

Assignment 2: Data Warehouse.

Story Analysis of Louisville Kentucky's Pre, Mid and Post Covid Restrictions Crime Rate

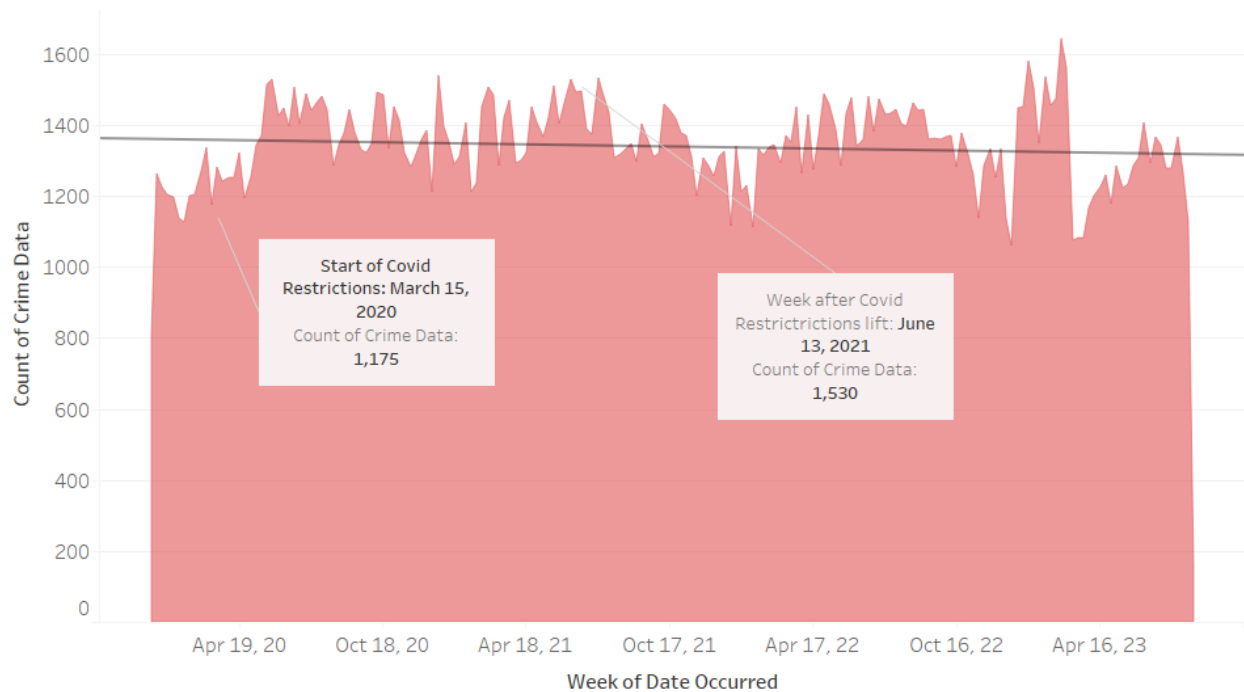
Data Warehouse: A data warehouse is a centralized store of data from diverse sources, organized and standardized to provide a consolidated view of an organization's data while supporting business intelligence analysis and reporting.

I decided to use crime data from the last four years (2020-2023) to analyze the effect of covid restrictions on crime occurrences in Louisville Kentucky. I downloaded four datasets from the Louisville KY data warehouse detailing the yearly crime data for 2020, 2021, 2022, 2023. The data was then compiled vertically resulting in a dataset of about 256000 rows. Some preprocessing was also done with some attributes dropped while others were transformed to be more suitable for analysis. Namely the Date occurred and reported variables, the dates and times were separated into their own individual variables for easier analysis. Also a new variable was calculated called report time that shows how much time passes between a crime's occurrence and its report.

Based on the info, I believe the warehouse need to be improved in terms of format standardization. The dates used and inputted in the warehouse need to use the same format so they can be easily compared and analyzed. Also, the dates and times should be separated. Some variables need to be broken down into smaller variables. There should also be descriptions on what each variable in a dataset is.

Moving onto the analysis part, Tableau was used to analyze and present findings to the new mayor on the crime in the city of Louisville Kentucky.

