

Southwest DC City Council

Spring 2023 CS506 Data Science

Harshitha Tumkur Kailasa Murthy	- U00683580	- harshutk@bu.edu
Vishwas Bhaktavatsala	- U74206902	- vishwasb@bu.edu
Kristopher	- U36116166	- krischou@bu.edu
Yeh-Min Lu	- U70603446	- leolu@bu.edu

SL No	Particular
1.	The Big Picture
2.	Base Analysis
3.	Project Goal & Overview
4.	Extension Analysis
5.	Visualizations & Graphs
6.	Team Members Contribution
7.	Conclusion & Future Work
8.	Data Sources

1. The Big Picture

Southwest DC is a neighborhood that has seen tremendous change over the last 60 years. In our words, Southwest DC was the heart of the African American community until the 1950s. According to the National Housing Act of 1949 the Federal government decided and issued funds to restructure the whole of Southwest. This was the massive urban renewal project in Southwest DC which displaced over 20,000 households and changed the character of the area. The government attempted to eliminate the alley dwelling lifestyle and tried to improve the situation, which leads to many challenges as listed below post the URBAN RENEWAL of the area.



Fig 1. Old Southwest

1. Due the restructuring, the land value increased with the high rise and government buildings which unfortunately meant that the poor people could not afford to buy houses or live there and hence they got scattered.
2. This also meant that the street life was completely sucked out of SW DC with all the alley dwelling being removed.
3. people had to travel for a very long time for just a cup of coffee as the restaurant or a shop which was there right around the alley were removed.

Initial rebuilding in the 1970s focused on large-scale projects featuring distinctive mid-century modern architecture. Southwest became known as a quaint residential neighborhood along the waterfront with a significant federal government footprint. It wasn't until the 2000's that DC's development boom reached Southwest, bringing a redesigned Arena Stage, the new Waterfront

Station complex, the \$3.6 billion District Wharf project, the south campus of the Smithsonian (Hirschhorn, Freer, Sackler, Air & Space, American Indian, & African Art), the Rubell Museum, International Spy Museum, Museum of the Bible, Audi Field (home of the DC United), and the redevelopment of Buzzard Point, Southwest's southern peninsula that includes historic Fort McNair.

This "neighborhood within a neighborhood" is known as "Old Southwest," and faces many of the common problems associated with concentrated generational poverty – high rates of unemployment, community violence, substance use, chronic disease, teen pregnancy, and food insecurity.

2. Base Analysis

Initially we started off with exploring some basic data sets that were given by the client. We explored various parameters like population, Age, poverty, Income, crime, affordable housing, and method of preferred transportation. By looking at these parameters, we can get a better idea of the economic vitality of the district. Using visualizations such as graphs and maps we were able to find many correlations.

Few overarching questions that were asked in the initial stages of the project.

1. What are the priority parameters to be considered for Analysis?

Initially, all were focused on all the data sources that were given the project documentation, later the focus was set to only a few sources.

2. What impacted areas is the client really interested in focusing on?

Clients were interested to see the story of Southwest with respect to poverty and mainly in line with Economic vitality and Crime.

3. What are the regions in the southwest where the most focus must be applied?

Ward 6 was one of the wards we focused more on. This also happens to be the ward with the most concentrated affordable housing. We started off by analyzing the ward entirely but we wanted to look deeper, so we started focusing on specific tracts in the ward. The focused tracts are 105, 110.01, 110.02, and 64.

4. What are the challenges and disparities of the neighborhood?

The restructuring of the whole neighborhood during the 1970s caused challenges for the locals in the area. The new plan had clubbed all the restaurants together and the residential areas together. Alley dwellings disappeared causing the cost of the area to escalate. The main challenges were:

- Due the restructuring, the land value increased with the high rise and government buildings which unfortunately meant that the poor people could not afford to buy houses or live there and hence they got scattered.
- This also meant that the street life was completely sucked out of SW DC with all the alley dwelling being removed.
- People had to travel for a very long time for just a cup of coffee as the restaurant or a shop which was there right around the alley were removed.

3. Extension Analysis

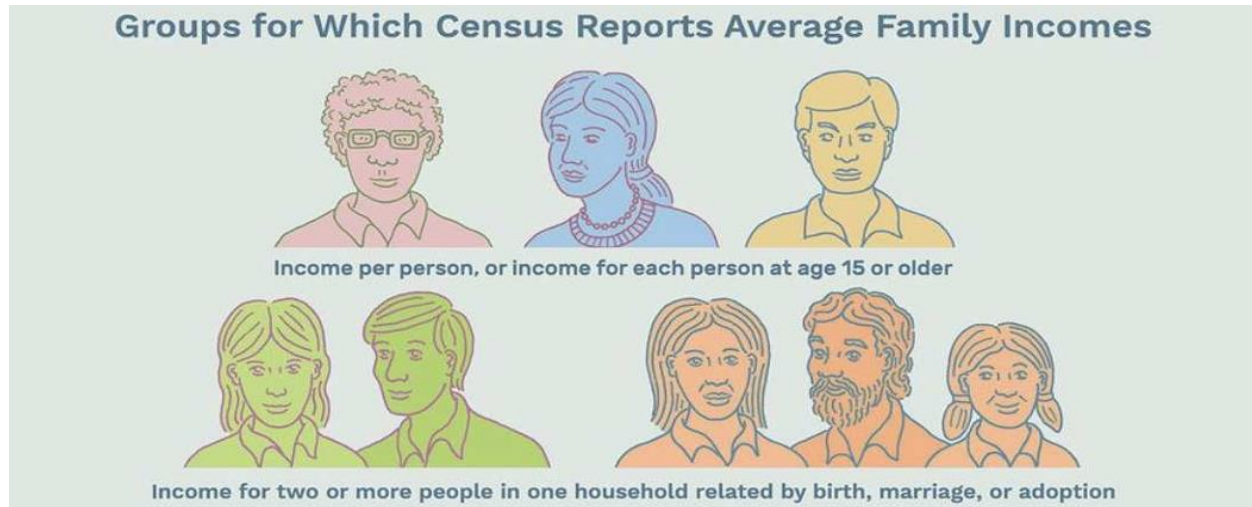
One of the overbearing questions which the clients wanted us to answer was that of the crime and what it looks like in the Southwest region of Washington DC. The Initial approach for us was to find some correlation between the crime which occurred in the region with other important aspects like, poverty and housing.

Our initial correlation focused on trying to find the relationship between affordability and crime in Southwest DC. Crime was one of the main focuses for us as DC as it was one of the important areas which the client wanted to improve on to make Southwest DC a more equitable, inclusive and thriving community.

Initial analysis involved us looking at each of the aspects of population, age and crime individually. For further analysis, we investigated how crime was distributed over different age ranges and how the population of the area has contributed to the crime statistics. Adding onto this, we also looked at whether nurturing factors such as affordable housing, child abuse, and child neglect influenced crime in the area.

4. Visualization and Graphs

1. Family Income Trends



Considering the various backlashes of the restructuring of Southwest DC, many individuals and families were displaced. The economic stability of DC was shaken and with this new restructuring, the cost of the residential areas shot up quite high. To investigate the effects of this, we started with analyzing the income of each of the families in the main regions or tracts of Southwest DC. We tried to look at the distribution of income across Households, Families, Married-Couple Families and Nonfamily Households.

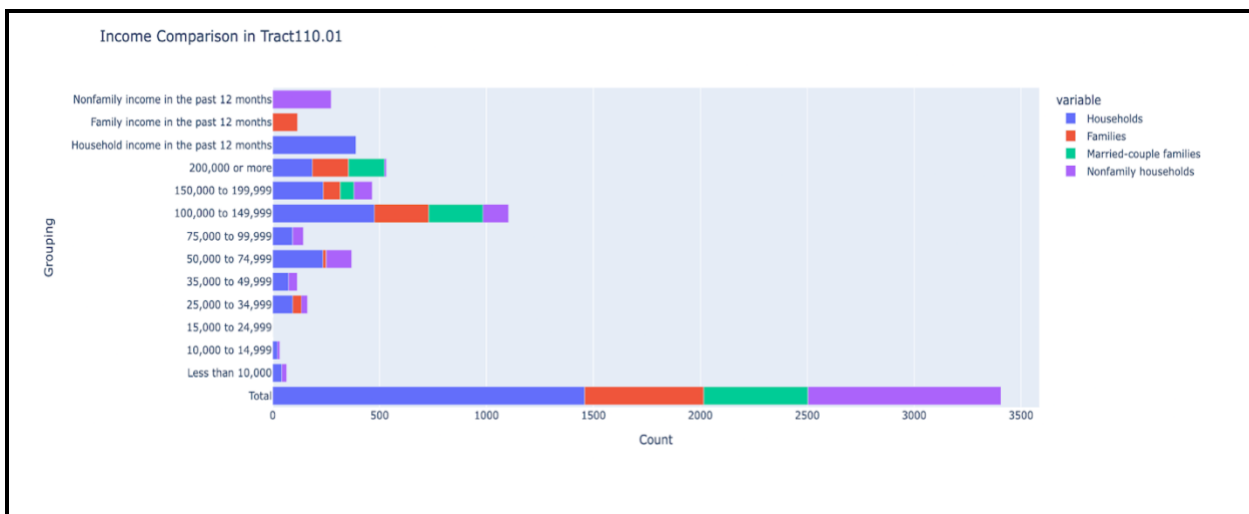


Fig 2. Income Comparison in Tract 110.01

Here, Fig 2 clearly shows the distribution for Tract 110.01. In this tract, most of the households and families have an income within the range of \$100,000 to \$149,999. Although this seems to be a better case in Tract 110.02 where the income distribution is evenly spread over the ranges

\$50,000 to \$200,000 or more. This tract has households and families which have incomes over \$200,000. Figure 3 below, shows this visually.

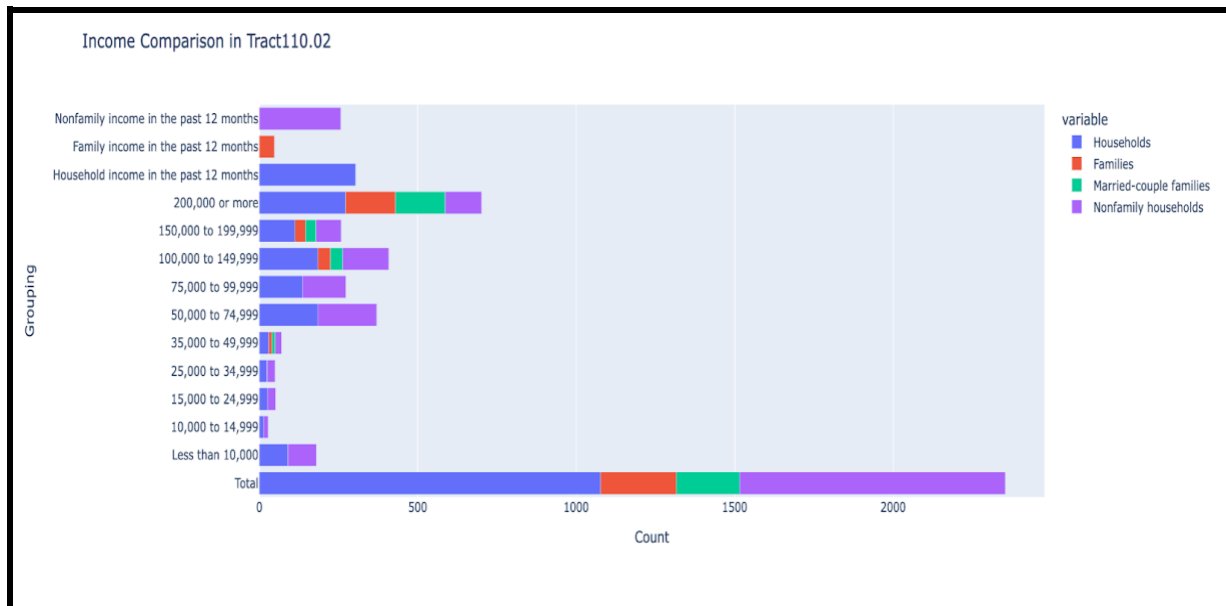


Fig 3. Income Comparison in Tract 110.02

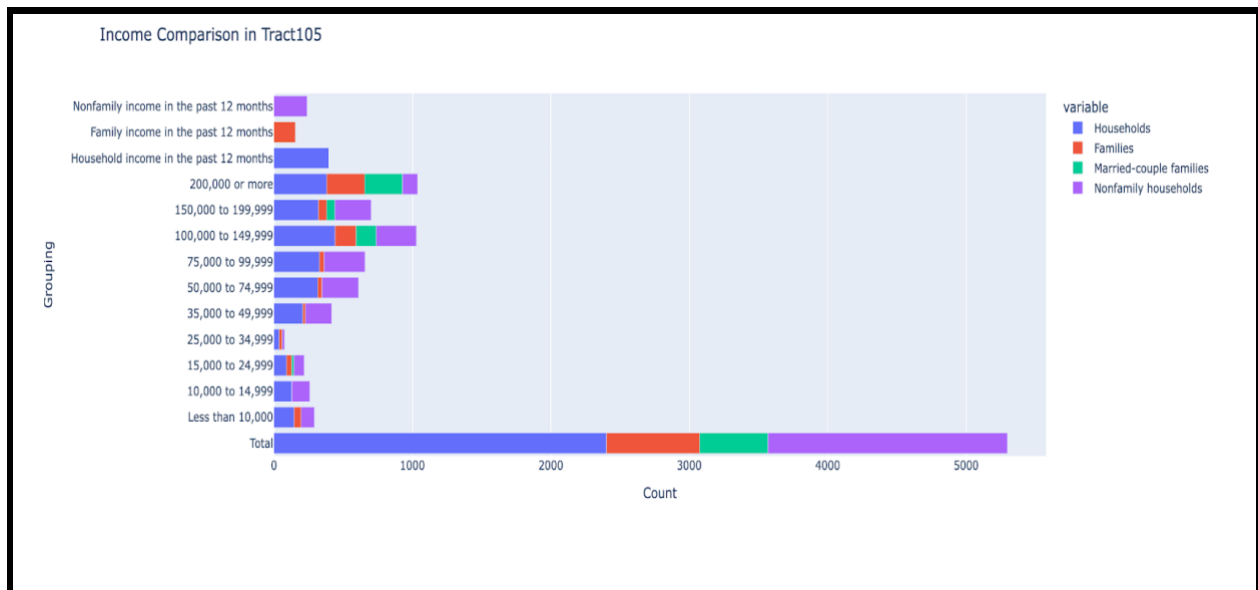


Fig 4. Income Comparison in Tract 105

Figure 4, which shows the distribution for Tract 105, showed us that this tract had very few families and households which have less incomes. Overall, all the tracts which we have seen until now had decent incomes to support the updated lifestyle of the neighborhood.

