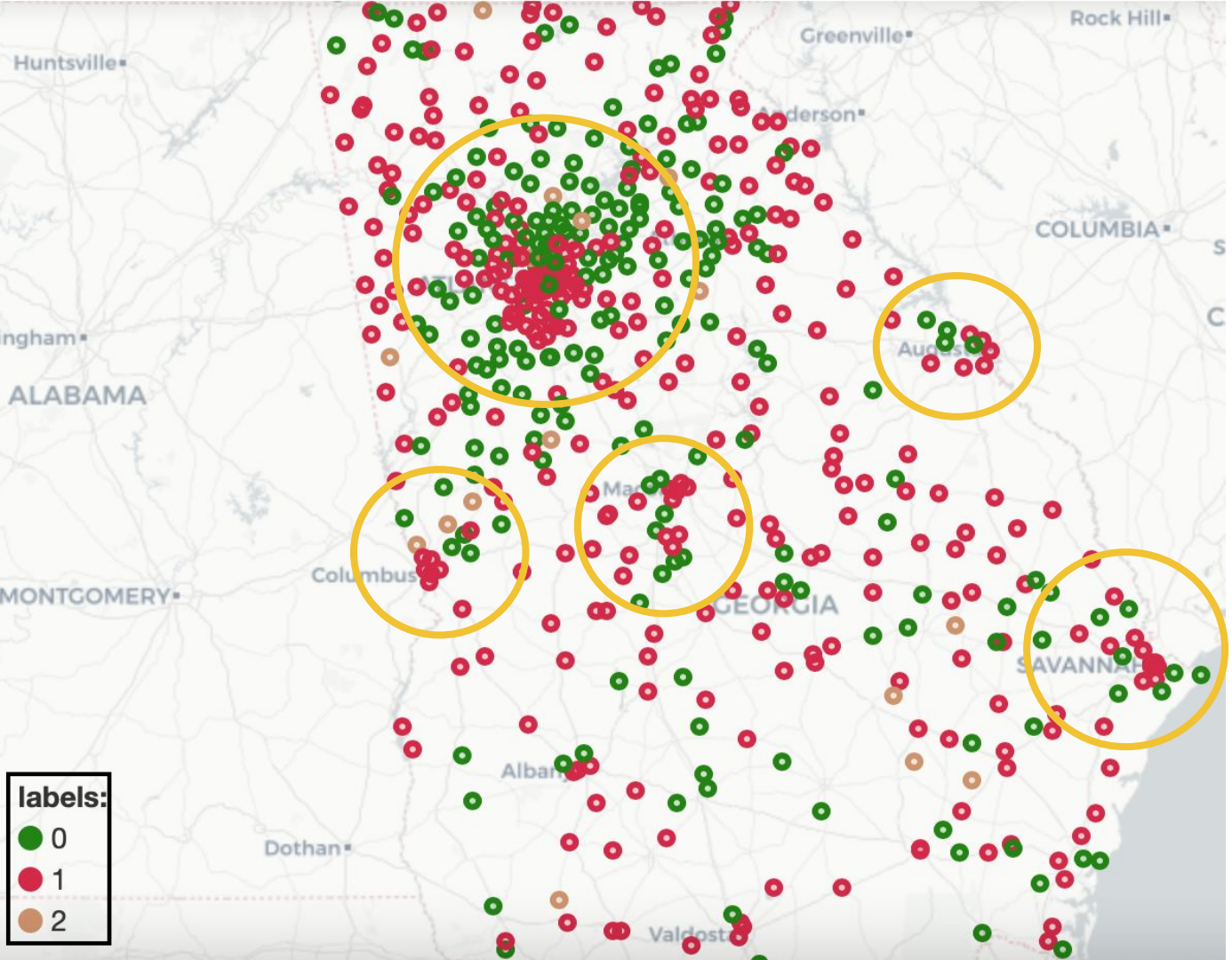
Hackstack Ai Project
Clustering for Real Estate



Atlanta

Atlanta is the capital of Georgia, and with a population of about six million in the metro area, is also the state's largest city

Columbus

A city with good schools, has ranked among the top cities in which to live for people of all ages and top 100 places to raise a family

