Stage 1:

Detailed Project Description

Tianshun Gao, Alex Ostrowski, Alex Capps, Jacob Stolker

1. Describe what data is stored in the database. (Where is the data from, and what attributes and information would be stored?)

The data being stored in our database from the Kaggle “Crime in Los Angeles Data from 2020 to Present '' is the crimes committed in LA with their division of records number, date occurred, time occurred (in 24 hour military time), the name of the area where the crime was committed, a code that represents the area where the crime was committed, a crime code that indicates what crime was committed, and a crime code description. The data that is taken from the “Crime in Los Angeles Data from 2020 to Present '' is the data transcribed from the original crime reports that are typed on paper.

The division of Records Number will be a number made up of a 2 digit year, area ID, and 5 digits. The areas will be based off of the 21 community police stations in LA and will be coming from the Kaggle area database itself. The areas will also be labeled sequentially from 1-21. Each of these areas will have a specific name designation that references a landmark or the surrounding community which will also be taken from the Kaggle database. The date occurred will look like MM/DD/YYYY and the time occurred will look like XXXX and both of these will also be taken from the Kaggle database. There will also be a crime code that is taken from the Kaggle database ranging from 110-956 and each of these codes will have a description associated with them. An example description of a crime description would look like this: “VANDALISM - MISDEMEANOR ($399 OR UNDER)”.

2. What are the basic functions of your web application? (What can users of this website do? Which simple and complex features are there?)

Can search for crimes that occurred in Los Angeles by:
Specific time and date or date ranges. Specific location (accurate only to the nearest hundred block.) Type of crime. A combination of the previous traits.

You would get the type(s) of the crime(s) committed, the location name, and the date/time. It would make most sense for the output of the data to differ slightly depending on what the user is searching for. If searching for crime of a certain type, data would be grouped first by location, then time. If searching for crimes at a certain location, data would be grouped first by category,

then time. If searching for crime at a certain date/time, data would be grouped first by location, then type. However, this visualization is more meaningful with the creative component described below in #3.

3. What would be a good creative component (function) that can improve the functionality of your application? (What is something cool that you want to include? How are you planning to achieve it?)

A creative component would be map visualization. Basically it's like a google map with crime information on it. So users can click on different attributes to see the patterns of crime that happen in certain areas in Los Angeles. Right now, we propose that crime could be viewed on the map in terms of date, time, sub-area, or crime code. For example, if a user input a specific date, then all crimes show up on the map would be the ones that occur on that date. Same goes for other attributes. A user can also apply multiple attributes to see what crime fits the condition. We are planning to achieve this probably by using some existing libraries like geopandas in python.

4. Project Title

LA Crime Visualization

5. Project Summary: It should be a 1-2 paragraph description of what your project is.

Our project allows users to interact with the "Crime in Los Angeles Data from 2020 to Present" database in a meaningful way. With the data we have stored and our UI, users can search for crimes that occurred in Los Angeles by time and date, location (accurate only to the nearest hundred block for privacy,) type of crime, or some logical combination of the previous traits.

Our interface can provide a text-based output based on the input criteria that is sorted logically based on the inputs it has been given. More importantly, however, our interface can show a visual representation of the text-based output that will have more meaning to the user. The program will output a map of the LA area in which the crimes are distributed based on location, with a key that makes them distinguishable by density, type, time, or all three.

6. **Description** of an application of your choice. State as clearly as possible what you want to do. What problem do you want to solve, etc.?

One application could be that a researcher wants to find all the robberies that happened between January 7th and January 10th in 2021.

Another application would be to use the data to determine where to deploy more surveillance systems. Areas with higher crime rates should receive more cameras to catch/deter criminals.

7. **Usefulness**. Explain as clearly as possible why your chosen application is useful. Make sure to answer the following questions: Are there any similar websites/applications out there? If so, what are they, and how is yours different?

Our chosen application is useful because it can pinpoint a list of robberies between two specific dates in a specific year for analytical use. It can also visualize where those robberies happened, and at what times.

There are websites somewhat similar, such as https://know-your-community-2-lahub.hub.arcgis.com/pages/public-safety , but this only gives information on the relative safety of one location. You cannot input anything other than a location.

8. **Realness**. Describe what your data is and where you will get it.

Our data is essentially a very broad overview of crimes that are reported in the Los Angeles area. We will include the type of the crime that happened, when it occurred (date and time), and where it occurred. The data will also date back to the year 2020 and is present up until today.

