

Your Project Title

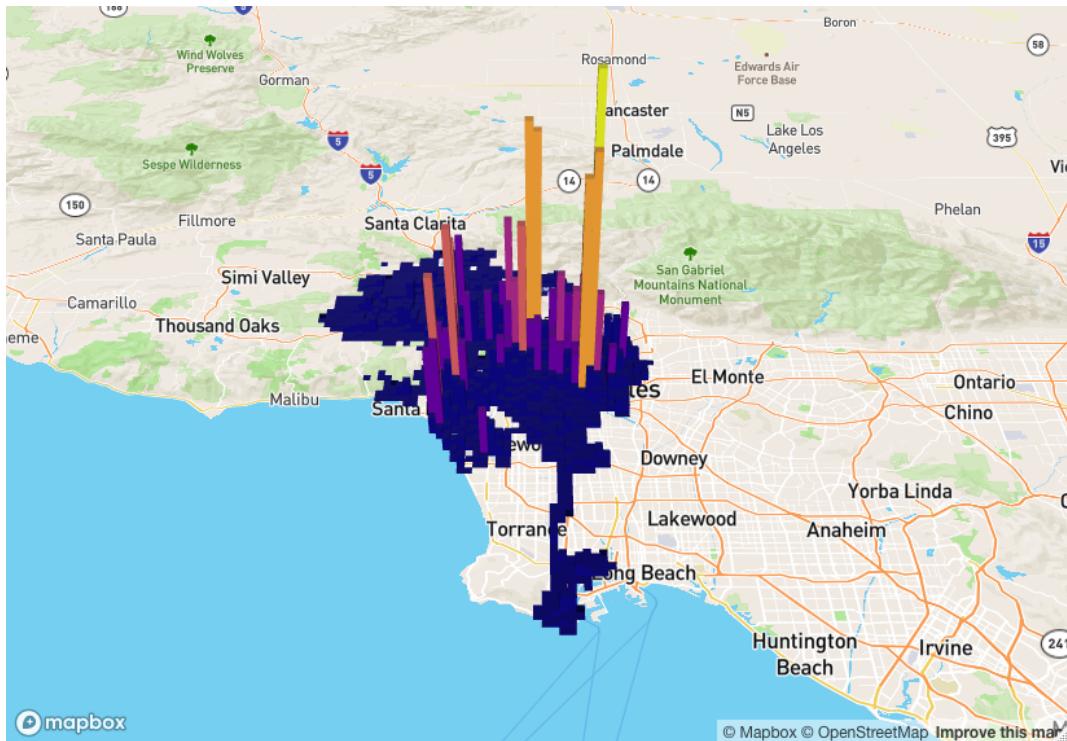
Matthew Barclay, Griffin Coccari, Alex Gallegos, Anushi Singhal

Contents

Introduction	2
Who receives parking citations?	4
Where do parking violations occur?	8
When do violations occur?	11
Conclusions	15
References	16

Introduction

Alex, Anushi, Matt and Griffin all come from different backgrounds and majors, so in discussions for which dataset we wanted to analyze, we explored a wide range of topics from public health to the New York City subway. We eventually decided to take on data from another large metropolis, but this time in sunny Los Angeles. In Los Angeles, the beautiful beaches, countless celebrities, and Hollywood homes are magical for visitors. However, LA traffic is often very bad and parking in tourist areas can be a nightmare. As the second largest city in the United States, there are over 6.4 million vehicles in the Los Angeles urbanized area¹. Our dataset, `Parking_Citations`, contains all the details of nearly 10 million parking violations in Los Angeles from 2010 to the present. That means we have access to the records that account for fines totalling close to \$600 million. With so much money at stake and the huge volume of data, the city of Los Angeles keeps track of these records electronically. We intend to take this massive amount of data and transform it so that the intricacies of parking violations can be easily understood.

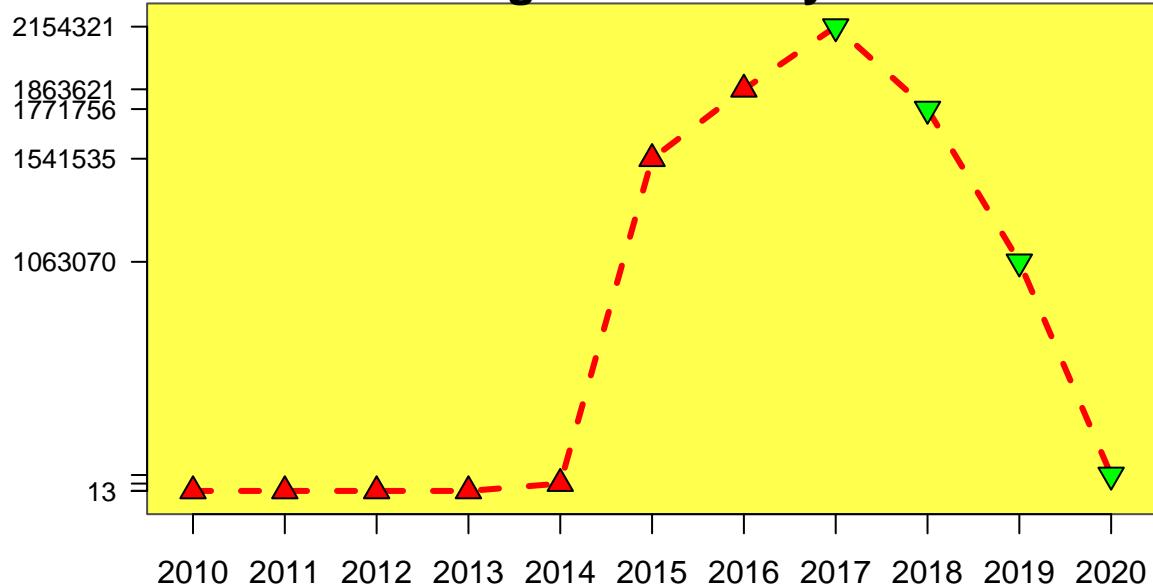


Big Data

The raw data was accessed directly from the City of Los Angeles Department of Transportation (LADOT) through the city's Open Data website². The original data set contained 9.97 million rows, each containing details on one parking violation. We identified the fine amount and location data as the most important variables and therefore removed all rows containing empty or NA values in those two columns. This narrowed the amount of rows to about 8.5 million. LADOT uses US Feet coordinates according to the NAD_1983_StatePlane_California projection, which is not easily comparable to standard latitude and longitude values, so our next step was to use `sp` package for R to transform our location data. Our final cleaned data set contained data on 8,502,692 tickets, including information of the time, date, location, vehicle information, parking offense, and more. Although well over the 1 million row minimum requirement, this large amount of data will allow us to explore the trends in parking violations in LA over a relatively large time frame.

