

Kaival J Panchal

Coursera Capstone Project

IBM- Professional Certificate

Best Place to Live in NJ – A 2020 Analysis

1. Problem Statement:

A customer approaches a real estate agent, and asks: Where is the best place to live in NJ where there are a good number of things to do, availability to grocery stores and necessities, and has a reasonable price? The real estate agent answers, well I need to search my database and get back to you. Most people don't have time to wait around, so they look for a quick guide to get an idea of how the state of NJ is! Lucky for the customer, this analysis is exactly that!

When thinking of buying a house or a place to live in the expensive state of New Jersey, there are many factors to consider: House prices, things to do, population, density, schools, environment, etc. It depends on how much money the buyer is willing to spend and what she wants in return. This analysis will serve as a guide/reference to whomever would like to purchase a home in the upcoming year in the State of New Jersey. We will be studying the house prices, the number of venues in a 5-mile radius, population, and density by towns and counties in NJ. In addition, we find a relation between the number of venues and county to see if there is any trade off between things to do and the average cost of homes in a specific county. Lastly, we recommend a list of towns and counties based on the amounts in the ranges of \$0-200,000, 200,000-400,000, 400,000- 600,000, 600,000 to 850,000 and 850,000+ by selecting venue counts greater than 75 in a 5-mile radius.

2. Data:

There are four sources of data that was used in this analysis. The first is geolocation data of each town in the state of New Jersey which provides us with the name of the town, zip code, and the Latitude and Longitude points. This will help us with visualization of trends and study of factors using the folium map. The second set of data is the County information from Wikipedia in which Beautiful Soup is used to scrape the data of County Name, its FIPS code, density, population and Area for each county in NJ. The third set of data is given by the State of NJ division of taxation which gives a breakdown of average residential sales prices for each town in NJ for the Year of 2020. Lastly, we utilize the Foursquare API to find venues in a 5-mile radius of each town.

Links:

1. GeoLocation Data: <https://public.opendatasoft.com/explore/dataset/us-zip-code-latitude-and-longitude/table/?refine.state=NJ>
2. County Data: https://en.wikipedia.org/wiki/List_of_counties_in_New_Jersey
3. Residential Sales Data: <https://www.state.nj.us/treasury/taxation/lpt/class2avgsales.shtml>
4. Foursquare API

3. Methodology- Exploratory Data Analysis

3.1. NJ Map and House Price Clusters

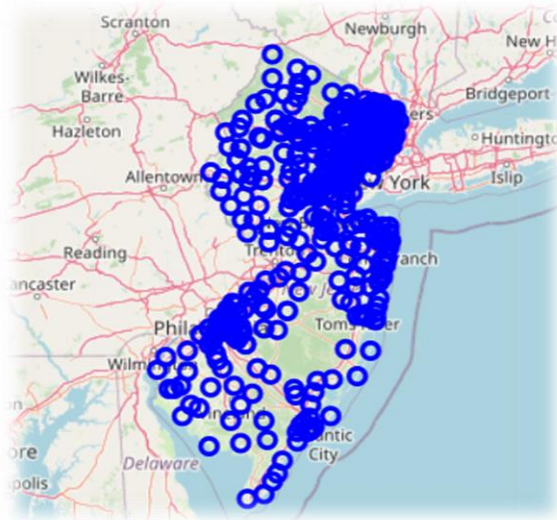
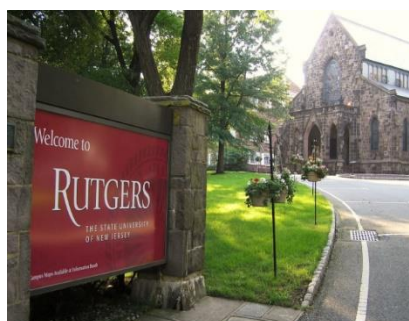
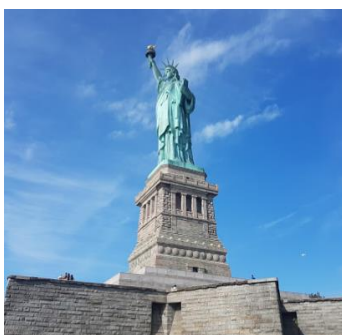


Figure 1. NJ Map with all towns of NJ highlighted with a Blue marker.

Utilizing the geolocation data and folium maps, a folium map was created using the Longitude and Latitude of each town in the state of NJ. With this we got an idea of what we are working with and provided a starting point of our analysis. Based on Figure 1 above, we can already get a sense of the closeness of each town, or the suburban areas. Two clusters that visually stick out are nearby two major cities of the United States: New York City and Philadelphia. We can also see that the North and Central Part of NJ have more towns and hence is a more suburban area than South Jersey. This is mainly due to the industry distribution. First there is NYC which is the center of Arts and Finance in the world, then within NJ where we have United States biggest Pharmaceutical Companies Such as Merck, Novartis, and Bristol Meyers Squibb and other large Manufacturing industries. The big empty spot in the middle of NJ is the Wharton State Forest which is famous for miles of unpaved road for mountain biking, horseback riding, and wildlife observation.

From North to South Jersey is approximately 130 miles of Atlantic coast. A famous city within NJ Jersey City located amidst the heavy blue area seen above, across the Hudson River from Lower Manhattan, is the historic Liberty State Park, where ferryboats embark for nearby Ellis Island, with its historically significant Immigration Museum, and the iconic Statue of Liberty. It is also home to one of the largest educational establishments: Rutgers University which is in New Brunswick but encompasses 3 whole towns! In addition, one of the largest gambling fore front lies at the tail of New Jersey: Atlantic City where you can spend time gambling In NJ's best Casinos and also have family time traveling the boardwalk and beach!



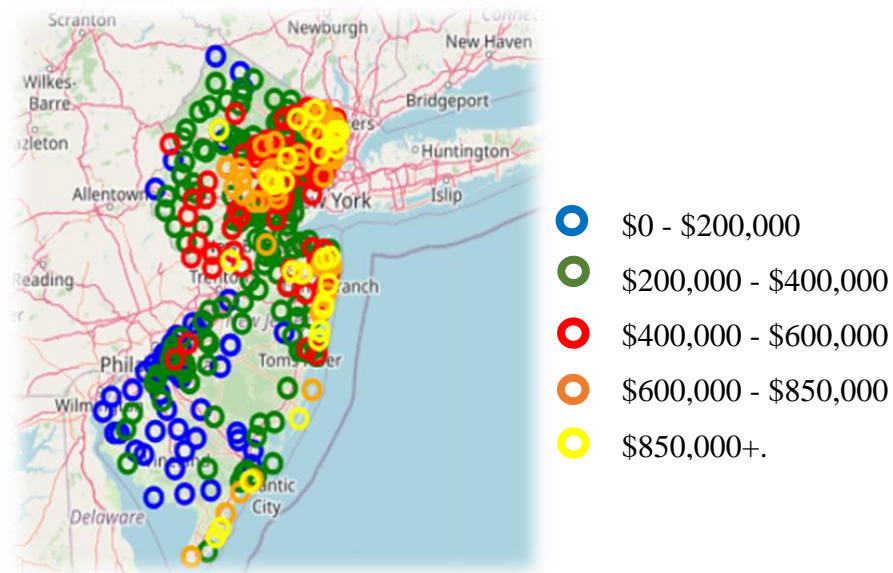


Figure 2. NJ by Average House Price.

As stated before, one of the goals of this analysis is to break down the towns of NJ with respect to their average house prices. Just to reiterate the ranges are: \$0 - \$200,000, \$200,000 - \$400,000, \$400,000 - \$600,000, \$600,000 - \$850,000 and \$850,000+. This was accomplished by combining the geolocation data set with NJ's average house price data for the year of 2020. This is visualized below again using Folium maps but also connected house price with each town's geolocation.

