Visualization of Crime Statistics for US

PROJECT PROPOSAL

Submitted by

Kaustubh Shrivastava (u1524432@utah.edu) - u1524432
Ashfaq Pasha (u1526443@utah.edu) - u1526443
Kaartheek Gottimukkula (u1526665@utah.edu) - u1526665
Sai Vignesh Alluri (u1527684@utah.edu) - u1527684

for

Visualization for Data Science (CS-5630)

September 2024

Table of Contents

1. Background and Motivation

Crime statistics in the US are collected from various sources, including the FBI and local law enforcement agencies, and cover a range of data such as crime rates, types of offenses, and geographical patterns. Traditionally presented in raw formats like spreadsheets, this data can be challenging to interpret. Visualizing crime statistics through charts, graphs, and maps simplifies the information, making it more accessible and easier to understand. This transformation helps to reveal trends, patterns, and anomalies in crime data that are crucial for effective decision-making and public safety planning.

Visualizing crime statistics is motivated by the need to make complex data more understandable and actionable. Clear visual representations help identify trends and patterns, aiding law enforcement and policymakers in strategic planning and resource allocation. It also enhances public awareness, fosters community engagement, and supports advocacy for policy changes by presenting data in a more compelling and accessible manner.

2. Project Objectives

The primary objective of this project is to leverage crime data visualization to enhance the understanding of crime patterns and trends across the US. The key questions we aim to address through our visualizations include:

- Explore the most prevalent types of crimes across different regions and their varying frequencies.
- Analyze how crime rates change over time to identify long-term trends.
- Examine the differences in crime rates between urban and rural areas to uncover geographical patterns.
- Investigate the relationship between socio-economic factors and crime rates in various areas.
- Assess how seasonal changes and time-of-day variations impact crime rates.

3. Data

The data utilized for this project is sourced from the Corgis Dataset Project, which compiled information from the Unified Crime Reporting Statistics. This dataset was developed in collaboration with the U.S. Department of Justice and the Federal Bureau of Investigation.

Key columns in the dataset include state, year, data population, property burglary, property larceny, violent assault, and violent murder, among others.

State	Year	Data.Population	Data.Rates.Property.All	Data.Rates.Property.Burglary	Data.Rates.Property.Larceny	Data.Rates.Property.Motor	Data.Rates.Violent.All	Data.Rates.Violent.Assault
Alabama	1960	3266740	1035.4	355.9	592.1	87.3	186.6	138.1
Alabama	1961	3302000	985.5	339.3	569.4	76.8	168.5	128.9
Alabama	1962	3358000	1067	349.1	634.5	83.4	157.3	119
Alabama	1963	3347000	1150.9	376.9	683.4	90.6	182.7	142.1
Alabama	1964	3407000	1358.7	466.6	784.1	108	213.1	163
Alabama	1965	3462000	1392.7	473.7	812.1	106.9	199.8	149.1
Alabama	1966	3517000	1528	527.5	869.6	131	230.3	177.7
Alabama	1967	3540000	1612.4	571.4	895	146	238.6	183.5
Alabama	1968	3566000	1766.6	628.2	967.7	170.7	232.4	168.5
Alabama	1969	3531000	1876.2	667.2	1037.8	171.2	250.4	181.7
Alabama	1970	3444165	2183.8	776.4	1184	223.5	295.7	215.2
Alabama	1971	3479000	2187	791.8	1173.9	221.2	311.4	219.7
Alabama	1972	3510000	2081.3	789.6	1096.7	195	313.2	211.7
Alabama	1973	3539000	2232.2	897.3	1107.8	227.2	350.1	236.3
Alabama	1974	3577000	2627.2	1057.9	1308.7	260.6	372.9	235.6
Alabama	1975	3614000	3079.6	1163.8	1645.5	270.3	392.9	233.5
Alabama	1976	3665000	3419.5	1170	1987.2	262.3	388.8	256
Alabama	1977	3690000	3298.2	1135.5	1881.9	280.7	414.4	278.3
Alabama	1978	3742000	3519.7	1229.3	1987.9	302.5	419.1	281.2
Alabama	1979	3769000	3830.5	1287.3	2223.2	320.1	413.3	263.1
Alabama	1980	3861466	4485.1	1526.7	2642.2	316.2	448.5	273.2
Alabama	1981	3916000	4428.3	1450.7	2693.3	284.2	470.5	306.1
Alabama	1982	3943000	4185.8	1256.2	2656.4	273.3	447.7	299.1
Alabama	1983	3959000	3685	1073.1	2381.4	230.5	416	284.9
Alabama	1984	3990000	3470.9	1001.8	2235.5	233.6	431.2	300.6
Alabama	1985	4021000	3484.6	1034.9	2191.2	258.5	457.5	315.5
Alabama	1986	4053000	3730.3	1159	2304.4	267	558	407.9
Alabama	1987	4083000	3892.2	1198.3	2431.1	262.8	559.2	409.9
Alabama	1988	4127000	4003.1	1233.8	2502.6	266.7	558.6	401.2
Alabama	1989	4118000	4037	1146.8	2592.8	297.5	590.8	415.7

