

Seattle Crime Analysis

Abhishek K., Isha D., Sanyam S.,
Siddharth P., Shinjini G.



Table of Contents

1 Data Cleaning

2 Note on Coordinates

3 Data Analysis

4 Insights



Data Cleaning

- ▶ The original Seattle crime data includes information about offense category, offense, sector, beat, start time, end time, location, etc.
(Size: 10334151 x 17)
- ▶ All rows with nulls in the start time were dropped to make the data more meaningful. Added year and month column for easy filtering (Size: 1032212 x 19)
- ▶ Filtered for start years \geq 2008 + removed unclear/faulty values from Beat, Sector, MCPP, Crime Against Category, Precinct + fixed spelling error in MCPP - (Size: 1021595 x 19)
- ▶ Dropped nulls in Precinct, Sector, and Beat
(Size: 1021590 x 19)



Note on Coordinates

- ▶ We noticed that ~37K rows have latitude and/or longitude that lies outside Seattle
- ▶ The latitude and longitude for Seattle is 47.6062° N, 122.3321° W (47.60, -122.33)
- ▶ Most of the incorrectly entered points have latitude greater than 47.6
- ▶ We can ignore these inconsistencies in our analyses as we are not using coordinates

Crime and Seasonal Trends (2008 - 2022)

At what time does most crime happen?

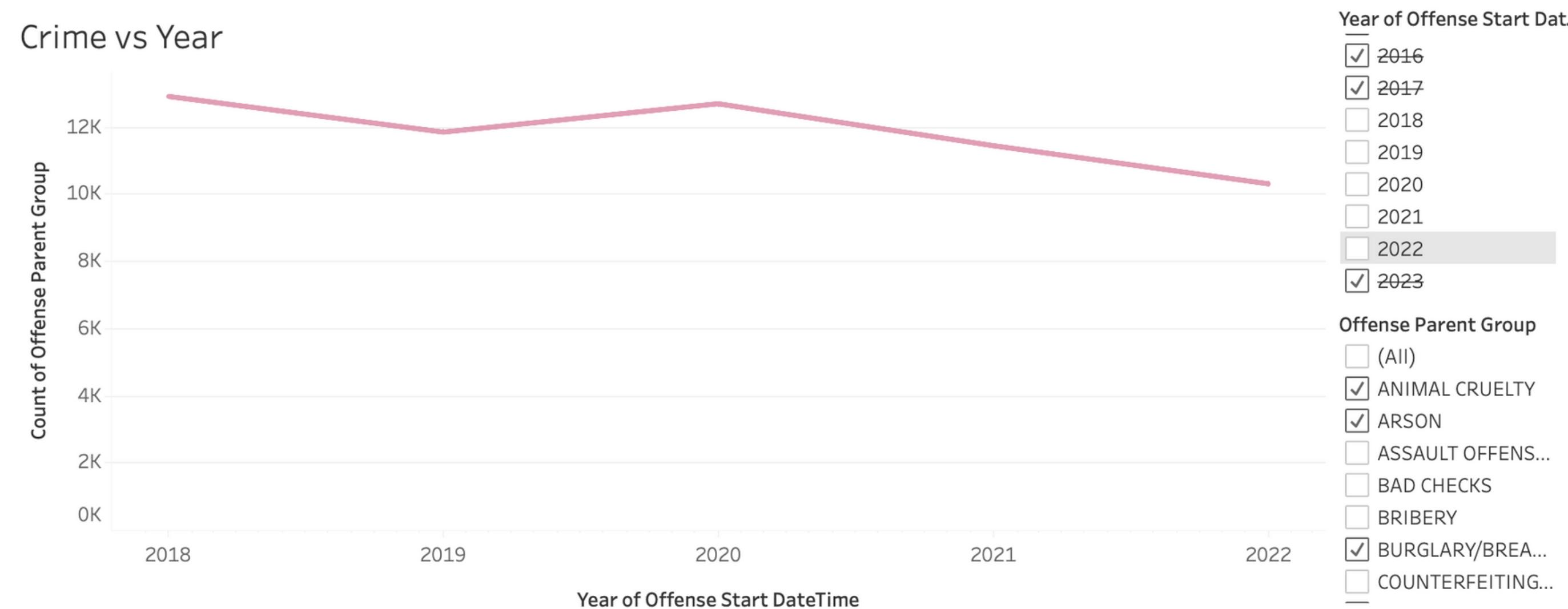
Does drunkeness, drugs and driving under the influence increase during the holiday season?

Does Assault, loitering, and drugs and narcotics increase in the holiday season?

Any specific interesting insights observed.

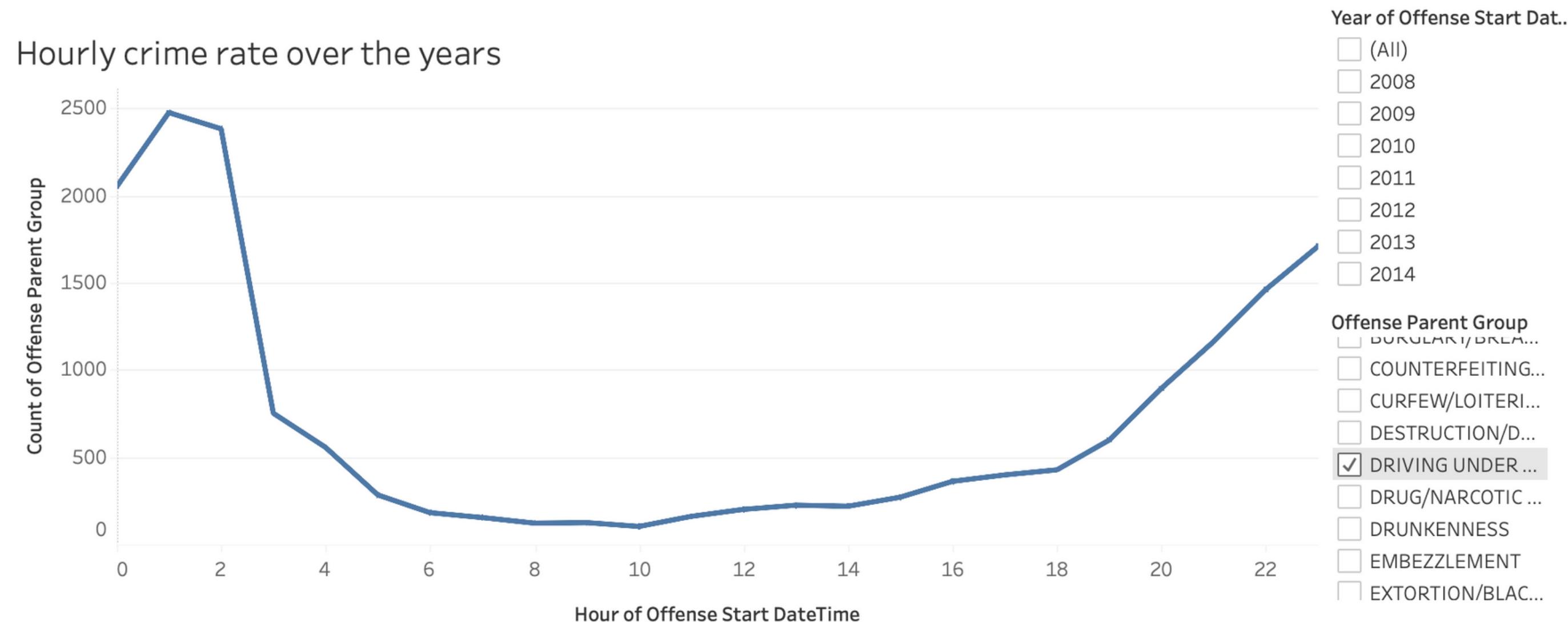
Insights

Animal cruelty, arson, bad checks, burglar, counterfeiting, curfew, embezzlement, fraud, prostitution, and trespassing has decreased in the past 3 years