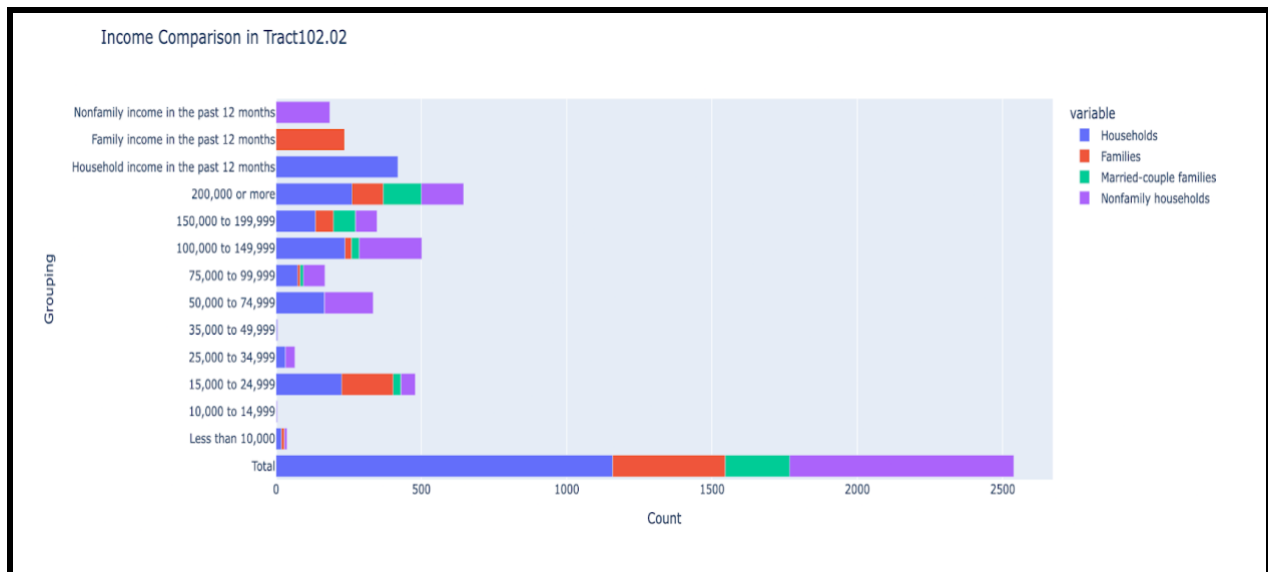


Fig 5. Income Comparison in Tract 102.02

Considering the placement of the following two Tracts, 102.02 and 102.01 as described by the two graphs, Figure 5 and 6, we were expected to find something of value. As expected, Tract 102.02 shows a considerable difference in the number of households and families which can be considered as being in the gray area. We found that some of them have a low overall income of \$15,000 to \$24,999. Although, this is not the same in Tract 102.01, we many families and households have comparatively lesser income when compared to others.

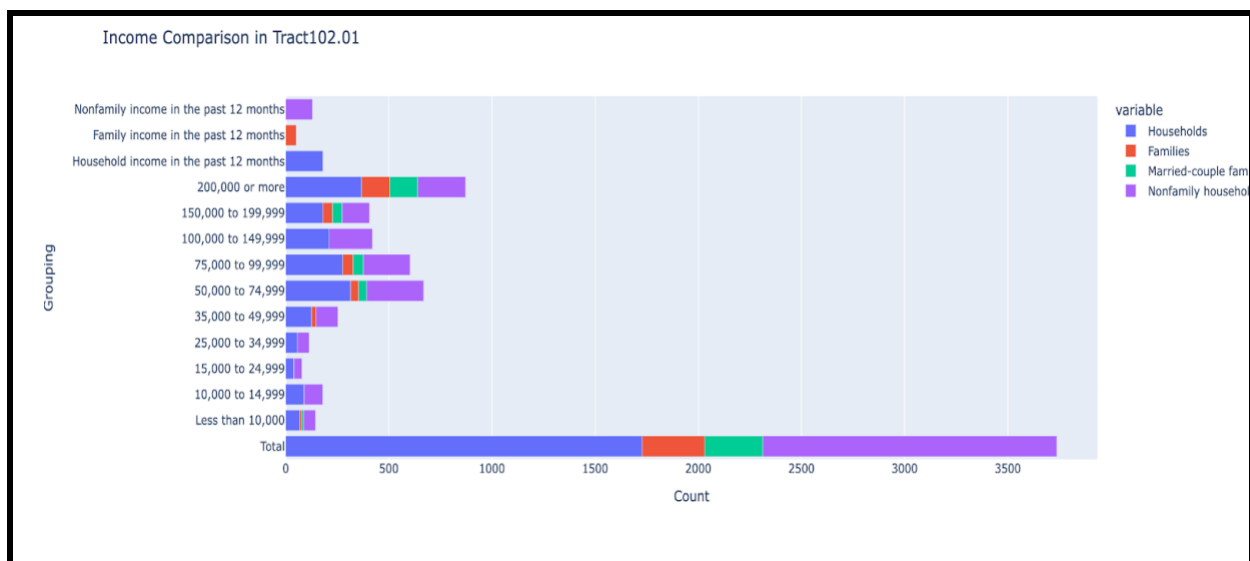


Fig 6. Income Comparison in Tract 102.01

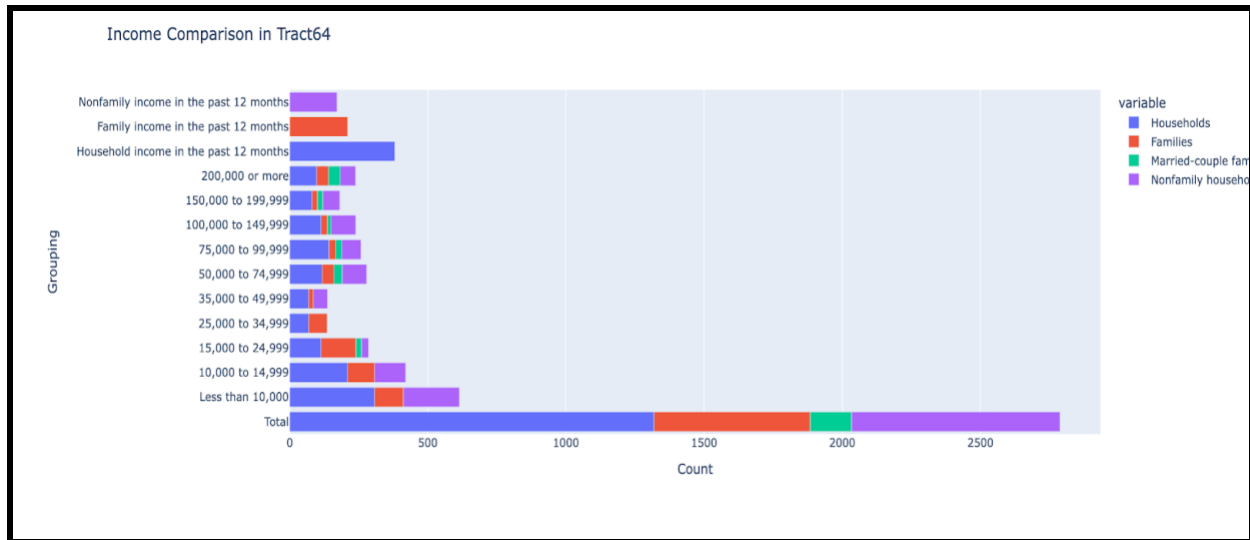


Fig 7. Income Comparison in Tract 64

Figure 7 depicts the same for Tract 64, This tract has most of the families which have income under \$49,999. The main red area has been in the income bracket of \$10,000 or lesser where, majority of the households have income problems.

As we can see from the plots, we can understand that majority of the households and families in Tract 64 and Tract 102.02 have less than normal incomes.

2. Population Trends



Our next focus to understand the situation in this community and neighborhood was to look into the population of each of the tracts and how they are distributed. We looked into this so that we can analyze and interpret the effects of population on the neighborhood. If we better understand the population of each of the tracts, we should be able to track the effects of Income and population on other aspects of the community.

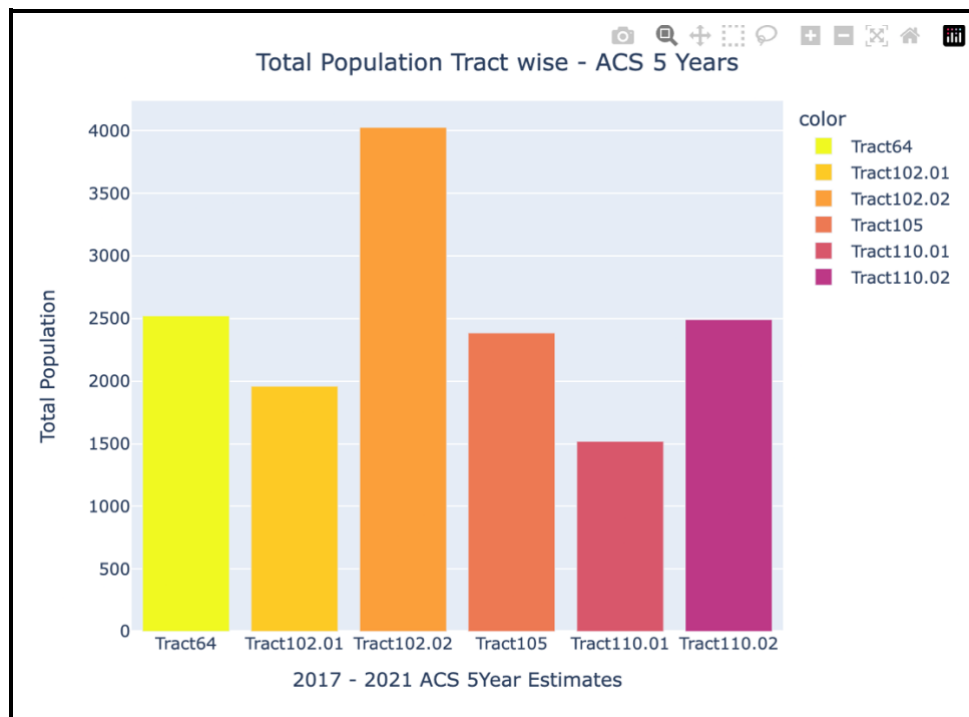


Fig 8. Total population Tract wise for past 5 years

Fig 8 tries to describe the population as provided by ACS sources, which is the average of 5 years, which spanned over 2017 to 2021. In this period Tract 102.02 is comparatively more populated over the other tracts under question. Tract 64 which was under focus in the previous sections, averages around 2500 units.

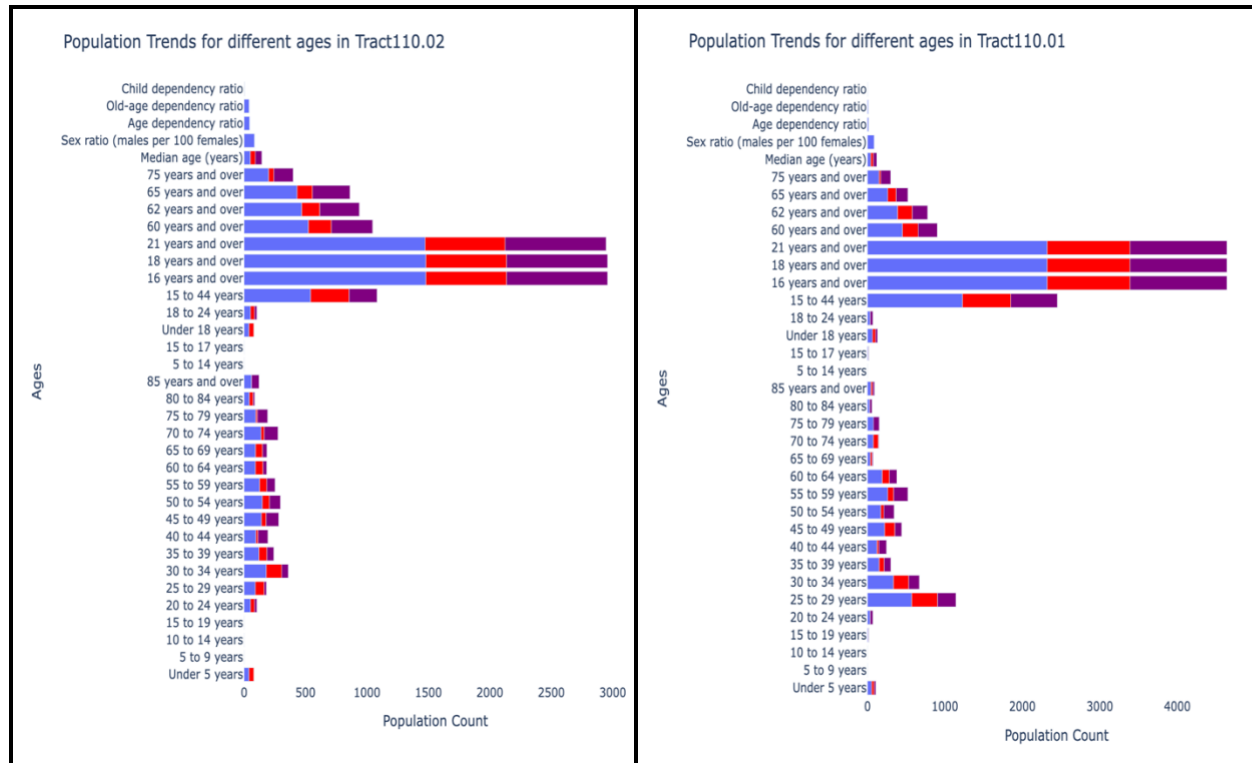


Fig 9. Population Trends for different Ages for Tract110.02 & Tract110.01

Figures 9, 10 and 11 show the population trends for all the tracts. One aspect which we observed, most in all the tracts, is the number of people in the age range of 15 to 44 years. We can safely say that, although all age groups are present, all the tracts are housing many people between the ages of 15 to 44. Upon further analysis, the granular age group of 25 to 29, the young adults are more in number when compared to others. This age is especially important as they can be influenced towards negative aspects. We can further investigate this trend in further graphs.

