Analysis of Crime Reports in Austin, TX

Danny Sheng, Andrew Kim, Rahul Lal, Anirudh Gupta

2024-10-27

Table of contents

Introduction Of Project And Dataset

Crime is and has been prevalent across the country for centuries. But what has the trend been over time as far as how much crime is happening? Are there certain areas where crime is more common? To take it a step further, do areas with high crime impact the quality of education in an area, impacting the students that go to that school? These were all questions that we had when looking at the topic of crime. As for why we chose to look at Austin specifically, we believe that with the rapid growth Austin has had over the past few decades, we believe we could find interesting trends and patterns as this growth occurred.

Our primary dataset has crime data reported since 2003. Some of the variables include the types of crime, the date/time the crime happened, and the location type (house, stores, etc.). Additionally, we have access to the districts where the crimes occurred and whether or not the crimes were closed. These data points will be integral in our analysis of crime rates in different parts of Austin against each other and across time, as well as whether some districts of Austin have a greater proportion of unsolved cases.

We also have a secondary dataset that provides insight into education quality in elementary, middle, and high schools across Austin. This insight, in conjunction with our primary dataset, will hopefully help us answer the question of whether high crime in areas negatively impact the education quality in an area, such that the students in those schools are negatively impacted.

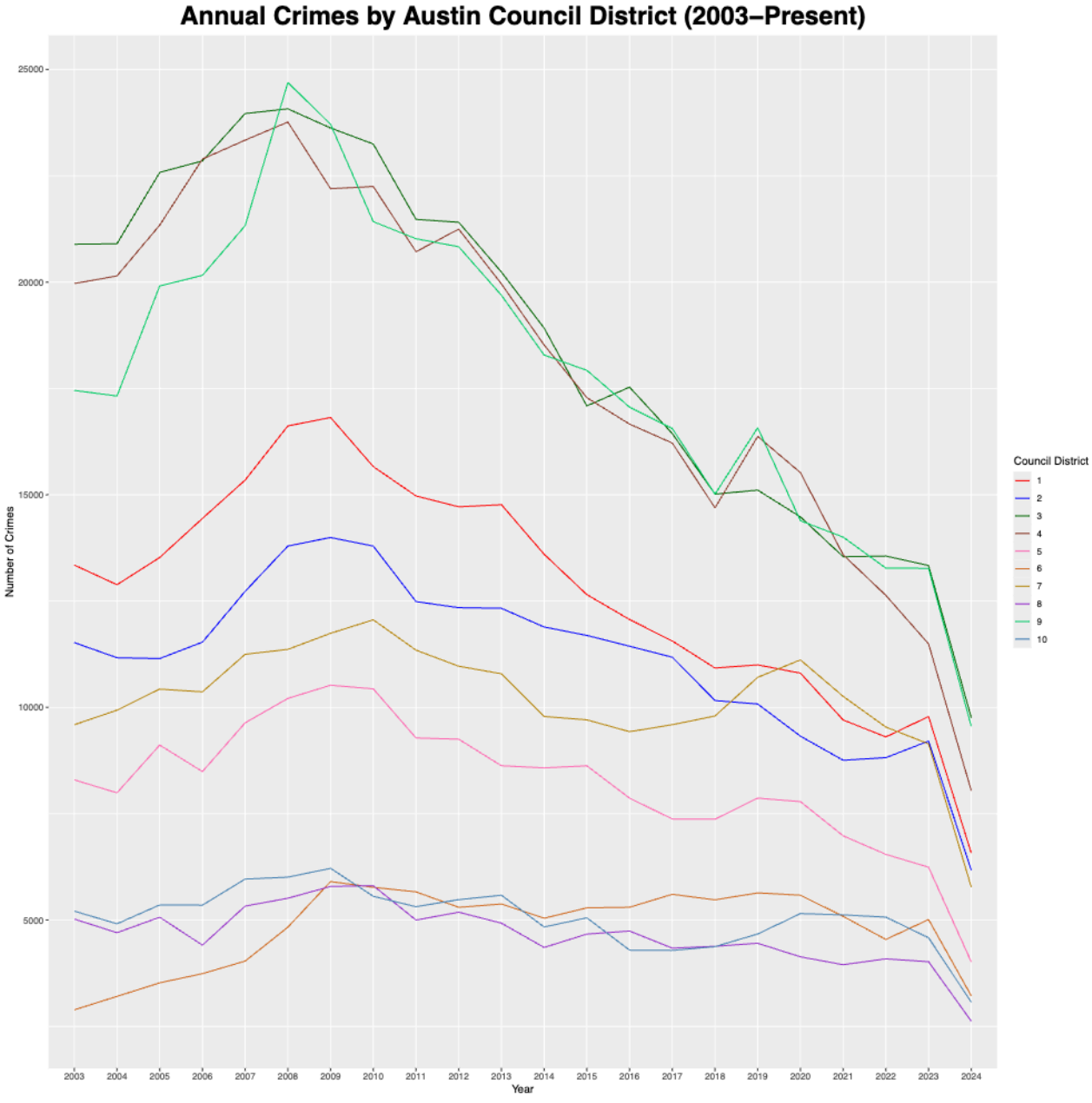
Lastly, we narrow in on the case study of family violence to tackle the incidence and potential causes of a specific type of crime. We seek to answer whether family violence occurs more during different times of the year and locations, as well as how local school quality might predict it.

Overview of Austin Crime

We first created a few plots to visualize our primary dataset, allowing us to grasp overall trends in Austin crime data that will be useful when conducting our more in-depth analysis.

Plot 1

The line graph below maps the trend of the total number of crimes by year in each council district in Austin from 2003 to the present.

