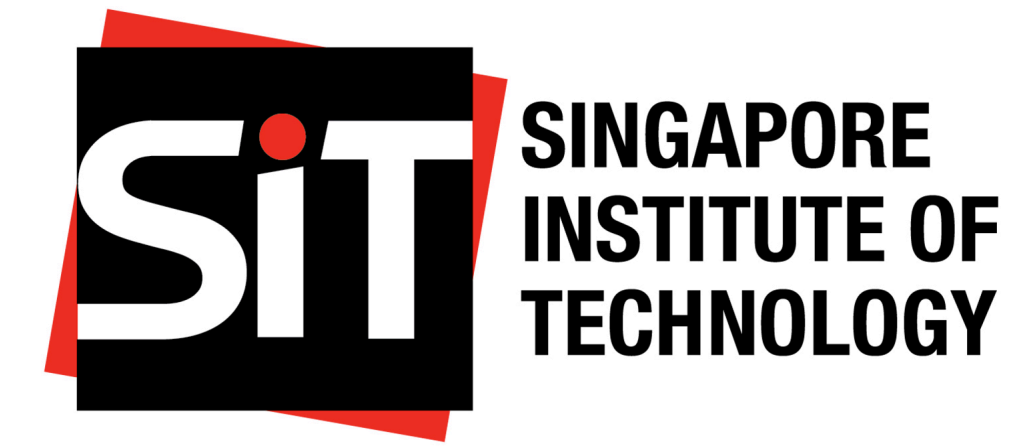


# Mapping Crime in the Los Angeles: A Visual Journey (2020-2024)

*Team Cyan* (Information and Communication Technologies)



# INTRODUCTION

Crime in america is a significant concern , the safety and well-being of residents, business and tourism will be impacted severely. There is a sharp rise in motor vehicle thefts is up 25% from 2019 to 2022<sup>1</sup> showing that there is still plenty of improvement required to curb the crime rate.

Understanding and identifying area with high type of crime are can help the effort of effective law enforcement training and perpetration. Data constantly shows that neighborhoods such as South Los Angeles are still prevalent with violent crime<sup>2</sup>. Using statistical past data , using quantifiable information to identify hotspots allows law enforcement agencies allowing efficient dividing of resource better while maximizing safety.

Taking a look at a crime distribution around the University of Southern California's University Park campus on medium<sup>3</sup> (Figure 1). This visualization is display see through blue dots to represent a hit in crime on a specific area, while straight to the point however there are several aspects of the plot can be refine.

