

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

Our first section of data (which represents the majority of information stored in our database) pertains to crime in LA from 2020 to the present. The data is from an online dataset found on [Kaggle](#) that was listed in the TA-proposed datasets. This data is from LA's official police department, the LAPD. For each crime listed in the dataset, we will uniquely identify it with the crime's official, unique file number in the LAPD's system. Additionally, we will store information such as the crime's date, time, location, and type (what specific crime was actually committed), the victim's age and gender, and what weapon was used.

Additionally, we will be storing information regarding schools in LA. This data is also from an [online dataset](#) compiled by the city of LA. From this dataset, we will store the school's name, the type of school (elementary, middle, high, charter, etc.), and the location of the school. Having this location will allow us to determine, from the locations of the crimes in the previous dataset, which crimes would impact the overall safety of a school. Each school will be uniquely identified by an ID that is provided in the dataset.

We will also be storing rating information for said schools in LA. This data will come from the application's users – they will be able to rate schools based on their overall quality on a scale of 1-10, and the average rating for each school will be displayed alongside its other information once searched for by a user. For each school in the database, we will store the sum of all the 1-10 scores it has received from users along with the number of people who have rated the school. This way, when a user searches for a school, we can divide these two numbers and show an average rating for the school (along with how many ratings it has).

Lastly, we will have data in the database pertaining to login information. This data will be derived from the user when they create an account on our site. We will store usernames and passwords and reference these when users attempt to log in. Furthermore, for each user in the database, we will store their other account information (such as their name and what school their child attends), their past ratings for schools (so that they can adjust or delete past ratings as they see fit), and their saved searches (this is further expanded on in the next section, but users will be able to save certain schools as “favorites” so that they can quickly reference them upon login without having to repeatedly search for them).

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

The basic functions of our web application are as follows. Firstly, users will be able to log in to their account or register for a new account. Once they have logged in, they will have the

option to review their settings. These settings include updating their name and what school their child attends. Additionally, users will be able to review their ratings of schools. One of the features of our application, further discussed in one of the following paragraphs, is that users can rate schools. From their settings, users will be able to see these past ratings and adjust them as they see fit – either editing the rating itself, or deleting the rating as a whole. The final part of the user settings is that users will be able to review their saved schools instead of having to repeatedly search for them. Once again, this is discussed further in one of the proceeding paragraphs.

Our main page on the website will provide users with the ability to search for different schools in LA. We will initially have a list of every school present on the page. Users will have the option to sort these by overall user rating, safety index, or name (all in ascending or descending order). There will also be a search bar for users to interact with, where they can search for a specific school's name. Once a user clicks on a specific school, they will be taken to a new page where they will be presented with various information about said school.

This information will include the school's name, type, and location (all of which are provided by our online dataset from the city of LA). We will also present users with the school's rating based on overall quality, which is on a scale of 1-10 and is an average of all the ratings for the school provided by users of the application. The school's safety index will also be present, which is generated by our website. This takes into account all crimes present in the LAPD's dataset that occurred close to the school and adds special emphasis in calculation to crimes that involved school-aged children. Finally, users will also be presented with a breakdown of the crimes committed around the school, including the total number of crimes, the percentage that involved school-age children, the percentage that involved certain weapons, and the percentage that were violent versus non-violent.

On a school's page, users will also have the ability to star it, thereby saving it in their settings so that they can quickly reference the school upon login as opposed to needing to repeatedly search for it every time (if the school being viewed is one that the user wants to keep checking). Moreover, users will be able to give schools a rating from 1-10 based on its overall quality, which will contribute to the school's rating that all users will be able to see. We also aim to include a map on the page that visualizes exactly where crimes were committed close to the school.

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 that we believe would improve the functionality of “SafeSchools” is including a map on each school’s page. This map would demarcate the district the school is located in and plot the crimes that occurred around the school to create a visualization for users that highlights where most crimes have been perpetrated. The map would also identify the district’s local police station to show parents where the police are in relation to the school. This would allow parents using “SafeSchools” to easily tell which schools have lower crime rates and also to plan out safe routes for their kids to commute to school.

We don’t have all of the information or research done yet regarding this map, but we could potentially leverage a Google Maps API to highlight crimes on a map since the longitude and latitude for each crime is given.

4. Project Title

Our project title is “SafeSchools”.

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

Our project is a web application that stores and displays information about schools in Los Angeles (LA), California. Using our website, users will be able to search for schools in LA and gain valuable insights about them. This information will include a general rating of the school’s overall academic quality (provided by other users of the website) on a scale of 1-10, a safety rating (generated by our website) that takes into account crimes that occurred near the school and also specially considers crimes that involved school-aged children, a breakdown of said crimes (displaying the total number of crimes that have occurred there since 2020, the percentage of crimes that were violent, the percentage of crimes that involved school-aged children, and the percentage of crimes that involved guns), and finally, general information about the school such as its name, address, and local police station. One of our stretch goals is also including a map for each school that pinpoints the coordinates of different crimes that occurred in the area. This will let parents very clearly visualize the most common crimes that are perpetrated near a given school and how close they were to said school.