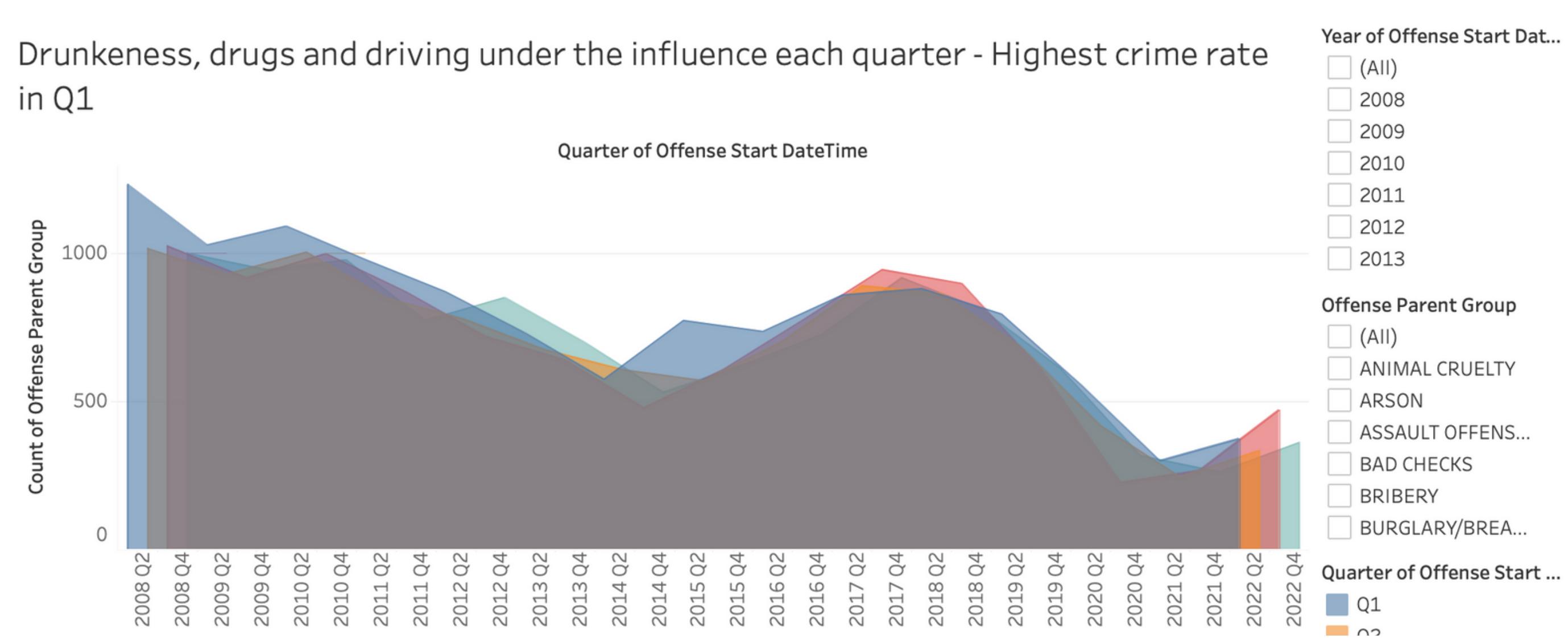
Insights

Overall, most crimes happen at midnight.
Most driving under the influence happens from 6pm to 2am



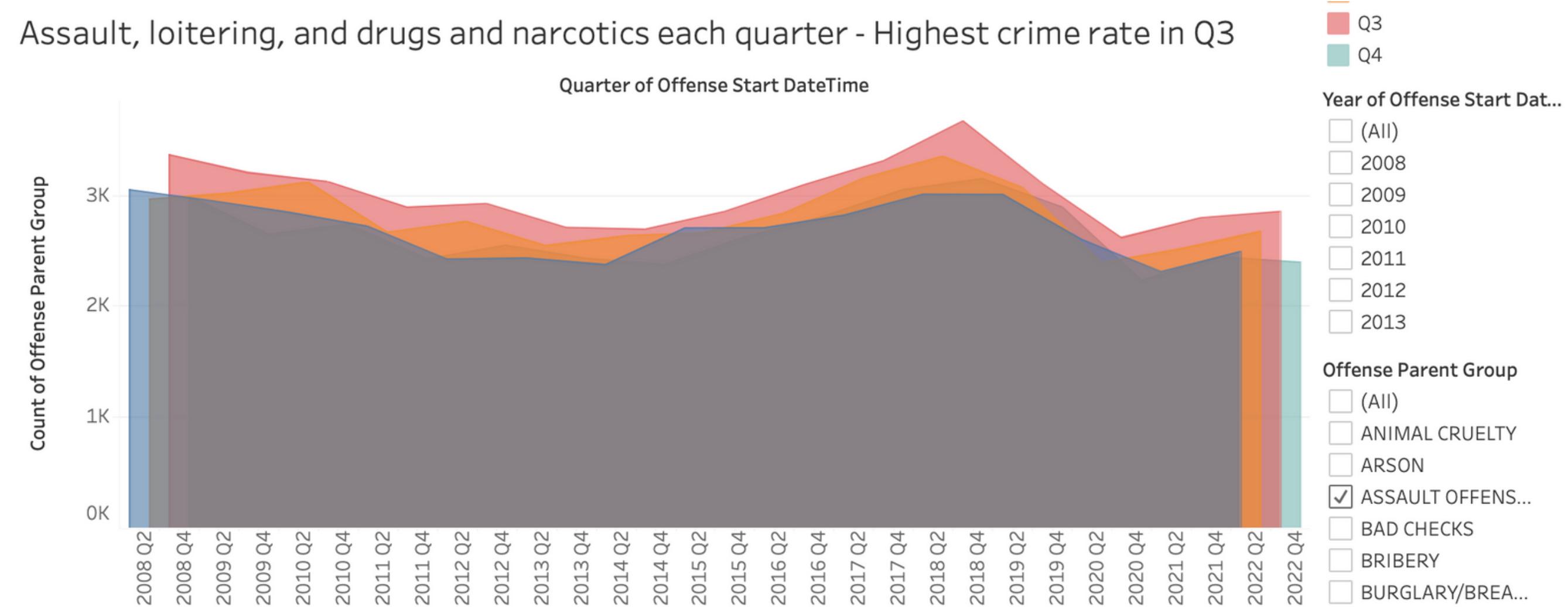
Insights

Drunkenness, drugs and driving under the influence each quarter is highest in Q1 (Jan-Mar)