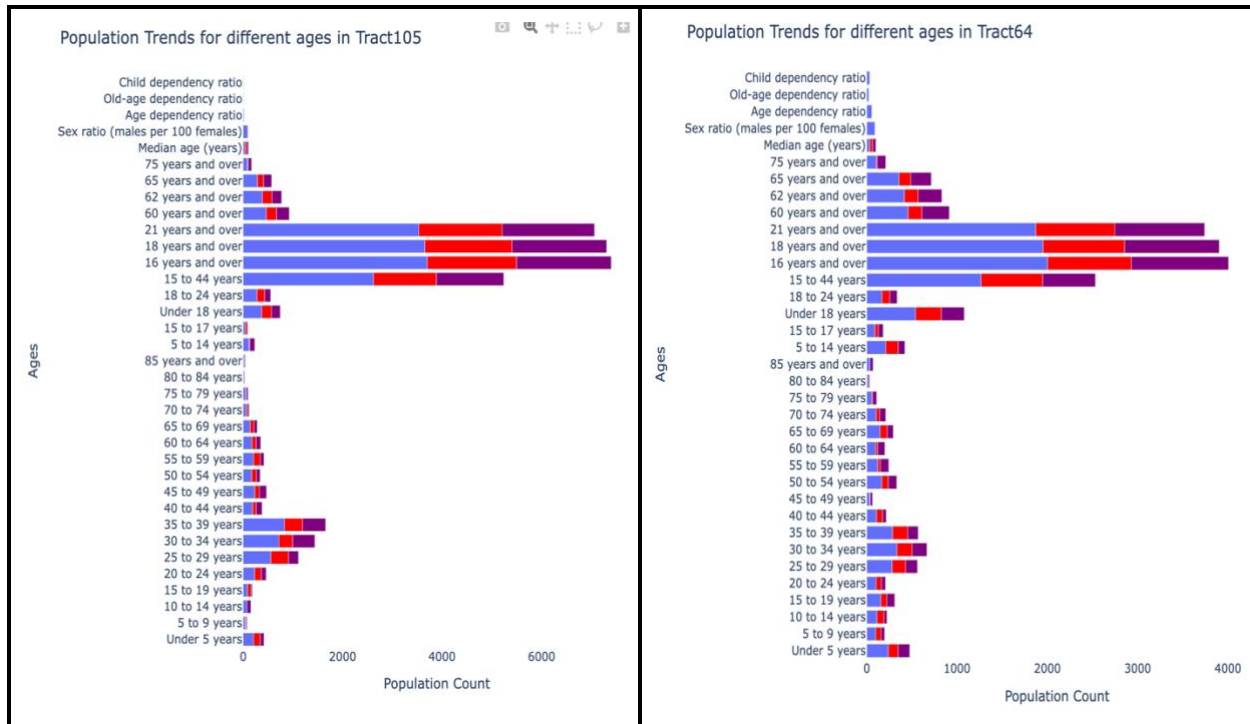


Fig 10. Population Trends for different Ages for Tract105 & Tract64

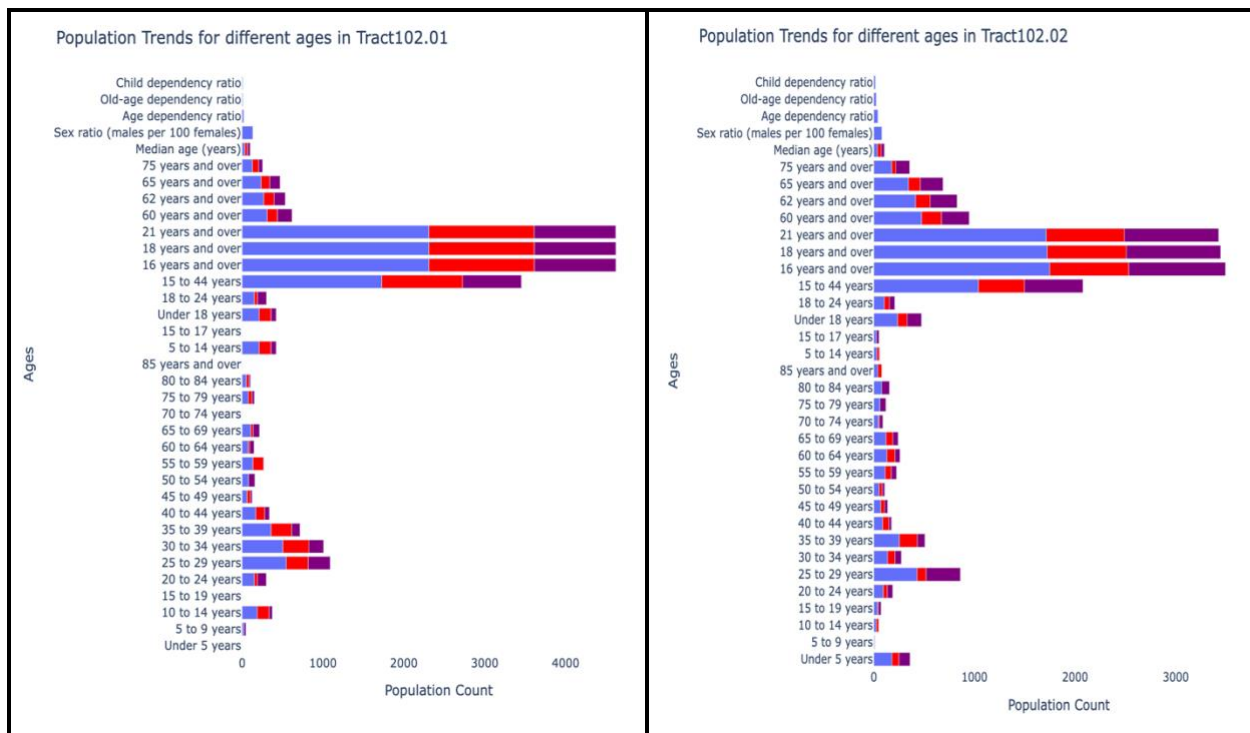


Fig 11. Population Trends for different Ages for Tract102.01 & Tract102.02

3. Poverty Trends

As we analyzed the Income trends and the population trends for all the tracts which are in focus here in the Southwest region of DC, we started to wonder what poverty looks like in these regions. Since most of the income trends looked good, we expected the numbers to look decent in a few of the tracts. Here, we were looking forward to seeing the main effect on Tract 64 and Tract 105.

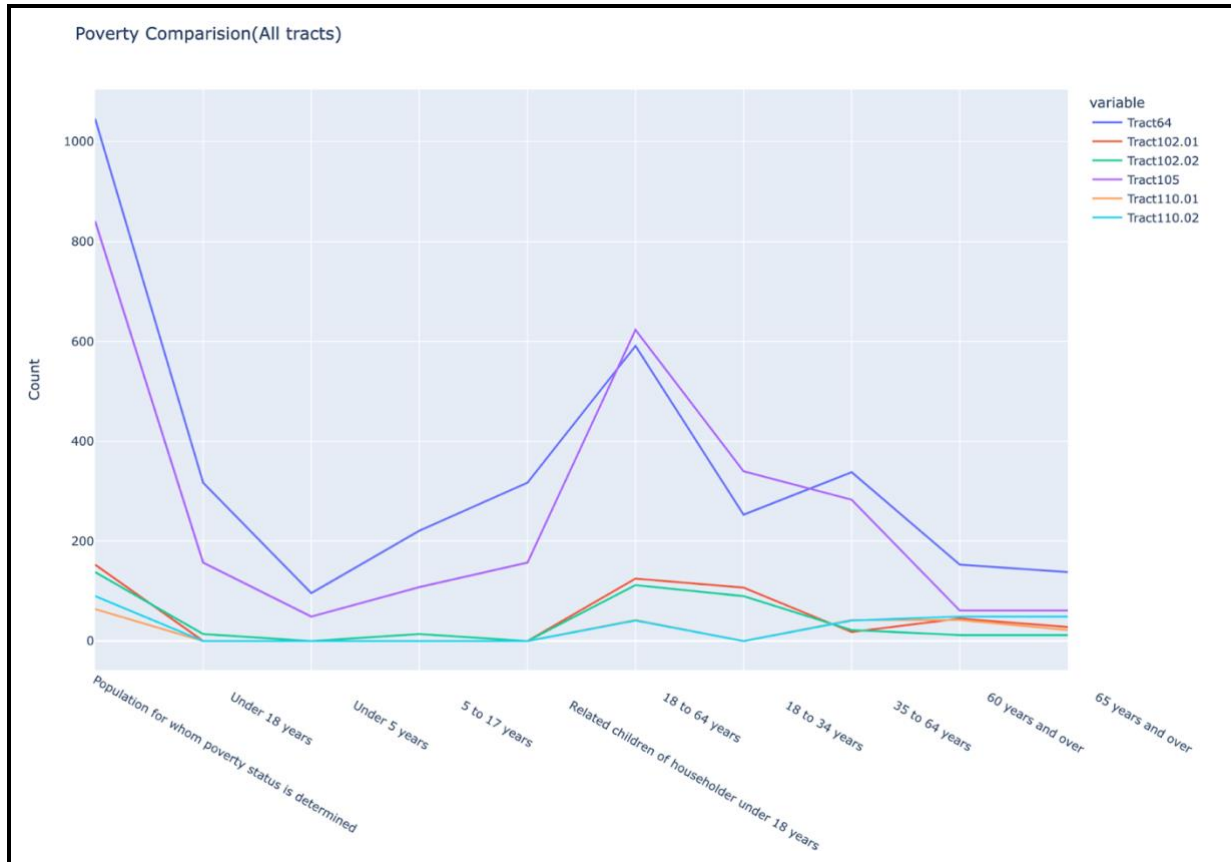


Fig 12. Poverty Comparison for all the tracts

As expected, Fig 12 clearly showed the pain areas for Southwest Washington DC. While the tracts 102.02, 102.01, 110.01 and 110.02 have a smaller number of people under the poverty level, this isn't the same case for Tracts like 64 and 105. These Tracts have an elevated number of people who are below the poverty line. Tracts 64 and 105 have a huge majority of people within the age ranges of 18 to 64. Although this is a very huge range, this is quite significant. On a more granular level, the age groups of 18 to 34 years have many people which must be the backbone of the society. This can mean high rates of unemployment and crime as the people of this age groups are influenced a lot when they have nothing to do and a family to look after.

