Subedi Final Project

Amisha

2023-11-16

Introduction

In our team's comprehensive exploration of crime and transportation dynamics, my specific focus centers on Washington D.C., a city known for its intricate urban challenges. As we collectively delve into the nuanced interplay between crime incidents and transportation infrastructure in the United States, this individual inquiry homes in on the nation's capital. The primary objective is to unravel distinct patterns or correlations within the local landscape, with a particular emphasis on examining how crime rates relate to the presence of metro bus stops. By incorporating demographic factors and counties characteristics, this investigation aims to shed light on the multifaceted dynamics of crime and transportation in Washington D.C.

Question 1

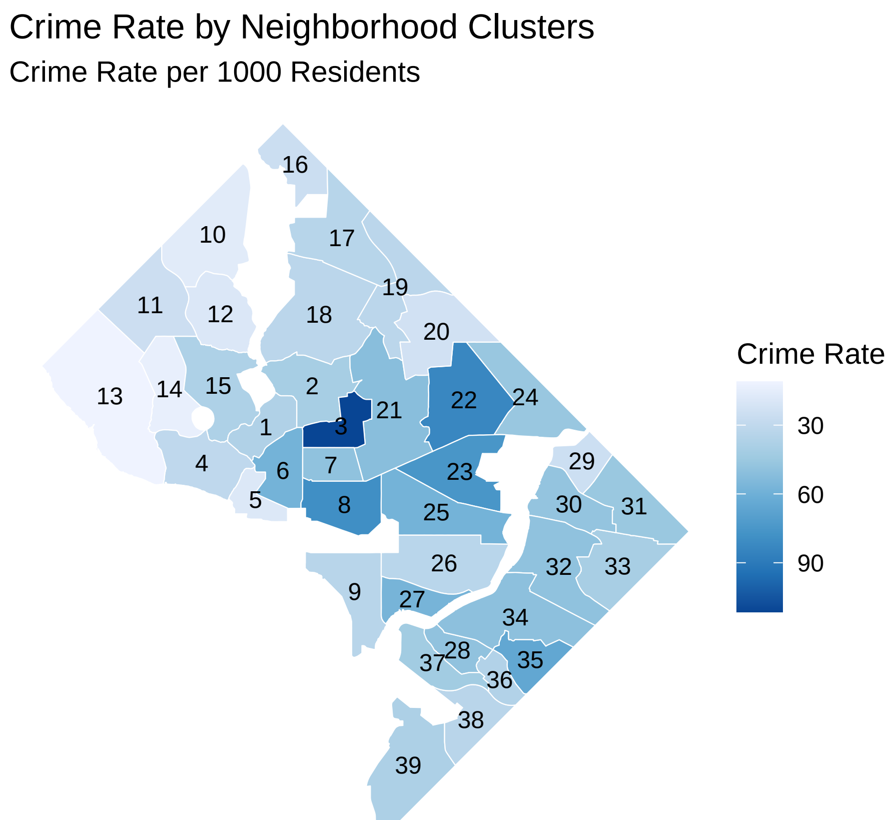
How are crime incidents and metro stations presence distributed across Washington D.C.'s neighborhood clusters, and what insights can be drawn from their spatial relationship?

Before we delve into the spatial analysis of crime incidents across Washington D.C.'s counties, it is pertinent to understand the composition of the Washington DC. DC is divided into smaller areas called "clusters" - compromising two or more areas called "neighborhoods", which I will interchangeably use as "counties" in this report. These clusters are groupings of counties that share similar characteristics and are crucial for our analysis as they form the basis of our geographic segmentation. Below is a table listing a random selection of these clusters alongside the corresponding counties they comprise. This will provide a reference point as we proceed to the choropleth map visualization, which illustrates the distribution of crime incidents across these clusters.

Table 1: Random Selection of 10 Clusters and Their Corresponding Counties

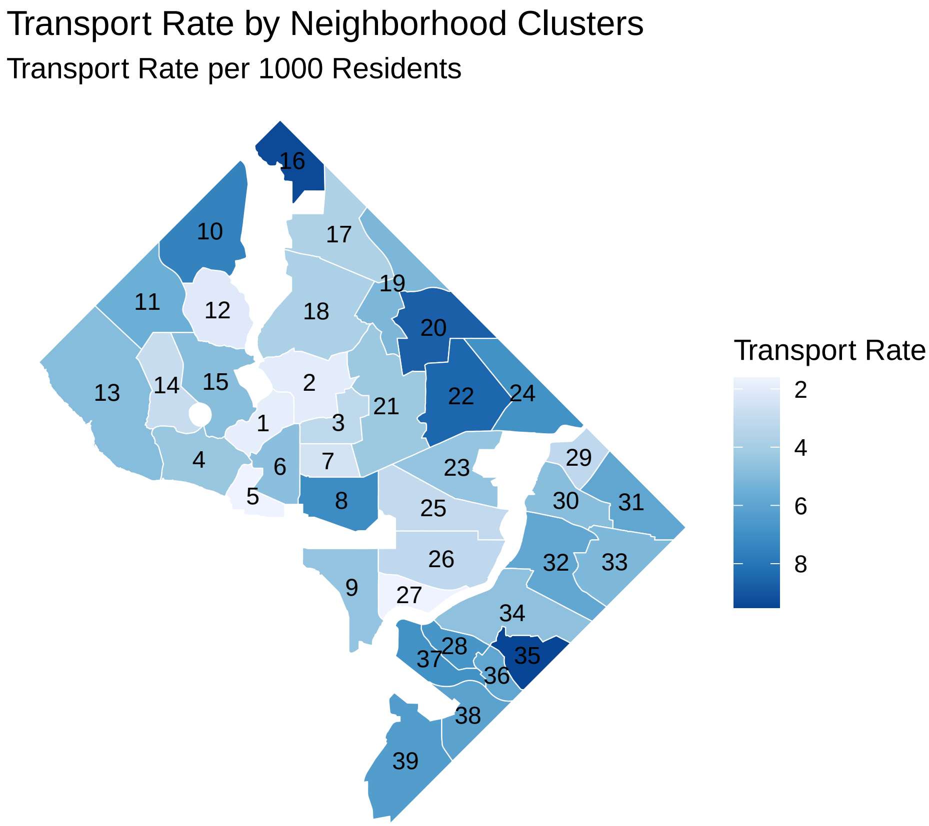
Cluster_Id	counties
cluster 11	Friendship Heights, American University Park, Tenleytown
cluster 16	Colonial Village, Shepherd Park, North Portal Estates
cluster 23	Ivy City, Arboretum, Trinidad, Carver Langston
cluster 29	Eastland Gardens, Kenilworth
cluster 32	River Terrace, Benning, Greenway, Dupont Park
cluster 33	Capitol View, Marshall Heights, Benning Heights
cluster 37	Sheridan, Barry Farm, Buena Vista
cluster 4	Georgetown, Burleith/Hillandale

