# DATS6103_DATA MINING

# Project Proposal


## Team 5

Harshith

Muhannad

Sissoko


## Analysis of the Crime Rate Incidents in Los Angeles: Focus on 2024 Data

## Background:

Crime remains one of the most critical issues in major metropolitan areas like Los Angeles, where social, economic, and environmental factors contribute to variations in criminal activity. Understanding crime trends is essential for improving community safety, optimizing police resource allocation, and guiding prevention programs.

In this project, we will analyze reported crime incidents in Los Angeles during **2024**, focusing on how patterns vary across time, geographic areas, and demographic attributes. The dataset contains key variables such as **date and time of occurrence, area name, crime type, weapon involvement, and victim characteristics (age, sex, and descent).**

Our study aims to answer key questions about spatial and temporal distributions of crime, the relationship between weapon usage and victim demographics, and area-wise crime intensity relative to population. By identifying patterns and correlations in the data, we will provide actionable insights that can support law enforcement planning and community awareness.

We will apply various Exploratory Data Analysis (EDA) to reveal hotspot regions and high-risk times of day, while inferential techniques will help evaluate statistical differences between groups.

## Source:

The dataset was obtained from **Data.gov**, published by the Los Angeles Police Department (LAPD). It contains detailed records of reported criminal incidents across Los Angeles dating back to 2020, including data transcribed from original paper crime reports.
https://catalog.data.gov/dataset/crime-data-from-2020-to-present

**GitHub Repository:** https://github.com/Ebrima170/G5_Data-Mining-Project-Repo

## Smart Questions

1. Which Areas had the highest crime concentration and crime rate per 10,000 residents in 2024?
   *Explanation:* Quantifies and compares crime intensity across areas while accounting for population differences to identify the most affected regions in Los Angeles.

2. Is there a significant difference in crime distribution between weekdays and weekends?
   *Explanation:* Tests whether temporal variations exist in crime frequency, providing insights into situational patterns.

3. Does weapon involvement vary by crime category and victim sex?
   *Explanation:* Examines demographic and situational influences on weapon-related crimes.