Community Planning in Washington D.C.

1 Introduction

1.1 Background

Washington D.C. has a racially and economically diverse population. While the current median income of D.C. as a whole is $85,203, this is hardly representative of the unique communities that compose the many neighborhoods of D.C. At 68 square miles with a population of 705,749 people, D.C. is made up of more than two dozen different neighborhoods with various cultures and personalities. Additionally, the racial and ethnic demographics of D.C. are reported to be 46% Black, 37.5% Non-Hispanic White, 11.3% Hispanic, 4.5% Asian, 2.9% Two or More Races, and 0.1% Native Hawaiian and Other Pacific Islanders[1], but this distribution may vary from neighborhood to neighborhood.

1.2 Problem and Approach

In order to determine whether the needs of the communities are being met by the businesses and services within those communities, one needs to look at the racial and economic distribution within those communities rather than as a sum of all of D.C. There are a number of factors that contribute to the culture and economy of a neighborhood including the racial makeup of the population as well as the median household income. Additionally, necessary resources for the community may include access to fresh and healthy food choices over fast food choices, athletic facilities or clean outdoor spaces for recreation, and educational resources such as museums and bookstores.

The goals of this project will be two-fold. First, each neighborhood in D.C. will be characterized by determining the racial distribution and median household income for each zip code in Washington D.C. Some questions that will be explored during this part of the analysis include:

- **Racial distribution across zip codes**
  - Does the racial distribution of each zip code mirror the racial distribution of D.C. as a whole?
  - That is, is the racial distribution fairly consistent throughout D.C. or do some neighborhoods have higher concentrations of some races/ethnicities than others?
- **Wealth distribution across zip codes**
  - How does the median household income vary across zip code?
  - Is the median household income for each zip code similar to the median household income for all of D.C.?
  - Or is the median household income higher in some zip codes and lower in others indicating a wealth disparity across zip codes?

Secondly, the businesses, services, and facilities in the community will be assessed and categorized to determine whether the needs of that community are being met. The venues will

[1] US Census Bureau Quick Facts on Washington D.C. https://www.census.gov/quickfacts/fact/table/DC/PST045219

be categorized into various categories representing different needs of the community. These include:

- **Access to fresh, healthy, and affordable food**:
    - Access to grocery stores and farmer's markets increase the likelihood that members of the community will be able to maintain a balanced and nutritious diet. In contrast, areas with high concentrations of fast food restaurants, which typically serve food high in fat, sodium, and calories, are more likely to inhibit community members' ability to maintain that balanced diet.
- **Access to recreational or fitness activities**:
    - Nearby fitness facilities, parks, or other outdoor areas increase the likelihood that members of the community will be able to lead an active lifestyle
- **Access to educational resources**:
    - D.C. is known for having a plethora of museums with free admission as well as bookstores which provide additional educational opportunities to members of the community outside of traditional schooling.
- **Access to fast food options**:
    - If a community has more fast food options than grocery stores, members of the community are more likely to choose these options rather than healthy, nutritious meals.
- **Diversity of restaurant choices**:
    - The selection of restaurants within a community may impact people's decisions when it comes to selecting nutritious and delicious foods.

1.3 Interest

Community organizations looking to improve the conditions of their neighborhoods will be very interested in understanding the demographics and needs of those neighborhoods. Finding that a neighborhood is lacking resources like access to healthy foods, outdoor areas or exercise facilities, or access to food from a desired culture, can impact how an organization decides to invest in the community. Additionally, when applying for government funding for community programs, those organizations need to explain why those funds are needed, and how they will be used effectively.

2 Data Sources

2.1 Demographic Data scraped from D.C. Health Matters

D.C. Health Matters[2] provides a one-stop resource for online access to community health indicators and related resources that impact the health of D.C. communities. You will find up-to-date demographic, health and social determinants data; hundreds of maps, tables and figures; and, promising practices. Data was scraped from D.C. Health Matters to obtain the racial demographics and median household incomes of 29 zip codes throughout D.C.:

[2] https://www.dchealthmatters.org/

D.C. Zip Codes: [20001, 20002, 20003, 20004, 20005, 20006, 20007, 20008, 20009, 20010, 20011, 20012, 20015, 20016, 20017, 20018, 20019, 20020, 20024, 20032, 20036, 20037, 20052, 20057, 20064, 20319, 20373, 20374, 20515]

The racial demographics were represented by providing the populations of White, Black, Asian, American Indian/Alaska Native, Native Hawaiian/Pacific Islander, Mixed Race, and Other races in each zip code. Additionally, each race was categorized as either Hispanic/Latino and non-Hispanic/Latino creating 14 groups of racial demographics for each zip code.