Now, I will begin my analysis by generating a detailed choropleth map to visualize the distribution of crime incidents across the various neighborhood clusters in Washington D.C. I initiate the inquiry with a choropleth map to visualize crime distribution across Washington D.C.'s 46 neighborhood clusters—simplified groupings from the original 131 counties. This initial step is crucial in understanding the landscape of urban crime before delving into the interplay with metro station presence. To ensure accuracy and relevance, my focus is specifically on types of crime incidents related to urban settings. Using a comprehensive dataset that includes crime locations and types, I was able to categorize and analyze these incidents effectively. This analysis is not just about pinpointing hotspots but understanding the broader patterns of crime distribution in relation to the city's fabric. This initial analysis seeks to highlight areas with higher crime rates and to set the stage for a deeper exploration into urban dynamics. The choropleth map below serves as a visualization tool to delineate the crime frequency by neighborhood, revealing patterns that may not be immediately apparent through raw data alone.

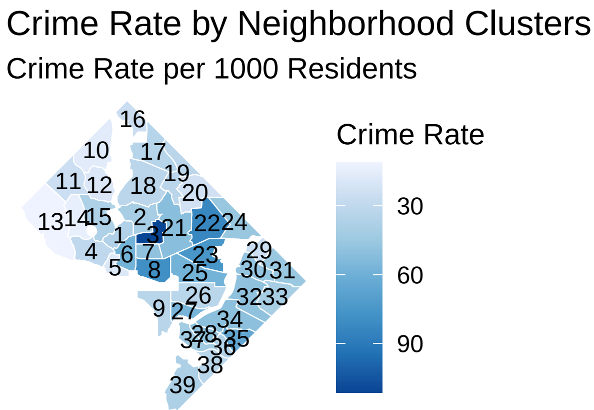


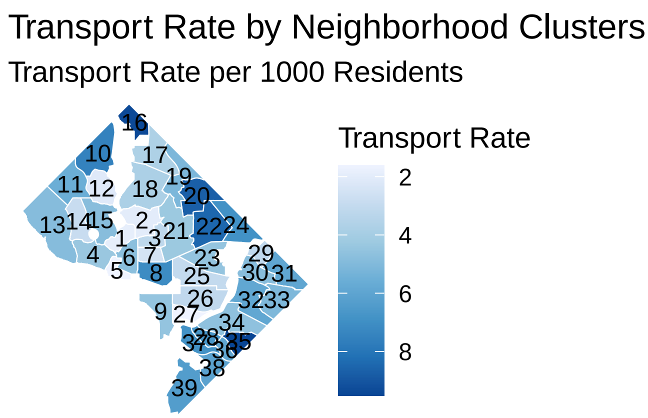
The choropleth map above illustrates the distribution of crime rates across Washington D.C.'s neighborhoods. Notably, central neighborhoods like Cluster 3 (Howard University, Le Droit Park), Cluster 22 (Brookland, Brentwood, Langdon) exhibit elevated crime rates, suggesting a potential link between urban density and crime incidence. This pattern underscores the need for targeted law enforcement and community resources in these areas.

Following the crime rate examination, I turn my attention to the presence of metro stations. By plotting a second choropleth map that represents the number of metro bus stops per 1,000 residents (which I'll refer by "Transport Rate" in this report) by neighborhood, we can overlay our understanding of crime distribution with public transportation accessibility. This comparison allows us to identify any apparent relationships between crime incidence and the density of metro services, thus providing insights into how transit infrastructure might influence neighborhood safety.