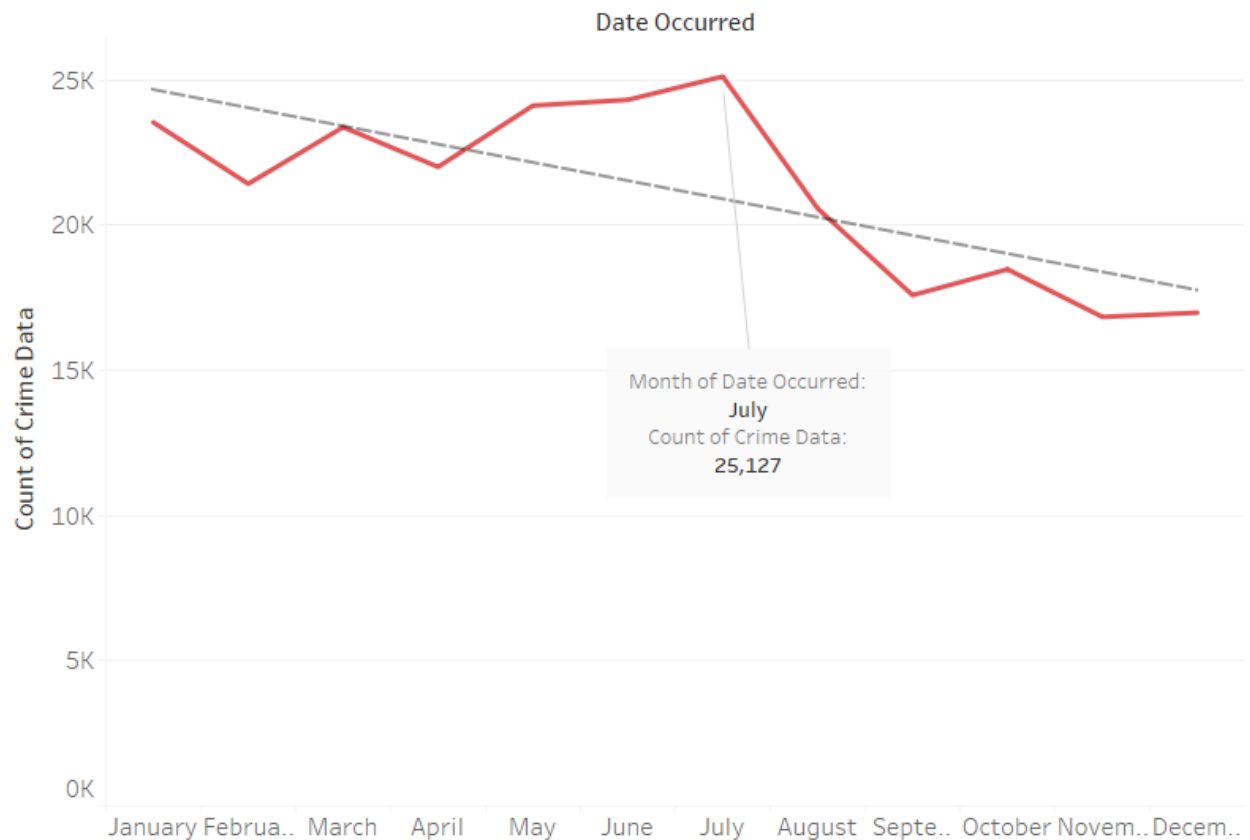
The Weekly Frequency of Crime Pre, Mid and Post Covid Restrictions



The plot of count of Crime Data for Date Occurred Week. The data is filtered on Date Occured Year, which keeps 2020, 2021, 2022 and 2023.

The graph above shows the weekly crime rate in Louisville KY from 2020 to 2023. As you can see annotated are the weeks when covid restrictions were officially put in place and lifted. As you can see it doesn't seem to have had a drastic or hugely noticeable effect on the crime rate. However, we can see that the overall trend of crime in the city is going down.

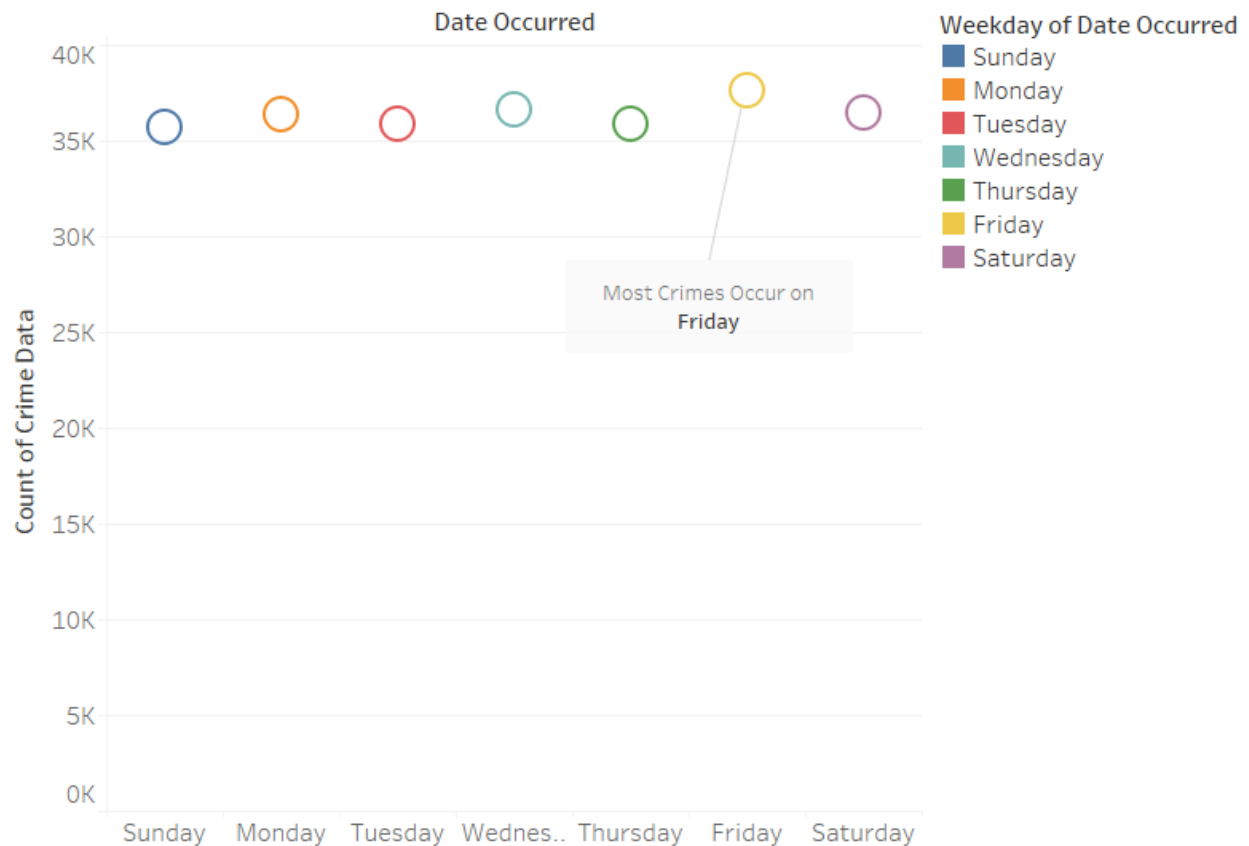
Which Month Does Crime Occur the Most?