Based on Figure 2, we can see the distribution of house prices are all mixed in North Jersey and the cheapest houses are in South Jersey nearby Camden and the city of Philadelphia in Pennsylvania. One interesting phenomenon that we see is that houses above the cost of \$850,000+ are all distributed along the coastline of New Jersey from North to South. This phenomena in the North is attributed to the proximity to New York City, and in Central and South Jersey to the amazing Shores we have with amazing views. The one county that stands out which holds average house prices of \$400,000 and above is Bergen County, which is known to be one of the most heavily dense and expensive counties in the Tri-State Area. Bergen county is also home to one of the famous Football Stadiums known as Giants Stadium, but now called Metlife Stadium. In addition it is now home to the American Dream mall, one of the biggest malls in NJ. For more information please see: https://en.wikipedia.org/wiki/American_Dream_Meadowlands



3.2. House Price Distribution

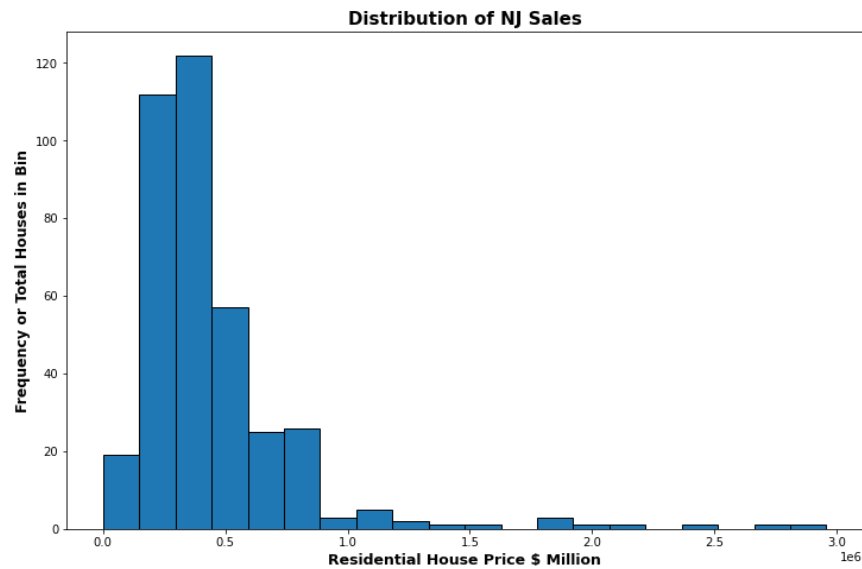


Figure 3. Distribution of NJ Sales

Figure 3 shows the distribution of all NJ house sales over the year of 2020. In this analysis we studied a total of 381 sales which occurred in NJ. Taking all the house prices including the minimum and maximum we found that the Median Cost of House prices in NJ is \$366,460 and a mean of \$443,232 with a standard deviation of \$350,000. The mean is greater than the median because we are considering the low number of houses that are greater than a million dollars, which as we saw before in the map are all located along the coastline of New Jersey. The max house price is \$2,958,280.00 and the minimum house price is \$65,000. Overall, there is a very wide spread of house prices, but it is all dependent on location. The most significant information we obtain from this figure is that most house are distributed within the \$200,000-\$450,000 range. Next, we will zoom in and see the distribution of houses less than a million dollars and compare with the total dispersal.



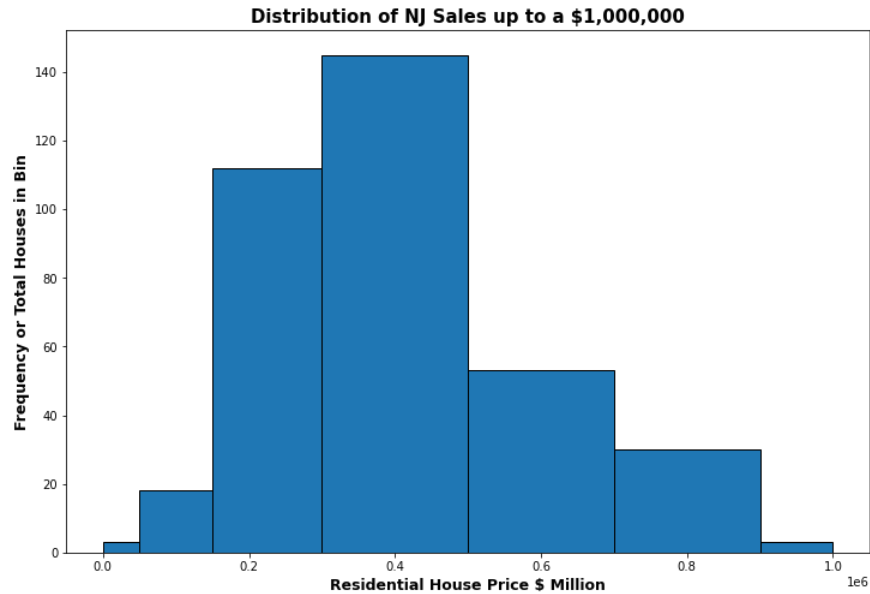


Figure 4. Distribution of NJ Sales less than \$1,000,000

Cleaning up the houses that cost more than a \$1,000,000 the total number of houses reduced to 364 sales and the median cost remained similar our first analysis at \$352,419 however the mean got corrected to \$385,470 with a standard deviation of \$191,463. The minimum house price considered here is \$65,000 and the maximum price is \$973,150. We can say the distribution is normally distributed but is more skewed to the negative side of the central point of \$500,000. The median cost of houses in NJ are usually located in the suburban areas nearby NYC. Typical houses of this cost would look like the figures shown below. Next, we will look at the distribution of houses greater than a million dollars.



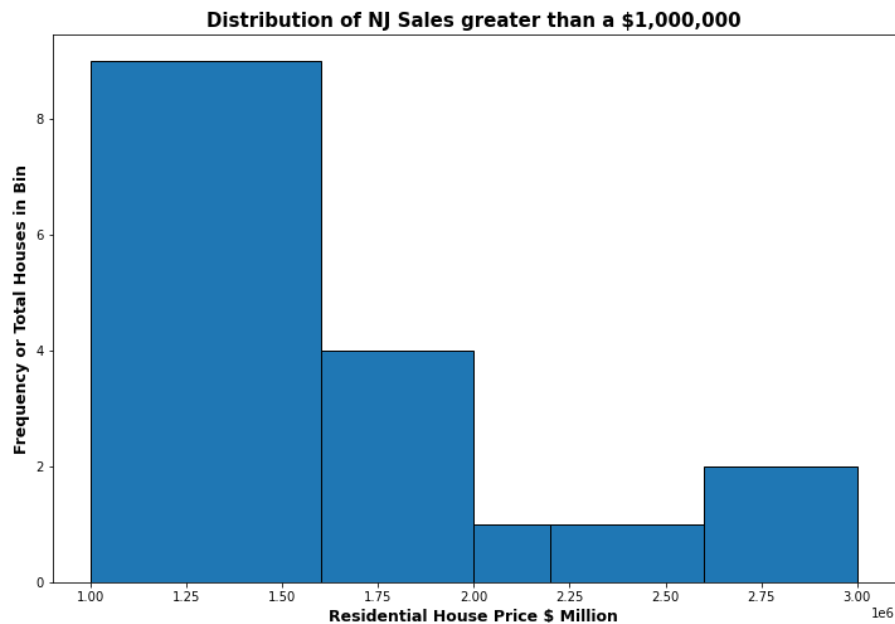


Figure 5. Distribution of NJ House sales greater than \$1,000,000.

Zooming into the right side of our overall distribution, Figure 5 shows all houses that were sold for a price of a million dollars or more in 2020. This accounts for 17 total sales with a median of \$1,523,616 and a mean of \$1,680,015 with a standard deviation of \$619,619. The minimum house price greater than a million dollars was \$1,038,792 and the maximum house price was \$2,958,280. Most houses sold were within the 1-1.6 million dollar range and anything over it is pretty rare. Typical houses at one and three million are shown below.