Overview graphs

Parking Citations By Year



When we began our analysis, we had to get some idea of what we were dealing with. With no instructions, we first boiled down the data into a more understandable form. The data gave us a good amount of information, and it also gave us a good starting point. The data showed a skewed distribution over 11 year span. From 2010 to 2014 there were less than 40 parking citations recorded. In 2014 that number increased to 34,000. After that, we saw a drastic 44 percent change into 2015, when 1,541,535 parking citations were recorded. The count peaked at 2,154,321 citations in 2017. This peak accentuated both the steep rise leading up to 2017 and also the surprisingly steep decline afterwards. While this graph gave us a good handle on the numbers, there were questions unanswered about the data collection. The drastic increase from 2014-2015 may be explained by an attempt by the City of Los Angeles to digitize their citations. But that explanation is contradicted by the equally sharp decline after 2017. This left the question of how many citations did they actually collect? The graph didn't tell us exactly, but it gives us a range of values that captures the true value or something close. Even without the actual answer, this graph highlighted the magnitude of the citations in LA.

Year	Dollars_Collected
2010	730
2011	630
2012	2276
2013	3062
2014	2371400
2015	107172874
2016	130124585
2017	151608139
2018	124738695
2019	75291586
2020	2275
Total	591316252

Los Angeles is not just a city with horrible traffic, it's also a city where it is terribly hard to park, especially in tourist areas, shopping districts, or other special zones. Over the past decade, the City of Los Angeles collected \$591,316,252 digitally.

Top_Violations	Count	fine_amount
NO PARK/STREET CLEAN	2389114	73
METER EXP	1634877	63
RED ZONE	635504	93

This table explains that the top three violations are parking in a street clean area, parking with an experied meter, and parking in a red zone. Street Cleaning happens every day in Los Angeles, albeit on different sides of the roads depending on the day of the week. Because the side of the street that is cleaned changes on a daily bases, many people are caught parking on the wrong side of the road and are penalized. People who are in a hurry, who do know the law, sometimes ignore it or assume that the street cleaner has passed or won't come and are also penalized frequently. Parking in a red zone is The City of Los Ageles's third most frequently cited violation. A red curb, or red zone, indicates no parking, standing or stopping for public safety reasons³. Parking violations vary in their severity. Violating a red zone parking law is a public safety issue, so the fine is more expensive than that of a meter experation ticket, which is stil very expensive! At the very least, do not get caught parking in a disabled vehicle designated spot without reason. That fine will cost you 1100 dollars.

Violation.Description	Fine.amount
DP- RO NOT PRESENT	1100

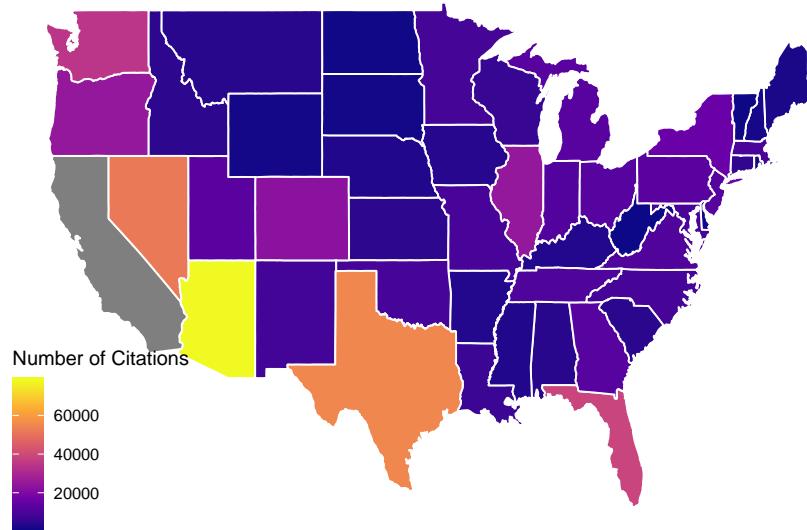
Location	count
1301 ELECTRIC AVE	11363
101 LARCHMONT BL N	7815
1600 IRVING TABOR CT	7346
2377 MIDVALE AVE	6113
5901 98TH ST W	5790
4301 TUJUNGA AV	5687
7000 HAWTHORN AVE	5606
2800 E OBSERVATORY	5378
4300 TUJUNGA AV	5279
100 LARCHMONT BL N	5236

Parking citations by location indicated where the most violations have occured. These violations happen in major hotspots of the city which have heavy tourist footraffick and very little space. The lack of parking forces people to park in residential areas, where there are high chances of fines. 1301 EElectric Ave is a street where people park to visit Long Beach, California's famous Seal Beach. Irving Court is very close to a the Venice Beach boardwalk and the Abbot Keny shopping district. Hawthorne Avenue is the spot just south of the bustling Hollyword Blvd. Among these hotspots is also The Griffith Observatory. These Los Angeles must-sees attract more cars and tourists and so are, unsurprisingly, the locations with the most cited violations.

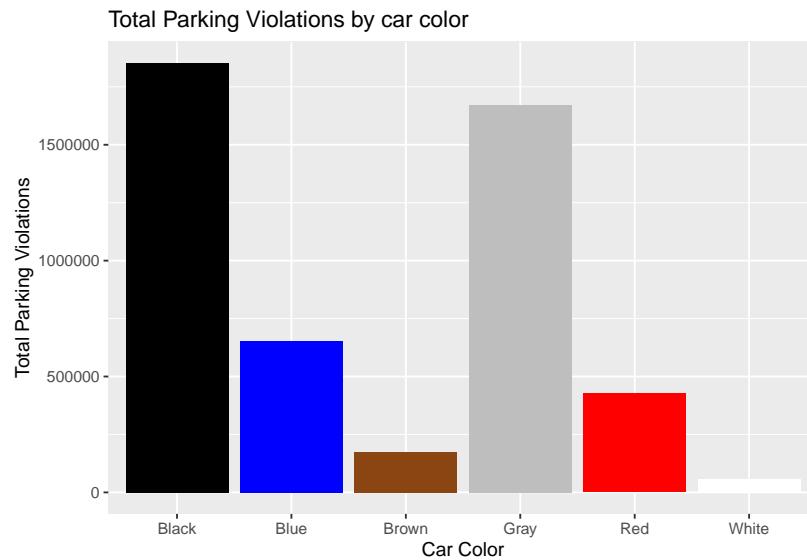
Who receives parking citations?

Parking citations and violations seem to be universal, affecting every vehicle on the road all the same. But with access to such large amounts of data, we wanted to know whether all vehicles had the same chance of recieving a parking violation.