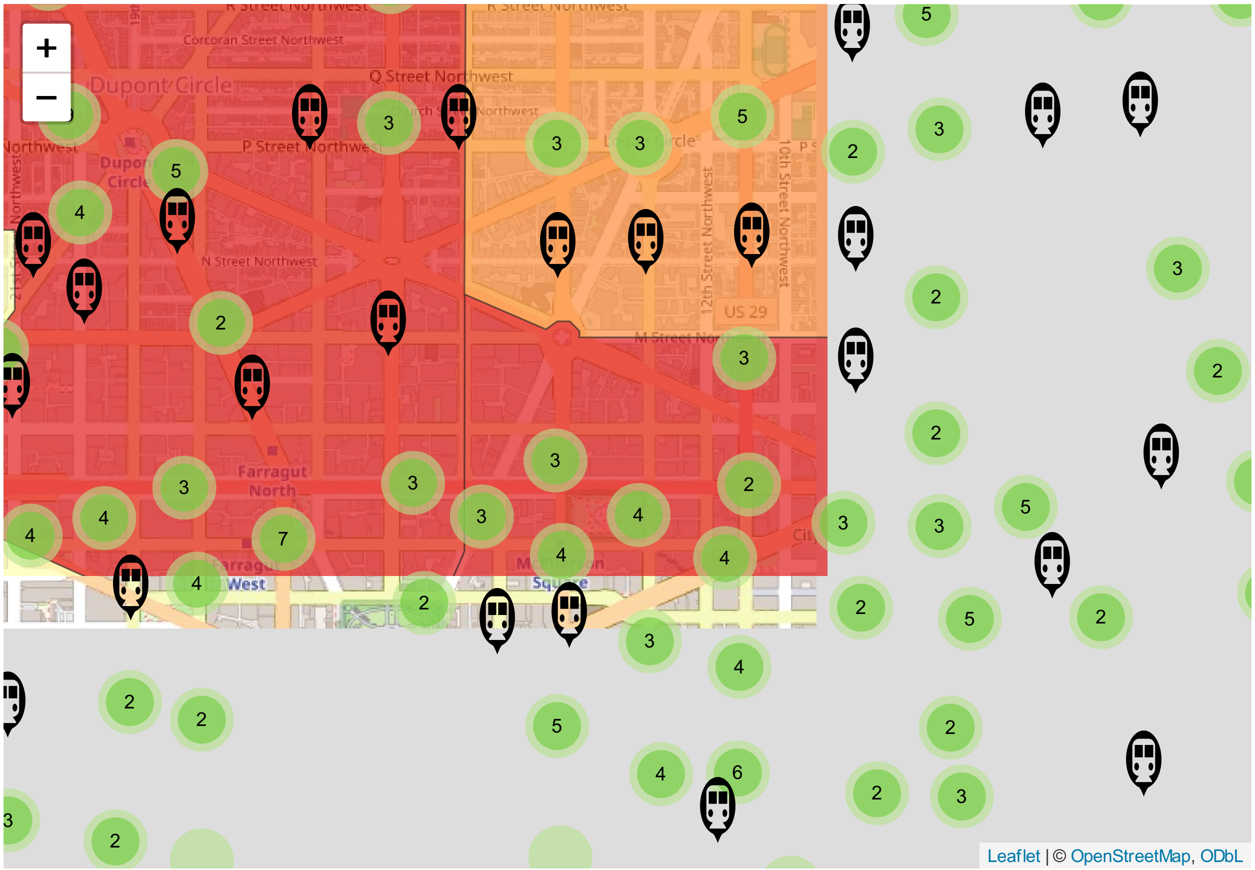
Upon analysis of the second map, the distribution is more darker on cluster 16, 35, 20, 22, and 8. To synthesize our findings, we utilize the grid arrange function to present a side-by-side comparison of crime and metro station distribution. This combined visual representation facilitates a comprehensive analysis of the potential interplay between these two critical urban factors. It provides a clearer narrative of how transit infrastructure may correlate with safety and security across neighborhoods.





The combined visualization raises intriguing considerations; while some areas (Clusters 22, 8, 20) with high metro rates align with increased crime rates, others (cluster 16, 13, 10) do not follow this pattern, suggesting the influence of additional, unexplored variables. Given the complexity of discerning clear patterns from static choropleth maps, I have employed an interactive Leaflet map for a more nuanced exploration of the spatial relationship between crime rates and transport rates in Washington D.C.'s neighborhoods.

In this interactive map, you can zoom out to view polygons representing crime rates in specific neighborhood clusters. Additionally, metro station locations are marked on the map for reference. When zoomed out, you may notice that markers are clustered together for a more organized presentation. This clustering allows for a clearer overview of the entire city, and as you zoom in, markers expand to provide finer details. The interactive map provides a dynamic and user-friendly way to explore the data, allowing for a more in-depth analysis of the urban dynamics in relation to crime and transportation accessibility.