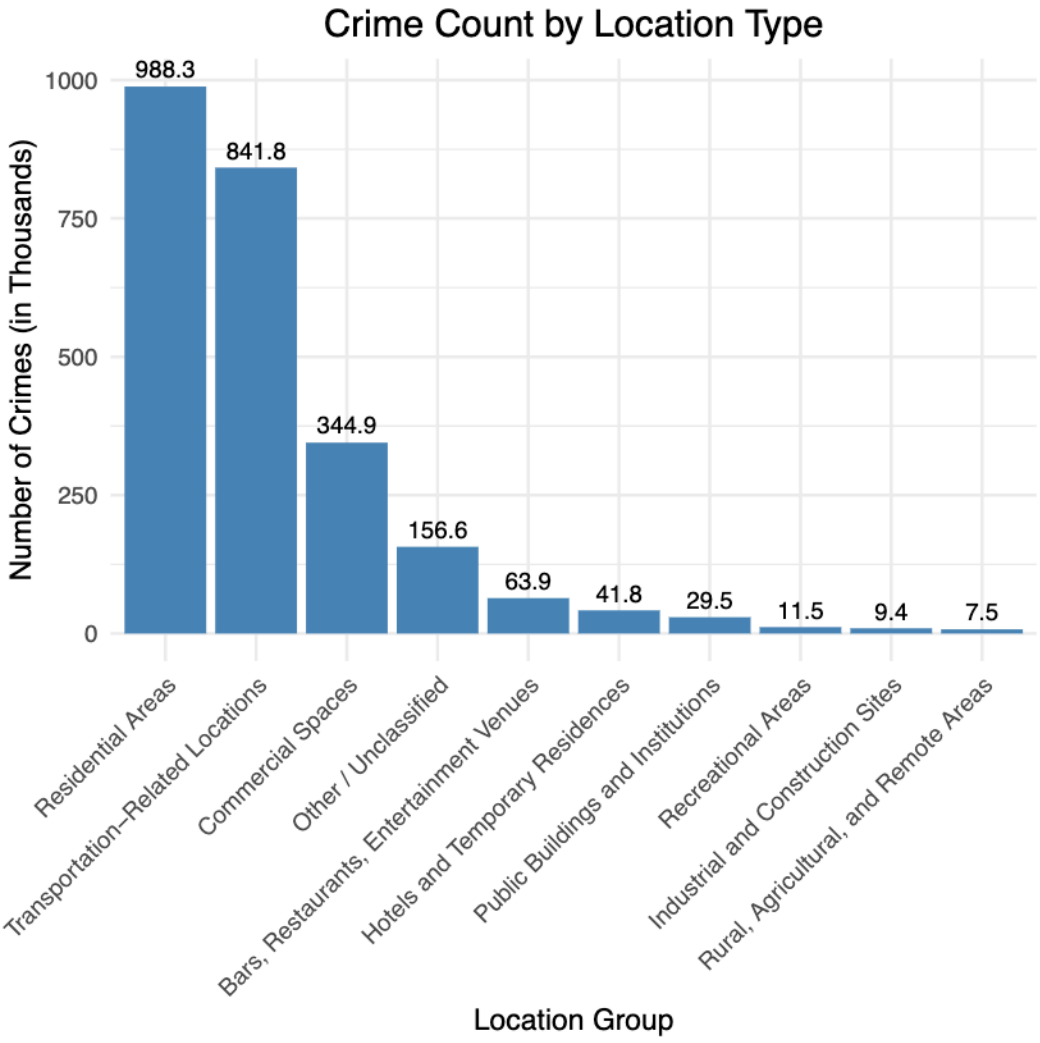


Overall, each council district has a general downward trend in crime. Crime appears to have peaked around 2008-2009 and has dropped since then for all districts, especially districts 3, 4, and 9, which appear to have seen the largest decrease in total crime. Despite these decreases, districts 3, 4, and 9 remain the districts with the highest number of total crimes. The three districts with the most crime are the most centrally located

within the city and contain a majority of the city's attractions. District 9 is home to Austin's downtown, the University of Texas at Austin, the 6th Street bars, and West Campus. District 3 is home to the South Congress area and East Austin, while District 4 is home to North Loop and Hyde Park, popular shopping areas within the city. Thus, these areas may be more prone to higher crime rates, especially petty crime, due to a combination of dense population, nightlife, and concentration of businesses and attractions.

Plot 2

In addition to looking at crime count by district over time, we also believed its important to look at where these crimes are happening. Specifically, we were curious if there were specific types of locations with an increased level of crime. One challenge we encountered was that there were too many different types of locations, resulting in many locations with relatively low frequencies. As a result, we consolidated them into ten different types of locations. The plot is shown below.



Two locations stood out from the rest in terms of the large amount of crime reported there: stores and terminals. Stores make sense when considering the vast numbers of stores, especially those that are high-traffic, like shopping malls and grocery stores. Terminals refer to transportation-related locations. The high frequency of crimes in this location type also makes sense, considering that transportation includes train/bus terminals and parking lots.

Other locations, like homes and fields, had a subdued amount of crime, with fields recording roughly 7,500 crimes since 2003. One location that we expected to be higher was bars because people can get disoriented

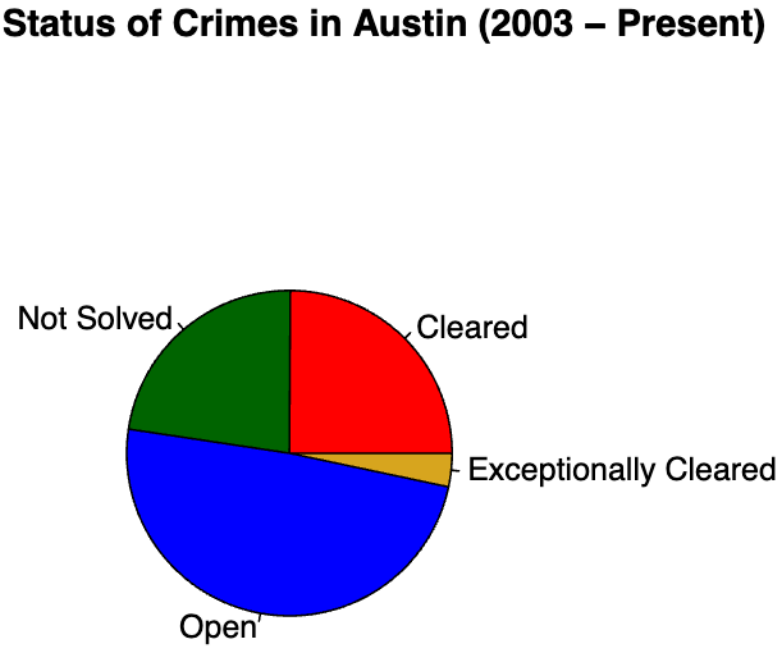
and agitated when intoxicated, but to our surprise, only 64,903 crimes have been recorded at bars since 2003.

