

Team 1 | CS506 | Deliverable 4

Team 1: Anton Barash, Henry Liu, Himanshu Patil, Animikh Aich

Project Manager: Yagev Levi

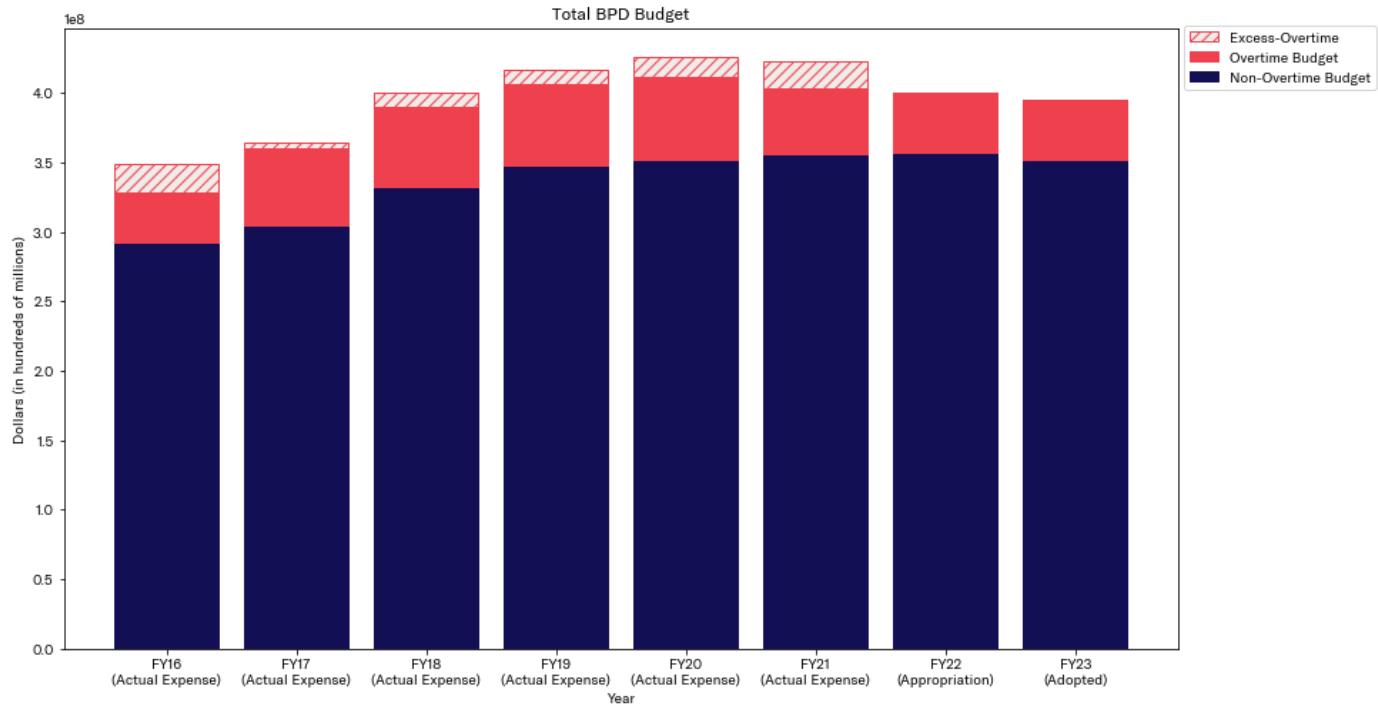
Team Lead: Himanshu Patil

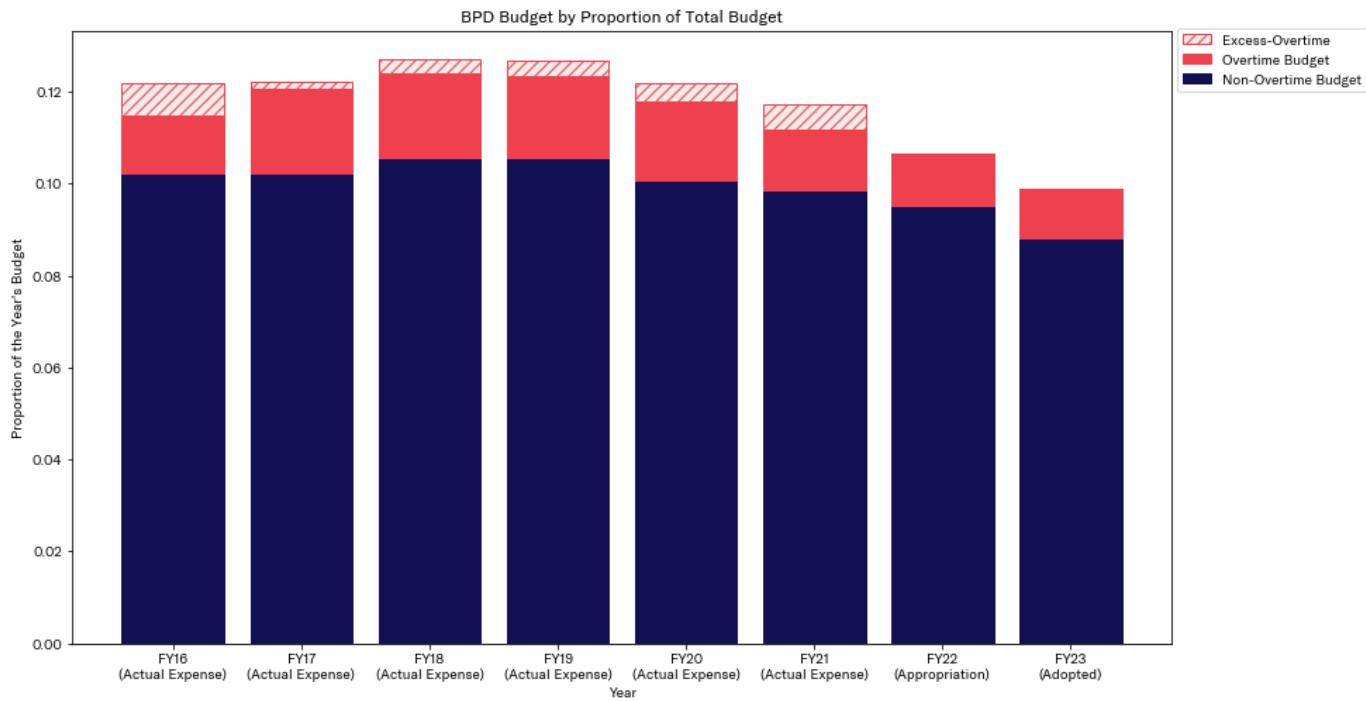
December 5, 2022

Base Project

1. Total BPD Budget vs. previous year

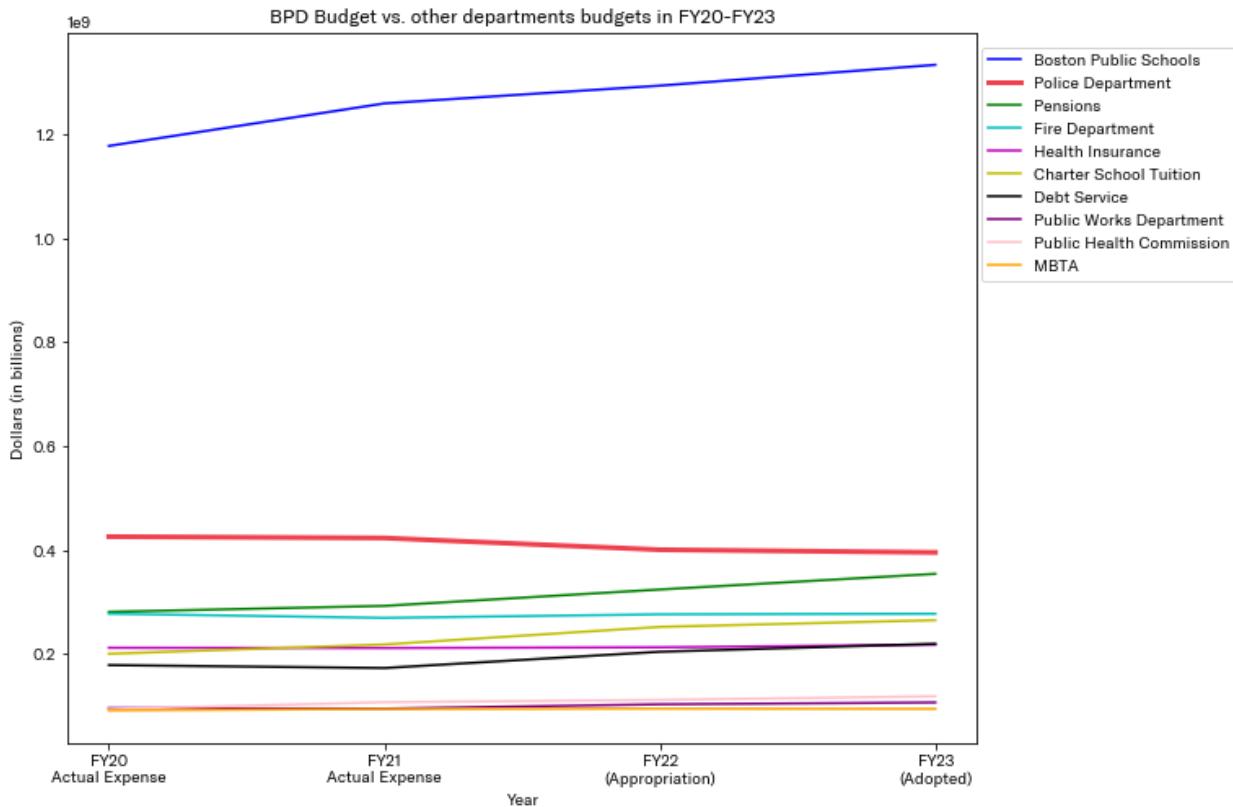
The total Boston Police Department Budget has been decreasing steadily over the past 3 years. In relation to the total budget of that year, around 10-12% of it is allocated to the BPD. When looking at the BPD budgets in proportion to the total budget, it has been decreasing over the past 5 years.

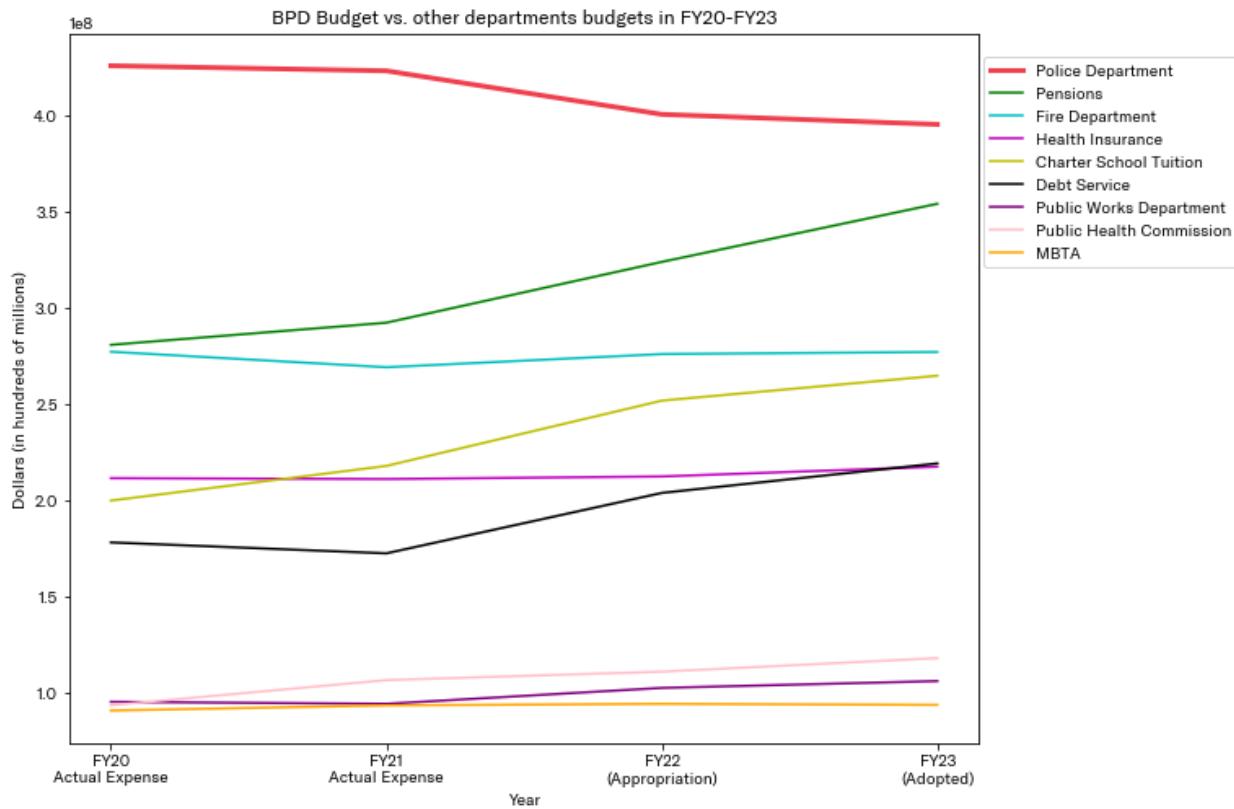




2. BPD Budget vs. other departments budgets + change over time

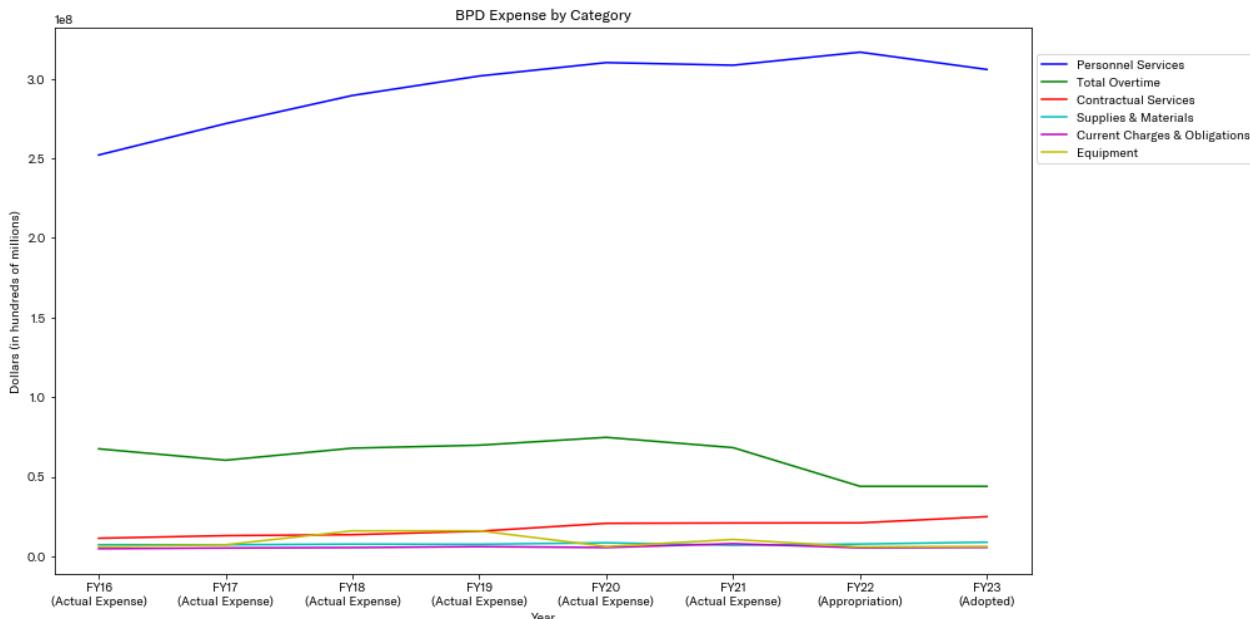
In comparison to all the other departments, the BPD took second most of the total budget throughout the past three years.