Macon

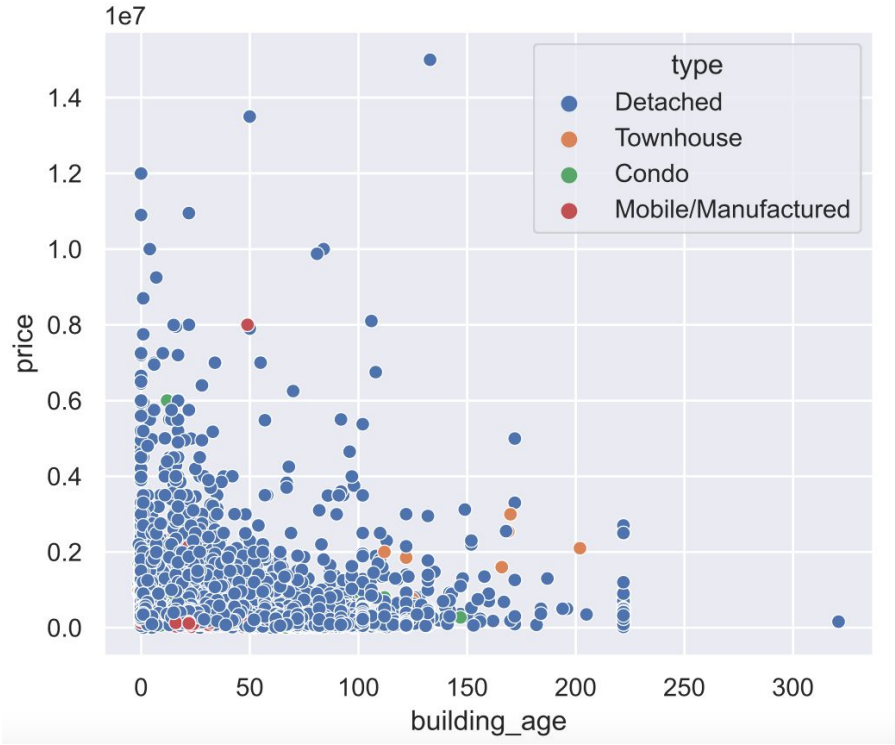
A city in central Georgia, has population of 153,095

Agusta

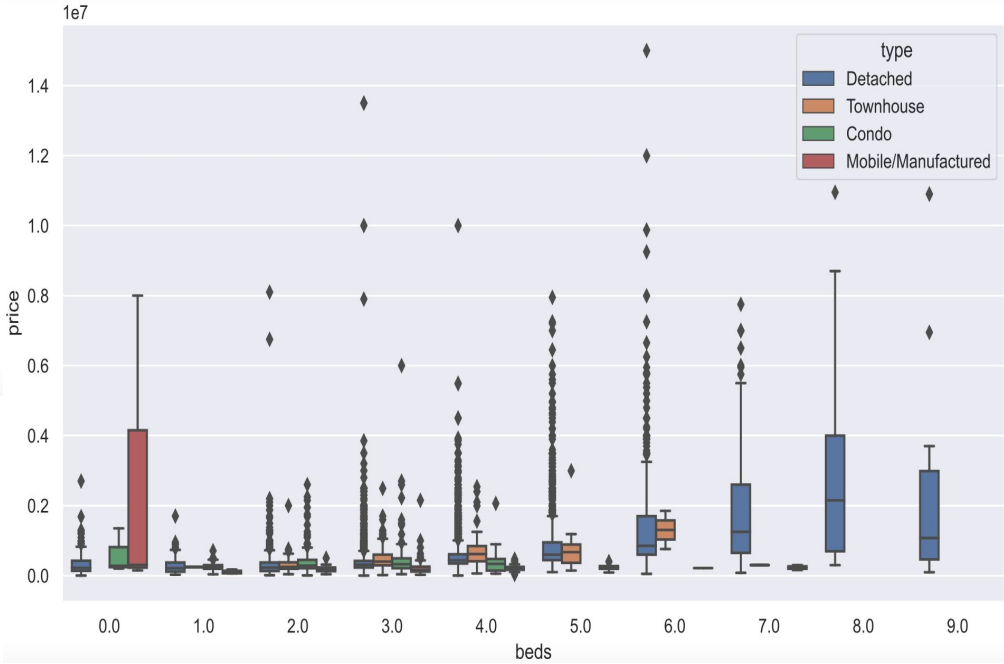
Located on the banks of the Savannah River and the South Carolina border, Augusta is the second-largest city in Georgia