The trend of count of Crime Data for Date Occurred Month. The data is filtered on Date Occurred Year, which keeps 2020, 2021, 2022 and 2023.

Further analysis shows that in terms of seasonality, in the past four years most of the crimes happen in the month of July and as we head towards December the crime rate goes down.

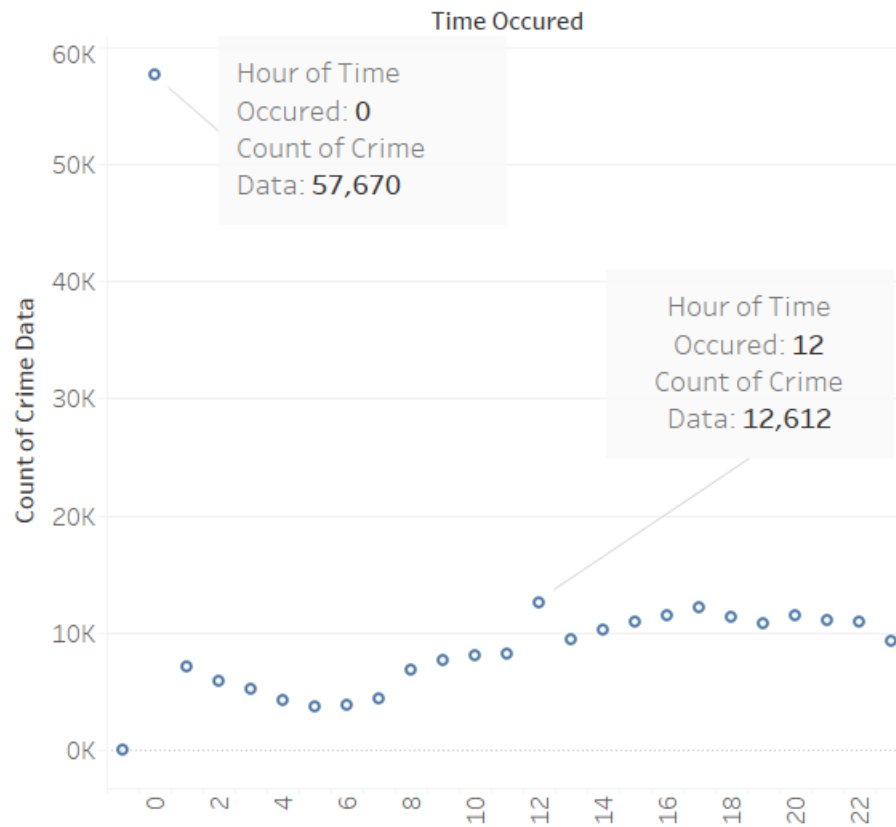
Which Day of the week Sees the most crime?



Count of Crime Data for each Date Occurred Weekday. Color shows details about Date Occurred Weekday. The data is filtered on Date Occured Year, which keeps 2020, 2021, 2022 and 2023.

Also, most crimes are likely to happen on Friday.

