EXECUTIVE SUMMARY

X

TEAM 51

CHICAGO: EMPOWERING A SAFER COMMUNITY THROUGH DATA-DRIVEN CRIME INSIGHTS

Presented by Team 51

Table of Contents

TEAM 51

EXECUTIVE SUMMARY

Problem Statement

Chicago is the third largest city in the United States. It is a hub for finance, commerce, telecommunications, transportation, and is home to some of the most competitive schools in the country. In 2023, **violent crime in Chicago was at an 11-year high** due to an increase in robbery and assault. **Chicago has the 6th highest homicide rate among top cities**, which is more than four times higher than New York City. Property crime of **automobile theft was at a 20 year high**, while burglaries were at historic lows.

Chicago's mayor has bolstered police funding and resources with the goal of reducing crime. **One of Chicago's main objectives for 2024 is crime prevention** through place-based and people-based approaches, which includes a range of intervention, socioeconomic, education, community involvement, and policing approaches. In addition, **the police chief supports data-driven policing[1]** as a part of the strategy.

> The City of Chicago retained Team 51 to discover and analyze data-backed drivers of crime and recommend ways to reduce it.

Overview and Deliverables

To support Chicago in achieving its crime reduction objectives, Team 51 developed and delivered crime prediction models, a suite of information access tools, and data-driven recommendations to put solutions in the hands of Chicago officials, police and residents.

Analytic Models: Descriptive and predictive analytic models provide historical crime information along with accessible data-driven tools to predict crime.

Information Access
- **Dashboards:** Provides city and community area crime information using 2023 data. Also provides predicted crime count trends and forecasts.
- **Mobile App:** Provides historical crime data at the community area and city levels as well as city level predicted crime trends.
- **Chatbot:** Provides quick access to all of the information developed for this project.
- **Website**: Provides easy access to the dashboards and chatbot where data can be monitored and analyzed.

Recommendations: Data-backed recommendations on which geographic crime hotspots and which crime drivers are the most impactful in crime reduction. Future business recommendations are also included.

---

[1] Ponce, Anthony. 2023. "Chicago police superintendent lays out 2024 goals for department". Fox News Chicago. (fox32chicago.com).Dec 11. https://www.fox32chicago.com/news/chicago-police-superintendent-lays-out-2024-goals-for-department

TEAM 51

Data and Modeling

Six publicly available datasets, primarily from government sources, were thoughtfully researched and obtained to support this important work. Through extensive data cleaning, a high-quality master dataset consisting of 250 thousand individual crime records was developed.

The primary modeling approaches included supervised and unsupervised methods to create predictive models and provide segmentation insights. **Supervised** methods were used for crime type and crime rate predictions. They classify outcomes based on input data by learning the relationship between inputs and corresponding outputs from labeled training examples. Team 51's supervised models include linear and logistic regression, decision tree, random forest, and gradient boosting. **Unsupervised** methods are intended to identify relationships, patterns, and clustered groupings in data that does not have a predefined target variable. Team 51's unsupervised methods include K-means, Clustering and Principal Component Analysis (PCA).

Due to the importance of maintaining transparency to Chicago officials, police, and residents throughout the entire project, **publicly available data and state-of-the-art "glass-box" machine learning and data science methods have been used**.

Key Insights

Team 51's analysis revealed useful insights in understanding crime, predicting crime, and segmenting the combination of factors that tend to cluster together and characterize certain crime profiles.

UNDERSTANDING CRIME

- Crime activity is dominated by property and violent crime (Figure 1).
- The top three crime hotspots are the community areas of Fuller Park, West Garfield Park, and Greater Grand Crossing (Figure 2).
- Higher crime rates correspond with higher temperatures in the summer months.
- Top factors that predict the type of crime (Figure 3) are latitude, longitude, distance to the closest police station, temperature, and crime location description (such as "street" or "apartment").
- Population density and the number of sex offenders in a community area are the top two predictive factors for crime rates.
- Environmental complaints regarding lead, snow removal, tree removal, and sewer cleaning are also associated with crime occurrences.

