

## **Crime Analysis in Chapel Hill from 2018 - 2020**

Link to Github repository with R code: <https://github.com/akunna1/PLAN--390-Final-Project.git>

### **Introduction**

Most new home buyers and renters would like to live in places that are safe i.e., places with low crime rates, so knowing the crime statistics and the rate of an area would be important to them. Also, crime statistics are important to law enforcement agencies because it provides them with data, which they can use in budget formulation, planning, resource allocation and assessment of police operations to help address the crime problem at various locations and levels. Crime rate is calculated by dividing the number of crimes reported in an area by the population of the area and then by multiplying it by a hundred. The purpose of this paper is to analyze the crime in Chapel Hill, North Carolina from years 2018 to 2020. The crime density in Chapel Hill, the frequency of arrests during each month of the year, the relationship between the number of arrests and race, the relationship between the number of arrests and age, the frequency of the use of weapons to commit crimes, the statistics of the number of crimes committed under the influence of drugs or alcohol, and the type of crimes committed during the year period would be analyzed.

### **Methods and Results**

#### ArcGIS Pro Portion

The shapefile of the for the cities in Orange County was obtained from the website, <https://www.orangecountync.gov/2057/Download-GIS-Data> and was opened in ArcGIS Pro. With the 'Select by Attribute' tool, the city of Chapel Hill was selected and a new shapefile was created out of the selection. The dataset of the Daily Police Arrests made in Chapel Hill, Carrboro,

Hillsborough, Durham, Burlington, and Pittsboro for the years 2018 to 2020 for all the twelve months in a year were obtained from the website, <https://opendata-townofchapelhill.hub.arcgis.com/datasets/townofchapelhill::police-arrests-made/about>. Most of the data were for arrests made in Chapel Hill. Year 2021 and 2022 were not used because the website only contains the months of January to October for year 2021 and none for 2022 because it had not yet been updated to include them. Each year has its own table. All the data contains spatial information i.e., longitudes and latitudes, so the csv file of the arrest record for each year was opened in ArcGIS Pro. Each table was right clicked on and ‘Display XY Data’ was selected in order to see each data as points. The ‘Select by Location’ tool was then used to select only the points within the boundary of Chapel Hill, and new points layers were created from the selection. The ‘Merge’ tool was used to merge the 2018, 2019 and 2020 data to make one new shapefile of points. The ‘Point Density’ tool (located under Spatial Analyst tools) calculates the density of point features around each output raster cell. The ‘Point Density’ tool was used to calculate the point density of the arrests made in 2018, 2019, 2020 and the three years combined. For each shapefile under symbology, the method was set to Geometric Intervals, and the number of classes was set to six. A point layer shapefile for the Old Well in the University of North Carolina at Chapel Hill (UNC-CH) was created and opened in ArcGIS Pro and it was added to the map.

Figure 1 below shows the map of Chapel Hill and the point location of the arrests made during 2018 to 2020. Figures 2 to 5 shows the point density for year 2018, 2019 and 2020 and the combination of the three years. The point indicating the location of the Old Well is shown in the figures. By zooming into the World Topographic Map layer on ArcGIS Pro, it was discovered that the regions with the highest point densities are places are within East and West Franklin Street,

South Columbia Street and East and West Rosemary Street, and some places within the university area. These places are a few miles from the Old Well.

## R Portion

### *Month Analysis*

I imported the .csv files for each year into R. Year 2018 contains 1,487 records, year 2019 contains 1,258 and year 2020 contains 648. With each dataset and the dataset of the three-years combination, which I created by binding together the three datasets using the `rbind` function, I did a Month Analysis using the `group_by` and `summarize` functions to count the number of arrests made during each month of the year. Then, I created bar plots of each of the months and their counts using the `ggplot` function and created a bar plot for the combination of the three years. With this method, it was easy to see how the number of arrests varied by month.

Figure 6 shows the number of arrests made in 2018 by month, Figure 7 shows for 2019, Figure 8 shows for 2020 and Figure 9 shows the combination of the three years. For 2018, the month of May has the highest number of arrests, August has the second, June has the third, and January has the least. For 2019, the month of April has the highest number of arrests, August has the second, October has the third and September has the least. For 2020, the month of January has the highest number of arrests, March has the second, February has the third and April has the least. After combining the three years, it was discovered that during the three years period, most arrests were made in May, August has the second, March has the third and the least was in November. The month of May and March are in Spring and August is in Summer.

### *Race Analysis*

For each year and the three-year combination, I did a Race Analysis using the `group_by` and `summarize` functions in order to count the number of arrests by the race of the arrestees. The races include Black, White, Hispanic, Asian, Islander, Other and Unknown. Using `ggplot`, I created bar plots for each year to illustrate the number of arrests by race, and I created a pie chart for the whole three years using the `pie` function.

Figures 9 and 10 were generated from the Race Analysis. Figure 9 shows the count for the number of arrests by race of the arrestees for 2018, 2019 and 2022, and Figure 10 shows a pie chart of the number of arrests by race for the years combinations. Black people have the highest number of arrests, followed by Whites, then Asians.

### *Age analysis*

For the 2018 to 2020 dataset, I did an Age Analysis by race and an age analysis by gender using the `group_by` and `summarize` functions in order to find the average, maximum, minimum and median age of the arrestees by race and gender. I created two tables with the data obtained and I calculated and added the average of the average, maximum, minimum and median age by race at the bottom of the table and did the same for the average, maximum, minimum and median age by gender.

Figure 11 shows the table of the Age Analysis by Race for 2018 to 2020, and Figure 12 shows the age analysis by Gender for 2018 to 2020. For all races, most of the people arrested were in their thirties and it is the same for gender

### *Weapon Analysis*

For the 2018 to 2020 dataset, I did a Weapon Analysis using the `group_by` and `summarize` functions in order to count whether each arrestee was harmed or unharmed and what type of

weapon he or she had on hand during or before arrest if he or she was harmed. I created a table with the data obtained.

Figure 13 shows the Weapon Analysis for 2018 to 2020. A huge majority of the arrestees were unharmed, then that was followed by the arrestees using handguns and then by those using lethal cutting instruments as weapons.

#### *Under the Influence of Drugs or Alcohol Analysis*

For the 2018 to 2020 dataset, I did an Under the Influence of Drug or Alcohol analysis using the `group_by` and `summarize` functions in order to count the number of arrestees that committed crimes under the influence of drugs or alcohol. I created a pie chart to display the data.

Figure 14 shows the Under the Influence of Drug or Alcohol Analysis for 2018 to 2020. 57.3% of the arrestees were under the influence of drugs or alcohol, 31.8% were not and it was unknown for 10.9%

#### *Crime Type Analysis*

For the 2018 to 2020 dataset, I did a Crime Type analysis using the `group_by` and `summarize` functions in order to count the number of crime types that the arrestees committed. I created a table to display the top ten crimes committed in Chapel Hill during that time period.

Figure 15 shows the table of the top ten crimes committed in Chapel Hill, which was obtained from the Crime Type Analysis for 2018 to 2020.

