Choosing the Location of a New Craft Brewery

Gabriel X. Martinez

gmartinez@avemaria.edu

Introduction

In the last few decades, there has been increasing interest in craft brewing. In this project I assume that I am advising a couple who want to start their own brewery. They can live anywhere in the US, but they want to raise their chances of success. They are looking for market research about potential markets. The goal is to find neighborhoods that are underserved by breweries: neighborhoods that, given their market characteristics, should have more breweries than they currently have.

Craft beer attracts people of certain demographic characteristics, both as consumers and as producers. Not every potential or current craft brewer can afford to re-locate purely on the basis of the market potential. So this analysis is interesting to those who have a great deal of freedom to pick a new location across the country. This analysis is interesting to craft brewers who are interested in … analysis – in careful, dispassionate, and rigorous thinking about practical problems – which seems to describe many craft brewers.

Regarding consumers, my hypothesis is that cities with a high proportion of craft-brew aficionados tend to have

- Enough average income to pay prices that are higher than those for mass-produced beer;
- An economy that is sufficiently vibrant, measured, for example, by the number of people with jobs out of the whole population;
- Many people with an interest in trying new things and in learning details and relationships, often correlated with education levels;
- Enough of a craft-beer culture to generate a critical mass of interest.

The last item is complex. A city that is oversaturated with a particular industry will produce too much competition. On the other hand, a sufficiently high concentration of that industry is necessary to ensure a sufficient number of suppliers. While a garage brewer can rely on the internet, a commercial brewer will need to keep costs down to protect its margin, and this is helped by the pre-existence of a large enough base.

Friendly government regulation is also important. The very first entrant might raise suspicion – perhaps only the well-connected will get in. As a few breweries are established, counties will develop adequate policies. If too many exist, a backlash may develop.

On the demand side, there is a "spillover" effect as an industry grows. The "Got Milk" campaign is a classic example: instead of promoting the products of a particular supplier, the industry association promoted milk in general, betting that if enough people drank enough milk, then drinking it would be socially

acceptable and bring more people in. Likewise, the first craft-brew place in a city might struggle to generate interest and be only patronized by people perceived by others as "geeks." But if there is a critical mass of breweries, it will be acceptable to go beyond the mass-market beer.

As an example of the latter point, note that Budweiser first reacted against the craft beer revolution: "brewed for drinking, not dissecting," said a 2015 commercial, to which craft beers responded in the same style. By now, Anheuser-Busch has caved, producing a line of "distinctive" beers, such as the Discovery Reserve, "its taste inspired by an archival recipe from that time in history. Featuring toasted Voyager barley malt, the lager's reddish color is reminiscent of the next frontier - Mars."

The goal, then, is to find Metropolitan Statistical Areas (a) that have a sufficient critical mass of people who are likely to have been exposed to craft beer and who have enough income to afford it, which is (b) demonstrated by the existence of some craft brewing, but (c) not enough to overwhelm a new entrant.

Data

where you describe the data that will be used to solve the problem and the source of the data.

An MSA is "a core area containing a substantial population nucleus, together with adjacent communities having a high degree of economic and social integration with that core. … Under the standards, the county (or counties) in which at least 50 percent of the population resides within urban areas of 10,000 or more population, or that contain at least 5,000 people residing within a single urban area of 10,000 or more population, is identified as a "central county" (counties). Additional "outlying counties" are included in the CBSA if they meet specified requirements of commuting to or from the central counties."

Because the definition insists on economic integration, MSAs vary widely in size. The data are drawn from the Census Bureau and include, for each Metropolitan Statistical Area (MSA) of the United States.

- Brewery Density. Two alternative measures of the Dependent Variable are used. From FourSquare, CategoryID = '50327c8591d4c4b30a586d5d',
    - breweries per 100,000 people. Divide number of breweries by the population of the MSA and multiply by 100,000.
    - breweries per 100 square miles. Divide number of breweries by the land area of the MSA and multiply by 100.

Explanatory Variables:

- Bar Density. From FourSquare, CategoryID = '4bf58dd8d48988d116941735',
    - bars per 100,000 people. Divide number of bars by the population of the MSA and multiply by 100,000.
    - bars per 100 square miles. Divide number of bars by the land area of the MSA and multiply by 100.
- Personal Income per person (Bureau of the Census)[1]
    - A better measure than GDP per person (which includes corporate retained earnings, etc.), this quantity measures the average income received by the residents of that MSA.