## PREVIOUS VISUALIZATION

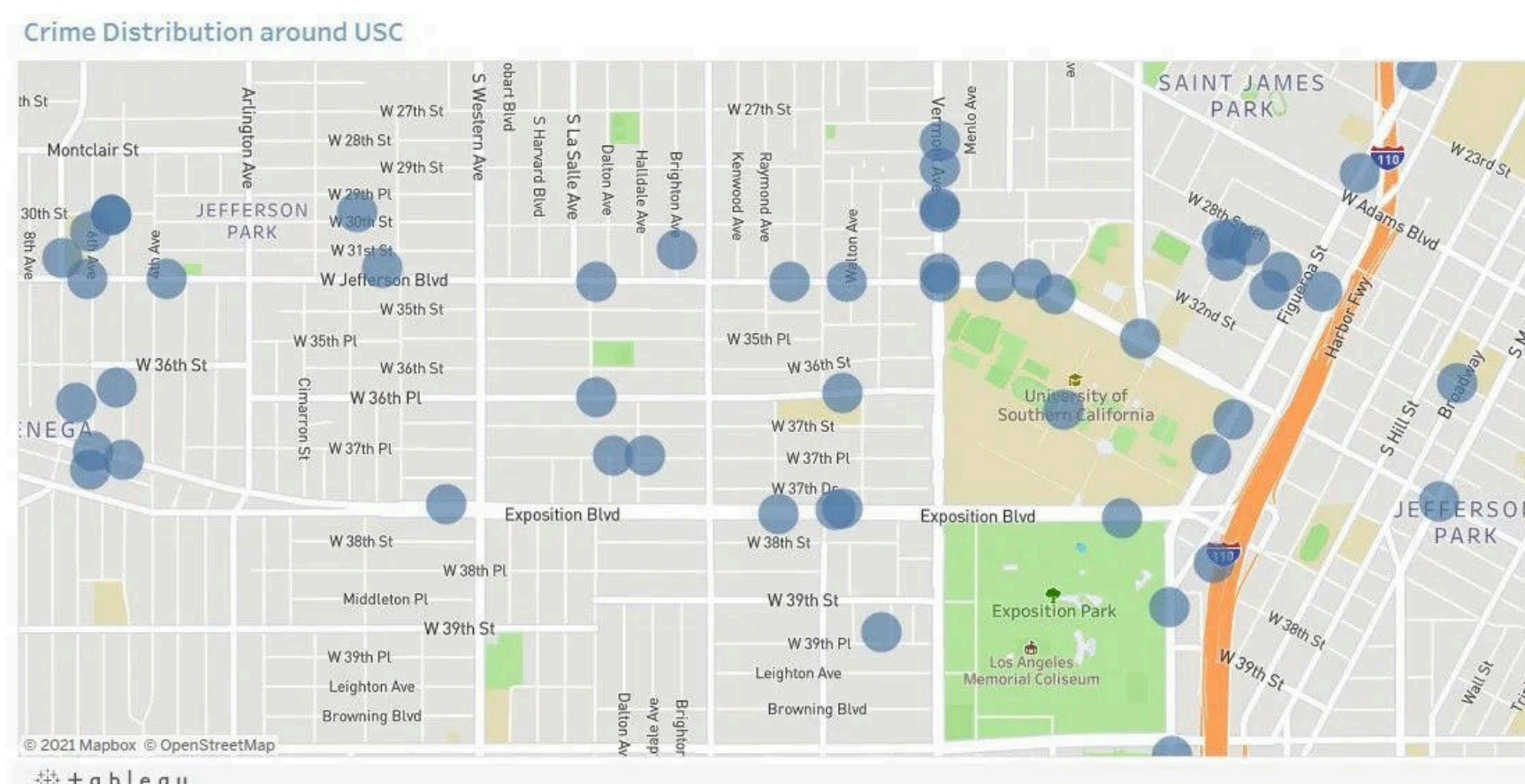


Figure 1: Crime distribution around USC, published by the Medium.

## STRENGTHS

1. Alpha was used on the circles allowing darker spots to appear if overlapped.
2. Gentle color blue was used, allowing the user to easily view spots affected by a crime. The color picked does not have a conflicting color with the background of the map, which was a nice touch.
3. An area with good distribution was picked, as there are clusters displayed on the map.
4. As the mouse is hovered over the circles, the exact number of crimes is displayed through a tooltip (Figure 2), allowing the user to see the exact number of crimes in that area .

