CS 506 | Police Overtime | Team E

DELIVERABLE 2 REPORT

1. Problem Statement

Recent analyses of the Boston Police Department's (BPD) overtime expenditures have highlighted potential inefficiencies and misconduct, particularly in the context of significant social events like the George Floyd murder and Black Lives Matter protests. This deliverable aims to delve deeper into these issues, examining the discrepancies in overtime spending and their broader implications.

2. Data Collection and Cleaning Steps

a. Data Collection:

- Data sources provided by the client:
  - Employee earnings data (search police)
  - Campaign contribution data
  - BPD field activity data
  - Overtime data from 2012-2022
- Data sources from self-research:
  - Internal Affairs Officer Data
  - Suffolk Brady List Data (2020)
  - BPD Personnel Data
  - Crime Incident Reports

b. Data Cleaning:

- Data Type Conversion:
  - Built a function to convert inconsistent data types (i.e: non-numeric characters in the "overtime pay" field) to a uniform numerical format.
- Handling Missing Values:
  - Used mean imputation for related data and row exclusion for sporadically distributed and scarce missing values in datasets.
- Timeframe discrepancies:
  - Addressed inconsistent time frames in datasets by applying mean imputation and row exclusion based on missing data distribution and relevance.
- Standardization and Normalization:
  - Normalized column names for clarity and consistency within data categories. Extended standardization efforts to maintain uniformity when integrating diverse data types.
- Dataset combination:
  - Organized yearly datasets by creating lists for each data category, facilitating easy retrieval with a single variable and index corresponding to the year.
- Integration of New Datasets:
  - Overcame challenges in integrating new datasets by meticulously aligning variations in structure and format, preserving the integrity of the combined data.

3. Exploratory Data Analysis (EDA)

3.1 Data Understanding

Datasets	Description
Employee Earnings	Provides names, departments, and earnings (regular, overtime, bonuses, etc.) of City of Boston employees in the years of 2011-2022
Campaign Contribution	Includes contributors to political campaigns (OCPF), in the years of 2011-2022, with names, occupations, employers, contribution amounts, recipients, and other personal information
BPD Field Activity	Includes interactions between BPD officers and private individuals, in the years 2011 - 2022, with names of individuals and officers, dates, times, locations, and reasons for interactions.
Overtime	Includes overtime hours and payments for BPD employees, over the years of 2012-2022, with names, IDs, ranks, work locations, clients, hours worked, and pay rates
Internal Affairs	Provides internal investigations involving BPD officers, with names, allegations, outcomes, and disciplinary actions
Suffolk Brady List Data (2020)	Includes BPD officers flagged for credibility or misconduct issues in the year of 2020, with names, reasons for listing, and other personal details
BPD Personnel	Provides comprehensive information on BPD personnel, including roles, ranks, tenure, demographics, and service years. Helps to analyze the workforce composition and departmental policies
Crime Incident Reports	Provides the initial details surrounding an incident to which BPD officers respond, including type of incident, when and where it occurred, location.

3.2 Data Exploration

4. Possible Answers To Clients' Questions And Hypotheses

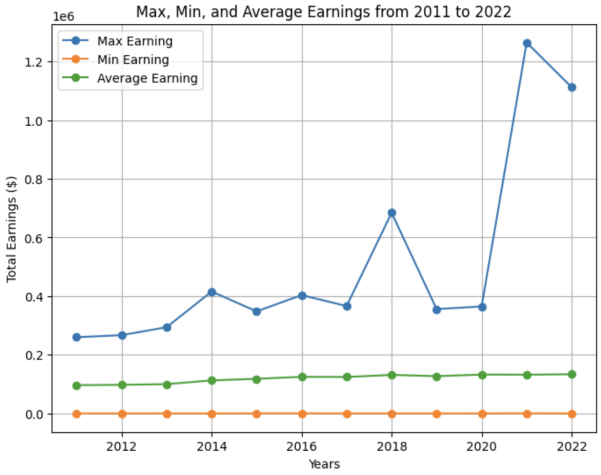
<u>Base Question 1:</u> How have BPD and paychecks changed year-over-year? Compare the average to non-BPD Boston city employees?