Number of Citations for Vehicles with Out of State Registration

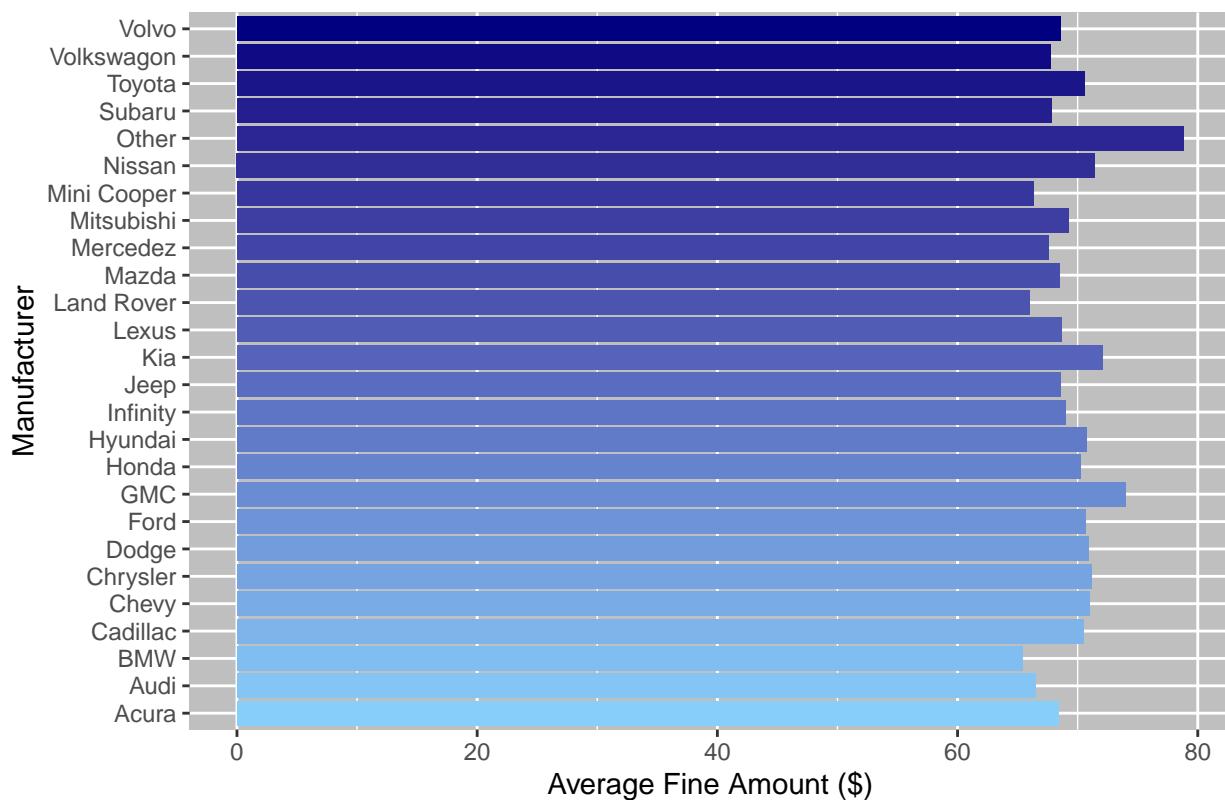


Although only around 1.5 million out of the 8.5 million violations we looked at were for out-of-state vehicles, we wanted to know what the distribution of these vehicles was. The fact that Arizona had the highest count seems logical given its close proximity to LA. This same reasoning could also explain the high concentration of vehicles from both Nevada and Texas. Although Oregon and Washington border the opposite side of California, the overall increased travel between the west coast states could account for these findings. On the other hand, the obvious outlier Florida was much higher than its neighboring states but this could either be coincidental or due to the higher volume of travel from this state.

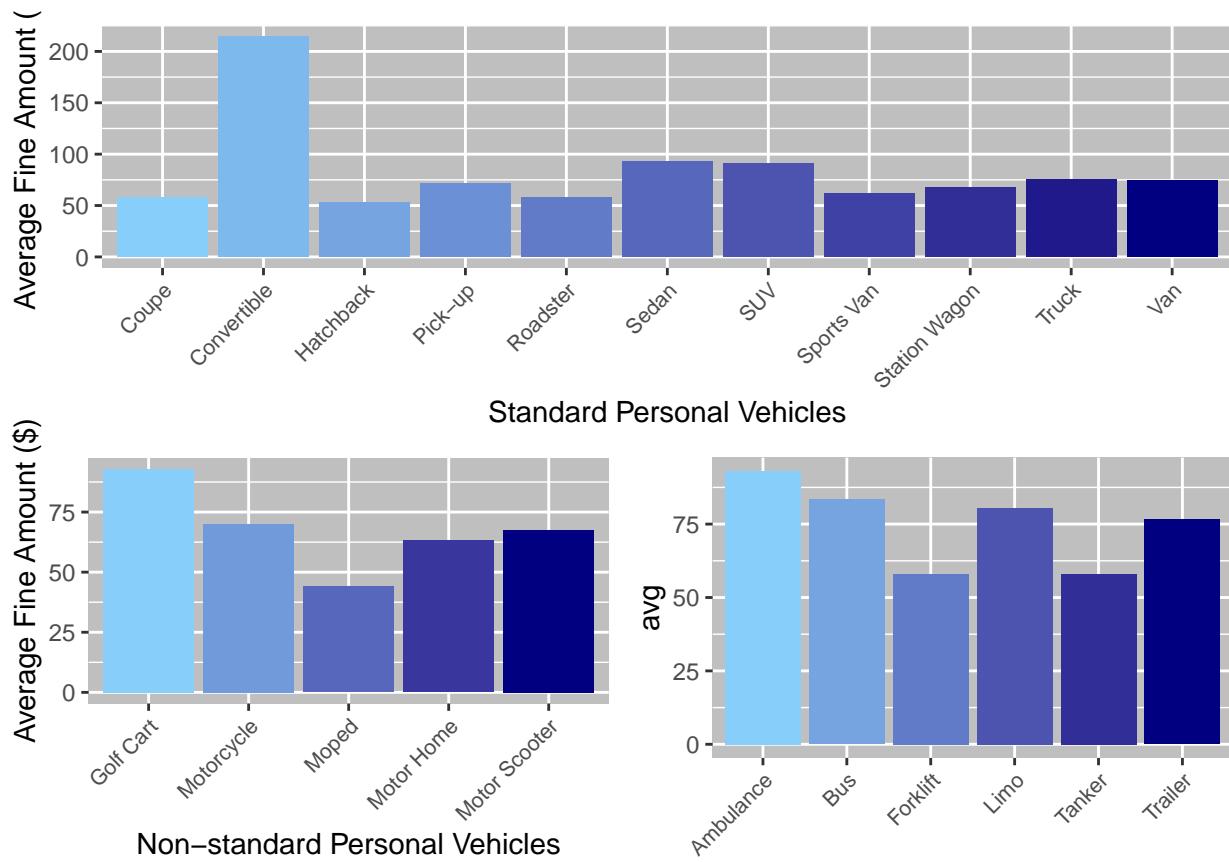


We have heard before that cars with bright colors, such as red, are more likely to get pulled over due to their high visibility on the road. In order to investigate whether there was any such color-bias when it came to parking violations we looked at the total number of violations by car color. In order to perform this analysis we manually picked 6 of the most common car colors. Since the distribution of count seemed to vary with the typical distribution of car colors we are used to seeing on the road, with colors like black and gray having the highest counts, we concluded that there was no significant impact on the likelihood of a parking violation based on the color of a car.