[1] https://apps.bea.gov/iTable/iTable.cfm?reqid=70&step=30&isuri=1&year_end=-1&acrdn=8&classification=non-industry&state=5&yearbegin=-

- Jobs-Population ratio (Bureau of the Census)
    - Calculated by dividing the number of jobs in the MSA by the population of the MSA. While (a) not all jobs are filled and (b) some people have more than one job, so this is not the same as an employment-population ratio. However, it is a good measure of the economic vitality of an areas.
    - Also notice that some residents may work outside of the MSA while some jobs are filled by people who commute into it. However, the definition of MSA (to include outlying commuting areas) reduces this bias.
- Proportion of college educated and graduate-school educated (Bureau of the Census via SSTI)[2]
    - Share of the population 25 years old or older with university degrees: associates, bachelors, graduate/professional, and "at least bachelors."
- Population Density and Housing Density (Bureau of the Census)[3]
    - Craft breweries often become a neighborhood business. Moreover, the nature of their product is such that driving long distances from or two them is unadvisable. Therefore, it makes sense to assume that higher population densities are correlated with a higher density of breweries.
- Percent Female and Percent of the Population between 25 and 44 years old (Bureau of the Census)
    - Craft brewing seems to be associated with a younger, "millennial" generation. There may also be an association with sex ratios.

Example Data

Here follow the first five observations:

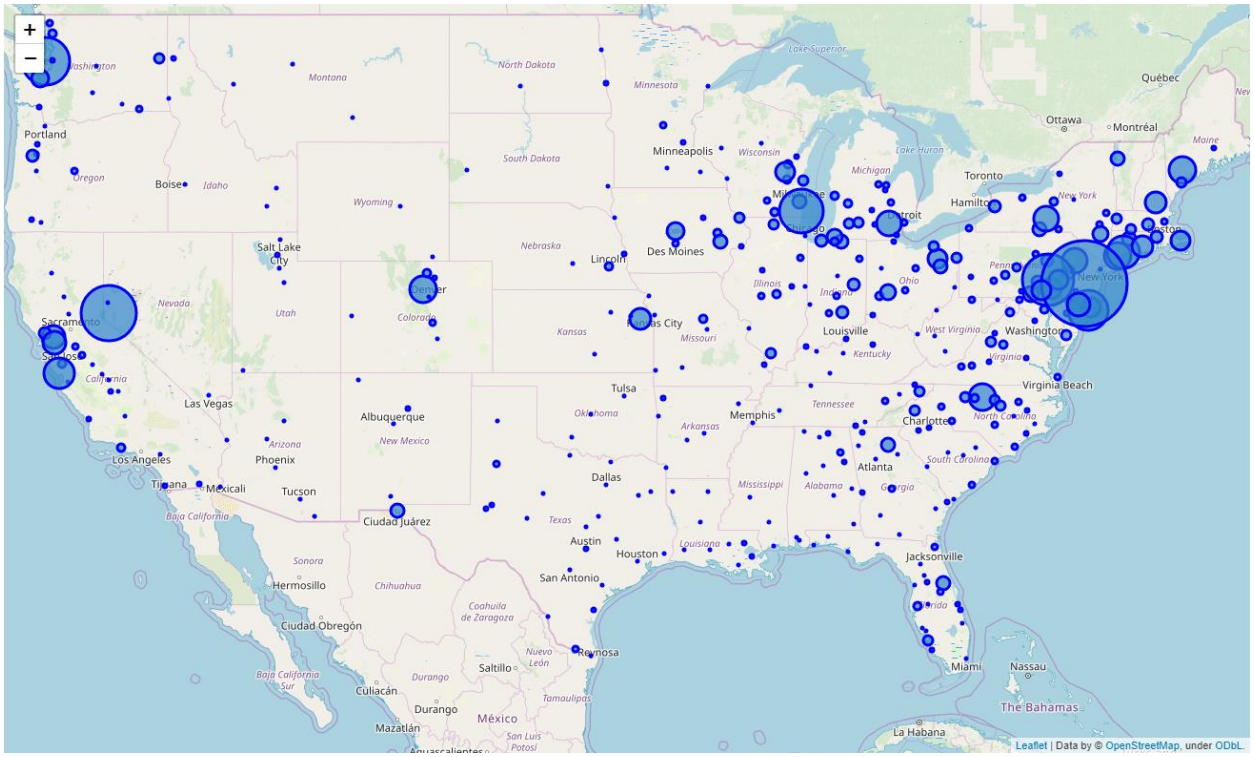
GeoFips	**10180**	**10420**	**10500**	**10540**	**10580**
GeoName (Metropolitan Statistical Area)	Abilene, TX	Akron, OH	Albany, GA	Albany, OR	Albany-Sch.-Troy, NY
Latitude	32.45	41.08	31.58	44.64	42.65
Longitude	-99.73	-81.52	-84.16	-123.11	-73.76
BrewDensity	5.87	5.26	1.32	12.8	3.27
BrewDensityLand	0.36	4.11	0.1	0.7	1.03
BarsDensity	18.8	7.11	15.85	30.39	5.64
BarsDensityLand	1.17	5.55	1.24	1.66	1.78
Per capita personal income	41326	47511	36805	40380	55848
Total employment (number of jobs)	103969	429591	83727	60760	569738
Jobs-Population Ratio	0.61	0.61	0.55	0.49	0.64
% 25+ pop with Bachelor's or more	21.5	30.8	19.5	18.6	36.2
PopDensity	60.23	781.22	81.4	50.95	309.69
HousingDensity	25.41	347.26	34.18	21.32	139.89
PercentFemale	49.37	51.41	52.63	50.64	51.09
Percent25_44	25.87	23.98	25.23	24.8	24.55