These results are significant as they show locations where an increased police presence could be beneficial in reducing crime. For example, based on our plot, we can conclude that increased police presence around stores and terminals has the potential to significantly decrease crime in Austin. On the other hand, having large amounts of police in neighborhoods and fields, although both important locations, may not be the best use of police resources if the goal is to reduce crime in Austin.

Solve Rates

Plot 3

Reducing the frequency of crime in Austin is important, but we were also curious to see of the crimes reported, how many have been cleared and how many haven't. The pie chart below uses Austin's crime data from 2003 to now to plot the distribution of crimes based on their clearance status. Crimes are sorted into four separate groups: Cleared, Open, Not Solved, and Exceptionally Cleared which essentially means that the case was not solved due extra factors that prevent arrest or prosecution.

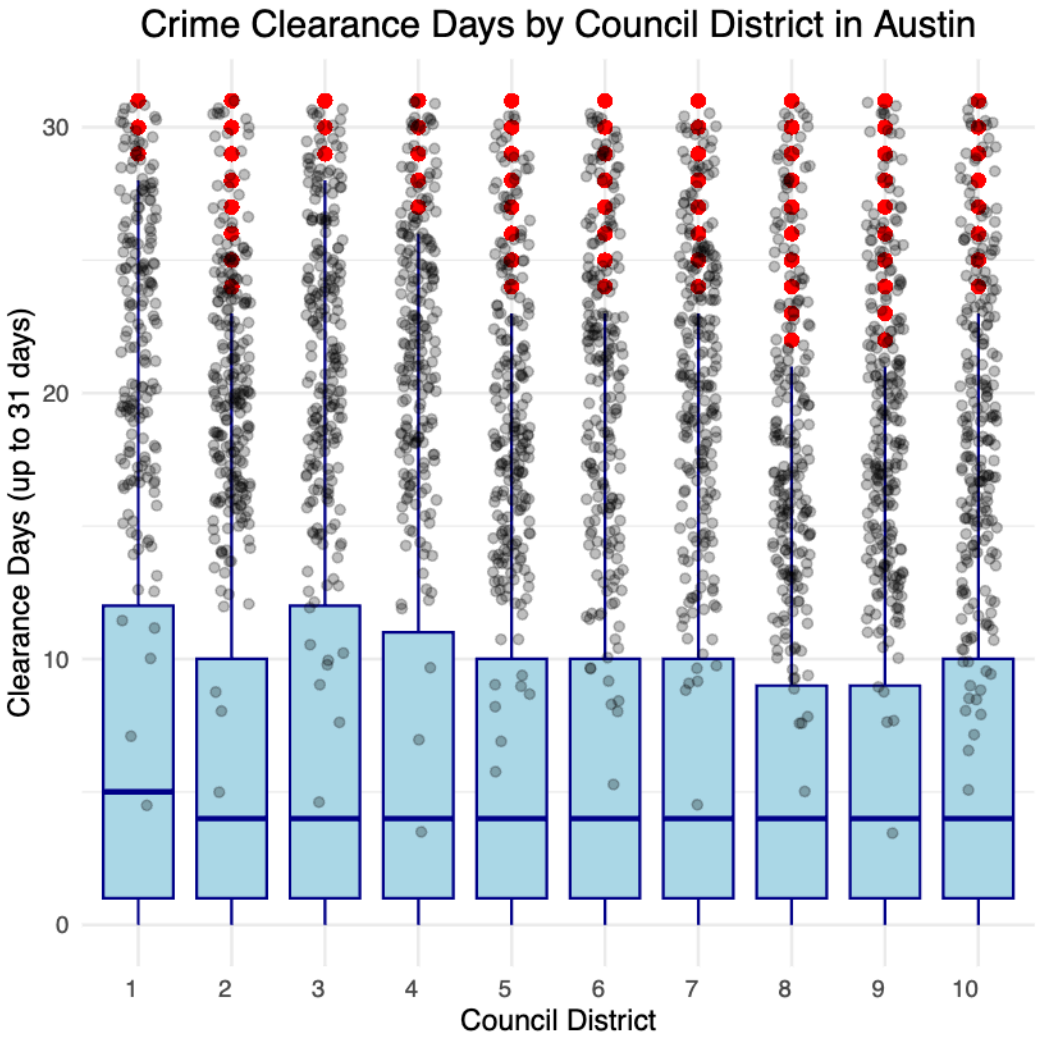


The pie chart suggests that more resources may be needed to close cases within the Austin community. This is because the blue and yellow sections, which are partitioned for cases that have either not been solved or are currently under investigation, make up over half of the pie chart. Cases that have not been solved make up almost half of the cases alone. Additionally, the red section, which stands for exceptionally cleared cases, tends to typically be situations where the officer knows who the perpetrator is but cannot proceed due to legal factors, making up about another eighth of the cases. Thus, very few cases have been resolved either from an arrest or having the legal process completed.

Plot 4

After looking at the results of the pie chart, specifically the large portion of crimes not yet cleared, we were curious what the solve rates looked like on a district by district basis. Namely, we were interested in looking at whether different districts in Austin have different clearance rates. If there are significant differences, that may be an indicator that there is too much crime in that area and/or the severity of the crimes are greater.