1. Hispanic/Latino Population: White
2. Hispanic/Latino Population: Black/African American
3. Hispanic/Latino Population: American Indian/Alaska Native
4. Hispanic/Latino Population: Asian
5. Hispanic/Latino Population: Native Hawaiian/Pacific Islander
6. Hispanic/Latino Population: Some Other Race
7. Hispanic/Latino Population: 2+ Races
8. Non-Hispanic/Latino: White
9. Non-Hispanic/Latino: Black/African American
10. Non-Hispanic/Latino: American Indian/Alaska Native
11. Non-Hispanic/Latino: Asian
12. Non-Hispanic/Latino: Native Hawaiian/Pacific Islander
13. Non-Hispanic/Latino: Some Other Race
14. Non-Hispanic/Latino: 2+ Races

2.2 Venue Data pulled from FourSquare API

The Foursquare API[3] provides information about venues, users, photos, and check-ins based on location. The API supports real time access to places, Snap-to-Place that assigns users to specific locations, and Geo-tags. Data was obtained from FourSquare by using the explore function to explore the venues within each zip code specified within the D.C. Health Matters data. The data obtained from the FourSquare API included the zip code, zip code longitude and latitude, venue, venue longitude and latitude, and the venue category.

The venue categories are used to further categorize the venues into five groups:

1. Educational Resources
2. Fast Food Places
3. Restaurants
4. Athletic or Recreational Facilities
5. Grocery Stores or Farmer's Markets

The number and types of facilities for each zip code were assessed with respect to the population and demographics of each zip code. Additionally, the diversity of restaurants is analyzed by

[3] https://api.foursquare.com/v2/venues/

grouping restaurants according to type and investigating how the diversity of restaurants reflects the neighboring community.

2.3 GeoJSON Data downloaded from Github user, Ben Balter

GeoJSON provides the location data for Washington D.C. and is necessary for plotting maps with the Folium library. Github user, Ben Balter[4] has collected GeoJSON data for various categories of venues through D.C.[5] from Open Data DC[6] and stored them in his account. The features of this data include the zip code and the boundaries in longitude and latitude values of the area that contains that zip code. This data will be utilized to understand the geographic location throughout D.C. for each zip code as well as for plotting where venues are located to better understand how venues are clustered throughout D.C. The raw GeoJSON data can be found here: https://raw.githubusercontent.com/benbalter/dc-maps/master/maps/zip-codes.geojson.

3 Methodology

3.1 Data cleansing

In order to reduce the noise in the dataset that can be caused by small population numbers in each race/ethnicity category, the racial demographic data was consolidated into eight categories by grouping all the Hispanic/Latino race data into one category representing the Hispanic/Latino community. The non-Hispanic/Latino race data was left as is so that the remaining racial demographics include White, Black, Asian, American Indian/Alaska Native, Native Hawaiian/Pacific Islander, 2+ Races (renamed Mixed Race), Hispanic/Latino, and Other Races. The median household income data and longitude/latitude data were then merged with the racial demographics.

3.2 Exploratory Data Analysis

3.2.1 Characterization of D.C. Demographics

The racial demographics of each zip code were summed to first understand the overall racial demographics and population distribution across D.C. Figure 1 shows the demographic distribution. The population of Washington D.C. is mostly non-Hispanic/Latino Black and White people. Overall, the demographics are 43.64% Black, 37.15% White, 11.97% Hispanic/Latino, 4.3% Asian, 2.47% Mixed Race, 0.23% Other, 0.19% American Indian/Alaska Native, and 0.05% Native Hawaiian/Pacific Islander.

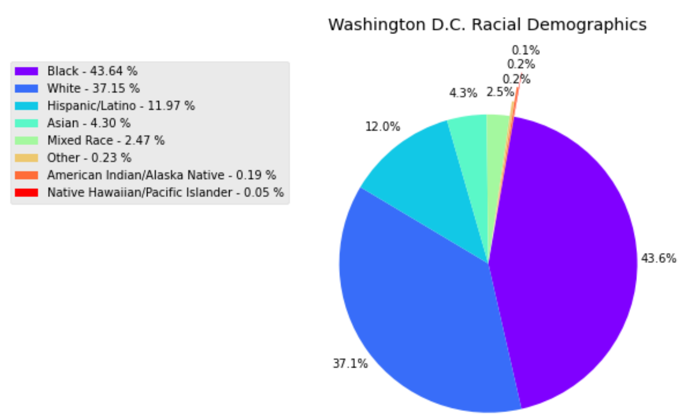

Figure 1. Racial demographics of Washington D.C.

[4] https://github.com/benbalter
[5] https://github.com/benbalter/dc-maps
[6] http://opendata.dc.gov/

The Folium library in Python was used to visualize the geographic details of Washington D.C. Figure 2 shows a map of D.C. colored by population with markers at the central longitude/latitude for each zip code in the dataset. The darker purple shows the more highly populated areas of Washington D.C. while the light blue colors represent the less populated zip codes. The map shows that the number of people who live in D.C. are not evenly distributed across zip codes. The most populated zip codes are 20002, 20011, and 20019. The area of the different zip codes varies greatly such that the population does too. The zip codes with the lowest population are 20374 and 20515, which have only 4 people each. Conversely, zip code 20002 has 73,910 people.