“SafeSchools” will also include basic account functions. Users will be able to register to use the website by creating a unique username and password. They can then use these same credentials to log back in. Once a user is logged into their account, they will be able to save certain schools so that they can quickly reference them again later as opposed to having to repeatedly search for them. Logged in users can also leave ratings for the general quality of schools which will be reflected when other users view the same school. Users will also be able to view their past school ratings and adjust them (editing or deleting) as they see fit.

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 of the biggest applications that we envision for “SafeSchools” is helping parents with younger children make the best decision possible about what school their kids should attend in LA. When moving cities / relocating, or just simply trying to find a new school for your child, one of the biggest concerns that parents have is overall safety. Especially given that LA is such a large city, children’s safety is of the utmost importance when it comes to schooling. Barring scouring online media platforms to try and discern information about the general safety of a school (that is often biased or skewed purely based on which individuals decide to post online), parents have no concrete way to determine how safe a school is.

Through “SafeSchools”, we aim to solve this. By using comprehensive data from the LAPD, we will be able to inform parents about the true safety of schools in LA. With an intuitive user interface, parents will be able to judge schools based on their safety and overall rating based on quality to find the school that is best suited for their children. Given search and save functionalities, parents can search through thousands of LA schools and save the ones that they are considering to quickly reference and compare them. Ultimately, we hope to provide a new way to give parents the most informed decision making as possible about schools in LA.

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?

We believe that “SafeSchools” would be a very useful application because of its relevancy and target audience. Parents are always searching for new tools that they can use to ensure the safety of their children, and “SafeSchools” presents a pertinent solution. While our website would be limited to analyzing schools in LA given the limitations on our dataset, this tool would ultimately prove invaluable for parents trying to research schools in LA.

There do exist web applications that are somewhat similar to what we are proposing for “SafeSchools”. However, these websites take schools into account and offer information relating to academic quality. Greatschools.org is an example of this – while it offers similar functionalities in terms of account login, starring schools, and search features, it offers no information about crimes committed near the school that would affect the overall safety of the institution. As a whole, “SafeSchools” is more centered around discussing the safety of schools, and no other application publicly available online seems to have the same centralized focus with data as relevant as LAPD reports.

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

We have obtained our first dataset, titled “Crime in Los Angeles Data from 2020 to Present”, from the Los Angeles Police Department, and are able to access it using Kaggle. According to the Kaggle page for this dataset, the data was copied into a database from the original, paper crime report file, and as a result, some of the data is inaccurate. In addition, missing location fields are marked as (0°, 0°) and address fields are to the nearest hundred block for privacy reasons. The dataset consists of 28 different columns. However, for our project, we will be concentrating on the following columns: Dr_No (unique ID), Date Occurred, Time Occurred, Area, Area Name, Report District Number, Crime Code, Crime Code Description, Victim Age, Weapon Used, Location, Latitude, and Longitude. All of the information described above is contained as a CSV file named “Crime_Data_from_2020_to_Present.csv” and has a size of 81.51 MB. To get direct access to the data, we will be downloading this CSV file from the Kaggle website onto our local machines for further use.

Our second dataset, titled “Schools (LAUSD)”, is obtained from geohub.lacity.org and is provided by the city of LA. The dataset was most recently updated on August 30, 2023 and has a file size of 101.6 KB. There is no definitive mention of inaccuracies or any missing information in the dataset. The dataset consists of 19 columns, but for the purposes of our application, we will be using: Id, Address, City, ZIP, Mpd_Desc (a description of what type of school it is), and Fullname. All of this information can be downloaded onto our local machines from the website as a CSV file.

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 stage 4 to see what other functionalities you want to provide to the users.