Based on this question, we decided to create a boxplot with the councils on the horizontal axis and the number of days it took to clear the crime on the y axis. It is worth noting that there are far more upper outliers that were cropped out of the graphic as there are crimes that took over 10 years to be solved. In addition, crimes that haven't been solved yet are not shown as the case is still open.

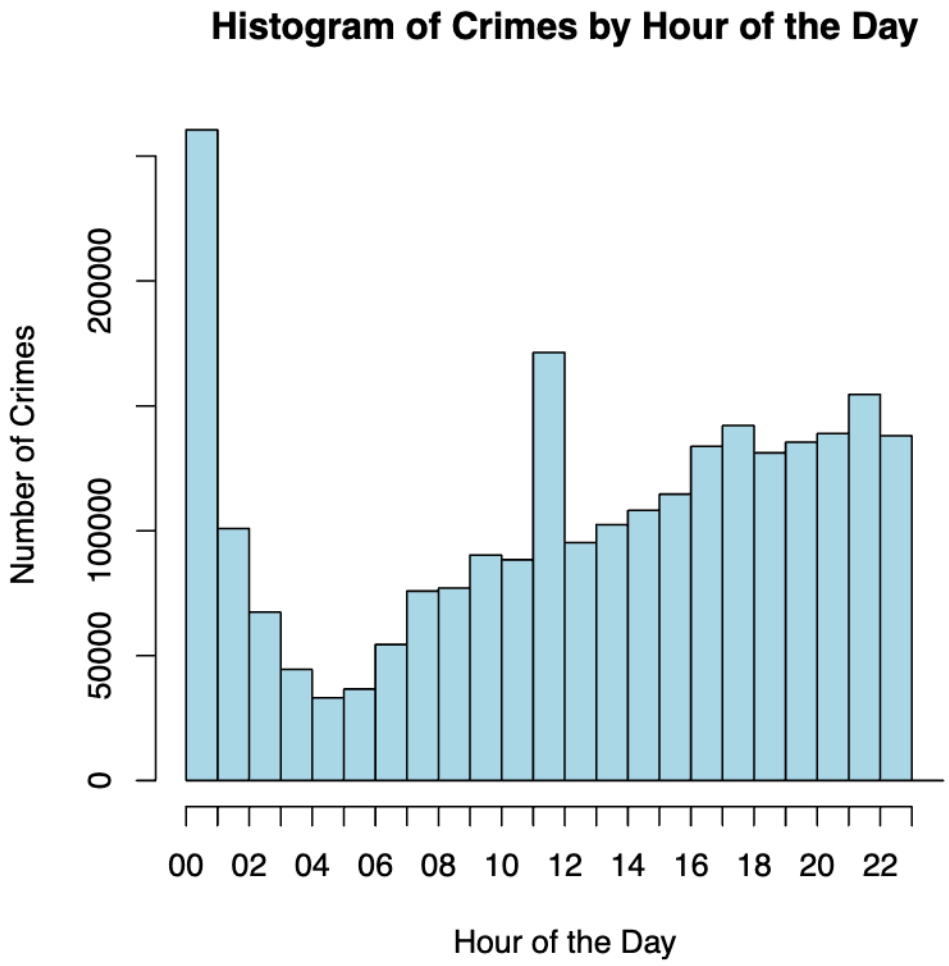


When looking at the boxplot, there is one striking similarity. The median time for a crime to be solved was relatively similar across the districts, at around four days. This result is particularly interesting when considering that districts 3, 4, and 9 have the largest amount of crime. Based on the boxplot, we can conclude that even the "safer" districts take similar amounts of time to solve the crimes. Perhaps there is red tape that the police department must follow that takes roughly four days. The IQRs of the boxplots for each district have a little more variability but are roughly eight days. Despite the large number of outliers, most of which are cropped out, one conclusion that can be made is that over 75% of the crimes cleared in Austin since 2003, were closed within two weeks of being opened.

How Can We Combat Crime in Austin

Plot 5

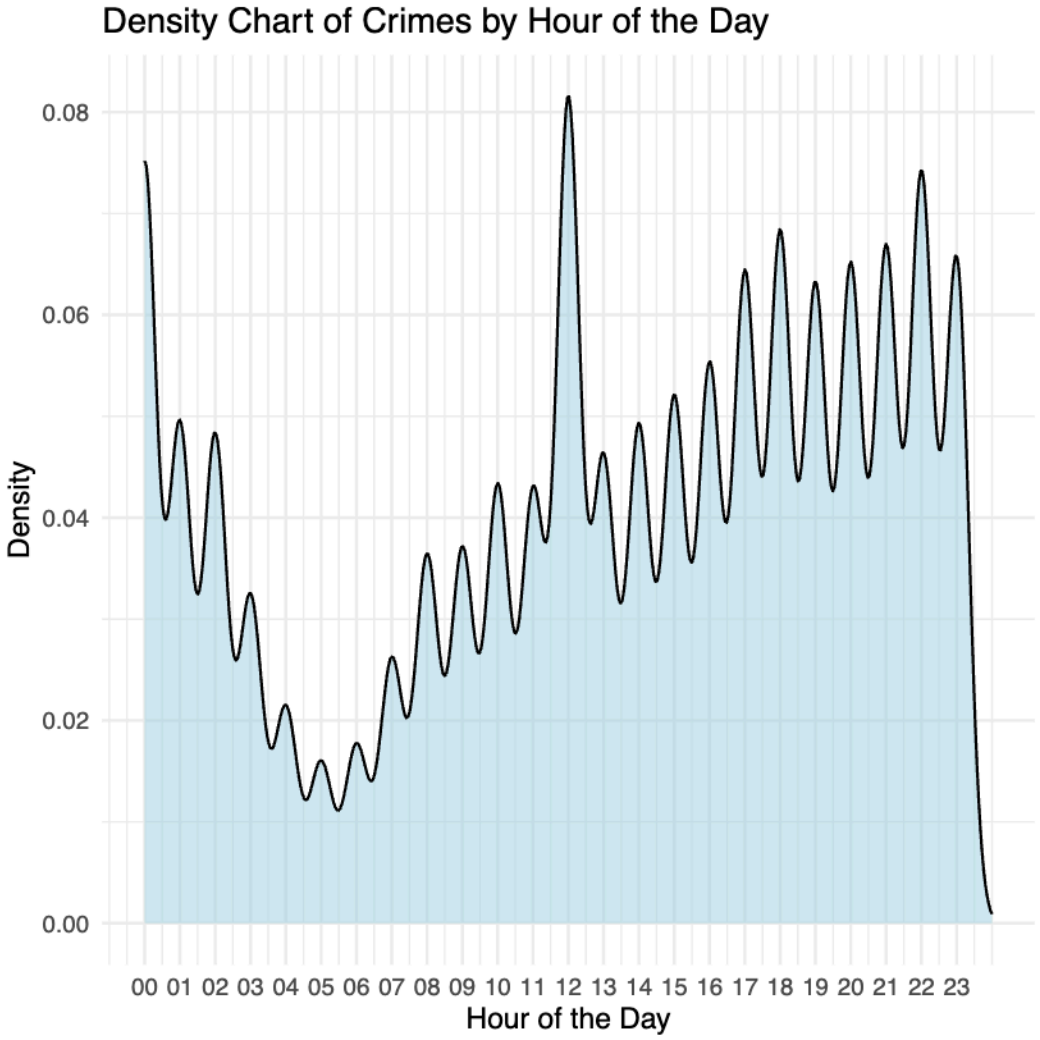
Now that we have more context as to the crime scene in Austin, we wanted to start to answer the question of how crime can be combatted. This is the histogram of crimes by hours of the day. Each bar reflects the number of crimes that happened in that hour. This histogram highlights the change of numer of crimes over time during a day, especially in an hour. The numbers of crimes are aggregated in the manner of hours, so that we can see cearly in which hour period the crimes happen the most.