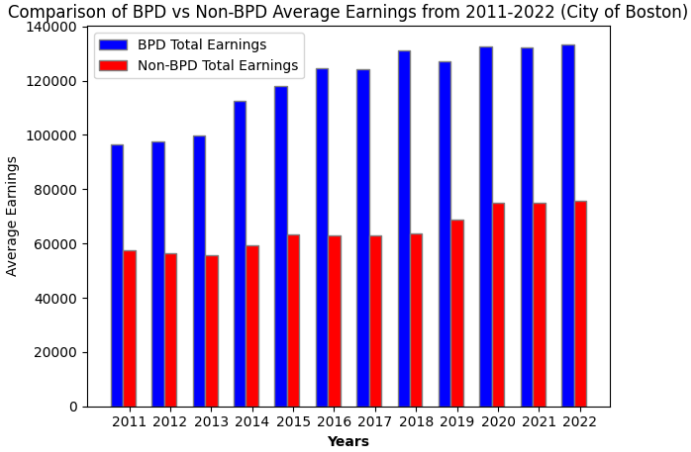
Assumptions:

- Average of "Total Earning" can be used as a measurement for paychecks.
- Non-BPD population includes all jobs in the City of Boston (i.e: cashiers, teachers, etc.).

Observations:

- Earnings for Boston Police Department (BPD) employees have been going up over time.
- Average salary for non-BPD city workers is approximately half that of BPD employees.
- Notice an abnormal increase in police earnings between 2020-2022, we looked into it and found out that the officer was actually awarded $2 million in a gender discrimination lawsuit by the Federal Jury.

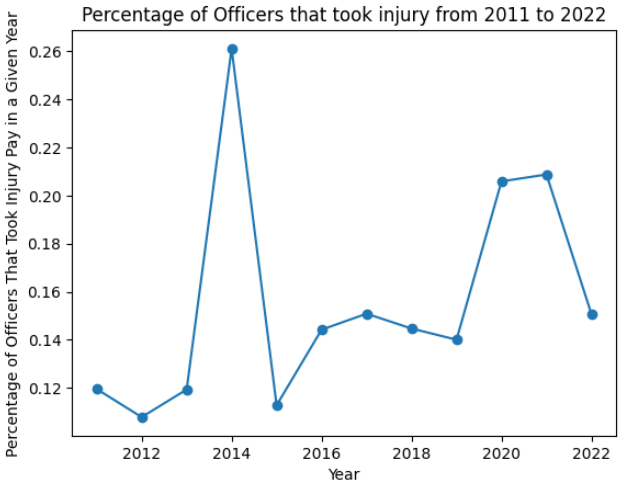


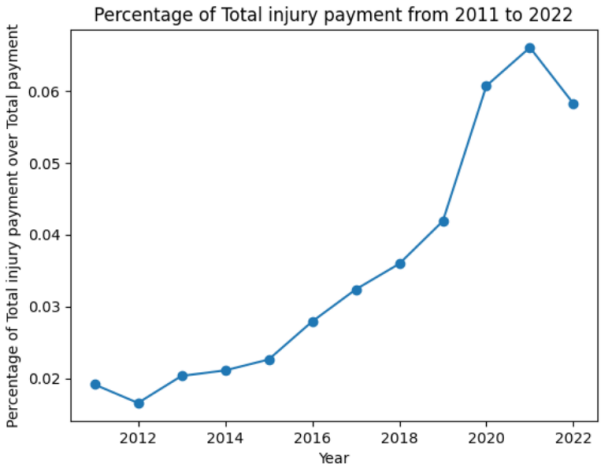


<u>Base Question 2:</u> How much BPD officer pay came from injury pay? What percentage of officers took injury pay in a given year?

Observations:

- The injury payment of BPD officers ranges from 5% to 25% of their overall payment and it has an increase overtime.
- About 2% to 6% of BPD officers got injured. This also has an increase overtime. Which corresponds to the injury payment increase.