The link for the database is here.

4. Data Processing

The data will be examined for invalid or missing values using the Pandas library in Python. Any discrepancies or empty values will be addressed to ensure the dataset is suitable for visualization. For instance, missing entries for property burglary will be replaced with zeroes to facilitate clear visualization. The CSV file will be imported using Pandas, which will then be employed to clean and manipulate the data effectively for analysis.

From the current data-set, segregation of the data based on the geographical area, timeline of the crimes and the various types of crimes committed will be the second step after the data cleansing is done.

From the subsets that are obtained from the dataset, visualization of the user inputted custom selection will become easier.

5. Visualization Design

The website design pattern will be as follows:
Navigation Bar:

- Home (Default landing page with key statistics and visuals)
- Explore by Crime Type (Allows users to filter and view data for specific types of crimes)
- Explore by State (Interactive map or dropdown to select a state)
- Yearly Trends (Graphs showing trends over time)
- About (Explanation of the dataset and methodology)

Main Visual Elements:

- Overview Banner: A succinct statement outlining the website's goal and the dataset's source. a background picture of the USA map or graphics depicting crime.
- Main Graphs Section: The visuals have to be in the main section. You can categorize graphs into tabs or sections such as Overview of Crime Type, Geographic Distribution, and Trends Over Time to make them easier to use.
- Overview of Crime Types: Pie and bar charts that show the total number of various crimes (e.g., violent crimes, property crimes).
- Geographic Distribution: State-by-state crime statistics are displayed on a USA Choropleth map.
- Trends Over Time: This section displays line graphs that illustrate changes in crime rates over time for either the entire nation or a subset of states.

Graph Types for Different Insights:

- Bar Diagram of Crime Categories. From 2000 to 2020, crime patterns are seen in this graph. It has a Y-axis for crime rates or totals and an X-axis for years. Different crime kinds (such as violent crime and robbery) are represented byocus on certain patterns, users can switch between crime types and click over data points to get exact statistics.
- Annual Crime Trends Line Graph. From 2000 to 2020, crime patterns are seen in this graph. The years are displayed on the X-axis, and crime rates or totals are shown on the Y-axis. Different forms of crimes, such robbery and violent crime, are represented by multiple lines. To view precise changes over time, users can swap between crime types and hover on specific values.
- State Crime Data Choropleth Map. We will create a USA map and color-code each state according to its property crime statistics for 2020 in order to graphically communicate crime data. An explanation will be included to explain the color gradient used on the map, which will range from light blue

(showing low crime rates) to dark red (indicating high crime rates). To improve user engagement and comprehension of the data, an interactive hover feature will show further details such as the year, crime rate, and name of the state.