### **Discussion**

East and West Franklin Street, South Columbia Street and East and West Rosemary Street, and some places within the university had high point densities. East and West Franklin Street,

South Columbia Street and East and West Rosemary Street have several stores, housing facilities, restaurants and businesses located on them, which attract many visitors, staff and students from the university, making it a populated place. There is a positive relationship between crime rate and population size. Higher populated regions tend to report higher crime rates. From the analysis of the 2018 to 2020 dataset, it was discovered that crimes are more likely to occur during the warmer months. During the warmer months, people are outside of their homes more and are in contact with more people, so therefore the number of crimes committed during those months would be higher than the colder ones. The current race demographics in Chapel Hill is White: 71.74%, Asian: 13.03%, Black or African American: 10.89%, and Two or more races: 2.87% The white population is about seven times the black population; however, 46.7% of the people in Chapel Hill that were arrested were white, 50.5% of the people arrested were black and 2.25% of the people arrested were Asians. In other words, Blacks and Asians are the minorities and Whites are the majorities in Chapel Hill, but black people are the most arrested race there. This is due to racist policing practices and the over policing of black communities. Plus, there are large observed racial differences in urban crime. “This is because urban blacks and whites tend to be highly segregated from one another living in communities that are ecologically and economically distinct” (Krivo and Peterson 622). Poor disadvantaged urban communities tend to have higher crime rates. “Predominantly black urban neighborhoods are characterized by average levels of poverty, joblessness, family disruption, and other aspects of deprivation that are higher than in white communities” (Krivo and Peterson 622). Furthermore, young people are more likely to commit crimes than old people because old people experience declining physical strength and energy, which makes crime for them too dangerous or unsuccessful. Most of the people that committed crime in Chapel Hill from 2018 to 2020 did it unharmed, and most were drug or alcohol related crimes. Five out of the

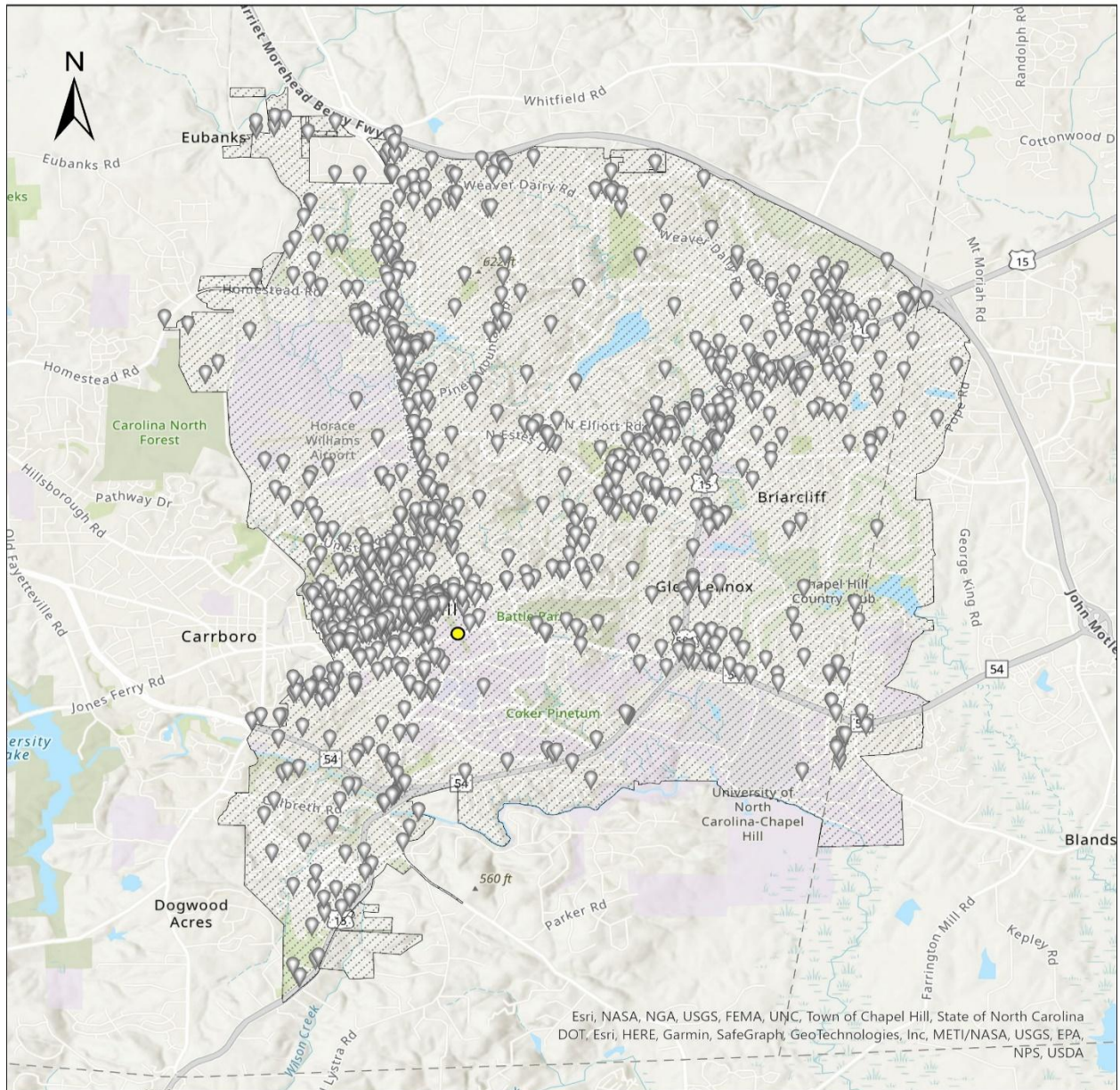
top ten crimes committed were drug and alcohol related. Assault was the second common. Also, most of the crimes were committed under the influence of drugs or alcohol. One's friendship network and peers can determine if he or she would use and abuse drugs and alcohol and to what degree. The college setting has provided a transitional environment for young people to move beyond the immediate influence of their family and to explore new activities primarily with peers (Perkins 15). Alcohol and a variety of drugs have become part of the social scene, and they have been used for personal experimentations on college campuses.

The results of this analysis could help predict the crime in an area and could help law enforcement agencies know where to allocate their resources. Most of their resources should be allocated to the regions that have high crime point densities. For example, law enforcement agencies can send more police officers to a street that has been found to have multiple larceny incidents over the years, in order to take care of the situation. From the analysis, it can be seen that the use and abuse of drugs and alcohol is an ongoing issue at UNC-CH and there are high causes of assault. Education and continual law and rule enforcement can help reduce the number of cases. A possible limitation for this analysis is the lack of data for 2020. Year 2018 contains 1,487 records, year 2019 contains 1,258 and year 2020 contains 648, which is about less than half than that of 2018 and 2019. The lack of data in 2020 could be the reason why the highest three months for the month analysis of 2020 include January and February, which are winter months. Meanwhile, crimes occur the most in hotter months.

