# Using Data Science to Find the Ideal New Jersey Shore Town for a New Sustainable Microbrewery Business

Submitted as the Capstone Project for the IBM Data Science Certificate

Maureen Newman, June 2019

### 1. Introduction

This report outlines the data science methods used to compare three popular New Jersey coast towns for the purpose of starting a new sustainable microbrewery business.

### 1.1 Background

In the last several years, microbreweries have been springing up across New Jersey, and throughout metropolitan cities throughout the United States, at a steady pace. A brewery is a business that makes and sells beer. A microbrewery, or craft brewery, is independently owned, produces small batches of beer, and generally emphasize quality, flavor, and brewing technique.

The microbrewing movement originated in the United Kingdom in the 1970's, however, historians believe that people have been brewing beer commercially since at least 2,500 BC. Today there are over 5,000 microbreweries in the United States--100 just in the state of New Jersey.

New Jersey's Atlantic shoreline in New Jersey is ranked among the highest-income areas in the United States, and its many beach towns are a popular tourist destination for both locals and tourists alike. Among these beach towns is hip Asbury Park, casino mecca Atlantic City, and charming Cape May. Each town is home music venues, restaurants, bars, and a beach with a boardwalk.

Environmental sustainability is important to many craft brewery owners. This involves considering the environmental impact on beer production. The Brewers Association (<a href="www.brewersassociation.org">www.brewersassociation.org</a>) provides manuals outlining design and strategies for energy sustainability, solid waste sustainability, and water/wastewater sustainability in brewing.

### 1.2 Problem and Interest

Three friends wish to open a sustainable craft brewery in a New Jersey coastal town. They commissioned me, a data scientist, to use data science methods to analyze three towns in particular: Asbury Park, Atlantic City, and Cape May.

The brewers would like to start their business in a large, densely populated area with plenty of venues nearby (music, nightlife, etc.), but with no or few similar establishments—maximizing draw, while minimizing competition. They would also like to choose a community that has a proven commitment to sustainable practices.

To solve this problem, I used the Foursquare API to leverage location data to a.) find other breweries in, and around, each town, b.) explore popular venues in each town, and c.) explore trending venues on Saturday evening. I then analyzed a dataset for each county's and city's square footage, population, population density, and sustainability rating. These data science techniques will arm the three entrepreneurs with data, enabling them to make this important data-driven decision--determining the most ideal location for their business endeavor.

### 2. Data

#### 2.1 Data Sources

For the first part of the analysis, I used a Foursquare API (<a href="https://developer.foursquare.com/">https://developer.foursquare.com/</a>) to find nearby breweries, other nearby venues, and trending venues in the area.

For the second art of my analysis, I used a dataset called "Sustainable Jersey Certified Municipalities," available from the New Jersey Geographic Information Network (NJGIN) Open Data Site: <a href="http://njogis-newjersey.opendata.arcgis.com/datasets/sustainable-jersey-certified-municipalities">http://njogis-newjersey.opendata.arcgis.com/datasets/sustainable-jersey-certified-municipalities</a>. This dataset shows the New Jersey municipalities that were certified under the Sustainable Jersey Program in 2015. More information on the Sustainable Jersey program is available at <a href="www.sustainablejersey.com">www.sustainablejersey.com</a>. The population and population density data I used is from 2010, as it was the most recent data available.

Throughout this report, I refer to municipalities as towns or cities.

## 2.2 Data Cleaning

The Foursquare data did not require cleaning, as I pulled the data I wanted using an API query. However, I did experiment with several query terms: Beer, Brew, Breweries, Brewery, Bars, Pubs, Brewpub.

The sustainability dataset required some cleaning. I first used python to download and explore the dataset, and transform the data into a pandas dataframe. To simplify the dataframe, I dropped the unneeded columns, renamed the remaining columns, and dropped the unneeded rows.

### 2.3 Feature Selection

Although this dataset includes 26 attributes, I only used County, Municipality, Sq\_Miles, Pop2010, Popden2010, and Certification Designation in my analysis. The others were not pertinent to the problem this project was designed to address.