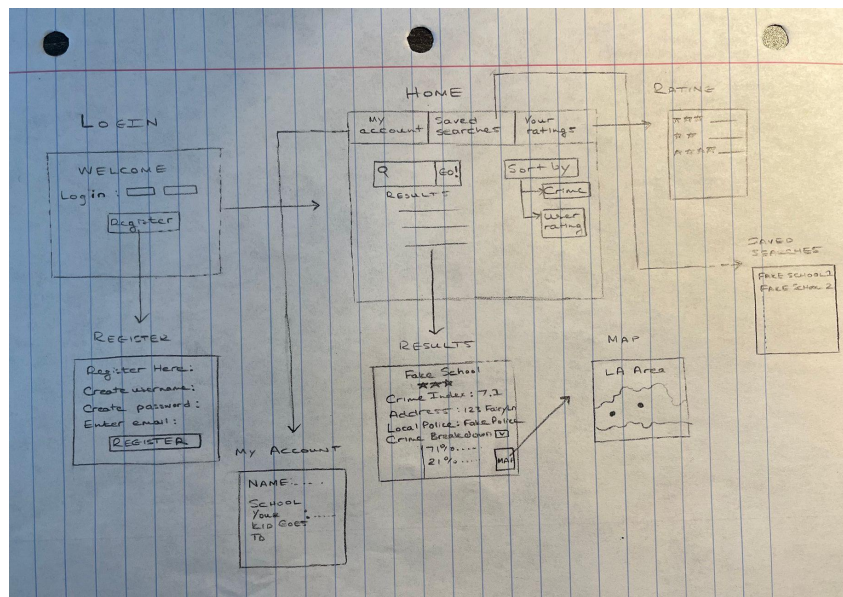
The user begins at the login page. Here, the user can log in with their information, which would involve searching the database. Alternatively, they can register to create an account. This would allow them to create new entries in our database with their new credentials. Once logged in, users are taken to the home page, which contains four key components of our application: the “My Account” page, the “Saved Searches” page, the “Your Ratings” page, and the “Home” page with search functionality.

Our “Home” page will offer the users a search bar to search for different schools in LA. This would involve searching through the database to produce relevant results. Furthermore, users can sort schools by safety index or user rating which would involve ordering search results from the database. Once a user clicks on a school, they will be redirected to a new page with information about the school which would involve searching the database once again. On this page, we have included a map component that highlights the district the school is in and crimes

committed in the area – this component is a creative extension which we may or may not be able to implement. We will also offer users the option to star the school for future reference or add a rating to the school. When starring the school, we would insert data into the database indicating that a user has saved this school. Additionally, when rating the school, we would update the database to add the user’s rating to the school’s already existing rating.

The “My Account” page will offer basic database searching to show the user their current account information along with updating database information if the user wishes to change their information. Furthermore, the “Saved Searches” page would also use database searching to query information about the user’s saved schools. If a user opts to un-star a school, this would require updating / deleting information from the database. Finally, our “Your Ratings” page will act similarly to the “Saved Searches” – we search the database to get information about the user’s ratings and offer delete and updating functionalities if the user wishes to update or delete their ratings.

1. **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!



2. **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 6. 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.

For Stage 2, Eeshan and Dhruv will be responsible for the database design. This includes creating the ER (Entity Relationship) diagram, as well as specifying the relationships, in English, between the various components of our database. Charlene and Rhea will be responsible for the normalization of our database. Based on our team's strengths, we plan to normalize the database using 3NF (3rd Normal Form) in order to reduce the amount of redundancy as well as optimize the performance of our database. In addition, Charlene & Rhea will also use the ER diagram developed by Eeshan and Dhruv in order to define the database's relational schema.

For Stage 3, each of us will individually set up the MySQL cloud database in Google Cloud Platform (GCP). Afterwards, Eeshan and Dhruv will be responsible for creating the primary tables to store our application's core information using the DDL (Data Definition Language) to define each table's structure. Rhea and Charlene will populate each of these tables with data. As a group, we will develop and execute the advanced SQL queries for our application. We have yet to decide which of the following queries (join of multiple relations, set operations, aggregation via GROUP BY, and subqueries) we will use. Once our advanced queries have been decided, for each query, we will all work on measuring advance query performance via the EXPLAIN ANALYZE command, adding different indices to different attributes on the advanced queries, and performing analysis on our index design.

The frontend design work, involving creating the HTML and CSS for the Login, Register, Home, My Account, Saved Searches, Your Ratings, and School Result pages, will be done by Dhruv. Additionally, establishing the connection between the frontend and backend will be done by Dhruv. The crime breakdown analysis for each school, involving displaying the number of crimes and associated percentages discussed in Sections 1 & 8, will be completed by Dhruv as well. Finally, Dhruv will implement the account settings handler (when a user changes / updates their account settings).

Eeshan will assume responsibility for both the login functionality and the registration functionality. Additionally, on the main page, Eeshan will implement the search bar functionality and collaborate with Dhruv to ensure it displays and works as expected on the main page. Furthermore, should we opt to implement the map functionality, Eeshan will work on displaying the search results using the Google Maps API.

Charlene's responsibility is to handle user ratings for each school, and ensure that all ratings added update in the database properly. She will be working closely with Dhruv to ensure that the frontend and backend connection handles this functionality correctly. In addition, Charlene will also be responsible for querying the school information to display all schools on the "Home" page initially and storing information in the database relating to user's saved searches.

Rhea is responsible for developing a formula to calculate the crime index. This formula will take into account the number of crimes committed, the amount of violent crimes, and especially the number of crimes that involved school-aged children. This crime index calculation will be used directly in her implementation of sorting the search results by overall user rating, safety index, or name, where she will work closely with Eeshan. In addition, she is also responsible for saving each user's ratings in the "Your Ratings" page. This also entails ensuring that users have the ability to update and delete their ratings as they see fit.