## Figures

**Figure 1.** Arrests made in Chapel Hill, NC from 2018 to 2020

### Arrests made in Chapel Hill from 2018 to 2020



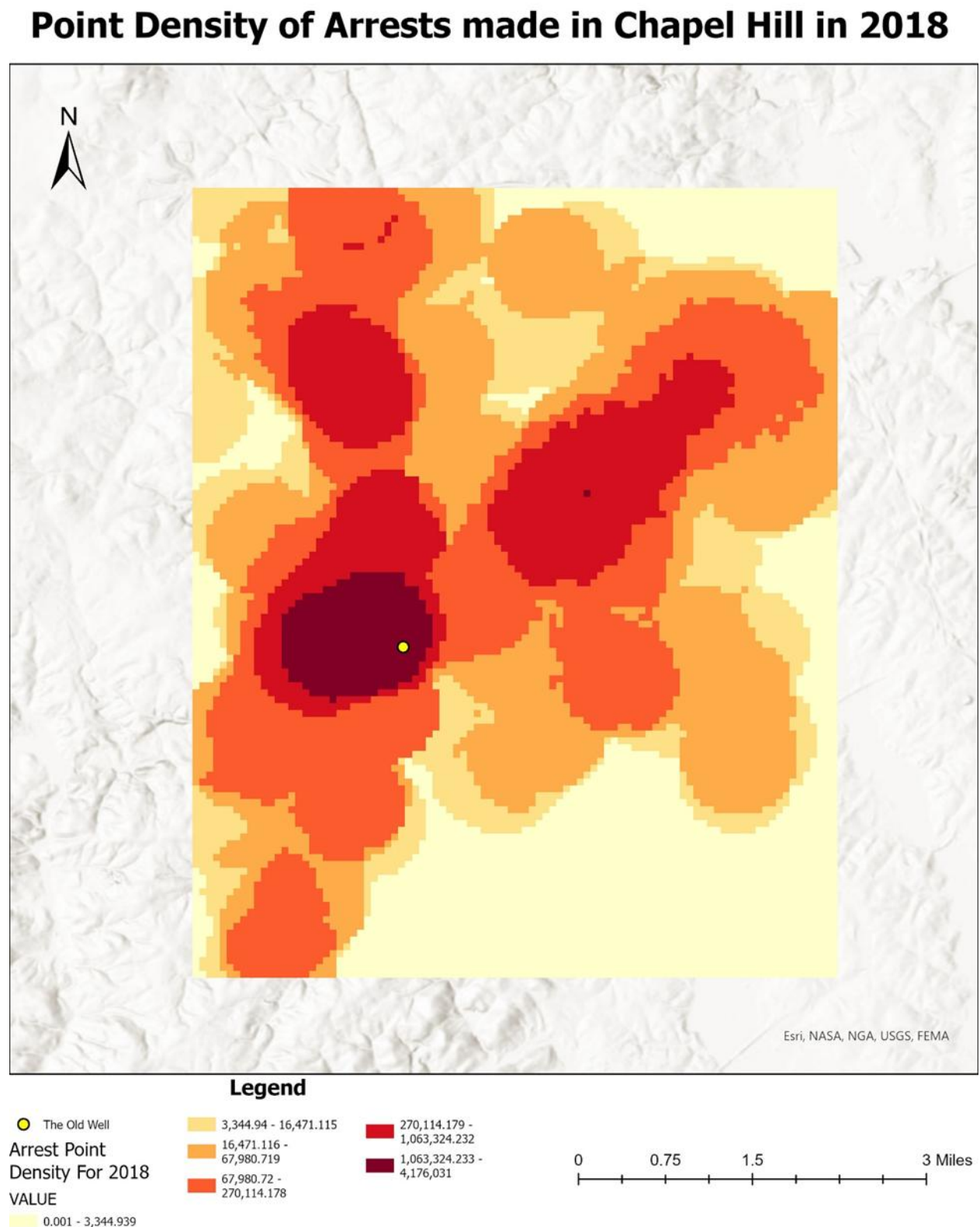
### Legend

-  2018-2020 Arrests (Chapel Hill)
-  The Old Well
-  Chapel Hill

0 0.75 1.5 3 Miles

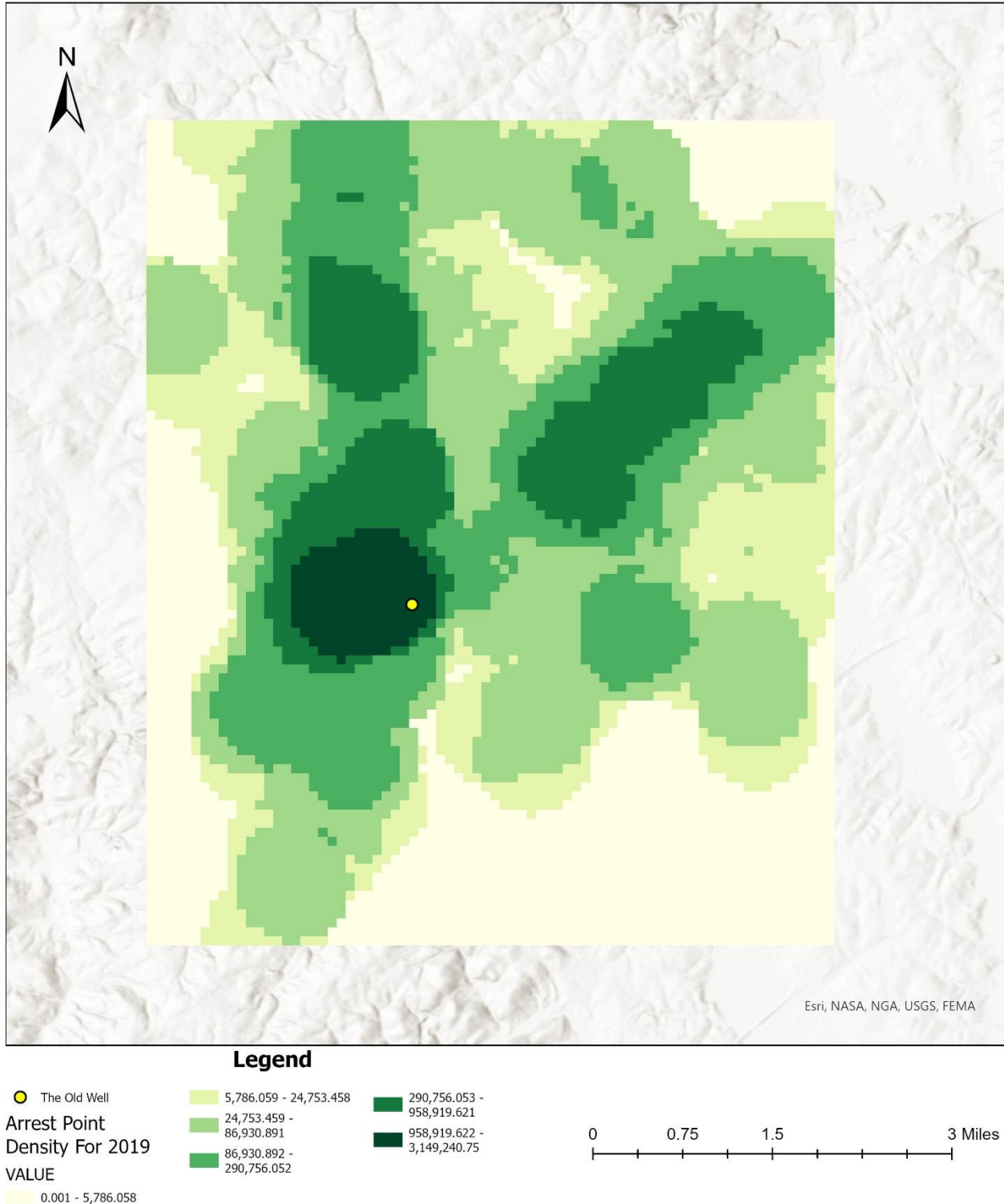