### 3. Methodology

### 3.1 Foursquare Data

Using the Foursquare API, I explored the towns of Asbury Park, Atlantic City, and Cape May. My python code is available on Github at:

https://github.com/maurnew/coursera\_capstone/blob/master/MN's%20IBM%20Capstone%2 0Project%20Code.ipynb

### Setup

In IBM Watson Studio, I imported the python libraries and defined my Foursquare developer credentials. I then converted each town's central Convention Hall address, into its latitude and longitude coordinates.

# Identifying Similar Businesses in Town

To identify other breweries in and around each town, I retrieved the venues containing the word "Brewery" within a radius of 8,000 meters. I defined the corresponding Foursquare URL ('https://api.foursquare.com/v2/venues/search?client\_id={}&client\_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'), set the GET request, and viewed the results in JSON format.

I created a pandas dataframe from the JSON data and filtered the dataframe to include only data of interest. I then used the python folium library to create a map with the venues superimposed on top.

### Identifying Popular Venues in Town

To identify popular venues in each town, I repeated the above steps, but used the URL specific to retrieving venue results

('https://api.foursquare.com/v2/venues/explore?client\_id={}&client\_secret={}&ll={},{}&v={}&radius={}&limit={}').

### Identifying Trending Venues in Town

To identify trending venues in each town, I repeated the above steps, but used the URL specific to retrieving venue results

('https://api.foursquare.com/v2/venues/trending?client id={}&client secret={}&ll={},{}&v={}').

### 3.2 Sustainable Jersey Certified Municipalities Data

For the second part of my analysis, I explored a dataset called "Sustainable Jersey Certified Municipalities." I determined the size in square miles, the population, and the population density of the towns of Asbury Park, Atlantic City, and Cape May, and the counties in which they reside—Monmouth County, Atlantic County, and Cape May County. I then looked at the sustainability award given to each town by the Sustainable Jersey Program in 2015.

#### Setup

In IBM Watson Studio, I imported the python libraries, downloaded the csv dataset (<a href="https://opendata.arcgis.com/datasets/4999349529e64474abcaeefa03f63d0c\_18.csv">https://opendata.arcgis.com/datasets/4999349529e64474abcaeefa03f63d0c\_18.csv</a>), and created a pandas dataframe.



# Data Cleanup

To clean up the data, I dropped the unneeded columns:

'OBJECTID','MUN','MUN\_LABEL','MUN\_TYPE','GNIS\_NAME','GNIS','SSN','MUN\_CODE','CENSUS2 010','ACRES','POP2000','POP1990','POP1980','POPDEN2000','POPDEN1990','POPDEN1980','GLO BALID','SHAPE\_Length','SHAPE\_Area.'



I renamed the remaining columns, leaving me with COUNTY, CITY, SQUARE MILES, POPULATION, POPULATION DENSITY, and SUSTAINABILITY AWARD.



I dropped the unneeded rows, keeping only the counties of Atlantic, Cape May, and Monmouth, which contain the three cities my clients have asked me to examine.

### Analyzing the County Data

To determine the largest county, I created a new dataframe containing only the columns COUNTY and SQUARE MILES. I grouped and sorted the dataframe by COUNTY, and summed each county's total square miles.

To determine the county with the largest population density (the number of people per square mile), I created a new dataframe containing only the columns COUNTY and POPULATION DENSITY. I grouped and sorted the dataframe by COUNTY, and summed each county's total population density.

To determine the county with the largest population, I created a new dataframe containing only the columns COUNTY and POPULATION. I grouped and sorted the dataframe by COUNTY, and summed each county's total population.

### **Analyzing the City Data**

To determine the largest city in the three counties, I created a new dataframe containing only the columns CITY and SQUARE MILES. I grouped and sorted the dataframe by CITY, and summed each city's total square miles.

To determine the city in the three counties with the largest population density (the number of people per square mile), I created a new dataframe containing only the columns CITY and POPULATION DENSITY. I grouped and sorted the dataframe by CITY, and summed each city's total population density.

To determine the city in the three counties with the largest population, I created a new dataframe containing only the columns CITY and POPULATION. I grouped and sorted the dataframe by CITY, and summed each city's total population.

