**Project Title: Exploring Crime Patterns in Urban Areas (Pakistan)**

**Team Members (Solo Project):**  
Syeda Munazzah Fatima (23p-0005)

# **Project Description**

## **Origin:**

* Urban areas in Pakistan face significant crime rates, impacting citizen safety and law enforcement efficiency.
* There is a lack of comprehensive, accessible data on crime patterns, making it difficult to understand and address the issue.

## **Objective:**

* Analyze crime data in urban areas of Pakistan to identify patterns in crime types, locations, and times.
* Provide insights that help authorities optimize resource allocation and improve public awareness of crime hotspots.
* Use **data analysis** and **visualization techniques** in Python to present findings clearly and effectively.
* **Web scraping** will be employed if needed to gather additional data from news articles and reports that highlight crime incidents in various urban areas.

**Phase 1: Data Collection and Preparation**

## **1. Data Collection:**

* Crime data will be collected from reliable public sources such as the **Pakistan Bureau of Statistics** and crime news reports from sources like **Dawn**, **Geo News**, and **The Express Tribune**. These sources often provide information on crime incidents in major cities across Pakistan.
* If available, official reports from city police departments or open government data platforms will also be utilized.

## **2. Data Cleaning and Preprocessing:**

* The raw data will likely have inconsistencies such as missing values, incomplete crime descriptions, and format issues.
* Using Python libraries like **pandas and NumPy**, the data will be cleaned, standardized, and categorized by crime type, location, and time (e.g., time of day, day of the week).
* The data will be formatted for easy analysis, including handling geographical data (e.g., city neighborhoods) to enable spatial visualization.

# **Phase 2: Data Analysis and Visualization**

## **1. Data Analysis:**

* The cleaned dataset will be analyzed to identify trends in crime occurrence. These analyses will focus on:
  + **Crime frequency**: Identify which crimes are most common and where they occur most often.
  + **Time-based patterns**: Examine how crime rates change during the day or week to reveal peak times for criminal activity.
  + **Spatial analysis**: Determine which areas of the city are most affected by different types of crime.

## **2. Data Visualization:**

* Using libraries like matplotlib, seaborn, and folium, visualizations will be created to illustrate the findings.
  + **Heatmaps** will show the concentration of crimes in different urban areas.
  + **Bar charts** will represent the most frequent types of crimes.
  + **Line charts** will reveal trends over time (e.g., times of day when crime rates peak).
* These visualizations will help both law enforcement agencies and the public better understand crime patterns and make informed decisions to improve safety.

# **Expected Outcomes:**

* Provide data-backed insights into crime trends in urban areas, focusing on identifying hotspots and vulnerable periods.
* Help authorities optimize their resource allocation for better crime prevention.
* Increase public awareness of crime patterns, allowing citizens to take precautions based on data.