The Battle of Neighborhoods

An Analysis of Postal Zones in Montgomery County, Maryland

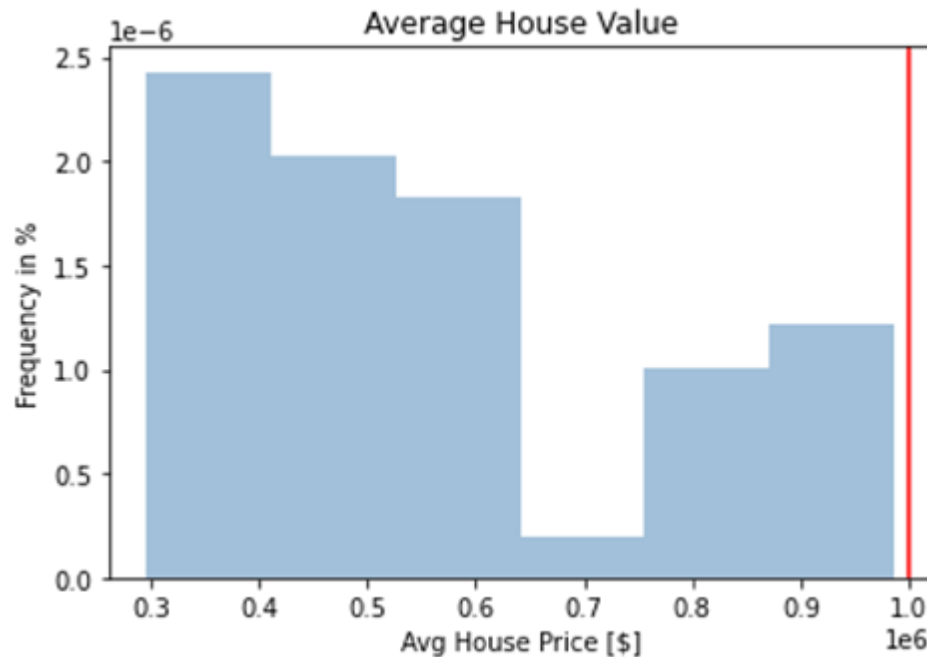Capstone Project

Introduction

- **Montgomery County is the most populous county in the U.S. state of Maryland, located adjacent to Washington, D.C.**

- **Eric is a biological scientist who is moving to Montgomery County. Eric and his family have their own expectations and budget for this new home.**

- **To help Eric start the house hunting easily, I designed this project to let him know the house market and common feathers in Montgomery county quickly.**

- **This analysis can provide help for other people who have the similar needs as Eric to find a good neighborhood in Montgomery county.**

Data Preparation

- **To develop the analysis, I collected the following data of Montgomery county through web scraping and downloading:**
    - **Zip Codes and City Names**

      https://www.zip-codes.com/county/md-montgomery.asp
    - **Demographic Data and Housing Data**

      https://www.zip-codes.com/county/md-montgomery.asp
    - **Montgomery's GeoJSON**
    - **Venues**

