**Capstone Project - The Battle of Neighborhoods**

* **Introduction:**

The City of New York, is the most populous city in the United States. It is diverse and is the financial capital of USA. It is multicultural. It provides lot of business opportunities and business friendly environment. It has attracted many different players into the market. It is a global hub of business and commerce. The city is a major center for banking and finance, retailing, world trade, transportation, tourism, real estate, new media, traditional media, advertising, legal services, accountancy, insurance, theater, fashion, and the arts in the United States. This also means that the market is highly competitive. As it is highly developed city so cost of doing business is also one of the highest. Thus, any new business venture or expansion needs to be analyzed carefully. The insights derived from analysis will give good understanding of the business environment which help in strategically targeting the market. This will help in reduction of risk. And the Return on Investment will be reasonable.

* **Business Problem:**

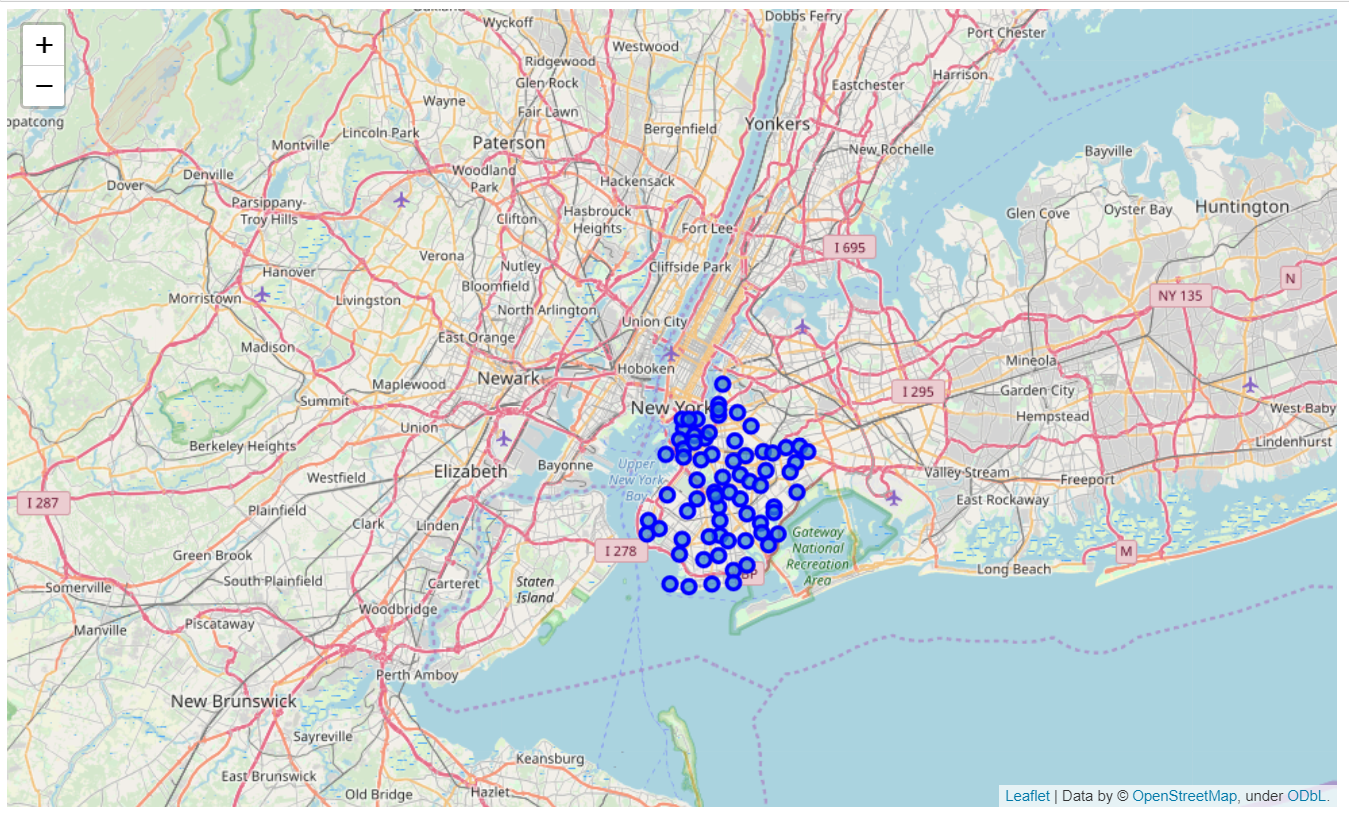
The City of New York is famous for its entertaining culture. The dance culture comprises of various international clubs inclined by the city's immigrant history. Dance studios became very popular in the United States and it spread across various location across the geographical area. The rich culture with lively people makes thing more attractive for international travels.