1&unit_of_measure=levels&major_area=5&area=xx&year=2017&tableid=82&category=712&area_type=4&statistic=100,110,240

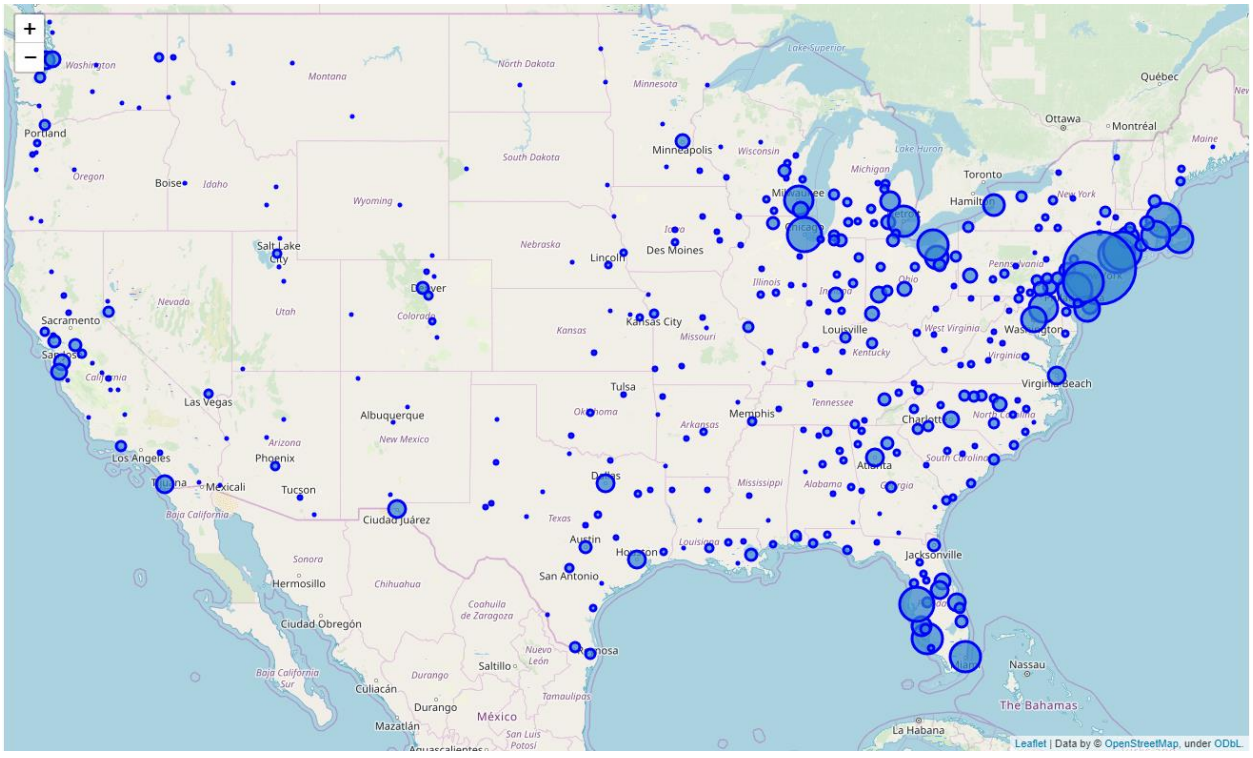
[2] https://ssti.org/sites/default/files/Educational%20Attainment%20by%20Metro.xlsx