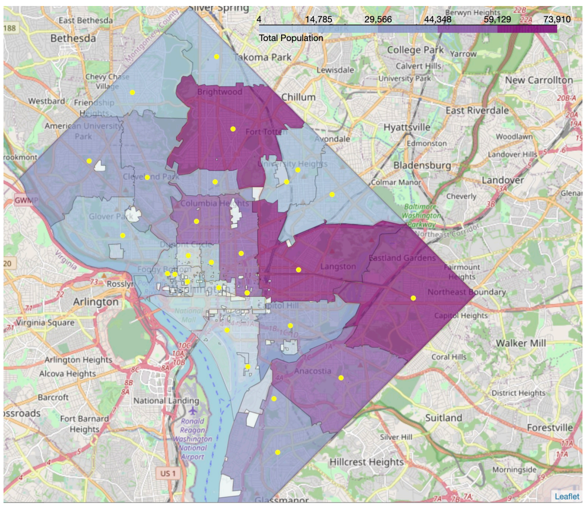

Figure 2. Map of D.C. colored by population

3.2.2 Descriptive Statistics

Using descriptive statistics, it is revealed that the demographics across zip codes vary greatly. Table 1 shows the descriptive statistics for each population and the median household income data. It shows that the standard deviation for any population is almost as much as or even greater than the mean population across zip codes. Similarly, the median household income has a dramatic range from the minimum of around $15K to a maximum of $200K.

The number of people per zip code will be an important consideration when determining community needs as a zip code with only 4 people will have fewer needs than one with tens of thousands of people. The median household income is also relevant since businesses may gravitate to wealthier communities for the investment while denying these resources to poorer communities.

Table 1. Descriptive statistics of racial and economic demographics of D.C.

	WHITE	BLACK	AMERICAN INDIAN/ ALASKA NATIVE	ASIAN	NATIVE HAWAIIAN/ PACIFIC ISLANDER	OTHER	MIXED RACE	HISPANIC/ LATINO	TOTAL POPULATION	MEDIAN HOUSEHOLD INCOME
COUNT	29	29	29	29	29	29	29	29	29	29
MEAN	9186.93	10792.72	47.38	1062.41	13.52	57.76	610.38	2959.55	24730.66	97238.90
STD	10361.02	16383.42	48.16	1172.74	16.24	56.67	528.85	4283.63	23191.77	45180.64
MIN	1	2	0	0	0	0	0	0	4	14999
25%	986	262	5	191	1	3	166	334	3392	73771
50%	4702	1782	31	648	6	47	532	1557	17258	95997
75%	12902	11246	71	1589	20	85	971	3056	36931	125361
MAX	34890	56818	157	4584	68	204	1933	20169	73910	200000

3.2.3 Characterization of Each Zip Code

The next step of analyzing each community is to determine the racial demographics of each zip code to see if they mirror the demographics of D.C. as a whole. Figure 3 shows the demographics by zip code by looking at the percentage of each race with respect to the total population. Using the percentages rather than raw numbers normalizes the populations to account for the very small populations of some zip codes. Expectedly, the most prominent populations are White and Black, however, one notable observation is that the zip codes tend to be dominated by *either* White or Black populations. Only zip code 20002, which again has the largest total population, has a balanced population of both Black and White people.

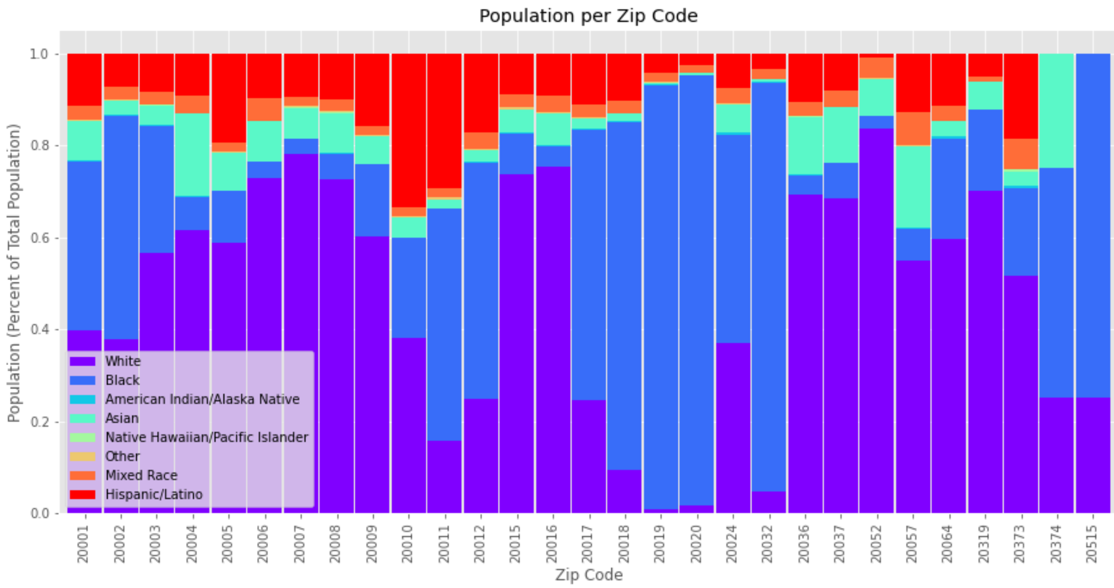

Figure 3. Plot of the racial distribution of each zip code in D.C.

