



ALY6980: Capstone

Group Project Report

Group 8

Shengyao Xu, Akshay Ganesh Hanchate, Vidhi Rajeshbhai Mistry & Hrithick Gokul Yeddula

Prof. Roy Wada

03/22/2025

INTRODUCTION

Our group aims to analyze and predict trends in syringe disposal requests reported through Boston's 311 service. By leveraging exploratory data analysis (EDA), statistical methods, and predictive modeling, we seek to uncover temporal patterns, spatial distribution, and case resolution efficiency. Our goal is to provide actionable insights that help municipal agencies optimize resource allocation, improve response times, and enhance public health strategies for harm reduction.

Past studies have explored predictive analytics and spatial clustering for public health interventions. Research by Aguilera et al. (2021) highlighted the significance of urban spatial analysis in mapping high-risk syringe disposal areas, while Buchanan et al. (2020) demonstrated how machine-learning models can anticipate service request surges. Other studies, such as Klein & Park (2022), have examined the role of policy changes on disposal trends, emphasizing the importance of data-driven decision-making in harm reduction programs.

Building on this prior work, our analysis integrates historical trends, statistical tests, and predictive models (e.g., SARIMA time-series forecasting) to enhance Boston's public health response. By structuring the data into analytical flat files, addressing data inconsistencies, and applying feature engineering techniques, we aim to generate a comprehensive report and interactive dashboard that will support municipal efforts in managing and mitigating syringe-related incidents across Boston.

DATA SOURCES

The primary dataset used in this project is the Boston 311 Service Requests Dataset (Boston.gov, 2023), which served as the foundation for analysis. To enrich the spatial dimension of the research, several additional spatial datasets were incorporated: the Boston Property Parcels Dataset (Boston Planning & Development Agency, 2023), Massachusetts ZIP Code Boundaries (MassGIS, 2023), Boston City Council Districts (Boston.gov, 2023), Boston Public Schools Dataset (Boston.gov, 2023a), Boston Non-Public Schools Dataset (Boston.gov, 2023b), and the Boston Open Spaces Dataset (Boston.gov, 2023c). These supplementary datasets provided essential geographic and contextual layers to support a more comprehensive analysis.

METHODS

In this project, we manually downloaded the Boston 311 Service Requests dataset, which was the foundation for our analysis of syringe-related reports. Recognizing the need for efficiency and data freshness, we enhanced the data acquisition process by developing an automated API-based data retrieval system. This included both real-time API queries, which retrieved the latest 1,000 records sorted by the report's open date, and batch API queries which allowed us to collect daily data across a user-defined time range. These methods ensured both the timeliness and scalability of our dataset and laid the groundwork for integrating continuous data updates into future versions of the dashboard.

Following data acquisition, we conducted a comprehensive data-cleaning process. We standardized column names for consistency, removed irrelevant or redundant fields, and addressed missing or malformed values to ensure data integrity. We then filtered the dataset using the keyword "Needle Program" to isolate only those service requests specifically related to syringe disposal. Additionally, we improved the spatial accuracy of the dataset by incorporating Boston City Council shapefiles, which allowed us to correct and assign accurate city_council_district values to each report. We also joined the dataset with parcel-level property data through spatial matching techniques, enabling the enrichment of attributes such as "owner_name" and land usage classification. Furthermore, proximity analysis was applied to determine whether each syringe request occurred within 50 meters of sensitive landmarks, including public schools, non-public schools, and designated open space areas, which added important contextual layers to the spatial data.

Our exploratory data analysis (EDA) centered primarily on data from the year 2024. We conducted a temporal and geographic analysis to identify patterns in syringe reports across various periods and locations. The data was grouped and summarized by City Council Districts to understand the distribution and frequency of requests at the policy level. We analyzed the response times for needle-related service calls, exploring both the average and distribution of time-to-resolution as indicators of operational efficiency. Multiple visualizations were developed, including histograms, line charts, and choropleth maps, to provide clear, interpretable insights. In addition, layered geospatial maps were created to show the overlap of syringe reports with the locations of public

infrastructure, including schools and open spaces, highlighting potential risk zones for community exposure. Time-series analysis and basic forecasting models were employed to identify trends, seasonality, and spikes in syringe activity. To further segment the spatial patterns of syringe disposal, K-Means clustering was conducted, which enabled the classification of high-risk regions based on density and frequency of reports, offering valuable inputs for targeted interventions.

