Group 19:

Khyathi Balusu

Pooja Dawada

Hannah Li

Rachel Meade

**Project Goals**

**Description**

Using a dataset of traffic collisions in the Los Angeles area between January 2010 and July 2019, we looked for patterns in when and where collisions occur, as well as the demographics of people involved in collisions.

**Importance of Problem**

Isolating the most dangerous traffic circumstances is the first step to making the streets of Los Angeles safer for drivers and pedestrians. The data indicates not only the location of each collision, but also the city council district and neighborhood council district for each incident. The inclusion of this information facilitates direct action from local government and municipalities responsible for the most dangerous areas. Los Angeles can use this information to better allocate resources to redesign or modify the most dangerous intersections in the city, and to better police dangerous areas that become dangerous and specific times.

Beyond helping to direct governmental action, this analysis would be helpful for drivers in the Los Angeles area. With the knowledge of when and where collisions occur frequently, they can better choose their routes, avoiding dangerous areas, intersections, or times if they so desire. Rideshare companies like Uber and Lyft, and taxi and car services could also make use of this analysis, programming their navigation systems to avoid high-frequency collision areas whenever possible. For companies which rely on driving to generate revenue, better information on what conditions increase likelihood of collision represents an opportunity to better protect their drivers, not only from dangers that might affect their wellbeing, but also from potential delays that could harm revenues or company reputation.

Although they make up a very small percentage of the dataset, we also have information on drivers who were involved in more than one collision. Along with the place and time collision information, insurance companies may be able to improve pricing policies for specific risk profiles.

**Exploratory Analysis**

The first goal was to uncover patterns in where collisions occur. At the broadest level, there are 12 Council Districts in Los Angeles, and the highest number of accidents occur in districts 12, 9, and 14. Within each Council District, we also found the Neighborhood Council responsible for the areas in which the highest number of collisions occur. For example, within council district 12, more collisions occur in the area which belongs to Neighborhood Council number 89 than any other Neighborhood Council district. Neighborhood 89 is the East Hollywood Neighborhood Council, which is led by president Arasele Torrez. This is precisely the kind of information that Los Angeles residents can use to demand better traffic management and policing.

Apart from the Council Districts and Neighborhood districts, the dataset also divides the area of Los Angeles into 21 distinct areas. Of these, the highest number of collisions occurred in the 77th Street area, followed by the Southwest, and Wilshire.

Although this broad information might be useful for city planners, it is unlikely that individual drivers will be able to avoid entire areas of the city at will. They may, however, have the power to avoid the most dangerous intersections. With this in mind, we found the intersections at which the highest numbers of collisions occurred. Some of these intersections had over 400 unique collisions – meaning that on average they had about one collision every week over the time frame captured in the dataset.

Another primary objective was to look for patterns in when collisions occur. The number of collisions appears to be increasing between 2010 to 2017. However, 2018 may indicate a more hopeful outlook for the city, as it had a slightly lower number of collisions compared to 2017. Between 2010 and 2018, there was a 26% increase in collisions, but during the same period, the city’s population only grew by an estimated 5% (World Population Review).

The highest number of collisions occurs on Friday, followed by Thursday and Saturday. The highest number of collisions occurred in the 5:00pm hour and was followed by the surrounding hours, between 3:00pm and 7:00pm. This is likely not a surprise to anyone who has been stuck in an afternoon rush hour before. The places where collisions occur most frequently change depending on the time of day. For example, between 11:00pm and 2:59am, more accidents occur in Hollywood. In the early mornings (3:00am to 6:59am) and later evenings (7:00pm to 10:59pm), more accidents occur in the 77th Street Area, and from 9:00am to 12:59pm, more accidents occur in West LA.

Finally, to meet our final objective, we identified demographic patterns among people involved in traffic collisions. Collisions involve men more frequently than women. In Los Angeles, people of Hispanic descent are involved in the highest number of accidents, followed by white people, then black people and people identified as “other.” More collisions involve people in their 20’s than any other age group. In fact, 28% of all collisions involve people in their 20’s. The next highest numbers of collisions involve people in their 30’s and 40’s.

Although there are distinct trends in the ages of people who are involved in collisions overall, the times at which people of different age demographics are involved in collisions are not uniform. For example, people in their teens and 20's are in proportionally more collisions between 9:00pm and 4:00am than they are overall. It is the opposite for people in their 50's: more collisions that involve this age demographic occur between 5:00am and 5:00pm. They are involved in proportionally fewer collisions between 6:00pm and 4:00am. Finally, people in their 70's are in proportionally more of the collisions between 9:00am and 4:00pm, and are in proportionally fewer of the collisions between 10:00pm and 5:00am.

Within the Los Angeles Police Department, it is standard procedure to use 9999 for certain values in a report when they are unknown or referred from an outside law enforcement agency (LAPD Online). Based on the disproportionately high occurrences of collisions involving people with 99 years of age, we believe that police officers or administrators may be occasionally applying the same guidance to age. If, when a person's age cannot be determined, police departments report age as 99, this could explain the relatively higher numbers of collisions.

We also found that of all the collisions over the eight-year period recorded, only 954 of them, or 0.2% involved the same DR Number. No DR Number was associated with more than 2 collisions.

**Solution and Insights**

**Sources**

<http://worldpopulationreview.com/us-cities/los-angeles-population/>

<http://www.lapdonline.org/lapd_manual/volume_5.htm>