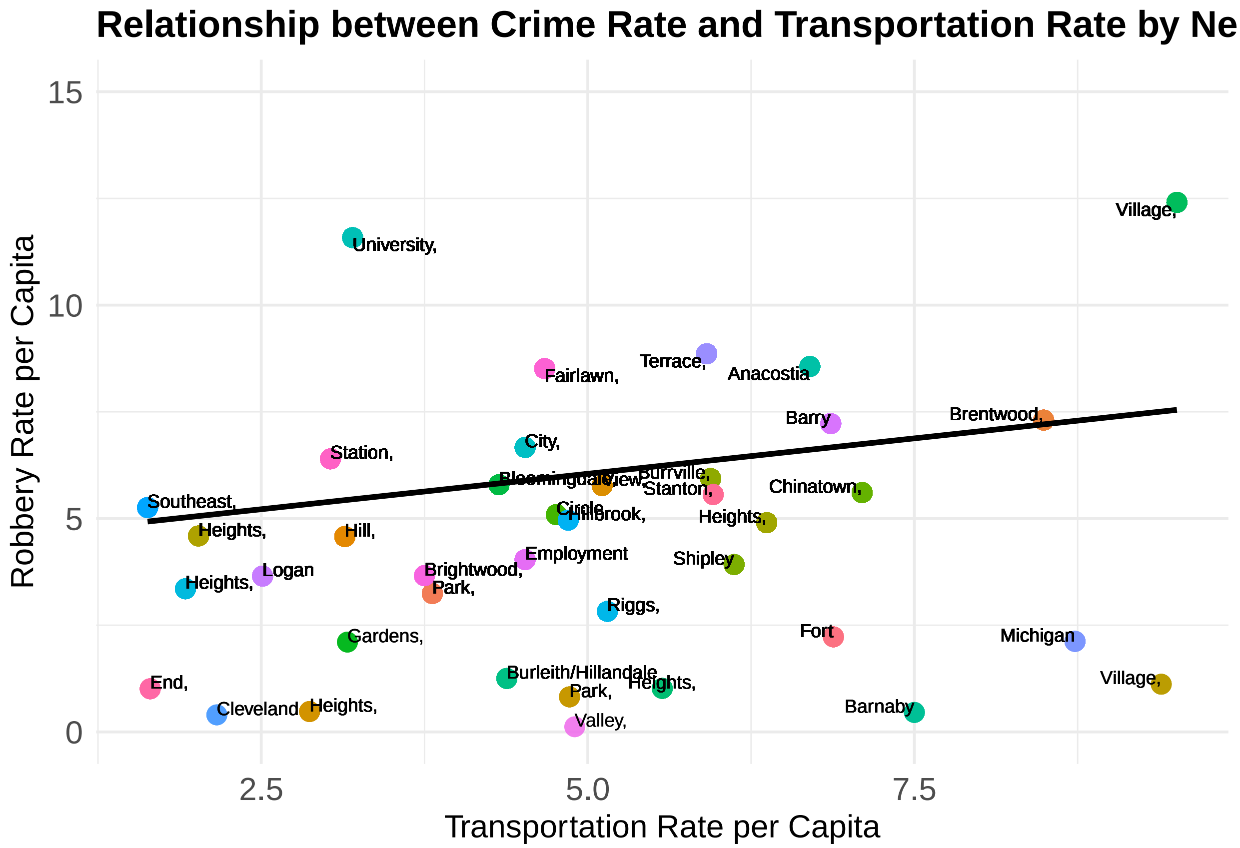
From the Leaflet visualization, it is evident that such as Cluster 25 (Union Station, Stanton Parkm and Kingman Park), Cluster 23 (Ivy City, Arboretum, Trinidad, Carver Lang), Cluster 8 (Downtown, Chinatown, Penn Quarters) which encompasses parts of Northeast and Southeast D.C., display a higher concentration of metro bus stops along with the most intense red shading, indicating the highest crime rates. In contrast, Cluster 9 - Southwest Waterfront, which includes the National Mall and surrounding areas, shows moderate transportation clustering with a notably lower crime rate, as suggested by the lighter coloration. Furthermore, Cluster 38 (Douglas, Shipley Terrace),despite having a dense transit network, do not exhibit the darkest shades, suggesting that factors other than transportation availability might play a role in influencing crime rates. This interactive mapping approach reveals the spatial nuances and complexities within Washington D.C., challenging simple correlations and highlighting the multifaceted nature of urban crime dynamics

Question 2

What patterns emerge for robberies and other offenses around transportation hubs, and which crimes are most strongly correlated with these locations?

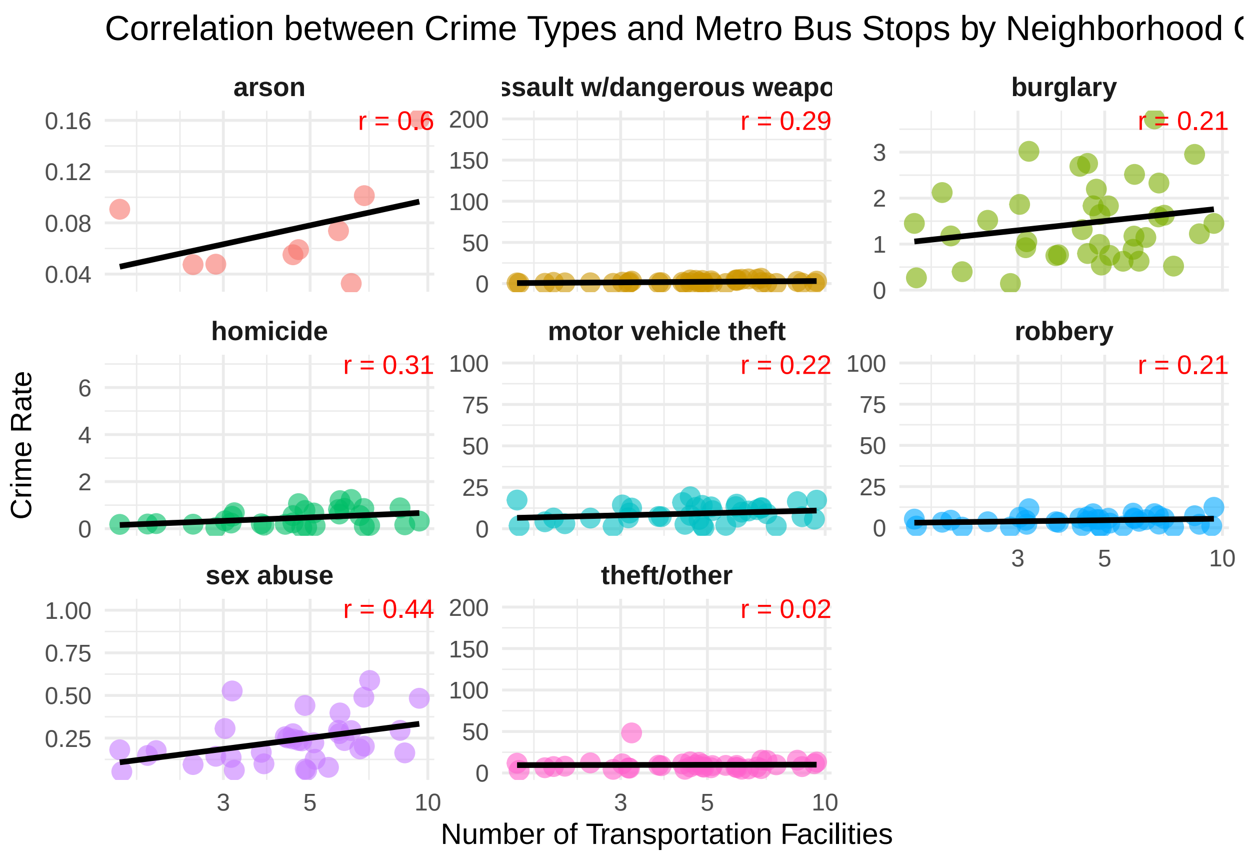
I was interested in discovering whether neighborhoods with extensive transportation facilities experienced a disproportionately higher rate of robberies compared to those with fewer transit options. The key term here is 'disproportionately' because my objective was not to simply identify if robberies occur but to ascertain if they occur at a rate that outpaces what might be expected given the neighborhood population. Such an understanding is vital; without it, one might mistakenly attribute a city-wide rise in robbery rates solely to increased transportation options, overlooking other socio-economic factors that influence crime.