[3] https://www2.census.gov/geo/docs/maps-data/data/rel/ua_cbsa_rel_10.txt

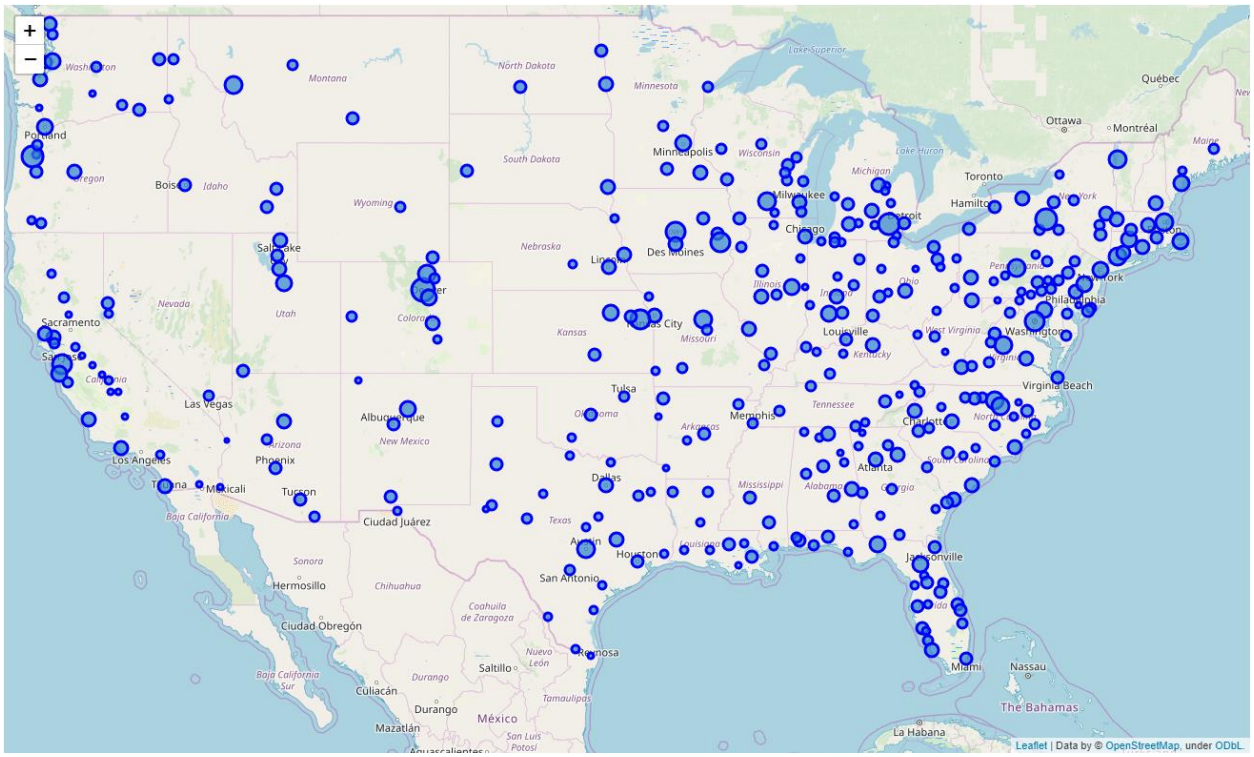
Geographical Distribution of Brewery Density (Number of Breweries per 100 square miles in MSA)



Geographical Distribution of Housing Density (Number of Houses per 100 square miles in MSA)



Geographical Distribution of Education (proportion of 25+ year olds with more than a BA)