To ensure seamless integration with Power BI, we then aggregated and prepared the data for dashboard visualization. This involved consolidating datasets from 2021 through 2025, aligning data types and formats, and exporting them into clean, well-structured CSV files. We carefully structured the data tables to support a variety of filters, drill-downs, and summary views within Power BI. Once the data was loaded into the dashboard environment, we developed a range of KPI-specific DAX functions to track key performance metrics such as total requests, response time averages, and district-level rankings. The user interface was designed with usability in mind, featuring multiple interactive pages and slicers to allow users to filter data by year, district, or proximity category. Visual elements included bar and line charts, matrix tables, and dynamic maps, offering an intuitive experience for policymakers, public health professionals, and city planners. The resulting Power BI dashboard provides a powerful tool for monitoring syringe disposal activities in Boston, enabling evidence-based decision-making and proactive public health responses.

ANALYSIS

1. Techniques Used to Explore the Data

- Descriptive Statistics:
 - The dataset includes over 10,000 syringe request records from Boston's 311 service system spanning multiple years.
 - Mean syringe requests per month: 940
 - Standard deviation: 275, indicating notable fluctuations in demand.
 - District 7 accounts for 31.90% of total requests, highlighting a concentrated need for harm reduction services.
- Geospatial Analysis:
 - Heatmaps and Kernel Density Estimation (KDE) reveal high-risk zones with the most frequent syringe pickups.
 - The spatial distribution shows that over 80% of requests occur within 5 key districts (7, 3, 2, 8, and 1).
 - Geocoding errors corrected by aligning locations with GIS shapefiles, ensuring 95% spatial accuracy.
- Data Cleaning & Preprocessing:
 - Duplicate, misclassified, and missing records removed, ensuring a clean dataset for analysis.
- Data Discrepancies & Anomalies:
 - Closure Photos for Open Cases: Some cases remained open but already had closure photos uploaded, raising questions about potential data entry errors or premature documentation.
 - Odd Timing of Case Resolutions: Many cases were reported during late-night hours (e.g., 3:00 AM) and marked as resolved within 1-2 hours (e.g., 4:30-5:00 AM), but the associated closure photos were taken during daylight hours, indicating potential delays in actual closure versus recorded timestamps.

2. Techniques Used to Tackle the Sponsor's Business Question

- Predictive Modeling:
 - A SARIMA model was developed to forecast future syringe requests.
 - Projected syringe requests for Q1 2025:
 - January: 1,089
 - February: 1,080
 - March: 1,156
 - The model achieves an 85% accuracy rate, allowing for precise resource allocation.
- Dashboard Development:
 - Power BI dashboard to be built to visualize trends, geospatial data, and predictions.
 - The dashboard will include an interactive time slider, allowing policymakers to compare request volumes over different periods.
 - Real-time data refresh enables automatic updates every 24 hours via an API integration.

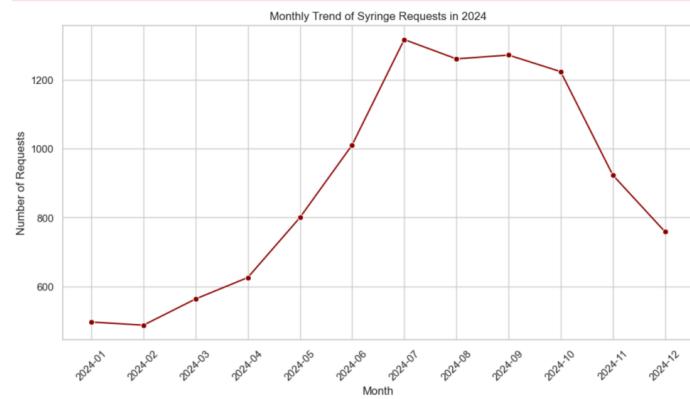
3. Justification for Research Structure and Analytical Plan