TEAM 51

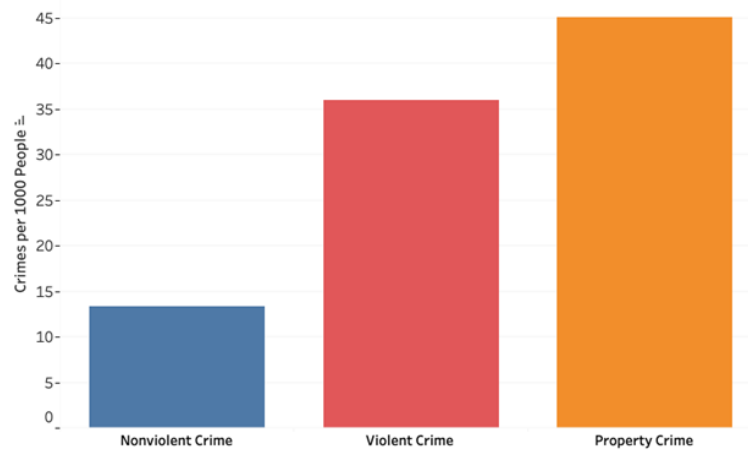*Figure 1. Crime Rate by Crime Type – Crimes Per 1000 People for 2023*

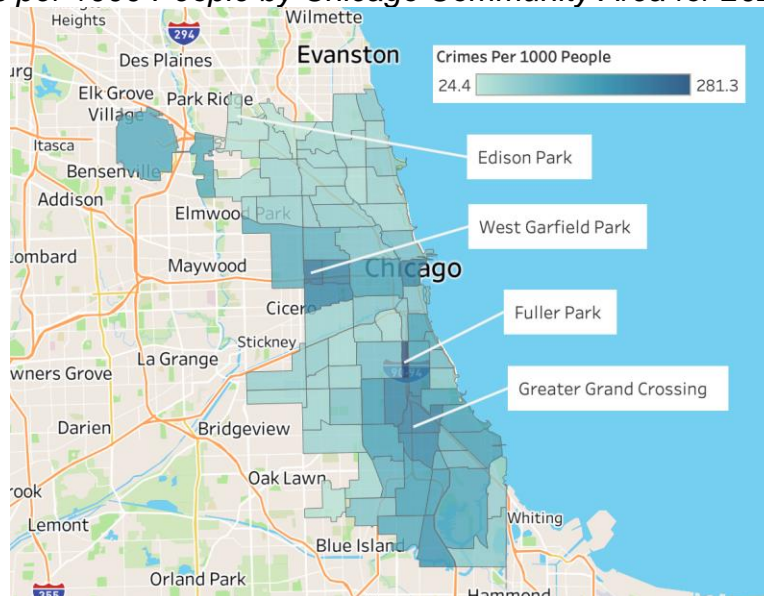*Figure 2. Crimes per 1000 People by Chicago Community Area for 2023*
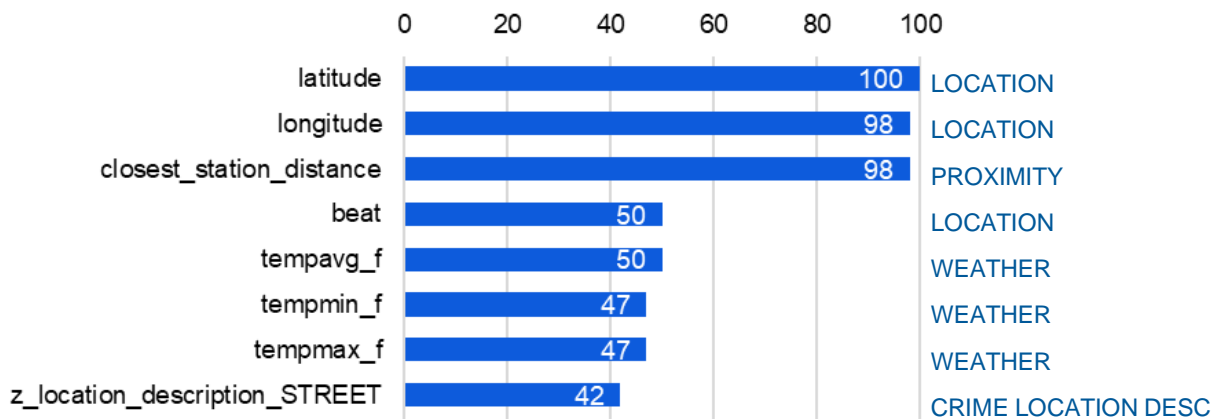