Insights

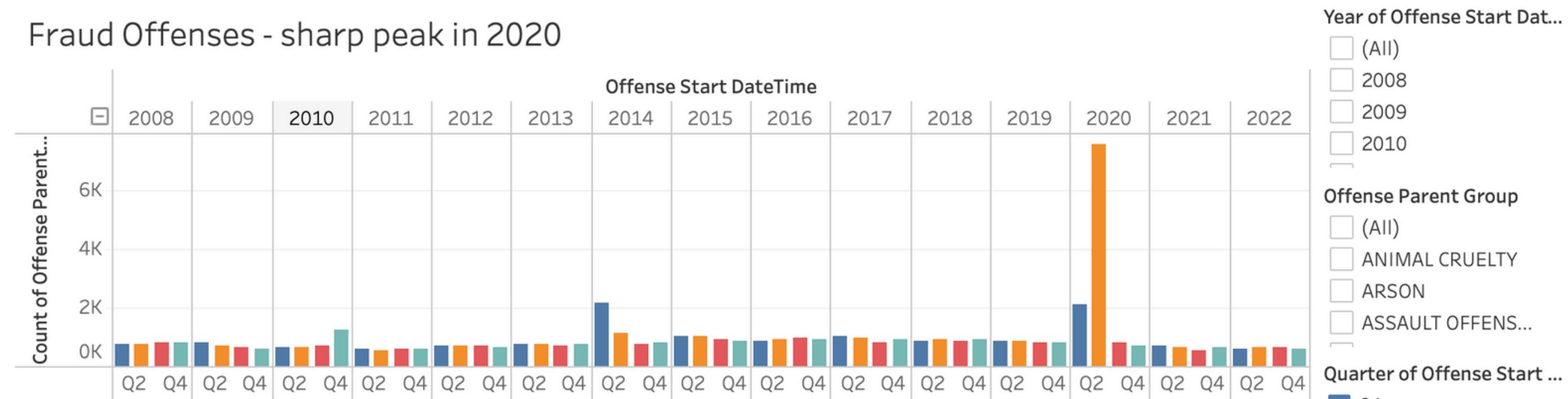
Assault, loitering, and drugs and narcotics each quarter is highest in Q3 (Jul-Sept)



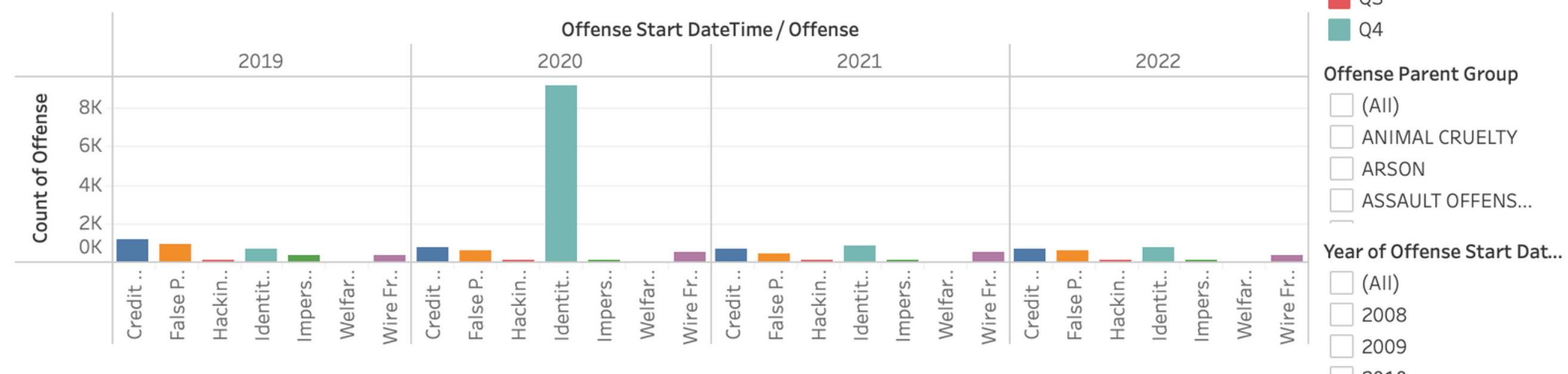
Insights

Identity theft increased sharply in 2020

Fraud Offenses - sharp peak in 2020



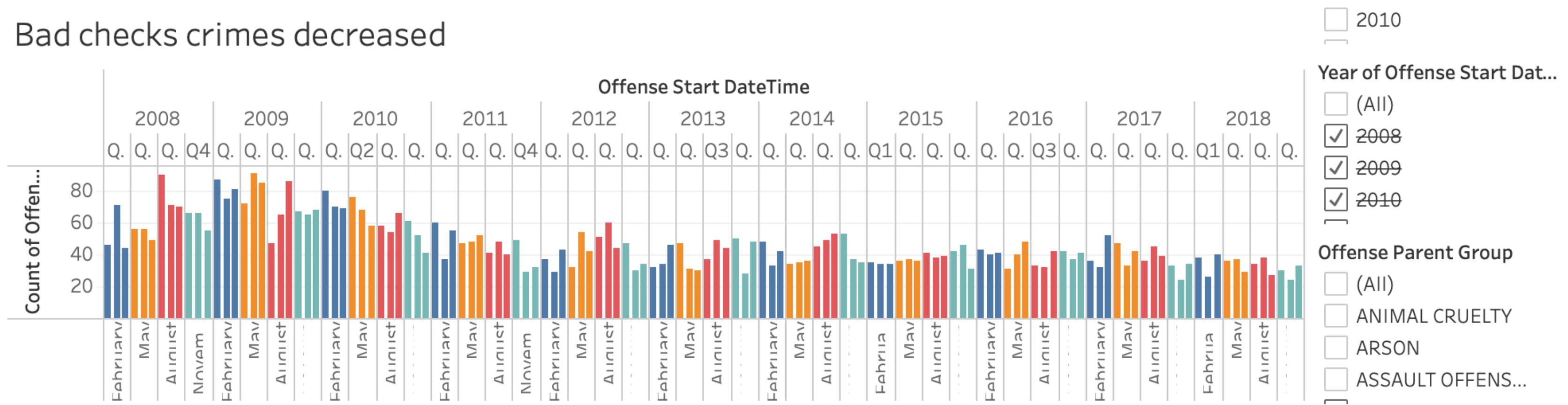
Identity theft increased in 2020



Insights

Bad checks crimes have decreased over the years

Bad checks crimes decreased



Insights Summary

- Animal cruelty, arson, bad checks, burglar, counterfeiting, curfew, embezzlement, fraud, prostitution, and trespassing has decreased in the past 3 years
- Overall, most crimes happen at midnight.
- Most driving under the influence happens from 6pm to 2am
- Drunkeness, drugs and driving under the influence each quarter is highest in Q1 (Jan-Mar)
- Assault, loitering, and drugs and narcotics each quarter is highest in Q3 (Jul-Sept)
- Identity theft increased sharply in 2020 - <https://www.justice.gov/usao-wdwa/pr/seattle-man-who-used-stolen-identities-seek-nearly-2-million-covid-19-benefits-pleads>
- Bad checks crimes have decreased over the years

Geographic Hotspots (2016-2023)

What neighborhoods witnessed maximum crime rates in the given year frame?