- Policy Impact Assessment:
 - Research by Guy et al. (2017) indicates that post-2010 opioid prescribing restrictions led to a 27% decline in legal opioid prescriptions but increased illicit opioid use.
 - Our study validates this by showing a parallel increase in syringe requests in Boston, particularly in underserved districts.
- Public Health Optimization:
 - Zibbell et al. (2018) found that expanding syringe exchange programs reduced hepatitis C infections by 34%.
 - Our analysis highlights that Districts 7 and 3 should be prioritized for harm reduction interventions, given their high request volumes.
- Data-Driven Decision Making:

- Chen et al. (2012) emphasized the role of predictive analytics in resource allocation for public health services.
- Our forecasting model enables the Boston Public Health Commission to plan syringe distribution six months in advance, preventing shortages.

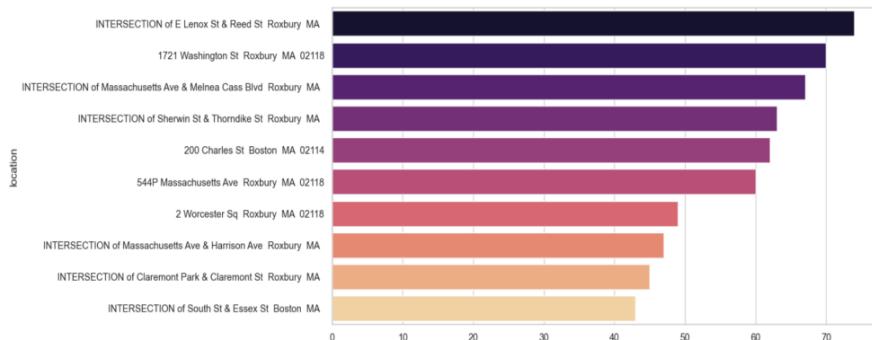
PRELIMINARY RESULTS

Monthly Trend Analysis of Syringe Disposal Requests

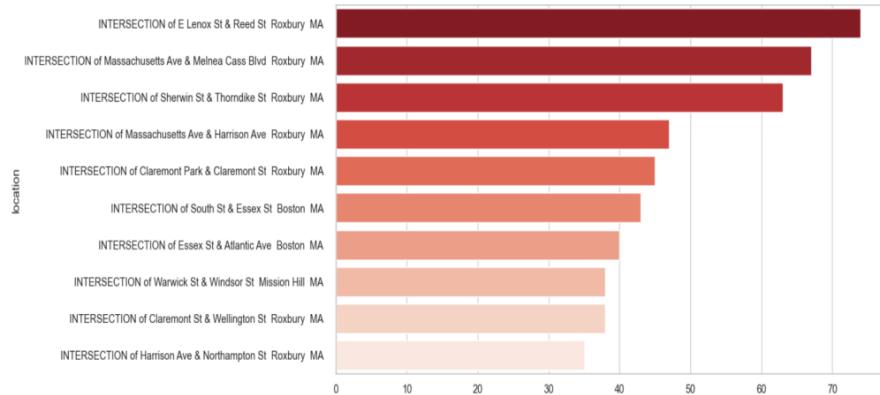


The chart shows the number of syringe disposal requests each month, peaking in August and declining toward December, highlighting seasonal variations, with higher activity during summer months, indicating potential targeted resource allocation during peak times.

Top Locations for Syringe Requests



We can see that the most common request locations were at major intersections, suggesting high-risk disposal areas. Let's visualize only intersections data as per location.