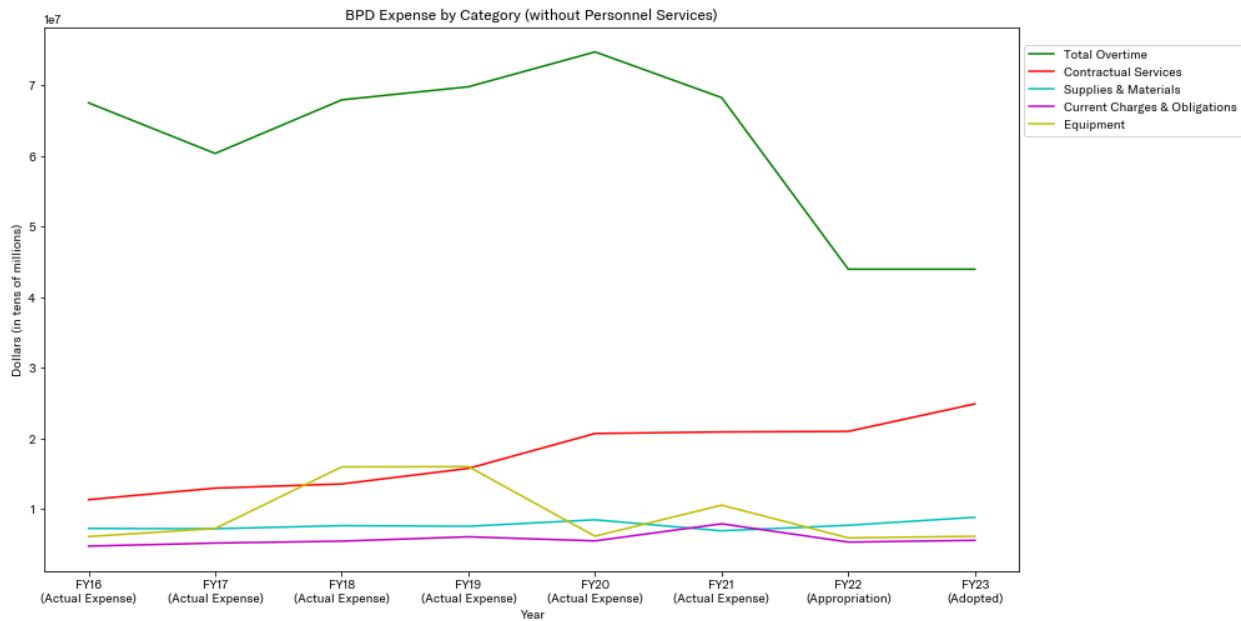




3. BPD Budget spend categories + change over time

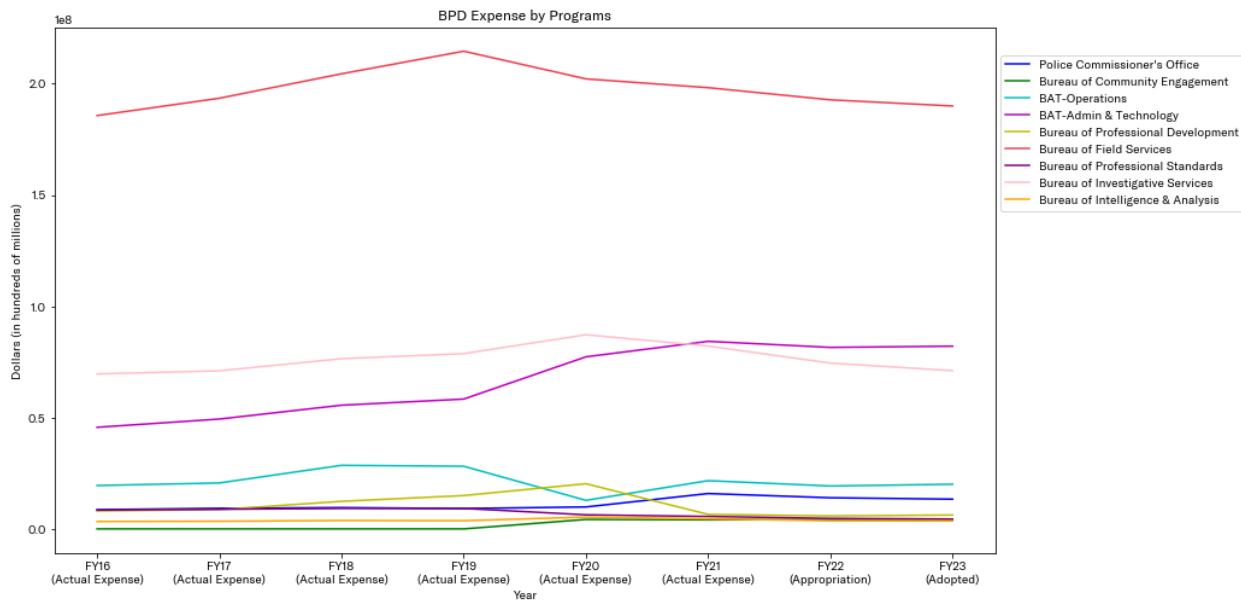
Most of the budget allocated to BPD is spent for personnel services. The graph on the left shows Overtime as the next highest expenditure for the BPD, which is a subcategory of personnel services. The rest of the categories are minuscule in comparison.





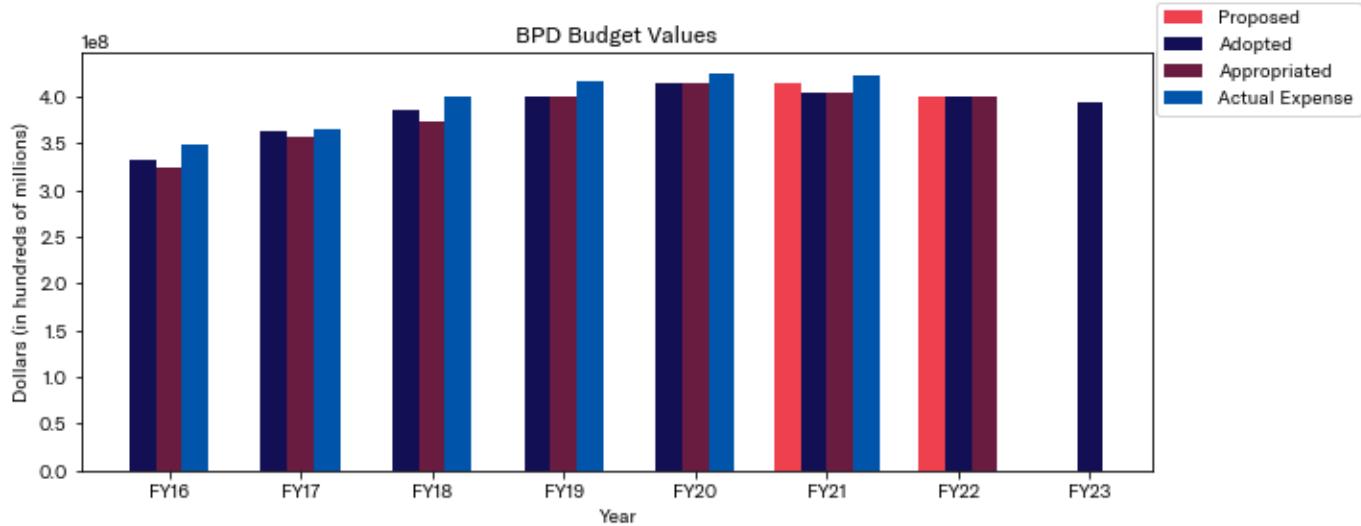
3.5 BPD Budget spendings per program + change over time

The Bureau of Field Services spends a majority of the BPD budget. Bureau of Investigative Services and also Professional Standards, which only recently shot up in spendings, are the next most spending programs.



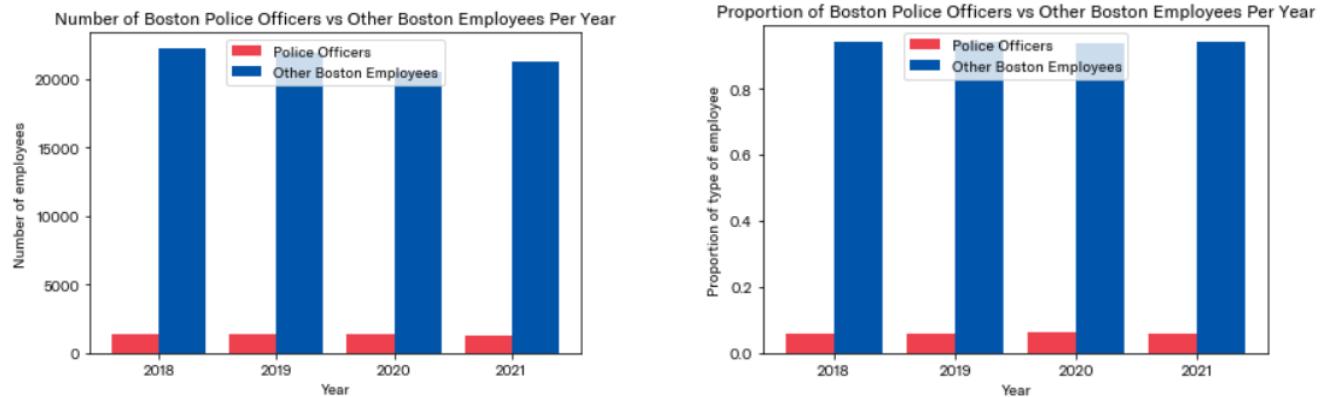
4. BPD adopted vs proposed budget

The proposed and adopted amounts for FY21 AND FY22 are around the same, demonstrating that the proposed amount has been accurate as a basis when used to determine the allocation of the budget. There is also a correlation seen between the adopted and appropriated budget, such that the adopted budget has been a determining factor for the appropriated budget. As seen from the past years, the actual expense always exceeds the appropriated budget.



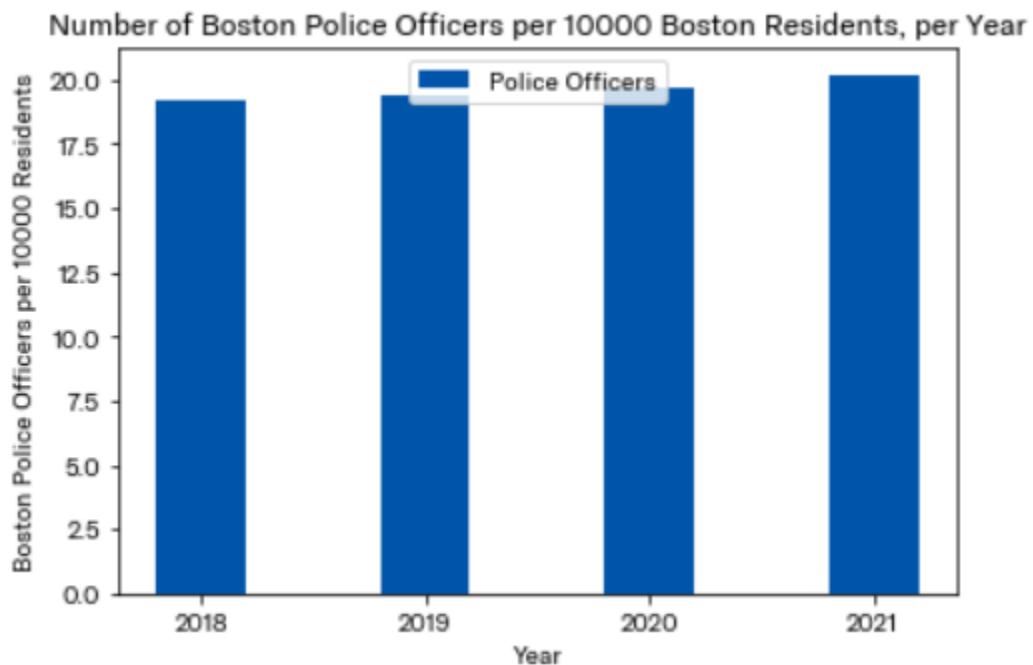
5. Number of police officers (compared to other Boston employees)

The number of police officers compared to the other Boston employees stays pretty stagnant throughout the years 2018-2021.