Savanna

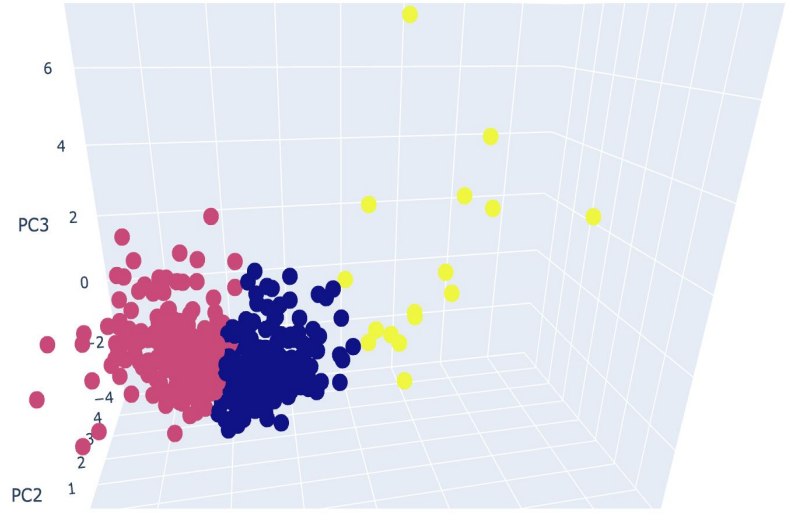
A growing city on the coast.its metropolitan area has grown by 100,000 people over the past decade.

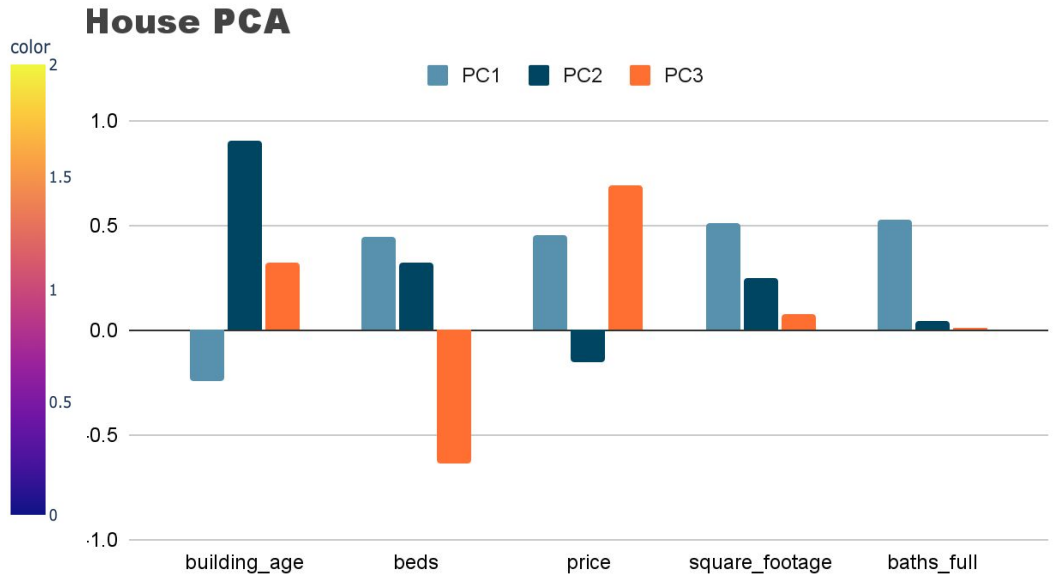


mean	583,955	mean	5.677241e+05
50%	3.990000e+05	50%	3.899000e+05
max	4,525,000	max	1.500000e+07
Beds	71 bedrooms	Beds	9 bedrooms



- The newer the house, the higher the price
- Single-family house/detached house is the majority in GA real estate market.
- The townhouse has a higher price for 3 and 4 bedrooms.





labels	beds	baths_full	square_footage	price	building_age	PC1	PC2	PC3
0	3.914634	2.624390	2416.619512	5.053961e+05	25.497561	1.135523	-0.068214	-0.359351
1	2.996894	1.877329	1764.192547	2.886798e+05	44.459627	-1.007825	0.011238	0.159266
2	4.875000	3.687500	4613.468750	1.614544e+06	24.937500	5.733594	0.647826	1.398957