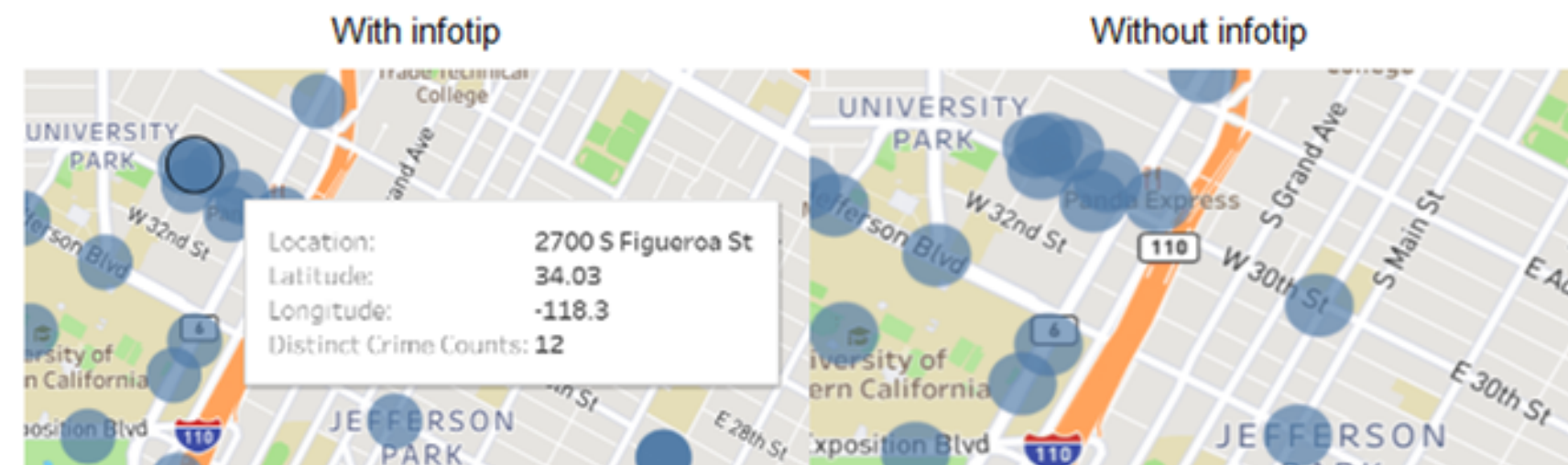


Figure 2: Crime distribution around USC, published by the Medium.

## SUGGESTED IMPROVEMENTS

1. *State layer view*: separated with area code, performs better visualization of the crime distribution. A localized view does not represent a crime spread well for meaningful actions.
2. *Identify type of crime clearly*: A clearer view of top crime should be labeled and displayed.
3. *Add missing title and guides*: Title should be used for clear description.
4. *Add missing guides*: Guides should be used to show clear distinction of color shade to total crime count.
5. *Use a saturation color palette*: Shows a meaningful progression through color space. Saturation palettes show cold to hot zone allowing humans to see intensity of an area.
6. *Label locations*: Labeling popular city centers allow enforcers to see crime distribution and easily identify the areas with high crime rate.
7. *Add crime types for each area*: Displaying the top 3 crimes in each area allows for better understanding of the crime distribution and provides additional information.

## IMPLEMENTATION

## Data

- Los Angeles year 1st January 2020 to 7th June 2024.<sup>4</sup> The data used is the universal data while Figure 1 use a subset of the data ending at 2021. The data set are broken apart to 10 years data set 2010 to 2019<sup>5</sup> however different format might be implement hence not used.

## Software

We used the Quarto publication framework and the R programming language, along with the following third-party packages:

- *dplyr* for data manipulation
- *tidyverse* for data transformation, including *ggplot2* for visualization based on the grammar of graphics
- *readxl* for data import
- *lubridate* for date and time manipulation
- *DT* for interactive data tables
- *knitr* for dynamic document generation
- *pals* for color palettes
- *RColorBrewer* for color palettes