### Analyzing Sustainability Award Level

In 2015 some New Jersey municipalities were certified under the Sustainable Jersey Program. Municipalities that achieve the certification are considered to be among the leaders striving for sustainability. There are two award levels: Silver (the highest) and bronze. For details, visit http://www.sustainablejersey.com/.

To determine the sustainability award of each of the three towns, I created a new dataframe containing only the columns SUSTAINABILITY AWARD, COUNTY, and CITY. I grouped and sorted the dataframe by SUSTAINABILITY AWARD, then by COUNTY.

### 4. Results

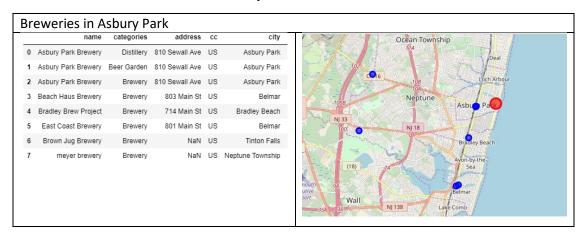
In this section I will present the results of the foursquare data analysis and the sustainability dataset analysis.

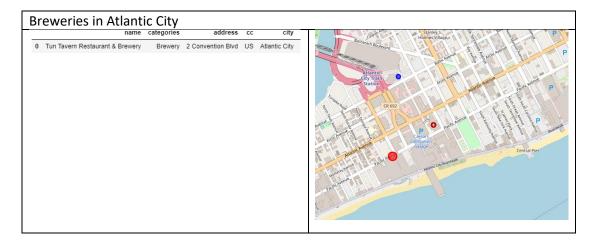
### 4.1 Foursquare Data

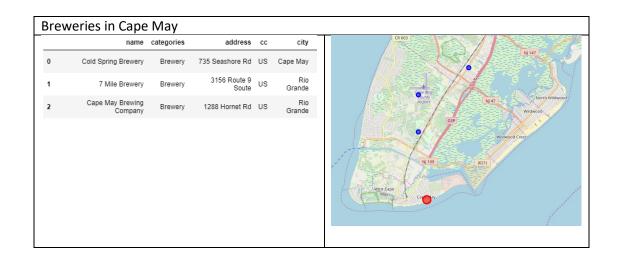
The results of the foursquare data analysis are as follows:

# Identifying Similar Businesses in Town

Foursquare data shows that Asbury Park has 1 brewery in town, and 5 in nearby towns. Cape May also has 1 in town, but 2 in adjacent towns. Atlantic City has 1 brewery in town and none within 8000 meters of the city center.

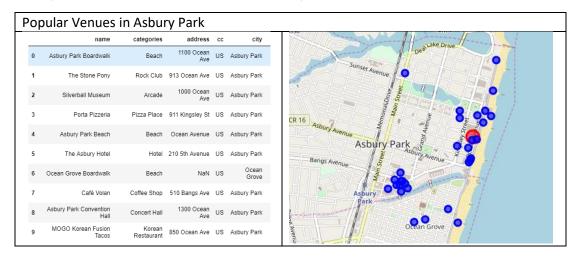




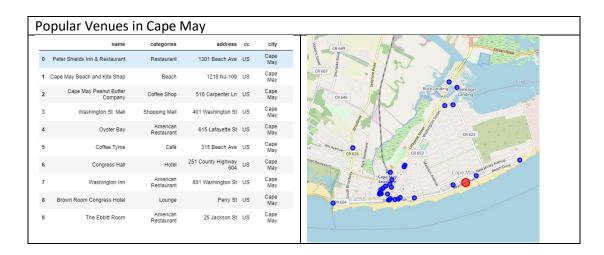


# Identifying Popular Venues in Town

Foursquare data lists the top 10 most popular Asbury Park venues as: 3 beaches, 2 music venues, 2 restaurants, a coffee shop, a hotel, and an arcade. Atlantic City's 10 most popular venues are: restaurants (including a sandwich shop and a steakhouse), a lounge, a bar, and a sporting goods store. Cape May's top 10 venues are: 4 restaurants, a beach, a hotel, a lounge, a café, a coffee shop, and a shopping mall.



