Project1

R. Harmer

Data set: Traffic Crashes from the City of Chicago Data Portal. Covers crashes since 2015. Data set had 162,055 rows after initial cleaning.

Traffic Crashes - Crashes - Crash data shows information about each traffic crash on city streets within the City of Chicago limits and under the jurisdiction of Chicago Police Department (CPD).

Data from E-Crash are available for some police districts in 2015, but citywide data are not available until September 2017. About half of all crash reports, mostly minor crashes, are self-reported at the police district by the driver(s) involved and the other half are recorded at the scene by the police officer responding to the crash.

Many of the crash parameters, including street condition data, weather condition, and posted speed limits are recorded by the reporting officer based on best available information at the time, but many of these may disagree with posted information or other assessments on road conditions. A traffic crash within the city limits for which CPD is not the responding police agency, typically crashes on interstate highways, freeway ramps, and on local roads along the City boundary, are excluded from this dataset.

We chose the data set since it was a good size at 94MB and had about 162K rows. It also had a lot of fields we thought would provide interesting data for analysis such as Day of Week, Hour of Day, Type of Crash, Weather and Lighting conditions, Latitude, Longitude and Type of Crash.

First, we look at the days of the week and found weekdays were when the most crashes happened with Friday and Tuesday having the most. When combined with days of the week, we found not surprisingly that Rush hours were when the most accidents happened, especially during the afternoons.

Since we were using data from a city, we wanted to visualize the areas where the crashes happened and see if any patterns emerged. The all crashes for heatmap has 161,397 crashes which was all the rows that had Lat and Lng data. With that may crashes the map was very saturated but did show some less saturated areas in the South.

The Fatal crashes map had just 155 crashes which we attributed to the data just being for the city and did not include highways. There was a bit of clustering around Central Chicago and South Shore areas.

Wanted to investigate the relationship between Day of the week vs. Hour of the day on traffic incidents. Heatmap shows most accidents happen during weekday Rush hours especially in the afternoons with Friday having the most and Tuesday 2nd most. On weekends, there is some significant activity on Saturday afternoons. Also, there are more crashes on Friday and Saturday nights between 9 and 12 than any of the other nights.

Wanted to investigate the relationship between Weather vs. Lighting Conditions on frequency of crashes. The results show that most accidents occurred with Clear weather conditions and majority had lighting conditions classified as either Daylight or Darkness on a Lighted Road.

We next looked at the type of crashes and found the most prevalent types were Rear end, angle, side swipes and most surprisingly crashes with parked cars.

For the next project, I think a clearer division of duties and more planning would be beneficial. We did some exploring on our own and wound up with a bit of overlap.