Predicting Crime in Los Angeles

CAPSTONE 1 - Project Proposal

by Oliver Bohler

The Problem: The past four years have been exceptionally challenging for the Los Angeles Police Department (LAPD). Beyond the global COVID-19 pandemic, the department has faced large-scale protests, incidents of looting, calls to defund the police, and increased scrutiny of police brutality. These factors have contributed to a significant shortage of police officers, primarily due to resignations and decreased hiring rates. Consequently, residents are experiencing longer response times for 911 calls, depending on the crime's severity. In response, many communities have initiated neighborhood watch programs to compensate for the perceived lack of security.

The objective is to develop a predictive model for crime patterns, times, and victim profiles within the LAPD's jurisdiction. This model will provide valuable insights for both the police force and residents, enabling proactive crime prevention measures and stopping crime before it happens.

The Client: The Los Angeles Police Department needs to adopt new strategies to address the officer shortage and enhance public trust. By understanding crime hotspots and trends, the LAPD can more effectively deploy resources, increase arrest rates, and ultimately reduce crime. Additionally, starting in March 2024, the LAPD will implement a new Records Management System for reporting crimes and arrests, in compliance with the FBI's mandate to collect NIBRS-only data.

The Data: The two datasets used reflect incidents of crime in the City of Los Angeles dating back to 2013. This data is transcribed from original crime reports that are typed on paper and therefore there may be some inaccuracies within the data. The data is taken from the following links:

https://catalog.data.gov/dataset/crime-data-from-2020-to-present

https://catalog.data.gov/dataset/crime-data-from-2010-to-2019

Approach to the Problem: The goal of this project is to develop a predictive model that classifies crime hot-spots and trends in Los Angeles. To facilitate effective classification, crimes and jurisdictions will be grouped into relevant categories.

Steps:

1. **Data Wrangling:** Cleaning and preprocessing the data to ensure accuracy and consistency.
2. **Exploratory Data Analysis (EDA):** Conduct a thorough EDA to uncover patterns and insights, which will inform the modeling process.
3. **Pre-Processing and Modeling:** Identify and implement the most suitable machine learning model for predicting crime trends and hotspots.

Key Questions to Address:

- What are the general trends for different types of crimes?
- How effective is the LAPD in making arrests?
- Where and when are specific crimes most likely to occur?
- Is there evidence of racial disparity in arrests?
- Which areas are experiencing an increase in particular types of crimes?

Deliverables:

The project will be documented in Jupyter Notebooks, with the code made available on GitHub. Additionally, a slide deck will be created, and the project will conclude with a comprehensive final report summarizing the insights gained. Throughout the project, I will ensure that my coding is well-documented and explained.