Data Visualization
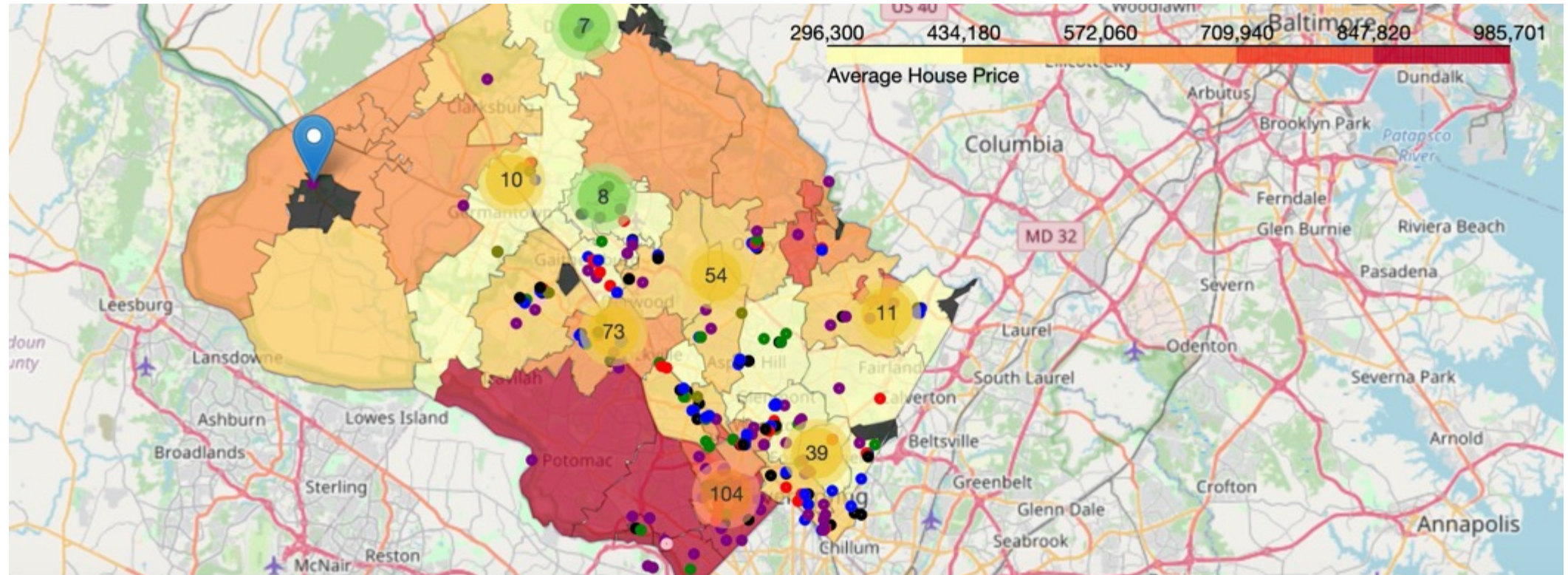
Histogram of **Average House Value**

The red line is Eric's budget of $1 million for his new house, which is higher than the average house price in any postal zone. There, house price budget is not a constraint in selecting a neighborhood.

Methodology
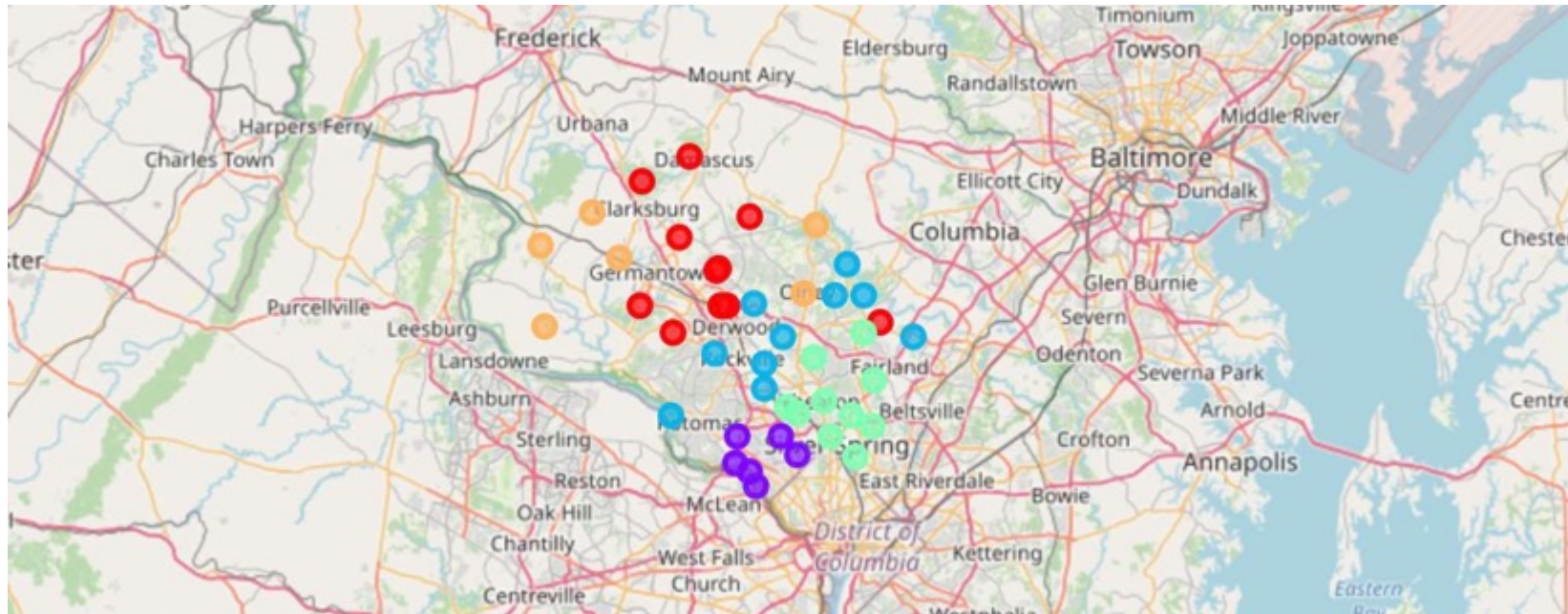
- Eric's Preferred Category Analysis