To this end, I analyzed robbery rates on a per capita basis, juxtaposing them with transportation rates within the same neighborhoods. This per capita approach ensures a consistent comparison, mitigating the risk of skewed perceptions that could arise from raw data comparisons.



In examining the scatter plot, my hypothesis was that neighborhood clusters with higher per capita transportation rates would exhibit increased robbery rates. However, the visualization unveiled a nuanced reality. The spread of data points across the scatterplot, along with the best-fit line, indicated a moderate positive correlation overall. The correlation coefficient was approximately 0.25. Yet, this wasn't universal; some neighborhoods with dense transportation networks did not see a corresponding spike in robbery rates, hinting at a more complex interplay of factors at work.

Building upon the initial analysis of robbery rates, I extended my investigation to encompass a wider array of crimes. My interest was to discern whether other offenses—like assault, theft, and motor vehicle theft—exhibited a similar or divergent pattern when considering their proximity to transportation hubs.



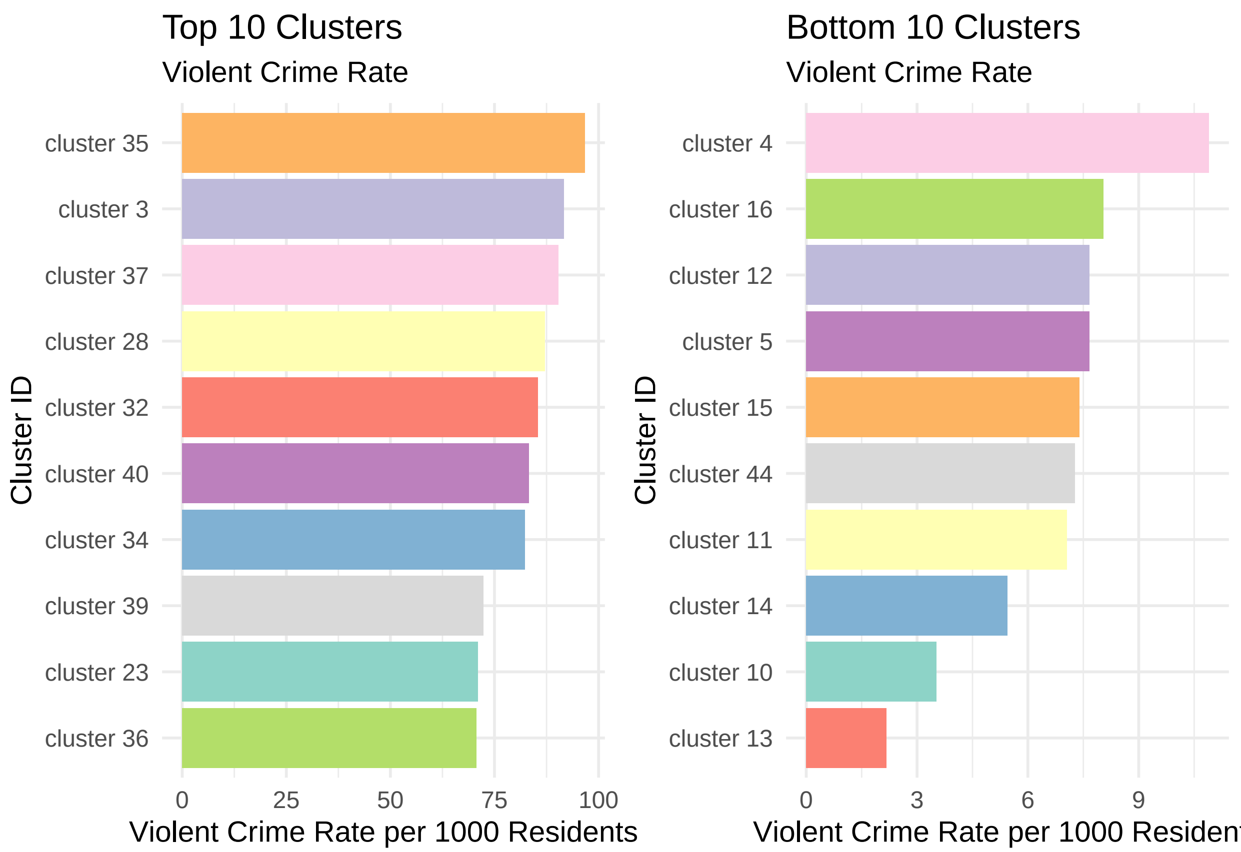
The analysis of crime types in relation to metro bus stop density in Washington D.C. neighborhoods reveals varying correlations. Arson shows a strong positive link, suggesting increased fire risk in areas with more transit facilities. Assault and burglary exhibit moderate associations, while homicide and motor vehicle theft are also positively correlated. Sex abuse demonstrates a significant connection, emphasizing the need for enhanced safety measures. However, theft and other offenses have a weaker relationship with metro bus stops. These findings highlight the nuanced impact of transportation accessibility on different crime types, calling for tailored approaches to urban safety planning.