Average Fine Amount by Manufacturer

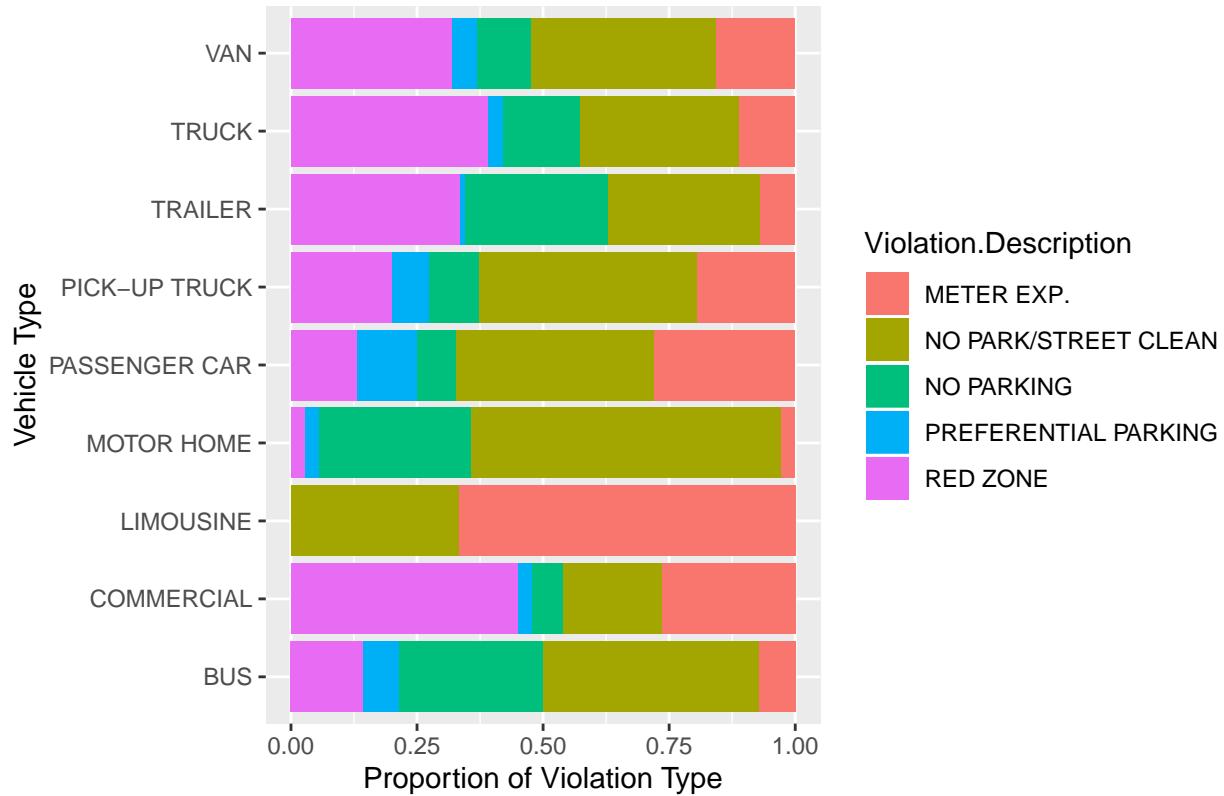


Although so far it seemed like there was a surprising amount of equality in the chances of various cars receiving parking violations, we wanted to look deeper into the actual manufacturers of the cars. We wondered if perhaps more expensive cars might be getting higher fines. So, grouped the cars by manufacturer and looked at the average fine amount of all the cars in groups greater than 50,000 violations. Given the fact that we represented over 8 million rows of data, even small fluctuations could be pretty significant. However when we looked at the highest averages of fine amounts, we saw that the Other category was highest followed by GMC, Kia and Nissan, none of which were considered very expensive. The lowest average was also BMW, which was known to be a more expensive manufacturer. The peak of the Other category could mean something given that this category might contain smaller less standard cars but this would also make sense if we assumed that the less standard cars were also more expensive.



Although a large majority of the vehicles was relatively standard personal vehicles, the dataset also contained non-standard personal vehicles, such as Golf Carts or Motorcycles and also more professional vehicles, such as ambulances or limos that were typically not personal vehicles and used for a more professional purpose. In order to analyze the fine amount at from another angle, we manually categorized 22 of the vehicle types into these three categories and looks at the trend. While convertibles were a clear outlier in the standard personal vehicles category this could likely be a coincidence due to the low overall count of these vehicles. Furthermore, mopeds had one of the lowest averages which could be explained by the fact that they may incur less significant violations due to the nature of the vehicle.