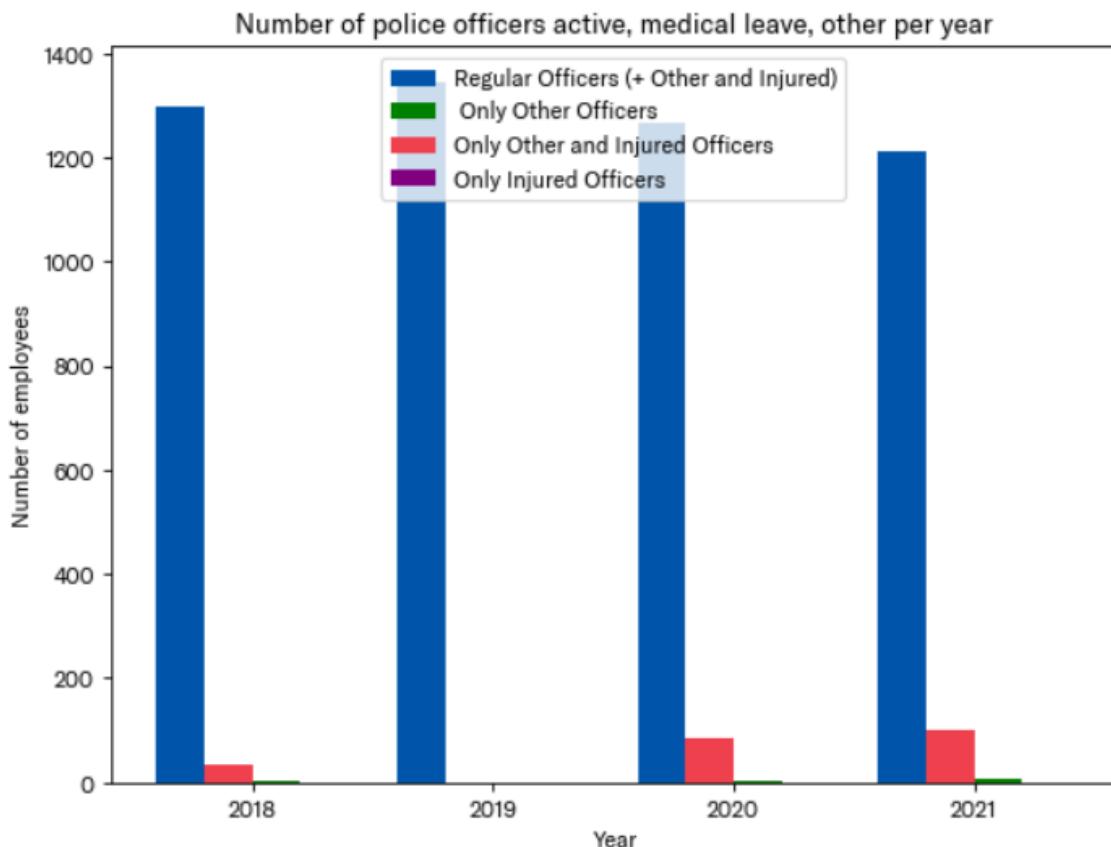
6. Number of police officers per 10,000 residents

The number of police officers per 10000 Boston Residents steadily increased from 2018-2020 until 20, then went down for 2021.



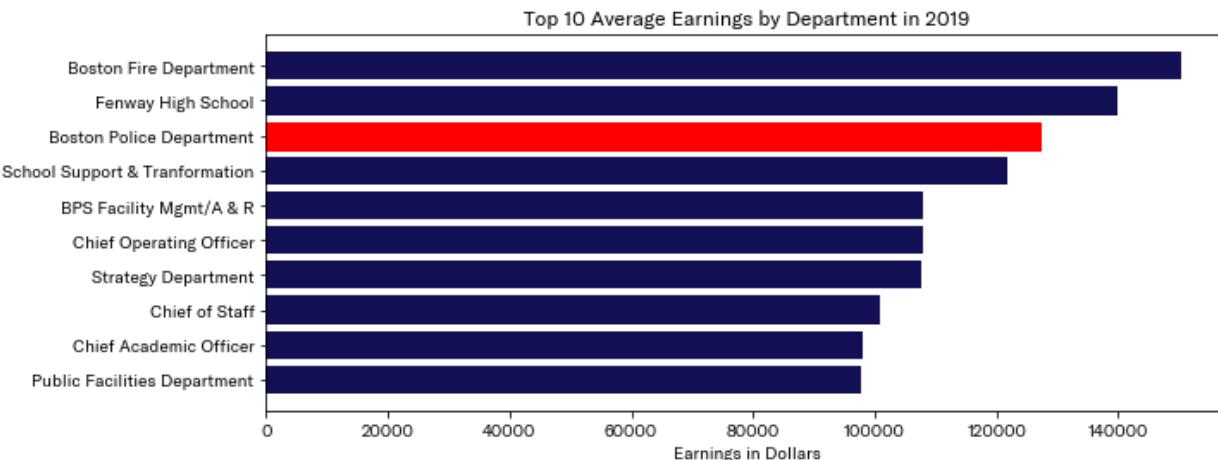
7. Number of police officers active, medical leave, etc.

The number of other and injured officers went up in 2020 and then further up in 2021. It is also notable that there are no officers who are only collecting injured pay, all officers at least collect Other + Injured pay, or collect Regular pay.

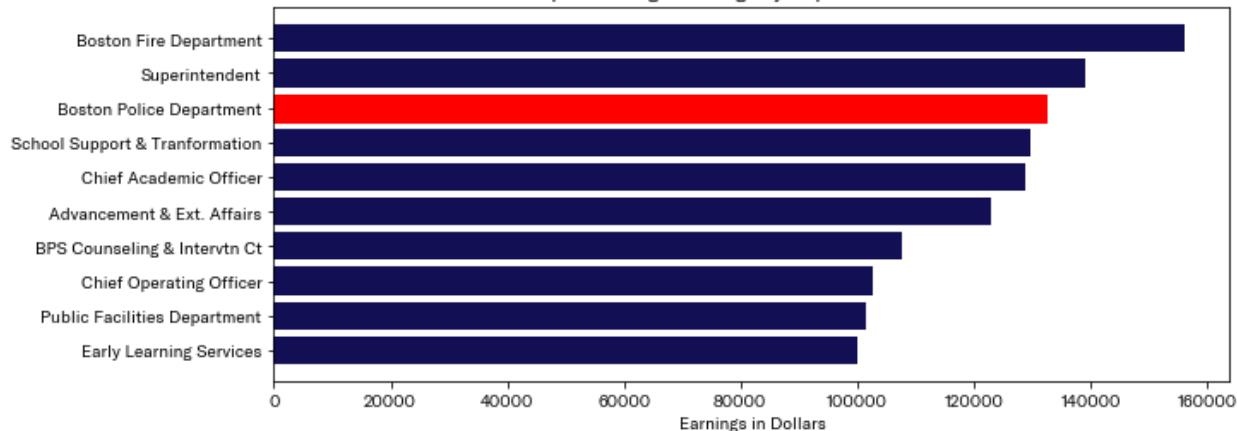


8. Employee compensation comparison to non-BPD employees

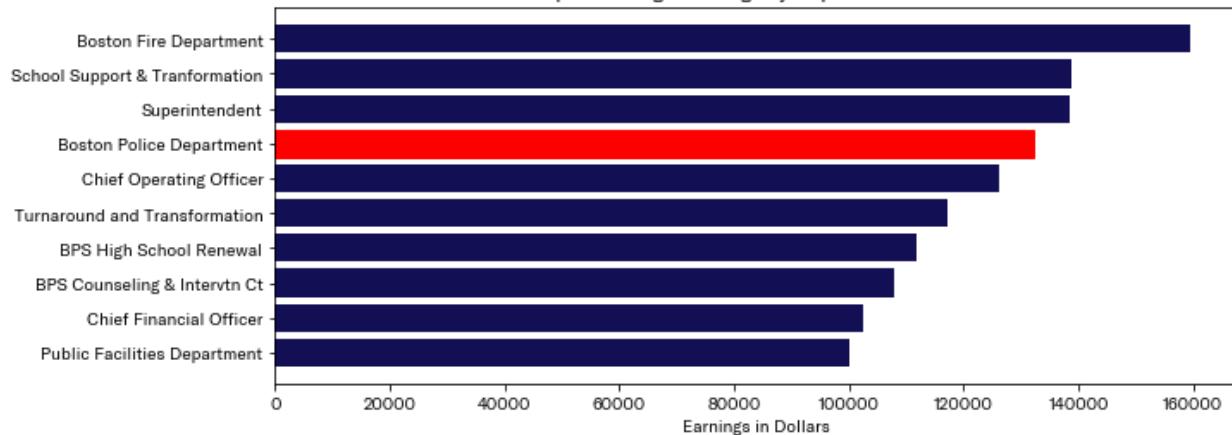
The average employee compensation stayed stagnant for police officers at around 130000 for all 3 years. As a rank, the police department goes down from 3 to 4 for 2021 (school support and transformation surpasses it).



Top 10 Average Earnings by Department in 2020



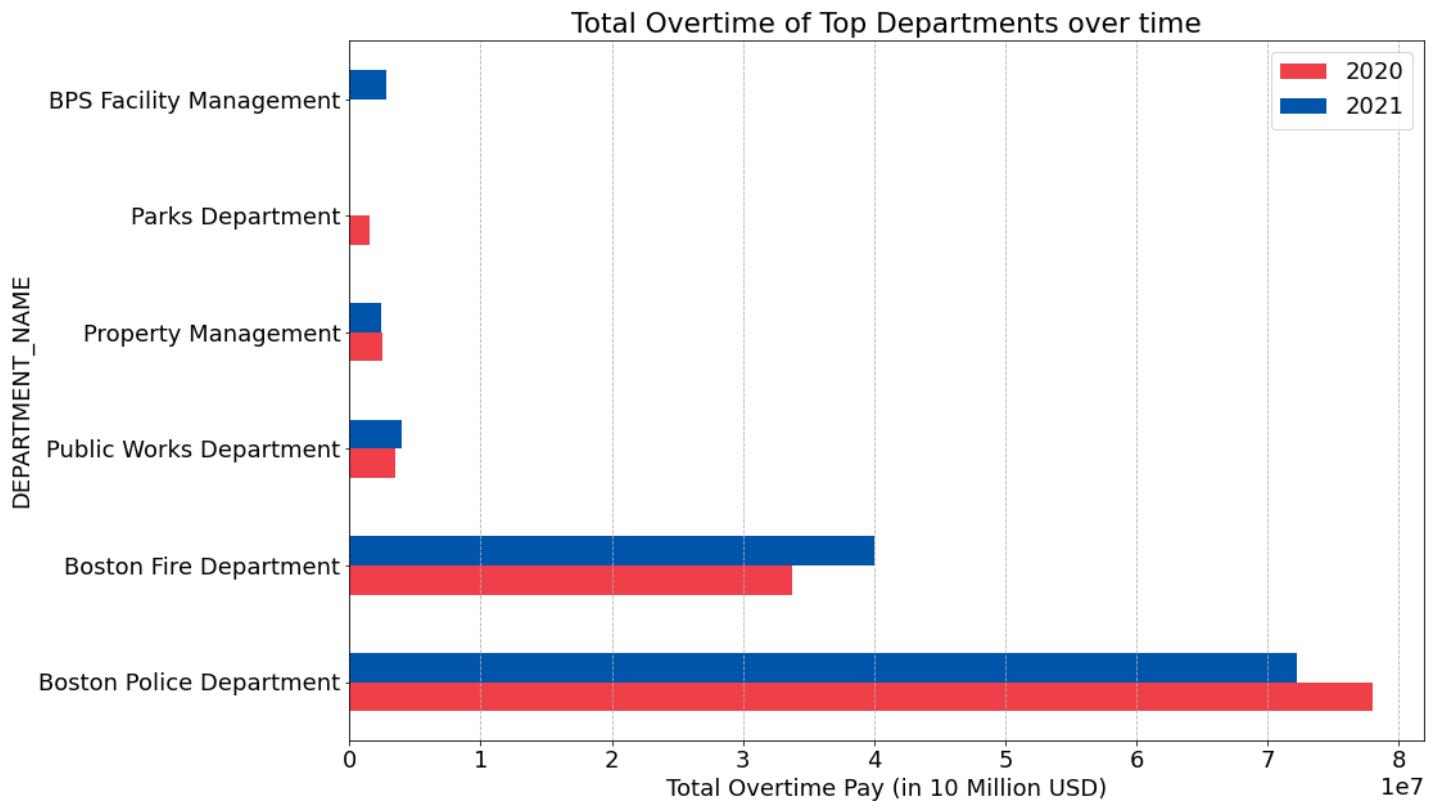
Top 10 Average Earnings by Department in 2021



9. Total overtime of top departments + over time

From the graph it can be observed that Boston Police Department has the highest total overtime spends among all departments, and is nearly double that of Boston Fire Department, which comes second.

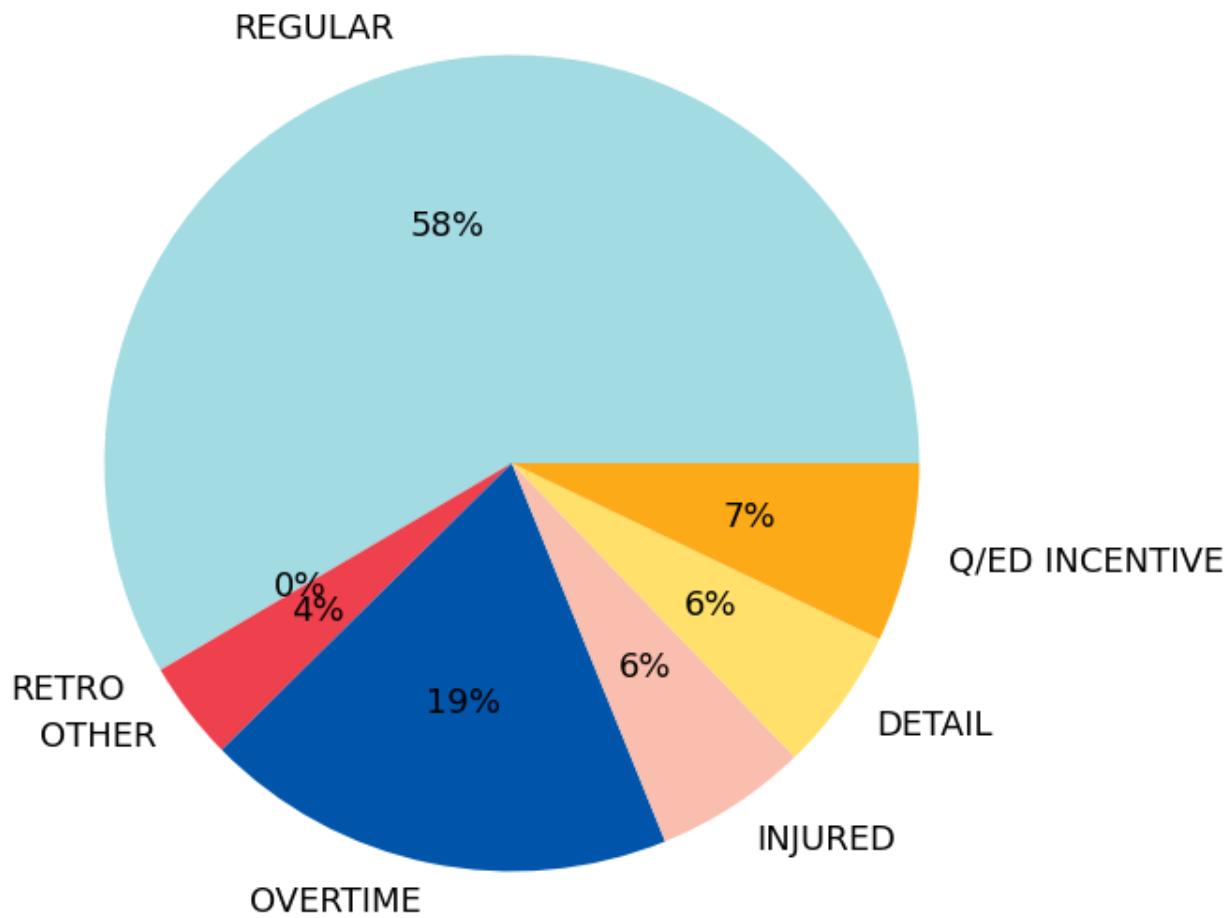
However, we also observe that the total overtime compensation for Boston PD has reduced between 2020 and 2021.