**Figure 2.** Point Density of Arrests made in Chapel Hill, NC in 2018



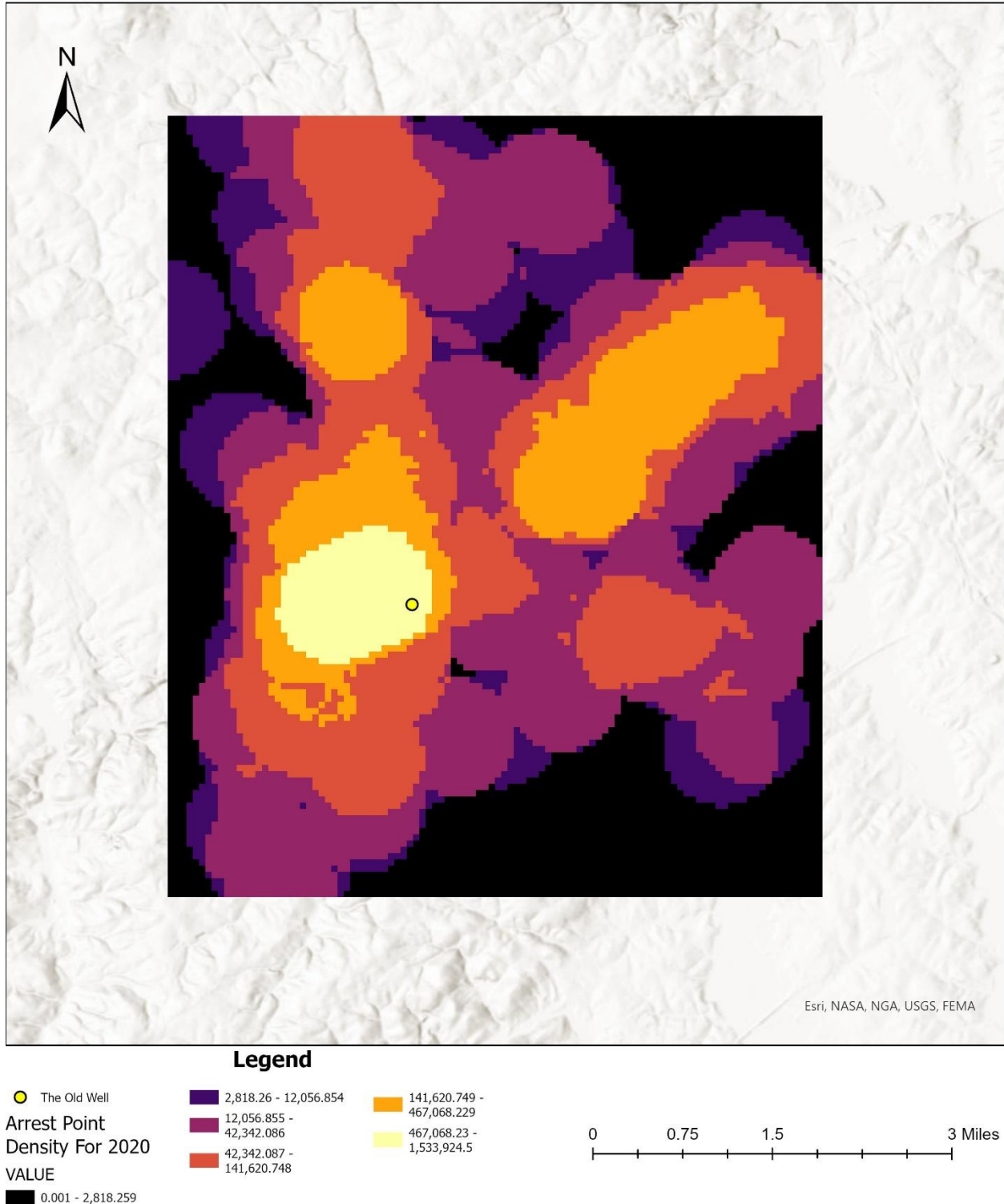
**Figure 3.** Point Density of Arrests made in Chapel Hill, NC in 2019

## Point Density of Arrests made in Chapel Hill in 2019



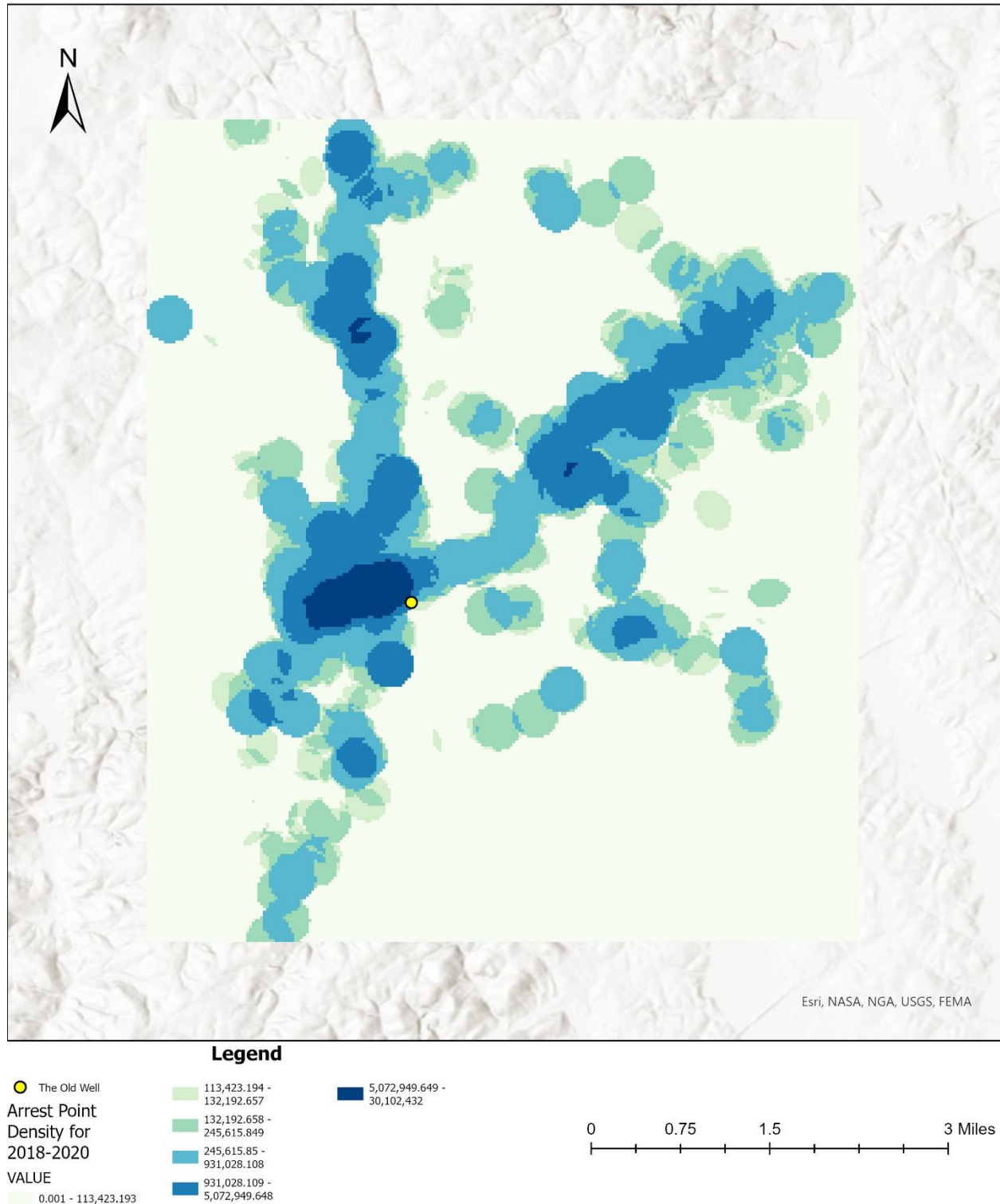
**Figure 4.** Point Density of Arrests made in Chapel Hill, NC in 2020