In this histogram, we firstly modified the time into a more readable format. In the original dataset, for example, "2030" is used for "20:30", which is not good for showing on the histogram. The histogram shows how crimes in Austin are spread throughout the day. The highest number of crimes happens right around midnight, with another noticeable increase around noon. This makes sense, as midnight is often when nightlife is active, while noon represents busy midday hours. Early morning hours, like around 4:00 AM to 6:00 AM, have fewer crimes. This could be because there are fewer people out and about, reducing the chances of crimes happening. As the day goes on, there's a steady increase in crimes from morning to evening, which aligns with people getting out for work, errands, or social activities.

Plot 6

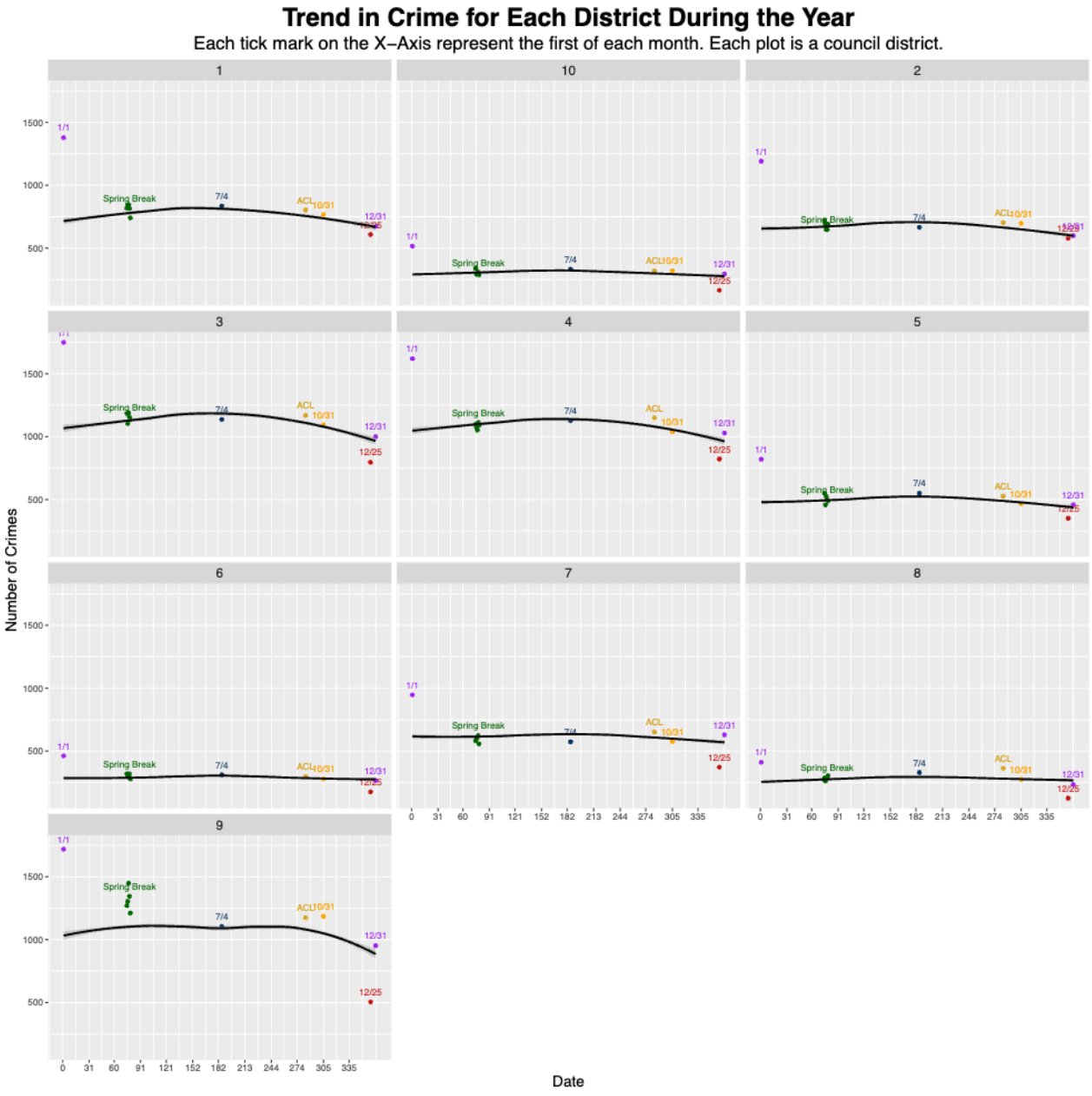
This is a density chart of crimes by hour of the day. The horizontal axis represents the 24-hours in a day, and each space is one hour. The vertical axis represents the density of the number of crimes. This graph is a more statistical version for the histogram, which provides us the density distribution.