3.3. Top 30 Highest Sold Prices in NJ and their Locations

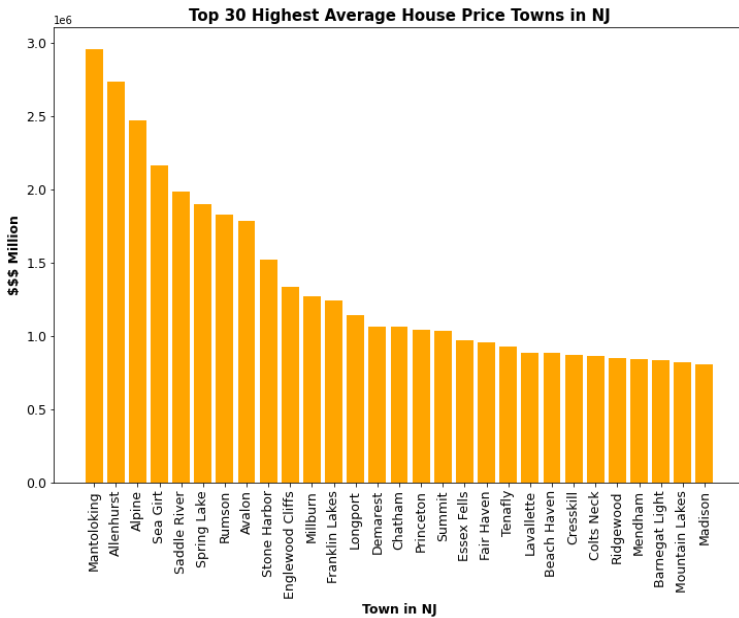


Figure 6. Top 30 Highest House Priced Towns in NJ

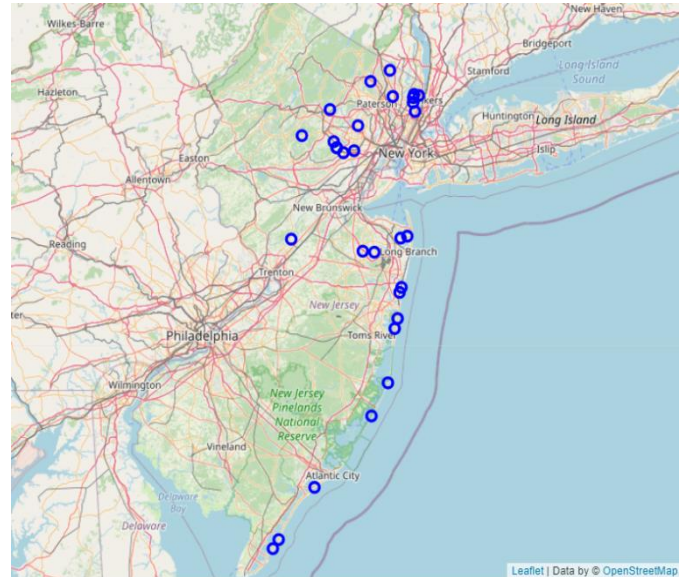
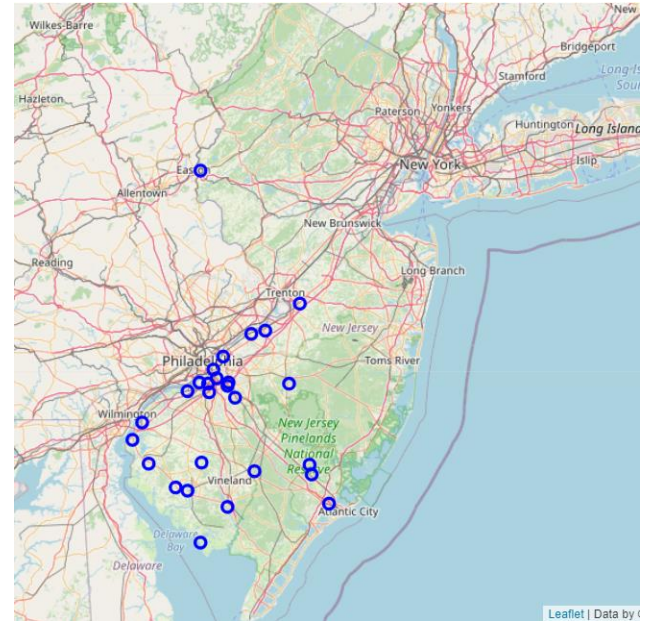
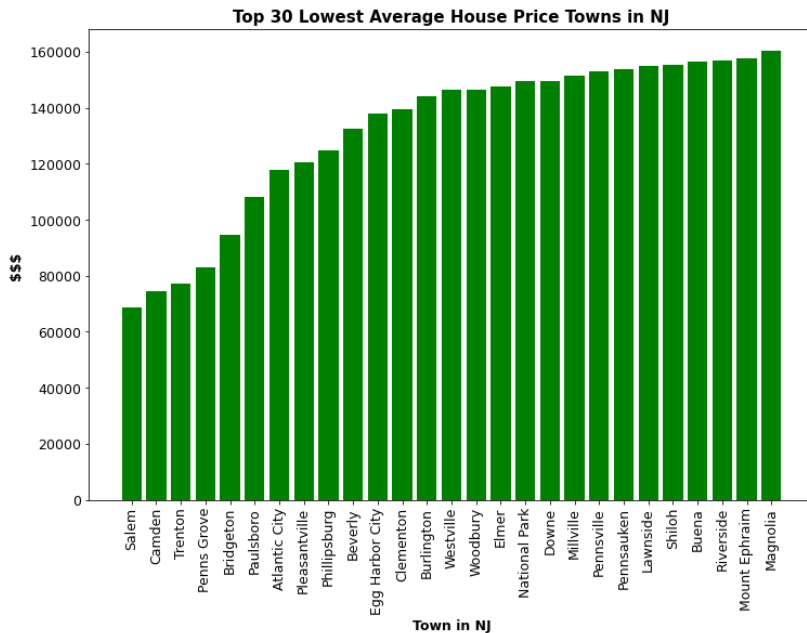


Figure 7. NJ Map of High Priced Towns

As mentioned before, the most of the highest house priced towns are distributed along the coast of New Jersey. The Highest priced house was sold in the Town of Mantoloking, NJ at \$2,958,280. An interesting cluster having high house prices greater than \$900,000 are the Bergen County Towns including Alpine, Saddle River, Englewood Cliffs, Franklin Lakes, Demarest, Tenafly, Cresskill, and Ridgewood. This can be attributed to the proximity to New York City. One interesting thing to note is that the house prices in North Jersey are higher but the average square footage of the house is lower compared to south jersey due to The north side being more urban and having a higher density house distribution in a certain town or area compared to south jersey where there is a lot more space and open area. See figures below to see a typical high-priced house in the north on the left and a typical high-priced house on the right.



3.4.



Most of the cheapest houses in NJ are located in the South west corner of nearby Philadelphia and Closer to the State of Delaware. This can be attributed to the open space and not many things to do except go travel in the Wharton State Forest. The cheapest houses sold at an average of \$65,000, however the house size is projected to be extremely small or are townhouses. The higher-level houses such as the ones greater than 100,000 are all located nearby the city of Philadelphia. Camden is one of the cheapest towns here as well, it has Rutgers 3rd campus there and also stated as one on the most dangerous cities in NJ in terms of Crime rates after Paterson and Newark NJ. Towns with house prices under 100,000 include Salem, Camden, Trenton, Penns Grove, Bridgeton, and Paulsboro NJ. House prices increase as the town gets closer and closer to the East Side nearby the coast of NJ.



3.5. NJ County Analysis

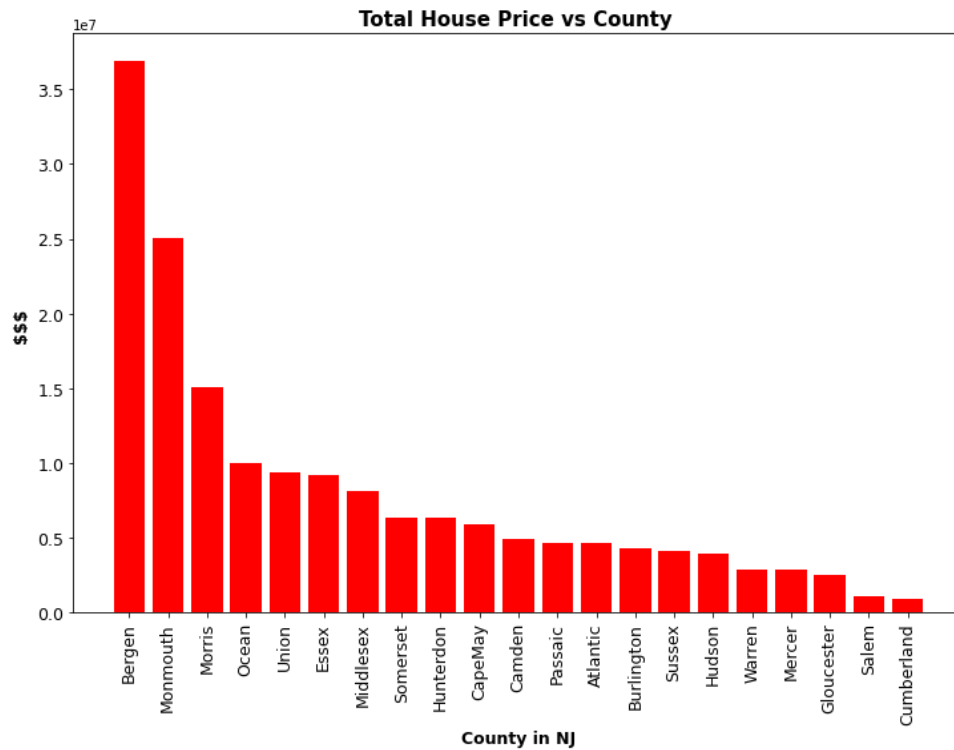


Figure 11. Total House Price vs County

This visualization was prepared by adding all the house price sales for each county up and totaling them. As mentioned before, and reasons also stated before the most expensive County in NJ is Bergen county mainly due to its proximity to NYC. The cheapest county in NJ is Cumberland County which is located in South Jersey. The second most expensive county is Monmouth County due to its historical nature, and the median household income is high at around \$175,000. It can be attributed to the proximity to NJ's pharmaceutical and industry in proximity.