Percent of Different Violations by Vehicle Type

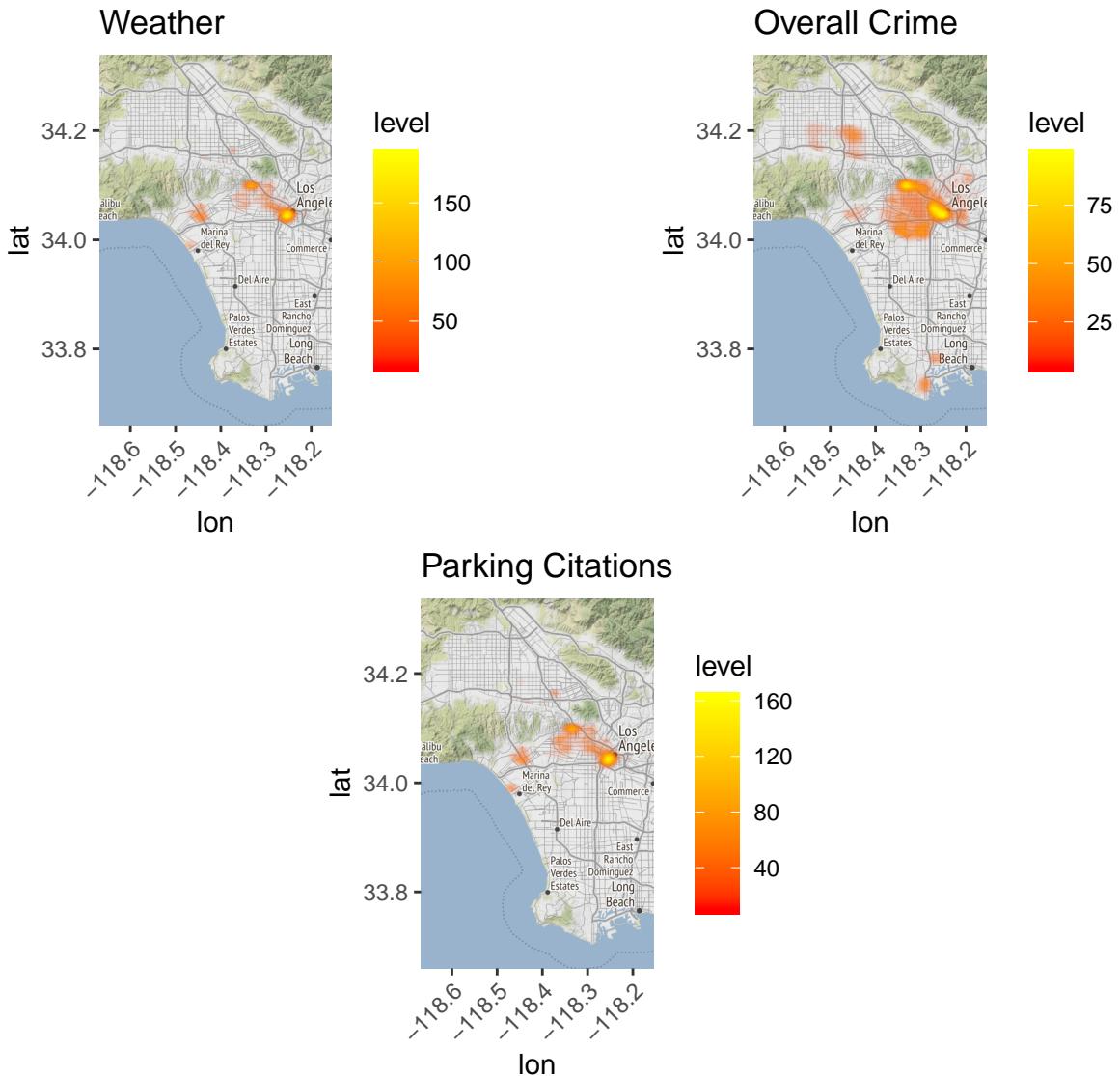


Finally we wanted to look at the various types of vehicles on a higher level. We wondered if perhaps the type of vehicle would impact the frequency of different types of violations. In order to make this analysis more feasible, we took subsetted the rows with the 6 most frequent types of violations and looked at these.

Where do parking violations occur?

Does weather or crime affect the parking convictions overall?

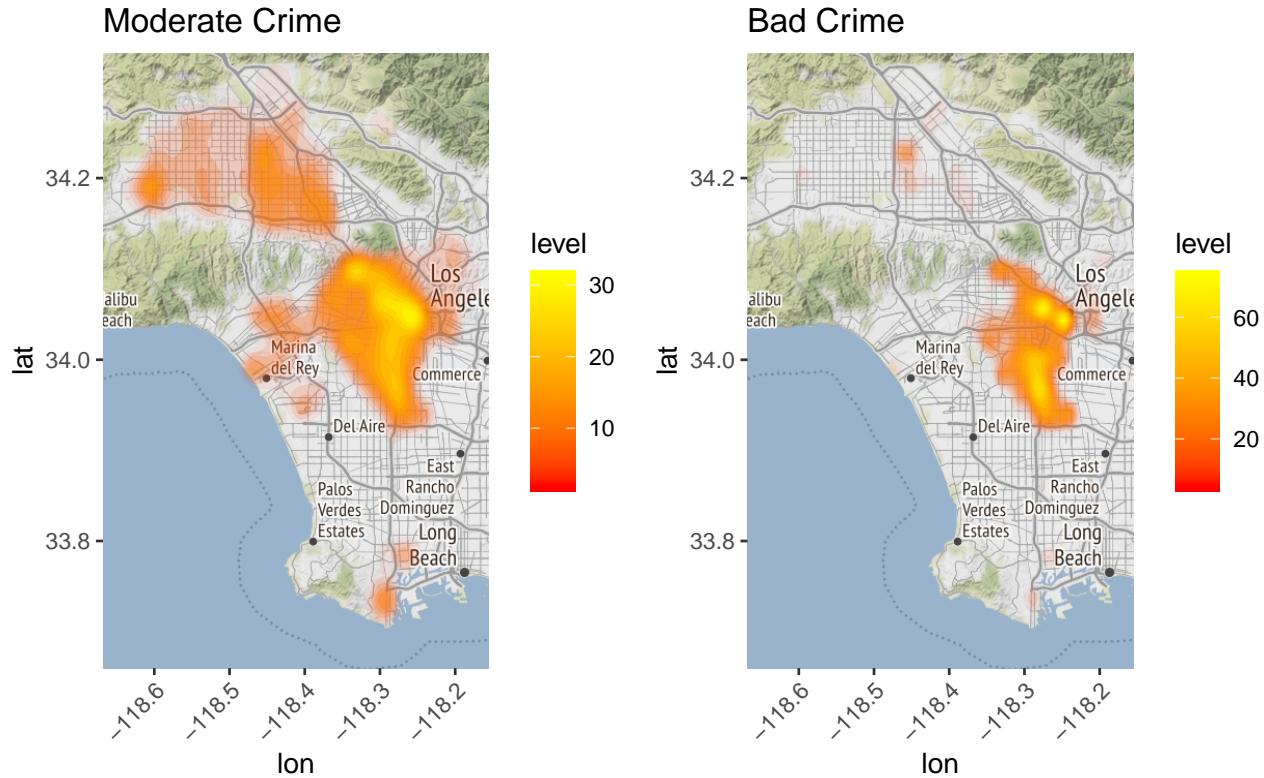
In our investigation of how parking citations are distributed across the city of LA, we wondered how crime data and potentially weather data would fit in with our parking citations data. We used 2 SQL databases of both LA weather and LA crime to compare with our initial data. By using a heatmap on the city of LA with out data we can see different concentrations of crime, weather, etc., and in conjunction with our initial dataset we can see if there are any correlations between crime, weather, and our parking citations.