Additionally, while the Hispanic/Latino population is less than the populations of Black and White people, there are certain zip codes where the Hispanic/Latino population is much greater than others such as zip code 20009 and 20010.

3.3 Venue Categorization

Venue categorization was necessary to identify venues corresponding to the various resource needs of the communities. In order to categorize each venue, keywords were selected that represent the five community needs that were identified at the beginning of this project. The keywords selected for each community need are listed in Table 2. The venue categories pulled from the FourSquare API were then searched to find these identifying categories, and if a keyword was found, the venues were labels as a member of that category. One important note is that some venue categories were listed as “Fast Food Restaurant”, which then were labeled as belonging to both the “Fast Food” and “Restaurant” categories. These venues only belong in the “Fast Food” category, so after the initial keyword search, any venues in the “Restaurants” dataset containing the keyword “Fast Food” were removed from that dataset.

Table 2. Community Need Categories with Identifying Keywords

Community Need	Keywords
Athletic/Recreational Facility	'Bike|Tennis|Recreation|Yoga|Field|Skate|Pool|Park|Martial Arts|Gym|Dance'
Restaurants	'Restaurant|Steakhouse|Noodle|Diner'
Fast Food Places	'Fast Food|Food Truck|Joint|Place'
Grocery Stores	'Grocery|Supermarket|Fish Market|Farmers Market|Bodega'
Educational Resources	'College|Student|Bookstore|Museum'

Using the labels assigned to each venue based on the keywords, it was possible to then calculate how many venues corresponding to each type of community need were found in each zip code as well as do a correlation analysis as to whether race and/or median household income were relevant factors in where businesses are developed.

3.4 Correlation Analysis

A correlation analysis was performed between each race and the median household income as well as the number of each type of venue. Correlation is the measure of the extent of interdependence between variables. The Pearson Correlation measures the linear dependence between two variables and assigned a coefficient with a value between -1 and 1. A value of 1 means that the two variables have a total positive correlation. A value of 0 means that there is no linear correlation and that the two variables likely do not affect each other, and a value of -1 means the two variables have a total negative correlation. In addition to the correlation coefficient, the statistical significance of each correlation can be determined by calculating a P-value. The P-value is the probability value that the correlation between these two variables is statistically significant. In other words, the P-values represents the confidence level we can have that the correlation between two variables is significant. The larger the P-value, the less confident the evidence there is that a correlation is significant.

By convention:

- P- value << 0.001: 99.9% confidence; strong evidence that the correlation is significant
- P-value << 0.05: 95% confidence; moderate evidence that the correlation is significant
- P-value << 0.1: 90% confidence; weak evidence that the correlation is significant
- P-value >> 0.1: Less than 90% confidence; no evidence that the correlation is significant

While correlation does not equal causation, meaning that the value of X variable may not necessarily have caused the variations in Y, it can present a starting plan to look for causation by investigating the reasons for those correlations.

3.5 Venue Clustering

An initial plot of the categorized venues on the map of D.C. shows a number of clusters of venues spread around Washington D.C. with the largest cluster around central D.C. (Figure 4). Two methods of clustering were selected to analyze the clusters of venues: **K-means** clustering and **Density-based** clustering. The number of clusters found via each method was compared, and then the final clusters were analyzed for their composition of each type of venue (i.e. Restaurants, Fast Food, Education, Grocery Stores, or Athletic/Recreational Facilities), the diversity of restaurant types, and the racial demographics. The racial demographics for the area surrounding the clusters was calculated by determining what zip codes fall within those clusters, summing the populations of those zip codes, and finally calculating the percentage of each race with respect to the total population.

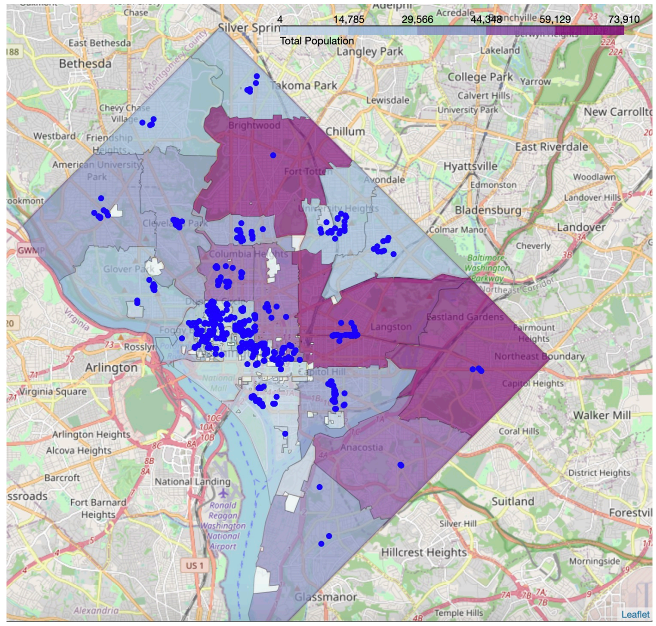

Figure 4. Map of D.C. with markers representing each venue

4 Results

4.1 Venue Categorization

The dataset pulled from the FourSquare API returned 1081 total venues within 233 unique categories. Venue categorization resulted in finding 537 venues that fall into the chosen categories with most venues being non-fast food restaurants. The number of venues in each category are as follows:

- Restaurants: 245
- Fast Food Places: 131
- Athletic/Recreational Facilities: 77
- Educational Resources: 42
- Grocery Stores: 42

The number of venues of each category was then calculate for each zip code and used for the correlation analysis.

