Impact of a dog park on preoperty values

Value of a dog park

Members of a neighborhood know their HOW fees go towards the maintenence of any park installed in their neighborhood.

- Having said that, those same neighbors also want to maximize the values of their home.
- The job here is to attempt to find a coorelation between the housing market in a given zipcode and the presence of a dog park.

Data Acquisition

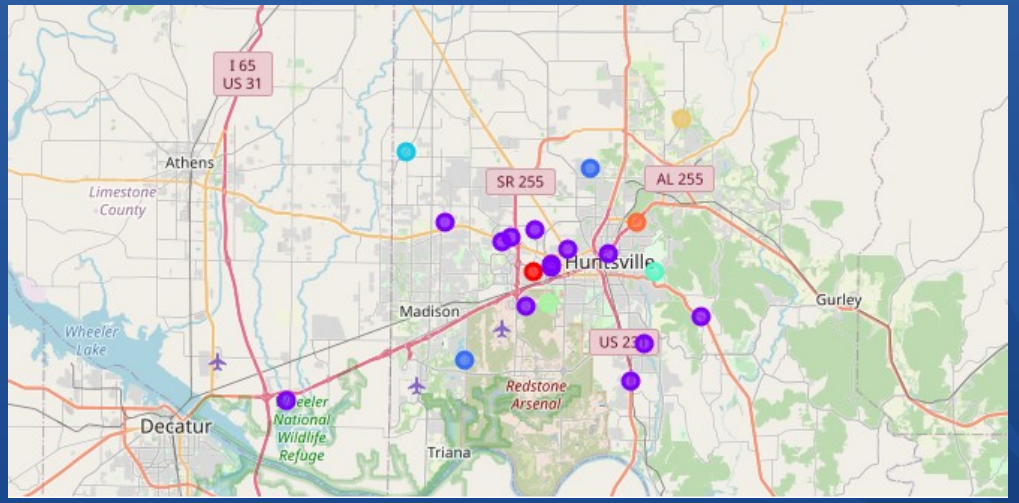
- Data for clustering and finding zipcodes with similar venues is collected using the Foursquare API
- Data for home values in those zip codes is collected using Zillow‘s online database.

Clustering Zip Codes

- Initial work is done to cluster neighborhoods within the subject city of Huntsville, AL.
- To do this, the different zip codes are stripped from wikipedia, geopy is used to find specific latitudes and longitudes and then we use foursquare‘s api to get the nearby venues.
- The goal is to find similar zip codes to compare to one another while at least partially controlling for the different types of venues present.

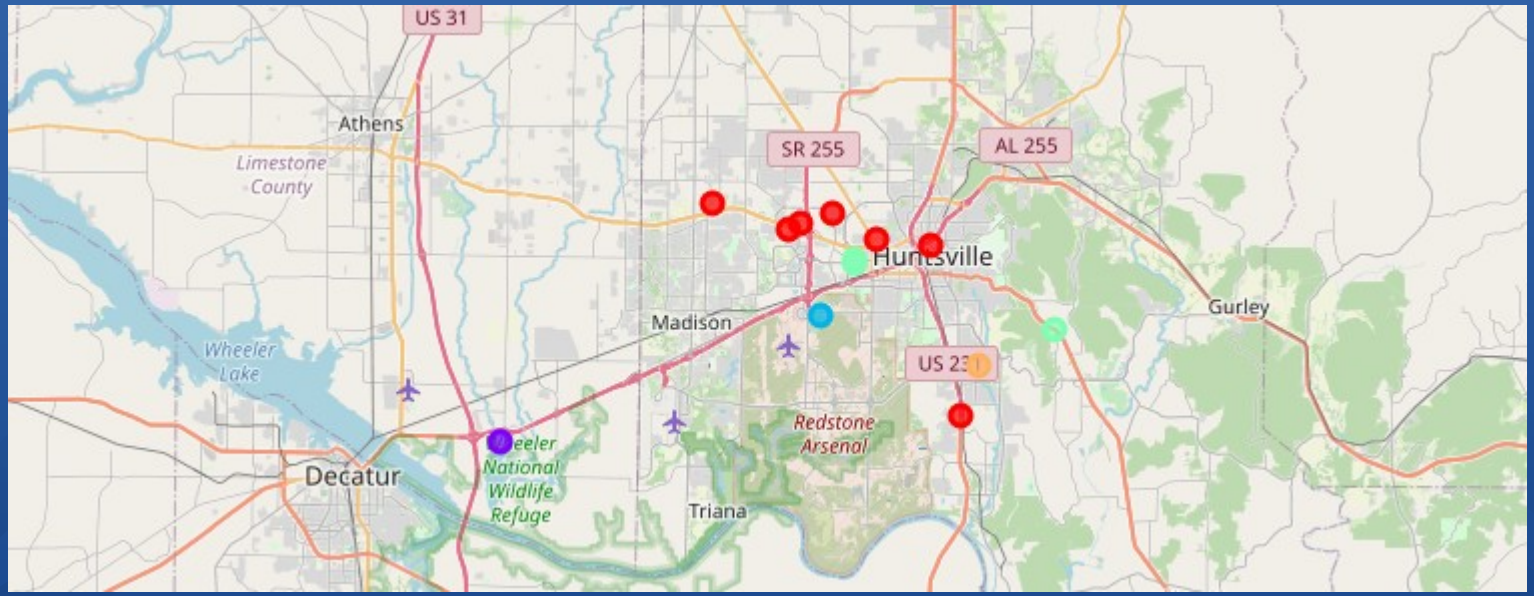
Cluster Data

- I ended up running clustering twice. The first time I ran the cluster, I got a lot of hits in the same cluster so I took those clusters and ran s econd set of clustering algorithms to further isolate truly similar neighborhoods.
- Initial cluster Data ->



Second Round of Clustering

- Our focus will be on the clusters marked in red as they give us a few data points to work with.