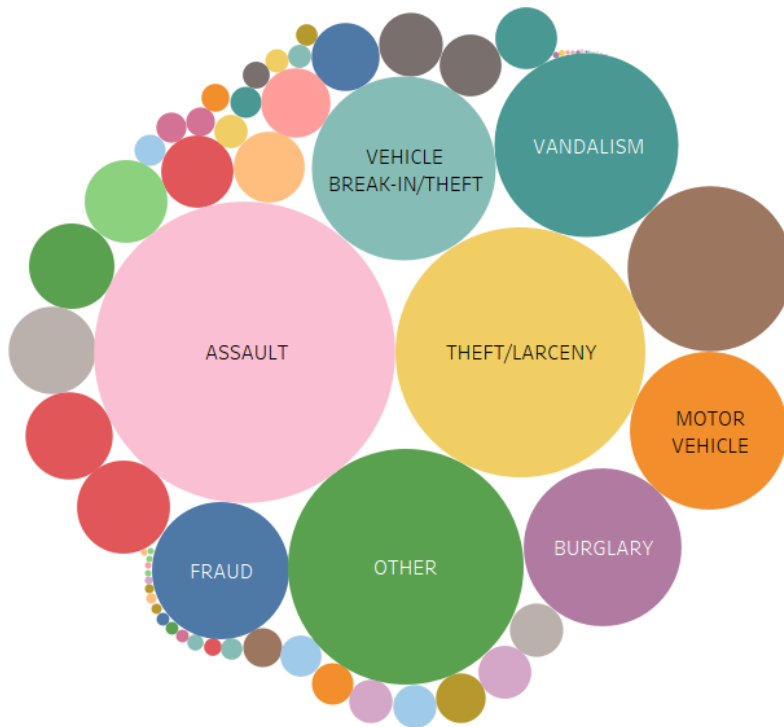
What Hours of the Day do the most Crimes Occur?



Count of Crime Data for each Time Occured Hour. The data is filtered on Date Occured Year, which keeps 2020, 2021, 2022 and 2023.

With the majority of crimes being committed at midnight and noon.

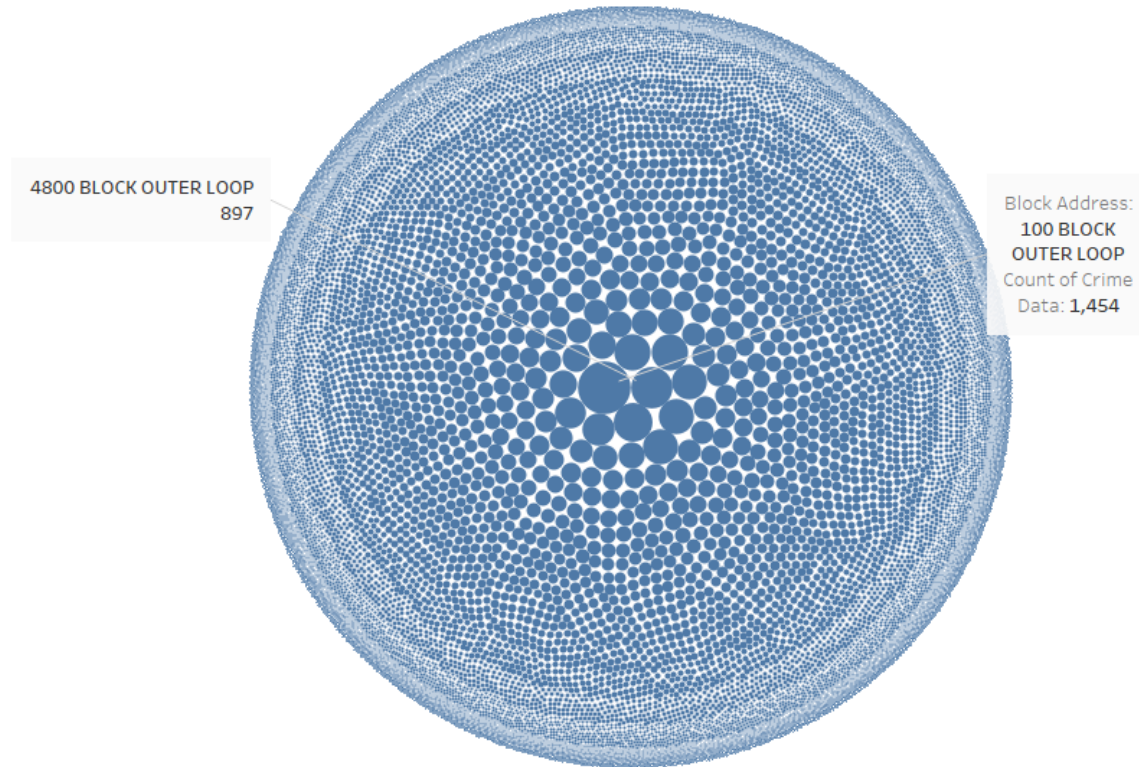
What Type of Crime Occurs the Most?



Crime Type. Color shows details about Crime Type. Size shows count of Crime Data.
The marks are labeled by Crime Type. The data is filtered on Date Occured Year, which keeps 2020, 2021, 2022 and 2023.

