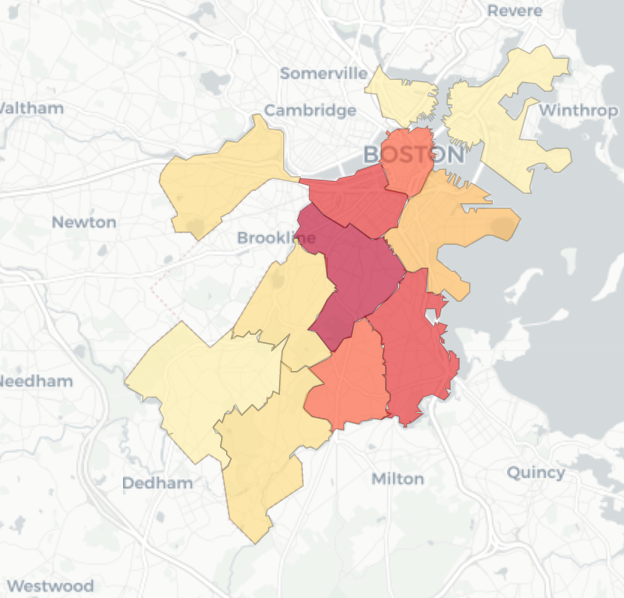
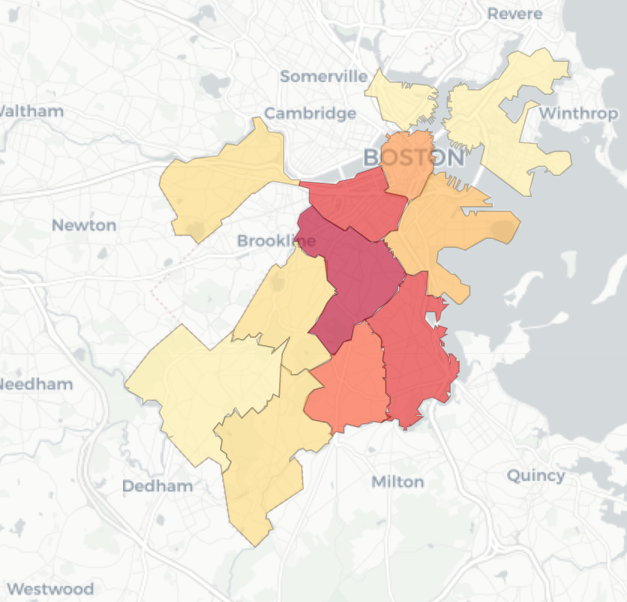
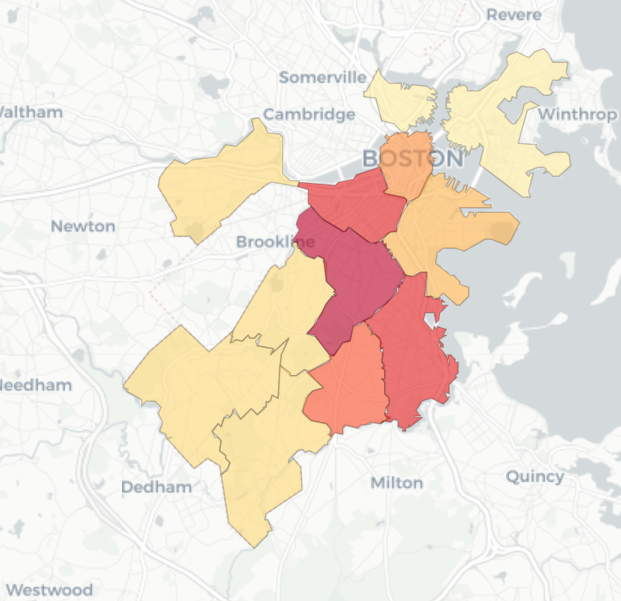
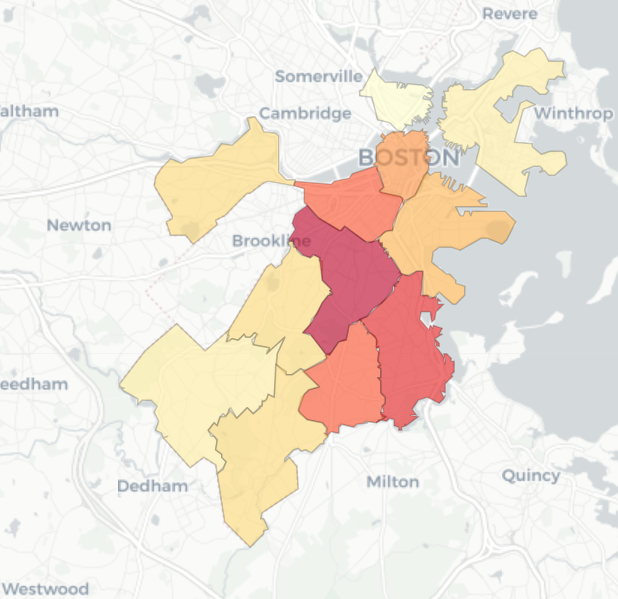
Hunter Boles