The graphs show the density of parking citations in LA, the density of parking citations when weather was above half an inch of rain, and all crime in LA. From these heatmaps we can see that though there is a greater spread with crime over the downtown LA area, there is a big concentration around the central downtown area. It should be noted that the concentration of crime is much lower as shown by the legend on the graph, but the heatmap does still show that there is still a concentration centred around downtown LA.

Crime and its effect

The crime map does account for overall crime regardless of severity, so we thought that it would be a good idea to rate the crime by severity. In our data for crime, we are supplied with crime codes that indicate the type of crime committed. Because the crime codes are known and this is a database of reported crimes for LA we can assume that this would give a fairly accurate map of crime activity in LA. We also made the assumption that if there is more crime in an area it is more likely that there is increased police activity in that area, and therefore more citations can be given. Crime codes are on a numeric scale where the lower the crime code number, the more severe the offense, so we thought that it would be an interesting idea to

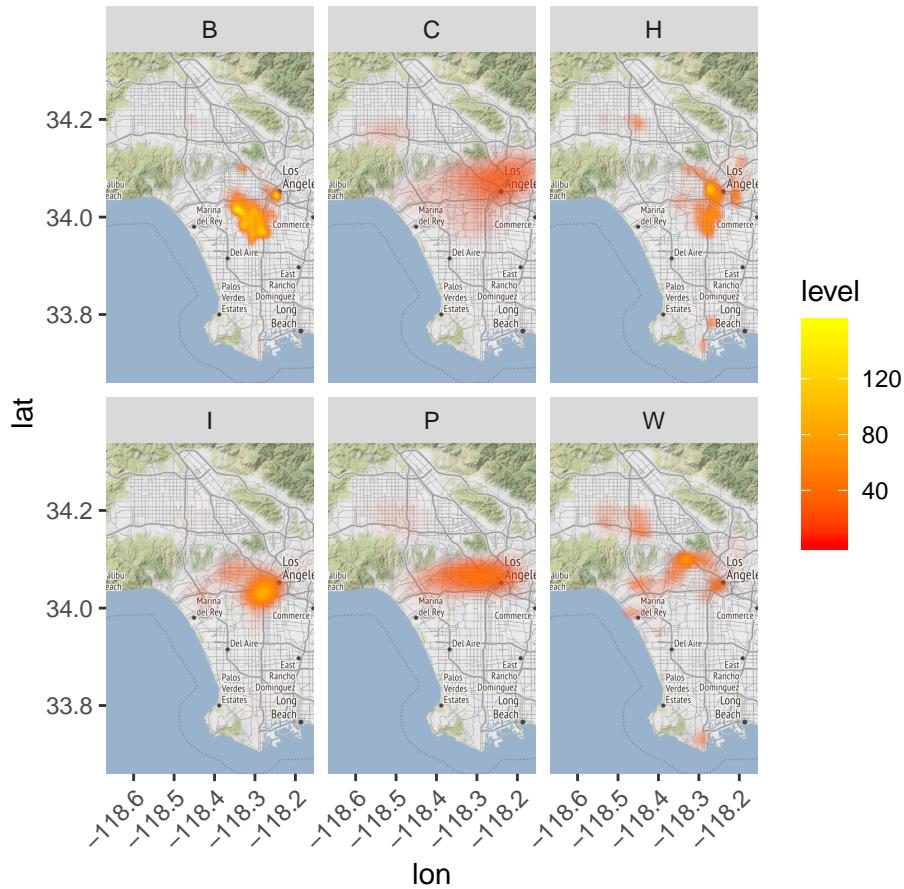


As shown in the graphs moderate crime with a moderate crime code tends to be way more spread out over the city when compared to the severe crime. Looking at the legends also gives credence to this as the density is way less than that of the severe crime with severe crime being extremely concentrated and much more dense than the moderate crime. When we compare to our parking citation data, we can see that there is an overlap in the central downtown area, but besides that there is still a lot of crime happening outside of the area regardless of severity suggesting that crime does not necessarily correlate to parking citations.

Demographic areas

Using our crime database we can find the demographics by crime and see in which areas crime happens based on demographic. By plotting these maps we can find demographic data by region which we can use to see whether or not there are demographic differences in parking citations frequency.

Demographic Data



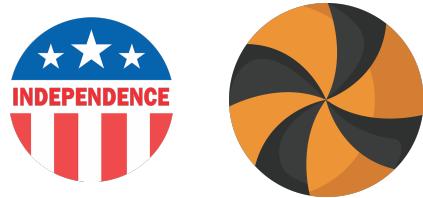
(Where B = Black descent, C = Chinese descent, H = Hispanic/Latinx descent, I = Amer-Indian descent, P = Pacific Islander descent, and W = White descent) Using this data, we can see that every one of our ethnicities does go around the central downtown LA area. We also see certain trends with our severe crime data. People of Black descent appear to be closer in proximity to the severe crime hotspot and people of Chinese descent appear to be further from the more severe crimes in the area. When we compare to parking citations it would suggest not only that crime has no effect on parking citations, but also that demographics does not necessarily contribute to frequency of citations and certain areas may not more likely to receive citations because of their ethnic makeup.

When do violations occur?

Holiday Parking

Major holidays account for some of the busiest travel days in the year, so we decided to investigate how the total amount of money collected on each day was distributed. We selected New Year's Eve, Super Bowl Sunday, Valentine's Day, St. Patrick's Day, July 4th, Halloween, Thanksgiving, and Christmas and summed up the fines issued for each day. The totals were then plotted proportionally for each holiday seen below.

New Year's	Super Bowl Sunday	Valentine's Day	St. Patrick's Day
Total: \$1,965,659	\$554,469	\$1,497,590	\$1,841,597