Cluster 1

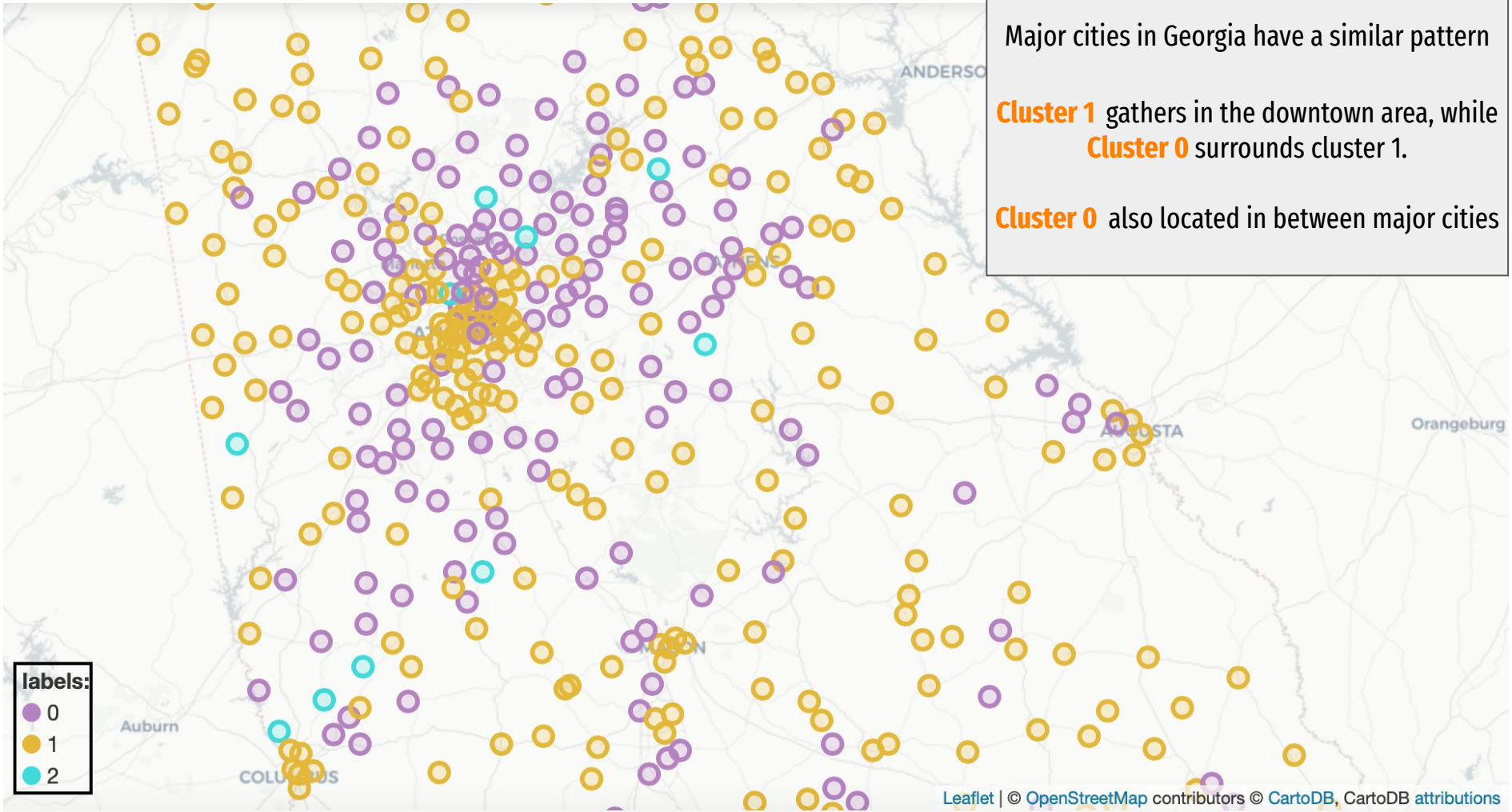
Older house, Smaller size, Lower Price

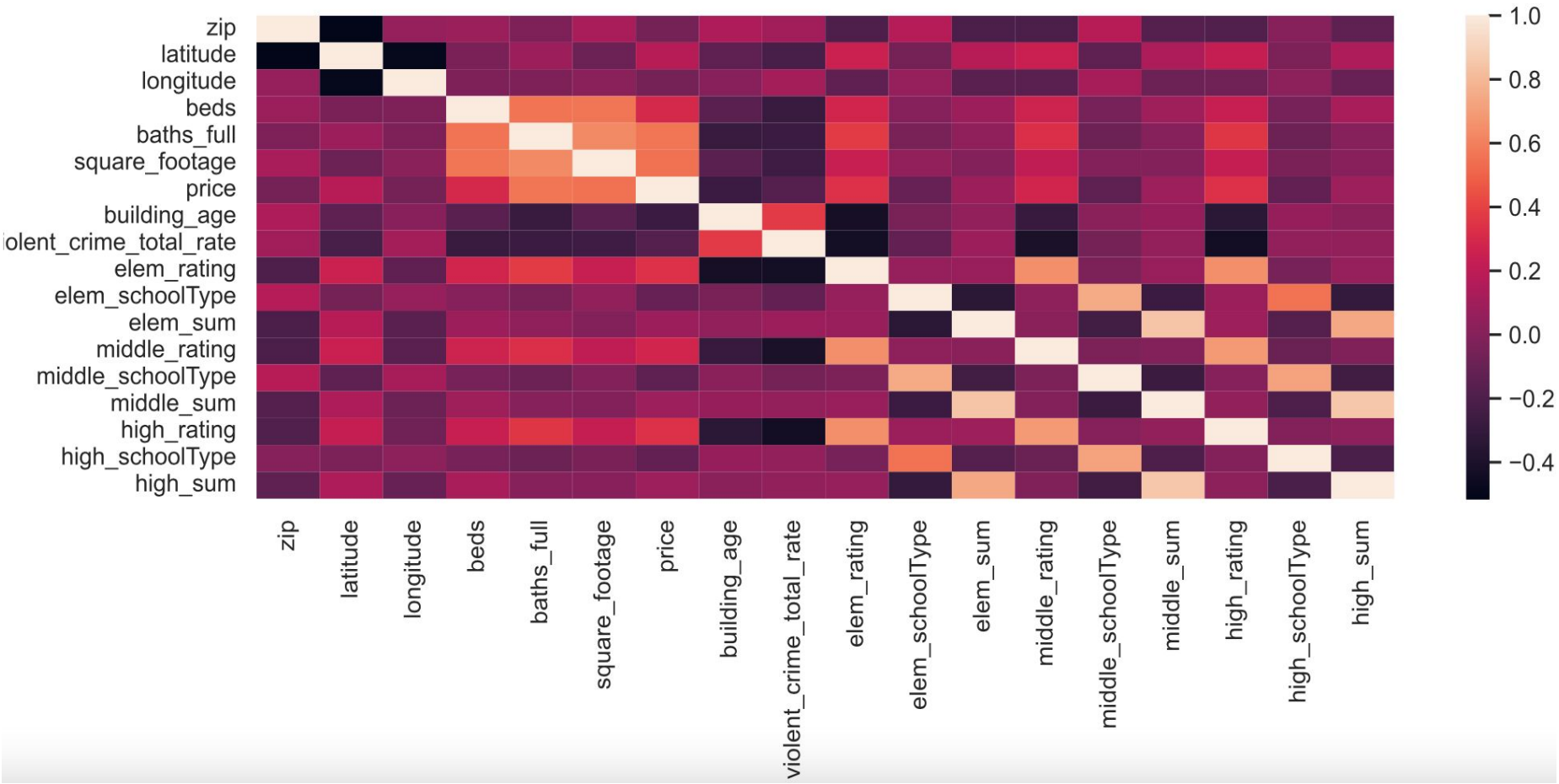
Cluster 0

In between cluster 1 and 2

Cluster 2

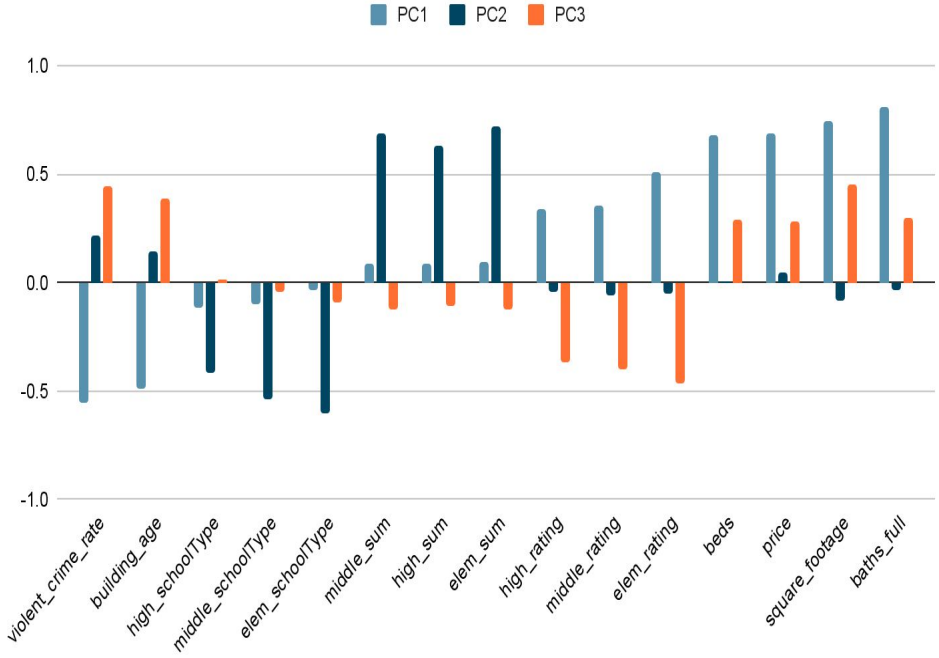
Higher Price, Low supply