## Point Density of Arrests made in Chapel Hill in 2020

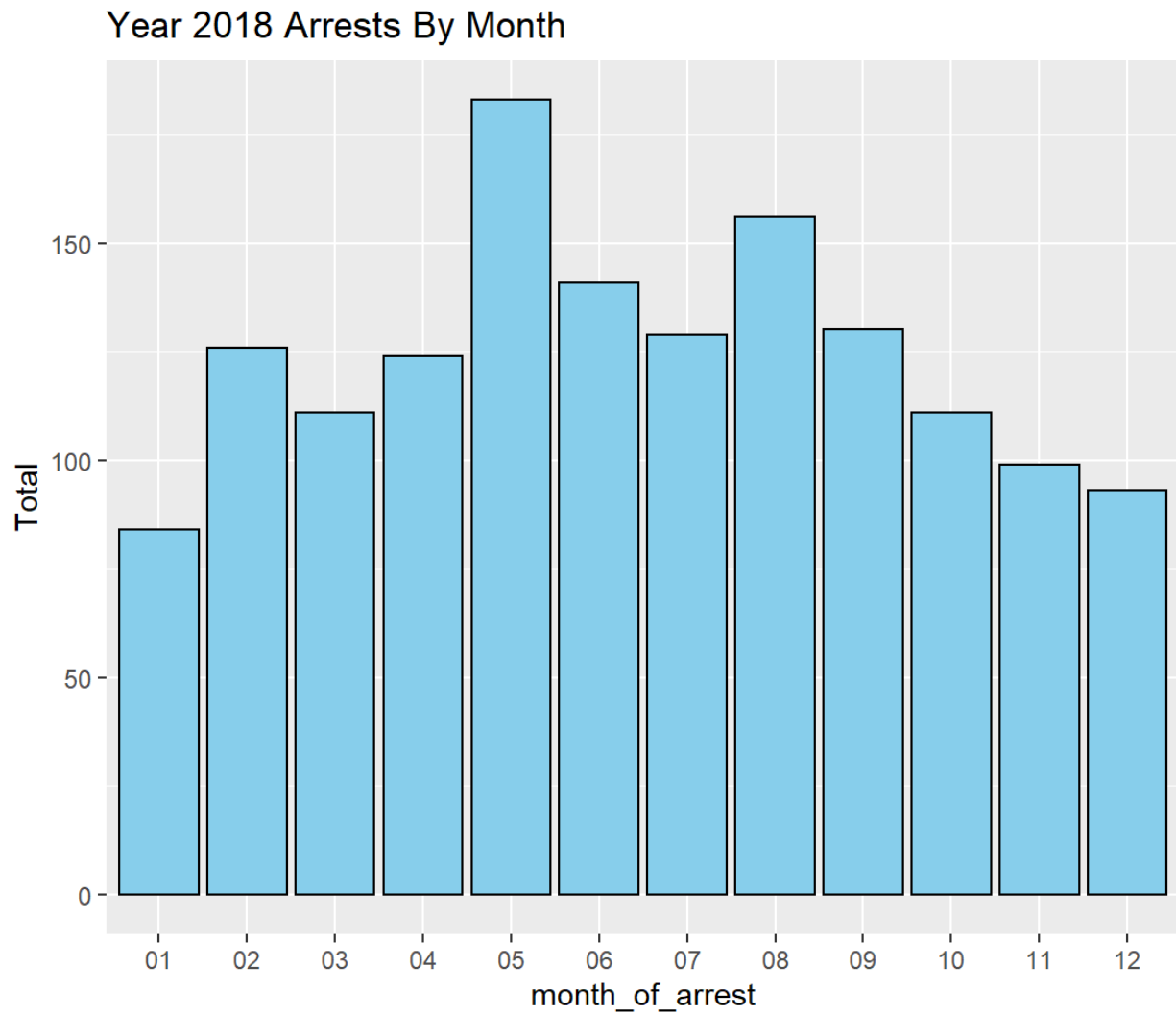


**Figure 5.** Point Density of Arrests made in Chapel Hill, NC from 2018 to 2020

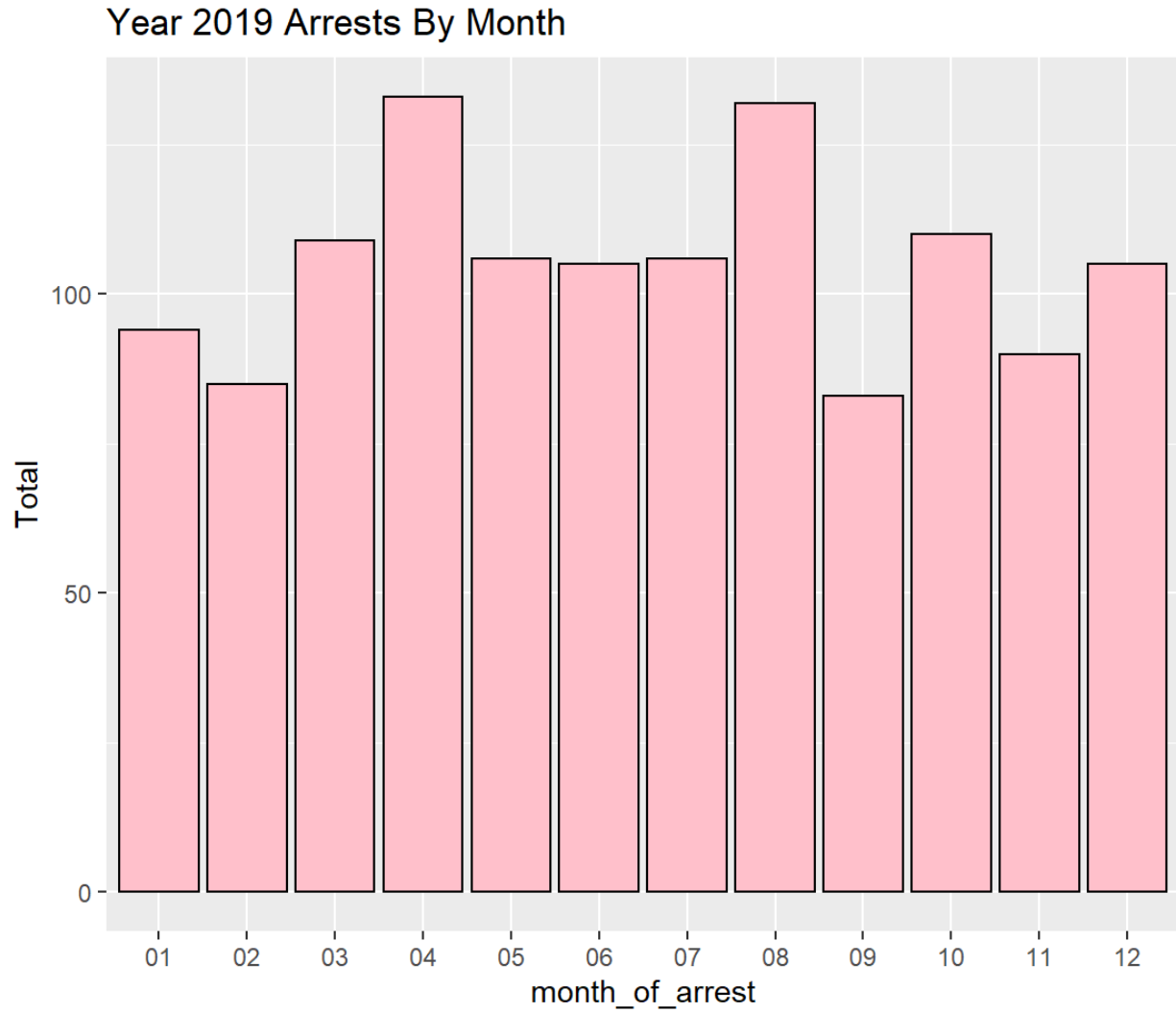