Our data is from the Kaggle database “Crime in Los Angeles Data from 2020 to Present''. That data is taken from the official Public Safety reports and is identical to the Kaggle database. The data is transcribed from the original crime reports from the City of Los Angeles dating back to 2020. Since the data is transcribed from the original crime reports which are typed on paper, there is opportunity for some errors or inaccuracies but the data should still be accurate and correct most of the time. Some location fields will also have missing data due to privacy and other reasons.

9. Description of the **functionality** that your website offers. This is where you talk about what the website delivers. Talk about how a user would interact with the application (i.e., things that one could create, delete, update, or search for). Read the requirements for stages 4 and 5 to see what other functionalities you want to provide to the users. You should include:

A user should be able to search for crimes that occurred in Los Angeles by time and date, location (accurate only to the nearest hundred block), type of crime, or a combination of the previous traits, and get a text-based list of data organized logically based on the queries, as well as a map visualization of the data.
Examples of information a user can search for:

Crimes committed in June in Sub-area 0377
Crimes committed on December 12th from 3:00 - 5:00 PM.
Crimes under the category BATTERY - SIMPLE ASSAULT in September.

A user should also be able to add or delete the record of a crime from a table of information they have generated and have this reflected in the results as well as the map.

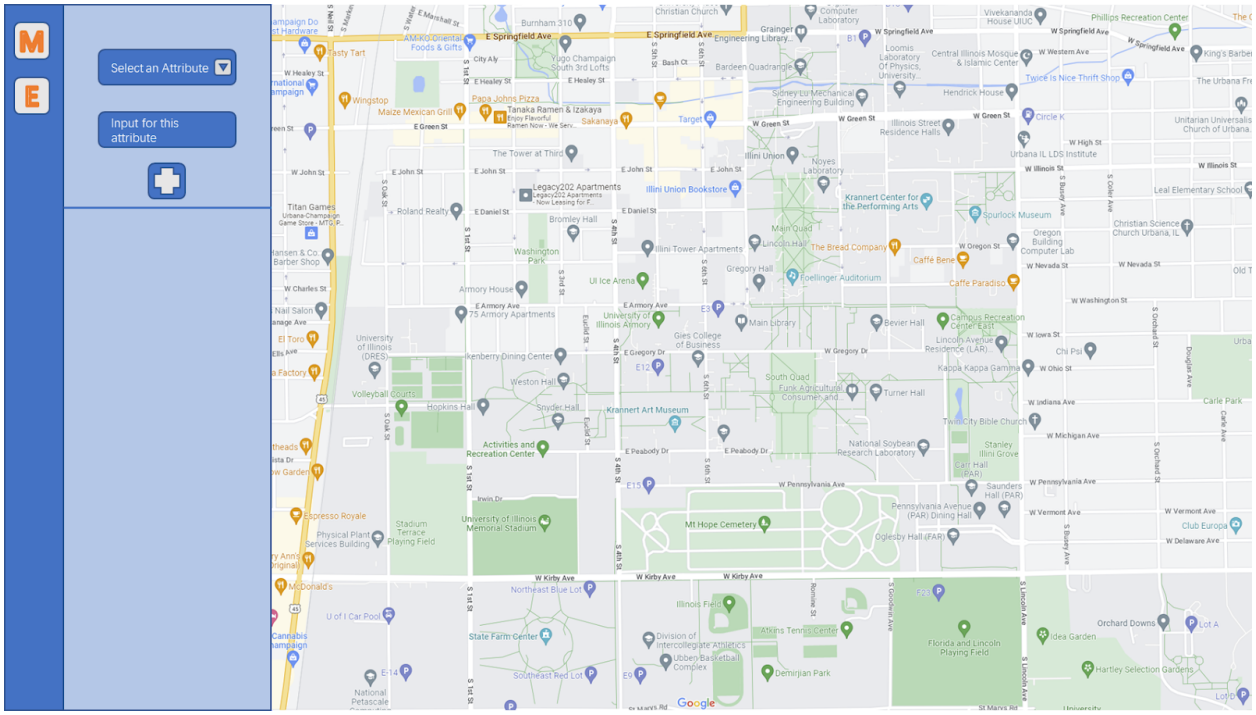
To ADD a record, user should have to specify the date (e.g. 01/08/2020 12:00:00 AM), the 24-hr military time (e.g. 2230), the Report District Number (e.g. 0377) whose first two digits provide info on the area code and name, the number of crimes committed (cannot be 0), and the crime code (e.g. 624) or the crime description (e.g. BATTERY - SIMPLE ASSAULT.) To DELETE a record, user

must specify the record by the above criteria, the division of records number, or some other variable (like location, date, time, etc.)