Figure 3. Top Predictive Variables for Crime Type

PREDICTING CRIME

Predictive modeling using supervised methods has shown that a Gradient Boosting model is the most accurate overall for predicting crime type; however, the Random Forest model is nearly as accurate (Figure 4), is easier to explain, and has more useful predictive variables. Therefore, **Random Forest is the selected model** for predicting both crime type as well as the number of crimes by community by type of crime (Figure 5). Additionally, a simplified Tableau exponential smoothing model was also used to provide indicative 2024 forecasts of crime count trends by crime type.

Figure 4. Accuracy of Top Three Models Based on Test Data[2] for Crime Type Prediction
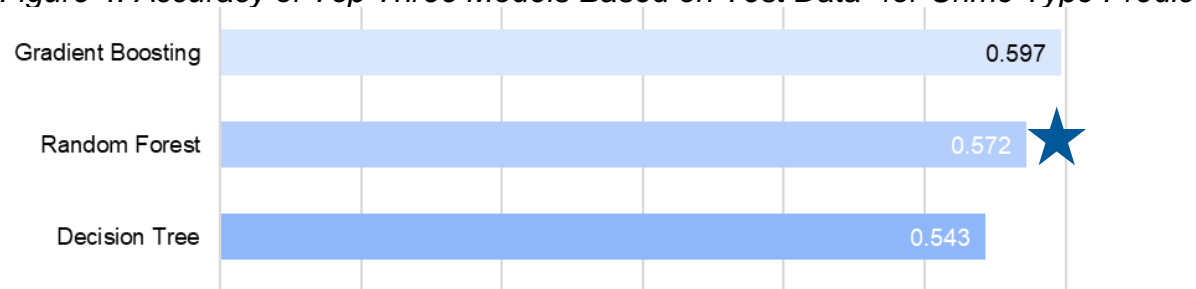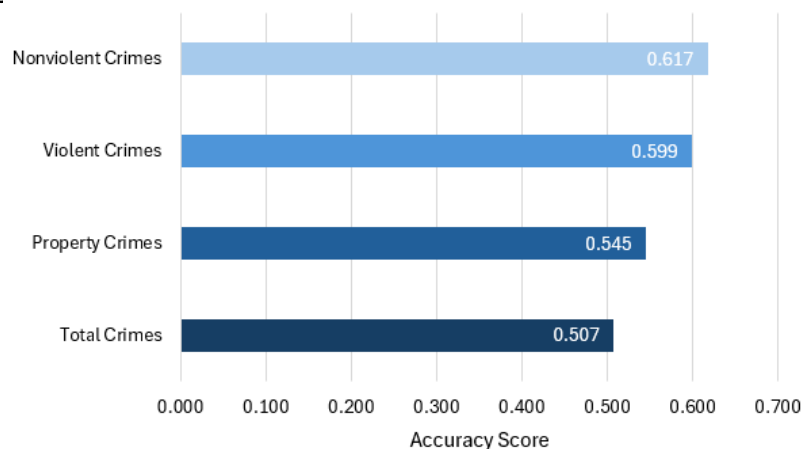

Figure 5. Accuracy of Random Forest Model for Predicting the Number of Crimes for Total and Categorized Crimes Based on Test Data


---

[2] "Crime type" prediction accuracies for related, but less sophisticated studies: Philadelphia 27%; Chicago 37%; San Francisco 43 & 64%; Vancouver 44%.See Initial Findings Reference Section for information.