### Point Density of Arrests made in Chapel Hill from 2018 - 2020



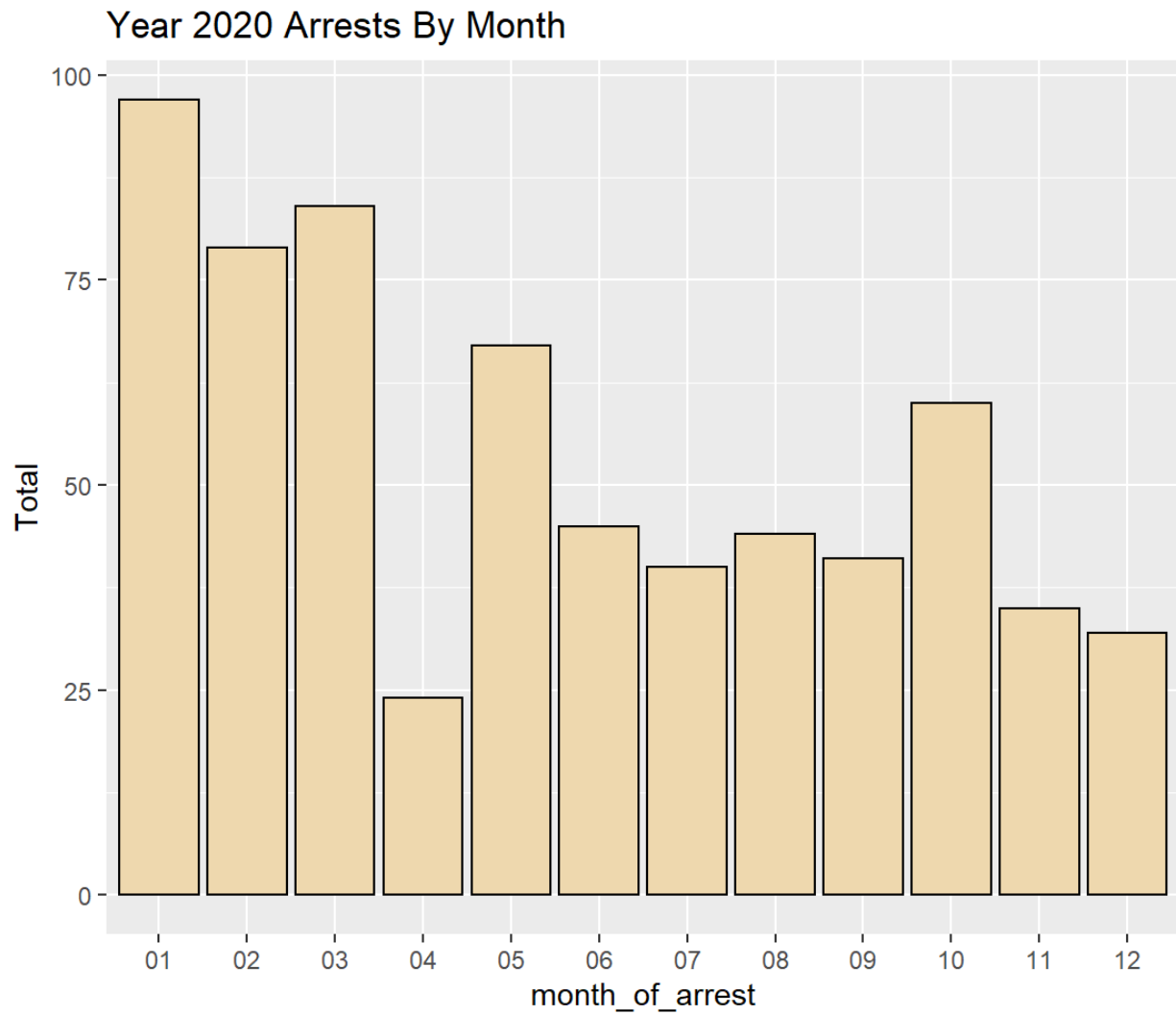
**Figure 6.** Number of arrests made in 2018 by month



