CSE-564: Visualization

PROJECT PROPOSAL

Analysis and Visualization of Crime Data in United States

ANKIT AGRAHARI (110946823)
Tarun Lohani (110921666)

Problem Statement:

The newly elected government wants to eradicate the crime in US. The government wants to appoint police force based on the number of criminal cases as well as based on the severity of the crime. We will develop an Interactive web visualization which will be helpful for the government to gain insight about the recent trends in crimes and take appropriate measures.

Objectives:

Our analysis will primarily be focussed to analyze the pattern of crimes in the country in recent times so that the police force could be used as efficiently as possible. We will visualize the regions in United States where the crime rates are high along with the type of crime so that the appropriate police staff could be appointed accordingly. We intent to evaluate the following:

1) Figure out the regions with low, medium and high crime rates densities.
2) Filter different regions according to the type of crime.
3) Draw different plots for different regions with the frequency of different kind of crimes occurred in that region.
4) Seek some correlation or pattern among different areas and type of crimes.

Literature Review:

The reported U.S. violent crime rate includes:

- Murder
- Rapes and sexual assault
- Robbery
- Aggravated assault

Crime rates are necessarily altered by averaging neighborhood higher or lower local rates over a larger population which includes the entire city. Having small pockets of dense crime may lower a city's average crime rate. US violent crime rates includes:

- Homicide
- Gun Violence
- Property Crime
- Crime Against Children

Crime rates vary in the United States depending on the type of community, within metropolitan statistical areas both violent and property crime rates are higher than the national average; in cities located outside metropolitan areas, violent crime was lower than the national average, while property crime was higher for rural areas, both property and violent crime rates were lower than the national average.

Dataset:

We have the following datasets available with us:

1. uci.fbi.gov data set: We have fbi crime data of recent years (2012 - 2016) for all the cities of United States.
2. There are around 10000 rows and 14 columns in the data.
3. Important Categories of Crimes recorded in dataset are:
    - Population
    - Violent Crime
    - Murder and nonnegligent manslaughter
    - Rapes and sexual assault
    - Robbery
    - Aggravated assault
    - Property crime
    - Burglary
    - Larceny theft
    - Motor vehicle theft
    - Arson

All datasets have only raw fields with none/few of them being an interpreted value.
Also, many columns have a set type of values they accept and can be interpreted accordingly.

Analysis:

We have crime data along with the type of crime for all the cities. To get a meaningful visualization from this data, first we will perform following operations with the data:

- Transform the data and find correlation among the rows and columns.
- Clean the data to remove inconsistency and redundant data.
- Transform the rows from city wise to state wise by adding up the crime data of all the cities in a state to get a state level picture in the visualization.

Once having both the city wise and state wise crime data, we will start our analysis with

- Putting different weights to different crime types which will denote the severity of that crime.
- Figuring out the correlation between different type of crimes and areas.
- Compare the crimes among cities and states by using different visualization techniques.

Goals\Deliverables:

Our final goal will be to deliver the following -

- Show the low, medium and high crime density areas on the US map.
- Provide two modes of analysis - city-wise and state-wise.
- Provide user interaction so that they can set the severity of different type of crimes.
- Provide user interaction to increase and decrease a particular type of crime to filter the areas on the map.
- Draw plots and graphs to compare and correlate among interesting features of the crime data.

Methodology:

- We will use python for data cleaning and processing.
- Server: We will run server on localhost through python Flask framework.
- Client: Our Client side implementation will be on javascript, html and d3 libraries.
- We will use different python libraries for our data analysis

References:

Dataset References:

- https://ucr.fbi.gov/

Literature Reviews:

- https://en.wikipedia.org/wiki/Crime_in_the_United_States#International_comparison
- https://crime-data-explorer.fr.cloud.gov/