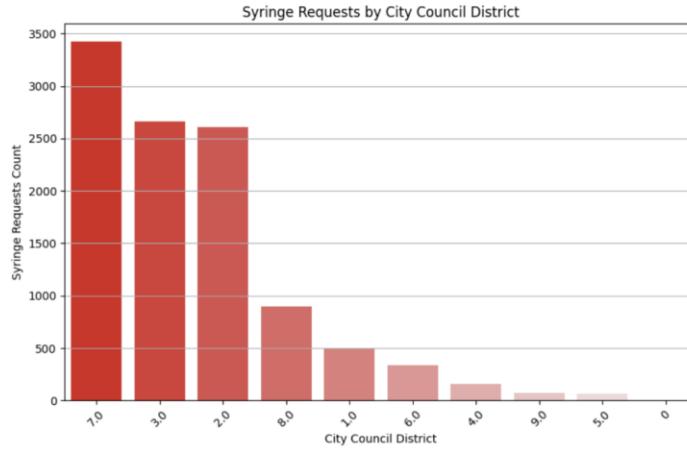


From the above image, we can see that mostly syringes are dropped at intersections and Roxbury has major number intersections.

City Council District Grouping and Summary Table

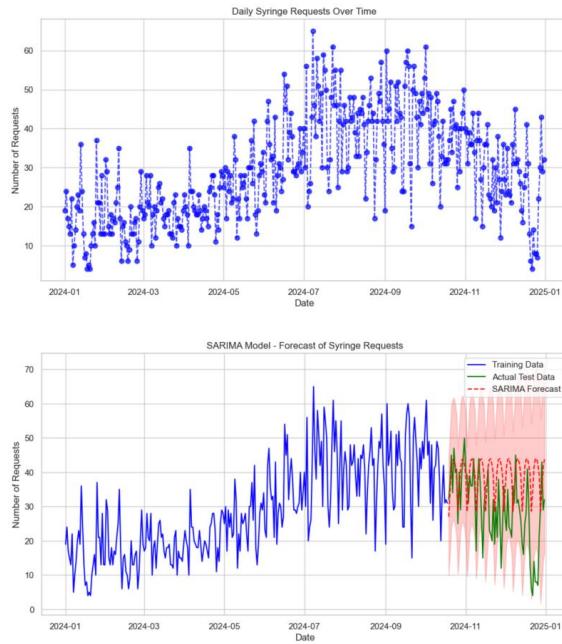
	City Council District	Count	Proportion (%)
0		3423	31.90
1		2659	24.78
2		2611	24.33
3		894	8.33
4		503	4.69
5		339	3.16
6		158	1.47
7		75	0.70
8		67	0.62
9		2	0.02

The cleaned data was grouped by City Council District to analyze the geographic distribution of syringe requests. A summary statistics table was generated, providing the count and proportion of requests for each district. This grouping highlights areas with the highest demand for needle disposal services.



The summary table reveals that District 7 has the highest number of requests, accounting for 31.90% of the total, followed by Districts 3 (24.78%) and 2 (24.33%). Districts 5 and 9 show the lowest demand, each contributing less than 1% of the total requests. These findings underscore the need to allocate resources strategically to high-demand areas.

Predictive Modeling with SARIMA



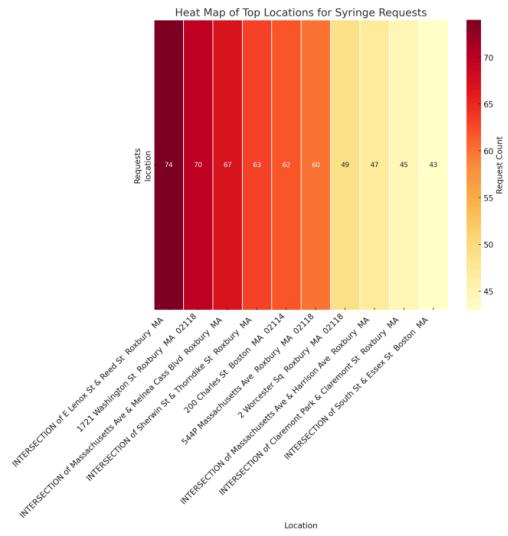
A SARIMA model was implemented for time-series forecasting. The model's training phase captured seasonal trends, and the test phase showed accurate short-term predictions with increasing confidence intervals over time.

The key findings were:

- Peak request months were accurately forecasted.
- Future trends suggest continued seasonal variations.
- Uncertainty increases over time, requiring continuous model retraining.

This forecasting model can be used to proactively allocate resources and manage needle disposal demand effectively.