TEAM 51

SEGMENTATION

Segmentation modeling using unsupervised methods has shown that the K-means model identified four crime profile clusters which highlight a combination of factors that influence crime (Figure 6). Clusters two and three are clearly marked by location in the South and North. One of the more interesting insights was that cluster three had high environmental complaints around lead and snow removal requests. Other groupings appear to be around days of the week such as crimes that were committed on Saturday and Sunday. The Principal Component Analysis (PCA) model identified seven components associated with crime and generally relate to community location, temperature, geographic location, day of week, and environmental complaints around snow removal and sewer cleaning.

It is noteworthy that unsupervised modeling picked up an association between environmental complaints and crime occurrences that supervised modeling did not emphasize. This might indicate a link between potentially underserved/neglected parts of the city and crime.

*Figure 6. Map of Chicago Based on Crime Profile Clusters*

TEAM 51
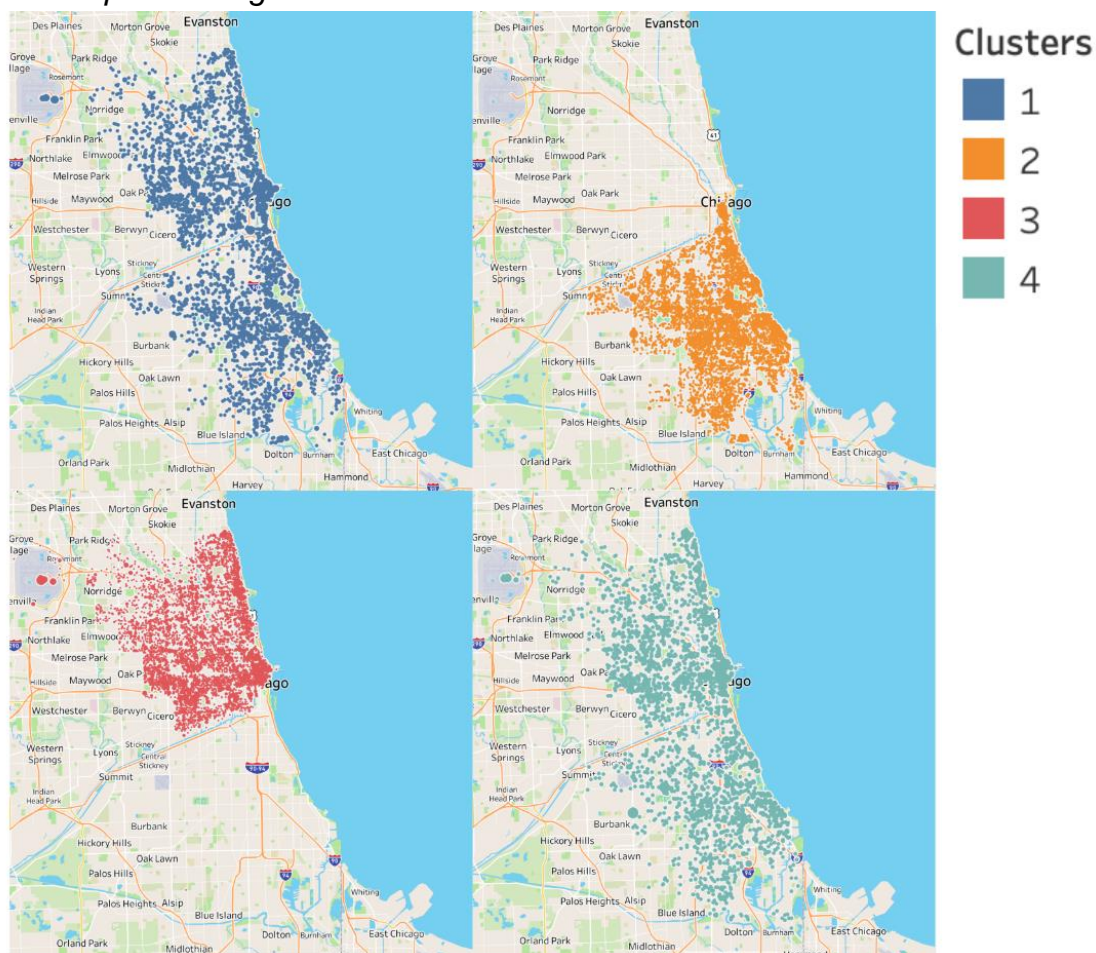
INFORMATION ACCESS

