

Final Project Milestone 2

Group 24

CIS 5500 Spring 2024

03/09/2024

1. Motivation for the idea/description of the problem the application solves.

As the crime rate started to drop in 2023 compared to the past years, we would like to enhance the safety of our users with an accessible and user-friendly data platform designed to aggregate safety information for Philadelphia based on the user's latitude and longitude coordinates to further reduce the risk.

Understanding which areas are prone to certain types of crime can help residents and visitors avoid potentially dangerous situations.

This information can be instrumental in allocating resources more effectively, developing targeted crime prevention strategies, and enhancing the overall efficiency of policing efforts.

2. List of features you will definitely implement in the application

- a. A home page that summarizes the application's function with quick access to all other features
- b. Based on the latitude and longitude coordinates, notify the users of all information about different types of crime in this area
- c. The crime report search page, which allows sorting by crime type, date, gender, dispatch time, offers both brief summaries and in-depth accounts of incidents. Users can look up all crime reports by category.

3. List of features you might implement in the application, given enough time

- a. Interactive/Static map that shows all available crime data by region
- b. User customize personal alert(based on crime frequency, severity, time)
- c. Police engagement page (police station, police activity)
- d. School location shown on map based on (Lon,Lat), with crime incidents.
- e. Safety rating (1-10 scale) for each area
- f. Optimization of UI aesthetics

4. List of pages the application will have and a 1-2 sentence description of each page

Root page: feature part 2.(b): Home page

A user interface Home page that directs to the following pages.

Page 1: feature part 2.(a):

Show on map that crime is located in different areas of Philly.

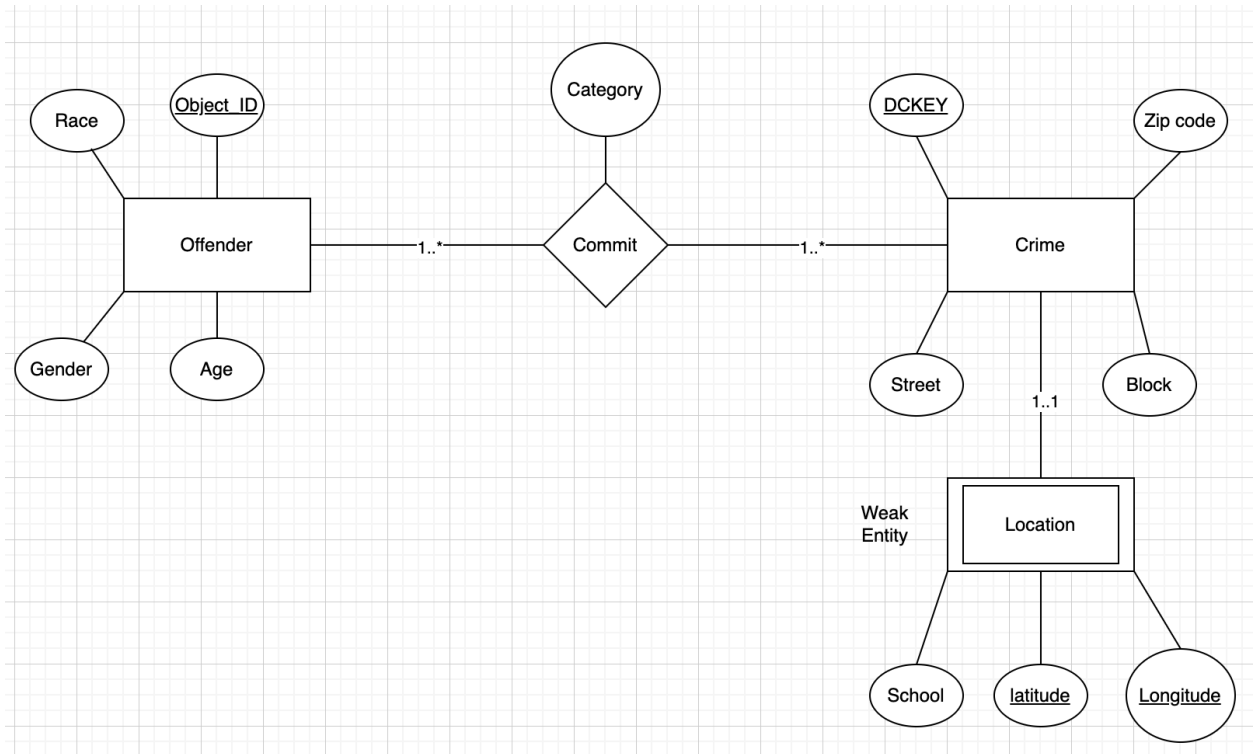
Page 2: feature part 2.(a):