	name	categories	address	CC	city
)	Chef Vola's	Italian Restaurant	111 S Albion PI	US	Atlantic City
1	White House Subs	Sandwich Place	2301 Arctic Ave	US	Atlantic City
2	Buddakan	Asian Restaurant	1 Atlantic Ocean	US	Atlantic City
3	Dock's Oyster House	Seafood Restaurant	2405 Atlantic Ave	US	Atlantic City
4	Ruth's Chris Steak House - Atlantic City, NJ	Steakhouse	2020 Atlantic Ave	US	Atlantic City
5	Angelo's Fairmount Tavern	Italian Restaurant	2300 Fairmount Ave	US	Atlantic City
6	Wet Willie's	Bar	The Quarter at Tropicana	US	Atlantic City
7	Bass Pro Shop	Sporting Goods Shop	30 N Christopher Columbus Blvd	US	Atlantic City
8	Gordon Ramsay Pub and Grill	Restaurant	2100 Pacific Ave	US	Atlantic City
9	Kelsey's	Lounge	1545 Pacific Ave	US	Atlantic City



# **Identifying Trending Venues in Town**

There were no trending venues on Foursquare for any of the three locations at the time of this report.

### 4.2 Sustainable Jersey Certified Municipalities Data Results

The results of the dataset analysis are as follows:

### **County Analysis Results**

### **Square Miles**

Atlantic County (Atlantic City), is the largest county in square miles at 366. Coming in second is Monmouth County (Asbury Park) at 218, and third is Cape May County (Cape May) at 188.

COUNTY	SQUARE MILES
ATLANTIC	366.472246
MONMOUTH	218.181419
CAPE MAY	187.837991

# **Population Density**

Monmouth County (Asbury Park) at 52,771, has more than double the highest population density of the second highest, Atlantic County (Atlantic City) at 21,051. Cape May County (Cape May) has the lowest population density at 7,843.

COUNTY	POPULATION DENSITY
MONMOUTH	52771
ATLANTIC	21051
CAPE MAY	7843

### **Population**

Monmouth County (Asbury Park) at 302,028, has almost twice the population than the second highest, Atlantic County (Atlantic City) at 195,499. Cape May County (Cape May) has the lowest population at 58,734.

COUNTY	POPULATION
MONMOUTH	302028
ATLANTIC	195499
CAPE MAY	58734

# City Analysis Results

# **Square Miles**

Of the cities in Atlantic, Cape May, and Monmouth counties, Atlantic City is the 12<sup>th</sup> largest at 15.90 square miles. Cape May is 24<sup>th</sup> at 2.88 square miles. Asbury Park, at number 33, is the smallest, with 1.52 square miles.

CITY	SQUARE MILES
Hamilton Township	112.924090
Galloway Township	111.251374
Middle Township	82.710932
Upper Township	68.414485
Howell Township	61.170255
Middletown Township	43.538520
Buena Vista Township	41.574907
Hammonton	41.310832
Wall Township	31.700737
Marlboro Township	30.433504
Holmdel Township	18.064123
Atlantic City	15.905108
Ocean City	11.801411
Egg Harbor City	11.426255
Ocean Township	10.985099
Brigantine	10.704999
Woodbine Borough	8.018742
Pleasantville	7.262592
Eatontown Borough	5.890035
Somers Point	4.975879
Avalon Borough	4.967886
Linwood	4.380875
Little Silver Borough	3.180725
Cape May	2.882549
Sea Isle City	2.757348
Ventnor City	2.549335
North Wildwood	2.490038
Stone Harbor Borough	2.312428
Shrewsbury Borough	2.176598
Red Bank Borough	2.160315
Roosevelt Borough	1.947676
Margate City	1.632162
Asbury Park	1.524615
Keyport Borough	1.449759
Atlantic Highlands Borough	1.236279
Sea Bright Borough	1.221324
West Cape May Borough	1.182092
Highlands Borough	0.855989
Bradley Beach Borough	0.645866
Longport Borough	0.594037
Cape May Point Borough	0.300080

# **Population Density**

Of the cities in Atlantic, Cape May, and Monmouth counties, Asbury Park is by far the most densely populated at 10,603. Atlantic City ranks 10<sup>th</sup> at 2,488. Cape May has the 21<sup>st</sup> highest with a population density of 1,252.