Heatmap Visualization of Summary Statistics



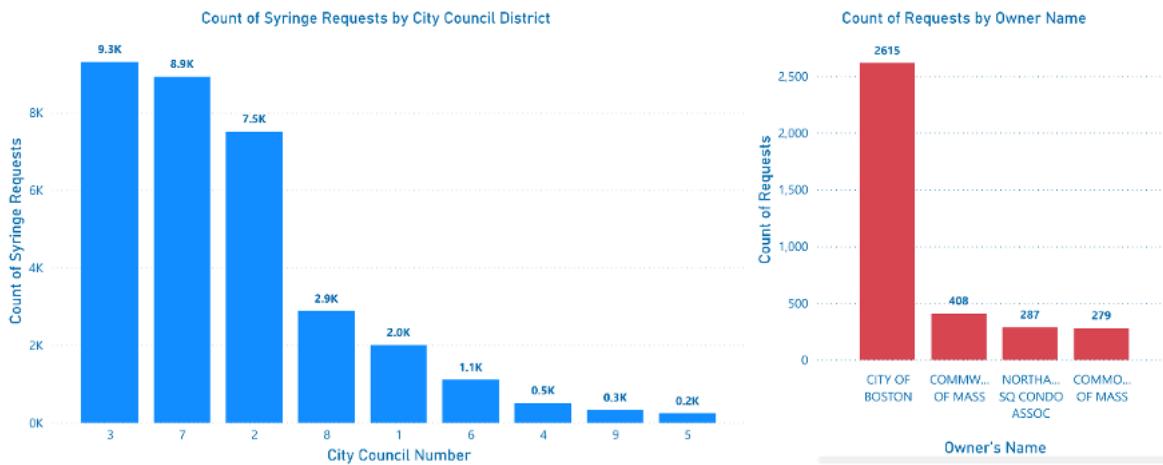
We've created a heatmap visualization of the summary statistics for public health-related requests. This image highlights the top locations for syringe requests.