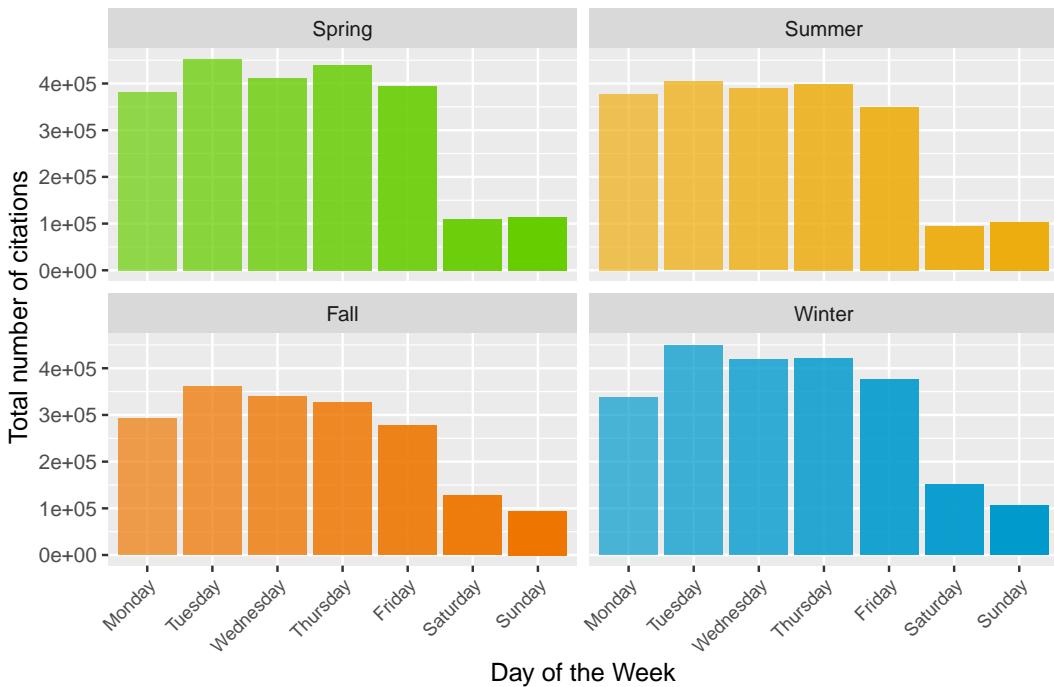
July 4th	Halloween	Thanksgiving	Christmas
\$1,373,611	\$1,619,969	\$284,933	\$209,549

Our analysis found a striking difference between the fine amounts for holidays. We observed that holidays typically associated with drinking, New Year's, St. Patrick's Day, and July 4th, tended to have more fines issued. We were surprised to find that Christmas and Thanksgiving had the two lowest total fines issued, as we associated these holidays with travel. These lower than expected values could possibly be explained by other variables, however, such as a decrease in LAPD/LADOT staff working on these days or increased leniency. We did expect Halloween to have a large fine amount, as finding legal parking while Trick-or-Treating can be difficult. We expected the Valentines Day's results as well, but were surprised that Super Bowl Sunday was so low, as traveling to parties and drinking often occurs on that day. Super Bowl Sunday, however, is the only holiday that is always on a certain day of the week, whereas all of the other holidays could fall on any day. This led us to investigate whether the day of the week influenced how citations were issued.

Days of the Week

While analyzing which day of the week elicited the most citations, we also decided to group by the season of the year to see if that would affect the total volume of citations issued.

Distribution of tickets throughout the week for each season

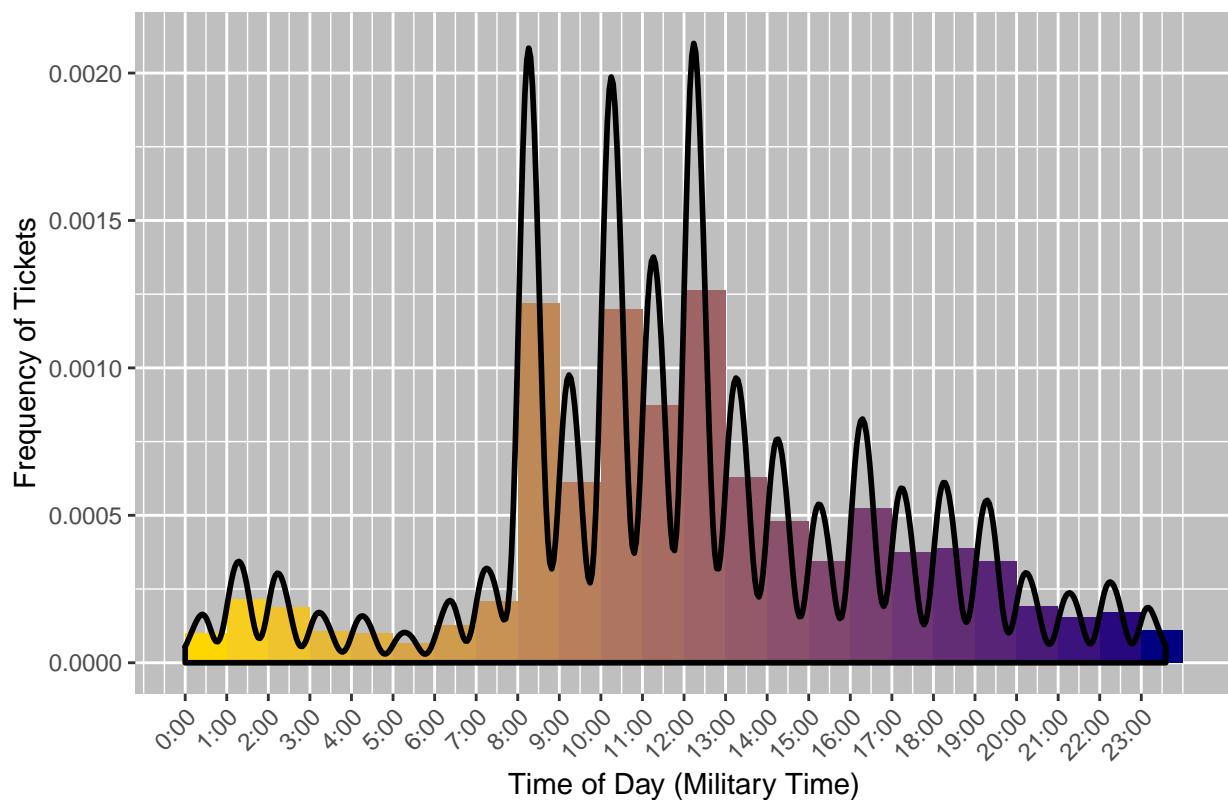


Across every season, Tuesday had the greatest volume of citations issued, while the weekends had the fewest. This finding was interesting, because the increased traffic usually generated on weekends did not translate into the volume of citations, but somehow created the opposite effect. Additionally, while we expected to see the citation volume skyrocket in the summer with the influx of tourists ignorant to LA parking laws, summer volumes were overall lower than both the spring and winter. This data did, however, validate the possibility that Super Bowl Sunday had relatively low fines, as Sundays in general received less citations than weekdays.

Time of Day

We also explored how citations were issued over the course of the day.