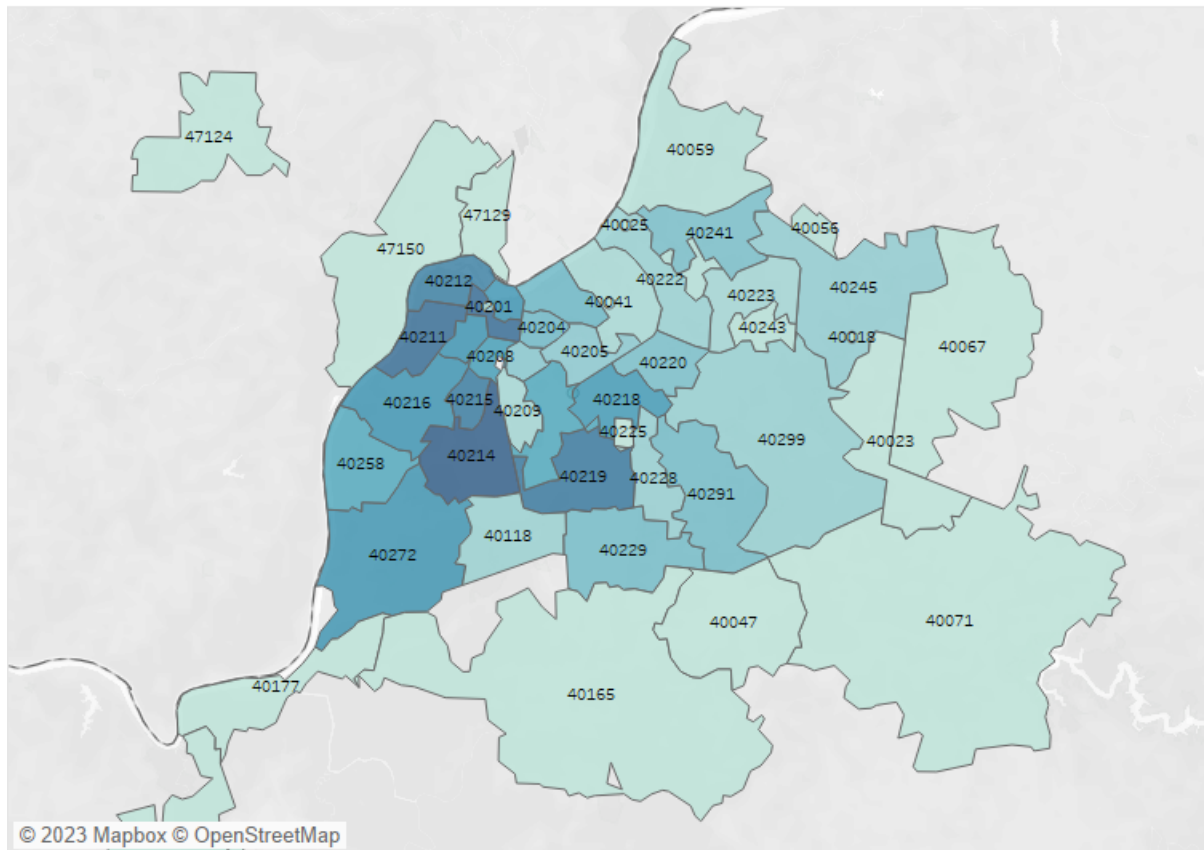
I found Assault, Larceny, Vandalism and Vehicle Theft were the most committed named crimes in the past four years.

Which Blocks Experience the Most Crime?



Block Address. Size shows count of Crime Data. The marks are labeled by Block Address. The data is filtered on Date Occured Year, which keeps 2020, 2021, 2022 and 2023. The view is filtered on Block Address, which excludes NULL.

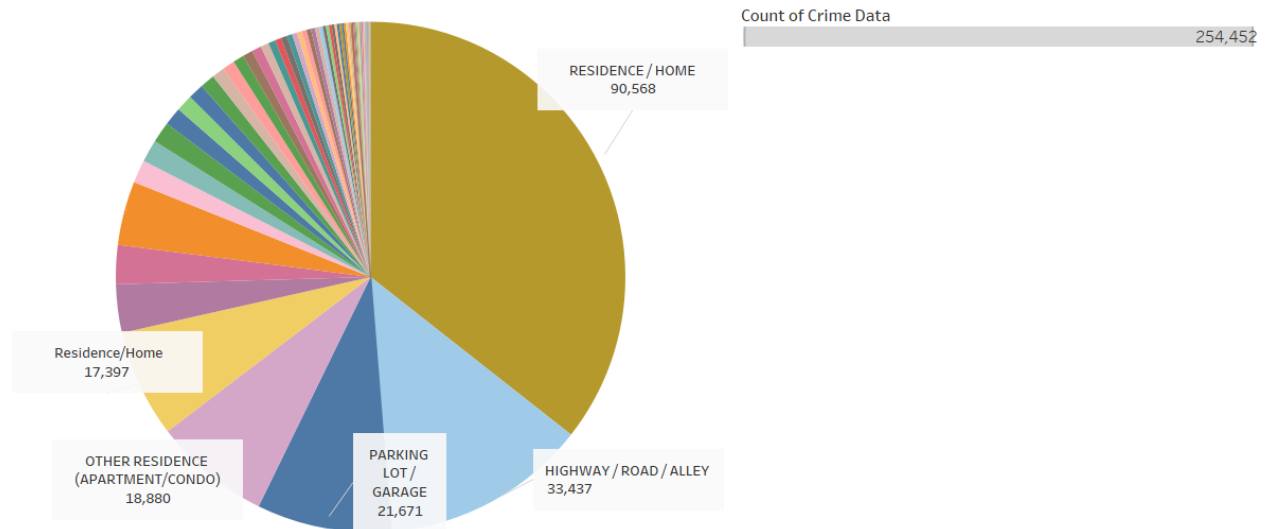
Which Zip Codes Experience the Most Crime?