One county that has surprised me in cost is Middlesex county mainly due to the number of things one is able to do in that area. There's Rutgers University one of the biggest state universities in NJ, route 1 has many venues such as Top Golf, Dave and busters, famous Malls, and many diverse restaurants. In addition, it is home to Edison NJ, also known as Little India, which is home to Oak Tree Road which has numerous south Asian clothing, food shopping, and extremely diverse food shops. Middlesex County also holds numerous manufacturing industries like All American Poly, a garbage bag manufacturer. RTape Corp, industrial Tape Manufacturer, Johnson and Johnson, Colgate- Palmolive, etc.

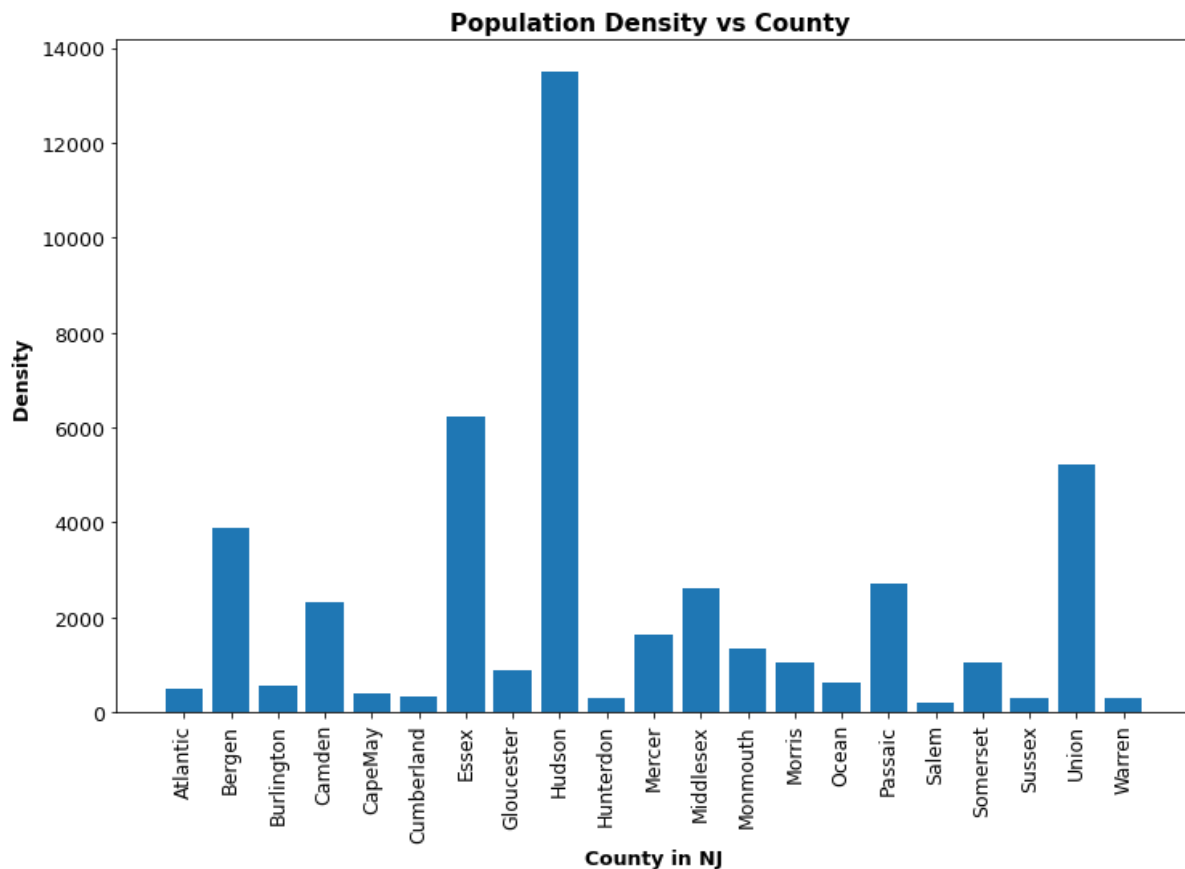
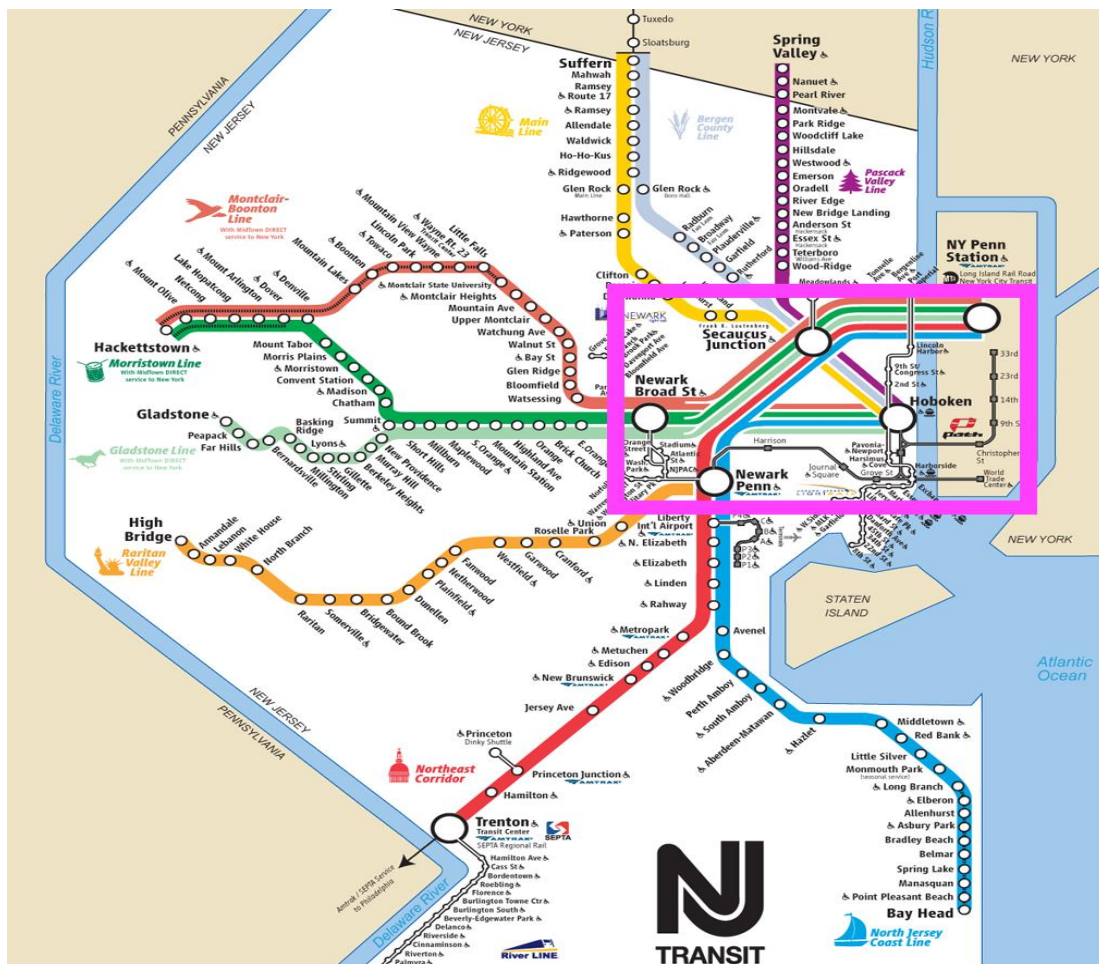
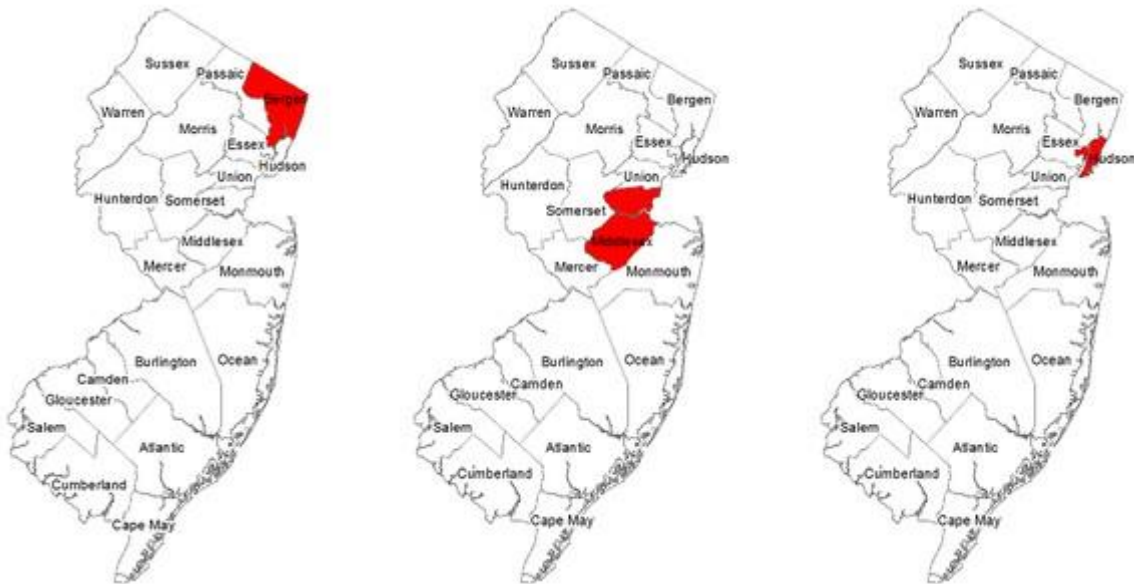


Figure 12. Population Density vs County.