Through development of dashboards, a mobile application, a chatbot, and a website, historical and near real-time crime information is available for the entire city of Chicago by community area. Additionally, crime prediction trends by type of crime (violent, non-violent, and property) for each community area are also available. The chatbot provides efficient access to all of the information contained in the Chicago crime project. A website provides easy access to the dashboards and chatbot where the data can be monitored and analyzed.

Dashboards: The "Chicago Crime Story" dashboard provides visibility to crime insights at a glance using 2023 data. The dashboard provides summary information and drill-down detail by community area (Figure 7). Two additional dashboards were also created called "Chicago Crime - Predictions" and "Chicago Crime - Tableau Forecasting" that show model accuracy and crime trends.

*Figure 7. 2023 Chicago Crime per 1000 People – Dashboard View*
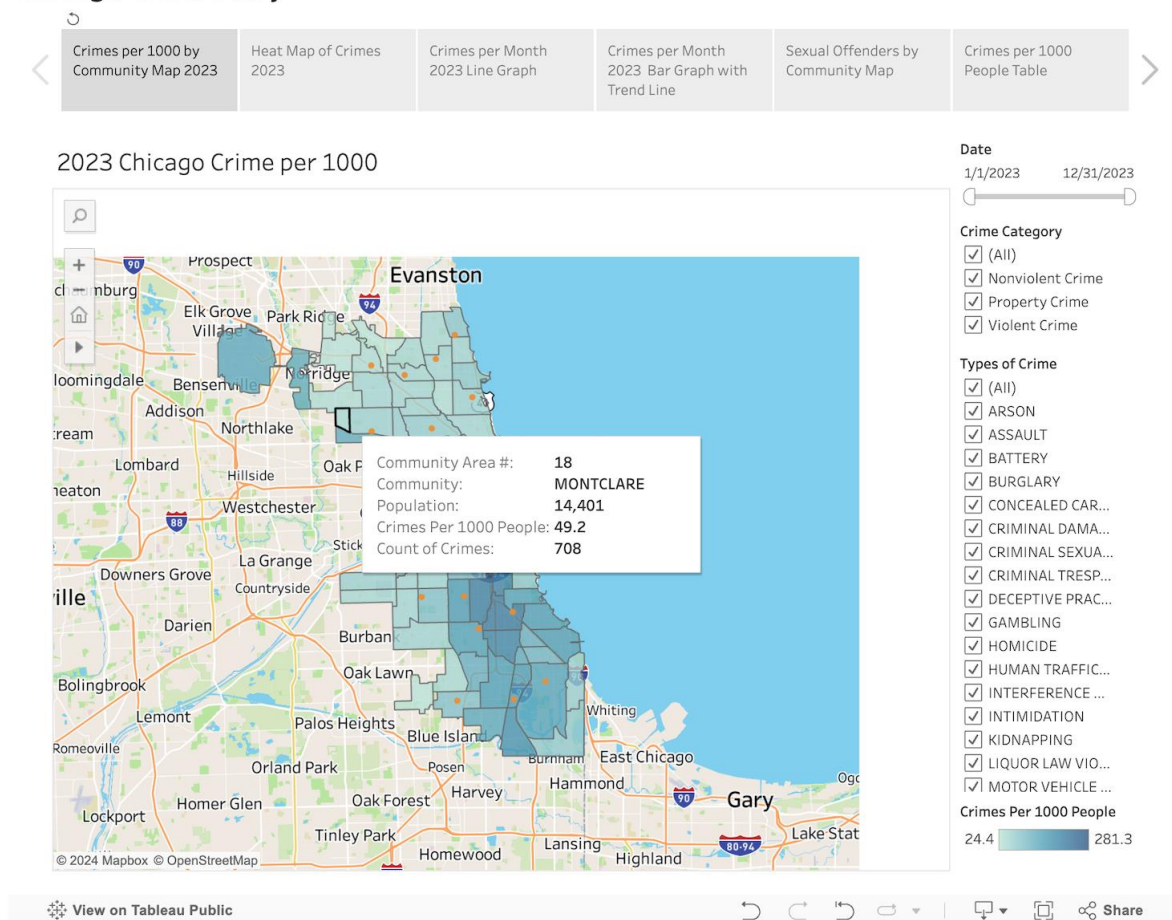
Mobile Application: The mobile app provides historical crime information by community area that is accessible anytime and anywhere and crime forecast at the city level (Figure 8).