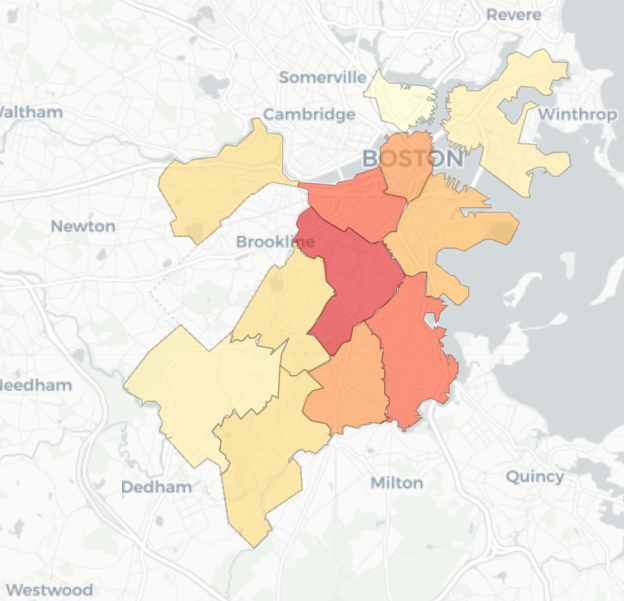
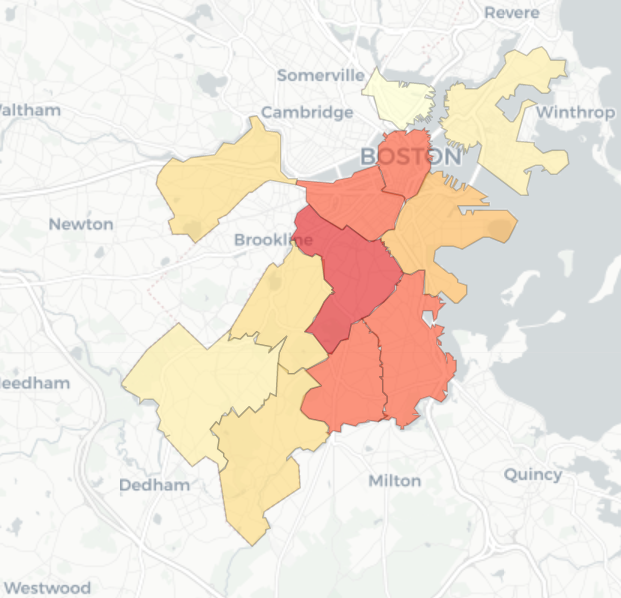
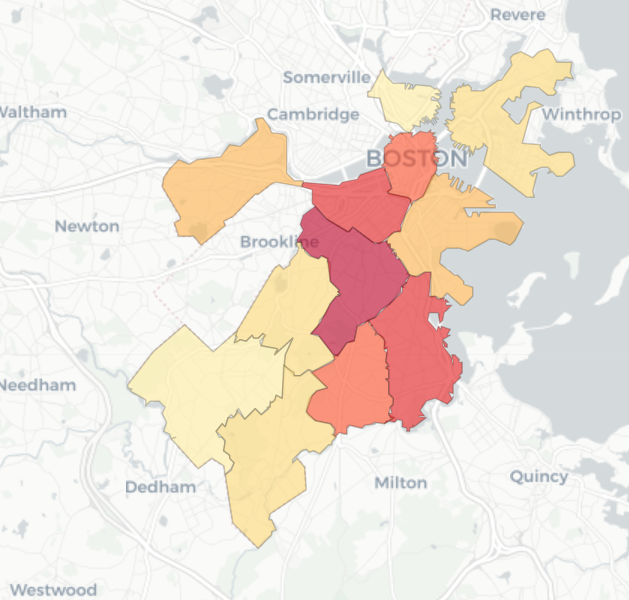
CSCI 444 – Data Visualization