Figure 12 Shows the population density of each county. The density units are in people per squared miles. The highest population density is given by Hudson County, home to towns like Secaucus, Weehawken, Hoboken, and Jersey City lying right at the border of NJ and NYC. Most people live here due to the ease and proximity into NYC for work purposes. There are approximately 13,500 people per squared mile in Hudson County. In terms of cost, Hudson county ranks 16 out of 21, meaning it is one of the cheaper places to live while still having ease and access into the city compared to Bergen County. This explains why the population density is very high compared to Bergen county, another county which is close to New York City. The second most populated county is Essex county which is home to the City of Newark, which is home to multiple universities and industries. In addition, Newark Penn station also has easy train access into NYC. Newark is home to NJ's largest international airport as well, making it the perfect spot for people who do international business to reside. In terms of cost. Essex county ranks 6 out of 21 and is 50% more expensive than Hudson county and has approximately 50% less population as well per squared mile.

For a Suburban Area a population density of less than 4000 people or less is considered comfortable for most people that way you do have people around you, but you do not feel claustrophobic at the same time. Bergen County is on the upper end of this spectrum but the cost of living here is tremendous. A county that continues to surprise me is Middlesex County mainly

due to its features, but now also its population density. It is a big county, but the spread of people is comfortable, and the best part is you get all these amenities for a reasonable price. Another great part about Middlesex county is that it also has numerous train stations for easy access into New York, North and South Jersey.



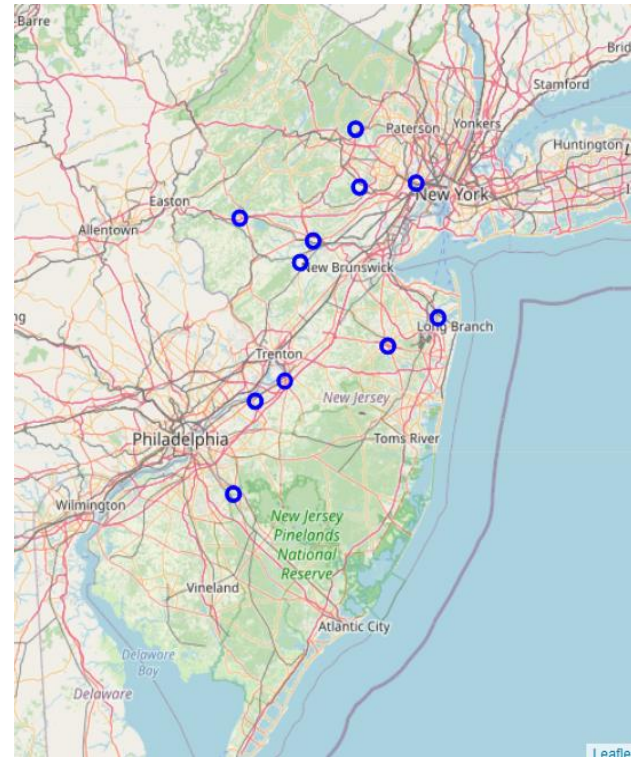
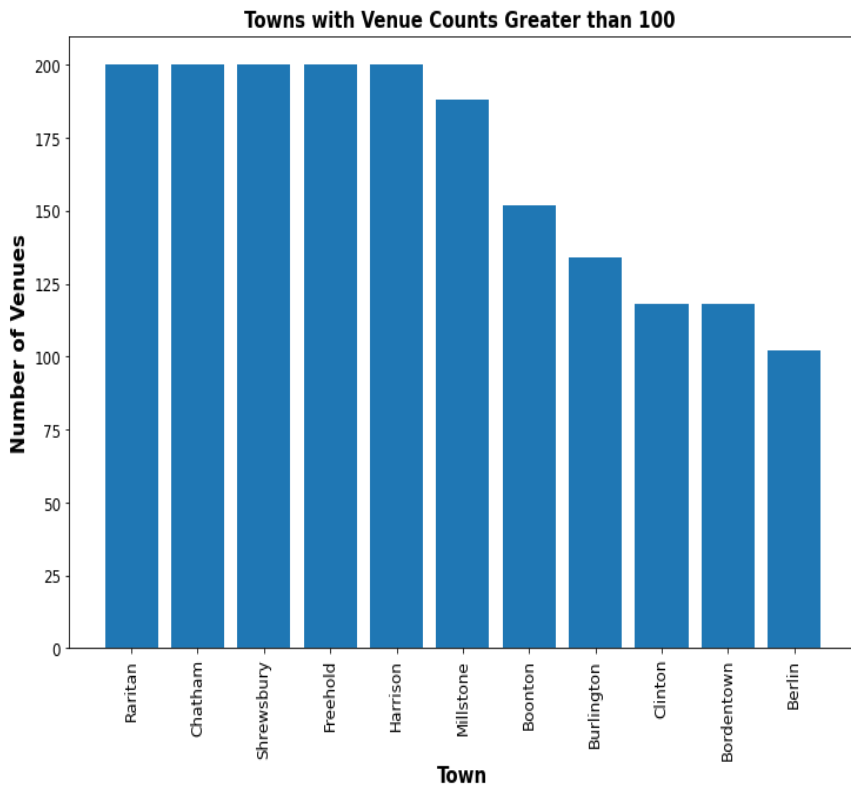
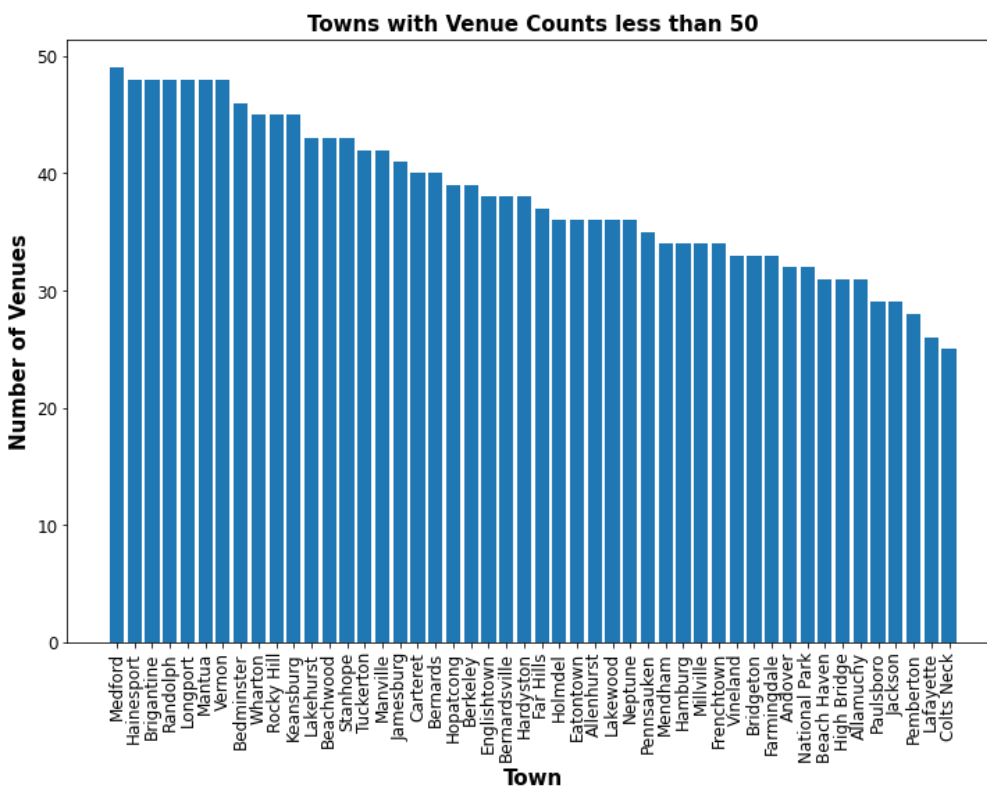
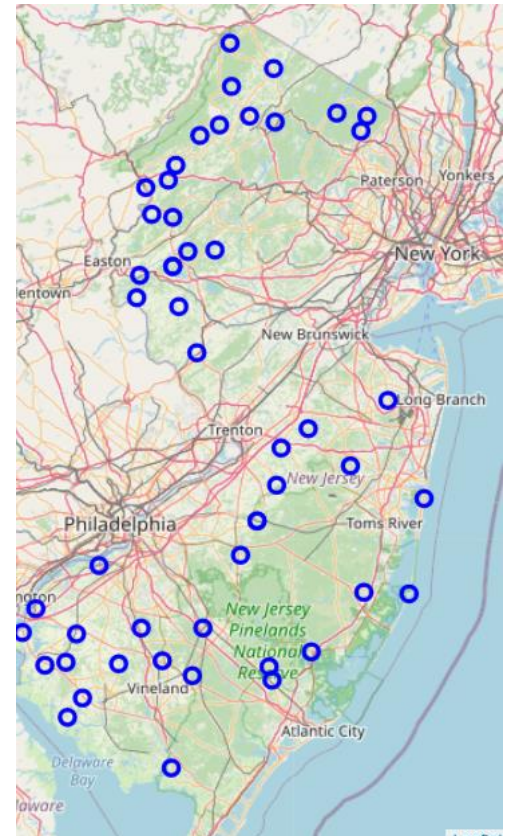
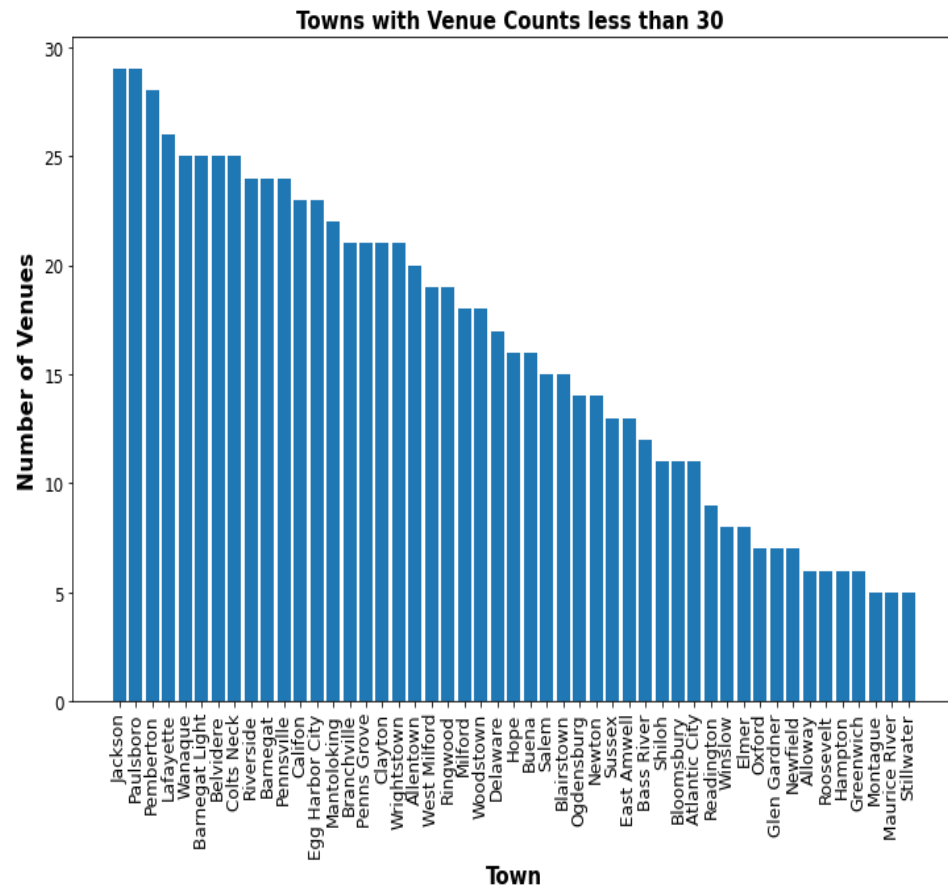