Methodology

- Eric's Preferred Category Analysis

From the top 5 postal zones, the 20815 has the greatest number of preferred categories Eric's family expect for and the house price is closest to their budget.

ZipCode	City	Avg_House_Val	Grocery Store	Gym / Fitness Center	Italian Restaurant	Mexican Restaurant	Park	Pizza Place	Theater	Total PreCate
20815	CHEVY CHASE	985700	3	2	1	4	6	3	1	20
20910	SILVER SPRING	556000	3	1	1	5	4	3	0	17
20902	SILVER SPRING	392100	3	1	1	2	5	5	0	17
20912	TAKOMA PARK	518400	3	0	0	1	5	6	0	15
20814	BETHESDA	759100	1	1	1	2	6	3	1	15

Methodology

- K-means Clustering Analysis

Methodology

- K-means Clustering Analysis

   The following cluster includes postal area 20815 which is the zone selected from preferred categories analysis. There are another four zip codes in this cluster.

ZipCode	City	Avg_House_Val	Grocery Store	Gym / Fitness Center	Italian Restaurant	Mexican Restaurant	Park	Pizza Place	Theater	Total PreCate
20815	CHEVY CHASE	985700	3	2	1	4	6	3	1	20
20814	BETHESDA	759100	1	1	1	2	6	3	1	15
20812	GLEN ECHO	898000	1	1	1	1	2	1	2	9
20816	BETHESDA	939700	0	1	0	0	4	0	2	7
20818	CABIN JOHN	884900	1	1	1	1	2	1	0	7

Results

From the clustering analysis, I found two more postal zones, 20814 and 20812, that can satisfy Eric's all expectations. After the analysis, zip codes 20815, 20814 and 20812 are all good choices for Eric's family which have all preferred categories and the house price close to the budget.

ZipCode	City	Avg_House_Val	Grocery Store	Gym / Fitness Center	Italian Restaurant	Mexican Restaurant	Park	Pizza Place	Theater	Total PreCate
20815	CHEVY CHASE	985700	3	2	1	4	6	3	1	20
20814	BETHESDA	759100	1	1	1	2	6	3	1	15
20812	GLEN ECHO	898000	1	1	1	1	2	1	2	9

Discussion and Recommendations

Two-step analysis provides the following results and recommendations.

- Firstly, after the preferred category analysis, I selected the top one zone 20815 as the first choice. It contains most number of categories that Eric likes.

- Secondly, through the K-means clustering analysis, I got another four zones similar to 20815 in both living environment and house price.

I finally recommend postal zones 20815, 20814 and 20812 for Eric and his family to move in.

Conclusion

- This analysis is helpful for not only Eric's family, but also other people new to Montgomery county to start their search for their new home.

- This analysis can only provide the results that satisfy their expectation for the venues, living environment and average housing prices, but did not give more information such as the house for sales.

- The project can help narrow down Eric's choice, and it is open to add more requirements, information to improve it.