Zillow Data

- The Zillow data was downloaded using their online web client. We are comparing the historical data of the Zillow Home Value Index (ZHVI) for a 3 bedroom house for each zip code.
- This data is then extracted based on a zip code that matches one of our zipcodes within our previous cluster.

	RegionID	SizeRank	RegionName	RegionType	StateName	State	City	Metro	CountyName	1996-01-31	...	2019-07-31	2019-08-31	2019-09-30	2019-10-31
3250	73619	3316	35806	Zip	AL	AL	Huntsville	Huntsville	Madison County	NaN	...	209981.0	211106.0	212688.0	213929.0
3539	73616	3609	35803	Zip	AL	AL	Huntsville	Huntsville	Madison County	NaN	...	163247.0	164195.0	165575.0	166718.0
6784	73629	6905	35816	Zip	AL	AL	Huntsville	Huntsville	Madison County	NaN	...	99667.0	100738.0	101480.0	101920.0
7366	73595	7500	35757	Zip	AL	AL	Madison	Huntsville	Madison County	NaN	...	171594.0	172544.0	173905.0	175054.0

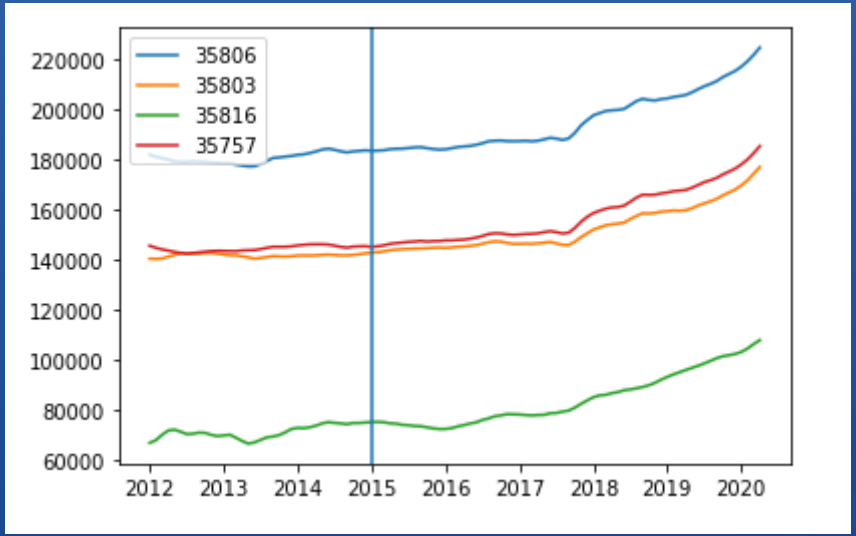
Finding Dog parks

- After the neighborhoods were clusterd, we searched for a list of all dog parks within the huntsville area if they were within one of our districts.
- Exactly one dog park was found in the Zillow data we managed to find for our cluster of zip codes.

	Venue	Venue Category	Venue Zip Code	Venue ID
0	Indian Creek Greenway	Trail	35806	4b896decf964a520713532e3
1	Huntsville Dog Park	Dog Run	35801	4be73b02910020a1b434d514
2	The Dog Spot	Dog Run	35801	4bbb23e1935e9521ebca2890
3	Dog Spot	Dog Run	35758	4c06f8900ed3c9280dea787d
4	Mill Creek Geeenway - Dog Park	Dog Run	35758	4dacdd0a93a0faeb8db7a326

Plotting the Data

- The Indian Creek Greenway park was found to have opened in early 2015 which we marked with a vertical line.
- 35806 is the neighborhood where the Indian Creek greenway dog park is located.



Results

- In the end, it seems that the addition of a new dog park did not have a significant upstream effect on the value of homes. It is interesting to see how consistently Huntsville rises as a city in the various neighborhoods throughout.
- This could be a result of how Zillow calculates ZHVI as it may be more heavily weighted to the cities average home values, more so than the specific areas.

Conclusion

- We couldn‘t find enough data to show that the installation of a dog park would directly lead to an increase in property vvalues for the homes in that zip code.
- Ideally, we would have been able to perform a much more speicfic search on data more focused down to a neighborhood level. Unfortunatly, we did not have a way to extrapolate neighborhood specific latitudes and longitudes.
- In the future, I would ideally have chosen a bigger city to work with as there would be more datasets and zip codes would be smallerwith a higher population density.