4.2 Correlation Analysis

After performing venue categorization, all the venues that were labeled with the five selected categories were summed and plotted against each zip code (Figure 5). Surprisingly, the number of venues per zip code did not follow the same pattern as the total population, meaning that zip codes with high populations sometimes had very few venues while zip codes with low populations often had a relatively high number of venues. In order to determine whether this was a significant correlation, the Pearson coefficient and P-Value was calculated between the Total Population and the Number of Venues. For this correlation, the Pearson Coefficient = -0.25 with a P-value of P = 0.19 indicating that there is no evidence of the significance of the apparent correlation.

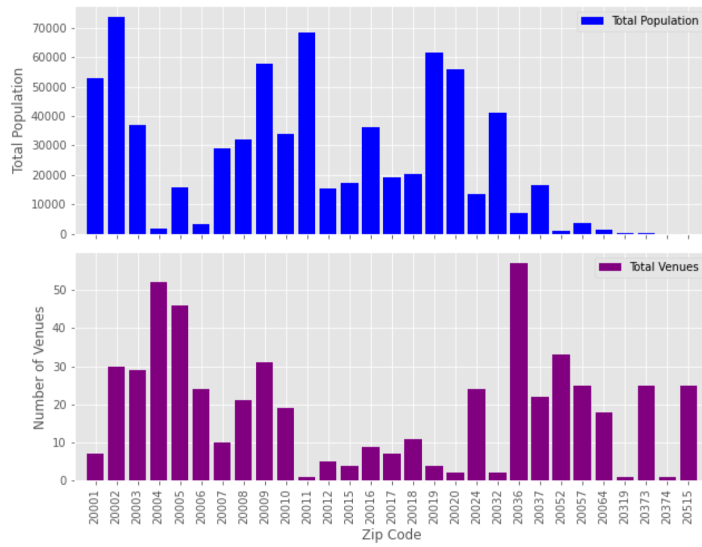

Figure 5. Total Population and Number of Venues per Zip Code

Since the sum of the values did not show evidence of a correlation, the correlation of each individual variable was calculated with respect to each other variable. While no correlations contained strong evidence, several had moderate or weak evidence. These are summarized in Table 3. A notable observation from these results is that only one variable combination has a positive correlation. The correlation between the White population and the number of Grocery Stores has a Pearson coefficient of 0.33, indicating slight correlation, and a P-Value of 0.076, indicating only week evidence of this correlation. Every other Pearson coefficient indicates a negative correlation with either moderate or weak evidence of correlation. In most cases, these correlations occur where the population is mostly Black.

Table 3. Pearson Coefficients for Race and Venue Variables

EVIDENCE	VARIABLE 1	VARIABLE 2	PEARSON COEFFICIENT	P-VALUE
WEAK	Black	Athletics	-0.348	0.0643
WEAK	White	Grocery Stores	0.331	0.0796
WEAK	Black	Grocery Stores	-0.354	0.0596
WEAK	Black	Fast Food	-0.337	0.0737
WEAK	Other	Fast Food	-0.35	0.0629
WEAK	Total Population	Fast Food	-0.342	0.0693
MODERATE	Black	Total Venues	-0.376	0.0444

The Pearson coefficients were also calculated for each race with respect to the median household income to identify if there were any racial disparities with regards to the wealth distribution of communities in D.C. Again, the correlation coefficients showed no strong evidence for correlation, however, there were three instances of moderate evidence indicating correlation. These are summarized in Table 4.

Table 4. Pearson Coefficients for Race and Median Houshold Income

EVIDENCE	VARIABLE 1	VARIABLE 2	PEARSON COEFFICIENT	P-VALUE
MODERATE	White	Median Household Income	0.424	0.0218
MODERATE	Black	Median Household Income	-0.415	0.0252
MODERATE	Asian	Median Household Income	0.391	0.0359

As with the venue data, there are positive correlations between the White and Asian populations and median household income, but there is a negative correlation between Black populations and median household income.

4.3 Venue Clustering

4.3.1 K-Means Clustering

To identify the optimal number of venue clusters using the K-means clustering method, the elbow method was used. The elbow method is a heuristic used in determining the number of clusters in a dataset. For K-means clustering, it runs the clustering algorithm on the dataset for a range of K values, and plots the sum of the square distances from each point to the assigned center of its cluster (Figure 6). The optimal K-value is where the bend in the curve occurs. In this data set, the optimal K-value falls at K=5.

