Northeastern University Police Department (NUPD)
Crime Analysis

Northeastern University
College *of* Engineering

10/14/2022

Mechanical and Industrial Engineering Department

NUPD (Northeastern Police Department Crime Data Analysis)

Introduction:

The Investigative and Special Operations Unit of the Northeastern Police Department plays a pivotal role in ensuring the safety and security of the university campus. In addition to safeguarding the community, it provides a comprehensive range of investigative and crime prevention services, serving as a vital asset to the entire university department.

Problem Statement:

This dynamic dashboard has been meticulously designed to address the strategic allocation of resources within the Northeastern Police Department. Its primary objective is to empower the department to take a more proactive approach to mitigating incidents across the university campus. By harnessing the power of data analysis, this dashboard offers the potential to significantly enhance the efficiency and effectiveness of resource allocation strategies.

Data Description:

Dataset Source — https://nupd.northeastern.edu/
Northeastern Police Department

The dataset featured in this analysis comprises reported incidents of crime that have transpired at Northeastern University within the timeframe spanning from 2021 to the present, as well as the most recent seven days. It is important to note that the data is sourced directly from the Northeastern Police Department's official website. Given the diverse nature of the information, it is imperative to underscore that these reports remain unverified and are provided to the Police Department. Consequently, neither expressed nor implied, is any guarantee of the data's accuracy, completeness, timeliness, or sequential correctness. It is vital to refrain from employing this data for the purposes of temporal comparisons.

S.No.	Fields	Description	Data Type
1	Incident Number	Incident Number for internal recording of incidents.	Numeric
2	Date Reported	Date when the incident was reported	Date Time
3	Time Reported	Time the incident was reported	Date Time
4	Incident Types	Different Types of Incidents	Character
5	Location	Location where the incident occurred	Character
6	Latitude	Latitude	Numeric
7	Longitude	Longitude	Numeric
8	Date Occurred	Date on which the incident occurred	Date Time
9	Time Occurred	Time the incident Occurred	Date Time
10	Disposition	Action taken by NUPD	Character
11	Narrative	Brief description of the Incident	Character

Table 1: Data Description

Inspiration

1. What are the major incidents happening across our university?
2. What categories of crime exhibited the greatest year-over-year increase between 2015 and 2016? Which month has the greatest number of incidents?
3. Which places across our university can be described as hotspots?

Design Process & Methodology:

We started with the NUPD crime statistics in this design approach. This dataset includes parameters such as Incident type, time of occurrence, time reported, and location. We utilized Python to transform the data into excel because it was gathered in pdf format.

Python packages & Libraries used:

pip install tabula-py
pip install pandas

import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import datetime

We then converted the data from the PDF to Excel. Most of the data is cleaned in Excel by changing the time from minutes to hours and utilizing the calculation field to add new columns to the tableau. The null values and outliers in the dataset were first removed. We also altered the time and date formats in Python to assist us plot the pattern.

Fig 1: Load Data - Python Code

Insights:

1. Monthly Incident Trends: The graph depicting incidents by month allows for a clear understanding of the volume of occurrences in both 2021 and 2022.

2. Crime Hotspots: Utilizing a heatmap, we've identified areas within Northeastern University most affected by various types of incidents, with a focus on the Boston location.

3. Weekday vs. Weekend Incidents: An analysis of incidents on weekdays and weekends offers valuable insights into the types and volume of incidents on different days of the week.

4. Hourly Analysis: A table comparing weekdays to hours reveals when and on which days incidents commonly occur, offering a valuable resource for safety measures in and around the university.

5. High-Incident Types: A heatmap illustrates incident counts for 2021 and 2022, with "Investigate Person," "Proctor Sign in Violation," and "Investigate Premises" emerging as the most frequently reported incident types during these years.

6. Neighborhood Analysis: We've explored the location-wise distribution of incidents, highlighting that the majority of crimes occur on Columbus Ave during the night. This proximity to the NUPD office simplifies regular monitoring.

7. Yearly and Monthly Incident Summary: Users can easily access information on incidents at Northeastern University in 2021 and 2022, simplifying the process of obtaining an overview of crime patterns during these years.

These insights underscore the potential for leveraging Text Mining and Analysis to gain deeper insights into the nature of crimes within the university. Furthermore, optimizing the categorization of each incident type and providing specific location details, such as zip codes, can enhance the analysis.

Future Scope:

1. Standardized Event Categories: The ability to focus on specific event categories is hindered by the current lack of data. We recommend standardizing event categories, potentially creating broader categories, like "Theft," to encompass various subcategories such as bicycle theft or electronic theft.

3. Geographical Focus: The generic manner in which locations are provided by NUPD often leads to dataset misuse. To rectify this, it's essential to incorporate suitable area names with zip codes, enabling more accurate geographical analysis. Future analyses could delve into specific locations and communities where incidents are concentrated, with a particular emphasis on utilizing zip codes for precision.

4. Enhanced NUPD Reporting: A more detailed and explicit account of NUPD's actions in response to incidents will provide users with a clearer understanding of which areas are relatively safer.

5. Culprit Identification: The addition of a new field to distinguish whether significant crimes involve faculty/students or external parties can be a valuable asset for increasing awareness and tailored security measures.

6. Severity Score: By introducing a "Severity Score," the seriousness of each incident can be more comprehensively assessed, allowing for prioritization of responses and resource allocation.

These future scope considerations aim to further refine and enrich the analysis, contributing to a more comprehensive understanding of campus safety and security.