. Our recommendation is to increase surveillance/ patrolling in the respective area by the police department.

2. It is recommended to implement continuous surveillance in certain neighborhoods according to the crime rates and the times at which crimes typically occur.
3. Throughout the years, burglary and assault have risen to the top of the list of crimes. The Police must monitor the plausible reasons behind these crimes.
4. Crime trends spike on certain days of the week which leads to tightened security and alerting the public.
5. The crimes reported were more as compared to resolved, this indicates the fast movement in solving the crimes asap.
6. Human trafficking commercial sexual act had to the chart in the year 2018 although the number went down this showed some precautions and actions were taken.
7. Most of the cases resolved were of malicious mischief and least resolved were of fraud and others were also resolved, but the pace for resolving fraud and warrants should be increased.

REFERENCE

- *Tableau: Business Intelligence and Analytics Software*. (n.d.). Tableau. https://www.tableau.com/
- *Free Training Videos - 2022.1*. (n.d.). Tableau. https://www.tableau.com/learn/training/20221
- *Tableau 10 for Data Scientists (2017) Online Class | LinkedIn Learning, formerly Lynda.com*. (2016, September 12). LinkedIn. https://www.linkedin.com/learning/tableau-10-for-data-scientists-2017?trk=lynda_redirect_learning
- *SFPD Incident Report: 2018 to Present - DataSF Dataset Explainers*. (n.d.). Unknown. https://datasf.gitbook.io/datasf-dataset-explainers/sfpd-incident-report-2018-to-present
- *SFPD Incident Report: 2018 to Present - DataSF Dataset Explainers*. (n.d.). Unknown. https://datasf.gitbook.io/datasf-dataset-explainers/sfpd-incident-report-2018-to-present