What type of offenses are most common in which neighborhoods?

What are the most/least safe neighborhoods for women and children?

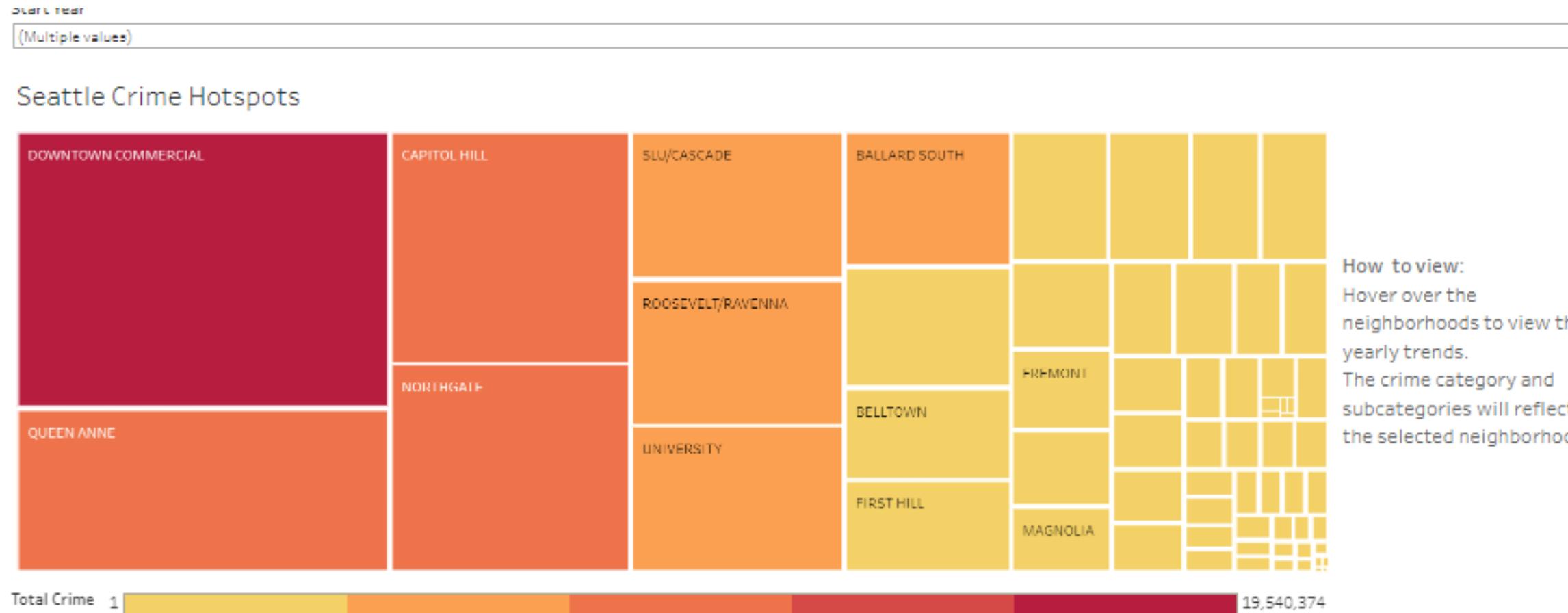


Data Pre-processing

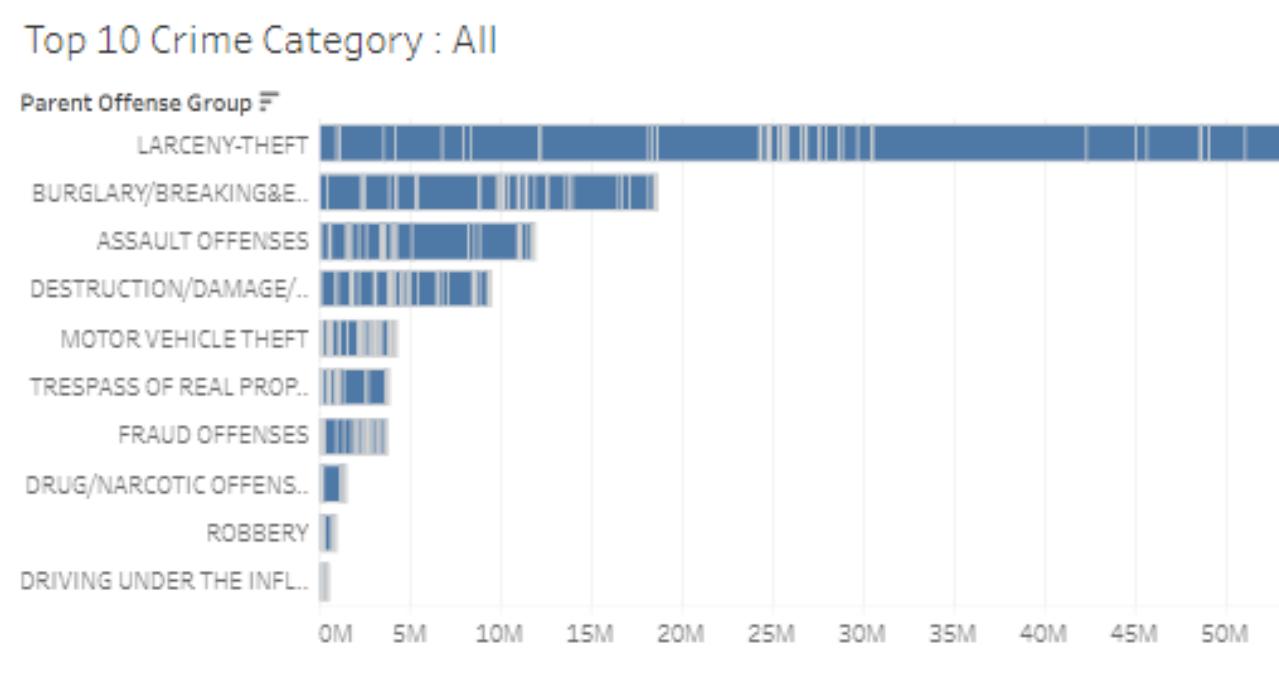
```
#read data and explore
spd_crime <- read.csv('crime_greaterthan2008.csv')
summary(spd_crime)
colSums(spd_crime == "")
#select required variables
crime_new <- spd_crime %>% select(Crime.Against.Category,
                                         Offense.Parent.Group,
                                         Offense, start_year,
                                         Precinct, Beat, MCPP,
                                         Report.Number) %>%
  filter(start_year > 2015)
# fix issues
crime_new$MCPP[crime_new$MCPP == 'CAPTIOL HILL'] <- 'CAPITOL HILL'
# calculate total crime per group
crime_new <- crime_new %>%
  group_by(MCPP,Offense.Parent.Group,Offense, start_year) %>%
  mutate(total_crime = n())
write.csv(crime_new, "final.csv", row.names=FALSE)
```

- For this analysis, we need the following variables: Offense, Offense Parent Category, Start Year, Precinct, Beat, MCPP
- Fixed the incorrect neighborhood
- Created another variable to get a count of crimes in each area, year, offense type.