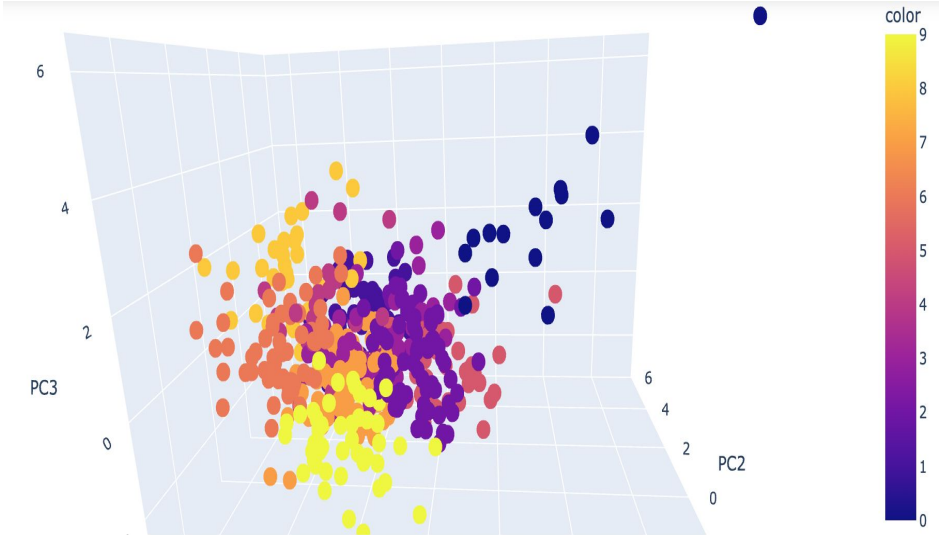

Major cities in Georgia have a similar pattern
Cluster 1 gathers in the downtown area, while Cluster 0 surrounds cluster 1.
Cluster 0 also located in between major cities
labels:
0
1
2
Auburn
Orangeburg