**Figure 7.** Number of arrests made in 2019 by month

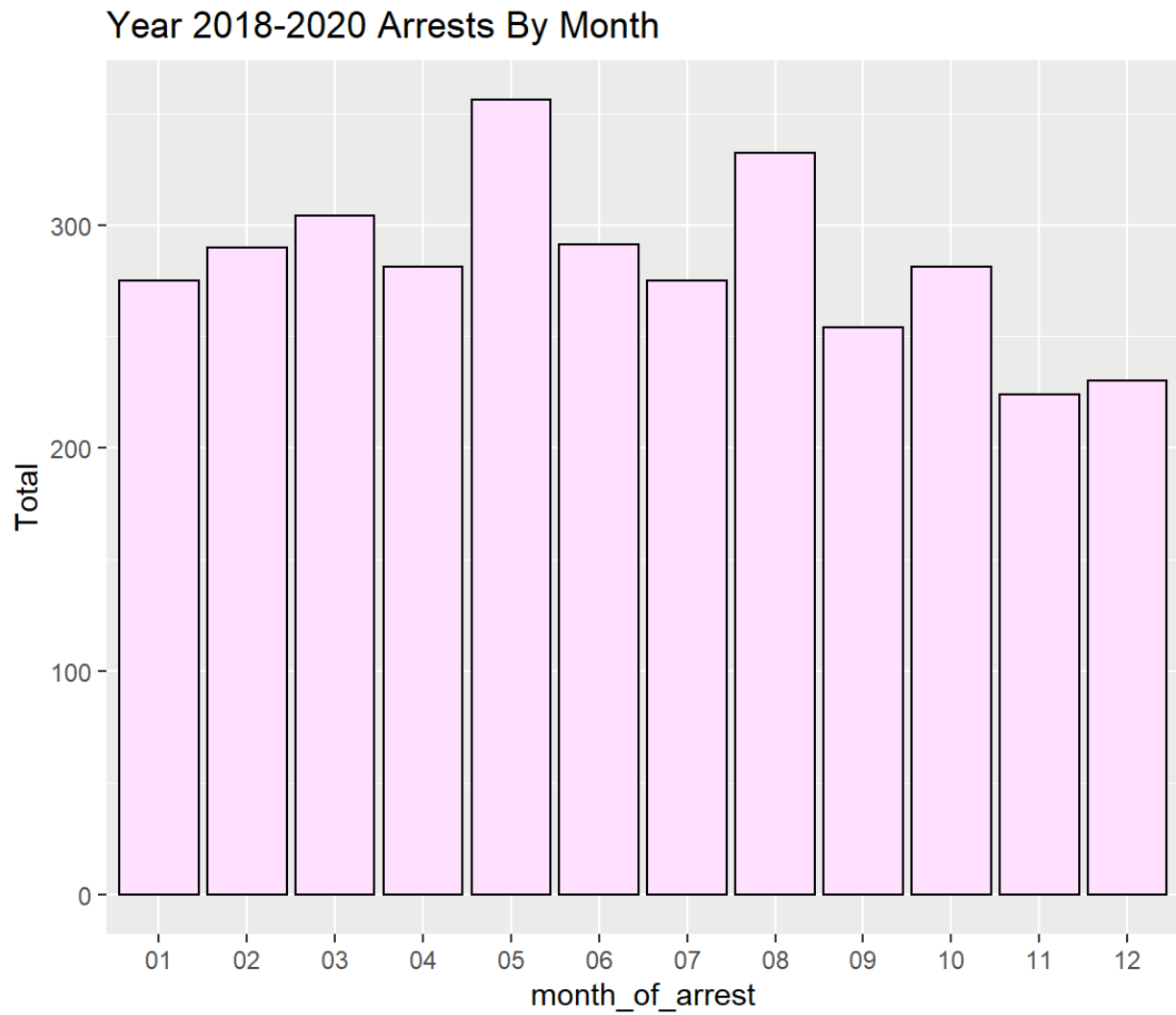


**Figure 8.** Number of arrests made in 2020 by month



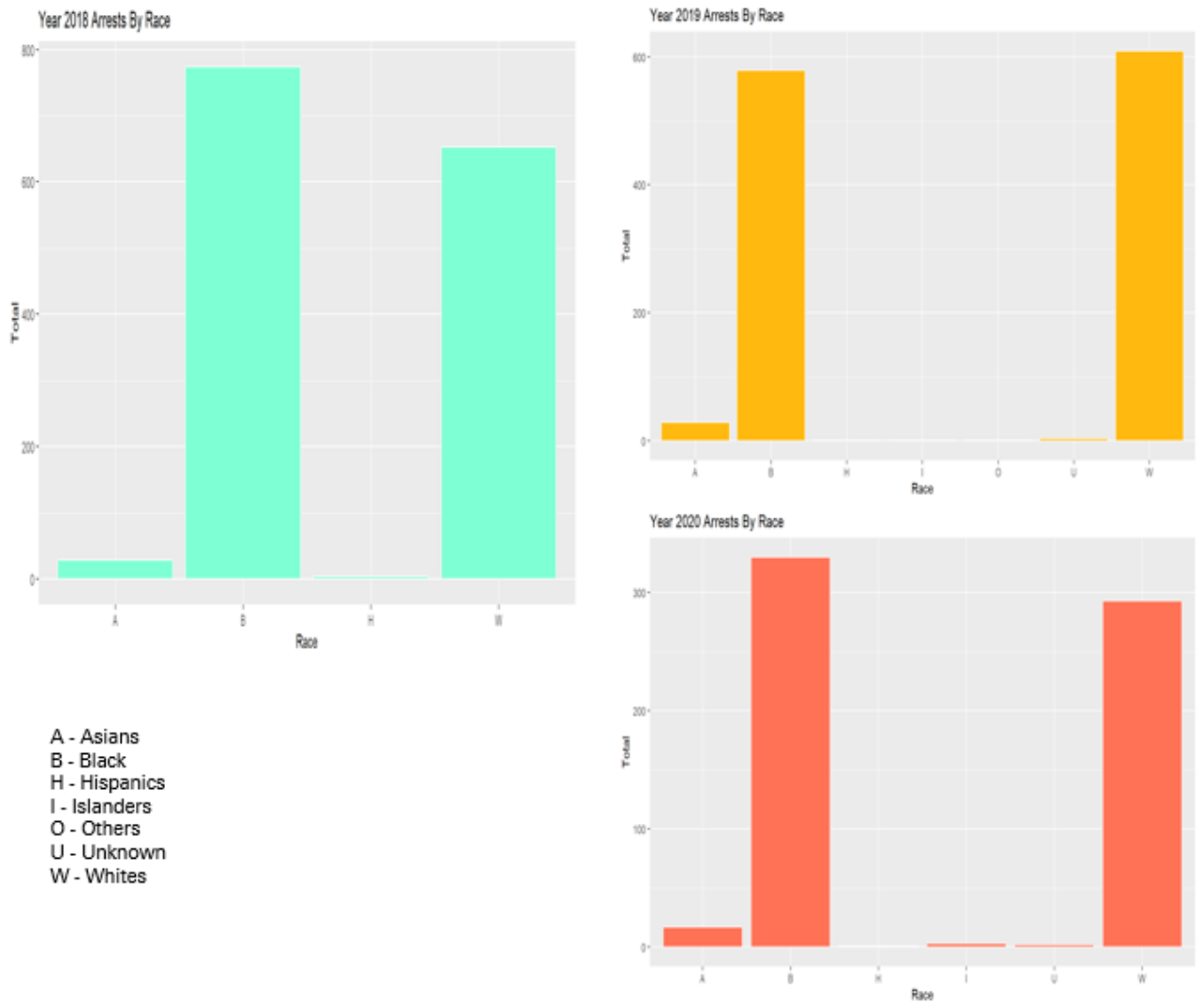


**Figure 9.** Number of arrests made in 2018 - 2020 by month

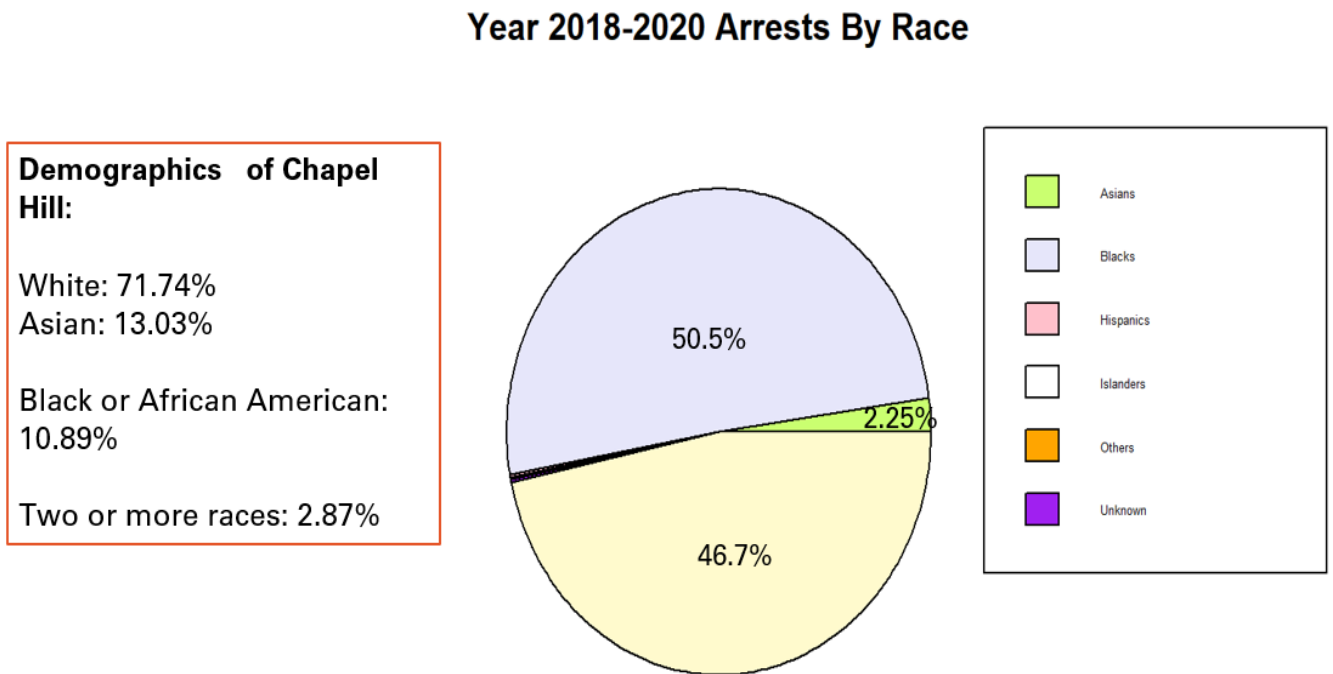




**Figure 9.** Number of Arrests by Race for each 2018, 2019 and 2020



**Figure 10.** Number of Arrests by Race for 2018-2020



**Figure 11.** Age analysis of the arrestees by Race for 2018-2020

<b>Race</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>Median</b>
A	30.48	71.00	18.00	25.00
B	38.51	79.00	18.00	35.00
H	28.00	46.00	18.00	27.00
I	32.00	65.00	20.00	21.50
O	34.00	34.00	34.00	34.00
U	32.17	45.00	18.00	34.00
W	37.19	87.00	18.00	35.00
<b>Average:</b>	<b>33.19</b>	<b>61.00</b>	<b>20.57</b>	<b>30.21</b>

**Figure 12.** Age analysis of the arrestees by Gender for 2018-2020

<b>Gender</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>Median</b>
F	34.26	79.00	18.00	31.00
M	38.60	87.00	18.00	36.00
<b>Average:</b>	<b>36.43</b>	<b>83.00</b>	<b>18.00</b>	<b>33.50</b>

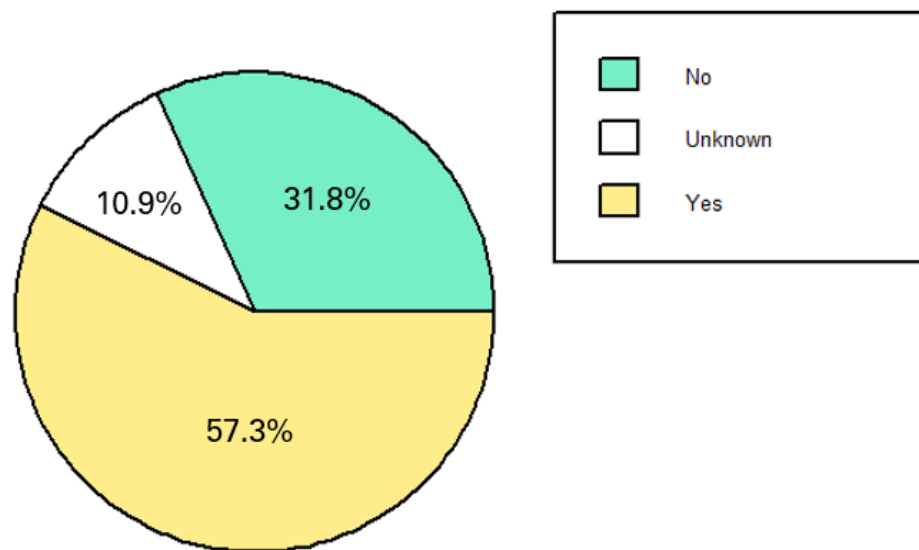
**Figure 13.** Weapon Analysis of the arrestees for 2018- 2020

Weapon Present	Total