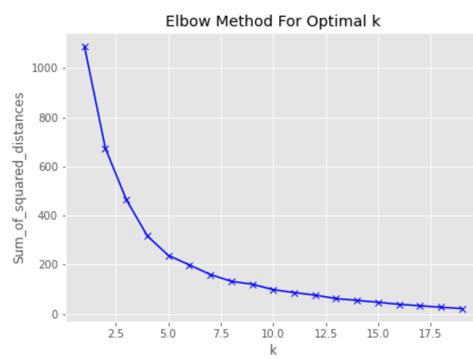

Figure 6. Sum of squared distances vs. K-value

Using K=5, the venue clusters identified by K-means clustering are plotted in Figure 7. As shown in the figure, the K-means method found some relatively dense clusters, such as clusters 1, 2, and 4, while clusters 0 and 3 feature smaller sub-clusters that are spread out.

4.3.2 Density-based Clustering

Density-based clustering finds the number of clusters based on the density between points, which means that the cluster will not be as spread out as those found with K-means clustering. The density-based clustering method found 10 clusters within the venue dataset as well as several outliers, which are individual venues or small groups containing less than seven venues.

The results of both the K-Means clustering and the Density-based clustering are shown in Figure 7.

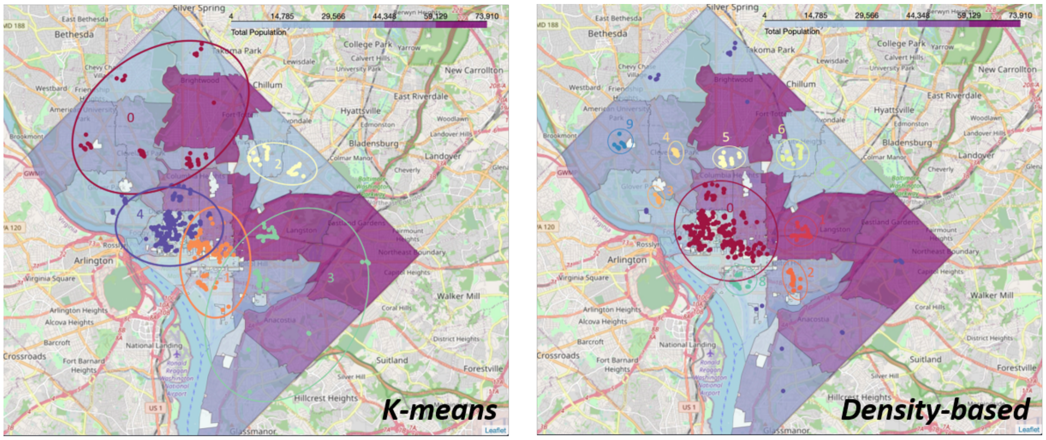

Figure 7. Results of venue clustering for both K-means and Density-based algorithms

In both cases, the densest cluster or clusters appears in central D.C. whereas the most populated areas of D.C. are shown to be in the northern and eastern parts of D.C. The density-based cluster analysis seems to more accurately represent the clusters observed on the map of D.C. In order to then further analyze these cluster, the venue distribution, racial distribution, and restaurant diversity was calculated. In order to calculate the racial distribution, the race population of each zip code that falls within that cluster was summed, and the percentage of each race calculated from the sum. A summary of the cluster analysis is provided in Appendix A. Density-based Cluster Analysis.

5 Discussion

5.1 Potential Bias in Resource Distribution

As observed in Section 4.2, positive correlations were observed for White and Asian populations with respect to the median household income and total number of venues. Conversely, Black populations often saw a negative correlation when it came to median household income, grocery stores, athletic and recreational facilities, and the overall total number of venues found in primarily Black communities. This trend could be indicative of potential bias when it comes to resource distribution. Figure 8 shows the correlation plots for White, Black, and Asian populations with respect to the Median Household Income.

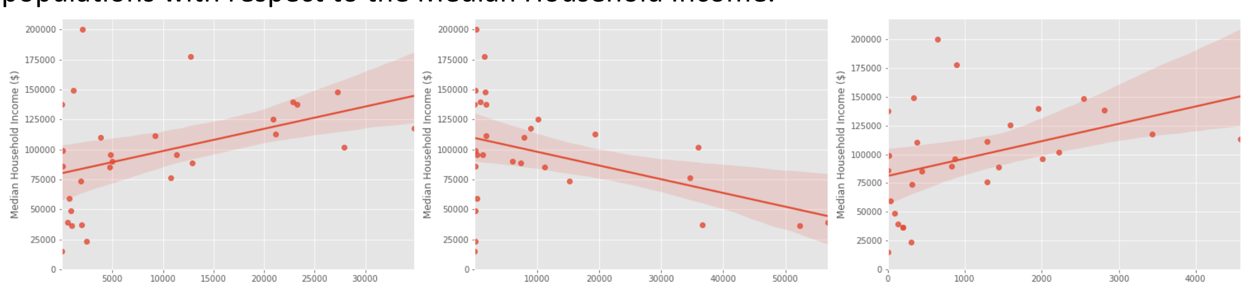

Figure 8. Correlation plots for White, Black, and Asian Populations with respect to Median Household Income

The data points for each plot are fairly scattered and noisy, which is why the evidence for the correlations are considered moderate. Therefore, it is difficult to determine from these correlations whether bias is present.

