Project-2 Proposal

Data Nexus

2024-11-13

Project Proposal: Public Safety Awareness Campaign Based on Crime Data

Objective:

To develop an impactful public safety awareness campaign by analyzing crime data in Los Angeles from 2020 to the present. This campaign aims to educate residents on prevalent crime types, identify high-risk areas, and promote proactive safety measures to enhance community well-being and reduce incidents.

Description:

Using crime data provided by the LAPD, this project will analyze trends and patterns, focusing on common crime types and their geographical distribution. The insights will help shape an awareness campaign tailored to community needs, encouraging residents to adopt practical safety measures and fostering a collaborative relationship between the LAPD and the community.

Data Source

The dataset is sourced from the LAPD, covering reported crime incidents in Los Angeles from 2020 to the present. The LAPD crime data is available online and includes detailed, anonymized records of various crime types. Due to the transition to a new NIBRS-compliant Records Management System in March 2024, updates are provided bi-weekly.

Source: LAPD Crime Dataset - Los Angeles Open Data Portal.

Data Dimensions

Each row is a crime incident

Rows –> 984K Columns –> 28

Here is the provided information converted into a markup table:

Column Name	Description	API Field Name	Data Type
DR_NO	Division of Records Number: Official file number made up of a 2 digit year, area ID, and 5 digits	dr_no	Text
Date Rptd	MM/DD/YYYY	date_rptd	Floating Times-tamp

Column Name	Description	API Field Name	Data Type
DATE OCC	MM/DD/YYYY	date_occ	Floating Times-tamp
TIME OCC	In 24 hour military time.	time_occ	Text
AREA	The LAPD has 21 Community Police Stations referred to as Geographic Areas within the department. These Geographic Areas are sequentially numbered from 1-21.	area	Text
AREA NAME	The 21 Geographic Areas or Patrol Divisions are also given a name designation that references a landmark or the surrounding community that it is responsible for.	area_name	Text
Rpt Dist No	A four-digit code that represents a sub-area within a Geographic Area. All crime records reference the "RD" that it occurred in for statistical comparisons.	rpt_dist_no	Text
Part 1-2		part_1_2	Number
Crm Cd	Indicates the crime committed. (Same as Crime Code 1)	crm_cd	Text
Crm Cd Desc	Defines the Crime Code provided.	crm_cd_desc	Text
Mocodes	Modus Operandi: Activities associated with the suspect in commission of the crime. See attached PDF for list of MO Codes in numerical order.	mocodes	Text
Vict Age	Two character numeric	vict_age	Text
Vict Sex	F - Female, M - Male, X - Unknown	vict_sex	Text
Vict Descent	Descent Code: A - Other Asian, B - Black, C - Chinese, D - Cambodian, F - Filipino, G - Guamanian, H - Hispanic/Latin/Mexican, I - American Indian/Alaskan Native, etc.	vict_descent	Text
Premis Cd	The type of structure, vehicle, or location where the crime took place.	premis_cd	Number
Premis Desc	Defines the Premise Code provided.	premis_desc	Text
Weapon Used Cd	The type of weapon used in the crime.	weapon_used_cd	Text
Weapon Desc	Defines the Weapon Used Code provided.	weapon_desc	Text
Status	Status of the case. (IC is the default)	status	Text
Status Desc	Defines the Status Code provided.	status_desc	Text
Crm Cd 1	Indicates the crime committed. Crime Code 1 is the primary and most serious one.	crm_cd_1	Text
Crm Cd 2	May contain a code for an additional crime, less serious than Crime Code 1.	crm_cd_2	Text
Crm Cd 3	May contain a code for an additional crime, less serious than Crime Code 1.	crm_cd_3	Text
Crm Cd 4	May contain a code for an additional crime, less serious than Crime Code 1.	crm_cd_4	Text
LOCATION	Street address of crime incident rounded to the nearest hundred block to maintain anonymity.	location	Text
Cross Street	Cross Street of rounded Address	cross_street	Text
LAT	Latitude	lat	Number
LON	Longitude	lon	Number