The density plot provides a smoothed-out view of how likely crimes are to happen at different times. It shows the same two peaks around midnight and noon, giving a clearer sense of the trend throughout the day. It also shows the drop-off during early morning hours more smoothly, then picks up again as the day progresses. Compared to the histogram, while the histogram shows the actual number of crimes by hour, the density plot reveals the general pattern of when crimes are more likely to happen. Together, they confirm the same story: crime peaks late at night and midday, drops early in the morning, and steadily increases into the evening.

Plot 7

We've looked at crime throughout the day but we now looked to see if crime occurs more frequently at different times of the year. In this plot, we plotted the trend line of how crime fluctuates throughout the year and separated the graph into 10 different plots in a grid with each plot representing a council district in Austin. We chose to highlight specific dates where crime may be higher in the city due to increased activity and tourism which includes New Year's, Spring Break, July 4th, Austin City Limits, Halloween, and Christmas.



The plot above suggests that crime in the city generally tends to peak during the summer and dips down during the winter months. This is likely due to the fact that people tend to spend more time out due to weather being warmer and hours of the day being longer. However, in council districts 6, 8, and 10, the trend of crime tended to have much less fluctuations compared to other districts. District 6, 8, and 10 are on the west side of Austin and are home to more of the upscale housing in the city and are less integrated within the city. Thus, tourists have less of a reason to travel to these areas as they are further away and there is likely less to do. Additionally, all districts showed a significant increase in crime in New Year's Day as people tend to spend the holiday drinking with friends and family and in some cases going to the bars. Because people are out late and may not have a plan to get back, this can lead to people making poor decisions such as driving home while drunk. Conversely, crime was below the trend line for all districts during Christmas as people tend to spend the holiday at home with their families, hence fewer people are spending time out. Looking at District 9, crime was well above the trend line during Halloween and Spring Break. Both Districts 8 and 9 saw increases in crime during Austin City Limits. These results are a reflection of the fact that District 9 contains Austin's downtown and UT-Austin's campus while District 8 contains Zilker Park where

Austin City Limits is hosted. Hence, during these specific holidays, more people from out of town and likely to travel to District 8 and 9 to celebrate and partake in these events. Because of the higher concentration of people within the areas, crimes such as pick pocketing are more likely to take place.

Root Up Via Education

Plot 8

An alternative approach to reducing crime in Austin is by addressing the causes for criminal activity far before it occurs, such as weaker education limiting students' future economic opportunities. To test this theory, we model a linear regression of annual crime counts by Austin council district between November 2014 to 2022 (when the council districts were in effect) on the average accountability rating of school quality awarded to the district's schools by the Texas Education Agency (standardized on a 0 to 4 scale). Given council districts are drawn with approximately equal populations, we assume population is held constant. A significant coefficient on the school rating variable indicates that school quality is related to annual crime rates in Austin.

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	26680.0517	8309.2686	3.2109	0.0148
Mean School Rating	-5783.6506	2556.9028	-2.2620	0.0582
Number of Schools	-31.8777	131.6311	-0.2422	0.8156

As anticipated, the coefficient of -5,783.6506 for the school rating variable is marginally significant at p = 0.0582. This suggests to some degree that, holding all else equal, for each increase in school rating (e.g., F to D, B to A), annual crime counts in the district are expect to fall by approximately 5,783 reports, a substantive drop from a baseline of about 26,680 crimes per year (the intercept of the regression representing an average F school rating). Additionally, the number of schools variable was not significant at a p-value of 0.8156. Although not conclusively causative, we confirm that higher school quality is somewhat associated with a precipitous decrease in crime rates, providing support for a ground-up approach to fighting crime: enhancing the quality of public schools to expand the economic opportunities of its students.

Plot 9

Another area that interested us was the distribution of types of crime occurring within each council district. Hence, we chose to use polar coordinates to create a radar graph of the number of crimes of each category occurring within each council district. Additionally, we wanted to look into how access to education impacted crime rates.

Indexed Violent Crime Values in Families vs. Average School Rating and Levels

Higher grade is better for schools but worse for crime. WMSR stands for Weighted Mean School Rating.

The type of crime that occurs the most in every district is theft with District 9 having the most theft. These trends are not surprising considering property crime tends to be more common than violent crime across the country. Additionally, violent crime is far less common as murder, rape, and robbery tend to fall close to the center of the graph for all districts. While burglary, aggravated assault, and auto theft are far less common than theft, they tend to more common than violent crime with District 1, 3, and 4 having higher rates of these specific kinds of crime as these point reach close to the first circle on the radar. District 9 only has higher rates of burglary compared to other crimes. This is likely because District 1, 3, and 4 are all in East Austin which is still continuing to develop as people move into the city. Hence, relative to other parts of the city, crime is likely to be higher. Additionally, there was a trend between quality of schools and crime rates as District 1 had the highest rates of violent crime and was on the lower end of school quality while District 6, 8 and 10 had a higher grade for WMSR and lower rates of crime. However, there seemed to be less of a correlation between types of education present and crime rate. That being said, it does appear that council

districts with more high schools tended to have less crime and fewer elementary and middle schools.