5.2 Resource Allocation

It is clear from the venue clustering that the majority of venues are located in the lower populated regions. This is likely due to the tourism industry of Washington D.C. Restaurants are the primary type of venue observed in the dataset, which would also be most attractive to tourists. In analyzing the individual clusters (Appendix A), there does not appear to be any bias towards development of venues in neighborhoods with respect to the racial demographics. Additionally, the diversity of restaurants seems dependent on the number of restaurants in that cluster rather than any racial demographic. For example, the largest cluster, which as 339 venues, contains 45% restaurants including 34 different ethnic choices as well as vegetarians/vegan and gluten-free options. Even the cluster with the least amount of venues, cluster 3, which has only 8 venues has 50% restaurants including options for French, Asian, Turkish, or Italian food. Conversely, cluster 9, which has only 9 venues and similar racial demographics, has no restaurants. Therefore, further investigation would need to be done to determine how resources are allocated across zip codes.

An interesting future study would be to analyze the distance of venues from the central tourism spots in order to confirm the conclusion that venue development is based on tourism rather than residence. Also, missing from this dataset are primary schools or tutoring services since the FourSquare API relies on "check in" information, and users typically do not check in to those services. In order to fully understand what sort of educational resources are being offered to the local community, another database should be mined to discover K-12 resources and determine whether those too are meeting community needs.

6 Conclusions

The economic and racial distribution across zip codes were analyzed to understand the specific needs within each community that represents Washington D.C. It was observed that different zip codes throughout D.C. tend to have either a primarily White or primarily Black population and that there may be a potential bias when it comes to wealth distribution. There was a moderate negative correlation when it came to Black populations and median household income and the number of venues in those communities. While there was only moderate to weak evidence of these correlations, it is worth investigating further in order to make sure resources are being allocated to communities in need.

Aside from a potential racial bias, there also seems to be a bias for developing venues in central D.C. while developing very little in the neighboring communities. This could be due to the increased tourism industry in central D.C. Most of the venues discovered were restaurants, which would also be favorable to the tourism industry over educational resources or grocery stores. Further study is needed to determine whether the tourism industry is causing this bias. Also, additional dataset should be sought out that contain information about the educational resources in each zip code as the FourSquare data was lacking in this area.