CITY	POPULATION DENSITY
Asbury Park	10803
Bradley Beach Borough	6716
Highlands Borough	5888
Red Bank Borough	5651
Keyport Borough	5028
Ventnor City	4193
Margate City	3898
Atlantic Highlands Borough	3565
Pleasantville	2789
Atlantic City	2488
Ocean Township	2486
Somers Point	2172
Eatontown Borough	2158
Little Silver Borough	1871
Shrewsbury Borough	1755
North Wildwood	1629
Linwood	1627
Middletown Township	1528
Longport Borough	1517
Marlboro Township	1321
Cape May	1252
Sea Bright Borough	1157
Ocean City	992
Cape May Point Borough	970
Holmdel Township	929
Brigantine	883
West Cape May Borough	868
Howell Township	835
Wall Township	825
Sea Isle City	769
Roosevelt Borough	455
Stone Harbor Borough	375
Egg Harbor City	373
Hammonton	358
Galloway Township	336
Woodbine Borough	309
Avalon Borough	269
Hamilton Township	235
Middle Township	229
Buena Vista Township	182
Upper Township	181

# Population

Of the cities in Atlantic, Cape May, and Monmouth counties, Atlantic City has the  $4^{th}$  largest population at 39,558. Asbury Park is  $12^{th}$  at 16,116, and Cape May is number 32 with a population of only 3,607.

CITY	POPULATION
Middletown Township	66522
Howell Township	51075
Marlboro Township	40191
Atlantic City	39558
Galloway Township	37349
Ocean Township	27291
Hamilton Township	26503
Wall Township	26164
Pleasantville	20249
Middle Township	18911
Holmdel Township	16773
Asbury Park	16116
Hammonton	14791
Eatontown Borough	12709
Upper Township	12373
Red Bank Borough	12206
Ocean City	11701
Somers Point	10795
Ventnor City	10850
Brigantine	9450
Buena Vista Township	7570
Keyport Borough	7240
Linwood	7092
Margate City	6354
Little Silver Borough	5950
Highlands Borough	5005
Atlantic Highlands Borough	4385
Bradley Beach Borough	4298
Egg Harbor City	4243
North Wildwood	4041
Shrewsbury Borough	3809
Cape May	3807
Woodbine Borough	2472
Sea Isle City	2114
Sea Bright Borough	1412
Avalon Borough	1334
West Cape May Borough	1024
Longport Borough	895
Roosevelt Borough	882
Stone Harbor Borough	866
Cape May Point Borough	291

# Sustainability Results

Only one of the three towns, Cape May has earned the highest sustainability award, Silver. However, Asbury Park and Atlantic City earned a Bronze level sustainability award.

SUSTAINABILITY AWARD	COUNTY	CITY
Silver	CAPE MAY	Stone Harbor Borough
Silver	CAPE MAY	Ocean City
Silver	CAPE MAY	Cape May
Silver	ATLANTIC	Hammonton
Silver	ATLANTIC	Galloway Township
Silver	ATLANTIC	Buena Vista Township
Bronze	MONMOUTH	Wall Township
Bronze	MONMOUTH	Shrewsbury Borough
Bronze	MONMOUTH	Sea Bright Borough
Bronze	MONMOUTH	Roosevelt Borough
Bronze	MONMOUTH	Red Bank Borough
Bronze	MONMOUTH	Ocean Township
Bronze	MONMOUTH	Middletown Township
Bronze	MONMOUTH	Marlboro Township
Bronze	MONMOUTH	Little Silver Borough
Bronze	MONMOUTH	Keyport Borough
Bronze	MONMOUTH	Howell Township
Bronze	MONMOUTH	Holmdel Township
Bronze	MONMOUTH	Highlands Borough
Bronze	MONMOUTH	Eatontown Borough
Bronze	MONMOUTH	Bradley Beach Borough
Bronze	MONMOUTH	Atlantic Highlands Borough
Bronze	MONMOUTH	Asbury Park
Bronze	CAPE MAY	Woodbine Borough
Bronze	CAPE MAY	West Cape May Borough
Bronze	CAPE MAY	Upper Township
Bronze	CAPE MAY	Sea Isle City
Bronze	CAPE MAY	North Wildwood
Bronze	CAPE MAY	Middle Township
Bronze	CAPE MAY	Cape May Point Borough
Bronze	CAPE MAY	Avalon Borough
Bronze	ATLANTIC	Ventnor City
Bronze	ATLANTIC	Somers Point
Bronze	ATLANTIC	Pleasantville
Bronze	ATLANTIC	Margate City
Bronze	ATLANTIC	Longport Borough
Bronze	ATLANTIC	Linwood
Bronze	ATLANTIC	Hamilton Township
Bronze	ATLANTIC	Egg Harbor City
Bronze	ATLANTIC	Brigantine
Bronze	ATLANTIC	Atlantic City