What neighborhoods witnessed maximum crime rates in the given year frame?



- Most unsafe - Downtown, Queen Anne, Capitol Hill
- Most safe - Commercial Duwamish, Pigeon Point, Eastlake(E)



Top 10 Crime Sub-Category

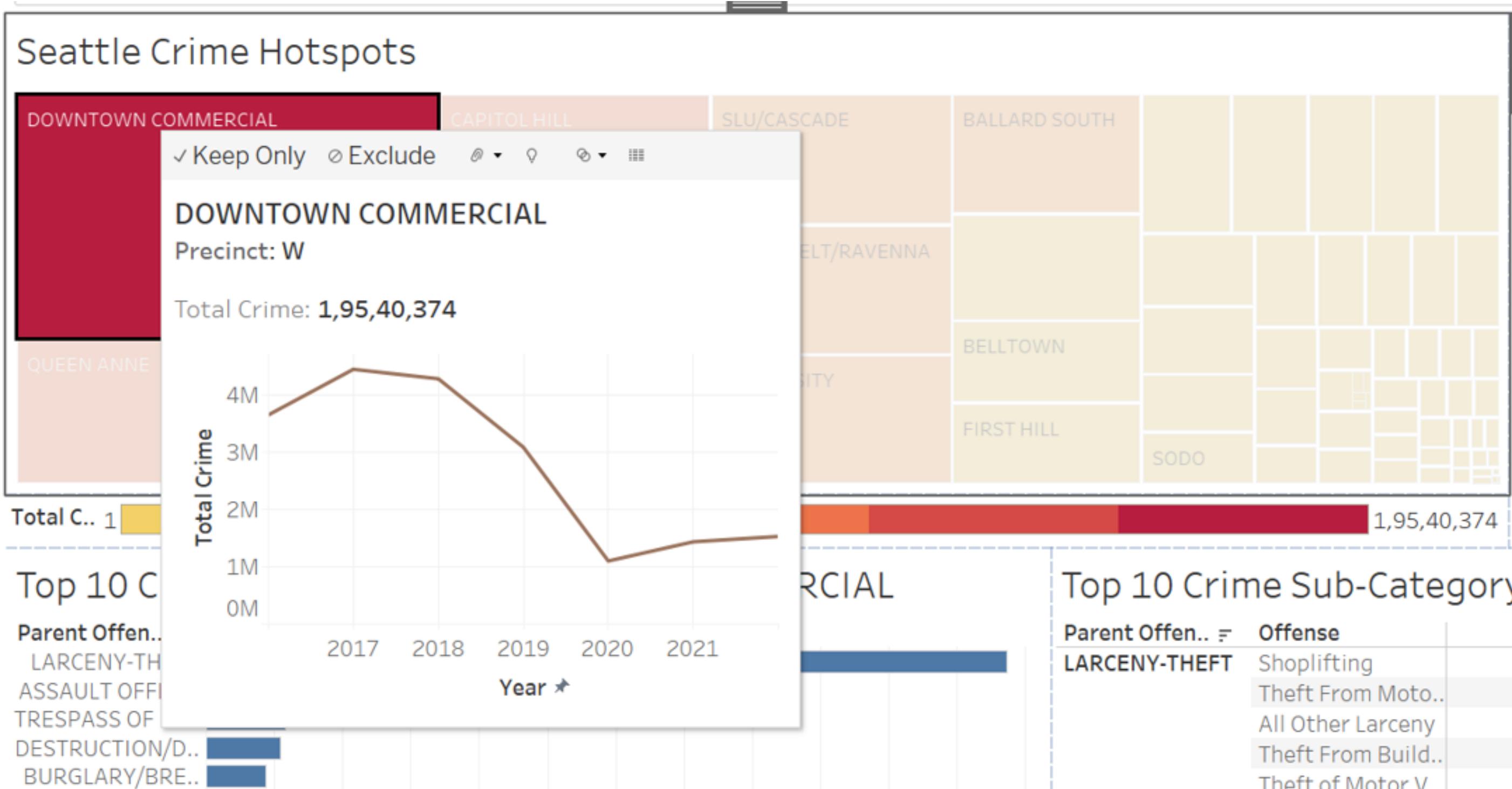
Parent Offense Group	Offense	Count
LARCENY-THEFT	Shoplifting	12,761,828
	Theft From Motor Vehicle	29,269,390
	All Other Larceny	8,460,083
	Theft From Building	1,316,453
	Theft of Motor Vehicle Pa..	2,579,019
	Pocket-picking	26,453
	Purse-snatching	675
	Theft From Coin-Operated..	344
BURGLARY/BREAKING&E..	Burglary/Breaking & Ente..	18,582,114
ASSAULT OFFENSES	Simple Assault	8,182,934
	Aggravated Assault	2,036,547
	Intimidation	1,605,396
DESTRUCTION/DAMAGE/..	Destruction/Damage/Van..	9,456,105
MOTOR VEHICLE THEFT	Motor Vehicle Theft	4,288,970
TRESPASS OF REAL PROP..	Trespass of Real Property	3,798,224
FRAUD OFFENSES	Credit Card/Automated T..	304,223

Link:

[https://public.tableau.com/app/profile/siddharth.purohit/viz/Final Datathon/SeattleCrimeRate](https://public.tableau.com/app/profile/siddharth.purohit/viz/FinalDatathon/SeattleCrimeRate)
(geo hotspot-treemap)