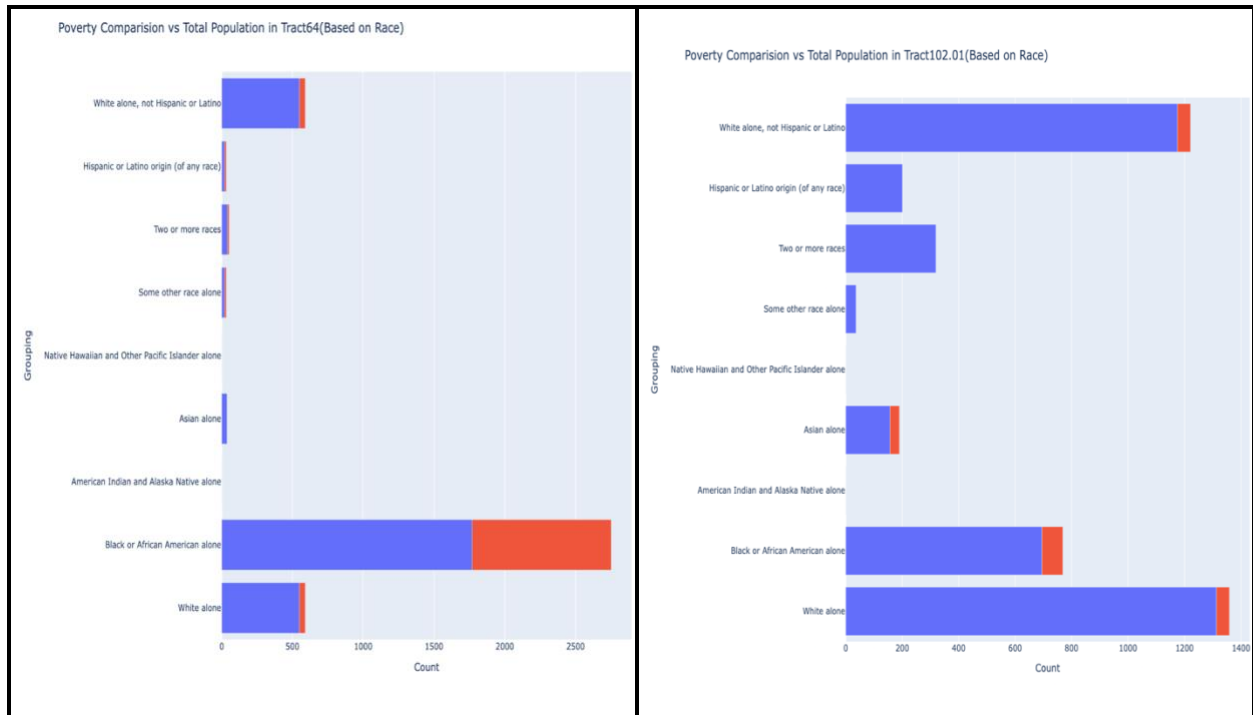


Fig 13. Population Trends for different Races for Tract102.01 & Tract102.02

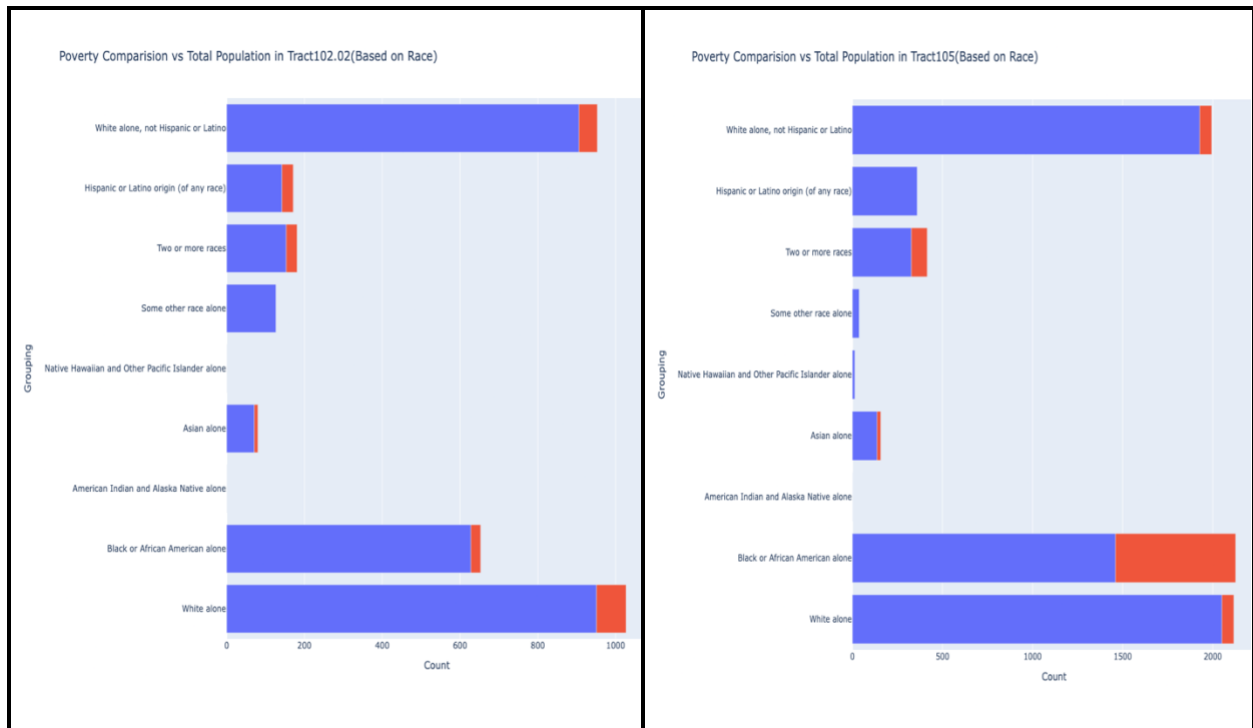


Fig 14. Population Trends for different Ages for Tract102.01 & Tract102.02

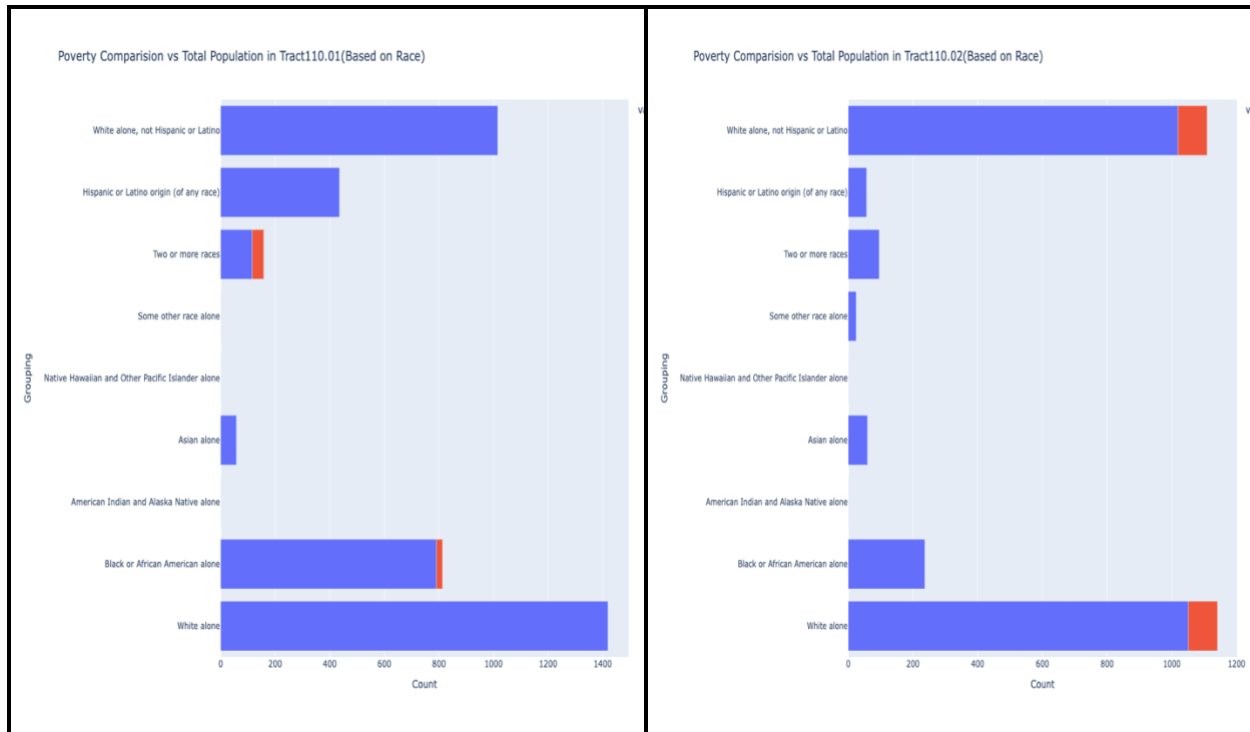


Fig 15. Population Trends for different Ages for Tract102.01 & Tract102.02

Figures 13, 14 and 15 show the distribution of poverty over all the races and their combinations when compared to the total population for them. Almost all the Tracts are dominated by white and African American Residents. One major takeaway from all these graphs is that Tract 105 has a lot of African Americans who are below the poverty line accompanied by Tract 64 as well. Additionally, Tract 64 is the only tract which is dominated by only African Americans. This tract also has a lot more people who are below the poverty line.

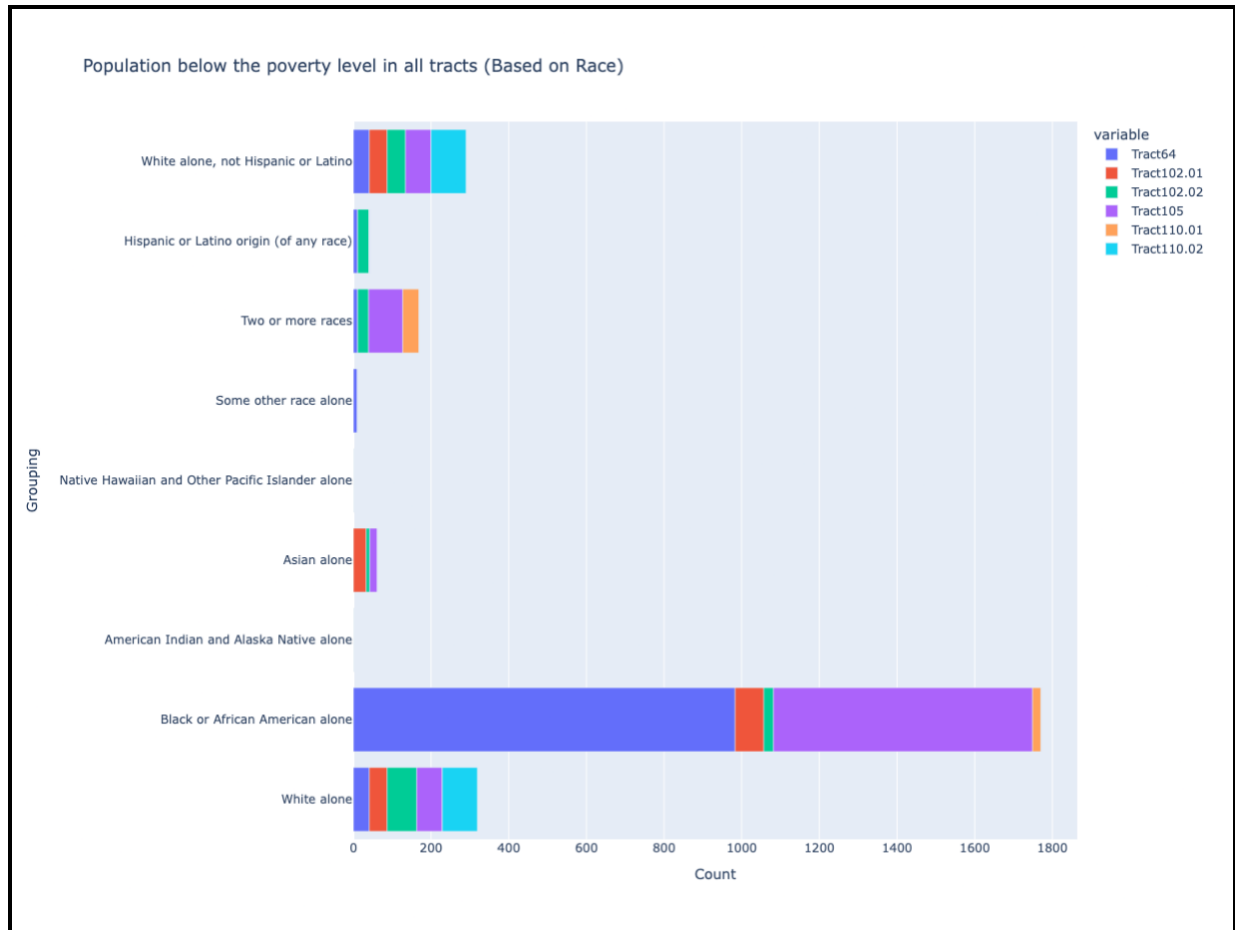


Fig 16. Population Trends below poverty level in different Tracts

Figure 16 clearly shows the population which is under the poverty line in all the tracts. All the bars depict different races. As discussed already, Tracts 64 and 105 are the main tracts which we can look into. Black and African American races are being dominated in these two tracts and as we know that due the restructuring of Southwest DC, these races were affected the most.

4. Crime Trends

Once, we had a clear understanding of the majority factors of which may have contributed to the current state of DC, the main focus shifted towards the elephant in the room, which is the crime stats of the area. For our initial analysis, we used the DC Crime cards data to get an overview of the crime in the area. Ward 6 being the main area of focus, we were able to understand the following.