Descriptive Statistics

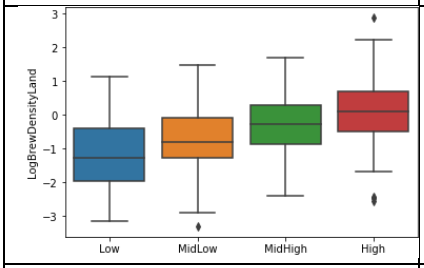
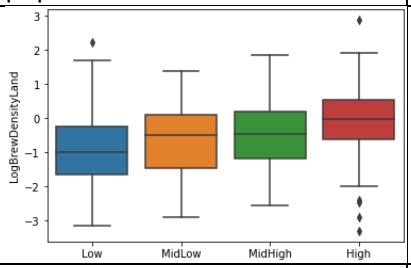
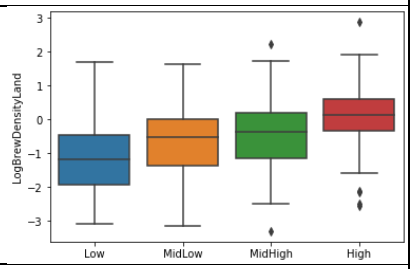
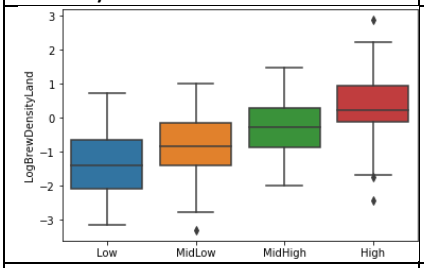
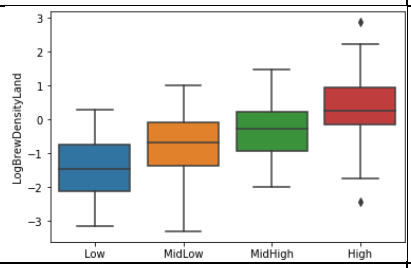
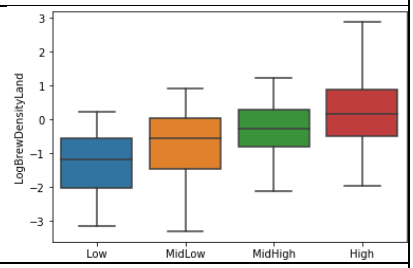
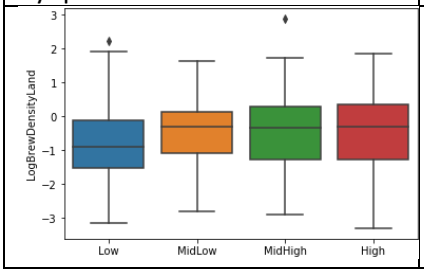
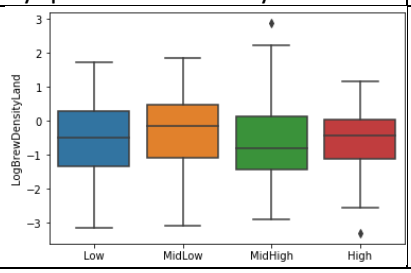
	count	mean	std	min	25%	50%	75%	max
BrewDensity	360	4.25	3.72	0.12	1.60	3.06	6.00	26.31
BrewDensityLand	360	1.01	1.43	0.03	0.26	0.64	1.19	17.81
Per capita personal income	360	45636.98	9178.26	25617.00	39896.25	43642.50	49120.50	110104.00
Jobs-Population Ratio	360	0.58	0.09	0.34	0.52	0.58	0.63	0.87
% 25+ pop with Bachelor's or more	360	27.83	8.23	12.30	21.50	26.95	32.85	60.40
PopDensity	360	270.52	287.41	7.22	113.49	187.09	333.51	2825.99
HousingDensity	360	117.15	121.52	3.40	47.46	81.16	147.36	1125.75
BarsDensity	360	16.93	11.85	0.25	7.69	15.14	23.56	54.93
BarsDensityLand	360	3.34	2.97	0.18	1.29	2.48	4.45	22.27
PercentFemale	360	50.62	1.08	44.34	50.07	50.78	51.33	53.18
Percent25_44	360	25.18	2.46	12.20	23.89	25.10	26.62	32.80

Many of these variables are right-skewed, and they were log-transformed before regression.

Descriptive Graphs

The following graphs divide the data into quartiles of selected explanatory variables, and plot it against (the logarithm of) the Land Density of Breweries (number of breweries per 100 square miles).

The pattern that arises is that Housing Density and Education levels are highly correlated with the Density of Breweries over land. Similar analysis was done for Brewery Density over hundred thousand people, with much weaker results.

By quartiles of income	By quartiles of the job-population ratio	By quartiles of Education
		
By quartiles of Population Density	**By quartiles of Housing Density**	**By quartiles of Density of Bars**
		
By quartiles of % Female	**By quartiles of 25-44 year olds**	
		

	Log Brew Density Land	Jobs-Population Ratio	Log Per Capita Income	Log Education	Log Pop Density	Log Housing Density	Log Bars DensityLand	Percent Female	Percent 25_44
Log Brew Density Land	1.00	0.32	0.40	0.45	0.61	0.61	0.58	0.13	-0.04
Jobs-Population Ratio	0.32	1.00	0.60	0.64	0.10	0.10	-0.01	0.00	0.36
Log Per Capita Income	0.40	0.60	1.00	0.64	0.34	0.36	0.03	0.01	0.15
Log Education	0.45	0.64	0.64	1.00	0.27	0.28	0.02	0.08	0.19
Log Pop Density	0.61	0.10	0.34	0.27	1.00	0.99	0.43	0.35	0.06
Log Housing Density	0.61	0.10	0.36	0.28	0.99	1.00	0.46	0.39	-0.01
Log Bars DensityLand	0.58	-0.01	0.03	0.02	0.43	0.46	1.00	0.10	-0.39
Percent Female	0.13	0.00	0.01	0.08	0.35	0.39	0.10	1.00	-0.27
Percent 25_44	-0.04	0.36	0.15	0.19	0.06	-0.01	-0.39	-0.27	1.00