Question 3

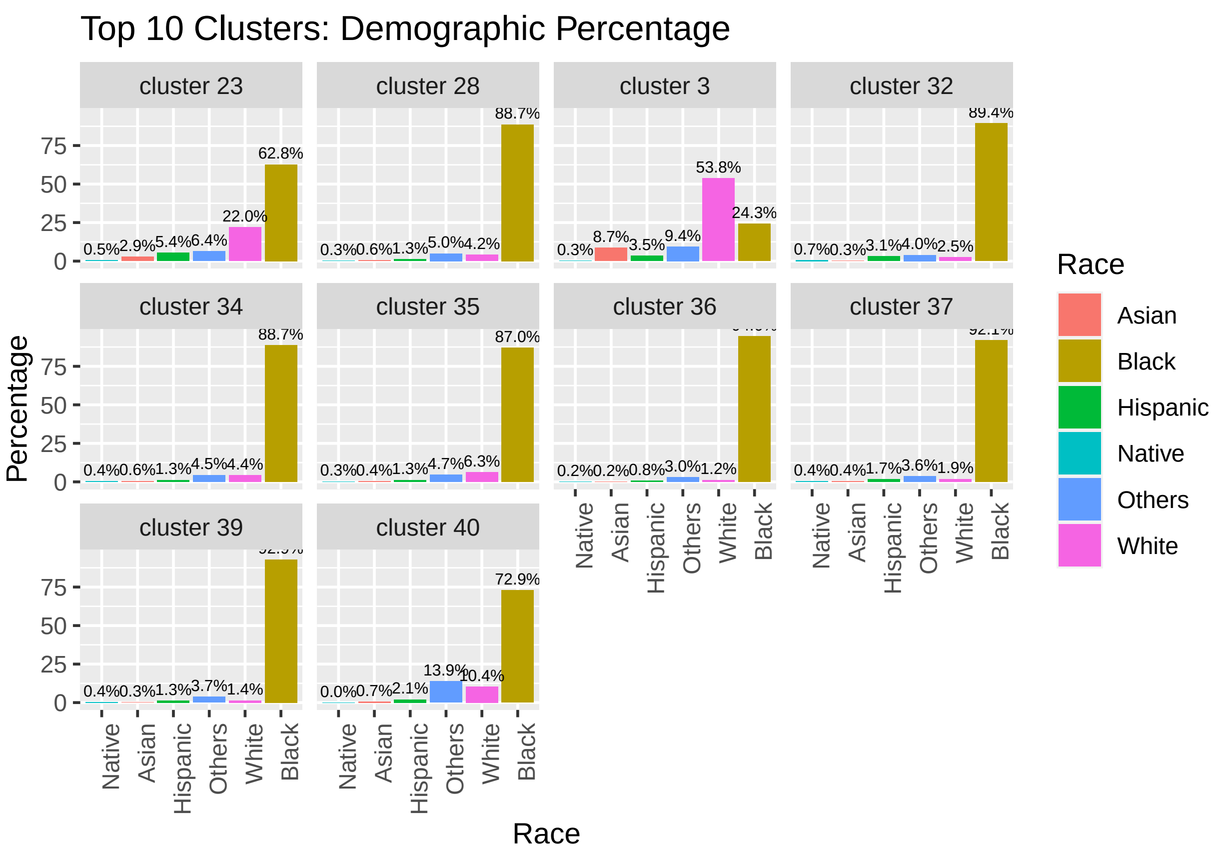
What are the observed patterns in racial demographics related to crime types across top 10 neighborhoods with highest violent and property crime, and What might be the underlying reasons for these patterns?

