**Introduction**

Growing up in India, my parents have always had a hard time letting me leave the house at night and I had a strict curfew at 8pm. Yet somehow at the age of 18, my parent allowed me to come to USA to pursue higher education, and they live in constant fear. I feel if there was a way of knowing whether the neighborhood I'm about to enter is safe at that time, it would greatly assuage their fears.

So for this project, I would like to create a model to predict whether a neighborhood is safe at a particular time of the day. The location chosen for this project is Oakland, since I frequently travel to Oakland and use the BART. I would like to use a prediction model, but haven't yet decided on a particular algorithm.

This app is targeted towards anybody who needs to check whether the area they are about to enter is safe during that time slot, especially students travelling alone. Ideally, I would have liked to create a general model which can be applied to any city but the form of data collection differs and each dataset would require a unique approach to data cleaning.

Data

I will be using the Geopy package to obtain the location of the neighborhoods in Oakland. The information extracted will also be used to create maps (using Folium package). The data about the crimes committed will be taken from the <https://data.oaklandnet.com/Public-Safety/CrimeWatch-Maps-Past-90-Days/ym6k-rx7a> site. I will be using the csv file dataset and making additional calls to the SODA API if needed. This contains the information on the crimes that have occured in the past 90 days so the information is not outdated. The data contains the date and time of the crime, description, and the address of the location of the crime. I will be able to obtain the latitude and longitude using that address. The columns of the csv file are as follows:

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Description** | **Example** |
| CRIMETYPE | Plain text (a brief description of the crime) | ‘Weapon’, ‘Domestic Violence’ |
| DATETIME | The date and time of the crime (yyyy-mm-dd-hh-MM-ss-p) | 04/02/2020 07:42:00 AM |
| CASENUMBER | An ID given to the case (string) | 20-017349 |
| DESCRIPTION | Contains more details about the crime | CARRY CONCEALED WEAPON IN VEHICLE |
| POLICEBEAT | Indicates the ‘beat’ where the crime occurred (beat is the smallest geographic area) | 19X |
| ADDRESS | The street address of the crime | 1500 E 12TH ST |
| CITY | Oakland for all rows |  |
| STATE | CA for all rows(since the city is the same) |  |
| LOCATION | The address along with the latitude and longitude of the location in string format | 1500 E 12TH ST\nOakland, CA\n(37.787943, -122.94586) |