Figure 8. Mobile App Screenshot Examples
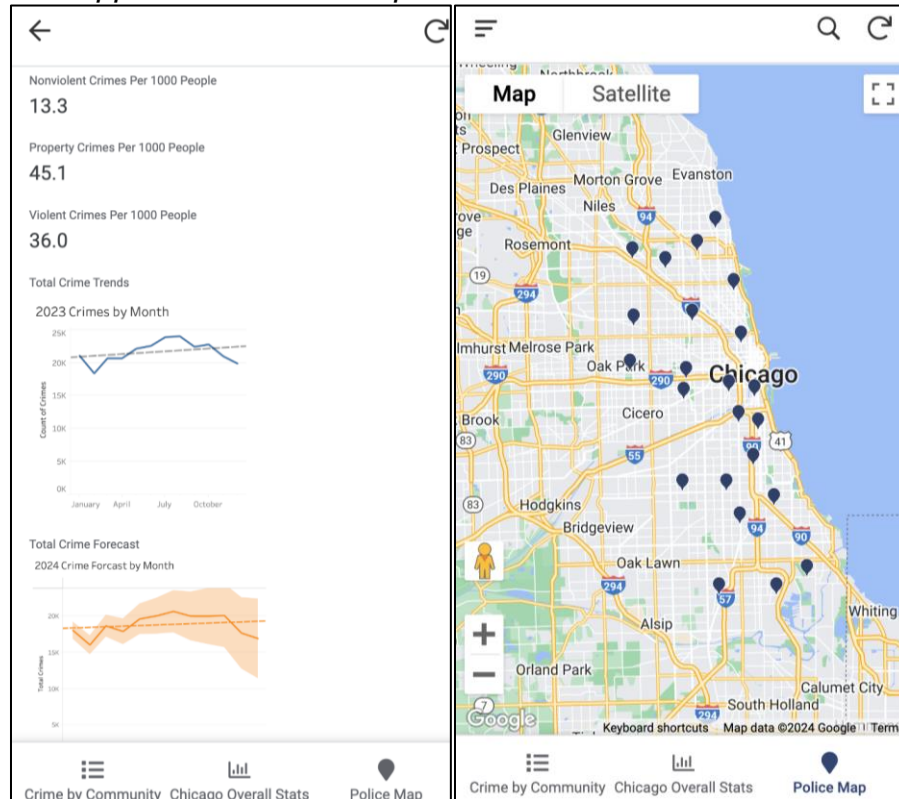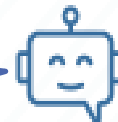

Chatbot: The Chicago Crime chatbot provides easy access to all of the information contained in this project through use of a question and answer chatbot.

USER QUESTION   **CHATBOT RESPONSE**
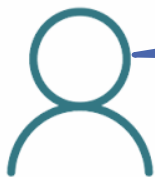
Where are the crime hotspots?

The crime hotspots in Chicago are Fuller Park, West Garfield Park, and Greater Grand Crossing communities.