Map based on Longitude (generated) and Latitude (generated). Color shows count of Crime Data. The marks are labeled by Zip Code. Details are shown for Zip Code. The data is filtered on Block Address and Date Occured Year. The Block Address filter excludes NULL. The Date Occured Year filter keeps 2020, 2021, 2022 and 2023. The view is filtered on Latitude (generated), Longitude (generated) and Zip Code. The Latitude (generated) filter keeps non-Null values only. The Longitude (generated) filter keeps non-Null values only. The Zip Code filter excludes 24251, 40403, 45424 and 47170.

I was also able to find which locations experience the most crime both in terms of block addresses and zip code.

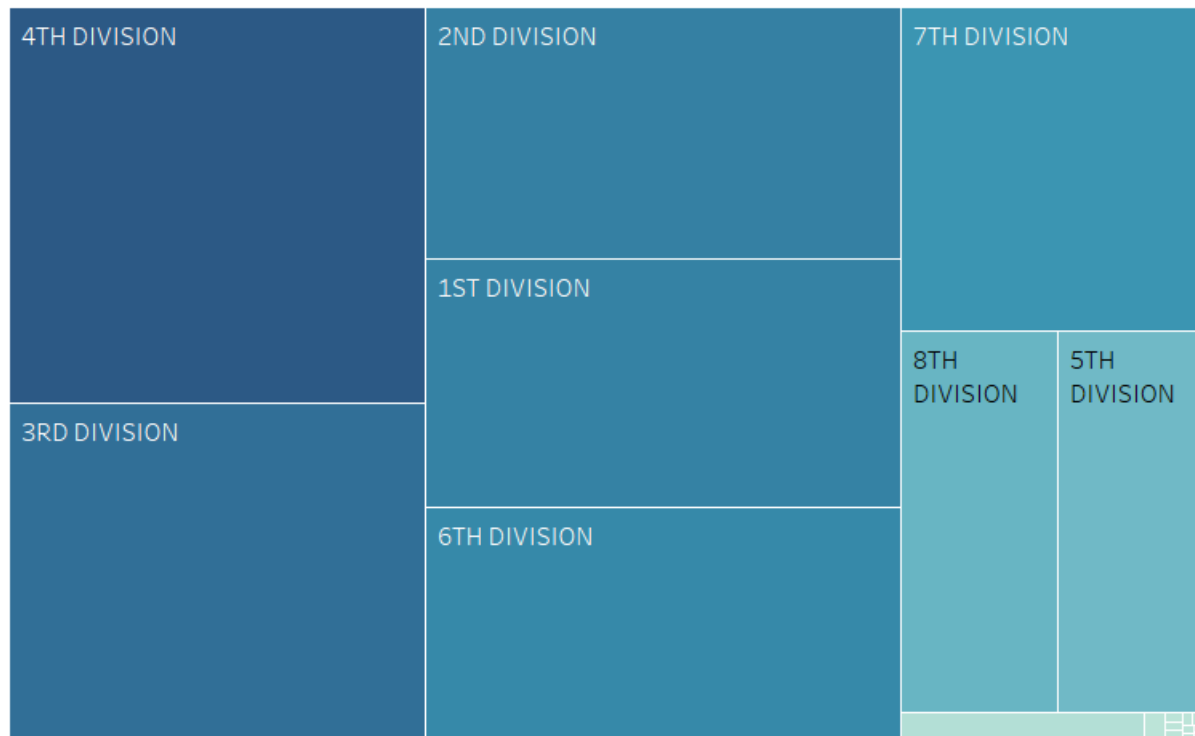
Where is Crime Most likely to Happen?



Premise Type (color) and count of Crime Data (size). The data is filtered on Date Occured Year, which keeps 2020, 2021, 2022 and 2023.