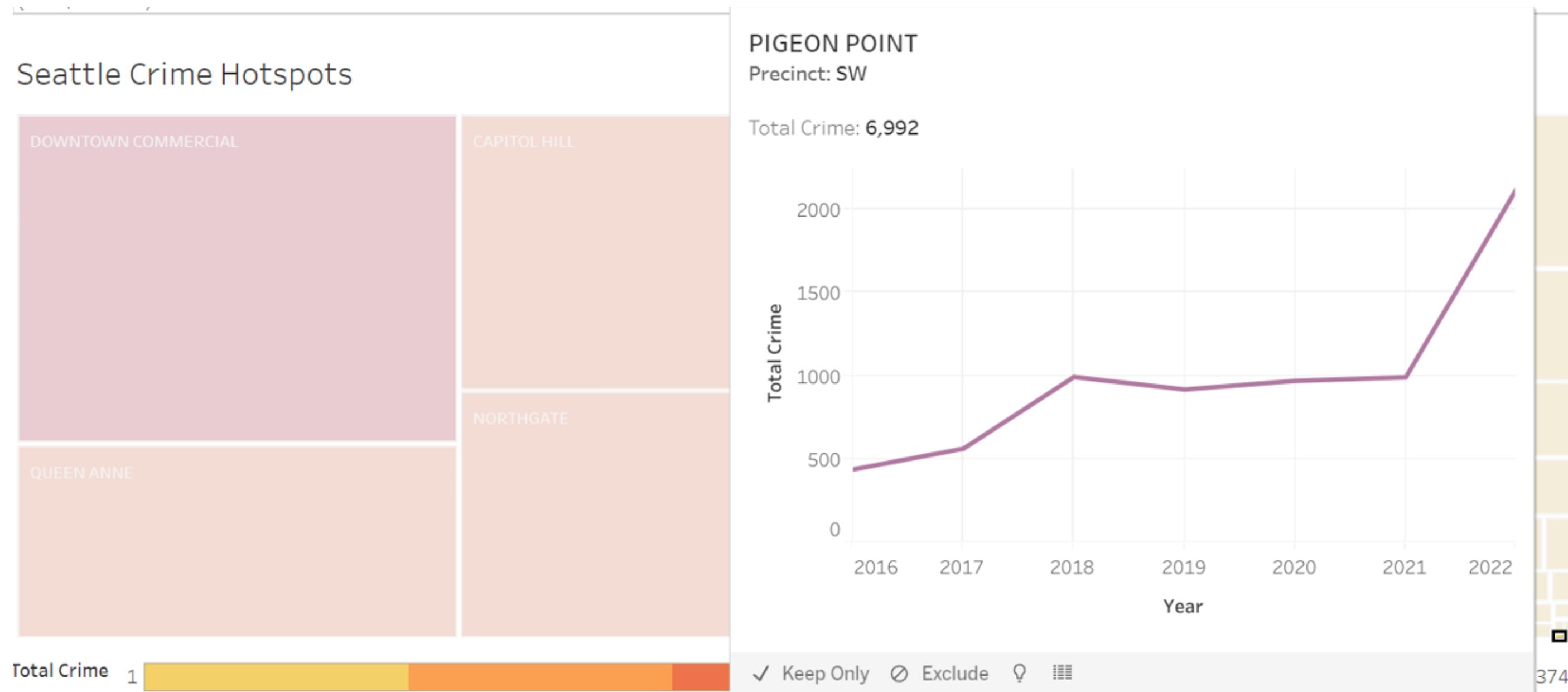
Insights - Neighborhoods with high crime rates are trying to reduce or are staying stable

Example: <https://public.tableau.com/app/profile/siddharth.purohit/viz/FinalDatathon/SeattleCrimeRate> (unsafe neighborhood)



Insights - Neighborhoods with low crime rates are seeing significant increase in the recent years

Example:



Top 10 Crime Category : PIGEON POINT

Parent Offense Group 🔍

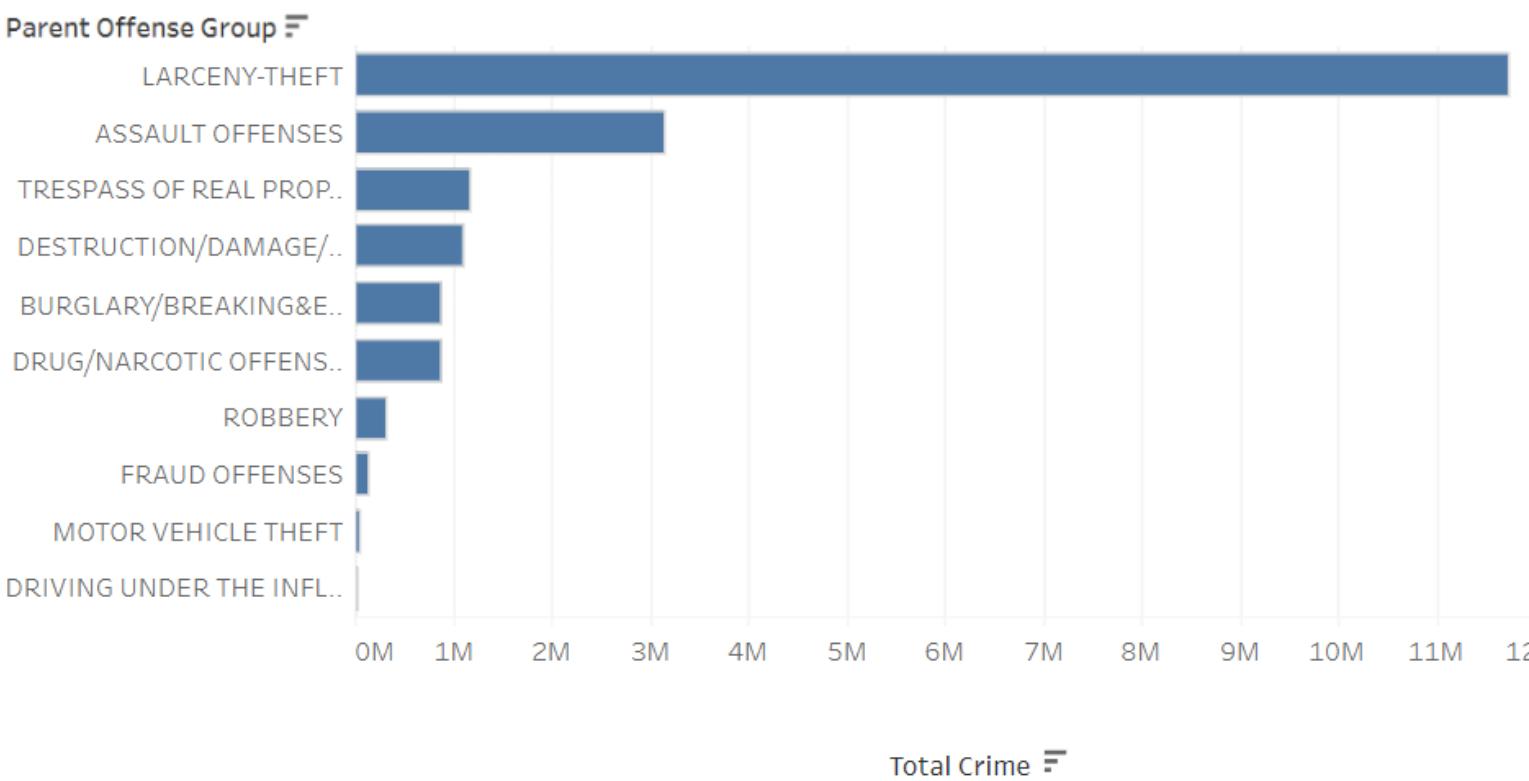
Top 10 Crime Sub-Category

Parent Offense Group 🔍 Offense

LARCENY THEFT Theft from Motor Vehicle

**Insights -
Mostly the
crimes in
the unsafe
areas are
non-life
threatening
and not
against
people**

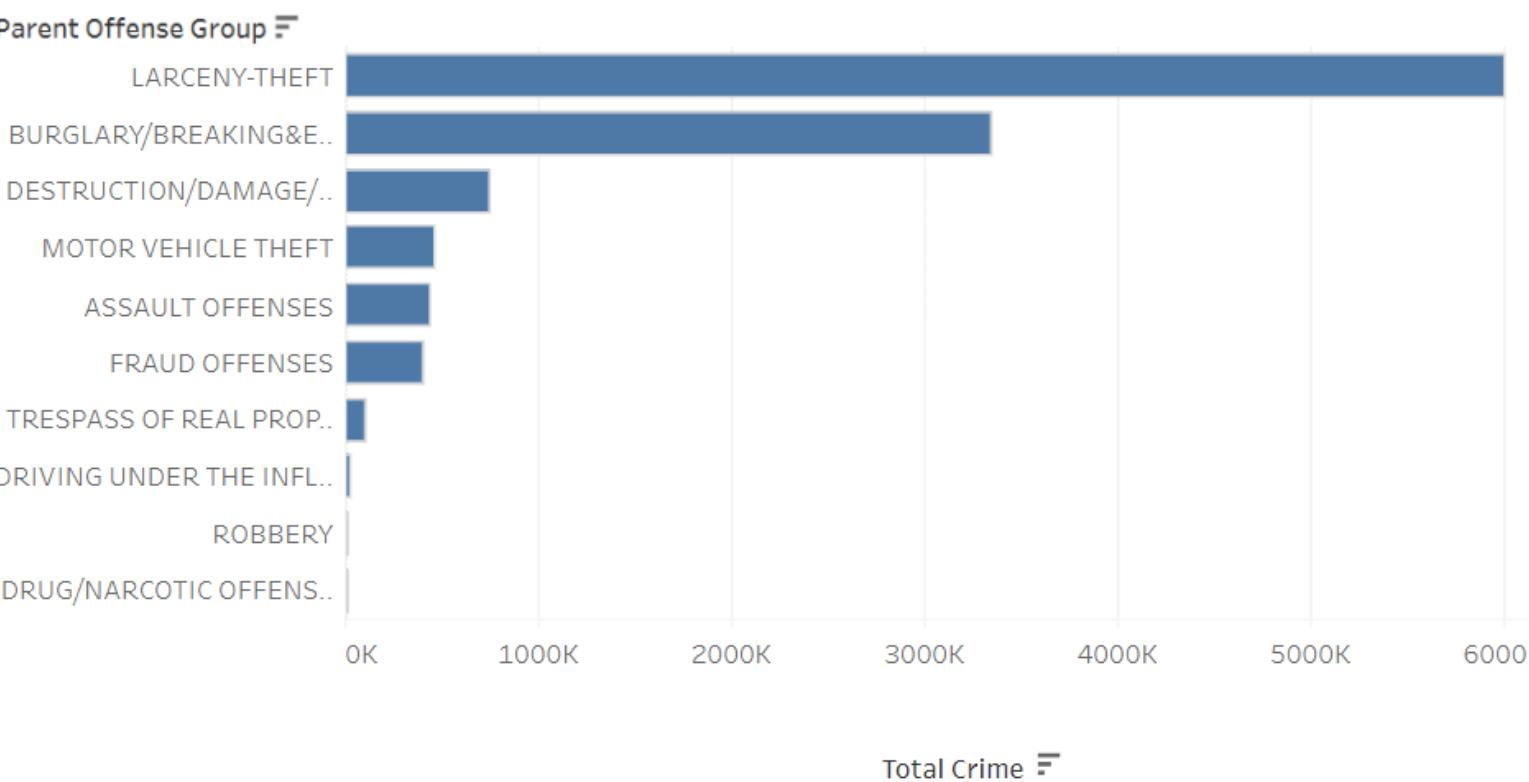
Top 10 Crime Category : DOWNTOWN COMMERCIAL



Top 10 Crime Sub-Category

Parent Offense Group	Offense	Total Crime
LARCENY-THEFT	Shoplifting	5,524,959
	Theft From Motor Vehicle	5,017,825
	All Other Larceny	845,768
	Theft From Building	310,793
	Theft of Motor Vehicle Pa..	17,886
	Pocket-picking	10,336
	Purse-snatching	224
	Theft From Coin-Operated..	2
ASSAULT OFFENSES	Simple Assault	2,415,695
	Aggravated Assault	453,489
	Intimidation	276,477
TRESPASS OF REAL PROP..	Trespass of Real Property	1,159,569
DESTRUCTION/DAMAGE/..	Destruction/Damage/Van..	1,092,357
BURGLARY/BREAKING&E..	Burglary/Breaking & Ente..	873,547
DRUG/NARCOTIC OFFENSES	Drug/Narcotic Violations	864,935
	Drug Equipment Violations	7,670
ROBBERY	Robbery	318,507

