

Team A: Taeheon Kim ([tkim293@emory.edu](mailto:tkim293@emory.edu)), Jae Ho Choi ([jcho427@emory.edu](mailto:jcho427@emory.edu))

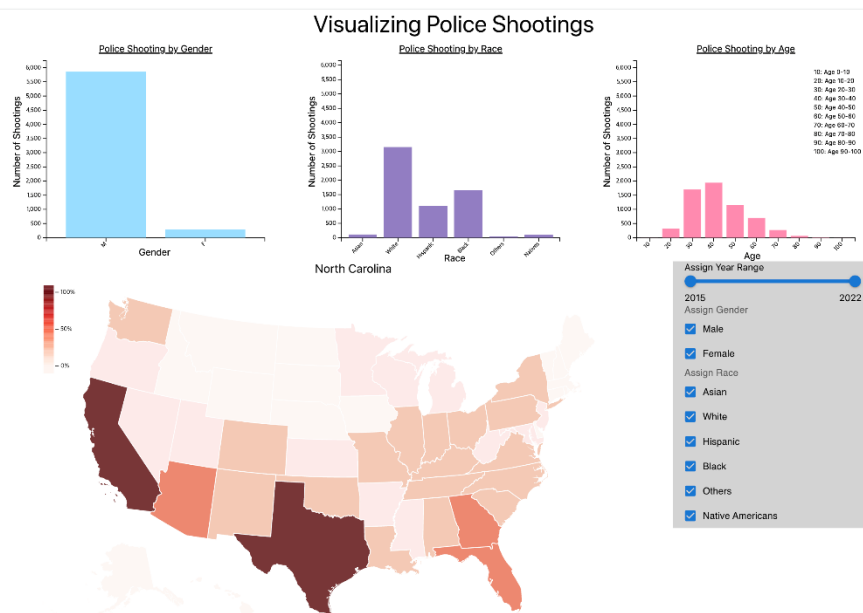
Yifeng Wu ([yifeng.wu@emory.edu](mailto:yifeng.wu@emory.edu))

## Final Deliverables

### Summary of Dataset Characteristics

The dataset that our group used in order to create our final data visualization was the “US Police Shootings from 2015- Sep 2022” from Kaggle (<https://www.kaggle.com/datasets/ramjasmaurya/us-police-shootings-from-20152022>). The dataset consists of incidents where people were killed by law enforcement in the United States during the time period from January 2015 to September 2022 and it has attributes related to the incidents such as (name of people killed by law enforcement, date, manner of death, the type of weapon the civilian was armed with, age, gender, race, city, and state). Prior to creating our data visualization, we first preprocessed the data to reduce the dataset dimensionality, get rid of missing values, and create attributes that would be useful for creating our data visualization. During our data preprocessing step, we started off by removing duplicate datasets and filling or dropping rows that had missing values. Moreover, in order to avoid bugs and to make the loading and creating process of our data visualization more efficient, we deleted a couple of attributes that we believed to be unnecessary and extracted new attributes that would be more useful. After data preprocessing the initial dataset, which originally had 17 attributes with 7,729 data points, was reduced to 15 attributes with 6,142 data points. Amongst the attributes from our data set, the ones that we mainly worked with for our data visualization were date, age, gender, race, state, and the count number for police shooting incidents.

### Images and Description of Our Visualization



Our final data visualization project has been successfully created and is currently hosted on the address:  
<https://638da6ea97124e5bad6382cd--beautiful-pie-a4edbf.netlify.app/>

The general layout as well as the default visual representation is demonstrated in the image above. Initially, when a user first sees our data visualization, they are presented with three bar charts that represent and compare the number of U.S. police shooting incidents for gender, race, and age. The three bar graphs were created based on the total number of U.S. police shooting incidents from the dataset (all incidents between (January 2015 to September 2022)). Below the bar graphs, we also have a heatmap of the U.S. that displays the number of police shooting incidents for each state. The color density was adjusted to a percentage scale and if the color was darker for a specific state, it means that more police shooting incidents had occurred for that state. Like the three bar graphs, the default map represents the total incidents. On the right side of the U.S. map, we have a functioning slider and buttons that allow the users to filter out the data based on what specific information the users would like to see. The slider allows the users to adjust the time frame and by checking/unchecking the gender or race categories, the users will be able to see more detailed information concerning police shooting incidents for the categories that they have chosen. When the filters are changed, the U.S. map will change according to the filters that have been applied by the users.



To see more detailed information concerning the U.S. police shooting incidents for a specific state, the users will be able to click on a specific state. Hovering over a specific state on the map will display the state name and once the state is clicked, the map will be zoomed in, and three bar graphs that are like the ones on the default page will be displayed. The three bar graphs will be detailed information concerning police shooting incidents for gender, race, and age for the specific state that the user had selected. The filtering functionality that had been selected will also be applied to the three bar charts.

### **Group Member Contributions**

In terms of contribution, everyone worked on the project altogether and so there was not any specific part that only a certain group member worked on separately. From the initial brainstorming process to the final creation of the data visualization, we all worked together and contributed equally. For every task concerning the project, we always got on Zoom to collaborate and worked together to constantly provide feedback to one another and help each other out whenever we faced difficulties with our project. For the creation of our final visualization, we initially discussed how we would be implementing the data preprocessing and cleaning and worked together to get our cleaned data. Moreover, for the coding process regarding our final visualization, we all worked together using GitHub to share our codes, and we constantly worked together through Zoom to learn and code using d3 and react. We worked together to develop and implement one functionality at a time and whenever we faced a bug, we helped each other out by looking up resources and debugging the problems.

### **Reflections**

Using our dataset, our initial goal was to create a visualization that portrays the frequency of U.S. police shooting incidents for each state and allows users to analyze whether there is a bias in police shooting incidents regarding various factors such as race, gender, and age. In regard to our initial goal, we believe that we have successfully created a visualization that covers all these factors. The heatmap allows users to easily see which states had more police shooting incidents compared to other states. Moreover, the bar graphs allow the users to see and compare the number of police shooting incidents for different categories (gender, race, and age) and see if there is any existing bias. Initially, we had planned to only create the U.S. heatmap for our project as we believed that it would allow users to easily see the frequency of police shooting incidents at once. However, after receiving feedback from our professor as well as our classmates, we realized that the map alone was not sufficient enough to deliver all the information that we intended to display with our visualization. Hence, we decided to include bar graphs in our visualization as the bar graphs for various categories will allow users to easily make comparisons and see if there is a bias or not. By incorporating bar graphs into our initial plan and displaying both the bar graphs as well as the U.S. map, we believe that our visualization displays all the information that we want to provide to the users. During our project, the biggest roadblock that we faced was the creation and implementation of our visualization. Since all of our group members did not have any prior knowledge or experience concerning d3, we had to look up various online resources and learn as we created our visualization. Besides learning how to utilize the d3 library, our group did not face any other substantial roadblocks. In terms of collaboration, our team did not have any issues as we synergized and worked well together. For each task concerning the project, we always divided up the work equally and everyone contributed to the project.