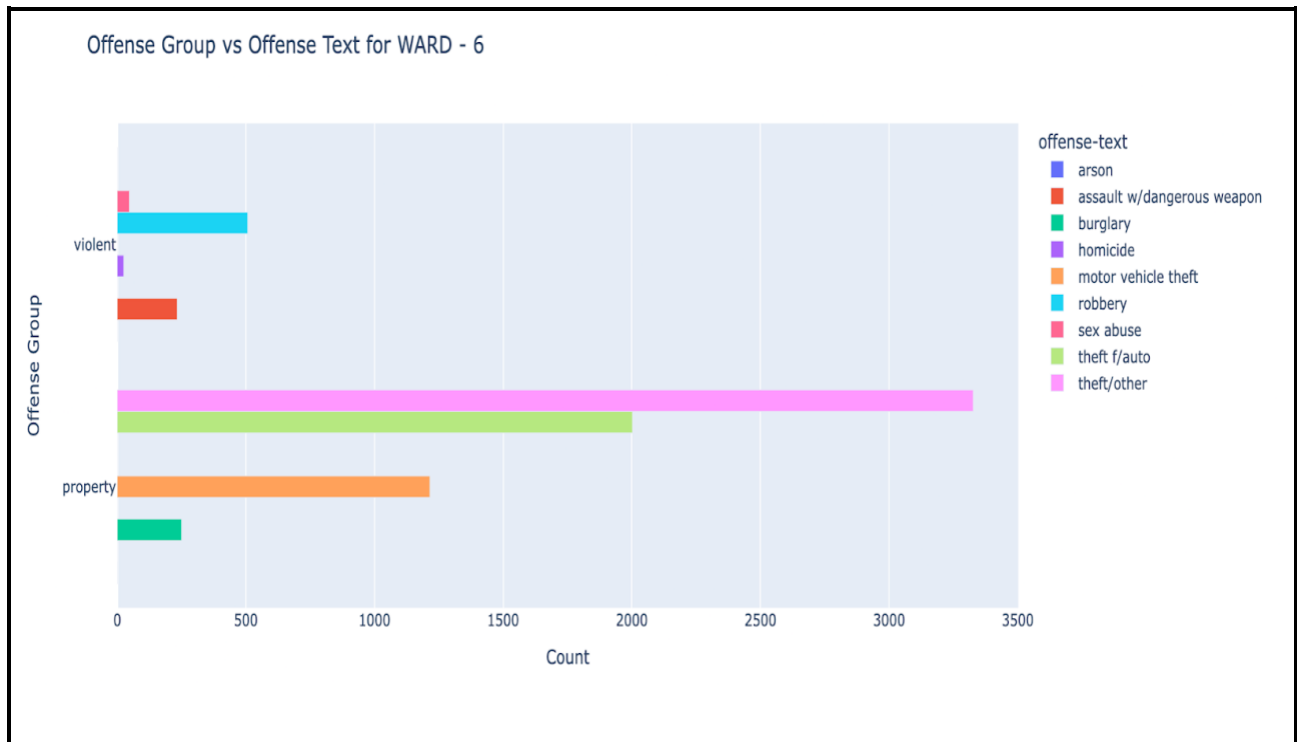
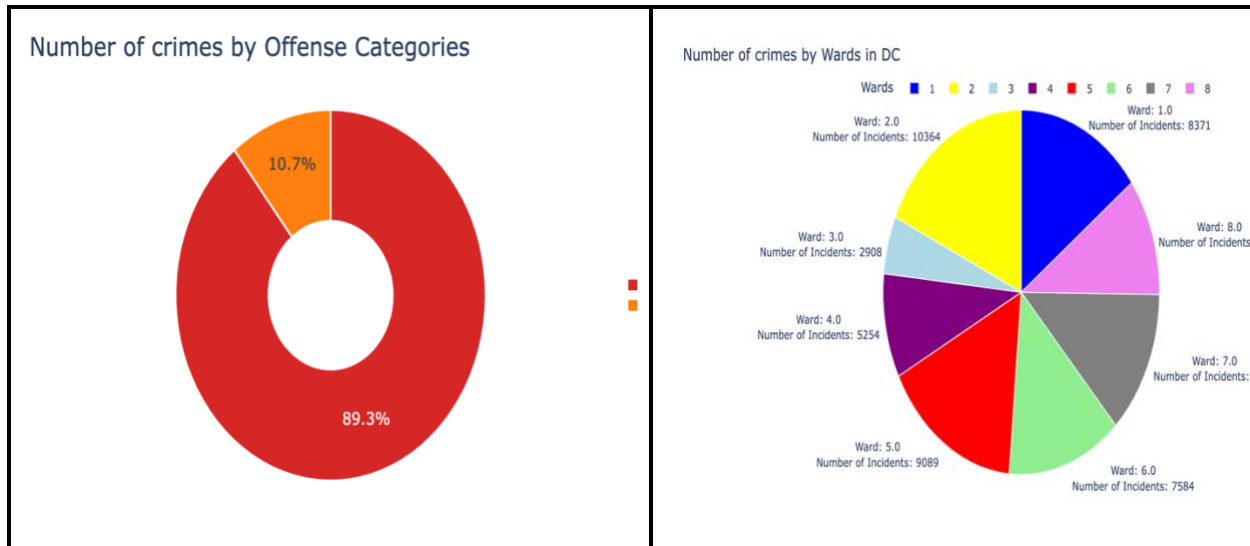


Fig 17. Offense Groups in WARD 6

In Figure 17, the major categories of crime i.e., property and violent crimes were used to investigate the crime in the area. Ward 6 had most property crimes which varied from theft to motor vehicle crime. Over the past two years, ward 6 has witnessed around 3300 theft incidents and over 2000 motor vehicle crimes. While property crimes have been quite high, violent crimes such as robbery and sex abuse have been quite high when compared to others in the region.



**Fig 18. i) Number of crimes by Property and Violent categories
ii) Ward wise Crime Trends in District of Columbia**

Figure 18 depicts the donut and pie chart for the number of crimes in ward 6 based on the offense category. As analyzed in the previous graphs, 89.3% of the incidents are property crimes while only 10.7% are violent crimes. Ward 6 contributes to around 7500 incidents of the whole crime in the past two years.

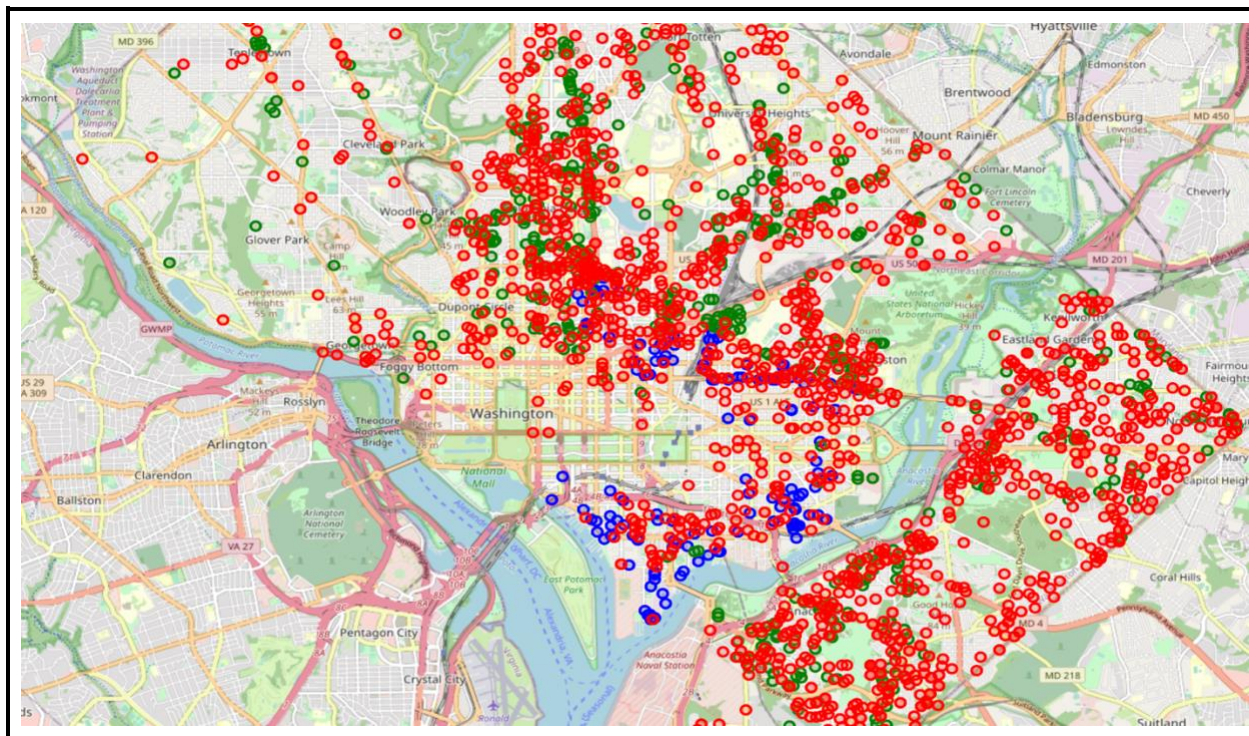


Fig 19. Map showing the location of crime, and affordable housing locations.

Figure 19 shows the location of crime (red markers) where the method of crime is by using a gun. It also shows the location of affordable housing in the district, labeled using blue and green markers where the blue markers are affordable housing in ward 6 and green markers are affordable housing in wards that aren't ward 6. We can see the correlation between affordable housing location and gun violence locations, as there are more gun violence situations near areas of affordable housing.

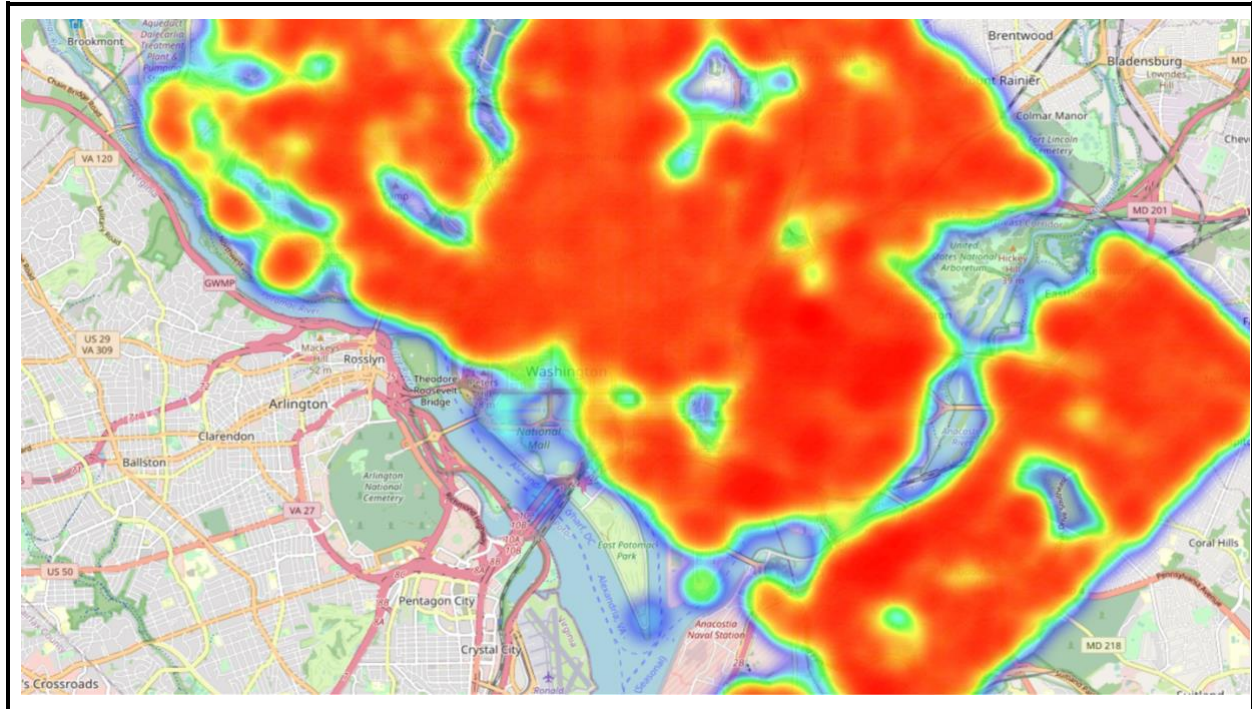


Fig 20. Heat map of all crime in Washington DC

Figure 20 map shows the concentration of all types of crime in the district. The redder the area of the map is, the more concentrated that area is. The map is interactive, allowing the user to zoom in and out of the map, seeing which areas are concentrated with crime. They are also able to move along the map and focus on areas that users want to put focus on.

5. Education

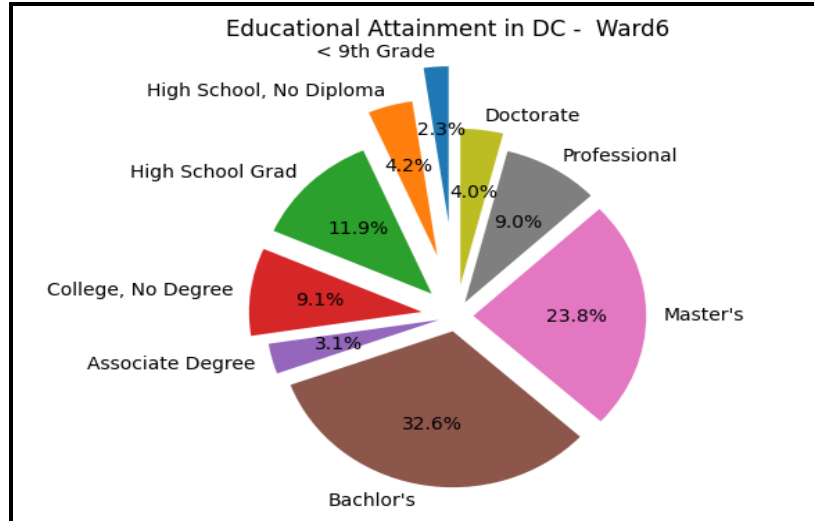


Fig 21. Educational attainment in DC - Ward 6

Fig 21 shows the educational attainment at the percentage of the population Age25+. According to the graph, over 93% of Ward 6 residents aged 25 years and older are high school grad or higher which is a little higher than the rate in DC: 92.2% and about 10% higher than the rate in the US: 88.9%. Furthermore, over 30% of Ward 6 residents hold a graduate or professional degree, which is significantly higher than the national average of around 13.1%. This figure suggests that ward 6 has a highly educated population, which may be attributed to the presence of numerous universities and research institutions in the area.