Figure 13 and 14 above shows towns with venue counts greater than a 100. These venues include stores, restaurants, and lure places such as hookah bars, arcades, and bars. The towns listed here are Raritan, Chatham, Shrewsbury, Freehold, and Harrison at 200 venues in a 5 mile radius. Towns between 100 and 200 include Millstone, Boonton, Burlington, Clinton, Bordentown, and Berlin. These towns are in counties very close to common cities like New York, Newark, New Brunswick, Trenton, and Philadelphia.



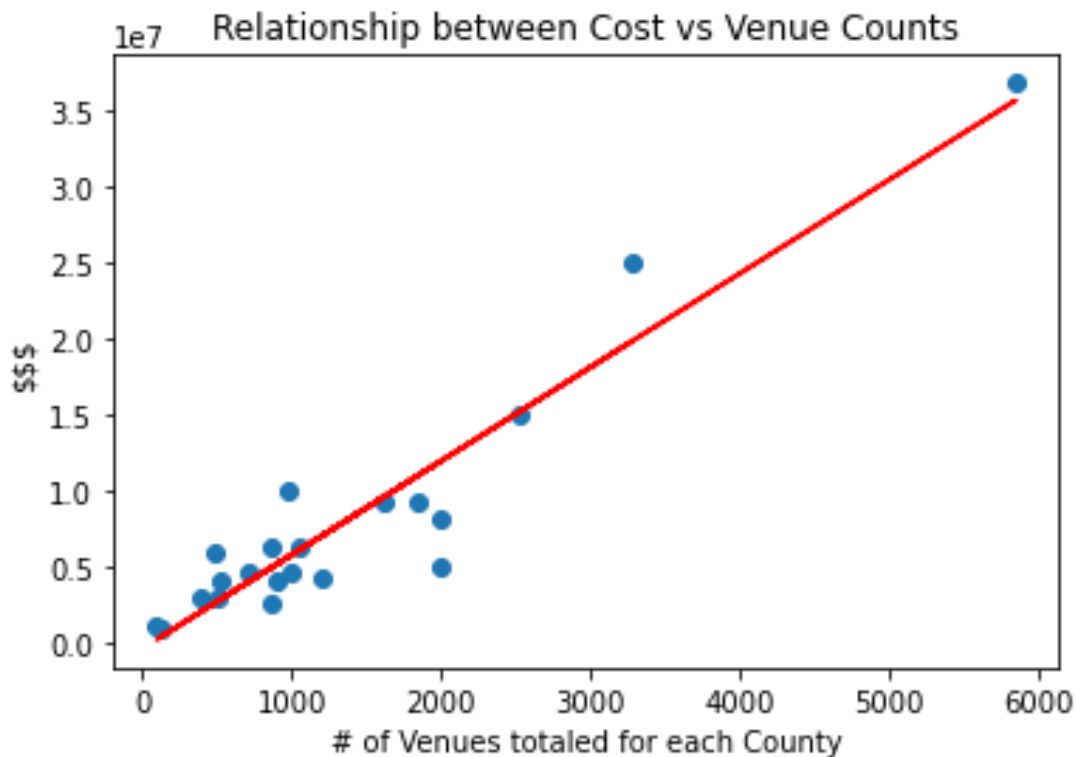
Figures 14 and 15 show the venue counts less than 50. These towns are mainly distributed in the non-suburban areas and have less than 50 but more than 35 venues in a 5 mile radius. Some spots that are greater than 50 include the whole north east area, close to NYC, the whole south west area near Trenton, and south jersey where the forest is. Most of these places are occupied by the industrial and manufacturing companies.



Moving to even lower venue counts, we can see that the towns are distributed away from the largest cities nearby NJ, and most of these towns are distributed towards West NJ nearby Pennsylvania where there are mostly farmlands, and south jersey where there is the forest and mainly wide open spaces.

Venue counts is an important factor to consider when purchasing a home, because you want to know what you have around you so you can socialize with people and have a better overall experience. You can try out different foods, have fun with your friends and family and also have access to diverse stores where you can buy groceries. Personally for me I would suggest purchasing a home in Central jersey because of the things to do in a close vicinity and also if you got 10 to 20 miles out you pretty much cover 3 major cities of New York and Newark, which is a

melting pot of many cultures and experiences. However, we have to study the relationship between venue counts and house prices to see if the price is justified is justified.

