**CS434 – Data Base Theory and Design**

**Project #1**

**Team Database Application (TDA): Part 1 - Relational Database Design**

**Team**

Lipika Baniya | 800794205 | [lbaniya@siue.edu](mailto:lbaniya@siue.edu)

The domain I would like to manage with the TDA is **Washington DC Crime Datasets 2024** by the District of Columbia Metropolitan Police Department (MPD)

**Overview**

The dataset provides comprehensive information on reported crimes within a subset of locations and attributes of incidents. The purpose of the application is to create a relational database system that tracks, stores, and visualizes reported crime incidents across Washington D.C. The application will use real-world dataset provided by the District of Columbia Metropolitan Police Department (MPD). The dataset will help discover crime patterns and aid in strategic decision-making for community safety and law enforcement.

**Entities**

The **“Crimes 2024”** dataset includes the following key entities:

1. Crime\_incident

* ccn VARCHAR PRIMARY KEY : A unique identifier assigned by MPD to each incident report.
* report\_date TIMESTAMP : The date the offense was reported, which may be later than the date the offense actually occurred.
* shift VARCHAR : Time the report was taken. Day shift generally runs between 0700 and 1500 (military time); evening shift between 1500 and 2300, and midnight shift between 2300 and 0700. If the shift is unknown, the field will say "UNK".
* start\_date TIMESTAMP : Crime incident start date and time.
* end\_date TIMESTAMP : Crime incident end date and time.

1. Method

* method\_id INT PRIMARY KEY : Unique key assigned to type of weapon.
* method\_type VARCHAR : Types of weapon used to commit crime.

1. Offense

* offense\_id INT PRIMARY KEY : Unique key assigned to crime offense.
* method\_id INT FOREIGN KEY : Unique key assigned to type of weapon.
* offense\_name VARCHAR : Crime offense.

1. Crime Location

* crime\_id VARCHAR PRIMARY: A unique identifier assigned by MPD to each incident report.
* location\_id INT PRIMARY KEY : Unique key assigned to location crime occurred.
* latitude FLOAT : Latitude (decimal degrees) of Crime Incident.
* longitude FLOAT : Longitude (decimal degrees) of Crime Incident.

1. Location

* location\_id INT PRIMARY KEY : Unique key assigned to geographical location.
* ward INT : Ward ID
* ans VARCHAR : ANC ID
* district INT : Police district
* psa INT : Police Service Areas
* neighborhood\_cluster VARCHAR : Neighbourhood cluster

1. Block Group

* X-block INT : Block X Coordinate
* Y-block INT : Block Y Coordinate

**Relationships**

Crime and Crime Location have many-to-one relationship, which means one crime occurs at one crime location, but one location may be associated with many crime.

Crime and Offense have many-to-one relationship, where one offense can have multiple crimes but one crime can belong to only one offense.

Crime and Method have many-to-one relationship, where one method can have multiple crimes but one crime can belong to only one method.

Crime Location and Location can have many-to-one relationship, where multiple crime locations is described with a single location

Crime and Block Group have many-to-one relationship, where many crimes can happen in one block group.

**Difficult Aspects**

* **Missing Data :** Some rows have x\_coord = 0, y\_coord = 0.
* **Timeline Confusion :** Difference between start\_date, end\_data and report\_date
* **Geographical overlap :** A single incident may belong to overlapping jurisdictional boundaries.

**Important Queries**

* Which neighbourhood clusters have the highest number of reported theft incidents?
* At what time during the day does the crime occur the most?
* Which voting precincts have experienced more than 3 assault reports at the same street block?
* Is there a certain crime types in specific wards or neighbourhoods?
* What are the top 5 blocks with the highest number of reports?
* **How many incidents involved a "Gun" as the method?**

**ERD**

**A diagram of a crime method

AI-generated content may be incorrect.**