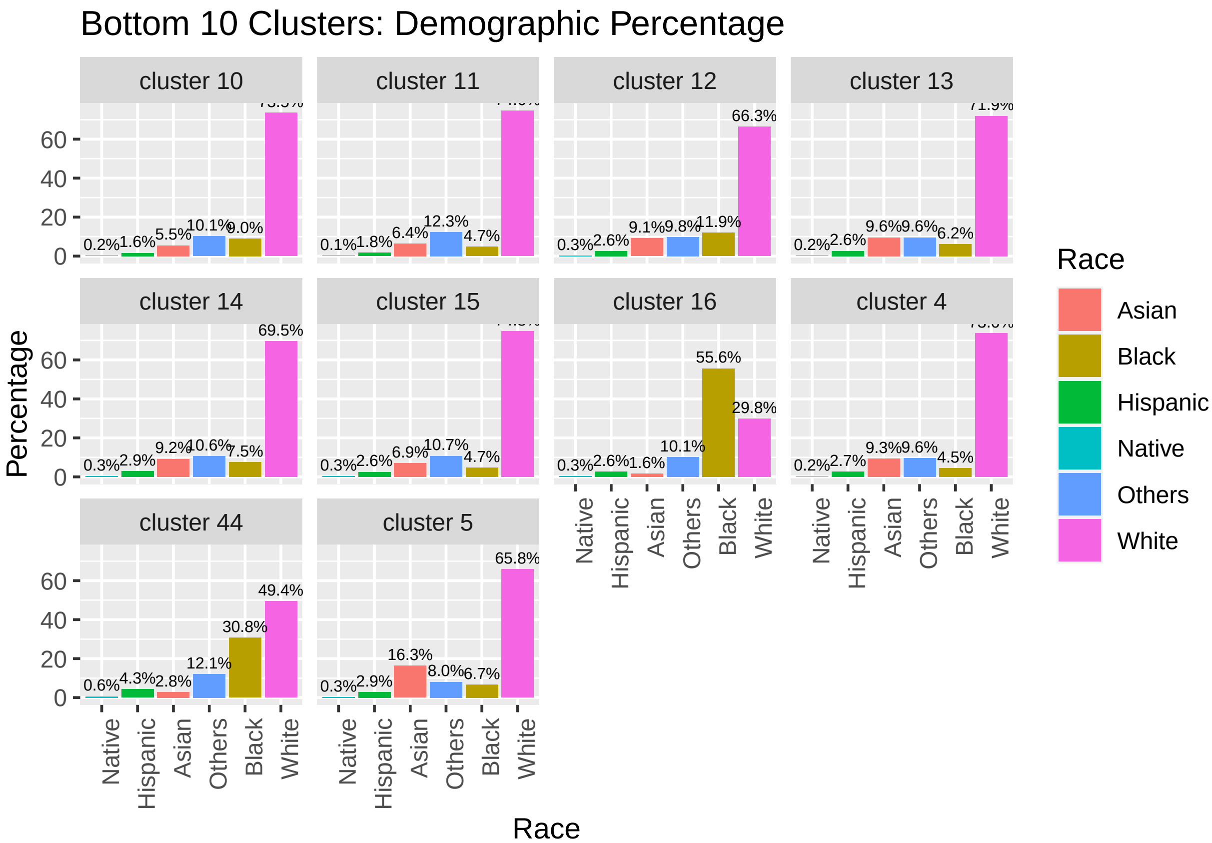
In our pursuit to understand the social underpinnings of crime within Washington D.C.'s varied landscape, Question 3 led us to an investigative analysis of the demographic composition within the neighborhood clusters most affected by crime. I was particularly interested in discerning whether the demographic profiles of these areas show any patterns. This curiosity was rooted in the hypothesis that certain socio-demographic characteristics might predispose neighborhoods to higher incidences of violent and property crimes. The subsequent findings reveal not just cold numbers but stories of communities, each with its unique socio-economic tapestry that potentially influences its crime narrative. To bring these stories to light, the ensuing sections will present tables and scatter plots that map out the demographic breakdowns against crime rates. These visual narratives are not merely illustrative but are integral to sculpting targeted, data-informed strategies for urban policy and community revitalization.

Before delving into the analysis of the top 10 and bottom 10 clusters for violent crime rates in Washington DC, it is important to note that certain clusters, specifically clusters 42, 45, and 46, 43 are not included in this and similar analyses. Cluster 42, consisting mainly of the U.S. Naval Observatory and Embassy Row, Cluster 45, encompassing the National Mall with its museums and cultural institutions, and Cluster 46, home to the U.S. National Arboretum, are all primarily non-residential areas. These characteristics render them less relevant for residential population-focused analyses such as crime rate studies.



The bar charts depict the top and bottom 10 clusters in Washington D.C. by violent crime rate per 1000 residents, with clusters 43 and 4 experiencing the highest and lowest rates respectively. To further understand the underlying factors, analyzing the demographic composition of the top 10 and bottom 10 clusters could provide valuable insights into the correlation between crime rates and demographic variables.


Top 10 Clusters: Demographic Percentage
cluster 23
0.5% 2.9% 5.4% 6.4% 22.0% 62.8%
cluster 28
0.3% 0.6% 1.3% 5.0% 4.2% 88.7%
cluster 3
0.3% 8.7% 3.5% 9.4% 53.8% 24.3%
cluster 32
0.7% 0.3% 3.1% 4.0% 2.5% 89.4%
cluster 34
0.4% 0.6% 1.3% 4.5% 4.4% 88.7%
cluster 35
0.3% 0.4% 1.3% 4.7% 6.3% 87.0%
cluster 36
0.2% 0.2% 0.8% 3.0% 1.2%
cluster 37
0.4% 0.4% 1.7% 3.6% 1.9% 92.1%
cluster 39
0.4% 0.3% 1.3% 3.7% 1.4%
cluster 40
0.0% 0.7% 2.1% 13.9% 10.4% 72.9%
Percentage
0
25
50
75
Native
Asian
Hispanic
Others
White
Black
Race
Race
Asian
Black
Hispanic
Native
Others
White



Analyzing the demographic composition in the top 10 clusters with the highest violent crime rates reveals disparities. Notably, the Black demographic is more prevalent in all clusters, while cluster 3 shows a significant White demographic presence. In contrast, the bottom 10 clusters have a distinct demographic profile, with the White demographic being more prominent all clusters, except for cluster 16. These observations suggest potential correlations between demographic variables and crime rates, which warrant further investigation.