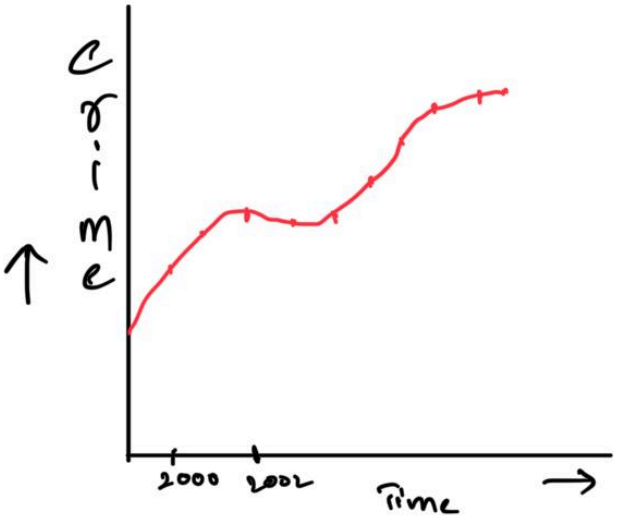
- Pie Chart for Crime Distribution. The percentage of each sort of crime for a given year, state, or the United States as a whole is visually represented by the pie chart. Users can quickly see the distribution of different crimes because each slice of the pie represents a particular criminal category. To better comprehend the frequency of various offenses in a certain state and year, readers might, for example, explore the breakdown of crime types in California for 2019.
- A bar chart with stacked bars to compare crimes across states. Comparing crime rates across several states is the goal. The design includes a stacked bar chart that shows the rates of crime by category of crime on the Y-axis and the states on the X-axis. To view the data, users can interactively choose particular years. A comparison of the crime rates in California, Texas, and New York, for instance, would show how the crime figures in each state have changed over time.

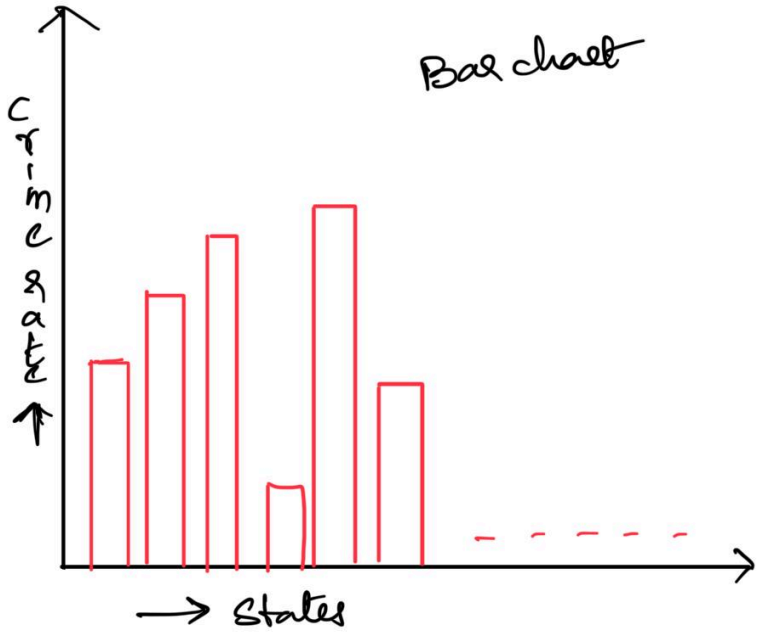
Interactive Elements

- Dropdowns/Filters: For more in-depth information, refine your search by year, state, and kind of offense.Permit users to choose between national and state-specific perspectives.
- Interactive Map: Clickable states on the map that open a pop-up window or separate section with comprehensive data or trends for that state.
- Hover/Tooltip Info: When a data point on a graph is hovered over, a tooltip with more specific information (such as the precise crime rate for that year or type of crime) should appear.

Image and Visual Enhancements

- Background Images: To keep users interested without being overly intrusive, utilize subdued images relating to crime, such as a faded picture of a map of the United States, law enforcement emblems, or cityscapes.
- Types of Crime Icons: In pertinent areas, use small images that symbolize certain crimes (e.g., a lock for burglary, an automobile for motor vehicle theft).
- Color Schemes: Choose a palette that highlights clarity and readability:
- Bar charts: Different colors for different categories of crimes.
- Choropleth map: A gradient representing a range of crime rates from low to high.