Methodology

The number of breweries within an MSA (as a ratio to 100 square miles in the area) is predicted with the following model:

$$\begin{aligned} Log\ of\ BrewDensityLand \\ &= \beta_0 + \beta_1 Log\ of\ Per\ Capita\ Income + \beta_2 JobsPopulation\ Ratio \\ &+ \beta_3 \%\ of\ over 25\ with\ more\ than\ Bachelor's + \beta_4 PopDensity \\ &+ \beta_5 HousingDensity + \beta_6 BarsDensityLand + \beta_7 \%Female \\ &+ \beta_8 \%\ between\ 25\ and\ 44\ years\ old + \epsilon \end{aligned}$$

Via basic OLS. The strategy is to estimate the number of predicted breweries per 100 square miles, and calculate the difference between that quantity and the actual number of breweries per 100 square miles.

$$Missing\ Breweries = (Predicted\ BrewDensityLand - BrewDensityLand) * \frac{LandArea}{100}$$

A positive result is an indication that a particular locality has fewer breweries per 100 square miles than would be expected given the locality's income, jobs, education, and population density.

We want to locate MSAs that are relatively underserved by breweries but not a craft-beer desert, in the assumption that a sufficient critical mass is necessary to support a new brewery, but that too many breweries would crowd out a new one. We will identify the MSAs that have at least 4 "missing breweries" but fewer than 20 actual breweries.

Results

The results are as follows.

OLS Regression Results

Dep. Variable:	LogBrewDensityLand	R-squared:	0.619
Model:	OLS	Adj. R-squared:	0.610
Method:	Least Squares	F-statistic:	71.31
Date:	Fri, 21 Jun 2019	Prob (F-statistic):	6.08e-69
Time:	17:38:36	Log-Likelihood:	-362.91
No. Observations:	360	AIC:	743.8
Df Residuals:	351	BIC:	778.8
Df Model:	8		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-9.8108	4.044	-2.426	0.016	-17.765	-1.857
Jobs-Population Ratio	1.0089	0.648	1.557	0.120	-0.266	2.284
LogPerCapita_Income	0.4200	0.301	1.396	0.163	-0.172	1.011
LogEducation	1.0298	0.185	5.568	0.000	0.666	1.393
LogPopDensity	0.7506	0.378	1.985	0.048	0.007	1.494
LogHousingDensity	-0.3624	0.387	-0.937	0.349	-1.123	0.398
LogBarsDensityLand	0.8188	0.083	9.874	0.000	0.656	0.982
PercentFemale	-0.0486	0.040	-1.213	0.226	-0.127	0.030
Percent25_44	-0.0061	0.022	-0.278	0.781	-0.049	0.037

Omnibus:	12.273	Durbin-Watson:	1.984
Prob(Omnibus):	0.002	Jarque-Bera (JB):	12.823
Skew:	-0.462	Prob(JB):	0.00164
Kurtosis:	3.040	Cond. No.	6.67e+03

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified
[2] The condition number is large, 6.67e+03. This might indicate that there are
strong multicollinearity or other numerical problems.

Notice the strong multi-collinearity. Therefore the model is re-estimated as

$$Log\ of\ BrewDensityLand = \beta_0 + \beta_1 JobsPopulation\ Ratio + \beta_3 \%\ of\ over25\ with\ more\ than\ Bachelor's + \beta_4 HousingDensity + \beta_7 BarsDensityLand + \epsilon$$