CLUB/BLACKJACK/BRASS KNUCKLES	8
FIREARM (TYP NOT STATED)	8
HANDGUN	64
LETHAL CUTTING INSTRUMENT	29
OTHER FIREARM	1
RIFLE	1
SHOTGUN	2
UNARMED	3105

**Figure 14.** Under the Influence of Drug or Alcohol Analysis for 2018-2020

## Year 2018-2020 Arrests By UI

UI = Under the Influence of Drugs or Alcohol



**Figure 15.** Top ten crimes committed in Chapel Hill for 2018-2020

	Primary_Charge	Total
1	IMPAIRED DRIVING DWI	350
2	OPEN CONTAINER	248
3	FAIL TO APPEAR/COMPL	247
4	2ND DEGREE TRESPASS	135
5	ASSAULT ON FEMALE	120
6	POSSESSION MARIJUANA	104
7	LARCENY - ALL OTHER	101
Showing 1 to 7 of 460 entries, 2 total columns		

8	ASSAULT-SIMPLE	77
9	DRUG PARAPHERNALIA	65
10	PUBLIC CONSUMPTION	57

### References

- Krivo, Lauren J., and Peterson, Ruth D. “Extremely Disadvantaged Neighborhoods and Urban Crime.” *Social Forces*, vol. 75, no. 2, 1996, pp. 619–48, <https://doi.org/10.2307/2580416>. Accessed 28 Apr. 2022.
- Perkins, H. Wesley. “Religious Traditions, Parents, and Peers as Determinants of Alcohol and Drug Use among College Students.” *Review of Religious Research*, vol. 27, no. 1, 1985, pp. 15–31, <https://doi.org/10.2307/3511935>. Accessed 28 Apr. 2022.