7 Appendix A. Density-based Cluster Analysis

CLUSTER	VENUES	VENUE TYPES	VENUE DISTRIBUTION	RESTAURANT TYPES	RACE DISTRIBUTION
0	339	72	[('Athletics: 12.98%', 'Grocery: 7.37%', 'Education: 8.26%', 'Fast Food: 25.96%', 'Restaurants: 45.43%')]	['Southern / Soul Food', 'Thai', 'Brazilian', 'Steakhouse', 'American', 'Indian', 'Seafood', 'Diner', 'Italian', 'Tapas', 'Peruvian', 'Gluten-free', 'Mediterranean', 'French', 'Asian', 'Mexican', 'Arepa', 'Vegetarian / Vegan', 'Latin American', 'Eastern European', 'Belgian', 'Sushi', 'Korean', 'Middle Eastern', 'Ramen', 'German', 'Japanese', 'Ethiopian', 'Falafel', 'Vietnamese', 'Spanish', 'Portuguese', 'Persian', 'Russian', 'Israeli', 'Turkish', 'Greek', 'Chinese', 'English']	[('White: 54.87%', 'Black: 20.16%', 'Hispanic/Latino: 13.28%', 'Asian: 8.44%', 'American Indian/Alaska Native: 0.16%', 'Native Hawaiian/Pacific Islander: 0.06%', 'Mixed Race: 2.79%', 'Other: 0.25%')]
1	34	26	[('Athletics: 17.65%', 'Grocery: 5.88%', 'Education: 8.82%', 'Fast Food: 26.47%', 'Restaurants: 47.06%')]	['Asian', 'Pakistani', 'American', 'Southern / Soul Food', 'Swiss', 'Burmese', 'Cajun / Creole', 'Sushi', 'Mexican', 'Thai', 'Ethiopian', 'Korean']	[('White: 37.76%', 'Black: 48.70%', 'Hispanic/Latino: 7.39%', 'Asian: 3.01%', 'American Indian/Alaska Native: 0.21%', 'Native Hawaiian/Pacific Islander: 0.09%', 'Mixed Race: 2.62%', 'Other: 0.22%')]
2	28	21	[('Athletics: 7.14%', 'Grocery: 3.57%', 'Education: 7.14%', 'Fast Food: 17.86%', 'Restaurants: 64.29%')]	['Chinese', 'American', 'Greek', 'Eastern European', 'Belgian', 'Seafood', 'Italian', 'Sushi', 'Asian', 'Mediterranean', 'Spanish', 'Mexican']	[('White: 56.58%', 'Black: 27.51%', 'Hispanic/Latino: 8.27%', 'Asian: 4.30%', 'American Indian/Alaska Native: 0.29%', 'Native Hawaiian/Pacific Islander: 0.05%', 'Mixed Race: 2.77%', 'Other: 0.22%')]
3	8	7	[('Athletics: 12.50%', 'Grocery: 25.00%', 'Education: 0.00%', 'Fast Food: 12.50%', 'Restaurants: 50.00%')]	['French', 'Asian', 'Turkish', 'Italian']	[('White: 78.17%', 'Black: 3.18%', 'Hispanic/Latino: 9.37%', 'Asian: 6.68%', 'American Indian/Alaska Native: 0.10%', 'Native Hawaiian/Pacific Islander: 0.07%', 'Mixed Race: 2.13%', 'Other: 0.29%')]
4	20	16	[('Athletics: 5.00%', 'Grocery: 10.00%', 'Education: 0.00%', 'Fast Food: 20.00%', 'Restaurants: 65.00%')]	['Steakhouse', 'Israeli', 'Italian', 'Mediterranean', 'Indian', 'Xinjiang', 'Thai', 'Asian', 'Mexican']	[('White: 72.53%', 'Black: 5.60%', 'Hispanic/Latino: 10.02%', 'Asian: 8.76%', 'American Indian/Alaska Native: 0.13%', 'Native Hawaiian/Pacific Islander: 0.06%', 'Mixed Race: 2.63%', 'Other: 0.26%')]
5	20	17	[('Athletics: 15.00%', 'Grocery: 5.00%', 'Education: 0.00%', 'Fast Food: 10.00%', 'Restaurants: 70.00%')]	['Mexican', 'American', 'Asian', 'Filipino', 'Vietnamese', 'Caribbean', 'Cuban', 'Indian', 'Mediterranean', 'Salvadoran', 'South American', 'Latin American', 'Seafood']	[('White: 37.99%', 'Black: 21.82%', 'Hispanic/Latino: 33.62%', 'Asian: 4.24%', 'American Indian/Alaska Native: 0.16%', 'Native Hawaiian/Pacific Islander: 0.06%', 'Mixed Race: 1.86%', 'Other: 0.26%')]
6	20	15	[('Athletics: 20.00%', 'Grocery: 5.00%', 'Education: 20.00%', 'Fast Food: 20.00%', 'Restaurants: 35.00%')]	['Vietnamese', 'Ethiopian', 'Latin American', 'American', 'Mexican', 'Chinese']	[('White: 26.74%', 'Black: 56.45%', 'Hispanic/Latino: 11.06%', 'Asian: 2.38%', 'American Indian/Alaska Native: 0.24%', 'Native Hawaiian/Pacific Islander: 0.02%', 'Mixed Race: 2.82%', 'Other: 0.30%')]

7	12	11	[('Athletics: 41.67%', 'Grocery: 8.33%', 'Education: 0.00%', 'Fast Food: 25.00%', 'Restaurants: 25.00%')]	['Chinese', 'American']	[('White: 9.28%', 'Black: 75.61%', 'Hispanic/Latino: 10.27%', 'Asian: 1.57%', 'American Indian/Alaska Native: 0.34%', 'Native Hawaiian/Pacific Islander: 0.01%', 'Mixed Race: 2.72%', 'Other: 0.19%')]
8	25	17	[('Athletics: 12.00%', 'Grocery: 12.00%', 'Education: 8.00%', 'Fast Food: 24.00%', 'Restaurants: 44.00%')]	['Spanish', 'Mexican', 'Mediterranean', 'Italian', 'Falafel', 'Caribbean', 'Seafood']	[('White: 37.03%', 'Black: 45.18%', 'Hispanic/Latino: 7.62%', 'Asian: 6.14%', 'American Indian/Alaska Native: 0.48%', 'Native Hawaiian/Pacific Islander: 0.07%', 'Mixed Race: 3.21%', 'Other: 0.26%')]
9	9	9	[('Athletics: 33.33%', 'Grocery: 0.00%', 'Education: 44.44%', 'Fast Food: 22.22%', 'Restaurants: 0.00%')]	None	[('White: 75.32%', 'Black: 4.50%', 'Hispanic/Latino: 9.33%', 'Asian: 7.02%', 'American Indian/Alaska Native: 0.12%', 'Native Hawaiian/Pacific Islander: 0.04%', 'Mixed Race: 3.36%', 'Other: 0.31%')]
-1	22	15	[('Athletics: 27.27%', 'Grocery: 18.18%', 'Education: 0.00%', 'Fast Food: 31.82%', 'Restaurants: 22.73%')]	['Caribbean', 'American', 'Seafood', 'Indian']	[('White: 18.57%', 'Black: 65.90%', 'Hispanic/Latino: 11.27%', 'Asian: 1.74%', 'American Indian/Alaska Native: 0.19%', 'Native Hawaiian/Pacific Islander: 0.05%', 'Mixed Race: 2.08%', 'Other: 0.21%')]