1. Data Analysis for Awareness

- **Objectives**: Identify and understand the most common crime types and the neighborhoods most impacted. Analyze demographic factors to tailor the campaign for specific areas.
- **Process**:

```
# Load necessary package
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
# Get a list of all CSV files in the directory
csv_files <- list.files(path = "Data/", pattern = "*.csv", full.names = TRUE)

# Read and combine all CSV files into one data frame
crime_data <- do.call(rbind, lapply(csv_files, read.csv))

# View the merged data
head(crime_data)
```

```
##       DR_NO Date.Rptd DATE.OCC TIME.OCC AREA  AREA.NAME Rpt.Dist.No Part.1.2
## 1 190326475     43891    43891     2130    7   Wilshire         784        1
## 2 200106753     43870    43869     1800    1    Central         182        1
## 3 200320258     44146    44139     1700    3  Southwest         356        1
## 4 211904005     44196    44196     1220   19    Mission        1974        2
## 5 210705560     44162    44162     1800    7   Wilshire         776        1
## 6 200412582     44083    44083      630    4 Hollenbeck         413        1
##   Crm.Cd                                Crm.Cd.Desc        Mocodes Vict.Age
## 1    510                           VEHICLE - STOLEN                      0
## 2    330                      BURGLARY FROM VEHICLE 1822 1402 0344       47
## 3    480                             BIKE - STOLEN      0344 1251       19
## 4    624                     BATTERY - SIMPLE ASSAULT            416       26
## 5    230 ASSAULT WITH DEADLY WEAPON, AGGRAVATED ASSAULT      1309 0400       31
## 6    510                           VEHICLE - STOLEN                      0
##   Vict.Sex Vict.Descent Premis.Cd                                  Premis.Desc
## 1        M            O       101                                       STREET
## 2        M            O       128            BUS STOP/LAYOVER (ALSO QUERY 124)
## 3        X            X       502 MULTI-UNIT DWELLING (APARTMENT, DUPLEX, ETC)
## 4        M            H       502 MULTI-UNIT DWELLING (APARTMENT, DUPLEX, ETC)
## 5        F            O       101                                       STREET
## 6                     1       101                                       STREET
##   Weapon.Used.Cd                          Weapon.Desc Status
```

```
## 1             NA                                                    AA
## 2             NA                                                    IC
## 3             NA                                                    IC
## 4            400 STRONG-ARM (HANDS, FIST, FEET OR BODILY FORCE)     IC
## 5            307                                          VEHICLE  AA
## 6             NA                                                    IC
##    Status.Desc Crm.Cd.1 Crm.Cd.2 Crm.Cd.3 Crm.Cd.4
## 1 Adult Arrest      510      998       NA       NA
## 2  Invest Cont      330      998       NA       NA
## 3  Invest Cont      480       NA       NA       NA
## 4  Invest Cont      624       NA       NA       NA
## 5 Adult Arrest      230       NA       NA       NA
## 6  Invest Cont      510       NA       NA       NA
##                                LOCATION Cross.Street     LAT       LON Year
## 1 1900 S  LONGWOOD                    AV              34.0375 -118.3506 2020
## 2 1000 S  FLOWER                      ST              34.0444 -118.2628 2020
## 3 1400 W  37TH                        ST              34.0210 -118.3002 2020
## 4 9000    CEDROS                      AV              34.2336 -118.4535 2020
## 5 4500    LOMITA                      ST              34.0452 -118.3351 2020
## 6                        200 E  AVENUE 28             34.0820 -118.2130 2020
```

```
# Get the dimensions and column names of the dataset
dim(crime_data)
```

```
## [1] 990293     29
```

```
colnames(crime_data)
```

```
##  [1] "DR_NO"          "Date.Rptd"      "DATE.OCC"       "TIME.OCC"
##  [5] "AREA"           "AREA.NAME"      "Rpt.Dist.No"    "Part.1.2"
##  [9] "Crm.Cd"         "Crm.Cd.Desc"    "Mocodes"        "Vict.Age"
## [13] "Vict.Sex"       "Vict.Descent"   "Premis.Cd"      "Premis.Desc"
## [17] "Weapon.Used.Cd" "Weapon.Desc"    "Status"         "Status.Desc"
## [21] "Crm.Cd.1"       "Crm.Cd.2"       "Crm.Cd.3"       "Crm.Cd.4"
## [25] "LOCATION"       "Cross.Street"   "LAT"            "LON"
## [29] "Year"
```

Insights to Gather:

- Crime types with the highest incidence rates.
- Patterns of crime over time and by location.
- High-risk areas and times for specific crimes.