6. Economic Well-being

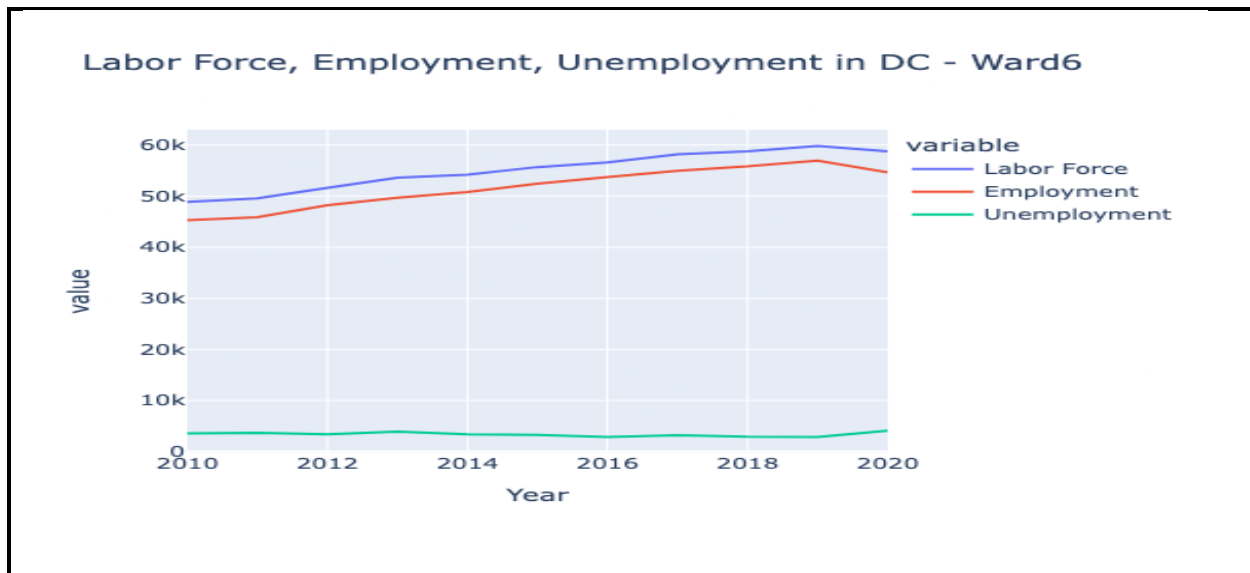


Fig 22. Labor Force, Employment, Unemployment in DC - Ward 6

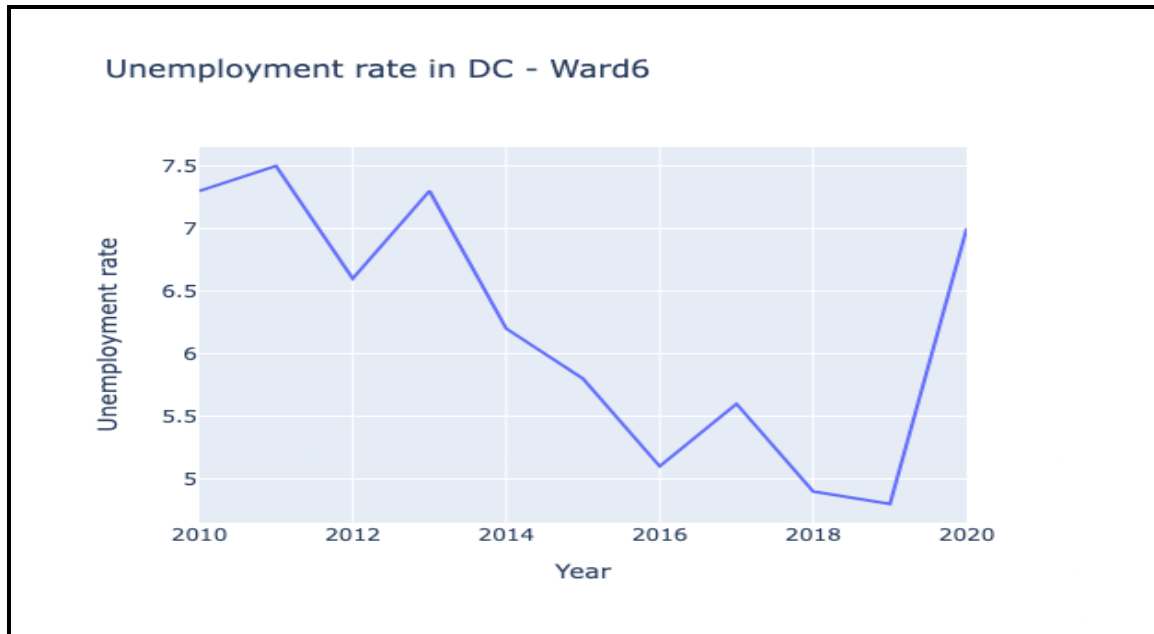


Fig 23. The unemployment rate in DC - Ward 6

Figs 22 & 23 show the basic economic well-being graphs including labor force, employment, and unemployment lines. As we can see that the unemployment rate rocketed during the pandemic year, as of September 2021, the unemployment rate in DC ward 6 was 5.5%. This is slightly lower than the overall unemployment rate in the District of Columbia, which was 5.7%. However, it's important to note that these rates may continue to fluctuate due to ongoing economic recovery efforts and the impact of the COVID pandemic on the labor market.

Moreover, changes in the unemployment rate reflect shifts in job creation or economic growth. Higher levels of unemployment may lead to higher levels of crime. This is because individuals who are unemployed may experience financial strain and social isolation, which can increase their likelihood of engaging in criminal activity.

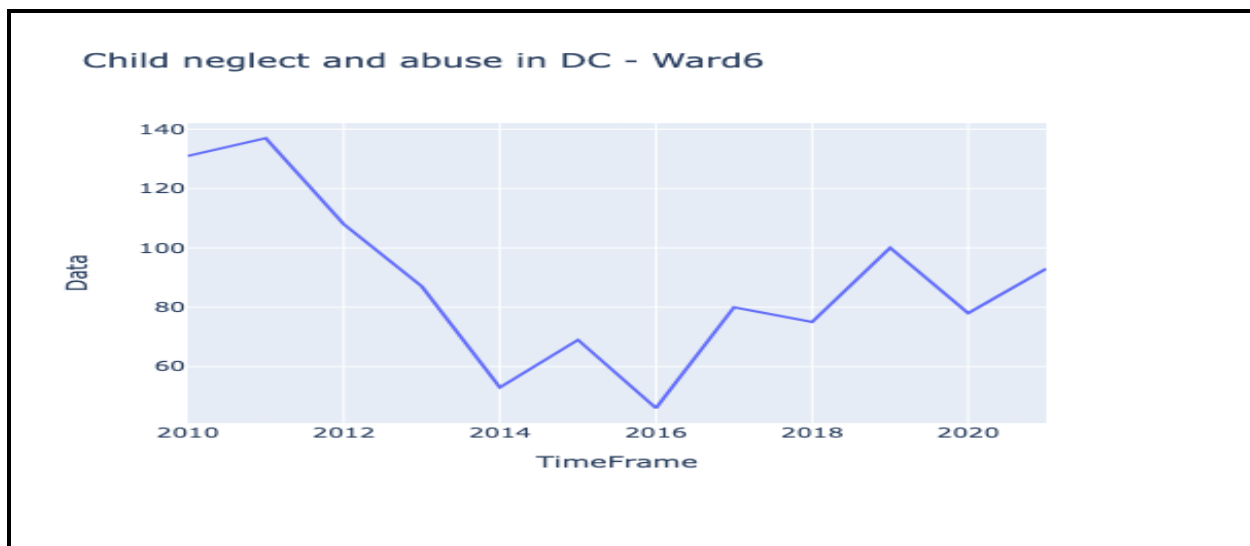


Fig 24. Child Neglect and Abuse in DC - Ward 6

Fig 24 shows the substantiated investigations of child neglect and abuse. Physical or mental injury, sexual abuse, negligent maltreatment, or maltreatment of any child under age 18 by a person responsible for the child's welfare under circumstances that threaten or harm the child's health or safety.

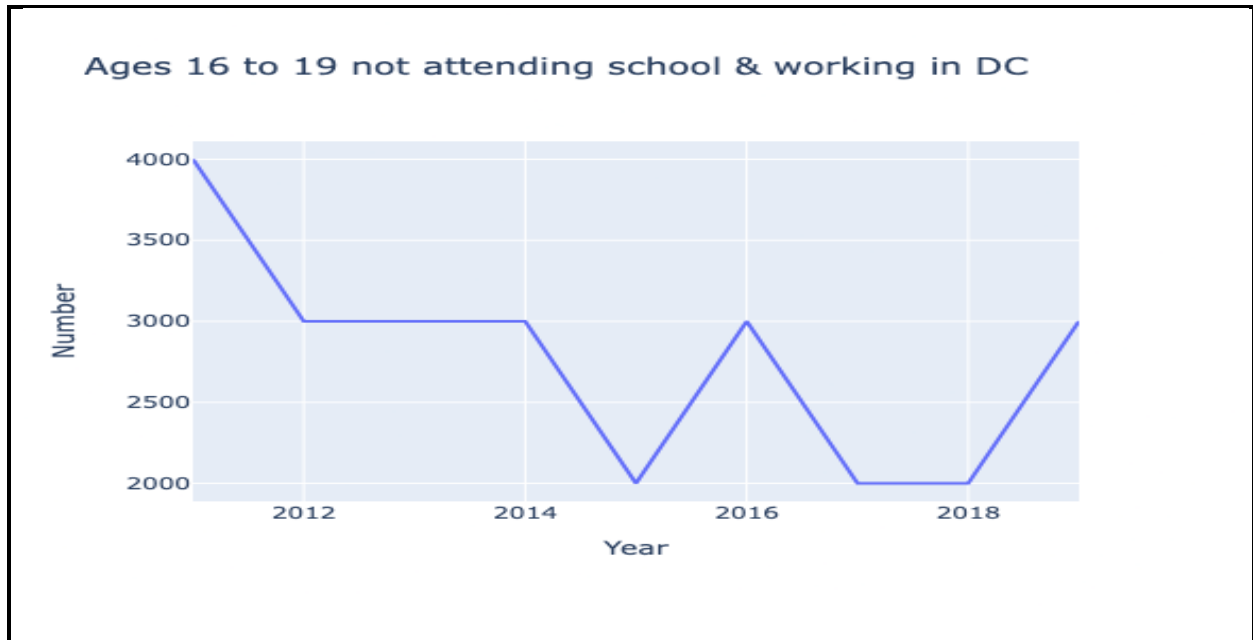


Fig 25. Teens aged 16 to 19 not attending school and not working.

Fig 25 shows teenagers between ages 16 and 19 who are not enrolled in school (full or part-time) and not employed (full or part-time). This measure is sometimes referred to as “Idle Teens” or “Disconnected Youth.” It’s worth noting that there may be various reasons why individuals in this group are not attending school or working, such as health issues, family responsibilities, or lack of job opportunities. It’s also possible that some individuals in this group may be pursuing alternative education or training opportunities outside of traditional school settings.

7. Transportation

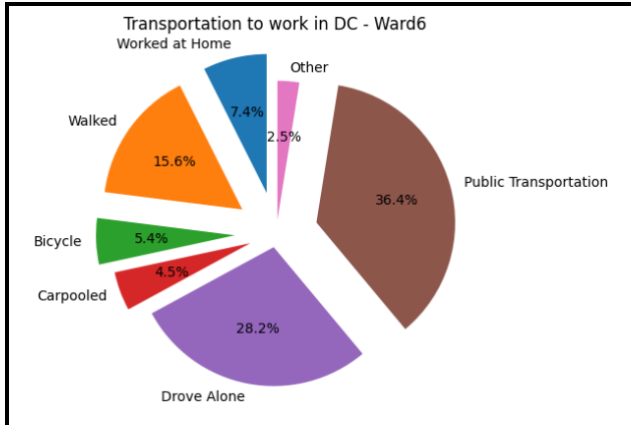


Fig 26. Transportation to work in DC - Ward 6

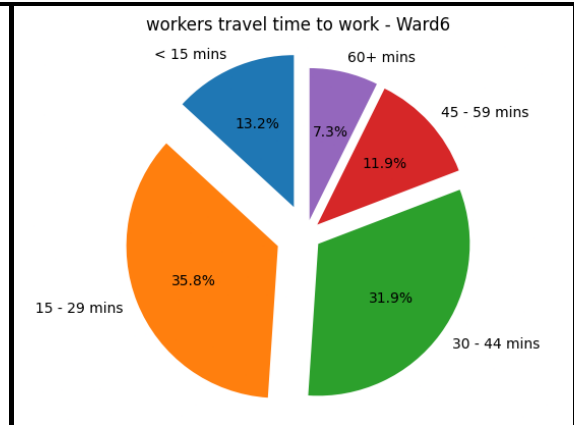


Fig 27. Workers travel time to work.

Fig 26 & Fig 27 shows the type of transportation methods and the travel time to work. We can see that the majority travel time was around 30 minutes which is slightly higher than the mean travel time in DC which was 27.1 minutes. In terms of mode of transportation, most workers traveled to work by public transportation followed by driving alone. Crime in public transport covers a wide variety of offenses committed in trams, buses, subways, commuter trains, and taxis and the crime targets can be the system itself, employees, and passengers.

5. Extension Analysis - Visualizations



As part of our extension analysis, we started to dive deeper into the crime data which we collected from Open Data and Metropolitan Police Department and tried to imprint it on map of Washington DC.