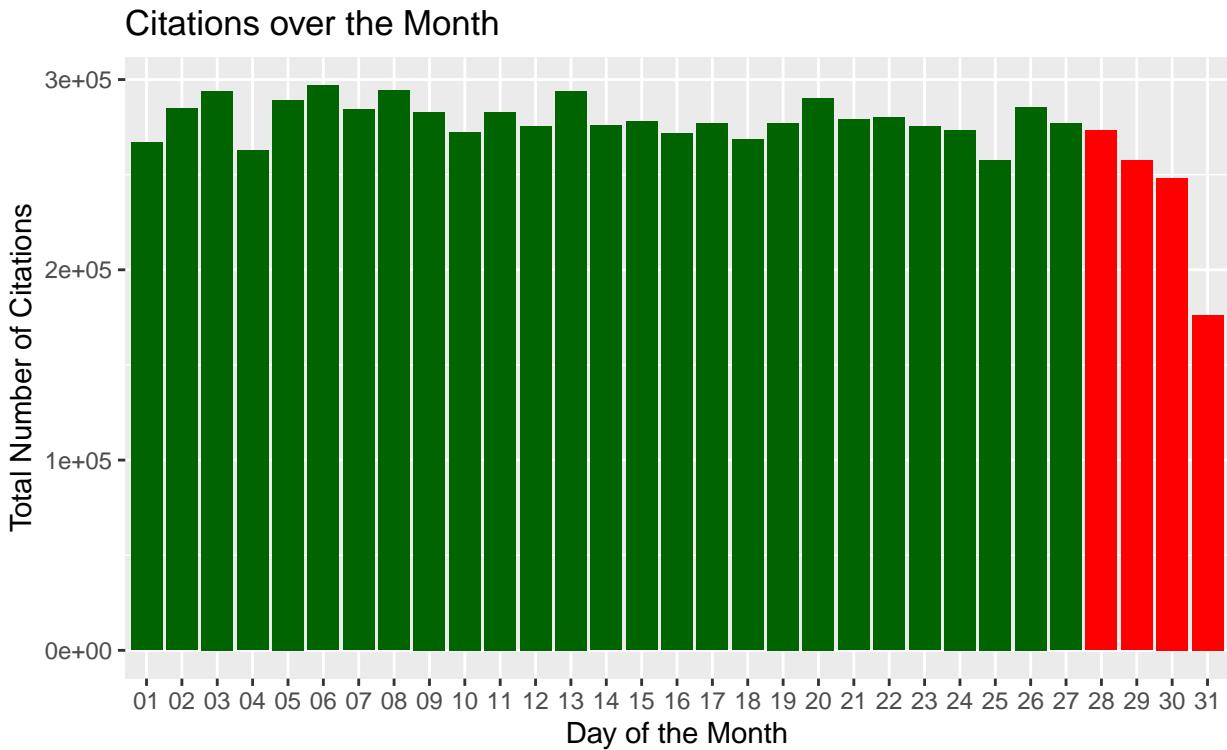
Frequency of Parking Tickets Throughout the Day



The greatest frequency of citations were issued between 8:00am and 2:00pm, with peaks occurring in the first half of each hour. Unsurprisingly, very few citations were issued in the early hours of the morning, wth the lowest frequency occurring between 5:00-6:00am.

Quotas

It is a common misconception that ticket frequencies increase at the end of each month as many people believe that officers have monthly quotas to fill. We explored whether an increase in ticket frequency was observed at the end of the month by plotting the total count of tickets for each day of the month for all months.



As seen in the plot, not only did we find no evidence of citation frequency ramping up toward the end of the month, we found that citation frequency tends to remain relatively stable. The extremely low frequency on days that are the 31st of the month is caused by only 7 out of 12 months having 31 days.

Conclusions

The analysis of our Parking Citations data evolved into more than just a matter of counting. Our aim was to do more than just viewing the dataset at face value. We wanted to see how parking citations explained larger picture characteristics about the city of Los Angeles, but especially in regards to traffic and parking citations. With this in mind, our analysis began with an overview of our data, focusing on prices and fines. This uncovered startling statistics about the total cost of these fines each year and the frequency of various violations. We then jumped into some of the more basic columns such as variables that defined the appearance of the car, what their manufacturers were and where they were from. We wanted to answer the overall question of whether or not the car itself would impact its likelihood of getting a violation, the type of violation it would get, and the price of the fine for this violation. When looking at the most cited color and car types, our analysis revealed that special vehicles like ambulances, golf carts, forklifts, and tankers also were subject to Los Angeles parking enforcement more frequently than one might think.

The times in which parking citations were issued also break some preconceived notions about police citations. Certain holidays that are usually celebrated with heavy drinking did have a corresponding increase in citations issued, but some holidays like Halloween and Super Bowl Sunday did not follow that pattern. Also, we found that weekdays generally had more fines associated with it than weekends and across all seasons Tuesday had the highest amount of issuings lending credence to the idea that going out on weekends may not lead to an increase in chance of getting a citation. We also found no evidence of quotas for citations affecting the issuing of citations as by day, the amount of issuings was relatively stable.

Excited by the potential of what our dataset offered, we used complementary datasets on weather and crime to dive deeper into the socio-economic, cultural, and meteorological effects on Los Angeles parking Citations. When observing the data shown as a heatmap it can be shown that crime hotspots do not necessarily correlate to a hotspot for parking citations; weather, on the other hand, can. Demographic data does not necessarily have an effect on the hotspot of parking citations. In fact, the only places where all demographic data overlapped is the hotspot of parking citations which suggests that demographic region does not necessarily

mean there is an increase in parking citations risk. Going forward, if it were possible to incorporate income data for each geographic region, that could be used to find correlations with parking citations.

What excited our group most about this project was finding relatable insights that have practical value. With a little more polishing, maybe we could present our insights to the city of Los Angeles or create a website to communicate our insights to the public. In order to reach this goal, our group aims to incorporate additional tools. Web Scraping could allow us to take our analysis further by depending less on common knowledge and incorporating real-world prices and locations to develop deeper insights from our data. We could also use Shiny to develop more interactive methods of communicating these insights. Furthermore, we want to apply these new techniques to build on our heatmaps and further generalize/verify our analyses with the goal of eventually being able to model parking violation trends in other cities based on geographic or other data.

References

1. Newton, Damien, et al. "Density, Car Ownership, and What It Means for the Future of Los Angeles." Streetsblog Los Angeles, 13 Dec. 2010, la.streetsblog.org/2010/12/13/density-car-ownership-and-what-it-means-for-the-future-of-los-angeles/.
2. "Parking Citations: Los Angeles - Open Data Portal." Data.lacity.org, 13 Feb. 2020, data.lacity.org/A-Well-Run-City/Parking-Citations/wjz9-h9np.3.
3. "Colored Curb Zones." LADOT, ladot.lacity.org/residents/colored-curb-zones.
4. Photos were obtained from the following websites: imgbin.com, pngguru.com, logodix.com, hiclipart.com, pngfuel.com, pikpng.com, and cleancpng.com