2. Campaign Development

- **Materials Creation**: Design accessible and informative materials such as flyers, social media graphics, and posters focusing on common crime prevention strategies.
- **Community Collaboration**: Partner with community leaders and local organizations to enhance reach and credibility. Materials will include crime prevention tips tailored to prevalent crime types and specific demographics.
- **Workshops**: Organize in-person and virtual sessions covering safety techniques and best practices, creating a platform for direct community engagement.

3. Engagement Strategies

- **Social Media**: Develop dedicated channels for sharing real-time safety tips, updates on recent incidents, and interactive content like quizzes and polls.
- **Community Feedback Mechanisms**: Establish feedback loops, such as online forms and QR codes on physical materials, allowing residents to report concerns, provide suggestions, and informally share safety-related experiences.

4. Feedback and Continuous Improvement

- **Data-Driven Adjustments**: Regularly analyze feedback and updated crime data to refine the campaign materials and engagement approach. This iterative method will ensure that the campaign remains relevant and impactful.
- **Ongoing Reporting**: Provide periodic reports summarizing campaign outcomes, community feedback, and any observable reductions in specific crime types.

Research Questions

Question 1: What are the most common types of crimes and high-risk areas in Los Angeles?

- **Variables**: Crime type, location, date, and time.
- **Plan**:
  - Clean and preprocess crime data, focusing on crime type, time, and location.
  - Perform trend analysis to reveal patterns over time and across neighborhoods.
- **Visualizations**:
  - A heatmap for spatial distribution of crime.
  - A time series chart for crime types by year(s).

Question 2: What is the distribution of victim ages, sexes, and descents across different crime types in the LAPD crime data?

- **Variables**: Vict Age, Vict Sex, Vict Descent, crm_cd_desc
- **Plan**:
  - Basis the top Crime types the bifurcation of visual for victims age, sex and descent
  - Clean and preprocess crime data, focusing on crime type, sex, age and descent
- **Visualizations**:
  - Parallel Coordinate Plot: Use a parallel coordinate plot to visualize the relationships among victim age, sex, descent, and crime type.
  - Bubble Chart for Victim Age and Crime Type: Use a bubble chart where the size of the bubble represents the frequency of crimes for a specific combination of victim age (vict_age) and crime type (crm_cd_desc).

Analysis Plan

Task	Status	Assignee	Due Date	Priority	Summary
Data Collection and Cleaning	Completed		Week 1	High	Collect and clean LAPD data for analysis.
Crime Trends Analysis	Completed		Week 1	High	Identify common crime types and locations.
Campaign Material Design	In Progress		Week 2	High	Develop digital and print materials.
Community Workshop Planning	Not Started		Week 3	Medium	Plan virtual and in-person sessions.
Feedback Mechanism Setup	Not Started		Week 3	Medium	Create channels for collecting feedback.
Campaign Rollout	Not Started		Week 4	High	Launch awareness campaign in target areas.
Impact Evaluation	Not Started		Week 5	High	Track crime rates and feedback post-launch.

Expected Outcomes

- **Increased Awareness**: Community members will gain knowledge of prevalent crimes and learn safety practices.
- **Crime Reduction**: The campaign may contribute to a decrease in frequently reported crimes in high-risk areas.
- **Improved Community Relations**: By actively engaging residents, the campaign fosters trust and cooperation with the LAPD.
- **Actionable Feedback**: Feedback collected from the community will support ongoing improvements and adjustments to the campaign.

This proposal presents a targeted safety awareness campaign that utilizes crime data to empower residents and promote safer neighborhoods. By fostering collaboration between the community and law enforcement, we aim to strengthen relationships and create a more informed and secure environment for all.