Based on the latitude and longitude coordinates, notify the users of all information about different types of crime in this area.

Page 3: feature part 2.(c)

The crime report search page, which allows sorting by crime type, date, gender, dispatch time, offers both brief summaries and in-depth accounts of incidents. Users can look up all crime reports by category.

5. Relational schema as an ER diagram.



6. SQL DDL for creating the database.

(1) CREATE TABLE Arrests (offence_categories varchar(25), day datetime, defendant_race varchar(10), count int, object_id int, **PRIMARY KEY (object_id)**)

(2) CREATE TABLE Crime (the_geom varchar(100), cartodb_id int, the_geom_webmercator varchar(100), object_id int, dc_dist int, psa int, dispatch_date_time datetime, dispatch_date date, dispatch_time time, hour int, dc_key int, location_block varchar(100), ucr_general int, text_general_code varchar(100), point_x float, point_y float, lat float, lng float, **PRIMARY KEY(dc_key), FOREIGN KEY (object_id) REFERENCES Arrests (object_id)**)

(3) CREATE TABLE GunViolence (dc_key int, race varchar(25), sex varchar(8), fatal varchar(15), date datetime, has_court_case varchar(5), age int, street_name varchar(50), block_number int, zip_code int, council_district int, police_district int, neighborhood varchar(25), house_district int, senate_district int, school_catchment varchar(50), lng float, lat float, **FOREIGN KEY (dc_key) REFERENCES Crime (dc_key)**)

7. **Explanation of how you will clean and pre-process the data. This tutorial demonstrates how to do simple pre-processing in Python.**

- a. Given that the original datasets comprise many columns, and our application requires only a selected few, we will tailor our dataset by removing the columns that are not needed.
- b. Once we have crafted our custom dataset, we will proceed to inspect the data, Handle missing or null values, imbalance values, eliminate the duplicate entries, and ensure data types are appropriate.
- c. Modify or remove some outliers; Normalize and standardize numerical data if necessary; Encoding categorical variables(to 0/1).
- d. Save the cleaned data, output as file for reusability.

8. **List of technologies you will use. You must use some kind of SQL database**

- a. Database(MySQL,PostgreSQL...)
- b. Back/Front-End(React, NodeJS,ExpressJs...)
- c. VS code, GIT,

9. Guidance for setting up a MySQL database (example can be found in WebDB Assignment posted on Canvas). Some groups in the past have had issues with MySQL, but Oracle is another option.

1. Download MySQL installer (MySQL web page or using Homebrew)
2. Access MySQL server, set up MySQL root password
3. Download MySQL GUI(MySQL Workbench, DataGrip)
4. May need download install driver for MySQL
5. Create database locally or cloud through AWS(RDS)
6. Connect to existing Data source use preferred MySQL GUI, Example:

Host: oracleimdb.cmkp3uvpymor.us-east-2.rds.amazonaws.com
SID: ORCL
User: student
Password: weakpassword
Port: 3030

10. Description of what each group member will be responsible for.

- a. Location mapping and location analysis: Yu Feng, Ruimin Yin
- b. Front End Design: Jiayi Wang, Luca Wu
- c. Algorithm decision-Yu Feng, Ruimin Yin
- d. Public school analysis: Jiayi Wang, Luca Wu