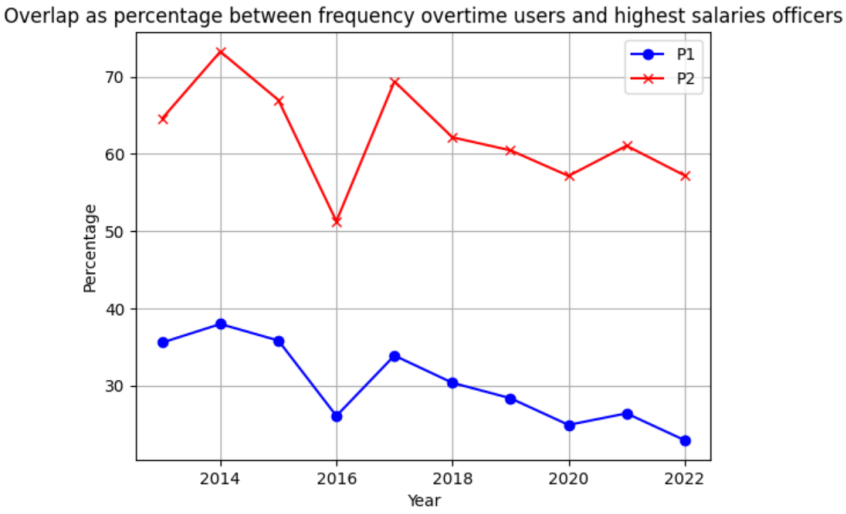




<u>Base Question 3:</u> How much overlap is there between frequency overtime users and officers who have the highest salaries on the force?

Assumptions:

- Most frequent overtime users set = the top 20% officers who have highest overtime taking hours (using HOURS_PAID)
- Highest earning officer set = the top 20% highest earning officers
- P1 = Percentage of officers that are in top 20% of overtime user given that they are in top 20% of highest income
- P2 = Percentage of officers that are in top 20% of highest income given that they are in top 20% of overtime user



Observations:

- An officer who had high income was very likely (>50%) to take overtime frequently.
- However, an officer who frequently took overtime didn't necessarily have a high income.

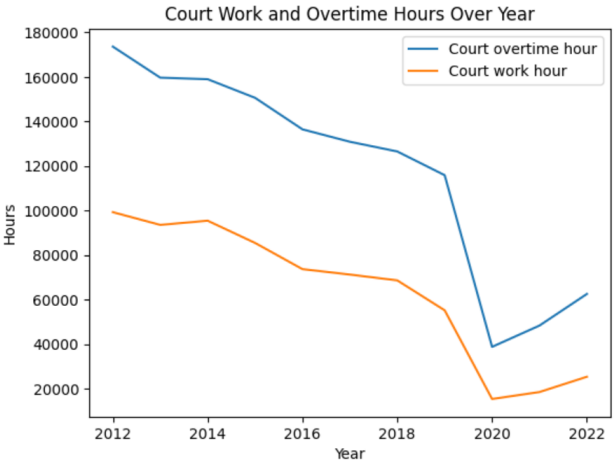
<u>Base Question 4:</u> How has overtime for court appearances changed year-over-year?

Assumptions:

- WRKHRS and OTHRS are used as a measurement for “appearances” in court

Observations:

- Overall, the total overtime hours consistently appeared to be twice the total work hours. This can be resulted from the 4-hr minimum court appearance policy.
- The year 2012 recorded the highest reported court overtime and worked hours.
- Conversely, 2020 witnessed the lowest reported court overtime and worked hours, potentially influenced by the COVID-19 pandemic and a surge in remote jobs.