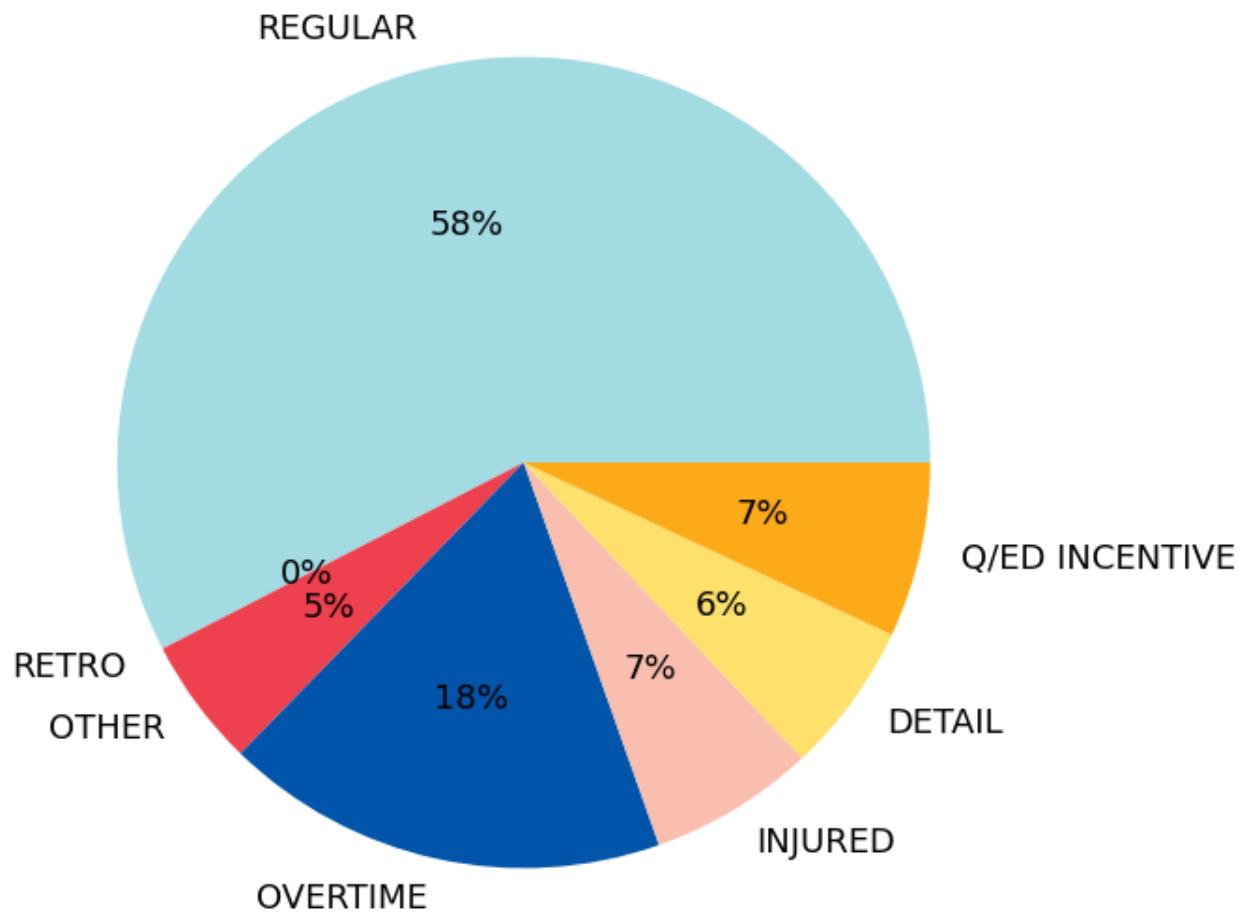
10. Payroll distribution over time (base, overtime, retro, detail, etc.)

From the charts we can conclude that the payroll distribution has remained consistent over the year with a slight decrease in overtime percentage in 2021 as compared to 2020.

Payroll Distribution 2020



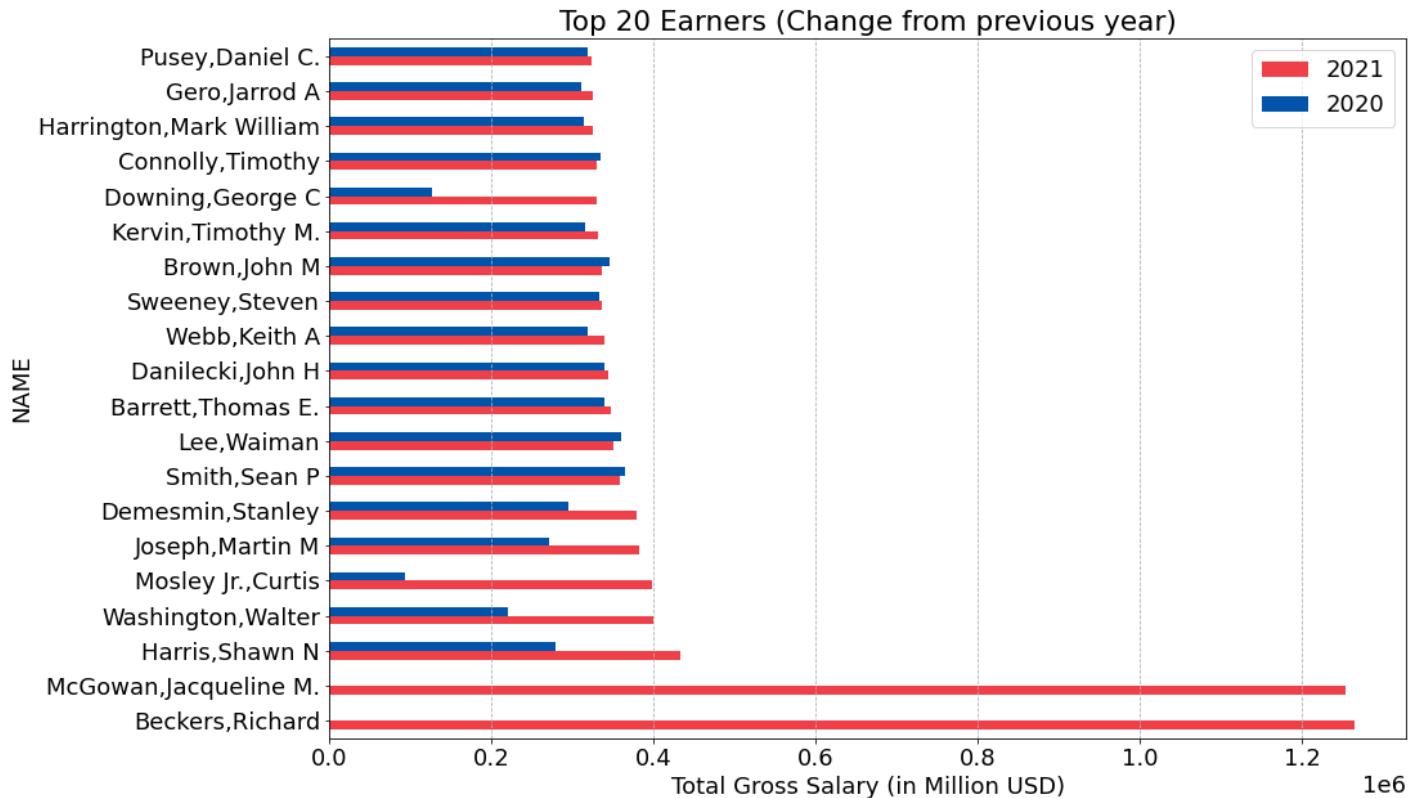
Payroll Distribution 2021



11. Top 20 Earners (change from previous year)

From the graphs The top two earners are earning significantly higher than the others.

Additionally, we also observe that the third to fifth top earners has had significant pay raise (gross salary) as compared to the previous year.



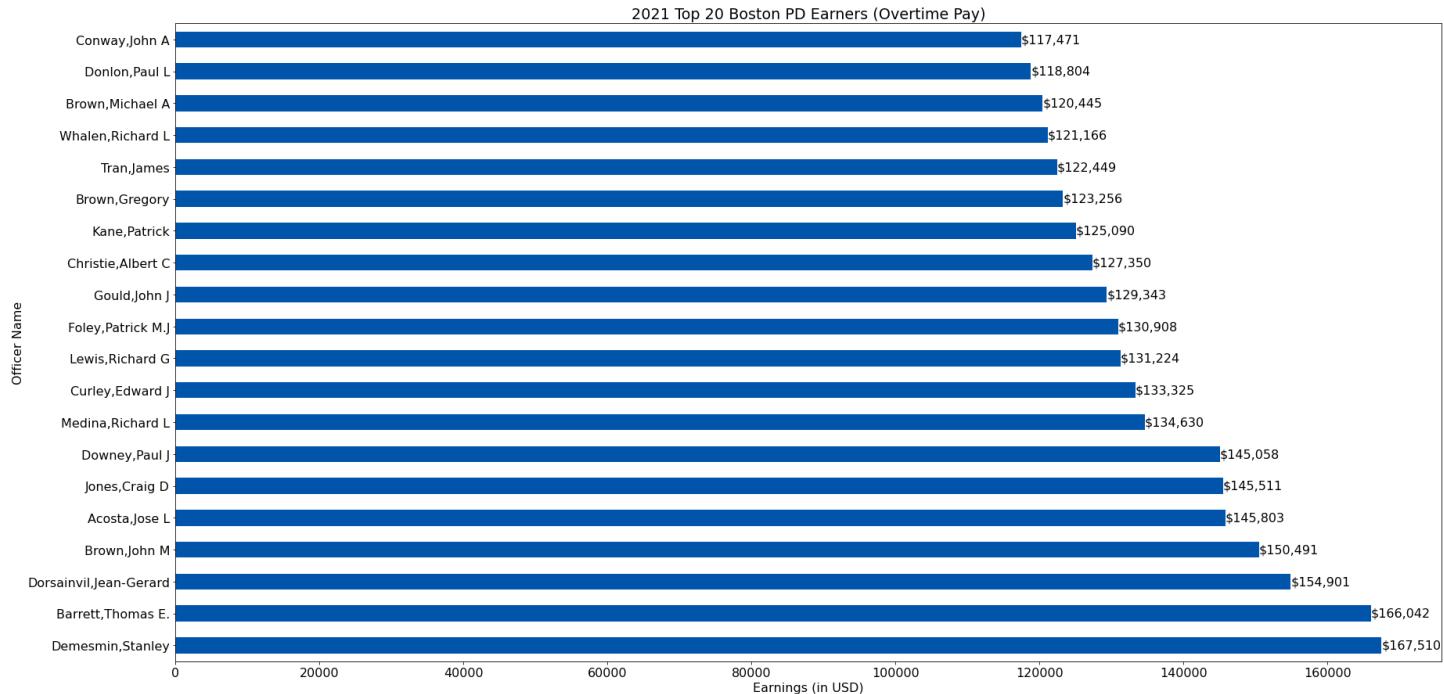
12. Top 20 Earners (OT, Base, Other)

The top 20 earners have related/similar pay scales and there does not seem to be any apparent pay disparity among them.

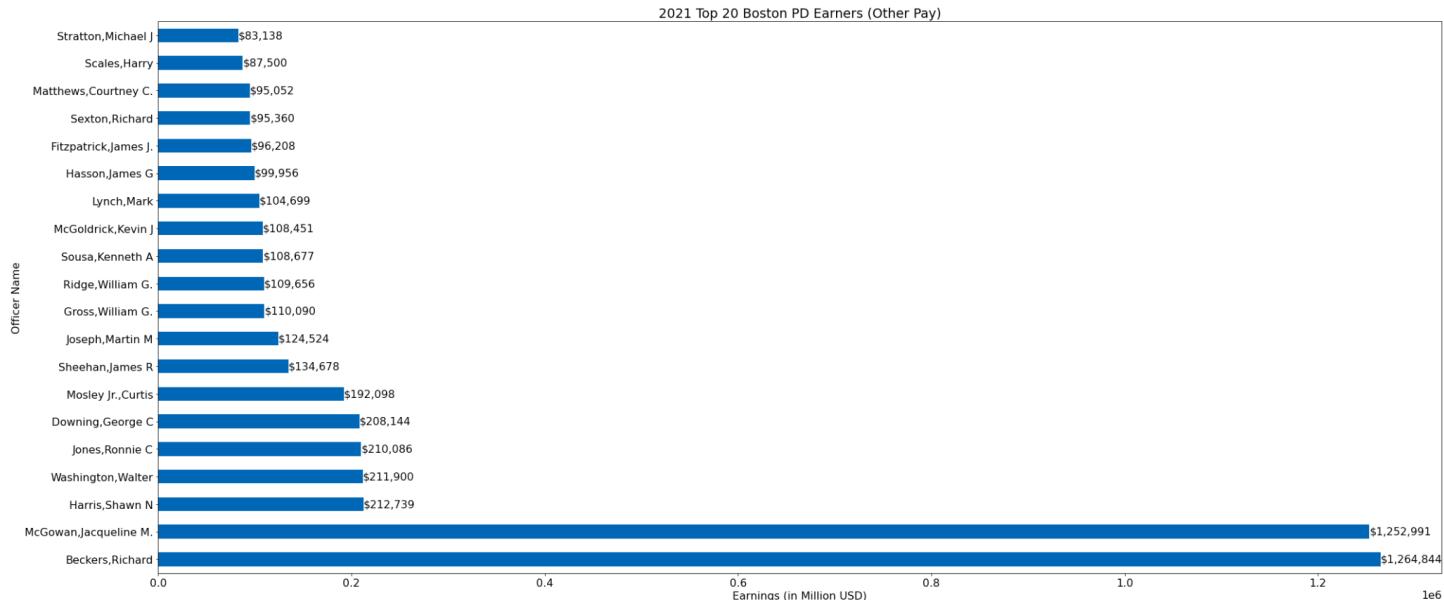
The base pay seems to be similar for officers with the same rank/experience

Based on the graph we can see a clear difference between the top earner and the top 20th earner in terms of Overtime Pay. This may be due to higher overtime hours or higher overtime rate, or both.