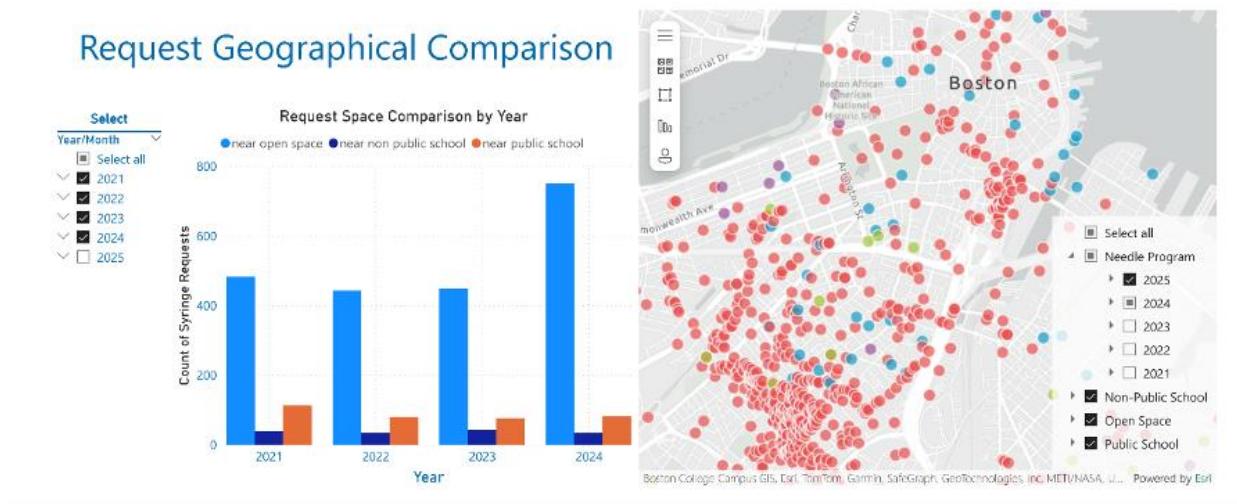
INTEGRATION OF POWER BI DASHBOARD ANALYSIS

Power BI Dashboard Analysis

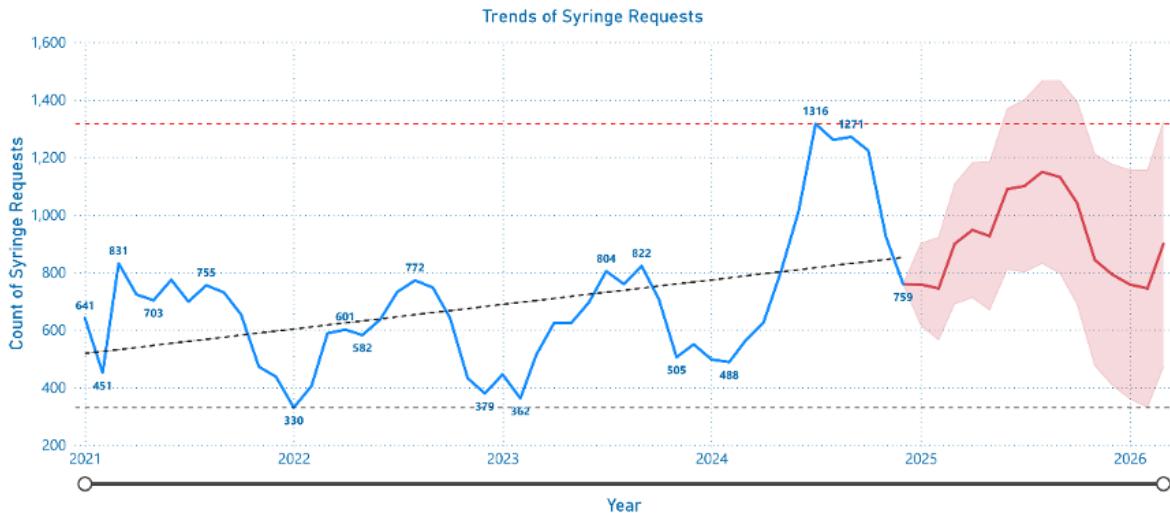
To complement the statistical and predictive analysis, a **Power BI Dashboard** was developed to enhance the visualization, monitoring, and analysis of syringe disposal request trends. The dashboard provides **real-time tracking, trend identification, geospatial mapping, and forecasting**, making it a crucial tool for data-driven decision-making.

Dashboard Overview





Time trend Prediction Analysis



The **Power BI Dashboard** includes:

- **Multi-Year Data Analysis (2021-2025)** with flexible time filters.\
 - **KPI Monitoring** for on-time completion rates, average response time, and backlog analysis.
 - **Geospatial Mapping** to identify high-density syringe disposal request areas.
 - **Trend Analysis & Forecasting** to predict future demand patterns.
 - **Automated Data Retrieval (API Integration)** ensuring real-time updates.

Key Metrics & Findings



KPI Monitoring:

- **On-Time Completion Rate:** 99.99%
- **Overall Completion Rate:** 99.89%
- **Requests Taking More Than 48 Hours:** 64 cases
- **Average Response Time:** 3.92 hours
- **Total Closed Cases:** 33,000+

These metrics indicate a **high operational efficiency** but also highlight the **need to optimize delays exceeding 48 hours**.

Trend Analysis:

- **Seasonal Peaks:** Requests peak during **summer months (June-August)**, reaching **1,316 cases in July 2024**.
- **Long-Term Growth:** A **consistent rise in syringe disposal requests** suggests an increasing demand for harm reduction services.
- **District-Wise Distribution:** **Districts 3, 7, and 2** account for the highest requests (~9.3K, 8.9K, and 7.5K cases, respectively), indicating concentrated areas for policy intervention.

Geospatial Insights:

- Requests **near open spaces and schools have increased** but remain a **small fraction of total requests**.
- The **majority of requests originate from urban core areas**, highlighting the need for targeted interventions in high-density neighborhoods.

Time Trend Prediction & Policy Implications

Forecasting Future Requests

Using **historical data (2021-2024)**, the **SARIMA forecasting model** projects:

- **1,007 requests in Jan 2025**, increasing to **1,519 by June 2025**.
- A **continued upward trend in syringe disposal demand**, requiring proactive planning.