Leaflet | © OpenStreetMap contributors © CartoDB, CartoDB attributions



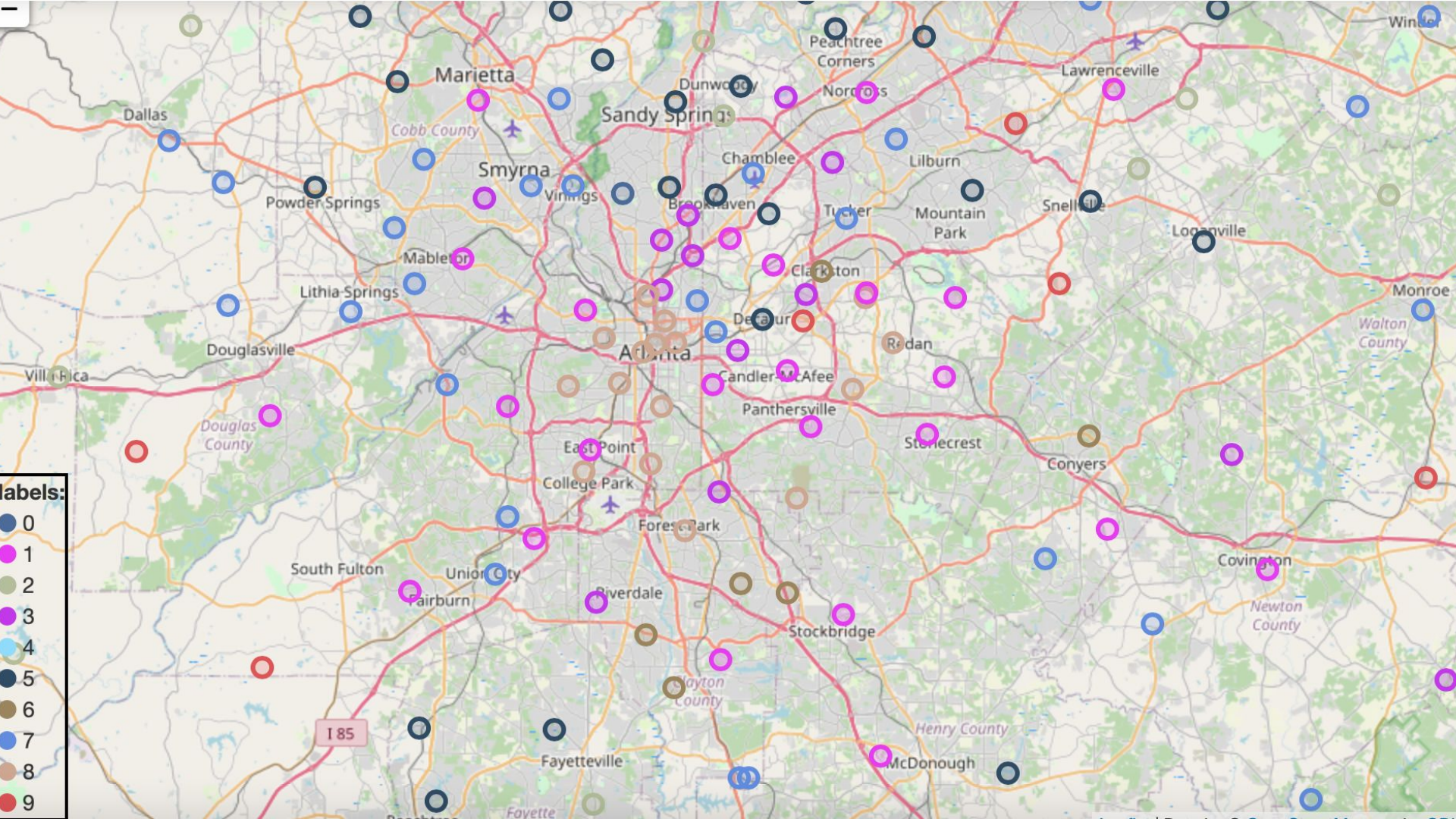
- **Building Age, Crime Rate and School Type have high correlation with House Price.**
- **Higher Sale Price - Newer house, lower crime rate, more private schools.**