Further analysis can be done by comparing the overtime pay with overtime hours.



There is a clear pay disparity between the officers in the “other” category, where the top two officers get paid over a million USD, where the third highest officer gets paid around 200K USD.

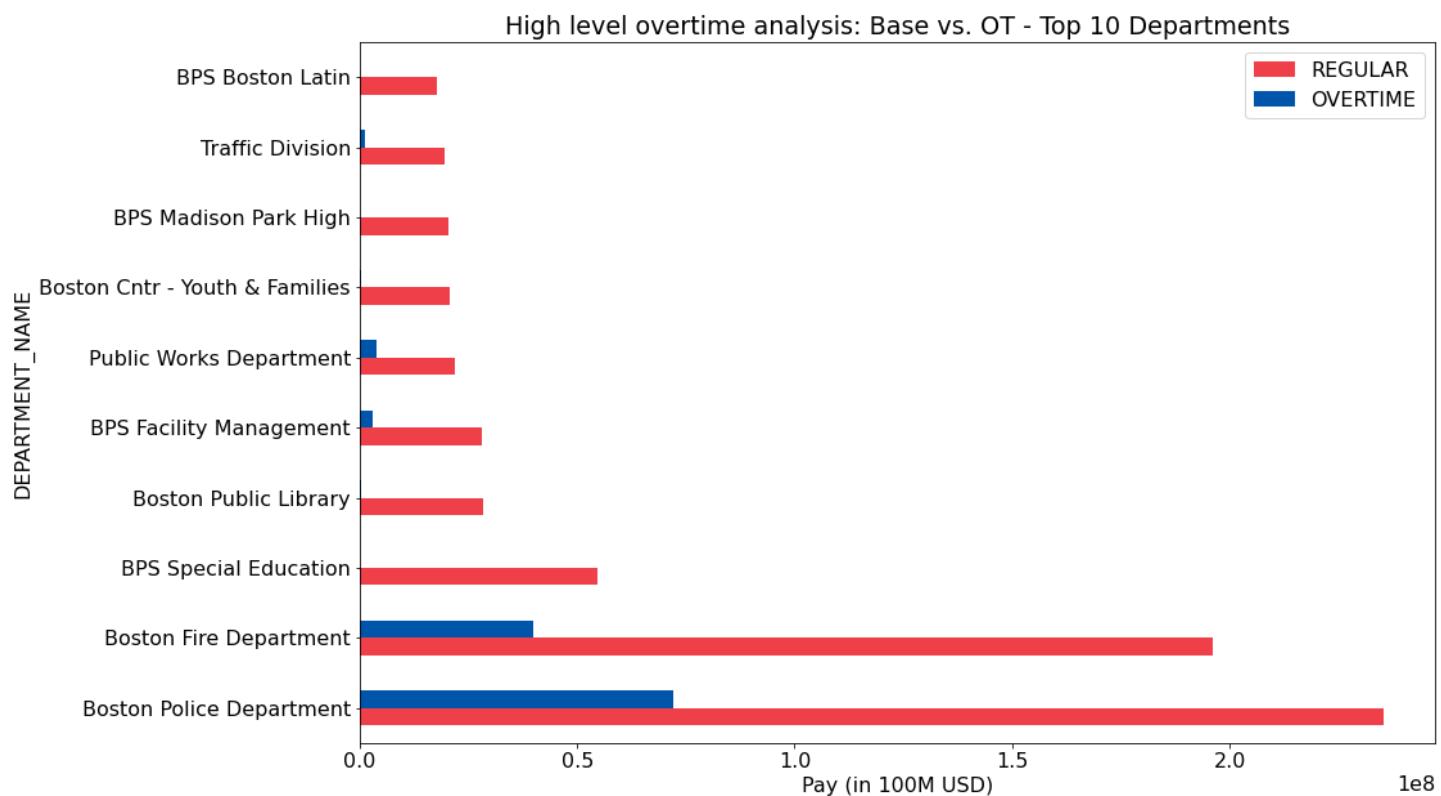


13. High level overtime analysis (base salary vs. overtime vs. other departments)

Among the top 5 departments in terms of Base Pay, Boston Police Department has the highest total compensation in terms of base pay and overtime pay.

It is comparable only to Boston Fire Department, which is second in the list.

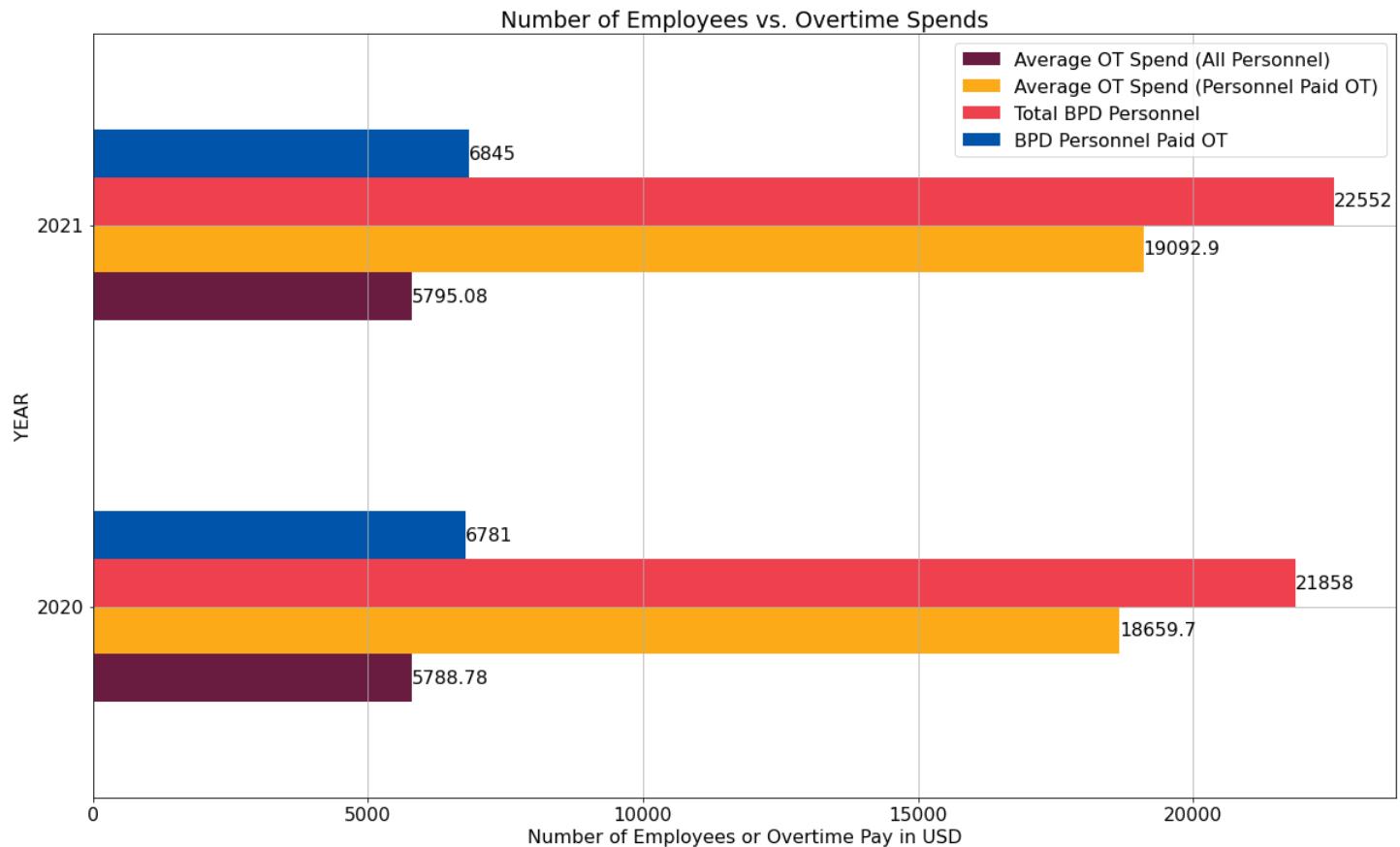
The other departments have several times lower compensation in comparison.



14. # of employees vs. OT spend

From the graph we can see that the total number of Boston PD personnel who are paid overtime is less than half of the total number of personnel in the department.

The distribution of overtime compensation and personnel has not changed much between 2020 and 2021.



Extension Project

Data Analysis:

Crime Incident Reports:

<https://data.boston.gov/dataset/crime-incident-reports-august-2015-to-date-source-new-system/resource/313e56df-6d77-49d2-9c49-ee411f10cf58>

The crime incident reports dataset contains information about incidents which a report has been filed that requires a police officer to go and investigate. Interesting data which comes from this dataset are the types of incidents, the time it occurs, and the location. The purpose of this dataset for our extension project is to use the location of each incident and compile a map to find which areas are most populated with crime reports. This will then be used to see the relationship between employees and crime reports.

Employee Earnings:

<https://data.boston.gov/dataset/employee-earnings-report>

The employee earnings dataset is the same dataset we used for the base project when analyzing questions 5-14. It contains data regarding the earnings of each worker for the City of Boston facilities, as well as their location in the form of a ZIP code. The main data being used from this data set will be the ZIP code and the earnings of the employees under the Boston Police Department.

FIO Reports:

[BPD Field Interrogation and Observation \(FIO\) - FIO Records 2021 \(Mark43\) - FieldContact Table - Analyze Boston](https://data.boston.gov/dataset/BPD-Field-Interrogation-and-Observation-FIO-FIO-Records-2021-Mark43-FieldContact-Table-Analyze-Boston)

The FIO report data is the dataset of all recorded FIO reports for a given year. It contains data pertaining to the report, like date, officer who was involved, and more, along with the ZIP code of the report. The main data being used from this data set will be the ZIP code of the FIO reports, seeing how many reports are from each ZIP code.

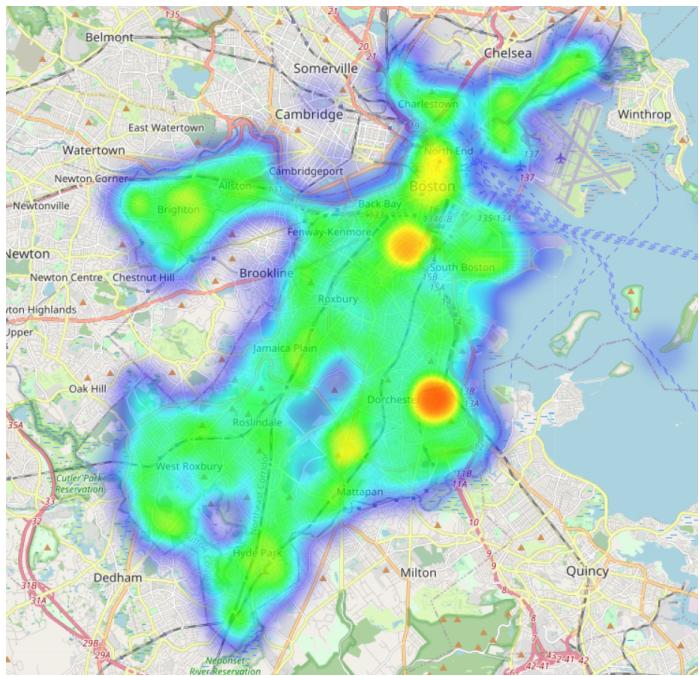
Connecting the datasets:

We can already connect the FIO and Employee Earning data using the ZIP codes in each, but this will be more tricky for crime incident reports. We found a potential module, geopy, which can convert the latitude and longitude in those reports to ZIP codes, but thus far running that to convert all records of crime incidents is taking far too long, so we're exploring other options.