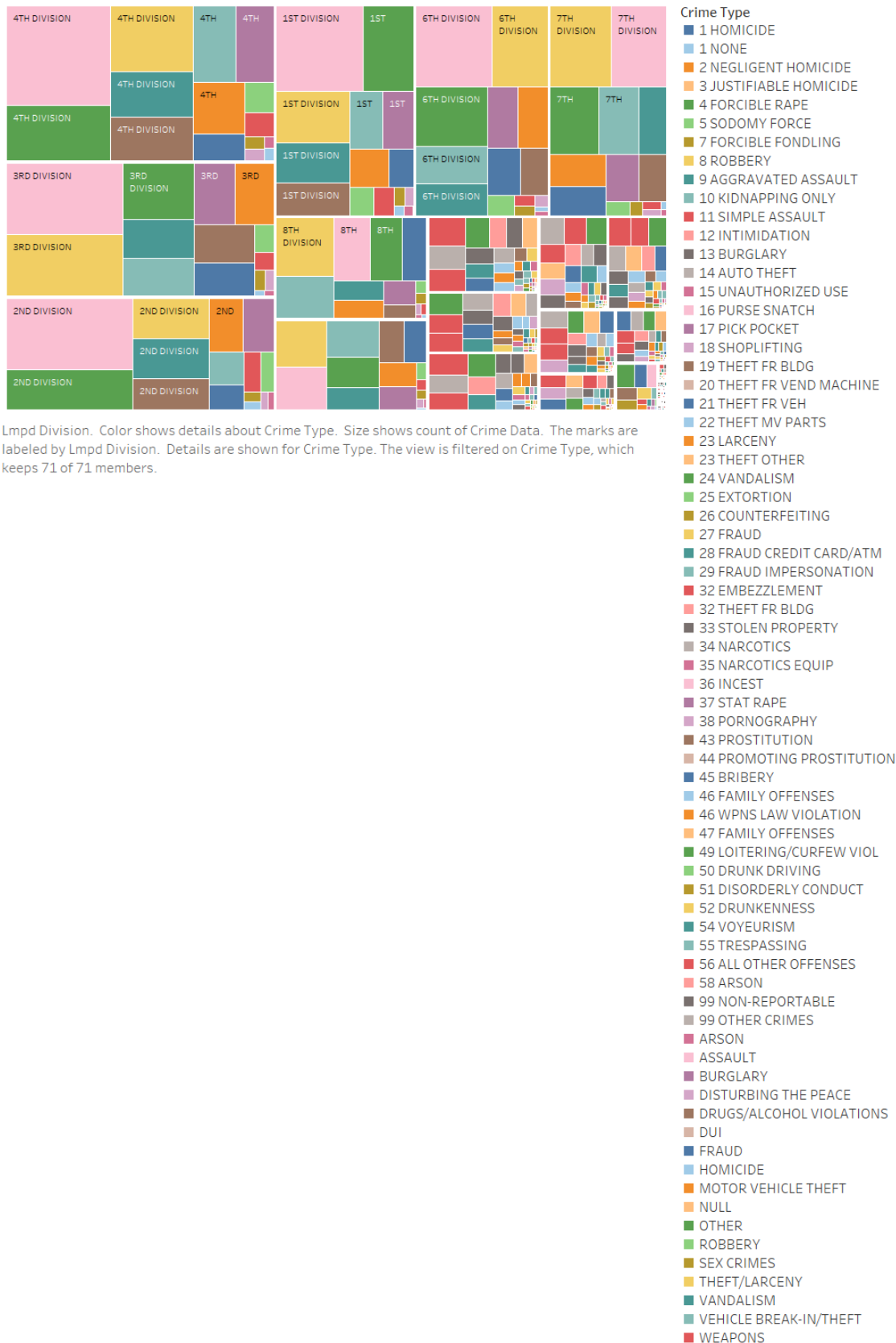
I was able to ascertain in the past four years in what type of location a crime was most likely to happen based on the data. Unsurprisingly it is the home or residence.

Which Divisions Handle the Most Crime?