There is a UI interface that shows the map visualization as well as the menus for editing and pulling info from the database.

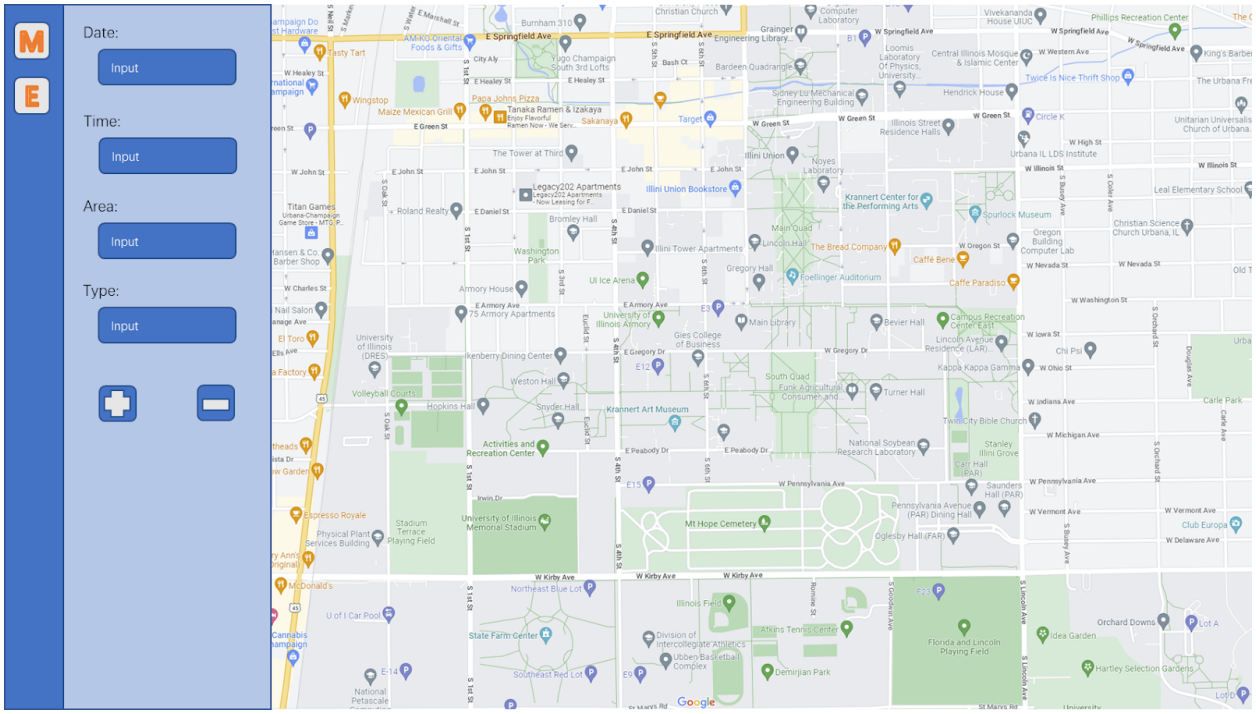
a. **A low-fidelity UI mockup**: What do you imagine your final application's interface might look like? A PowerPoint slide or a pencil sketch on a piece of paper works!

The map on the right is only for demonstration!!!



This is the UI for Map visualization by clicking "M" (Map).

A user can add an attribute by clicking the plus button.



This is the UI for adding and deleting records by clicking “E” (Edit).

After a user inputs all data, he can either add this record by pressing the plus button or delete this record on the map by pressing the minus button.

b. **Project work distribution**: Who would be responsible for each of the tasks or subtasks?
List of the person responsible for which exact functionalities in section 9. Explain how backend systems will be distributed across members. Be as specific as possible as this could be part of the final peer evaluation metrics.

Tianshun G

- Responsible for data preprocessing and communication between the server side and client sides.

Alex O

- Responsible for developing the UI/Front end
- Will work on the map visualization feature

Alex C

- Responsible for developing search queries specified by different combinations of: Date, time, location, type of crime. This will include ensuring that the organization of the search output is displayed in a logical manner.
- Will also assist with implementing the map visualization feature,

Jacob S

- Responsible for setting up the database and importing the records
- Will assist Alex O with UI design