House with Neighborhood PCA





- PC1 focus on the house features - baths, SF, price
- PC2 focus on school features
- PC3 on crime rate and building age


Marietta
Dallas
Cobb County
Sandy Springs
Dunwoody
Peachtree Corners
Norcross
Lawrenceville
Chamblee
Lilburn
Smyrna
Vinings
Brookhaven
Tucker
Mountain Park
Snellville
Powder Springs
Mableton
Lithia Springs
Clarkston
Loganville
Monroe
Walton County
Douglasville
Atlanta
Redan
Candler-McAfee
Panthersville
Stonecrest
Conyers
Douglas County
East Point
College Park
Forest Park
South Fulton
Union City
Fairburn
Riverdale
Stockbridge
Covington
Newton County
Clayton County
Henry County
I 85
Fayetteville
McDonough
Fayette
Labels:
0
1
2
3
4
5
6
7
8
9

labels	beds	baths_full	square_footage	price	building_age	violent_crime_total_rate	elem_rating	elem_sum	middle_rating	middle_sum	high_rating	high_sum	PC1	PC2	PC3
0	4.866667	3.666667	4659.700000	1.632247e+06	25.600000	2.727533	6.000000	4.000000	5.833333	3.000000	5.916667	2.714286	5.080081	-0.545831	2.872745
1	3.300000	1.987500	1838.862500	3.178920e+05	45.400000	4.345450	4.305755	14.675000	4.255617	7.850000	3.977477	5.875000	-0.789737	2.911821	0.101017
2	4.067568	2.918919	2576.222973	4.968261e+05	22.554054	2.586068	6.223898	3.818182	6.181373	2.447368	6.040000	1.843750	1.681117	-0.765070	0.198701
3	3.394737	2.315789	2035.763158	4.544221e+05	29.508772	2.918439	4.954286	2.672727	6.107143	1.600000	5.428571	1.437500	0.303169	0.259106	0.144907
4	3.369565	2.021739	2195.260870	2.304847e+05	106.673913	3.539087	4.237179	3.000000	4.750000	2.444444	4.714286	2.222222	-1.233806	-0.068622	1.000785
5	3.892857	2.630952	2457.738095	6.067100e+05	28.059524	1.851019	6.934382	12.833333	6.821250	6.380952	6.747917	4.928571	2.250801	1.900836	-1.085415
6	3.039062	1.781250	1755.851562	2.316598e+05	50.632812	4.103906	3.674881	2.936508	3.713836	1.875000	3.555556	1.703704	-1.889076	-1.174667	0.664934
7	3.119863	1.945205	1884.041096	3.522198e+05	27.428082	2.866786	5.314949	5.466667	5.546154	3.458333	5.101695	2.771429	-0.313022	-0.087004	-0.637919
8	2.871429	1.828571	1620.671429	2.773399e+05	60.257143	9.058314	3.759770	6.838710	3.870513	3.900000	3.591270	3.111111	-2.468077	1.209945	1.362392
9	3.287234	2.127660	1899.563830	3.358945e+05	29.893617	2.705319	6.188830	2.489362	6.240310	1.363636	6.223684	1.205128	0.011353	-2.109848	-0.949561