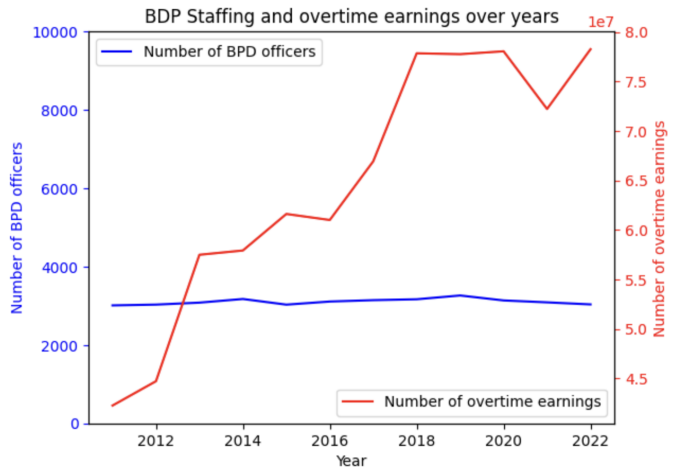


5. Extension Proposals

Extension Proposal: Boston Police Department Staffing Analysis	
Extension Pitch	*Investigate the correlation between staffing levels within the Boston Police Department (BPD) and the occurrence of overtime expenditures.* *This extension aims to understand if understaffing is a primary factor contributing to excessive overtime pay, providing insights into workforce management strategies.*
Rationale	*Examining the relationship between staffing levels and overtime spending is crucial for identifying the root causes of inefficiencies and potential misconduct highlighted in recent analyses.* *This extension will contribute to a holistic understanding of the dynamics affecting overtime expenditures during significant social events, helping the BPD optimize resource allocation and improve operational efficiency.*
Questions for Analysis	1. *How does the staffing level within the BPD correlate with the frequency and magnitude of overtime expenditures?* 2. *Are there significant differences in overtime spending patterns between BPD and comparable police departments?* 3. *Does the analysis reveal patterns in overtime spending during critical events, such as social protests or major incidents, and how are these patterns linked to staffing levels?* 4. *Analyzing the relationship between the number of police officers and the number of crime incident reports over the years*
Data Sets & Sources	1. *Boston Crime Incident Report (Aug 2015 - 2022)*
Data Visualizations	1. *Line graph: Illustrate changes in the number of officers and the crime incidents year-over-year* 2. *Scatter Plot: Illustrate the correlation between staffing levels and overtime expenditures over the analysis period* 3. *Departmental Comparison Chart: Compare overtime spending patterns across different departments or units within the BPD to identify specific areas of concern.* 4. *Event Timeline: Map out overtime expenditures concerning significant social events, providing a visual representation of staffing challenges during critical periods.*
Additional Information	*Factors such as population density, crime rates, and economic conditions can significantly impact budgeting and overtime practices. Additionally, understanding the context surrounding budgeting and overtime practices in police departments is crucial for a comprehensive analysis. Thus, we may need to double check our conclusion with TA and the Spark PM to make sure our conclusions are biased and out of context.*

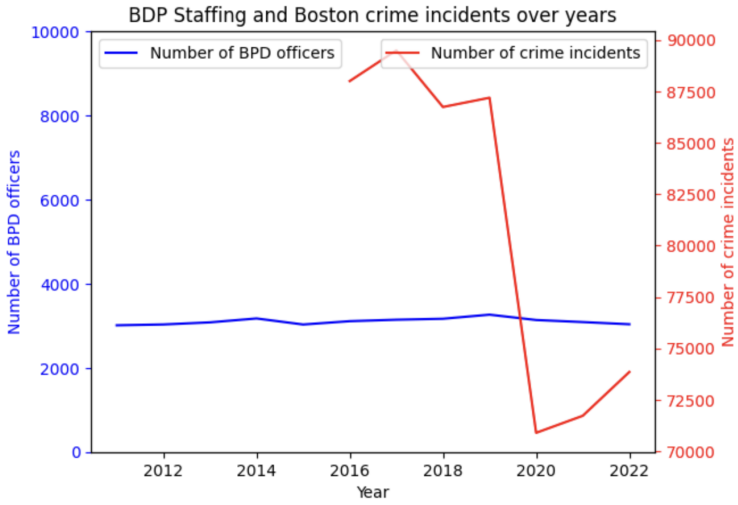
6. Visualization and insights for extension proposals

1. ***How does the staffing level within the BPD correlate with the frequency and magnitude of overtime expenditures?***



By the graph we can see that the number of BPD staffing stays around the same value year to year. But the number of overtime earnings grew significantly.
This indicates that each officer had more earnings on overtime. This can signal that BPD was experiencing short-staffing.