Family Violence Case Study

Family violence is a particularly important category of crime for the city of Austin to combat from its root causes, as it may be difficult to police crime that is less often reported. To determine if the involvement of family violence influences the time between a crime's occurrence and its report, we model a linear regression of time to report a crime after it has occurred on the involvement of family violence. A significant coefficient on the family violence variable indicates that changes in time to report are related to the involvement of family violence.

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	454.0605	3.7589	120.7965	0
Family Violence Involved	120.7881	6.5742	18.3731	0
Hour of Day	-11.9699	0.2248	-53.2377	0
Is It Nighttime?	-179.9785	3.2256	-55.7976	0

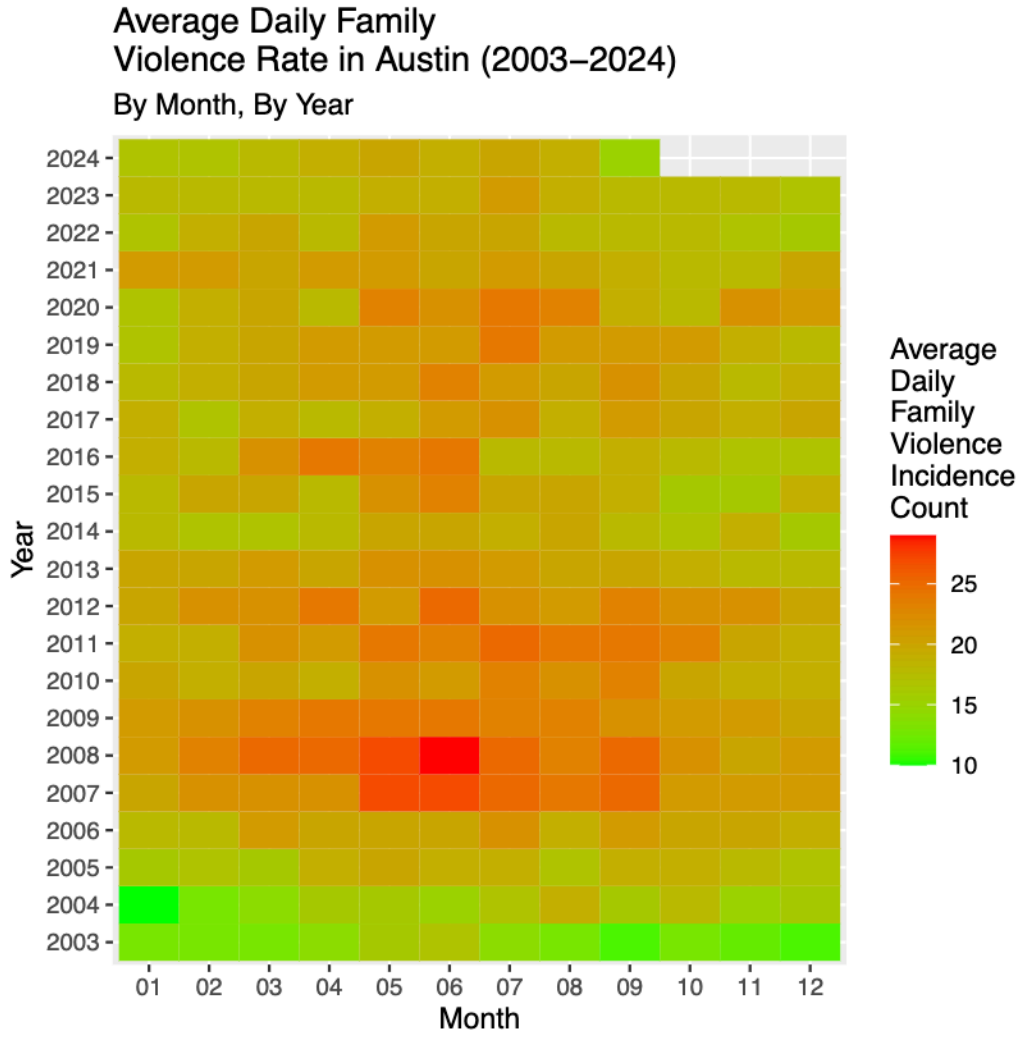
The coefficient of 107.8507 on level "Y" (indicating "Yes") for the family violence variable is statistically significant ($p < .001$). Holding all else equal, an Austin crime report that involves family violence is expected to take 120.788 hours longer to report than a crime report that does not involve family violence. The intercept, 454.0605, indicates a crime report that does not involve family violence takes an average of 454.0605 hours to be reported.

Additionally, the coefficient for the hour of the day is statistically significant ($p < .001$). Thus, as you increase the hour of the day (i.e. go from 8 AM to 9 AM), the hours taken to report the crime is expected to decrease by 11.9699 hours. If the crime occurs at night (Between 8 PM and 6 AM), the crime is expected to take 179.9785 hours less than a crime that is reported during the day.

Although this linear model is relatively primitive, the significance and positive sign of the coefficient supports a plausible explanation that individuals are more hesitant to report crimes committed by family members they are close to, or when they are otherwise in dangerous domestic situations that pose a physical threat to them if the perpetrator was made aware of the report.