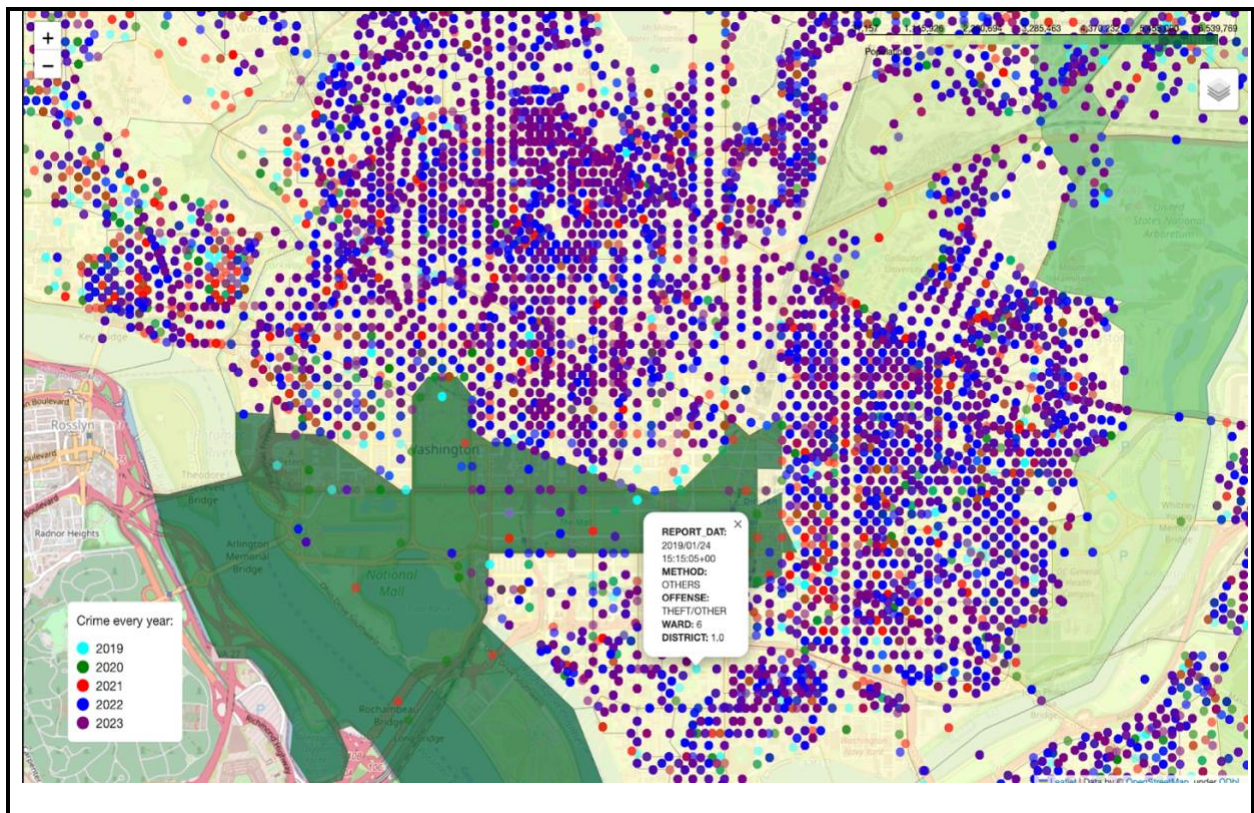


Fig 28. Crime in DC for all the year range 2019 to 2023

The above map which was created from folium and Open data showed us the exact locations of crimes in DC over the past 5 years. Each of the points show us more information about the crime

and the type of offense which is encountered. Different colors depict crime in different years. We can see that the crime is denser around Ward 6.

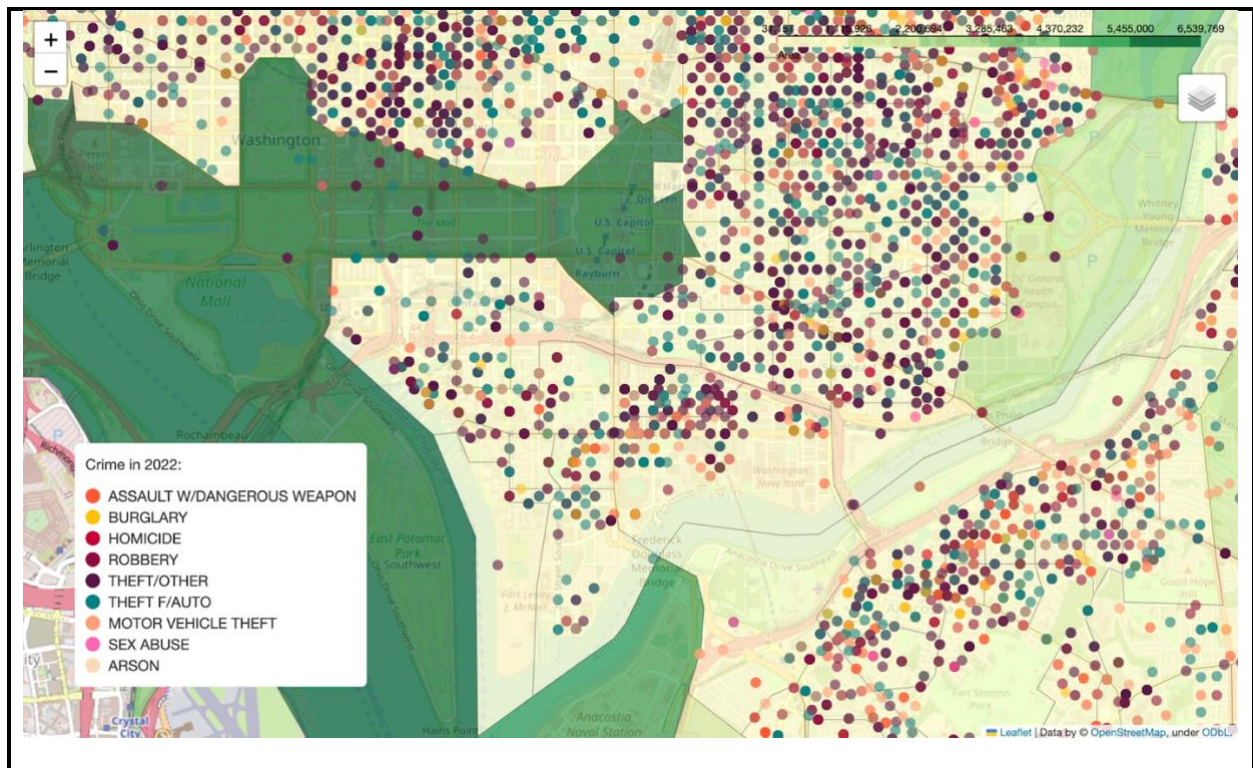


Fig 29. Crime in DC with respect to different Offense Categories

Figure 29 illustrates the different types of offense in the year 2022. The most common type of offense in Southwest as we can see is Theft. There have been a lot of such incident in this area which can be of some concern, and this should be investigated.

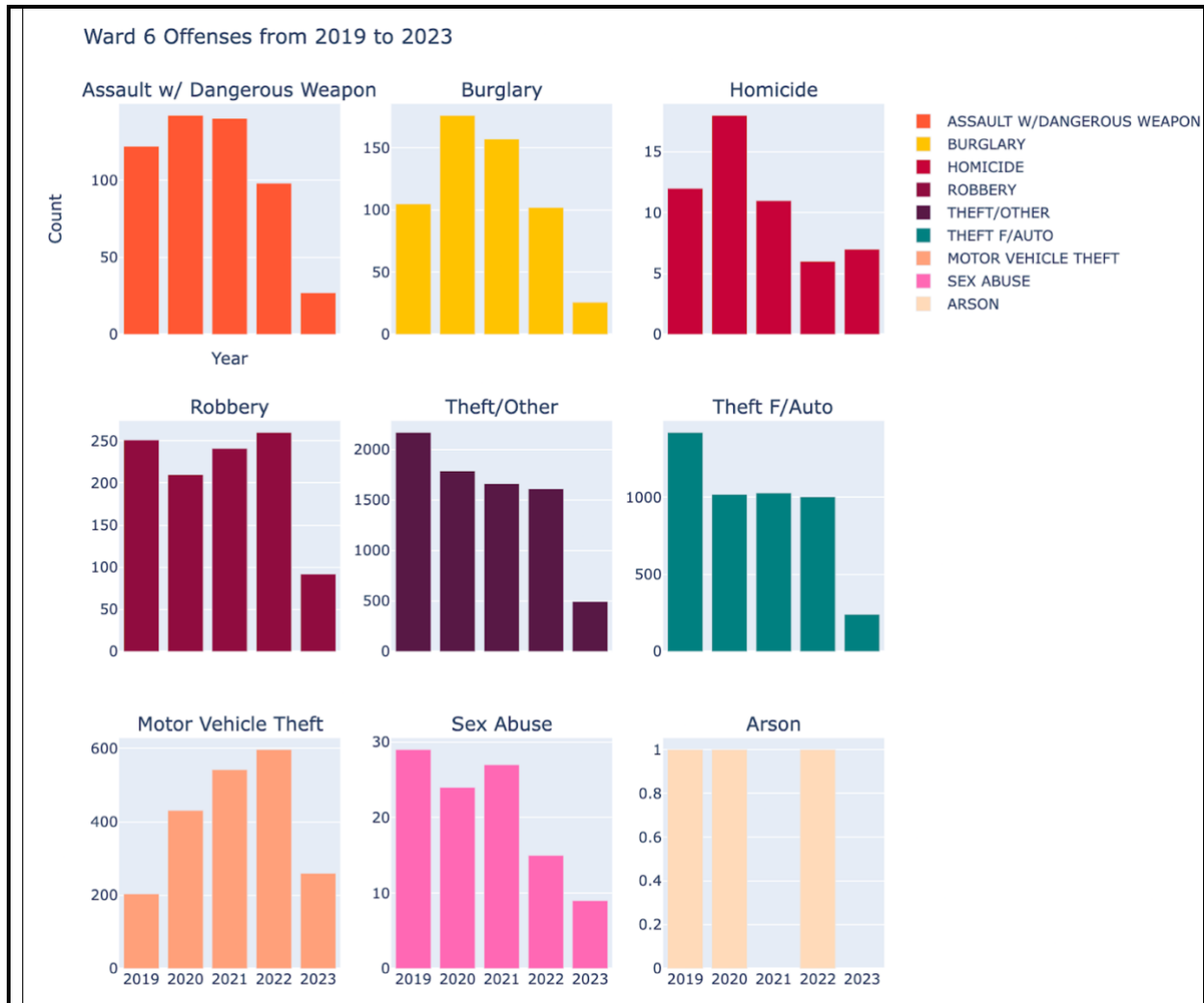


Fig 30. Crime in DC with respect to different Offense Categories

Figure 30 shows the distribution of all the different types of offense in different years. We can observe that over the past few years, the number of high-profile crimes has decreased, and it goes to say that there has been a good effort to tackle crime in the last few years.

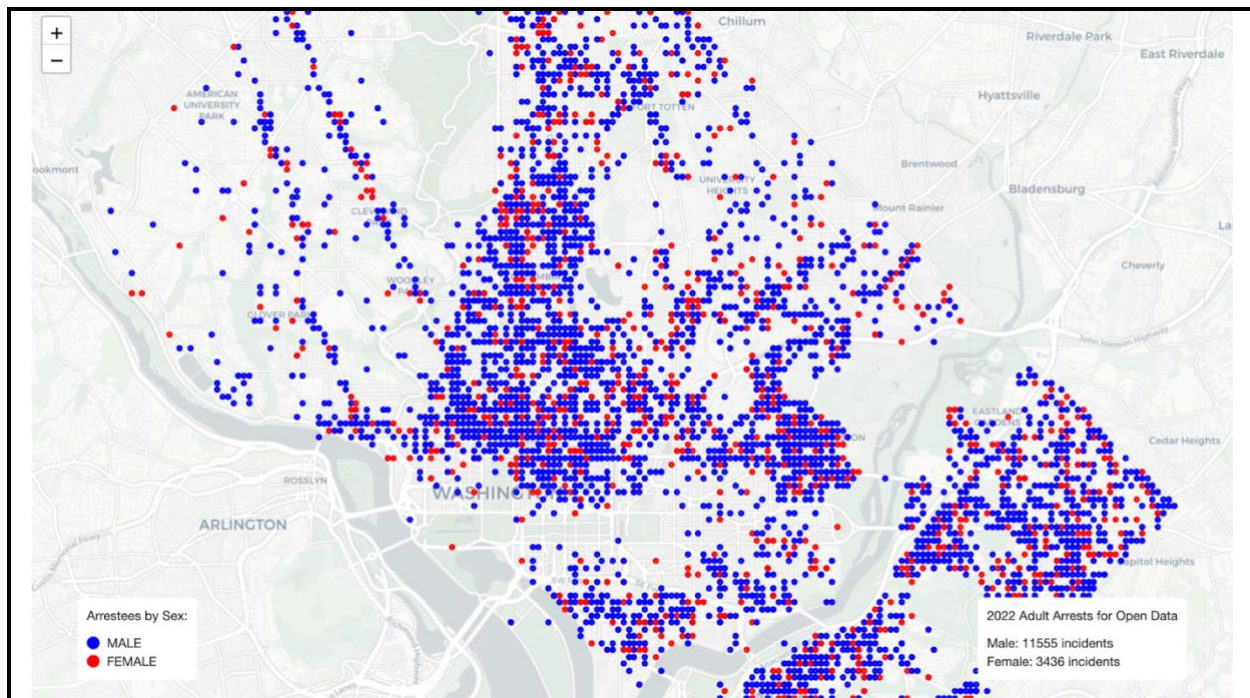


Fig 31. MPD Crime - Male vs Female

Additionally, we investigated the Metropolitan Police Department data which had more data on age and race along with other information. We investigated finding correlations between age and crime along with crime. When we saw the map, we found that although most of the crimes are being committed by men, there are a fair bit of incidents which are being committed by female.