Final Project

My hypothesis is that the day of the week, i.e. Monday, Tuesday, Wednesday, etc., influences what incidents occur.

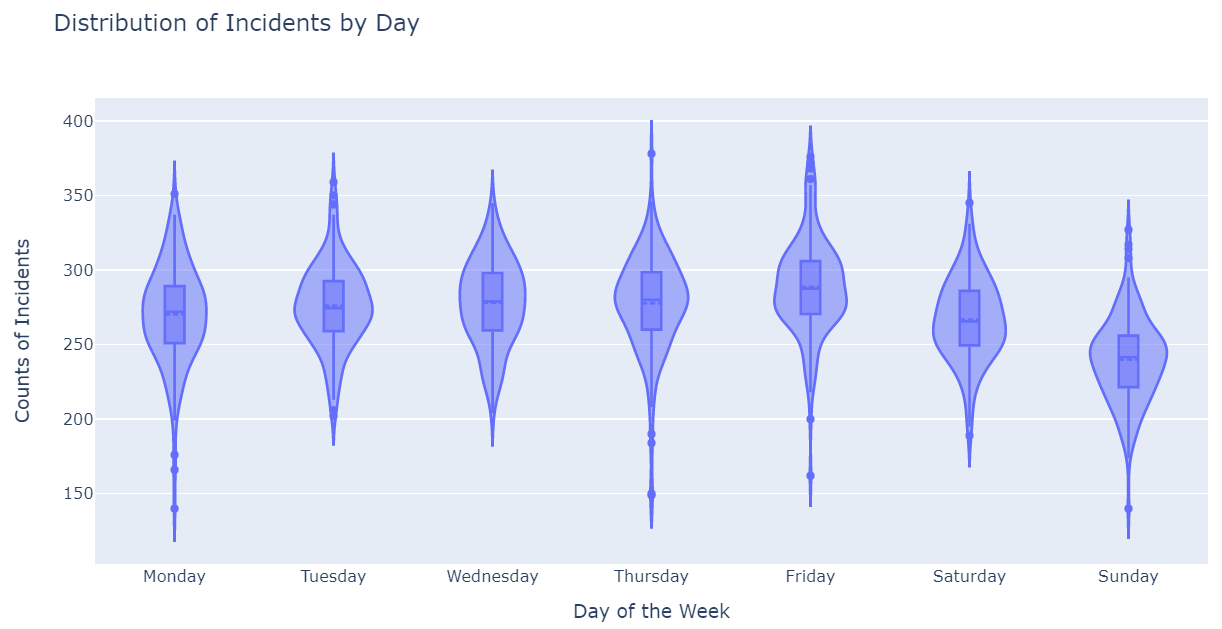
Figure 1: Choropleth maps for incident counts on each day of the week. (Monday, Tuesday, … Sunday)