And the results are

OLS Regression Results

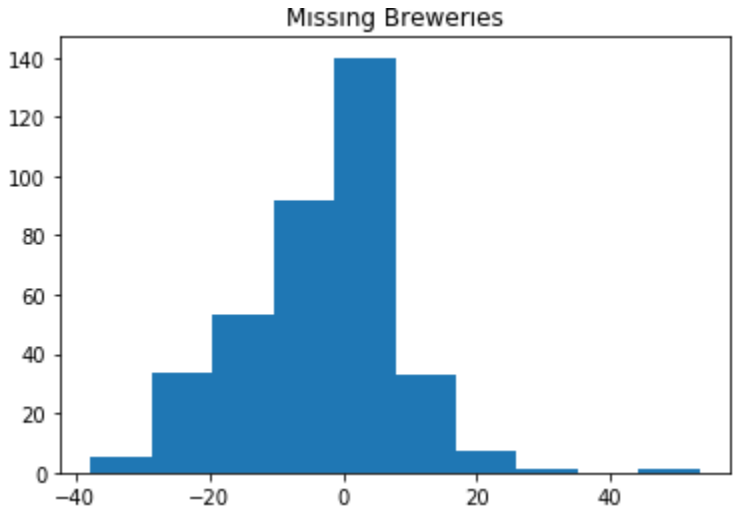
Dep. Variable:	LogBrewDensityLand	R-squared:	0.641
Model:	OLS	Adj. R-squared:	0.637
Method:	Least Squares	F-statistic:	161.0
Date:	Tue, 25 Jun 2019	Prob (F-statistic):	6.52e-79
Time:	13:07:33	Log-Likelihood:	-389.11
No. Observations:	366	AIC:	788.2
Df Residuals:	361	BIC:	807.7
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-6.8808	0.448	-15.349	0.000	-7.762	-5.999
Jobs-Population Ratio	0.6623	0.570	1.161	0.246	-0.459	1.784
LogEducation	0.9601	0.181	5.303	0.000	0.604	1.316
LogHousingDensity	0.3953	0.049	8.068	0.000	0.299	0.492
LogBarsDensityLand	1.0570	0.068	15.573	0.000	0.924	1.191

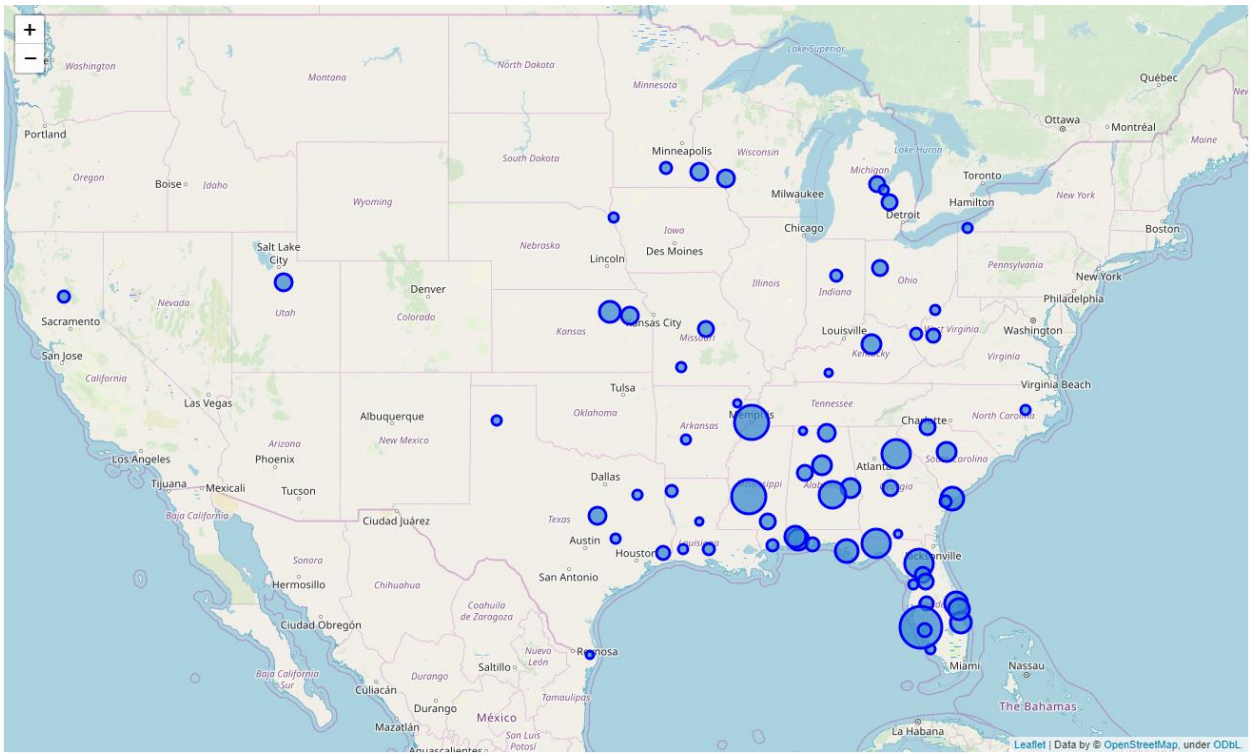
Omnibus:	24.925	Durbin-Watson:	2.026
Prob(Omnibus):	0.000	Jarque-Bera (JB):	28.506
Skew:	-0.617	Prob(JB):	6.46e-07
Kurtosis:	3.590	Cond. No.	93.1

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

On the basis of this model and the above formula, the Missing Breweries per MSA are calculated:



Focusing our attention on MSAs that are missing at least 4 breweries and have fewer than 20 breweries, we identify 68 MSAs.



