

New York Crime Dashboard

Team Panini Head

Data Science Capstone Project Launch Report

Date:

04/08/2021

Team Members:

Name: Ambrose Karella

Name: Janam Patel

Name: Naimish Bizzu

The System/Product

System/Product Name:

New York Crime Dashboard

Introduction:

Safety is a major consideration whenever people navigate life. Whether it's picking a place to live, a new car, or a new activity or sport, safety is usually one of the first things that are brought up. While crime is not necessarily an indicator of safety, it is something people usually look into.

With that, our team would like to explore one of America's most visited destinations, New York. We can utilize publicly available crime data in New York City to pull insights about crime trends. After collecting this information, we can cross reference spikes and dips with events, and investigation factors in high and low crime periods.

After compiling all of our information, we will deliver a report in with various forms of visualizations. Two tools we would like to integrate into our final report is ArcGIS and an interactive dashboard to visualize data.

Highlighted Features:

- Collect and clean crime-related data throughout New York.
 - Data Size may vary depending on time range and conditions set to pull the data.
 - [Data](#) – link to data
- Perform data analysis to investigate any time, events or location-based trends in crime.
- Trends around and during election times.
- After finding trends, investigate any relevant causes for increases and decreases in crime.
 - If this yields more data that can be collected and evaluated, we will integrate this into our project
- Create geographic visualizations on crime and contributing factors using ArcGIS

Sponsor or Proxy User:

We do not have a sponsor or end user for this project, but we intend to release our work via GitHub and/or Kaggle to allow the public to explore our methods and findings.

Issues:

- New Tools
 - While new tools like ArcGIS seem promising for this project, our team's experience with these tools are limited. Learning and implementing these new tools might slow us down.
- Correlation of Data
 - Crime can be a very touchy subject. Although our team may find some relationships, contributions to crime are dependent on many factors. No single factor will be the sole contributor to crime, even if a strong correlation exists.

- Finding all issues that contribute to crime would be a project of its own and is likely well outside the scope of this project.
- Quality of Data
- Storage
 - We're using MySQL to store data but we don't know how much storage it'll use.

The Team

Team Name:

Team Panini Head

Team Members and their specialties:

Ambrose Karella

As a graduate in game design, I have an interesting skillset that can help solve problems in creative ways. With my transition into the data science program, I would like to leverage my knowledge to add value to any project I work on. My outcome of this project is to challenge myself using python, GIS tools, and new visualization tools to create insights throughout our work these next 6 months.

Janam Patel

I graduated with a bachelor's in computer science, so I have experience in coding for C++ and Python. I have interned at a startup where I held the role of Sales and Data Operations Intern, where I learned a lot of data management and data cleaning. I'm looking to strengthen my coding skills in python and also learn about data management at a deeper level. From this project, I hope to strengthen my skills in EDA, visualization and ML.

Naimish Bizzu

My skills include python, visualizing data using Tableau, and project management. I am from an electronics and communications background. I also have solid skills in pulling via Rest API calls and EDA in both python and R. I have worked on visualization tools like TABLEAU and IBM COGNOS. I have some knowledge in SQL and cloud-based data systems after completing a course on database management systems.

Team Communication:

For communication and project management, our team will be using a variety of software, this includes:

- Trello
 - Used for project management and task organization
- GitHub
 - We will utilize GitHub for maintaining any code, programs, or non-database files that contribute to our project.
- MySQL
 - Over the course of last term, we built a server to host database information. We plan to use this to store any data to allow us to work synchronously without issues.
- Slack
 - For communication, our team has created a slack channel. This allows us to message each other more efficiently than email.

Team Issues:

- Consistency
 - Team members might explore vastly different factors to crime. This could make our project appear less thematically consistent.
- Time-zones / Location
 - Most team members reside outside of Philadelphia, with one student being out of country. This could create problems in communication.

Table of Contributions

The table below identifies contributors to various sections of this document.

	Section	Writing	Editing
1	Project	Ambrose	Janam
2	Team	Naimish	Naimish
3	Plan	Ambrose	Janam

Grading

The grade is given on the basis of quality, clarity, presentation, completeness, and writing of each section in the report. This is the grade of the group. Individual grades will be assigned at the end of the term when peer reviews are collected.