Top 10 Crime Category : QUEEN ANNE



Top 10 Crime Sub-Category

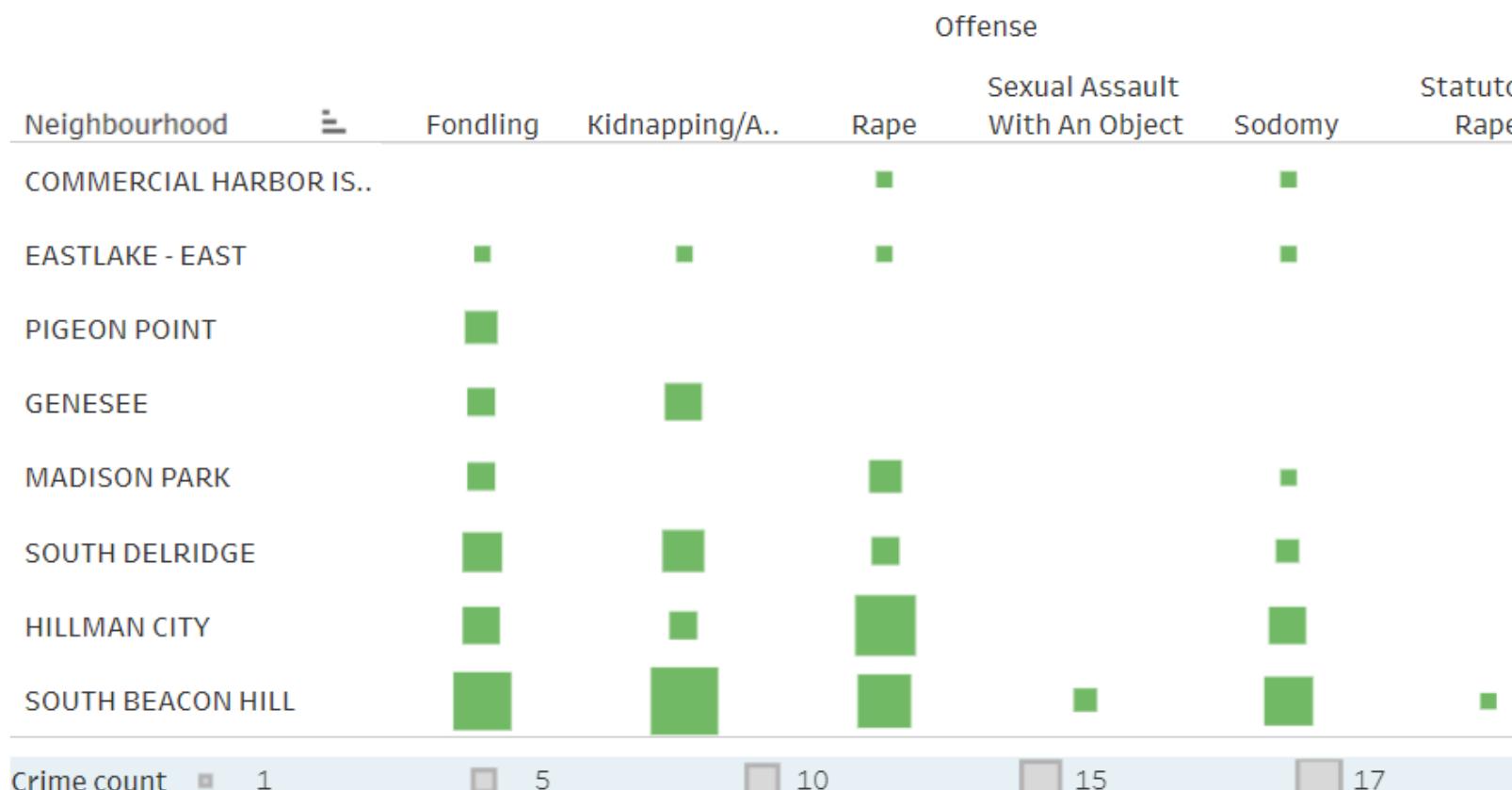
Parent Offense Group	Offense	Total Crime
LARCENY-THEFT	Shoplifting	99,161
	Theft From Motor Vehicle	4,642,075
	All Other Larceny	820,055
	Theft From Building	82,277
	Theft of Motor Vehicle Pa..	354,419
	Pocket-picking	430
	Purse-snatching	15
	Theft From Coin-Operated..	44
BURGLARY/BREAKING&E..	Burglary/Breaking & Ente..	3,345,076
DESTRUCTION/DAMAGE/..	Destruction/Damage/Van..	749,073
MOTOR VEHICLE THEFT	Motor Vehicle Theft	454,119
ASSAULT OFFENSES	Simple Assault	306,835
	Aggravated Assault	67,200
	Intimidation	67,043
FRAUD OFFENSES	Credit Card/Automated T..	21,349
	False Pretenses/Swindle/..	10,445
	Identity Theft	354,049

Neighborhoods safety for women and children

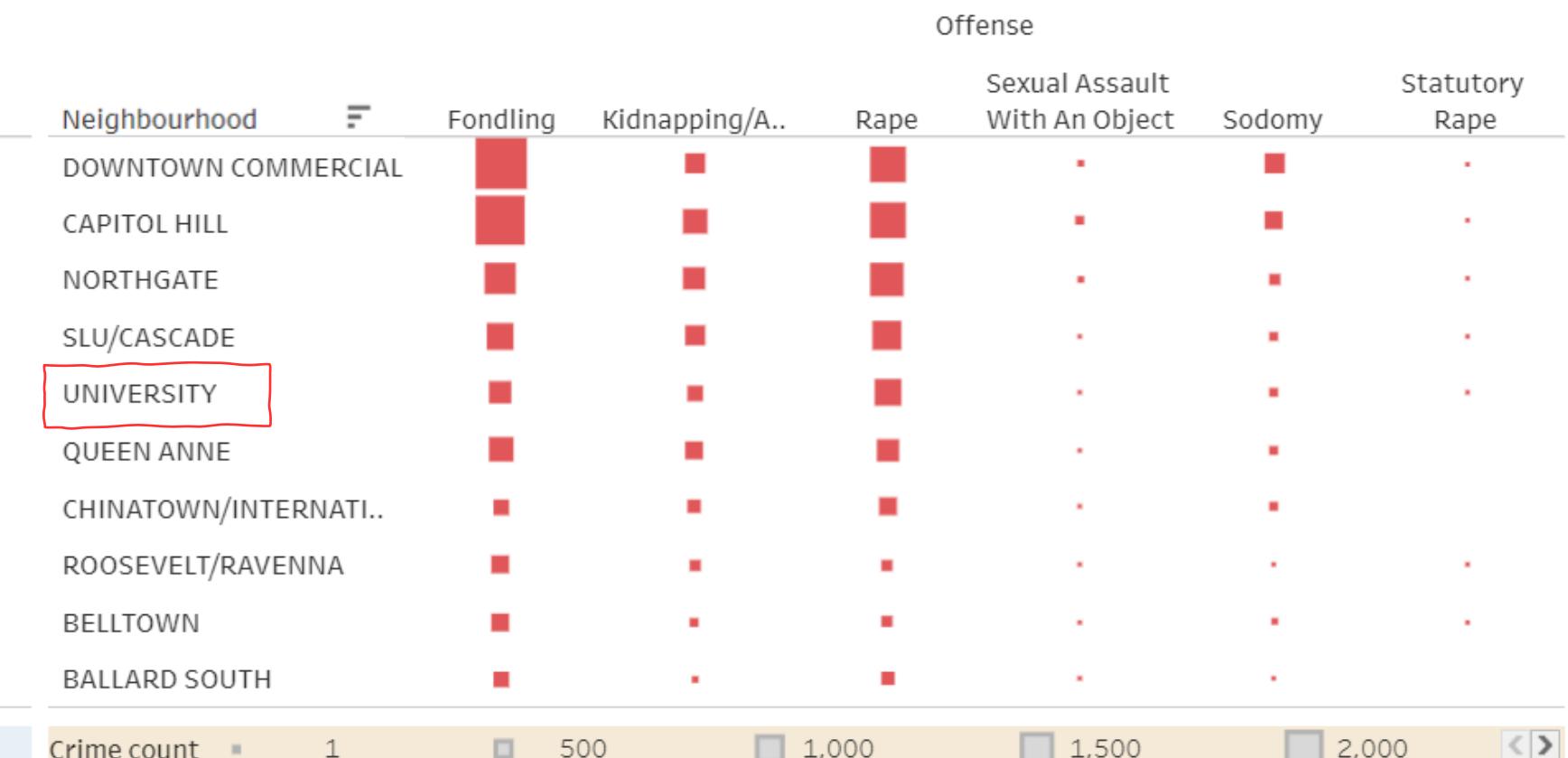
Neighborhood Safety for Women & Children

Based on 2022 data

10 Most Safe Neighborhood



10 Least Safe Neighborhood



Relevant crime categories - Fondling, Kidnapping, Rape, Sexual Assault, Sodomy

Link: <https://public.tableau.com/app/profile/siddharth.purohit/viz/FinalDatathon/SeattleCrimeRate>
(geo hotspot-treemap - bottom)

Thank you!