### 5. Discussion

In this section I will discuss the results and recommend a course of action.

### 5.1 Observations

## Identifying Similar Businesses in Town

Asbury Park has significantly more breweries in and around town at six. Cape May has three and Atlantic City has the least at only one.

### Identifying Popular Venues in Town

Asbury Park has three popular beaches, and Cape May has 1. Atlantic City's popular venues are primarily restaurants, and Cape May has an assorted mix of venues. Asbury Park is the only town with two popular music venues.

# **Identifying Trending Venues in Town**

The analysis did not show any venues that were trending at the time of analysis.

# **County Analysis Results**

Monmouth County (Asbury Park) is the 2<sup>nd</sup> largest of the three counties, has by far the highest population as well as double the population density of the other two counties.

Atlantic County (Atlantic City) is the largest is square miles, and has about half the population and population density of Monmouth County, but twice that of Cape May County.

Cape May County (Cape May) is the smallest in size, and in population and population density.

### City Analysis Results

Of the three cities analyzed, Asbury Park is the smallest in square footage, but has the highest population density. Its population is in the middle.

Atlantic City is the largest in square miles. It comes in  $2^{nd}$  in population density, but  $1^{st}$  in population.

Cape May ranks in the middle for size, and has the smallest population and population density.

### **Sustainability Results**

All three towns earned sustainability awards, showing that they are all committed to sustainable practices. Cape May is the only one with a silver award in sustainability.

### 5.2 Recommendations

All three towns received a sustainability award. Since a commitment to sustainability was important to the entrepreneurs, any of the three towns are eligible based on this criteria.

Based on my analysis, I do not recommend opening a microbrew in Cape May, the southernmost town analyzed. While there is relatively low competition, at one brewery in town and two in nearby areas, its low population and population density indicates that it may not be as bustling as the other two cities. However, perhaps the space rental is low enough and number of tourists is high enough to offset this. These factors may cause for further studies.

This leaves us with Asbury Park or Atlantic City. Atlantic City has the least amount of competition, at only one brewery. Based on my analysis, the most popular venues are restaurants. A microbrew might do well here, depending on space rental prices and local ordinances.

My top recommendation is Asbury Park, the most populated, smallest town, with several beaches and concert venues. This points to a lot of foot traffic. However, there are already six microbreweries nearby, so there is quite a lot of competition. But this competition could add to a microbrew mecca, which could pull in customers for all of the breweries nearby. If Asbury Park is chosen, I recommend choosing a distinct niche to differentiate the new brewery from the others. Competition could be further analyzed using Foursquare data and data science methods. A microbrew could do well here, depending on space rental prices and local ordinances.

#### 6. Conclusions

The three microbrew entrepreneurs asked me to start their business in a large, densely populated area with plenty of venues nearby (music, nightlife, etc.), but with no or few similar establishments—maximizing draw, while minimizing competition. They want a community that has a proven commitment to sustainable practices. I used data science methods to analyze two distinct sets of data. This enabled us to view the characteristics of each town, and the surrounding areas, to see how well the criteria met the entrepreneurs' needs. My number one recommendation is Asbury Park due to its small size, large population, and popular venues conducive to foot traffic. Atlantic City could also be a good spot, if space rental rates are significantly lower. Cape May is probably too large and sleepy to make an ideal spot for a new craft brewery.