2. ***Analyzing the relationship between the number of police officers and the number of crime incident reports over the years.***

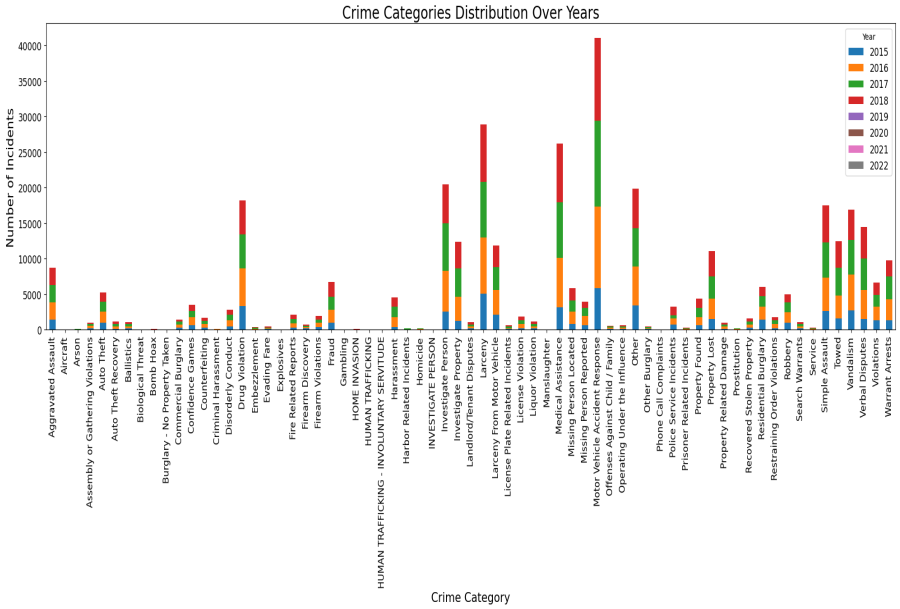


We can see, as the number of BPD staffing stays around the same value year to year, the number of crime incidents decreased significantly in 2020. But the decrease in crime incidents does not lead to lower overtime expenditure.
Observing the year 2019 to 2020, the number of crime incidents decreased significantly but overtime earnings still remain within the same range.

Thus, there is no obvious relation between understaffing and overtime payments, however having similar overtime payments. We will continue analysis on the factors that affect overtime earnings.

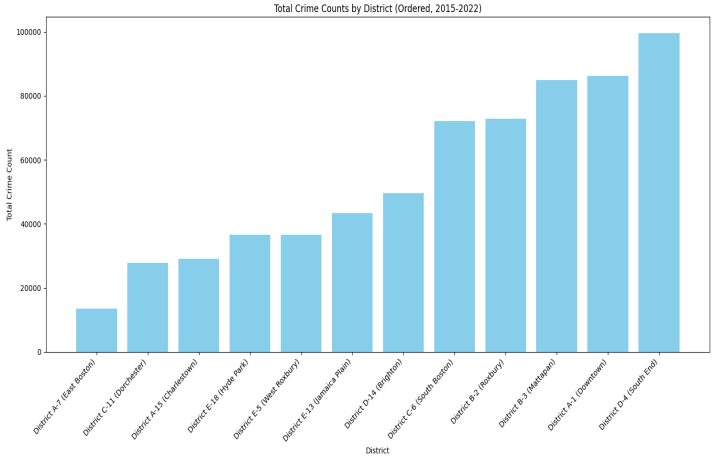
It should also be noted that our datasets do not accurately display the amount of BPD officers in each district. This is partly due to a nonuniform listing of districts in certain datasets (ie, an officer is listed as being based in District 11 when BPD Districts are assigned a letter followed by a number such as A1 or D14. However, it is very likely that districts with more incidents reported have more officers allocated.