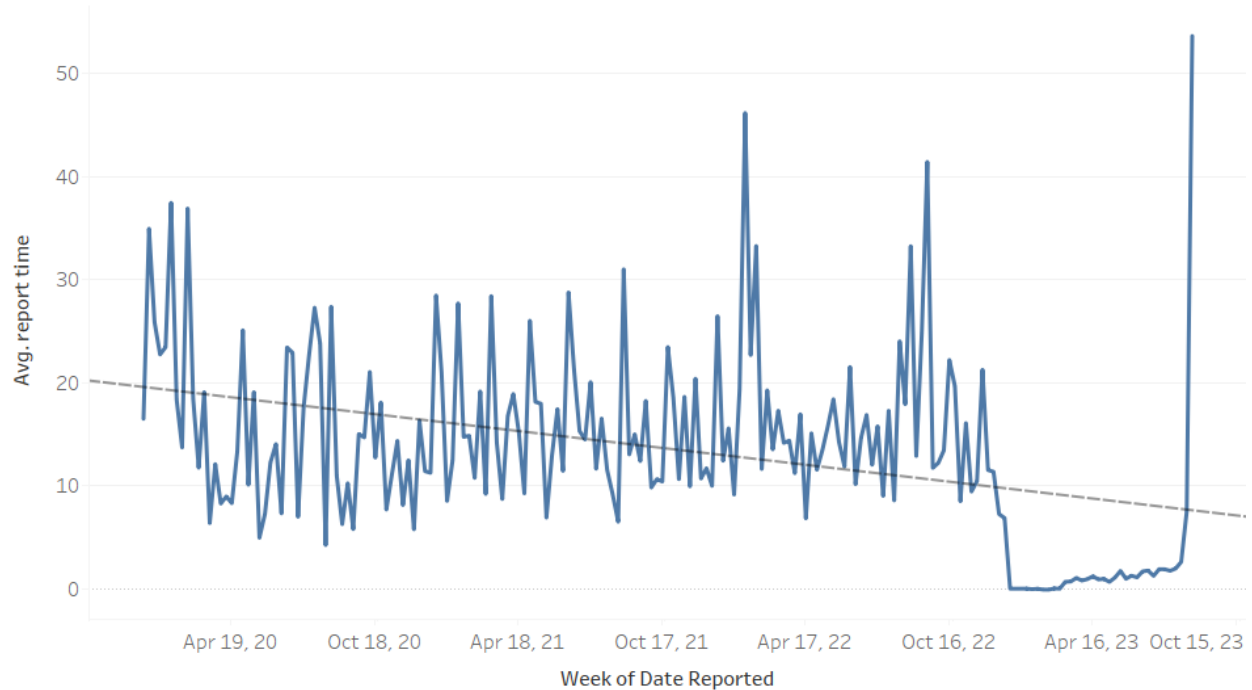
Lmpd Division. Color shows count of Crime Data. Size shows count of Crime Data. The marks are labeled by Lmpd Division. The data is filtered on Date Occured Year, which keeps 2020, 2021, 2022 and 2023.

What Kind of Crime Does Each Division Deal with The Most?



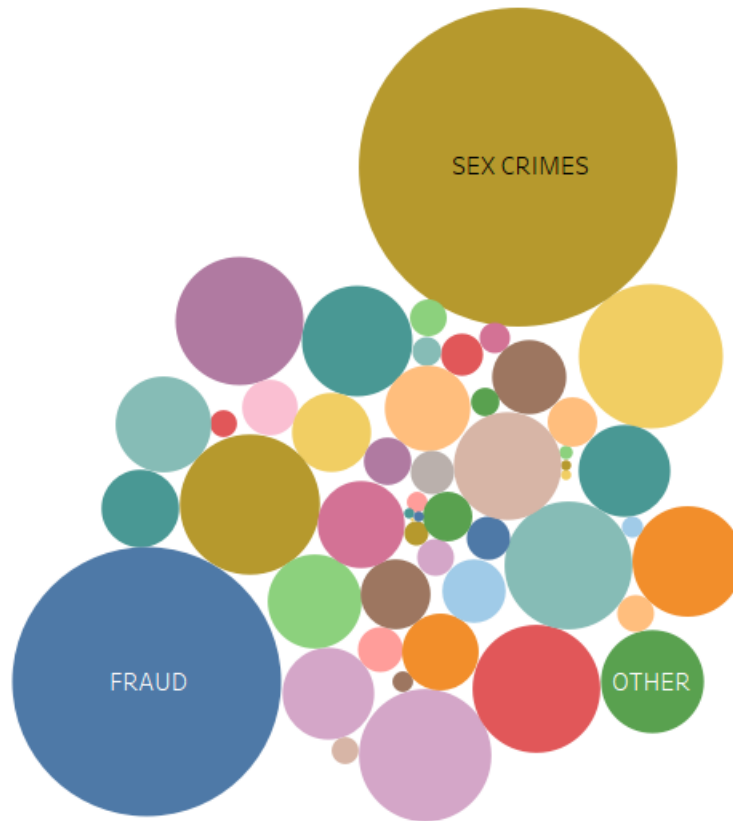
Based on the data it was discovered which division dealt with the most crime and the breakdown of what type of crime they dealt with. This could help with the allocation of resource amounts and types to the divisions.

How Long Does It take to Report A Crime?



The trend of average of report time for Date Reported Week. The data is filtered on sum of Days to report and Date Reported Year. The sum of Days to report filter keeps non-Null values only. The Date Reported Year filter keeps 2020, 2021, 2022 and 2023.

Which Crimes Take the Longest to Get Reported?



Crime Type. Color shows details about Crime Type. Size shows average of report time. The marks are labeled by Crime Type. The data is filtered on Date Occured Year, which keeps 2020, 2021, 2022 and 2023. The view is filtered on average of report time and Crime Type. The average of report time filter includes greater than and or equal to 0.00 and keeps Null values. The Crime Type filter keeps 71 of 71 members.

Finally, utilizing the calculated attribute of report time we were able to find out that report time on average was going down, which is a good thing, also that sex crimes tend to be the crime victims take the longest to report.