Cluster0

Luxury houses, some of the houses are mansion

Cluster 8

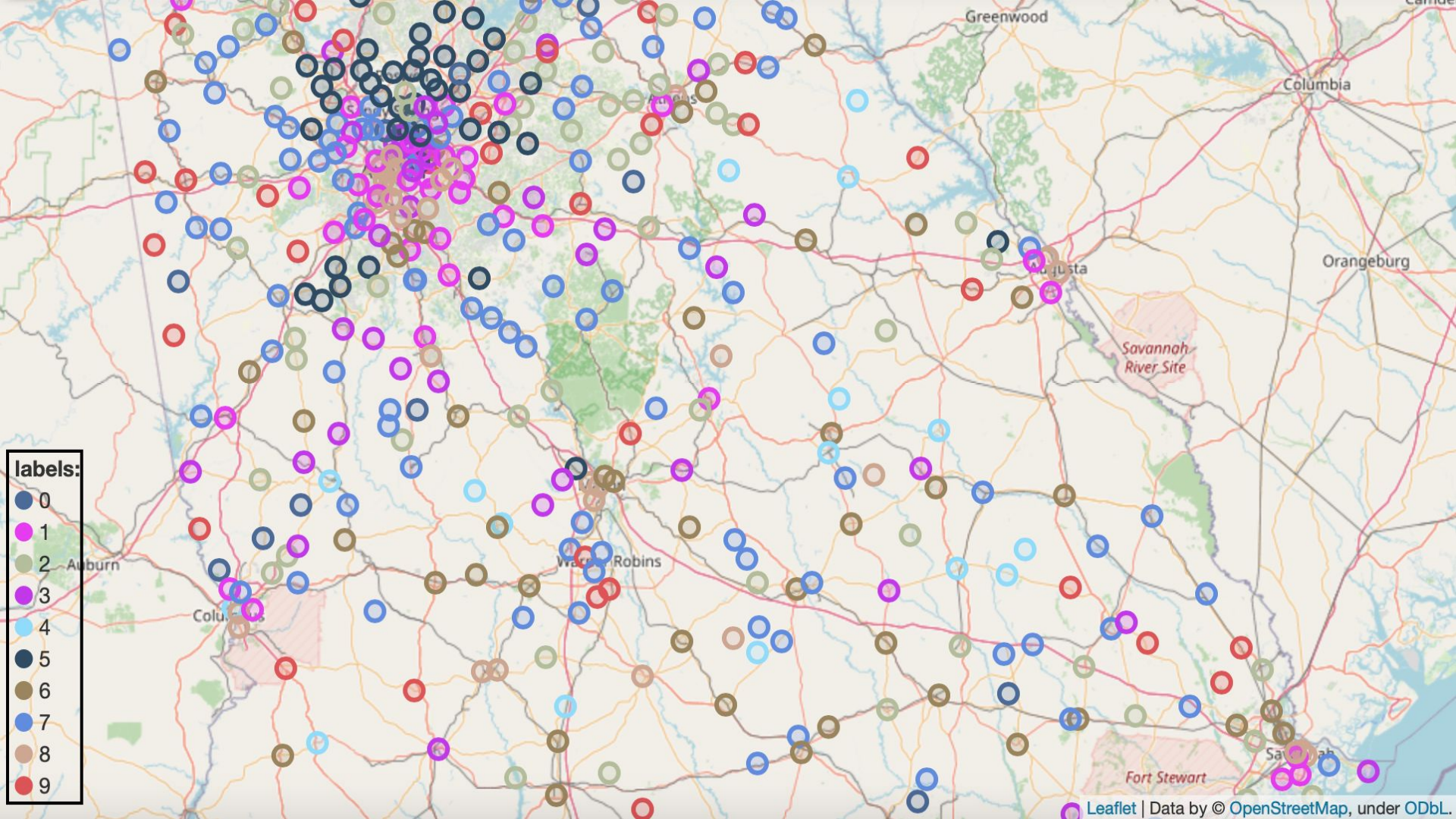
At the downtown area. Lower price, smaller house, higher crime rate, and lower school rating.

Cluster 1, Cluster 3, Cluster 7

Located around the downtown area. Similar as house PCA - better living quality yet not too far from the cities.

Cluster 5

Higher price (second highest), lowest crime rate, at the best school district of Georgia.


labels:
0
1
2
3
4
5
6
7
8
9
Greenwood
Columbia
Orangeburg
Savannah River Site
Auburn
Robins
Fort Stewart
Leaflet | Data by © OpenStreetMap, under ODbL.

Conclusion

- Georgia real estate does have the pattern, especially for those major cities.
- In Georgia, suburban houses have more value than houses in cities.

Future Work

- Take a closer look and investigate at those anomalies.
- Add more environment and neighborhood features.