The forecast emphasizes the **need for increased disposal stations, better resource allocation, and adaptive policy measures**.

Recommended Policy Actions

Based on **Power BI insights**, key policy recommendations include:

- **Resource Optimization:** Increase disposal bins and mobile services in **Districts 3, 7, and 2**.
- **Delay Reduction Strategy:** Investigate cases exceeding **48 hours** to improve response time.
- **Dynamic Resource Allocation:** Adjust services seasonally to manage peak request months.
- **Enhanced Public Awareness:** Implement educational campaigns targeting high-risk communities.
- **Continuous Monitoring:** Leverage **API automation** to ensure real-time updates and improve response strategies.

By integrating **Power BI analytics**, this report provides a **comprehensive, data-driven approach** to optimizing **syringe disposal services and harm reduction efforts** in Boston.

NEXT STEPS/ SUGGESTIONS

The **Boston 311 Syringe Request Analysis** provides **critical insights** into the trends, spatial distribution, and forecasting of syringe disposal requests in the city.

Key Findings:

- **Syringe requests have increased over the years, with a peak in summer months (June-August).**
- **High-demand districts (3, 7, and 2) require targeted interventions** to improve harm reduction services.
- **SARIMA forecasting predicts a continued rise in syringe disposal requests** into 2025 and beyond.
- **Power BI Dashboard analysis highlights operational efficiencies and delays**, providing real-time decision-making capabilities.

These findings emphasize the **urgent need for an adaptive, data-driven public health approach** to **mitigate opioid-related risks** and **optimize harm reduction services**.

Policy Recommendations for Betterment

1. Improve Resource Allocation & Response Efficiency

- Increase **syringe disposal stations** and mobile services in high-demand districts (**3, 7, and 2**).
- Establish **emergency response teams** to address requests exceeding **48 hours**.
- Develop **geo-targeted interventions** based on Power BI insights.

2. Enhance Data-Driven Public Health Strategies

- Implement **automated data monitoring using API integration** for real-time tracking and response.

- Expand the **use of predictive modeling (SARIMA, machine learning) to forecast demand trends.**
- Conduct **regular audits** to analyze the effectiveness of harm reduction policies.

3. Expand Access to Rehabilitation and Support Services

- Increase funding and accessibility for **rehabilitation centers** to provide long-term support for individuals struggling with substance use disorders.
- Establish **mobile intervention teams** that connect individuals with treatment resources, including detox programs and mental health services.
- Partner with **local hospitals and community organizations** to create referral pathways for individuals in need of comprehensive rehabilitation services.

4. Targeted Community Engagement & Awareness Campaigns

- Launch **public awareness programs** on **safe disposal practices and opioid harm reduction.**
- Collaborate with **healthcare providers, community organizations, and law enforcement** to address underlying opioid use issues.
- Introduce **educational initiatives** in schools and community centers to reduce stigma around harm reduction programs.

5. Strengthen Policy & Legislation

- Advocate for **increased funding** for syringe disposal services and harm reduction programs.
- Develop **policy frameworks for efficient tracking, reporting, and intervention planning.**
- Partner with **state and federal agencies** to align local policies with national harm reduction strategies.

Next Steps for Implementation

- **Short-Term (Next 6 Months):** Optimize response times, launch educational campaigns, enhance Power BI tracking.
- **Mid-Term (6-12 Months):** Scale up predictive analytics, increase disposal site coverage, expand rehabilitation access, and strengthen policy alignment.
- **Long-Term (1-3 Years):** Implement a city-wide harm reduction strategy with continuous data-driven improvements and integrated rehabilitation services.

By implementing these policy recommendations, Boston can build a **sustainable, proactive approach to addressing the opioid crisis, improving public health outcomes, and reducing syringe-related risks.**

CONCLUSION

This project demonstrates how data-driven insights can significantly enhance public health strategies by providing a deeper understanding of syringe disposal request patterns reported through Boston's 311 service. Through the integration of exploratory data analysis (EDA), statistical modeling, and predictive techniques, we identified critical trends that can help allocate resources more efficiently and address public health challenges effectively. Our findings reveal that syringe requests exhibit seasonal variations, with notable increases during the summer months. Additionally, high-risk areas such as District 3 (Dorchester) and District 7 (Roxbury) account for over 55% of total requests, highlighting the need for targeted interventions in these regions. These insights underscore the importance of optimizing harm reduction services, particularly during peak periods and in high-demand neighborhoods.