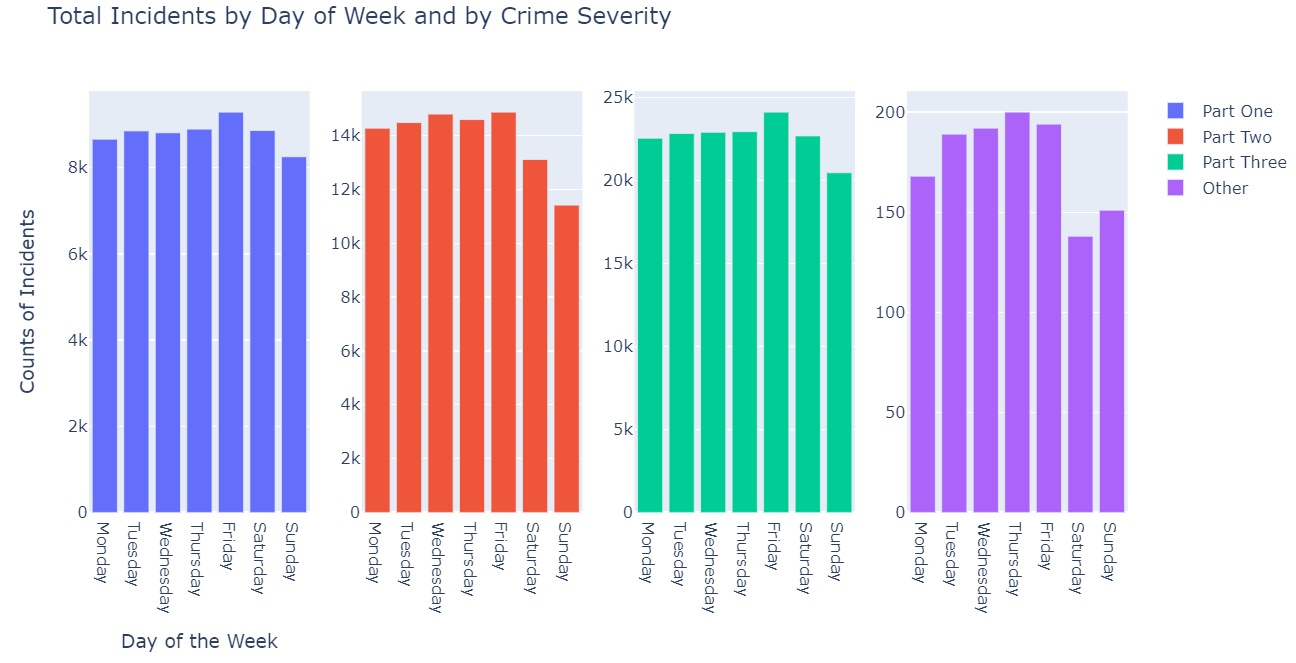
The first figure shows the different distributions of incident by district, by day. While there is mild variation in the number of incidents on each day, it is not enough to definitively prove that the day of the week has any effect on the police incidents that occur. For that reason, the plot disproves my hypothesis. It would have been beneficial to retrieve population data for the charts; however, the given densities are not in terms of the police districts. Merging the two GEOJSON files would be tedious, and out of the scope of this project.

Figure 2: Violin Plot for incident counts on each day of the week (Monday, Tuesday, … Sunday). Calculated across the entire 2.5-year span.