3. *Crime Incident Reports by Category from 2015-2022*



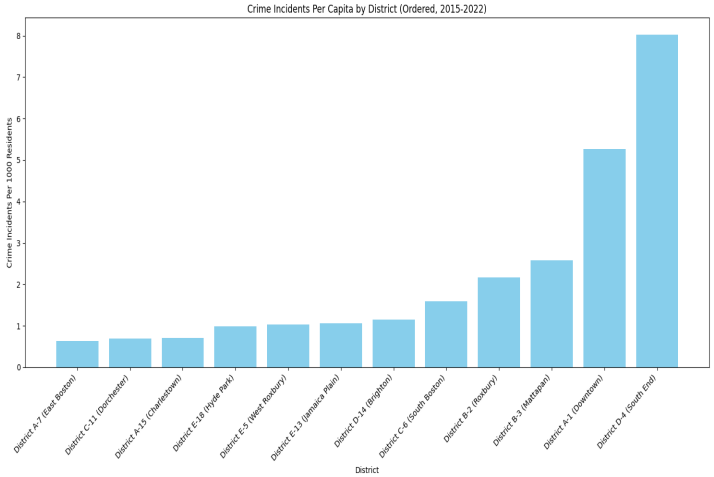
For crime incident reports by categories, it is important to note that most of the reports are for emergency dispatches such as car accidents or requests for medical aid. Incidents that are drug related, involve larceny/theft, or those having to do with investigations are the next most common reports. Although the period from 2016-2018 has the most reports, it isn't quite clear why. A possible explanation indicates a gradual rise in reporting that was curtailed by the global pandemic starting in 2019.

4. *Total Crime Incident Reports by District from 2015-2022*



For the total crime incident count for each district, it should be noted that the South End and Downtown districts have the highest total crime reports despite being the two districts with the lowest population. On the other hand, Dorchester and East Boston are the two most populated districts and have the lowest crime incident reports.

5. *Crime Incidents Per Capita by District from 2015-2022*



To better visualize this, here are the crime incident reports per capita (in this case, crime incident reports per 1000 people). Most districts have 2 or fewer crime incidents for every thousand residents. However, this isn't a measure of actual crime; just incident reports. Although the outliers of South End and Downtown suggest increased crime in the area, until further research is done, the only conclusion that can be made is that reports involving BPD are more frequent in these areas.

7. Individual contributions of each team member

[Nurassyl Medeu]

- Assisted with data processing and cleaning
- Reorganize the old workspace, condensing and moving the working code to a new code space
- Performed analysis on BPD and non-BPD paychecks 2011-2022, determining the early change of average paychecks of BPD employees and non-BPD employees
- Performed Statistical Analysis of Campaign Contribution data 2011-2022, calculating and plot the average amount contributed by BPD employees and non-BPD employees
- Initiated presentation slides and performed recording edits.

[Jiawei Sun]

- Assisted with data processing and cleaning.
- Conducted analysis on Total Earning Data from 2011-2022:
- Analyzed trends of total earning and overtime earning over years
- Calculated the percentage breakdown of individual earnings.
- Examined the percentage increment of individual earnings over time.
- Performed analysis on amount of overtime hours paid and overtime hours worked.
- Analyzed the overtime ratio from 2013 to 2022 and examined the outliers. (eg. 35000 overtime earnings per hour)

[Truc Duong]

- Assisted with data preprocessing and cleaning, performing research for new datasets (crime incidents, officers, court, etc)
- Performed Statistical Analysis on BDP Earning Data, Overtime and Regular Earning changes from 2011-2022 (determined average total earning of a police officer, analyzed changes in average total earnings, examined detailed statistics for police earnings)
- Analyzed differences in overtime type in the overtime dataset
- Initiated project extension, created line visualizations and produced preliminary observations

[Can Wang]

- Analyze data for injury payment and injury payment ratio 2011-2022.
- Analyze data for injury ratio of BPD officers per year 2011-2022
- Did research related to injury ratio and injury payment.
- Assesses if BPD officers have had better subsidies since 2011.
- Research on the ranks of officers and their payment.
- Adding analysis related to frequency of overtime and highest payment.

[Al Mbaye]

- Performed Statistical Analysis on Total Spending by Category from 2011-2022, determining the amount of total spending of each category, the average amount of spending of each category
- Performed Statistical Analysis on Total Earning by Category from 2011-2022 (determined the amount of total earnings of each category, the average amount of earnings of each category, a ranking of total earnings and a ranking of total earnings adjusted for inflation)
- Helped with data preprocessing for court overtime
- Imported and pre processed data for internal affairs officer complaints
- Found overlap between frequent overtime users and officers with internal affairs complaints
- Implemented analysis on the project extension