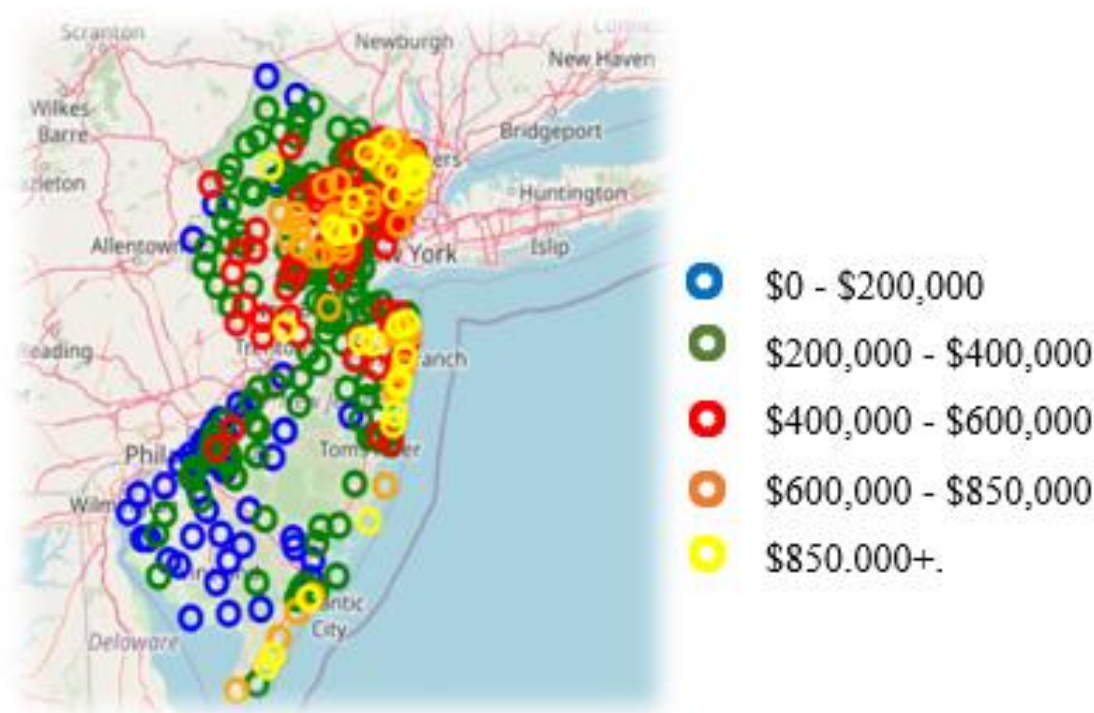


The R^2 is 0.899

The Figure above shows the relationship between the average costs of houses vs the numbers of things to do within a 5 mile radius of that location. We can see here that there is a positive trend, when the number of things to do increases the cost of buying the house there increases as well. The R-squared is .899, so even if we take out the two outliers at approximately 3500 and 6000 there is still an overall positive trend. Doing a visual cluster analysis we can see most of the houses lie within the Venue counts of 1000 and within \$1MM. There are also places which are the similar price and have even more things to do as seen at the 2000 count tick mark.

The Next Section will serve to provide recommendations based on minimizing the house price and maximizing venue counts and will recommend a list of counties that potential would be good satisfying these criteria. The recommendations made will follow the cluster analysis done at the beginning of this analysis.

Results- Recommendations of towns based on price and venue counts



Cluster A: Cluster A consists of Counties which have houses priced between 0 and \$200,000, and maximizes the venue counts. These counties include Warren, Gloucester, Camden, Burlington, Monmouth, and Essex County. The median house price in these counties is \$164,617, a mean of \$156,924 with a standard deviation of \$26,798. The county and list below shows a breakdown of towns and list of towns which satisfy these conditions.

		Avg Sale Price	Venue
County			
Warren	124587.318200	100.000000	
Gloucester	146464.322900	100.000000	
Camden	156910.045600	95.222222	
Burlington	159366.947367	109.666667	
Monmouth	165000.000000	200.000000	
Essex	184455.009700	100.000000	

Cleaned City	Avg Sale Price	Venue
Burlington	144096.7653	134
Maple Shade	168524.4565	100
Palmyra	165479.6203	95
Bellmawr	168120.3667	83
Camden	74312.2152	85
Lawnside	155007.6923	92
Magnolia	160177.2727	100
Mount Ephraim	157788.0328	100
Oaklyn	184802.8846	100
Runnemede	170032.7941	97
Somerdale	163463.2353	100
Stratford	178485.9167	100
Irvington	184455.0097	100
Woodbury	146464.3229	100
Shrewsbury	165000.0000	200
Phillipsburg	124587.3182	100

Cluster B: Cluster B consists of Counties which house houses priced between \$200,000 and \$400,000 and maximizes the venue counts in a 5 mile radius from each town. This makes up majority of NJ and is distributed all throughout NJ. The median house prices in this cluster are \$321,448 and a mean of 314,500 with a 56,412 Standard Deviation. The county and list below, including the next page shows the list of towns which satisfy these conditions.

	Avg Sale Price	Venue
County		
Atlantic	221513.546900	85.500000
Sussex	224871.934800	100.000000
CapeMay	236821.333300	100.000000
Mercer	242469.525000	91.000000
Camden	246265.039278	96.000000
Gloucester	257252.875933	117.666667
Burlington	280310.002850	126.000000
Essex	301801.893920	100.000000
Union	316083.200111	99.111111
Ocean	318474.007800	91.333333
Monmouth	323573.510725	120.750000
Hunterdon	327129.166650	100.000000
Somerset	327318.480175	142.750000
Middlesex	334460.893920	97.466667
Morris	334534.224150	100.000000
Warren	352024.137900	100.000000
Passaic	352487.102143	97.000000
Bergen	370410.664933	100.000000
Hudson	372438.587000	100.000000

Cleaned City	Avg Sale Price	Venue	County
Northfield	215845.0179	90	Atlantic
Somers Point	227182.0759	81	Atlantic
Bogota	375388.2353	100	Bergen
Elmwood Park	354441.8182	100	Bergen
East Rutherford	388250.0000	100	Bergen
Garfield	383247.6340	100	Bergen
Hackensack	271762.2340	100	Bergen
Little Ferry	392099.0000	100	Bergen
Lodi	396525.5814	100	Bergen
Ridgefield Park	396981.4815	100	Bergen
South Hackensack	375000.0000	100	Bergen
Bordentown	297853.3868	118	Burlington
Burlington	262766.6189	134	Burlington
Audubon	212134.2233	100	Camden
Barrington	220302.2000	100	Camden
Berlin	247236.7021	102	Camden
Berlin	200272.3684	102	Camden
Cherry Hill	269816.1894	89	Camden
Collingswood	254706.2529	100	Camden
Haddon Heights	293031.9101	82	Camden
Merchantville	227141.6867	100	Camden
Voorhees	291743.8406	89	Camden
Wildwood	236821.3333	100	CapeMay
Belleville	292866.4629	100	Essex
Bloomfield	372785.5581	100	Essex
East Orange	227646.1525	100	Essex
Newark	215940.6597	100	Essex
Nutley	399770.6364	100	Essex
Franklin	224871.9348	100	Sussex
Franklin	352024.1379	100	Warren
Glassboro	202028.9314	76	Gloucester
Harrison	368918.1579	200	Gloucester
Woodbury Heights	200811.5385	77	Gloucester
Kearny	372438.5870	100	Hudson

Kearny	372438.5870	100	Hudson
Clinton	344883.3333	118	Hunterdon
Flemington	309375.0000	82	Hunterdon
Raritan	334429.7200	200	Somerset
Union	345092.6902	100	Union
Hightstown	242469.5250	91	Mercer
Dunellen	300873.7595	100	Middlesex
East Brunswick	371686.1304	100	Middlesex
Edison	394132.8432	100	Middlesex
Highland Park	353850.0000	100	Middlesex
Middlesex	304125.9762	83	Middlesex
Milltown	353829.4118	100	Middlesex
New Brunswick	318155.3516	100	Middlesex
North Brunswick	318347.4698	100	Middlesex
Old Bridge	358597.3683	96	Middlesex
Perth Amboy	288998.0952	100	Middlesex
Piscataway	351443.8525	100	Middlesex
Sayreville	328016.2481	100	Middlesex
South Plainfield	352514.3173	83	Middlesex
South River	316610.8696	100	Middlesex
Woodbridge	305731.7153	100	Middlesex
Freehold	288321.6667	200	Monmouth
Hazlet	350855.3846	100	Monmouth
Howell	377876.9916	83	Monmouth
Keyport	277240.0000	100	Monmouth
Millstone	309250.0000	188	Somerset
Butler	347620.1724	100	Morris
Riverdale	321448.2759	100	Morris
Brick	340385.4507	94	Ocean
Pine Beach	361748.8788	100	Ocean
Bloomingdale	332538.0941	100	Passaic
Clifton	378157.2926	100	Passaic
Haledon	329518.5714	100	Passaic
Hawthorne	398396.5909	87	Passaic
Little Falls	387043.1000	92	Passaic
Passaic	331496.6095	100	Passaic
Paterson	310259.4565	100	Passaic
Bound Brook	326448.1481	99	Somerset
Somerville	339146.0526	84	Somerset
Elizabeth	339461.9002	100	Union
Hillside	302437.6923	100	Union
Kenilworth	373889.1045	100	Union
Linden	321751.7944	100	Union
Plainfield	298380.3061	92	Union
Rahway	292212.8372	100	Union
Roselle	266035.3500	100	Union
Roselle Park	305487.1261	100	Union

Cluster C: Cluster C consists of Counties which house houses priced between \$200,000 and \$400,000 and maximizes the venue counts in a 5 mile radius from each town. This makes up the second most majority of NJ and is distributed all throughout NJ. There are repeats of counties here, but there are more expensive towns within the same county. The median house prices in this cluster are \$464,766 and a mean of \$473,477 with a \$53,000 Standard Deviation. The county and list below, including the next page shows the list of towns which satisfy these conditions.