## IMPROVED VISUALIZATION

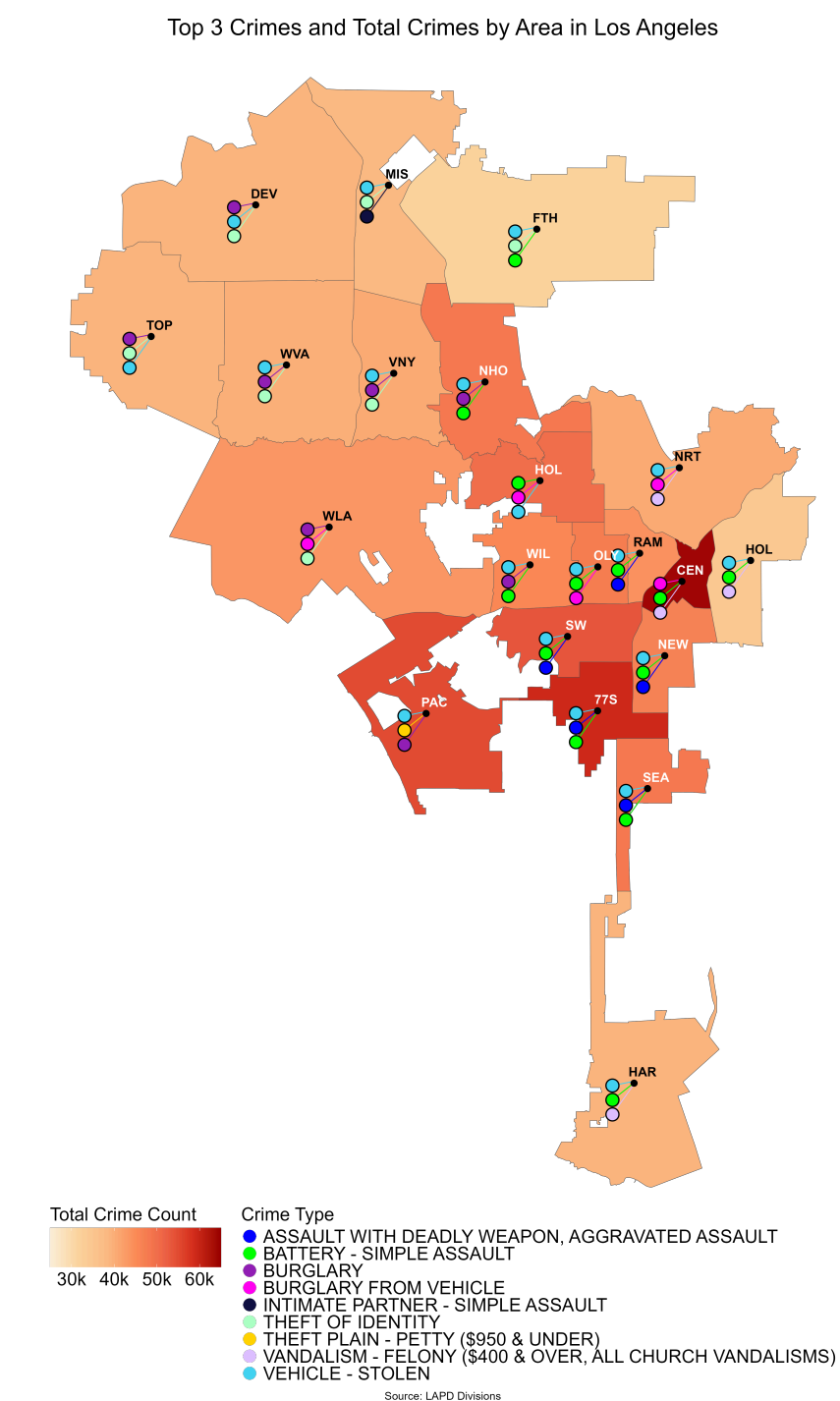


Figure 3: Top 3 crimes and total crimes by area.

## FURTHER SUGGESTIONS FOR INTERACTIVITY

While our visualization was designed for a poster, interactive features were not implemented. However, in an HTML document, these features can be achieved using various R packages. **ggplot2** allows for **hover**, **drag**, **zoom**, and **export**, which improves accessibility for people with sight disabilities by enabling zoom to increase text size. **Shiny** facilitates the **sorting of graphs** to clearly differentiate categories and provides **dynamic input**, such as displaying the distribution of only robbery crimes throughout the state. Additionally, **plotly** offers **customized tooltips** with ggplot2, expands long town names, and shows exact numbers without crowding. By **darkening borders** and adding shadows, plotly highlights areas when hovered, enhancing the overall user experience.

## CONCLUSION

To update not changed yet, We successfully implemented all suggested improvements for the non-interactive visualization. By labeling every state and choosing a colorblind-friendly palette, the revised plot is more accessible. The logarithmic color scale makes the decrease in incidence after the introduction of the vaccine less striking but enables readers to detect patterns in the low-incidence range more easily.

<sup>1</sup><https://www.marketwatch.com/guides/insurance-services/car-theft-statistics/>

<sup>2</sup><https://www.latimes.com/california/story/2023-10-12/violent-crime-is-down-fear-is-up-why-is-la-perceived-as-dangerous>

<sup>3</sup><https://towardsdatascience.com/visualizing-crime-in-los-angeles-14db37572909/>

<sup>4</sup>[https://data.lacity.org/Public-Safety/Crime-Data-from-2020-to-Present/2nrs-mtv8/data\\_preview](https://data.lacity.org/Public-Safety/Crime-Data-from-2020-to-Present/2nrs-mtv8/data_preview)

<sup>5</sup><https://catalog.data.gov/dataset/crime-data-from-2010-to-2019>