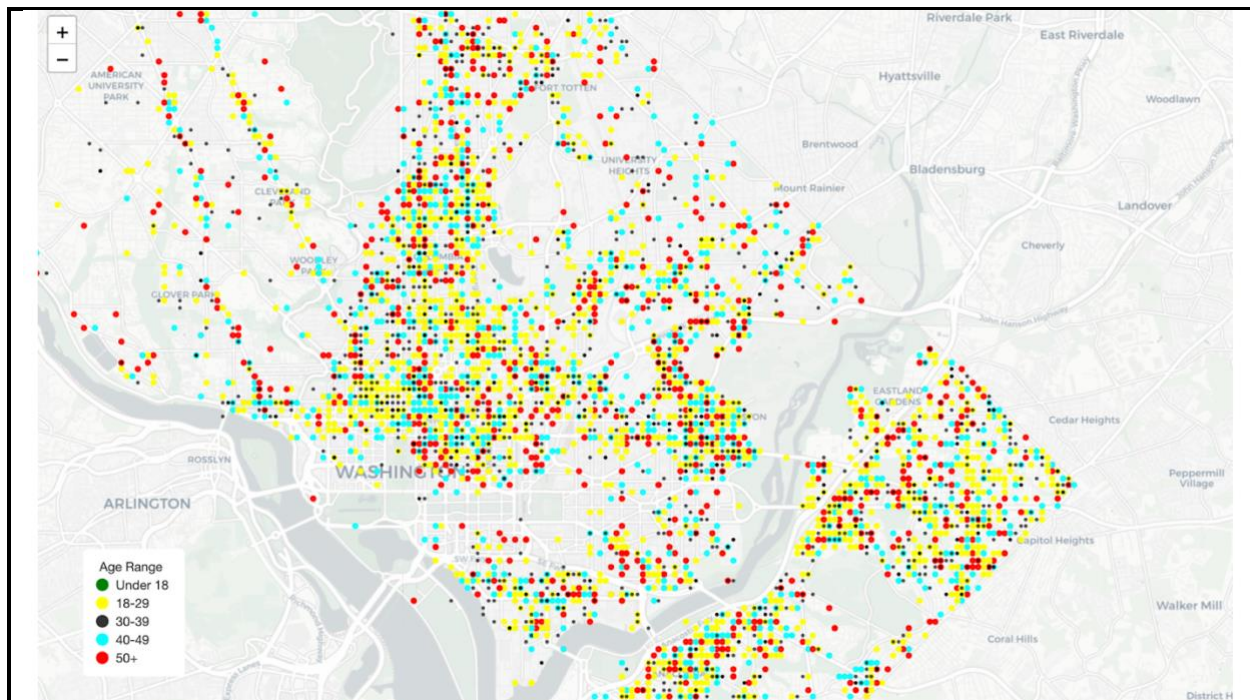


Fig 32. MPD Crime - Crime based on Age for the year 2022.

For further analysis, we looked to find the relations between Age and Crime. As we had discussed earlier in our older analysis and previous graphs, the age range which is most affected are the young adults. We also had seen that the age range had most of the people who were below the poverty level. The same is the case with crime as well, there seems to be young adults who have committed crime.

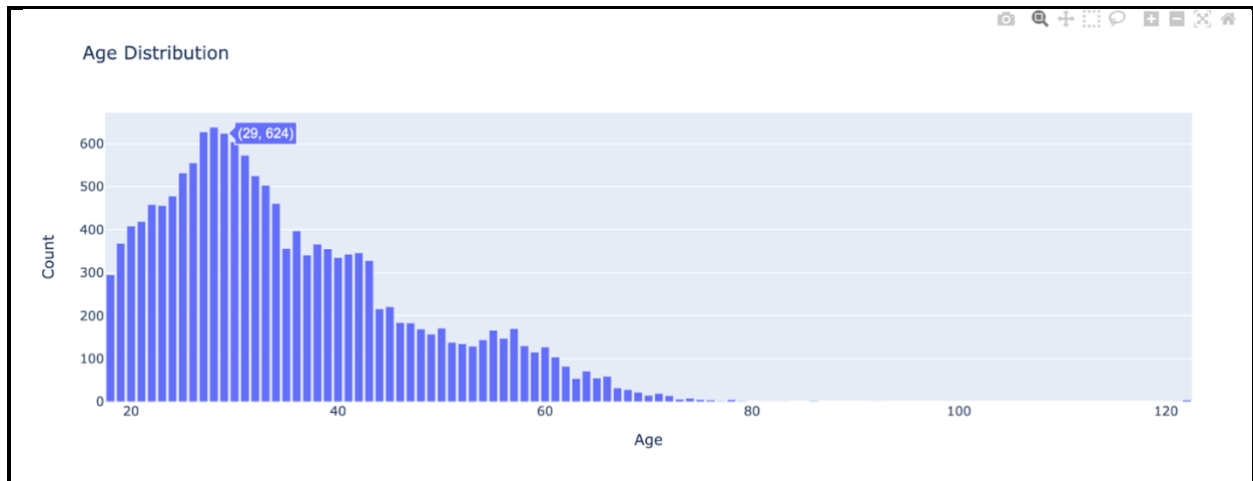


Fig 33. MPD Crime - Crime Age distribution for the year 2022

Figure 33 and 34 shows the age distribution according to crime in the year of 2022. As we analyzed earlier, we can see that the age of 29 is under the radar. There are 624 such incidents where the perpetrator is at the age of 29. The Below graph shows, Race distribution for the crime.

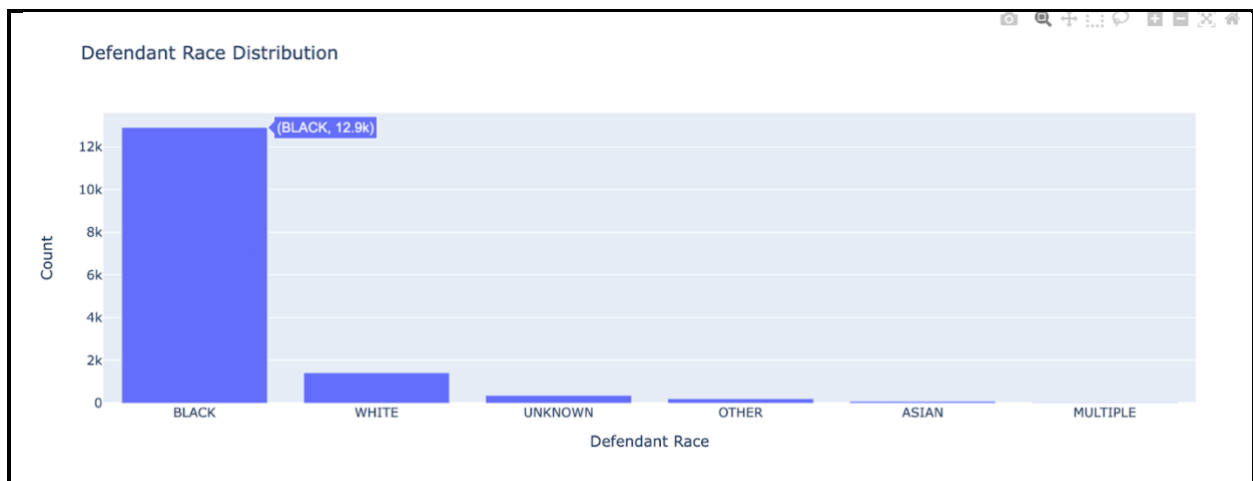


Fig 34. MPD Crime - Crime Race distribution for the year 2022

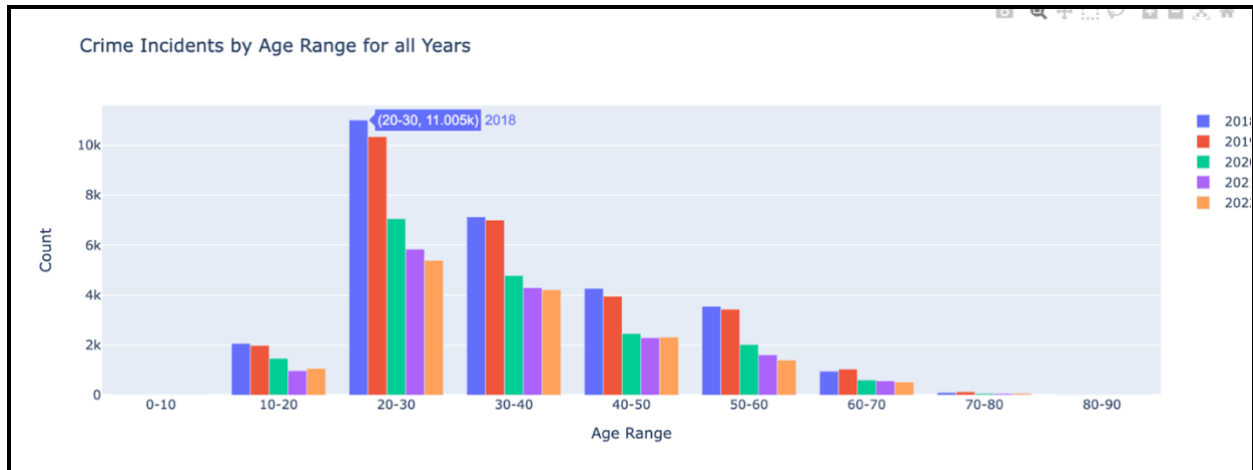


Fig 35. MPD Crime - Crime Incidents by Age Ranges for all years

Figure 35 clearly shows that there are many crime incidents in the age range of 20 to 40. As we had a similar graph previously, we can see that the crime overall and within the age range 20 to 30 has shown significant improvement. All the effort which has been invested is good and it is paying off.

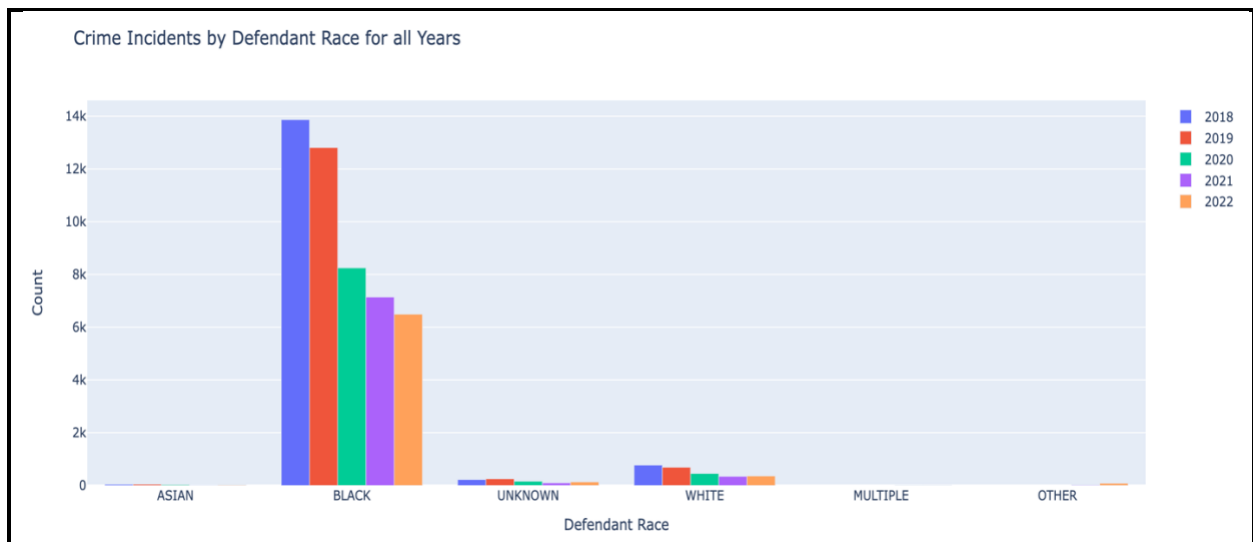


Fig 36. MPD Crime - Crime Incidents by Age Ranges for all years

Figure 36 is plotted to see the Defendant Race Categories and to investigate which race is involved more in Crime. As we see in the above the graphs for all the years from 2018 – 2022 people from Black race were seen to be involved more in various criminal activities. In the year 2018 we can see that there was higher number of crimes by this race that is 13.8K. Over the years even though there is decrease in crime rate, which is in 2022 it is 6.4k but majority of it is committed by black people.

5. Team Members Contribution

- **Harshitha Tumkur Kailasa Murthy -**

- As a Team Representative attended all Client meetings to understand the requirement better and plan how my team and I can work together.
- Exploratory data analysis for demographics data: Age, Income, and population.
- Analyzed data based on the above parameters to get plots for different tracts as per the client's feedback and inputs.
- Scheduled internal Team 1 meetings to discuss each member's update and plan and help each other.
- Prepared scrum reports and delegated by giving turns to each member to do scrum reports.
- Prepared the skeleton for all decks and reports submitted.
- As part of the correlation analysis, I tried to investigate deeper data regarding age and income. Worked towards finding new relationships which can help us understand how we can tackle crime in these regions.
- Worked on the graphs for this draft and documentation of the Final draft report.

- **Vishwas Bhaktavatsala -**

- Once the presentation was completed, in the latest meeting we communicated with what we would be requiring getting better visualizations.
- Started to look for more granular data regarding crime in Southwest DC. Looking into creating correlations between crime and other aspects which can help us understand the situation of Southwest DC better.
- Worked on mapboxes to visualize the crime statistics better and started to investigate various crime datasets which are spread over the past few years.
- looked into creating a specific map for Washington DC as that would be better to understand and visualize.
- Added the documentation graphs and certain explanations which can help understand our analysis better.

- **Kris Chou -**

- Used OpenData on affordable housing data and crime data. Found correlation between affordable housing and crime, heatmap of crime, map showing affordable housing and crime.
- Looked into alternative ways in showing and visualizing data.
- Created visualizations through a map tying in affordable housing and crime.
- Focused on Gun Violence data while trying to find locations of gun violence and whether they correlate to affordable housing.

- **Yeh-Min Lu -**

- Education: Educational Attainment,
- Economic well-being: Labor force, employment, and unemployment rate, Teens age 16-19 not going to school and not working, Child neglected and abused.
- Transportation: transportation methods and the travel time to work.

6. Conclusion & Future Work

Our analysis shows us that restructuring of Southwest Washington DC had its problems which is being clear. Tracts 64 and 105 have a very high percentile of people who are below the poverty level and most of them are young adults. Crime being our focus in the latter half of our analysis showed us various insights on the different types of crime as well as all the correlations of crime with other factors. We saw that, over the past 5 years the number of crimes has decreased, but there were a few crime categories which seem to increase by a big margin over the years. The correlations showed us that most of the crimes were committed by male, and they are mostly by the African American community. Communities which have affordable housing have a few incidents of crime concentrated.

Lastly, as part of the future work, more analysis on crime and the different types of categories can provide insights on problems in the area. Further correlations on various aspect can help understand the different ways to improve and support the displaced communities here in Southwest Washington DC.

7. Data Sources

OpenData : <https://opendata.dc.gov/>

Metropolitan DC Data: <https://mpdc.dc.gov/publication/mpd-adult-arrests-2013-2021>

Census US: <https://data.census.gov/>