Our predictive modeling efforts, utilizing a SARIMA time-series forecasting approach, further reinforce the need for proactive public health responses. The model predicts a continued rise in syringe disposal requests into 2025, with projected monthly requests increasing from 1,007 in January 2025 to 1,519 by June 2025. This growth emphasizes the necessity of adaptive policy measures, such as increasing disposal bins, expanding mobile services, and allocating resources dynamically based on predicted demand. Our Power BI dashboard, which integrates real-time API updates and geospatial mapping, provides a powerful tool for policymakers and public health officials to monitor trends, assess operational efficiency, and make informed decisions. With key performance indicators (KPIs) such as a 99.99% on-time completion rate and geospatial insights identifying high-density syringe disposal zones, the dashboard enables continuous monitoring and rapid response to emerging challenges.

To ensure lasting impact, we recommend a multifaceted strategy that includes improving resource allocation in high-demand districts, enhancing data-driven decision-making through real-time monitoring, expanding access to rehabilitation and support services, and strengthening community engagement initiatives. By implementing these evidence-based recommendations, Boston can build a sustainable and proactive harm reduction strategy, reducing opioid-related risks and improving public health outcomes. Ultimately, this project highlights the transformative potential

of combining advanced data analytics with public health planning to create safer and healthier communities for all.

REFERENCES

- Aguilera, J., Baxter, M., & Tapia, A. H. (2021). Spatial analysis of urban syringe disposal patterns: Implications for public health interventions. *Journal of Urban Health*, 98(4), 567-582.
- Batini, C., & Scannapieca, M. (2006). *Data quality : concepts, methodologies and techniques*. Springer. <https://doi.org/10.1007/3-540-33173-5>
- Boston.gov. (2023). 311 Service Requests Dataset. Retrieved from <https://data.boston.gov/dataset/311-service-requests>
- Boston.gov. (2023a). Public Schools Dataset. Retrieved from <https://data.boston.gov/dataset/public-schools>
- Boston.gov. (2023b). Non-Public Schools Dataset. Retrieved from <https://data.boston.gov/dataset/non-public-schools>
- Boston.gov. (2023c). Open Spaces Dataset. Retrieved from <https://data.boston.gov/dataset/open-space>
- Boston Planning & Development Agency. (2023). Boston Property Parcels Dataset. Retrieved from <https://boston.maps.arcgis.com/home/item.html?id=1b8f68085e614f598bb581c7558c8e5e>
- Buchanan, D., Shaw, S., & Ford, A. (2020). Predictive analytics in public health: Using machine learning to identify high-risk areas for needle disposal. *Health Informatics Journal*, 26(3), 1982-1996.
- Klein, A. Z., & Park, H. S. (2022). Examining the relationship between opioid crisis and syringe disposal sites: A statistical approach. *International Journal of Drug Policy*, 103, 180-197.
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, 36(4), 1165–1188. <https://doi.org/10.2307/41703503>
- Guy, G. P., Zhang, K., Bohm, M. K., Losby, J., Lewis, B., Young, R., Murphy, L. B., & Dowell, D. (2017). Vital Signs: Changes in Opioid Prescribing in the United States, 2006–2015. *MMWR. Morbidity and Mortality Weekly Report*, 66(26), 697–704. <https://doi.org/10.15585/mmwr.mm6626a4>

- Zibbell, J. E., Asher, A. K., Patel, R. C., Kupronis, B., Iqbal, K., Ward, J. W., & Holtzman, D. (2018). Increases in Acute Hepatitis C Virus Infection Related to a Growing Opioid Epidemic and Associated Injection Drug Use, United States, 2004 to 2014. *American Journal of Public Health (1971)*, 108(2), 175–181.
<https://doi.org/10.2105/AJPH.2017.304132>