Discussion

Most of the identified cities (as good candidates for a new craft brewery) are located in the Southeastern United States. These are cities where the proportion of highly-educated, younger people have been rising, and where housing booms have recently taken place. The number of craft breweries has not followed suit, making them attractive candidates.

GeoName	**Missing Breweries**	**Existing Number of Breweries**
North Port-Sarasota-Bradenton, FL	21.94	5
Jackson, MS	18.48	1
Memphis, TN-MS-AR	18.37	7
Gainesville, FL	15.48	7
Tallahassee, FL	15.16	6
Athens-Clarke County, GA	14.57	4
Montgomery, AL	14.31	2
Hilton Head Island-Bluffton-Beaufort, SC	12.07	7
Palm Bay-Melbourne-Titusville, FL	11.87	10
Panama City, FL	11.67	1
Mobile, AL	11.41	4
Port St. Lucie, FL	11.10	5
Manhattan, KS	10.94	4

Daphne-Fairhope-Foley, AL	10.88	4
Sebastian-Vero Beach, FL	10.57	6
Lexington-Fayette, KY	10.04	16
Auburn-Opelika, AL	9.78	2
Columbia, SC	9.64	14
Birmingham-Hoover, AL	9.63	16
La Crosse-Onalaska, WI-MN	9.22	6
Waco, TX	8.86	3
Huntsville, AL	8.85	16
Provo-Orem, UT	8.74	6
Rochester, MN	8.69	9
Topeka, KS	8.52	5
The Villages, FL	8.49	5
Midland, MI	8.29	7
Flint, MI	8.03	8
Kahului-Wailuku-Lahaina, HI	7.98	5
Hattiesburg, MS	7.98	2
Lima, OH	7.95	2
Tuscaloosa, AL	7.79	3
Warner Robins, GA	7.74	5
Spartanburg, SC	7.71	7
Jefferson City, MO	7.62	1
Ocala, FL	7.54	3
Pensacola-Ferry Pass-Brent, FL	7.39	9
Lakeland-Winter Haven, FL	7.31	6
Punta Gorda, FL	7.29	5
Beaumont-Port Arthur, TX	6.94	3
Charleston, WV	6.80	6
Chico, CA	6.40	5
Gulfport-Biloxi-Pascagoula, MS	6.34	5
Huntington-Ashland, WV-KY-OH	6.14	4
Lafayette, LA	6.10	7
Savannah, GA	6.01	13
Mankato-North Mankato, MN	5.68	7
Kokomo, IN	5.68	2
Shreveport-Bossier City, LA	5.57	8
Amarillo, TX	5.29	4
Erie, PA	5.18	12
Parkersburg-Vienna, WV	5.15	3
Tyler, TX	5.11	5
Sioux City, IA-NE-SD	5.07	4

Homosassa Springs, FL	5.03	4
Saginaw, MI	4.98	7
Hot Springs, AR	4.84	3
Naples-Immokalee-Marco Island, FL	4.73	17
Springfield, MO	4.69	12
College Station-Bryan, TX	4.63	11
Lake Charles, LA	4.56	3
Greenville, NC	4.52	11
Brownsville-Harlingen, TX	4.22	3
Florence-Muscle Shoals, AL	4.19	1
Valdosta, GA	4.11	1
Alexandria, LA	4.09	1
Jonesboro, AR	4.08	1
Bowling Green, KY	4.02	2

Limitations

The biggest limitation of this study is that it relied on a free version of Foursquare, which limits to 50 results per search. In practical terms, this means that we only get 50 breweries per MSA. Two solutions suggest themselves readily: one is to use a censored regression. Another is to redefine the geographical area to be a zip-code or a census-tract. Both of these corrections would produce a better estimate.

Conclusion

This report has examined the determinants of craft brewery location in metropolitan statistical areas of the United States, with the aim of guiding the location decisions of enterprising brewers. It was found that income, jobs, age, and gender were generally poor predictors, while housing (and population) density and educational attainment tended to predict brewery location fairly well. The basis of this information, we identified nearly 70 MSAs that would provide fertile ground for new-brewery location.