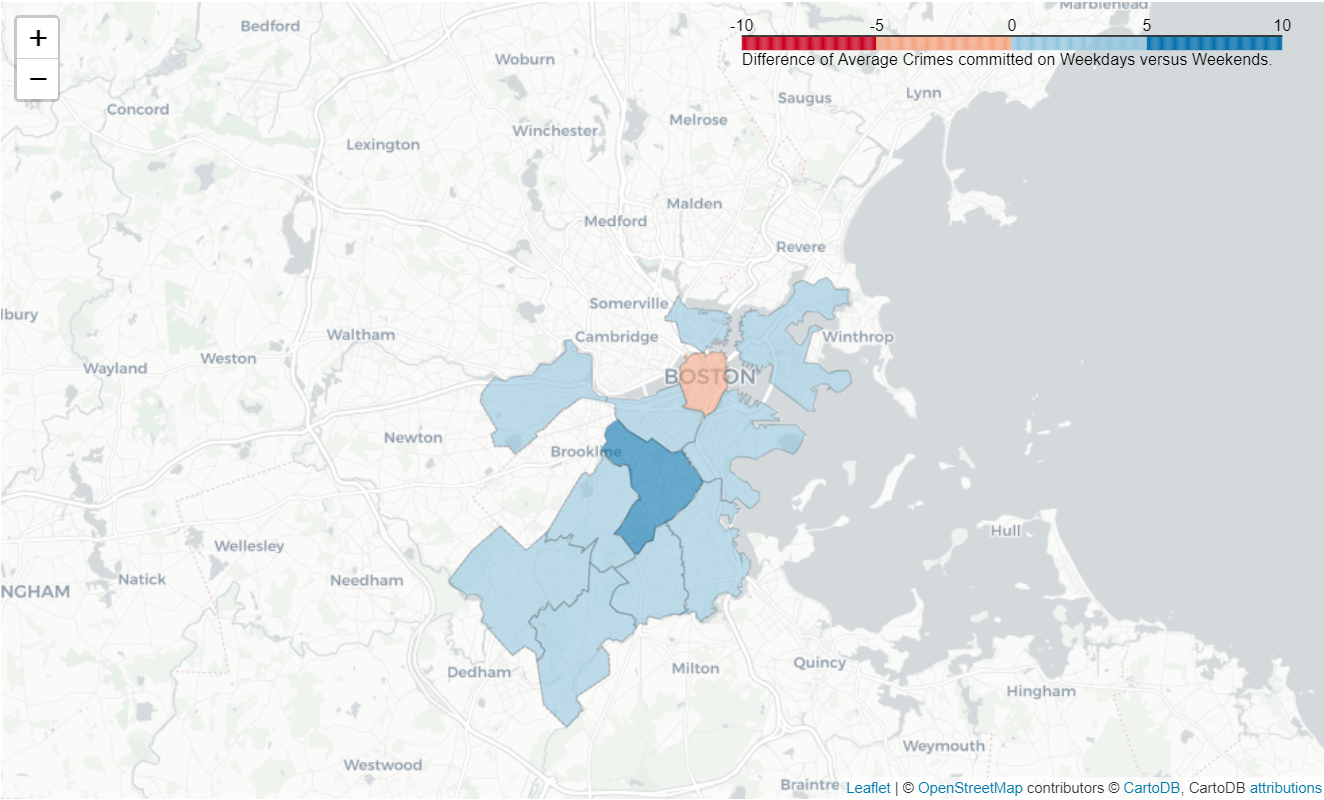
The plot helps to show how the distribution of incidents per day does vary drastically based on the day. As shown, Sunday’s 3rd quartile is significantly below Friday’s 1st quartile. This was a better solution than a line plot, which quickly became overwhelming and unreadable. There are also no unnecessary dimensions, such as color, which would confuse the reader. The axes are labeled, and there are only subtle y-axis tick marks to aid the reader in determining likely values.

Figure 3: Facet grid for incidents by day of week



After developing Figure 2, I wanted to drill down to see whether it was only a specific incident severity that was driving the narrative that incidents varied by day. I adjusted the axes for each facet and determined that all incident types were mildly affected by the day, further strengthening the hypothesis. This plot is consistent with the plots discussed in class. The bars were not grouped into one chart as it gave the impression that only certain groups were affected by day, and it made it extremely hard to compare across incident severities of the same type. The charts were colored in accordance to their severity and are ordered in the same way that you’d read them (left to right, top to bottom), and the days were ordered (they were originally ordered alphabetically).

Figure 4: Violin plot for Distribution for Motor Vehicle Accidents by Day of the Week



This plot adds another dimension to Figure 2, where it is possible to see if it is only a single district that is affecting the results. The graph takes the average number of incidents committed on a weekday and subtracts the average number of incidents committed on a weekend. As shown, the results seem modest, most have a difference of less than 5 incidents. However, given the fact that only 15-20 incidents happen per day in each district, the difference is about a 10% to 33% reduction in incidents. Moreover, nearly all the police districts reported more crimes on the weekdays than on the weekends. This strengthens the hypothesis that incidents are more likely to happen on a weekday, and thus prove my hypothesis that the day of the week influences the number of incidents. The plot is consistent with those shown in class. The color scale is divergent as 0 is in the middle of the color scale, and we want to see drastic differences between positive and negative values. Moreover, the tile set for the map was picked so that it didn’t overwhelm the data, as most built-in tile sets do not work well with choropleth maps because they are so colorful and detailed. The washed out version gives the reader context for the location, but does not point out all landmarks in the area.