Plot 11

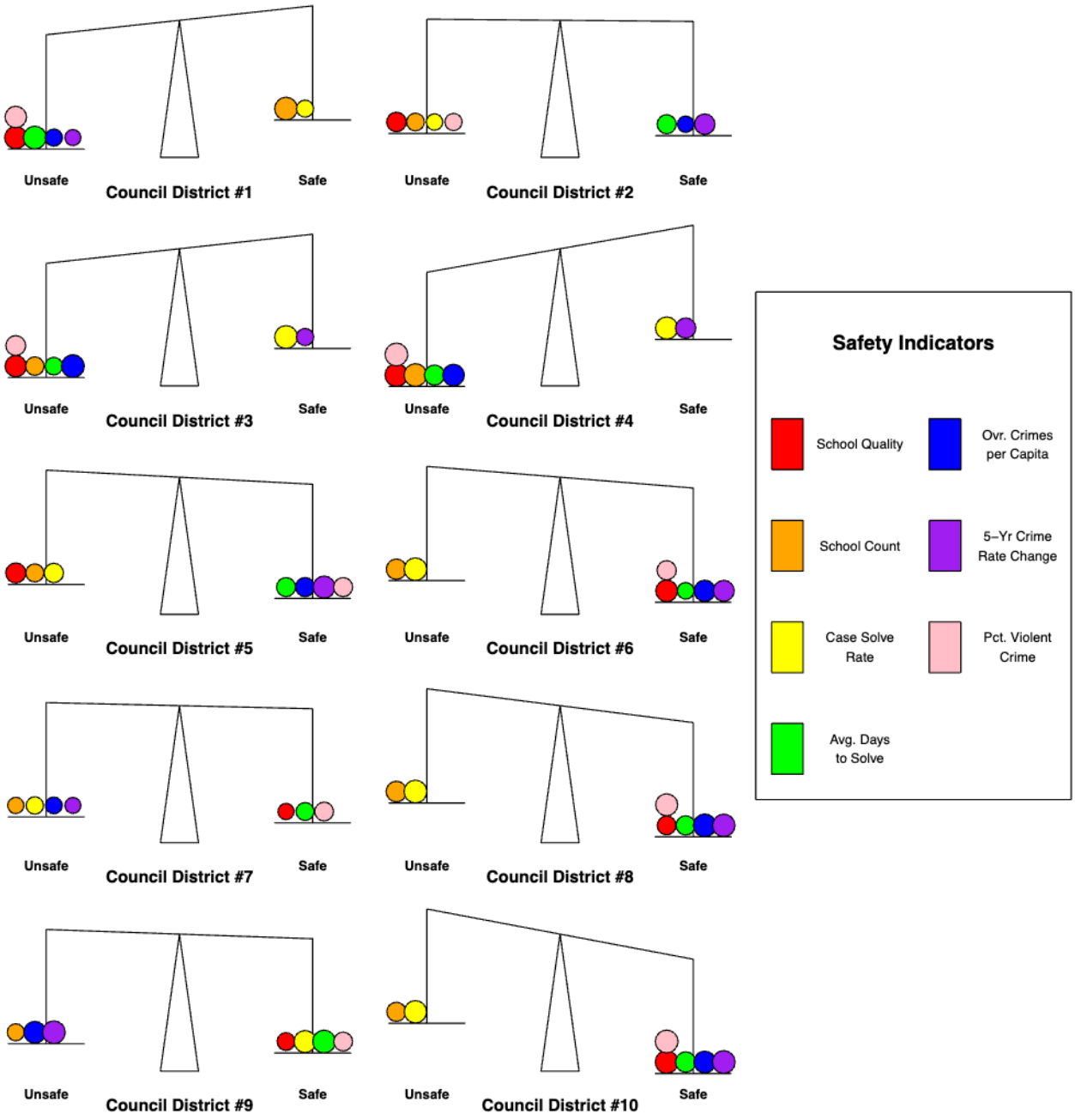
To better understand when family violence crimes occur, this heat map illustrates the average daily rate of reported family violence crime occurrences by month, by year in Austin. Each tile represents a specific month of a specific year and ranges from light green to red as the rate of family violence reports in our dataset increases. The plot provides more detail on whether these crimes occur at higher frequencies in certain months/times of year than others, as well as how crime rates have shifted between 2003 and 2024.



The plot is pronouncedly more red in the 2007 to 2009 range (the maximum crime volume of approximately 25-30 reports per month), coinciding with high crime report rates during this period we found earlier. Notably, there were anomalously fewer reports from 2003 to 2005 (around 15 per month). Assuming the cause wasn't a precipitous increase in family violence, one feasible explanation is an Austin PD infrastructural change after 2004 that encouraged or mandated greater report and documentation of family violence crime. Overall, the heat map illuminates a tendency across most years for family violence crime volume to be higher in summer months (05/May to 08/August) than in winter months (11/November to 02/February). This finding supports the general phenomenon that violent crime rises in hotter weather, potentially as Austin residents are more likely to spend time outdoors, around other residents, and with leisure time/alcohol consumption. Another explanation is that children attending school or college are more likely to be home over these summer months, signifying a greater need for law enforcement proactivity in this period.

Killer Plot

To consolidate our findings on crime and education in Austin's districts, we created a comprehensive visualization summarizing all variables. This plot leverages standardized z-scores to compare districts relative to one another. The size of each circle corresponds to the magnitude of the z-score, highlighting the relative significance of each metric. A negative z-score indicates that the district is below average and considered less safe for that metric, while a positive z-score suggests greater safety.



We derived a few key insights from this powerful visualization. As public perception often suggests, Council Districts 1, 3, and 4 emerge as the least safe. This makes sense as District 1, 3 and 4 are all on the Northeast side of Austin which is known to be less developed than the west side of the city. However, as Austin has continued to expand in population, East Austin, particularly District 3, has undergone development. This is likely a contributing factor as to why the 5-year change in crime rate is on the positive side for Districts 3 and 4. Although different variables drive this conclusion for each district, the overall assessment aligns with community sentiment. Districts 5, 6, and especially 10 are ranked safer. This aligns with expectations, as

District 10 encompasses areas like Westlake and Tarrytown, known for being affluent and well-resourced. No district is entirely on one side of the safety spectrum. Every district scores above average in at least one metric, indicating that each area has strengths and weaknesses.

Conclusion

In this report, we tackled the topic of Austin crime, utilizing various statistical graphs and performing analysis on numerous variables. We identified notable trends, such as an increased frequency of crime in Districts 1, 3, and 4, alongside a general uptick in incidents around midnight and noon. We also examined the intersection between crime and education, finding some evidence that districts with lower education quality have higher crime rates, disadvantaging students in these areas even more.

Crime remains a critical topic as it impacts everyone's quality of life. We hope our findings provide valuable insights into Austin's crime landscape and encourage actionable steps by Austin Police and local policymakers to enhance community safety.