Now, I'm curious to explore if similar demographic patterns exist for property crime rates. Let's delve into the demographic composition of property crime to understand if specific demographics have consistent impacts on both violent and property crime in different clusters.

The property crime rate bar charts for Washington D.C. reveal a stark disparity, with cluster 40 standing out for its notably high rate. This contrast to the much lower rates in the bottom clusters is striking and suggests a complex interplay of factors at play. It raises the question of whether demographic patterns observed in violent crime rates similarly influence property crime or if these rates are shaped by different dynamics altogether. So, I want to study the demographics of the top 10 and bottom 10 neighborhood clusters to see if their demographics follow any pattern.

I have utilized kableExtra to list the demographic details of top 10 and bottom 10 Neighborhoods below:

Table 2: Demographic Breakdown of Property Crime Perpetrators in top 10 neighborhood

Cluster_Id	counties	Asian	Black	Hispanic	Native	Others	White
cluster 22	Brookland, Brentwood, Langdon	2.18	61.59	4.82	0.52	8.10	22.78
cluster 23	Ivy City, Arboretum, Trinidad, Carver Langston	2.87	62.77	5.39	0.55	6.40	22.03
cluster 25	Union Station, Stanton Park, Kingman Park	4.95	27.60	2.06	0.27	8.45	56.67
cluster 27	Near Southeast, Navy Yard	5.24	17.06	1.79	0.33	8.09	67.50

cluster 3	Howard University, Le Droit Park, Cardozo/Shaw	8.69	24.30	3.49	0.31	9.44	53.77
cluster 35	Fairfax Village, Naylor Gardens, Hillcrest, Summit Park	0.40	87.04	1.29	0.26	4.69	6.32
cluster 40	Walter Reed	0.69	72.92	2.08	0.00	13.89	10.42
cluster 6	Dupont Circle, Connecticut Avenue/K Street	9.85	6.92	2.84	0.25	9.24	70.90
cluster 7	Shaw, Logan Circle	7.59	23.83	6.78	0.55	8.63	52.61
cluster 8	Downtown, Chinatown, Penn Quarters, Mount Vernon Square, North Capitol Street	11.93	23.82	3.27	0.39	8.72	51.87

Table 3: Demographic Breakdown of Property Crime Perpetrators in bottom 10 neighborhood

Cluster_Id	**counties**	**Asian**	**Black**	**Hispanic**	**Native**	**Others**	**White**
cluster 10	Hawthorne, Barnaby Woods, Chevy Chase	5.48	9.03	1.64	0.23	10.12	73.50
cluster 12	North Cleveland Park, Forest Hills, Van Ness	9.10	11.91	2.57	0.26	9.80	66.34
cluster 13	Spring Valley, Palisades, Wesley Heights, Foxhall Crescent, Foxhall Village, Georgetown Reservoir	9.56	6.17	2.62	0.18	9.56	71.92
cluster 14	Cathedral Heights, McLean Gardens, Glover Park	9.19	7.47	2.89	0.32	10.60	69.53
cluster 20	North Michigan Park, Michigan Park, University Heights	3.03	56.66	7.96	0.50	7.52	24.33
cluster 29	Eastland Gardens, Kenilworth	0.05	91.11	1.89	0.26	3.95	2.74
cluster 38	Douglas, Shipley Terrace	0.30	94.02	1.11	0.27	3.12	1.18
cluster 41	Rock Creek Park	4.64	4.00	2.40	0.32	8.48	80.16

cluster 44	Joint Base Anacostia-Bolling	2.85	30.79	4.28	0.59	12.09	49.40
cluster 5	West End, Foggy Bottom, GWU	16.34	6.69	2.86	0.28	8.03	65.80

The top 10 clusters with high property crime rates do not exhibit a consistent majority demographic, with clusters like 40 and 35 having a higher Black demographic, while cluster 25 has a predominantly White demographic. Comparing this to violent crime rates, where a higher Black demographic was more prevalent across high-rate clusters, the property crime data shows greater demographic diversity.

In the bottom 10 clusters for property crime, the White demographic is more dominant, (non-dominant in cluster 29, 38) a pattern that aligns with the violent crime data. These variances underscore that while certain demographics may appear more frequently in areas with higher crime rates, the relationship is complex and not uniformly applicable across different types of crime.

Conclusion

In summary, my comprehensive analysis of crime and transit patterns in Washington D.C. unveils a nuanced relationship between metro bus stop proximity and crime rates. Choropleth maps and Leaflet visualizations depict higher crime rates in central clusters, particularly around Howard University, with a significant correlation between crime prevalence and metro bus stop density in select areas. Arson followed by sex abuse emerges as having the strongest relationship among various crime types around metro bus stops, emphasizing the importance of considering crime categories individually. All of the crimes (robbery, sex abuse, assault) showed moderate positive correlation with the transport rate. Furthermore, The examination of property and violent crime rates in D.C. shows nuanced patterns: while higher Black demographics are prevalent in the top 10 high-crime clusters, this trend does not uniformly apply, as seen in clusters like 35. Conversely, the bottom 10 clusters, with lower property crime rates, generally exhibit higher White demographics, mirroring trends seen in violent crime analysis. These findings highlight the complexity of urban crime dynamics, where demographic factors are significant but not the sole determinants.