TEAM 51

Market Expansion

Team 51 has demonstrated its commitment to supporting the City of Chicago's success in reducing crime through use of our innovative, state-of-the-art "glass-box" machine learning and data science methods, along with our dashboard, mobile app, chatbot, and website. A value-added benefit from Team 51's Chicago Crime project is that it is transferable and scalable to other metropolitan areas such as New York City, Los Angeles and other major cities. This makes us an ideal partner in empowering safer communities in other large cities. At the same time, given that Team 51 has already invested in the analytic, data-science, dashboard, mobile app, chatbot, and website infrastructure to support the Chicago crime project, **we expect new projects will deliver a substantial profit for our company while still being of incredible value for future clients.**

Recommendations

Predicting crime is complex. Team 51 has explored and extensively analyzed numerous data-backed insights to crime in Chicago using machine-learning and other state-of-the-art data analysis techniques. Our analysis reveals that a combination of factors drive crime; therefore, multi-faceted efforts that leverage the most important crime drivers will have the greatest impact in helping Chicago meet crime reduction objectives. As a result of insights gained from the Chicago crime project, Team 51 developed three broad areas of recommendations: (1) for the City of Chicago, (2) for future project phases and model enhancements, and (3) for market expansion.

1. **RECOMMENDATIONS FOR THE CITY OF CHICAGO:**

   a. **Evaluate and align crime response resources in the crime hotspot community areas** to address staffing, shift design and police support with crime activity and volume. Geographic location (latitude and longitude) is the top feature that is the most predictive of the type of crime, while population density and sex offender numbers are the most predictive for crime counts.

   b. **Evaluate and promote crime awareness and prevention strategies targeted to of the top factors associated with crime:**
      - **Geographic location** - crime hotspot community areas
      - **Weather** – warmer summer months
      - **Crime location** – streets and apartments

      All of these factors are higher predictors of crime. Supporting awareness, police and community resources, and other prevention strategies that target these factors will have a positive impact on crime. Additionally, greater awareness by citizens alerts them to be mindful of these factors with the goal of preventing crime from happening in the first place.

TEAM 51

c. **Deploy and promote information access tools** for Chicago officials, police, and residents. These tools provide data availability and promote data and information transparency.

- A convenient **website** offers easy access to the "Chicago Crime story" **dashboards** and **chatbot**. The **dashboards** provide access and easy to understand data visibility to crime information by community area. Increasing awareness of crime information will help keep Chicago attuned to crime trends and changes. The **chatbot** provides easy access and reference to information contained in the Chicago Crime project.

- The **mobile app** puts historical crime data for 2023 at the community and city levels in the hands of Chicago officials, police, and residents. It also uses the Tableau Forecasting models to predict overall crime trends at the city level.

d. **Review mitigation of 311 environmental complaints**, especially related to requests for lead tests, snow removal, tree clean-up/removal, and sewer cleaning to assure that potential crime factors are also considered and addressed along with mitigation of the 311 complaints.

2. **RECOMMENDATIONS FOR FUTURE PROJECT PHASES AND MODEL ENHANCEMENTS**:
As more data is gathered along with resultant insights about Chicago crime mitigation strategies, the models and information access tools should be maintained and augmented with the newest machine-learning enhancements when appropriate.

3. **RECOMMENDATIONS FOR FUTURE MARKET EXPANSION:**
Team 51's innovations and investment from the Chicago project are scalable to other markets with minimal additional investment. Similar offerings should be marketed by our company to other large metropolitan areas in order to increase revenue from the intellectual property that has already been developed while offering future clients innovative solutions at a great value. Team 51 has already assessed publicly available data for Los Angeles and New York City. **We recommend that we engage Los Angeles and New York City officials to present and market our solution** that will equip police, businesses, and residents with the tools they need to proactively address and mitigate crime.

TEAM 51

APPENDIX

Appendix A: About Team 51

Nicole Crothers
Data Scientist, Data Exploration, Predictive Models

Booga Gilbertson
Data Scientist, NLP, Communications

Fatima Gower
Data Scientist, Data Exploration, Visualization

TK Kim
Data Scientist, Data Engineer, Cloud Platforms

Matt Riegsecker
Data Scientist, Project Manager, Data Analysis

PARTNERING WITH YOU TO EMPOWER A SAFER COMMUNITY THROUGH DATA-DRIVEN CRIME INSIGHTS