Crime
2000
2002
Time


Bar chart
Crime rate
States

6. Must-Have Features

Data Filtering

- Give consumers the option to filter data according to particular categories of crime (such as violent or property crimes).
- Year Range Selection: Give users the option to select a year or range of years for visualizations (2000–2020, for example).
- State Selection: Offer a state-by-state filter for crime data that is particular to a given area.

Visualizations

- Line Graph: Interactively hover to see data points for a certain year and see patterns in crime over time.
- Bar Chart: Show the total number of various crimes committed in a given year.
- Pie Chart: Show the percentage of various crimes in a state or country.
- Choropleth Map: Use a color-coded map to display crime statistics by state.
- piled high Bar Chart: Examine state-by-state crime rates along with a detailed description of each category of crime.

Interactivity

- Hover Tooltips: When a user hovers over a graph, the actual value of each data point is displayed.
- Toggle Options: Provide the option for users to alternate between several crime categories on the graphs (for example, from violent crime to robbery).
- Interactive State Map: Provide a clickable state map that displays comprehensive crime statistics for each state.

Data Summaries

- Crime Overview Cards: Display important data on the site, such as the top three states with the highest rates of crime.
- Give users the option to get the filtered data in a CSV file by selecting the "Download Data" option.

7. Optional Features

Advanced Data Analysis Tools

- Comparative Analysis: Permit visitors to view side-by-side comparisons of crime rates across various states or years.
- Tools for trend analysis and future crime rate forecasting based on past data should be made available.

Customizable Graphs

- Give users the option to alter the colors, line styles, and data point markers on the graph.
- Graphs that can be altered can be saved as pictures or PDFs for use in reports and presentations.

User Annotations

- Add Notes: Allow users to add their own observations or notes to graphs or maps.

8. Project Schedule

The project is organized over a 10-week timeline, with activities structured into two-week intervals to ensure steady progress.

Week 1: Project Planning and Tool Selection

- Define the project goals and objectives clearly.
- Choose appropriate development tools and visualization frameworks.
- Assign roles and responsibilities to team members.
- Initiate data collection and perform initial data cleaning tasks.

Week 2: Data Preparation and Task Allocation

- Continue with data cleaning and preparation processes.
- Refine and preprocess the dataset for analysis.
- Assign specific tasks and responsibilities to team members for streamlined workflow.

Week 3: Interactive Visualization Framework Setup

- Establish the framework for creating interactive visualizations.
- Develop initial prototypes for visualizing crime statistics.
- Conduct preliminary testing to ensure foundational functionality.

Week 4: Integration of Filters and Interactive Elements

- Integrate features such as date and state filters into the visualizations.
- Ensure smooth interaction between the filters and the visual data.
- Begin testing to identify and address any integration challenges.

Week 5: Development of Detailed Analysis Pages

- Start the development of pages for detailed crime statistics and analysis.
- Design and build the data viewer and summary pages.
- Perform ongoing testing to ensure the pages function as intended.

Week 6: Enhancement and Usability Testing

- Finalize the development of statistics pages.
- Enhance interactive features and improve user experience.
- Conduct comprehensive testing focusing on usability and functionality improvements.

Week 7: Refinements and Debugging

- Address and resolve any bugs or issues discovered during testing.
- Review and refine the user interface and design elements.
- Implement additional features if time allows.

Week 8: Final Testing and Quality Assurance

- Conduct thorough testing of all features and visualizations.
- Perform user acceptance testing to gather feedback.
- Optimize performance and make final design adjustments.

Week 9: Project Enhancements and Final Adjustments

- Apply improvements based on user feedback and testing results.
- Make final corrections and enhancements.
- Begin work on optional features if the schedule permits.

Week 10: Final Review and Documentation

- Complete all aspects of the project.
- Prepare comprehensive documentation, detailing data sources, methodologies, and project processes.
- Prepare for the official launch and presentation of the project.