# CS526: Understanding Traffic Accidents Summary

**Group 3: Jaini Patel (jp1891), Moriah Walker(maw305), Arun Subbaiah(as3590)**

Our project focuses on providing interesting information and insights to our website's users in regards to fatal car accidents provided by datasets from the National Highway Transport Safety Authority. We cleared the data and manipulated it using the OLAP cube concept to minimize the size it took it up. This manipulated data was then stored in an RDMS and retrieved using PostgreSQL. The connection between the database and the front end was done using Python's Django.

For the front end on the homepage of our website, our users are first given a welcome and brief introduction as well as a first look at one of our plots, which is a 3D car model that shows information about the percentages of fatal accidents that occurred because of the collision's impact with that part of the car. In order to switch easily to the next page, the user will notice an adjustable summary tab on the left side with a list of buttons that link to all the other pages on the website.

One of these pages is the *Stats* page, there are a variety of plots that cover the distribution of fatal accidents in relation to some variable which by default includes a map of the US, a fatality count, a year slider, a line chart following the time of day, a bar chart for the different days of the week, and a stacked bar chart that shows the age distribution across each sex. Upon selecting a state from the map, the user is also provided a new map of the selected state next to the US map. This also impacts what data is used for all the plots, so all of the plots on this page are updated with each user interaction. This brushing effect takes place when the values selected on the year slider are changed and when a sex on the legend for the age distribution bar chart is selected/deselected.

To give further insight into the conditions surrounding a fatal accident, a user can select the *Factors* page. There are two important line charts on this page related to time. One being based on fatalities throughout the years and another based on specific dates. Changing the time range on the former plot affects what data is shown on the other plots on the page. The page contains one more time plot, which is the wind rose chart, which resembles placing bar plots on a clock in order to show the AM and PM distribution of fatal accidents. A similar wind rose chart is used to depict the likelihood of an accident being fatal if the collision occurs at different points of impact. The regular bar charts on this page compare the likelihood of fatalities when alcohol was involved in the accident vs when it was not and a comparison of how different light conditions impact the likelihood of accident fatalities. The likelihood of getting in a fatal car accident when different weather conditions are present was shown in a pie chart.

If the user wants more information about the car model and analytical information, they can check out the *Analytics* page, which provides a WordCloud based on the frequency of car models that are involved in fatal accidents. Below this WordCloud is a radial bar chart that allows the user to view the count of fatal car accidents that are alcohol related based on the year the user selects from a slider. At the bottom of the page, there is a pie-of-pie chart with the larger pie chart displaying information about the frequency of car models and the smaller pie chart displaying the model age distribution of the specific model selected. This pie-of-pie chart is accompanied by a checklist of car models a user can select to determine, which car models are displayed by the pie charts.