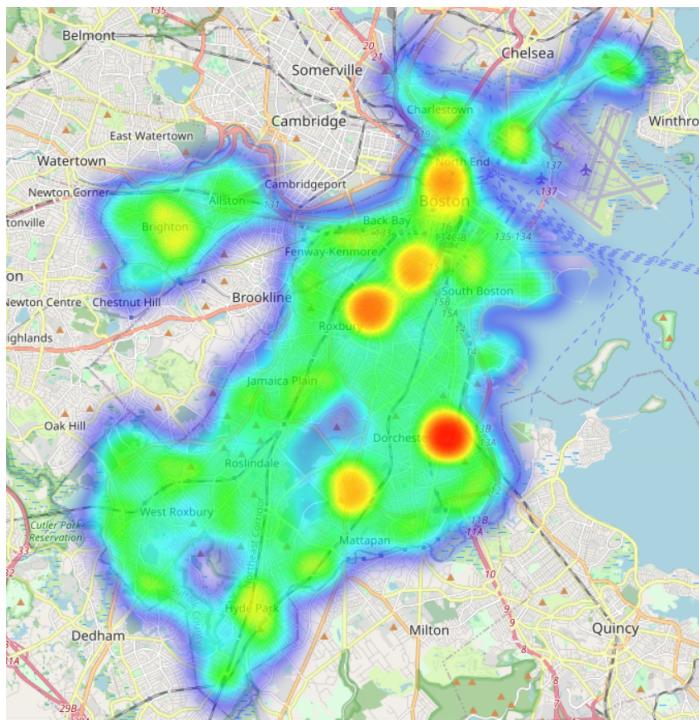
Analyzing the Questions:

1. Crime Incident Reports Heat Maps over Time

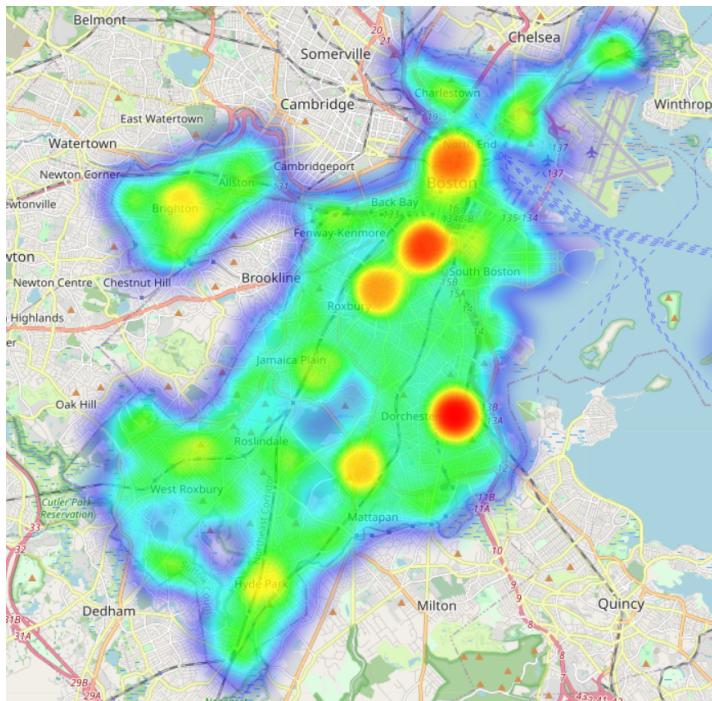
2022:

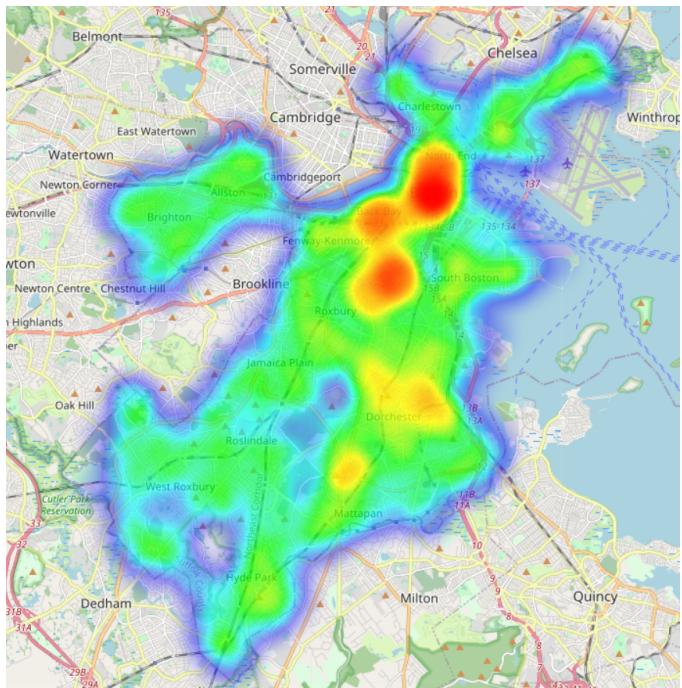


2021:



2020:

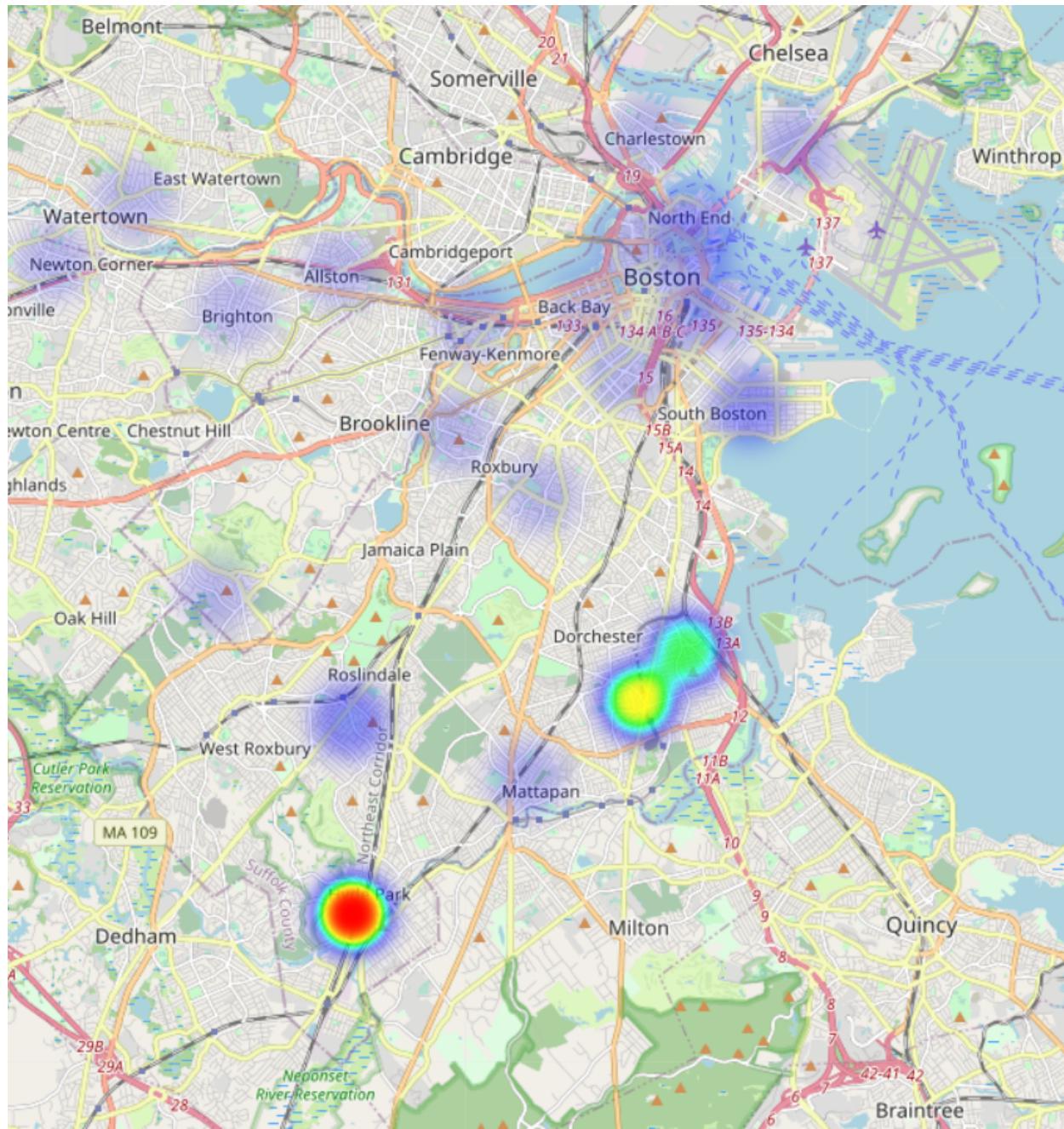




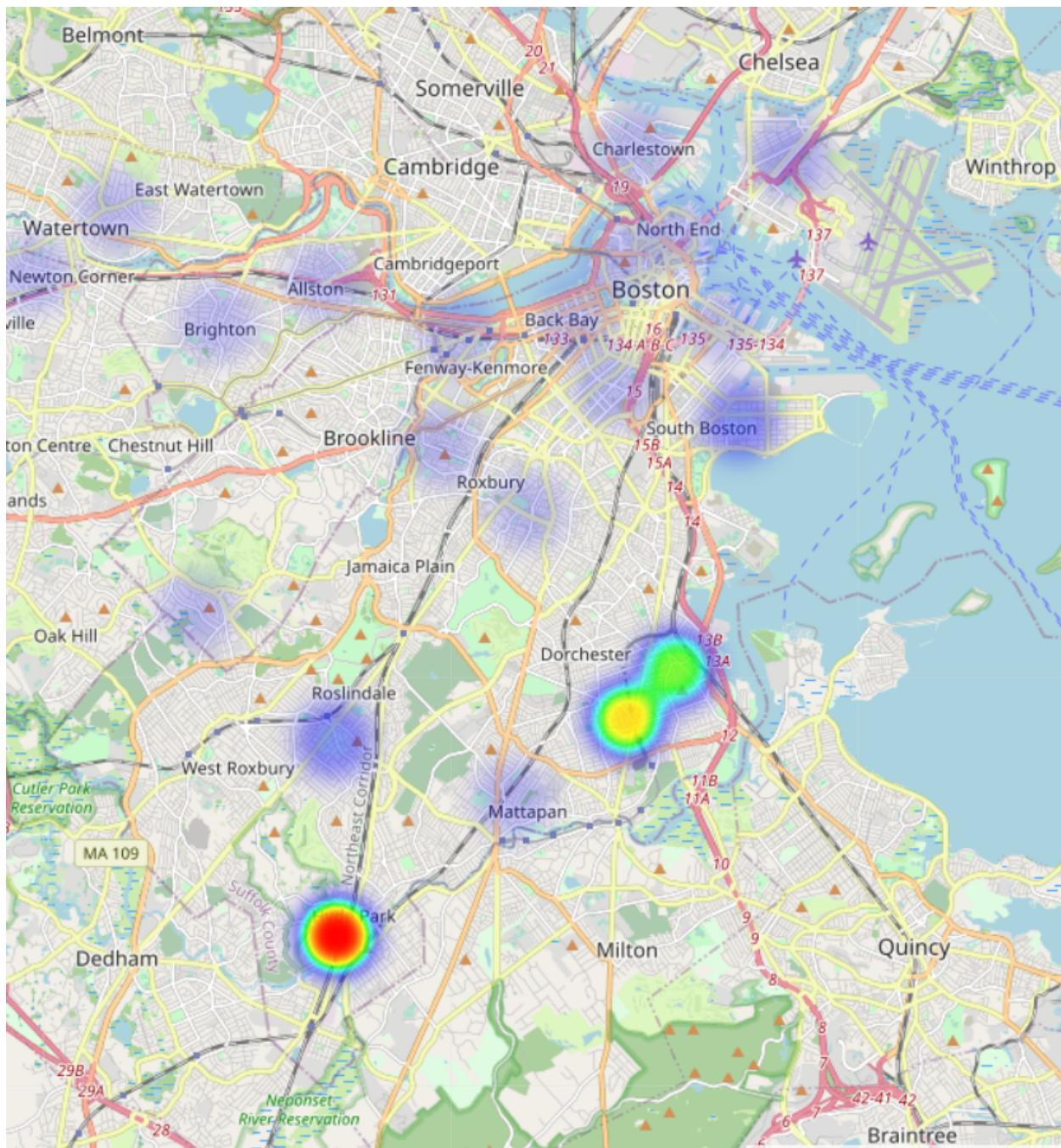
We can see that all the heatmaps are roughly similar, but that there appears to have been more active crime in high-crime areas during 2018, 2020, and 2021, due to more red hotspots during those years.

2. Police Officer Dispersion Heat Maps over Time:

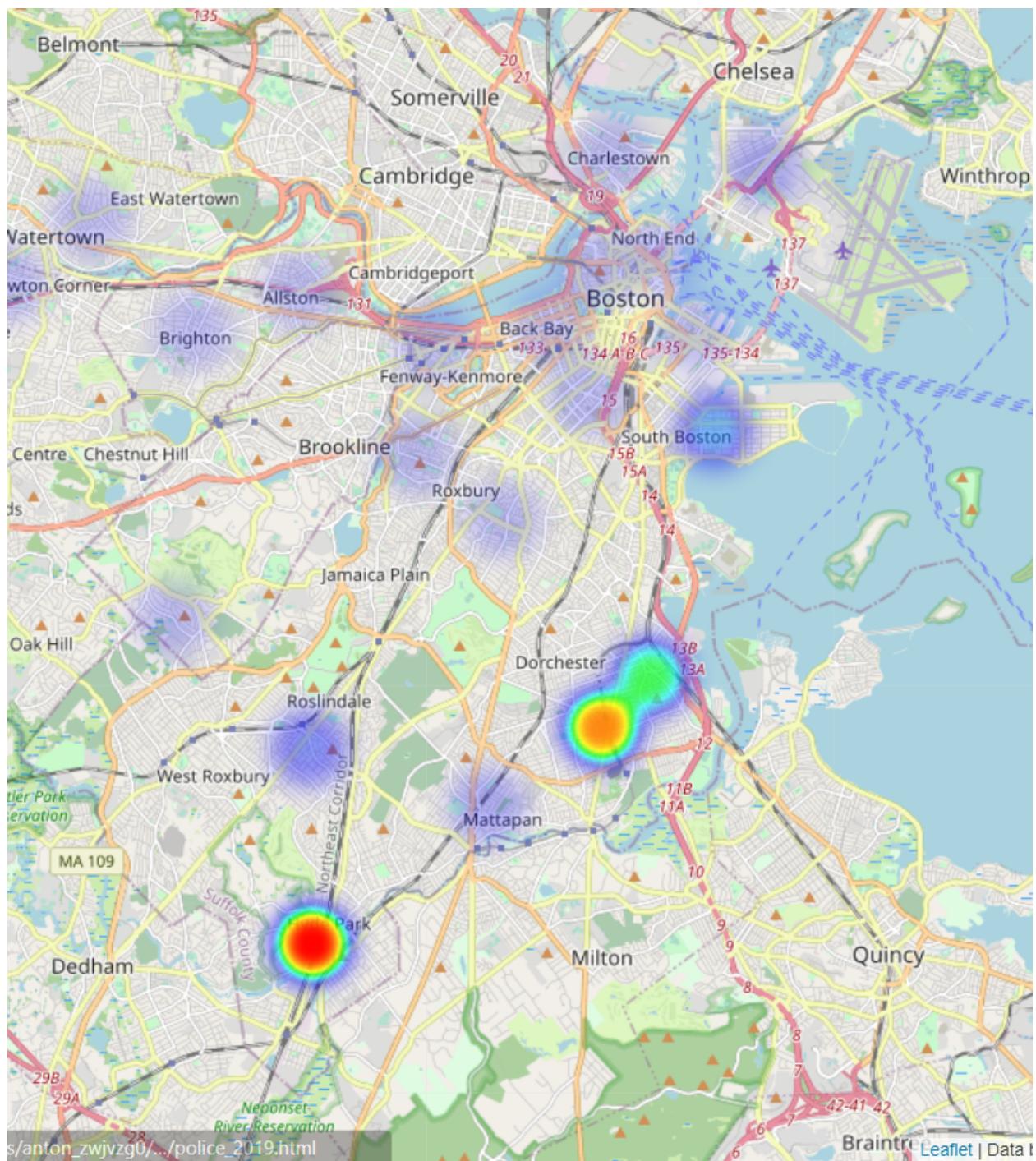
2021:



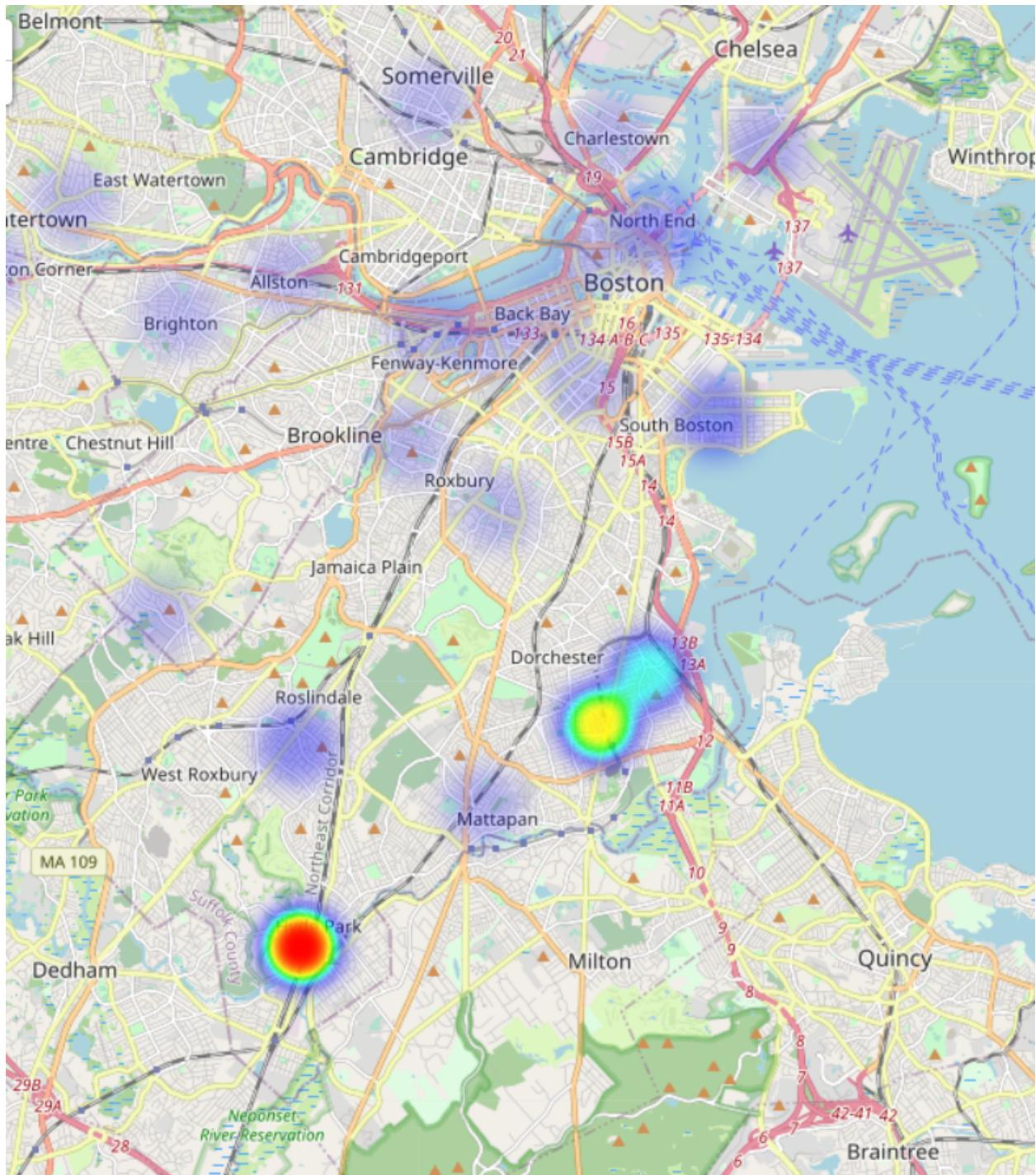
2020:



2019:



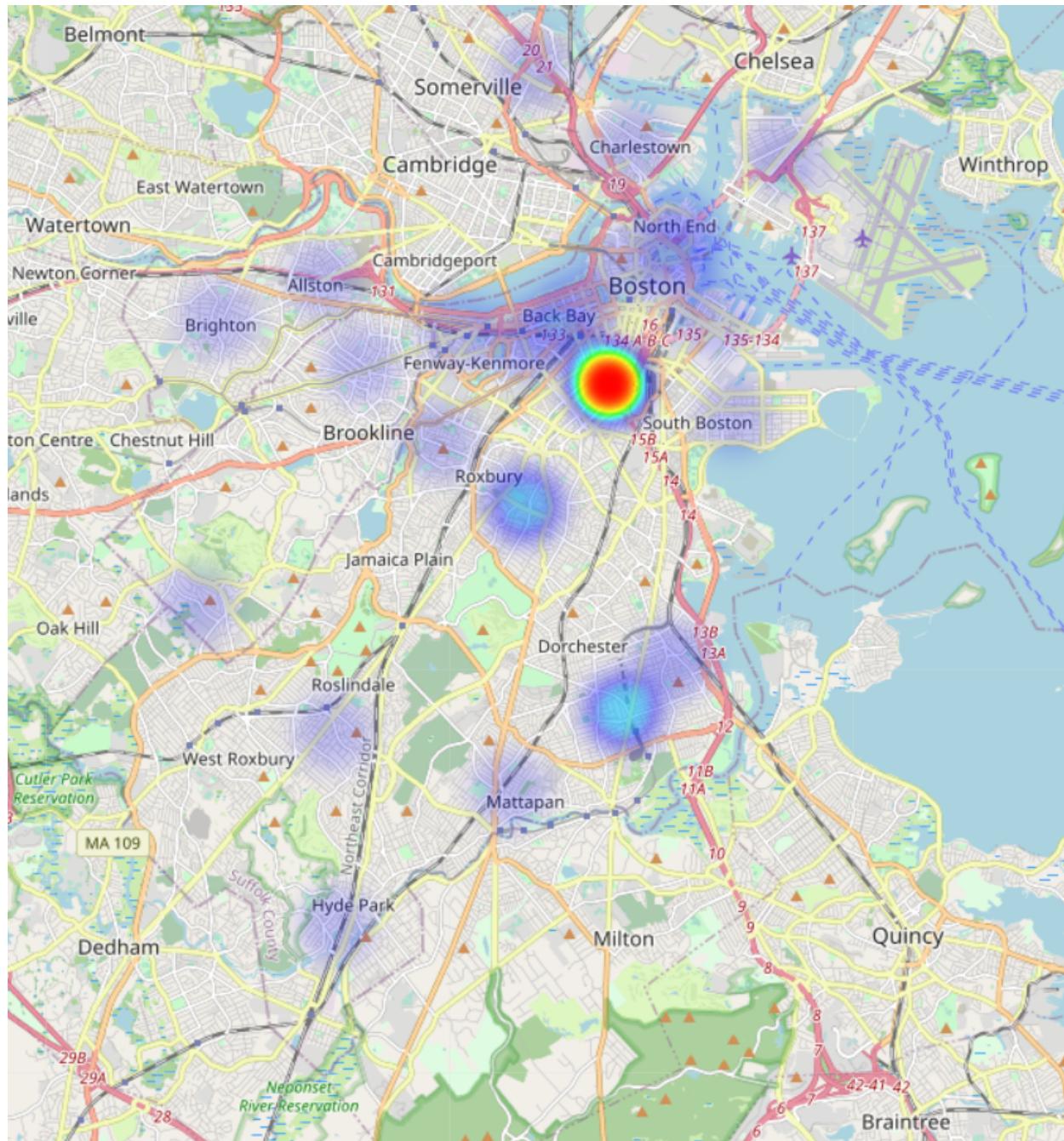
2018:

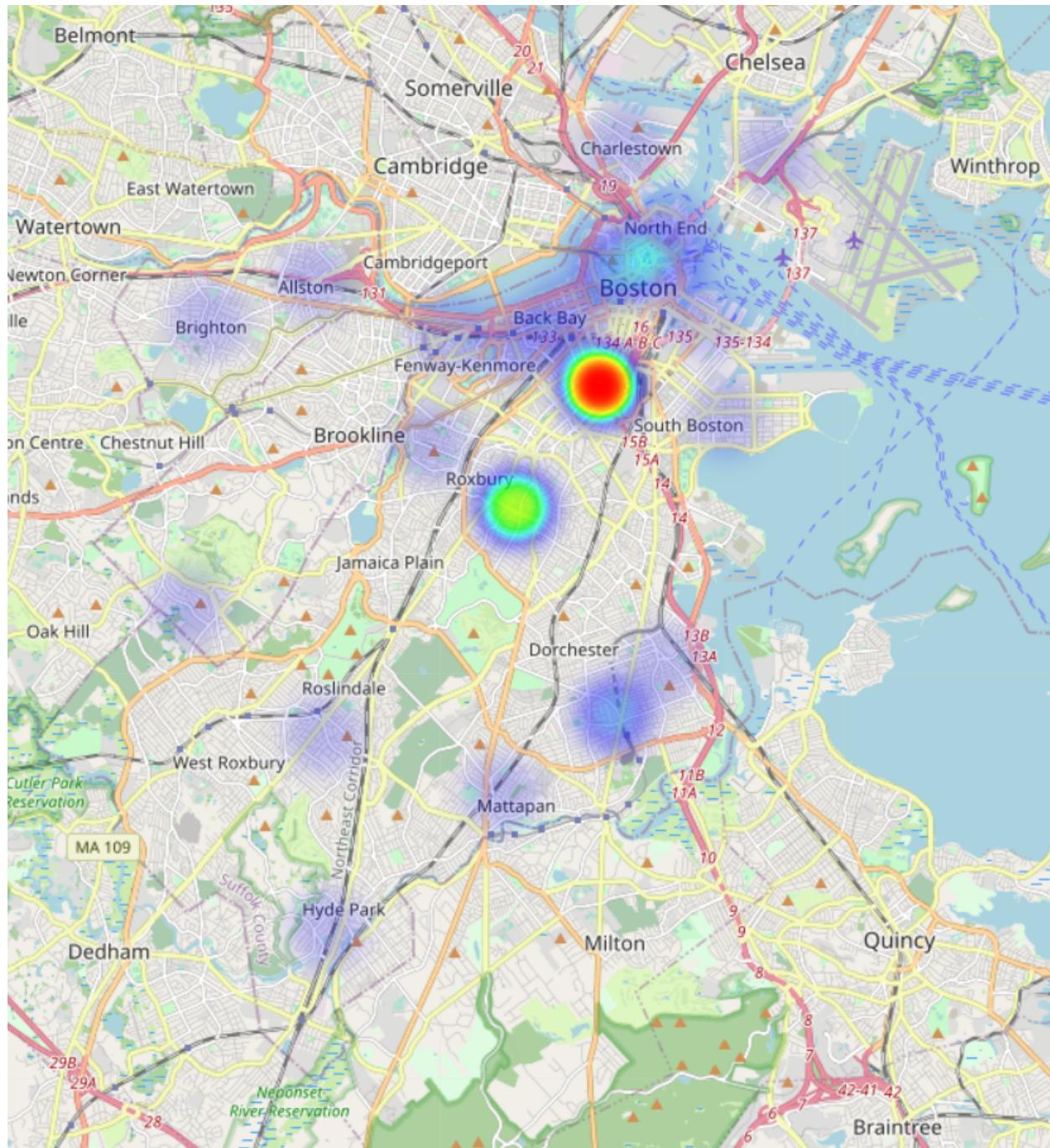


We can see that the police officer dispersion throughout the years doesn't change much, though most police officers live near Hyde Park due to the very red dot on that area. Also, several officers live near Dorchester. Although Dorchester makes some sense as an area where many police officers live, it's interesting that Hyde Park is so populated even though it's not an area which has an overabundance of crime.

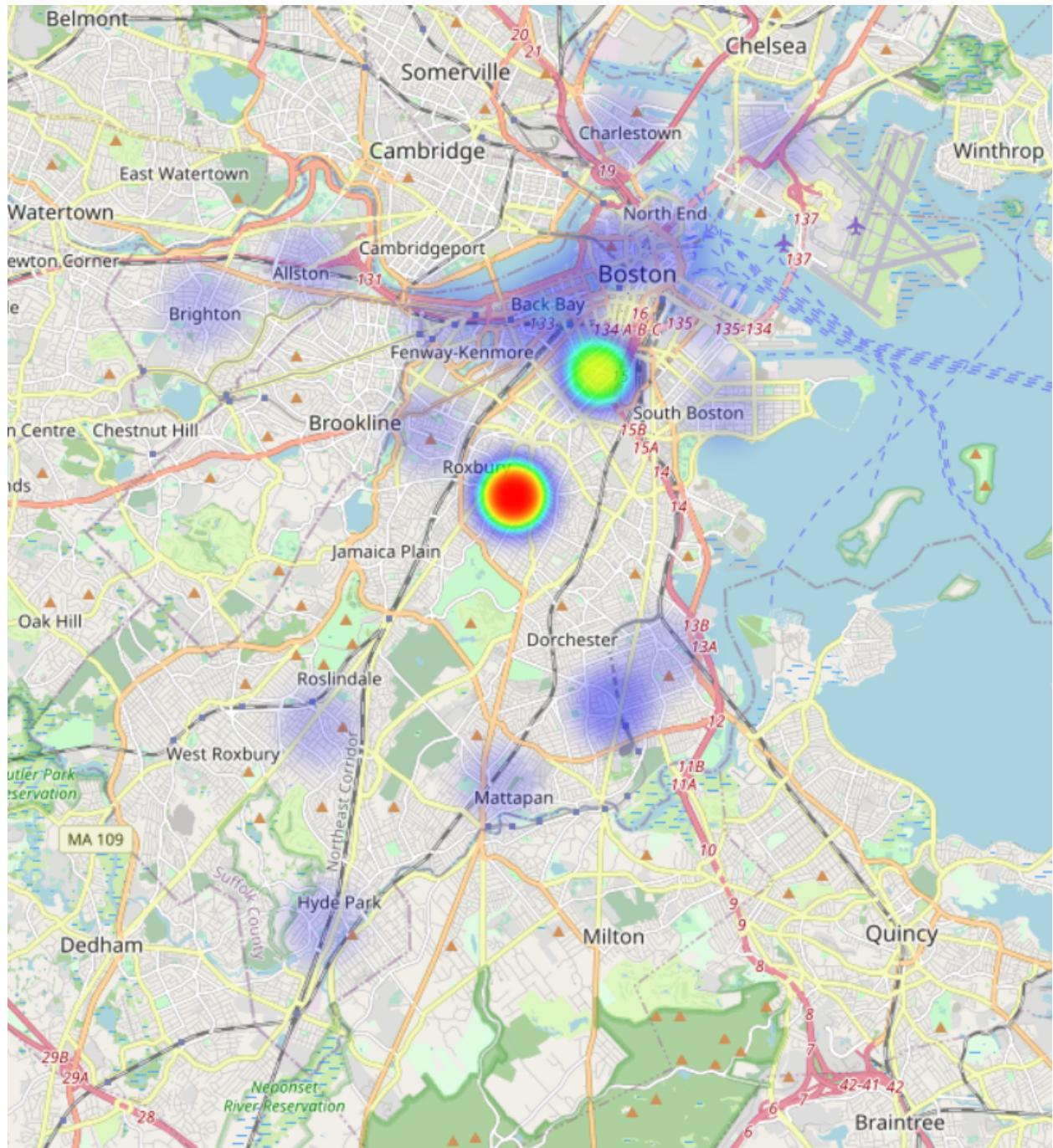
3. FIO Report Dispersion Heat Maps over Time:

2021:

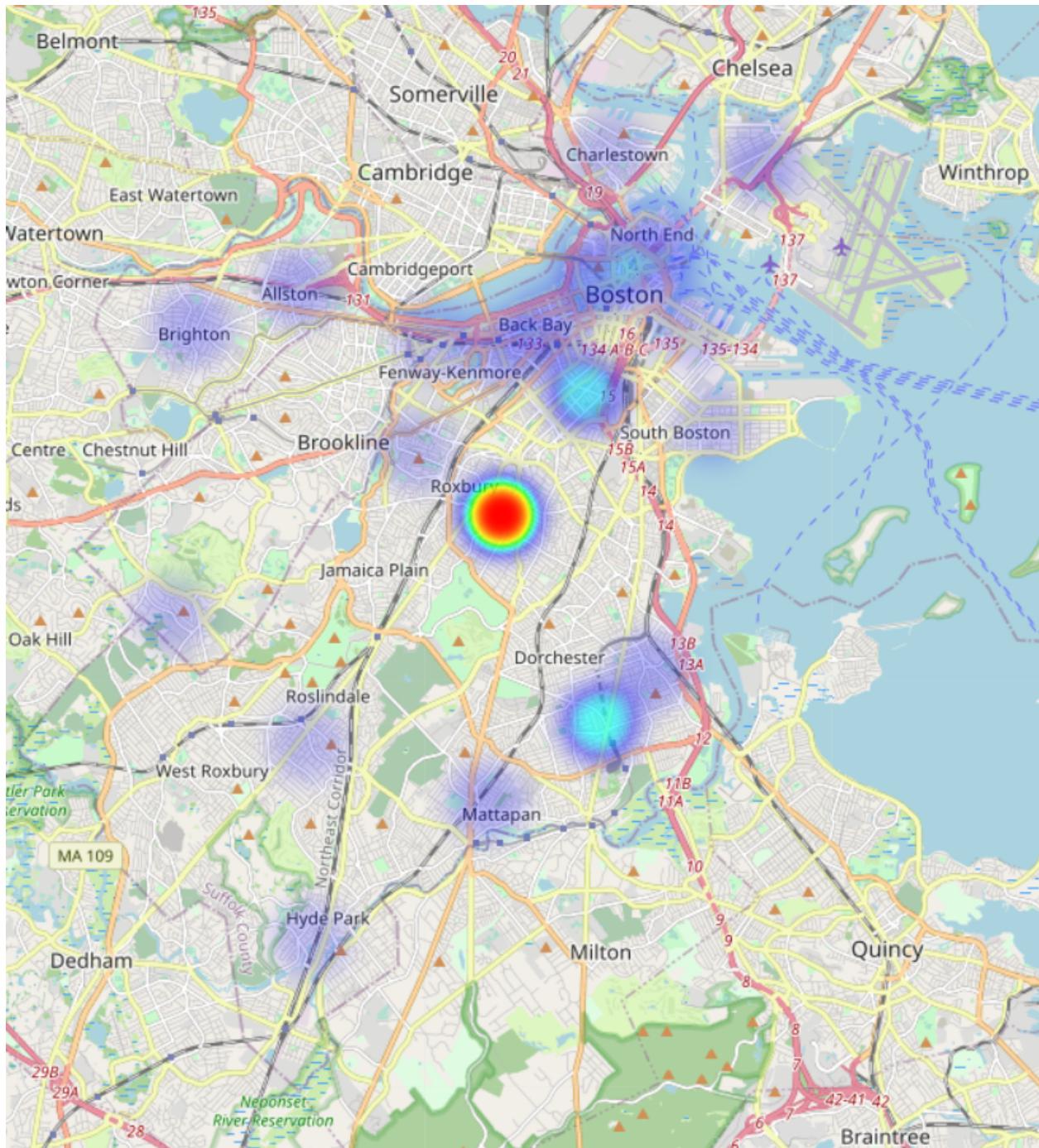




2019:



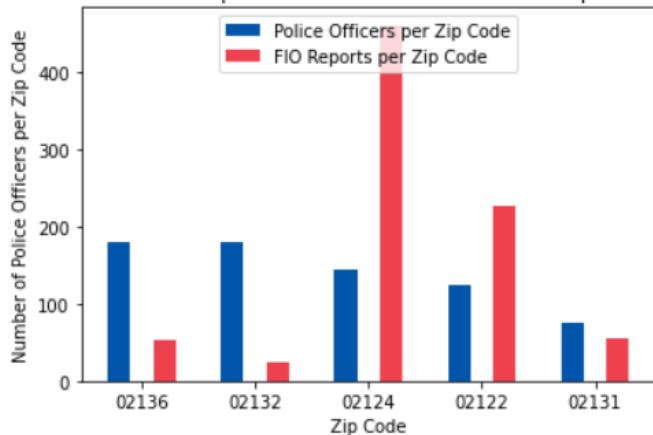
2018:



The FIO reports are dispersed throughout the city for all the years, but for all 4 Roxbury and South End (near South Boston) appear to be the most active FIO areas. Interestingly 2018 and 2019 Roxbury has more FIO reports, while in 2020 and 2021 South End now overtakes it due to the higher redness of the heat circle on that area.

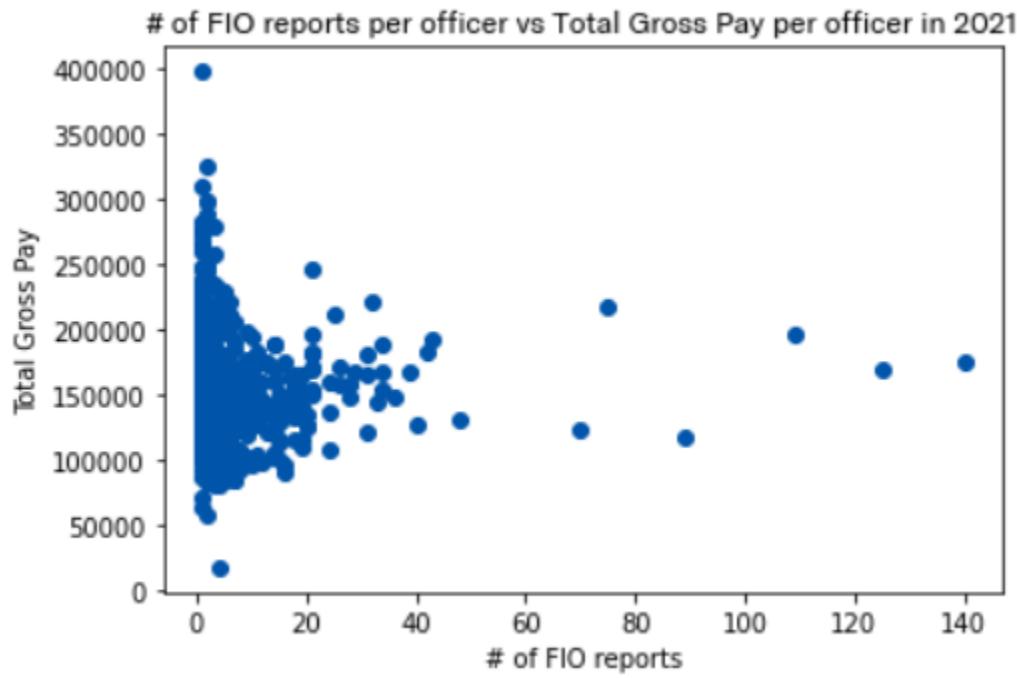
4. Top 5 Zip Codes by number of Police officers vs number of FIO Reports in those zip codes

Top 5 Zip Codes of number of police officers vs number of FIO Reports in those zip codes



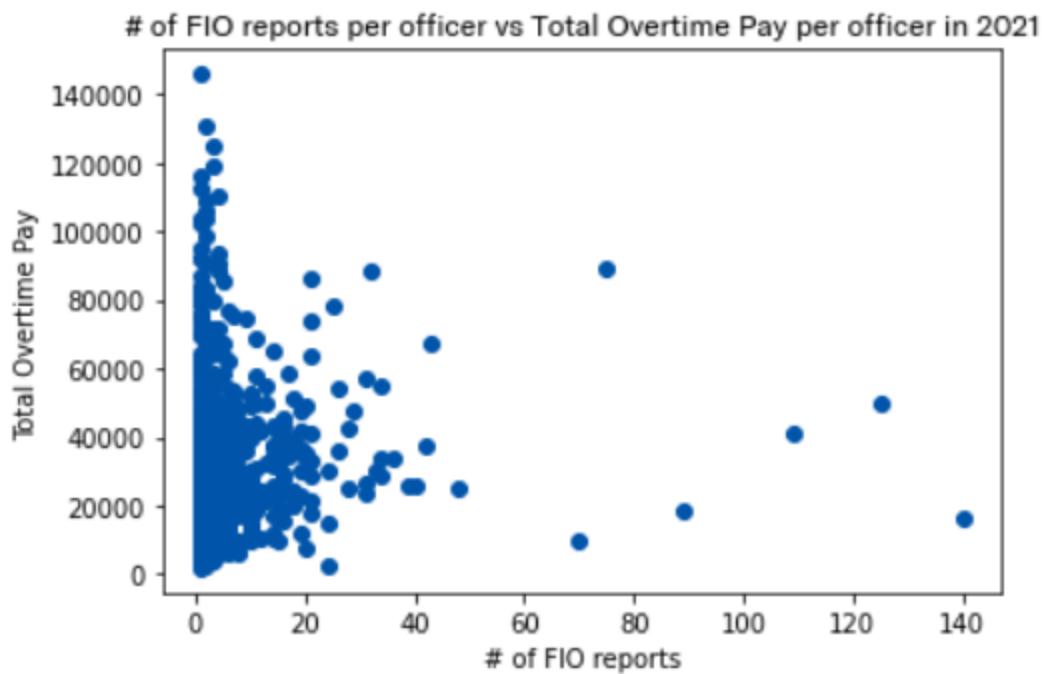
We took a look at the comparison between the top 5 zip codes by number of officers in them compared to how many FIO reports were generated for that zip code. This gave a weird looking bar graph that doesn't show much of a relationship, which is a bit strange since it's expected that the most populated zip codes by officers would have the most FIO reports.

5. # of FIO reports per officer vs Total Gross Pay per officer in 2021:



We can see that there's no clear trend between # of FIO reports and total gross pay per officer since all the data points are roughly in the same area, other than lots of outliers in the 0 to 20 range of # of FIO reports.

6. # of FIO reports per officer vs Total Overtime Pay per officer in 2021:



We can see that there's no clear trend between # of FIO reports and total overtime pay, and some individuals who have a lot of FIO reports assigned to them actually get paid less overtime than many who have very few.

7. Average salary per zip code of FIO report:



We can see that the gross salaries of the top 10 most reported FIO zip codes are not very stagnant, but most appear to be quite high above 100,000.

8. Average overtime per zip code of FIO report

Project Scope and Potential Risks:

Scope Changes:

Our scope of the project has been increased to include the FIO (Field Interrogation and Observation) dataset as recommended in our client meeting by the clients. This dataset gives us info on FIO reports, which will help to further analyze police compensation and number of police officers in different areas based on number and importance of FIO's. We'll also be able to compare the FIO's to the actual crime incidents to see how much the two are correlated, which we expect to be correlated a lot but this may not be the case.

As stated in our previous deliverable, there is no data about police officers other than ZIP code. Due to this problem, it will be difficult to represent their data in the form of a heatmap. Ideally we will convert ZIP codes to latitudes and longitudes to make these heatmaps but we're still in the process of creating this conversion. We will be constantly changing the scope of this project as we go along and see how we can best represent the data regarding police officers/employee compensation per area.

Another change in the scope of the project is that although we found the geopy module which can convert longitude and latitude values into zip codes, which would be useful for us to be able to compare the zip codes from the crime incidents to police officer zip codes, its runtime is far too long on the size of the dataset that the crime incidents has (about 63000 records in most datasets, each 90 records take around a minute to run, so would be around 700 minutes for just 1 of the years), so analysis on comparisons of police officer statistics to crime incident statistics hasn't been able to be run. This means that we will have to change our scope to likely either shrink the size of the dataset we're looking at, or find a different, more efficient way to convert latitude and longitudes to zip codes.

Potential Risks:

A potential risk of the project is the fact that in order to transform the crime incident data into zip code, we would have to remove some rows in the dataset because they were missing longitude and latitude values. These rows being missing from the analysis where we use the zip codes of the crimes may be a detriment to the accuracy of the analysis, so that is something important to consider when relaying our findings.

Another potential risk is that data for FIO reports may be difficult to compare effectively to some Employee compensation data. This is because we noticed that a decent amount of zip codes don't have FIO reports in

them, and so the amount of relationship graphing and data extraction we can do with FIO report numbers may be limited simply due to the limited number of FIO reports that were made.