1. **Data Cleaning protocol**

Our final dataset should be one object, crime. It should combine 2010 and 2019 data. The columns should be: month, primary type, description, location, arrest, domestic, violent, property.

In combining the datasets, we will mutate the data values to summarize them as months. We will also group crimes by whether they are violent crimes against persons, crimes against property, or something else. Finally, we will remove all of the columns that are not needed for our analysis.

1. **Formulate Your Question**

We have structured our data to explore were violent crimes and property crimes occur. Using this data, we can compare the proportion of violent and property crimes that happen at different locations. This data could be used by policymakers seeking to identify where and how they might take actions to improve the safety of citizens.

* Primary Research Question: Do certain types of locations have associations with different types of crimes?

1. **Exploratory Data Analysis**

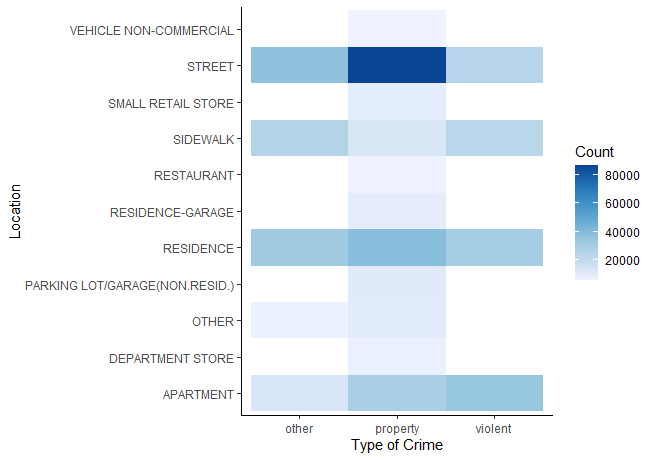
In our exploratory analysis of the combined dataset we created a number of tables to look at crime seasonality by type and crime by location. In doing so we identified a number of locations of interests that prompted us propose a two narrower research question:

* Sub-question 1 : Are violent crimes more or less likely to occur in commercial/rideshare vehicles than personal vehicles?
* Sub-question 2: What types of crimes are happening in vehicles that are not property or violent crime?

1. **Data Visualization**

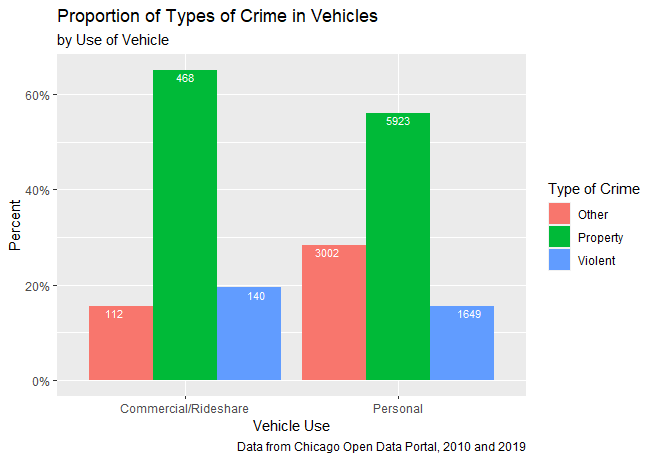
**Primary Research Question:** Do certain types of locations have associations with different types of crimes?

This plot is a heat map matrix where darker colors are associated a higher number of a particular time in each location. This plot shows us that property crime in general and, specifically, property crime that occurs on the street is the most common type of crime. This is represented by the dark blue share in the second row and middle column. The plot also shows us that most crimes of all types take place on the street or sidewalk and in apartments and residencies. This is represented by the blue bans across the plot in those four locations.



**Sub-question 1:** Are violent crimes more or less likely to occur in commercial/rideshare vehicles than personal vehicles?

This plot shows us that the there is a distinct difference in the types of crimes that occur in personal and non-personal vehicles.



We can also see that police are used much more often for property crimes.

**Sub-question 2:** What types of crimes are happening in vehicles that are not property or violent crime?