22.

	Avg Sale Price	Venue	Cleaned City	Avg Sale Price	Venue
County			Bergenfield	417595.4444	100
Hunterdon	422045.952950	159.000000	Carlstadt	433240.0000	100
Passaic	442781.825750	97.000000	Cliffside Park	479492.7012	100
Somerset	447745.312500	86.000000	Dumont	416009.5584	100
Bergen	462721.006048	97.655172	Emerson	500459.3750	100
Hudson	464103.134067	115.333333	Englewood	528094.0682	91
Middlesex	470645.919700	100.000000	Fairview	411477.2727	100
Union	473617.524917	98.166667	Fort Lee	538054.7354	100
Morris	486824.511614	110.285714	Hillsdale	502376.1463	100
Ocean	487820.065800	88.500000	Leonia	541650.7143	100
Monmouth	488449.655157	126.000000	Lyndhurst	408269.7183	100
Burlington	505679.597400	91.000000	Mahwah	524545.4780	100
Essex	510335.876633	93.500000	Maywood	420146.8493	100
Camden	553345.274300	100.000000	Midland Park	467752.5962	100
			Moonachie	441666.6667	100
			New Milford	433716.3043	100
			North Arlington	409210.5263	100
			Northvale	483787.8378	77
			Oradell	552360.0625	80
			Park Ridge	557903.0562	96
			Ridgefield	495748.4848	100
			Rochelle Park	405247.3469	100
			Rutherford	451802.7062	100
			Saddle Brook	423091.4286	100
			Teaneck	441417.1028	100
			Waldwick	463215.2542	88
			Wallington	402051.2821	100
			Westwood	451511.4582	100

Wood Ridge	417015.0000	100	Bergen
Moorestown	505679.5974	91	Burlington
Springfield	438795.0867	100	Union
Haddonfield	553345.2743	100	Camden
Fairfield	523807.1429	93	Essex
Cedar Grove	502501.9307	100	Essex
Maplewood	588843.7570	100	Essex
Roseland	549585.2941	89	Essex
Verona	471517.6220	100	Essex
West Orange	425759.5131	79	Essex
Harrison	575600.0000	200	Hudson
Bayonne	409113.5057	92	Hudson
Jersey City	508691.1941	100	Hudson
North Bergen	420117.6101	100	Hudson
Secaucus	459904.1976	100	Hudson
Union City	411192.2969	100	Hudson
Clinton	437938.2075	118	Hunterdon
Raritan	406153.6984	200	Hunterdon
Metuchen	470645.9197	100	Middlesex
Asbury Park	412867.0455	100	Monmouth
Belmar	551266.9643	100	Monmouth
Freehold	430096.0922	200	Monmouth
Long Branch	542441.1765	100	Monmouth
Middletown	466762.3784	100	Monmouth
Millstone	587964.3678	188	Monmouth
Red Bank	427749.5614	94	Monmouth
Boonton	408637.1548	152	Morris
Boonton	567868.1818	152	Morris
Chester	464723.4286	96	Morris
Denville	474683.0969	100	Morris
East Hanover	514290.2439	100	Morris
Morris Plains	503695.2935	82	Morris
Pequannock	473874.1818	90	Morris
Island Heights	451652.6316	86	Ocean
Seaside Park	523987.5000	91	Ocean
Totowa	420797.2545	97	Passaic
Wayne	464766.3970	97	Passaic
Bridgewater	447745.3125	86	Somerset
Clark	455331.1000	89	Union
Cranford	497696.7978	100	Union
Fanwood	490089.7073	100	Union
Garwood	428888.1875	100	Union
Scotch Plains	530904.2702	100	Union

Cluster D: Cluster D consists of Counties which house houses priced between \$600,000 and \$850,000 and maximizes the venue counts in a 5 mile radius from each town. There are repeats of counties here, but there are more expensive towns within the same county. The median house prices in this cluster are \$710,239 and a mean of \$712,002 with a \$70,000 Standard Deviation. The county and list below, including the next page shows the list of towns which satisfy these conditions.

	Avg Sale Price	Venue
County		
Monmouth	682523.237967	114.5
Bergen	682811.480410	99.6
Somerset	685200.000000	80.0
CapeMay	687965.152700	92.0
Union	713750.131750	82.0
Essex	748044.735350	100.0
Middlesex	758323.000000	100.0
Morris	777310.255680	116.4
Hudson	807229.013800	100.0

Cleaned City	Avg Sale Price	Venue	County
Allendale	767248.4848	100	Bergen
Edgewater	600201.2892	100	Bergen
Glen Rock	659864.2478	100	Bergen
Larrington Park	675482.3659	100	Bergen
Haworth	702390.3846	100	Bergen
Montvale	644095.3333	100	Bergen
Norwood	626394.5179	97	Bergen
Palisades Park	718089.4737	100	Bergen
Paramus	740036.2069	100	Bergen
Woodcliff Lake	694312.5000	99	Bergen
Cape May	741125.1462	84	CapeMay
Ocean City	634805.1592	100	CapeMay
Glen Ridge	769945.9101	100	Essex
Montclair	726143.5606	100	Essex
Hoboken	807229.0138	100	Hudson
Cranbury	758323.0000	100	Middlesex
Bradley Beach	672466.6667	100	Monmouth
Brielle	632900.0000	100	Monmouth
Little Silver	748982.9787	99	Monmouth
Manasquan	805730.1765	100	Monmouth
Oceanport	605328.5714	88	Monmouth
Shrewsbury	629731.0345	200	Monmouth
Chatham	816300.0000	200	Morris
Chester	693526.4868	96	Morris
Florham Park	743309.6154	100	Morris
Madison	810586.5333	86	Morris
Mountain Lakes	822828.6429	100	Morris
Watchung	685200.0000	80	Somerset
Mountainside	623233.9612	79	Union
Westfield	804266.3023	85	Union

Cluster E

Cluster E: Cluster E consists of Counties which house prices above \$850,000 and maximizes the venue counts in a 5 mile radius from each town. There are repeats of counties here, but there are more expensive towns within the same county. The median house prices in this cluster are \$1,063,323 and a mean of \$1,204,870 with a \$660,000 Standard Deviation. The county and list are shown below.

County	Avg Sale Price	Venue
Union	1.038972e+06	100.000
Mercer	1.039751e+06	82.000
Morris	1.063323e+06	200.000
Essex	1.120802e+06	100.000
Bergen	1.188781e+06	96.375
Monmouth	1.355091e+06	92.000
CapeMay	1.523616e+06	77.000

Cleaned City	Avg Sale Price	Venue	County
Alpine	2.489375e+06	83	Bergen
Cresskill	8.890314e+05	100	Bergen
Demarest	1.064515e+06	97	Bergen
Englewood Cliffs	1.337534e+06	100	Bergen
Ridgewood	8.522472e+05	100	Bergen
Saddle River	1.988882e+06	91	Bergen
Tenafly	9.308590e+05	100	Bergen
Teterboro	0.000000e+00	100	Bergen
Stone Harbor	1.523616e+06	77	CapeMay
Essex Fells	9.731500e+05	100	Essex
Millburn	1.268455e+06	100	Essex
Princeton	1.039751e+06	82	Mercer
Deal	0.000000e+00	78	Monmouth
Sea Girt	2.162083e+06	100	Monmouth
Spring Lake	1.903190e+06	100	Monmouth
Chatham	1.063323e+06	200	Morris
Summit	1.038972e+06	100	Union

Discussion and Conclusion

We have completed the study of the house prices, the number of venues in a 5-mile radius, population, and density by towns and counties in NJ. In addition, we have found a relation between the number of venues and county to see if there is any tradeoff between things to do and the average cost of homes in a specific county. Lastly, we recommended a list of towns and counties based on the amounts in the ranges of \$0-200,000, 200,000-400,000, 400,000- 600,000, 600,000 to 850,000 and 850,000+ by selecting venue counts greater than 75 in a 5-mile radius.

This guide serves as a starting point for people who would like to purchase a house in NJ for a certain price that they would like. I have maximized the venue counts here because I would like to get the best bang for my buck, and based on this analysis, I have concluded that purchasing a house in Middlesex county would be perfect since I would like to spend between 300,000-500,000 and make sure to be in vicinity of many stores and venues with close proximity to cities such as New York City, Trenton, Newark, and Pittsburgh. Further analysis of other factors are up to your discretion, such as schools, neighborhood, distance to commute to work, but I hope this analysis serves a good purpose and a great starting point to serve as a guide to your decision making.