The need of the hour is to get a dance studio nearby accompanied with security, low cost and hygienic environment. The Business problem depicted in this project is to address a user-friendly searching app which encapsulated with easy searching techniques and highly trained machine learning models.

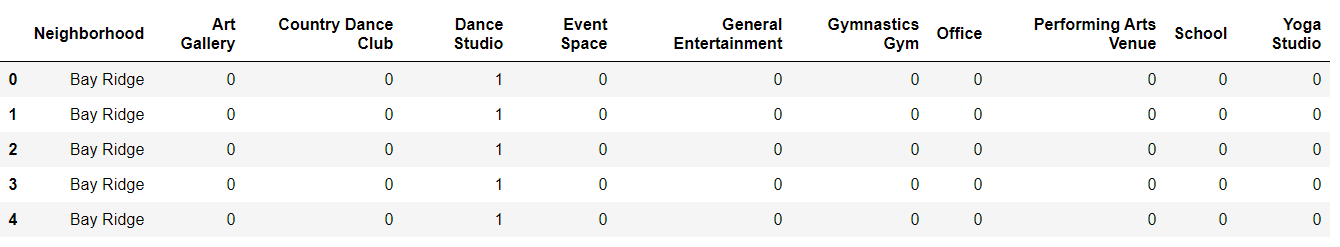
* **Data Selection:**

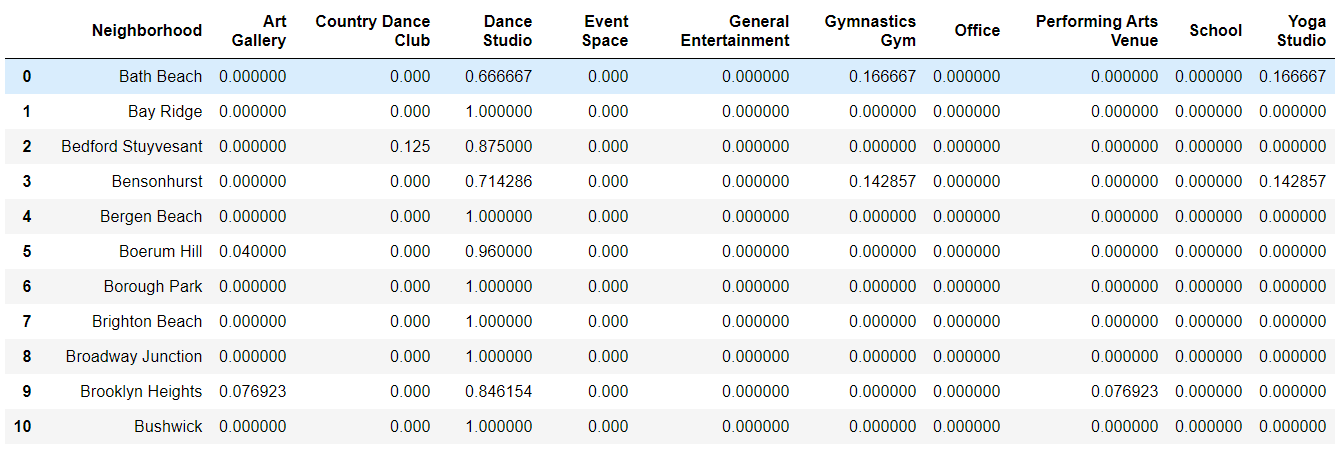
|  |
| --- |
|  |
|  |

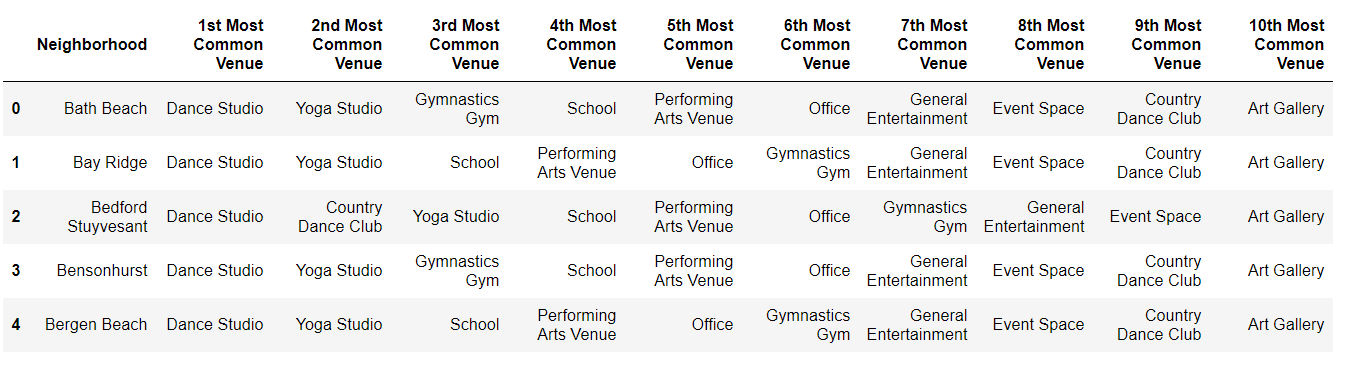
* **New York Map:**

****

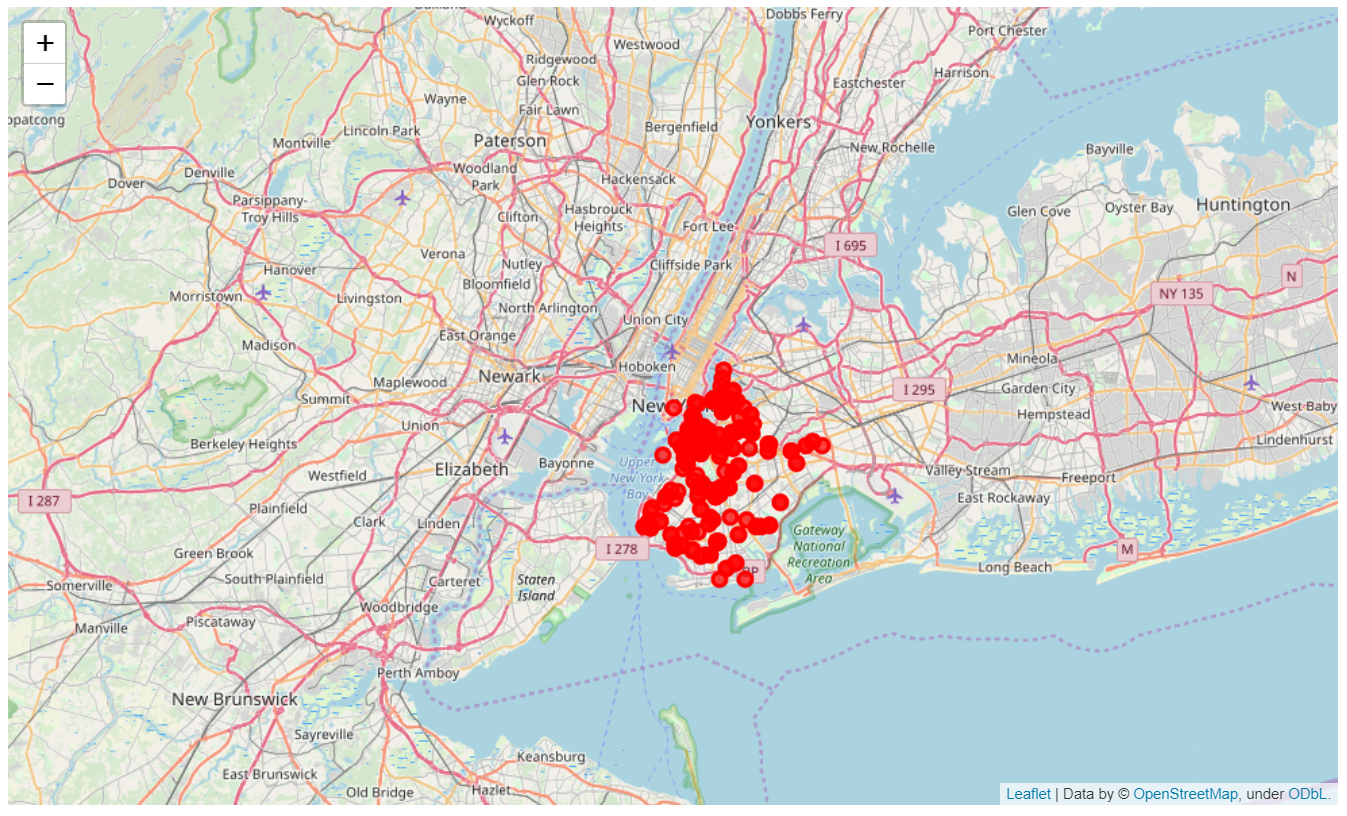
* **Methodology:**

****

****

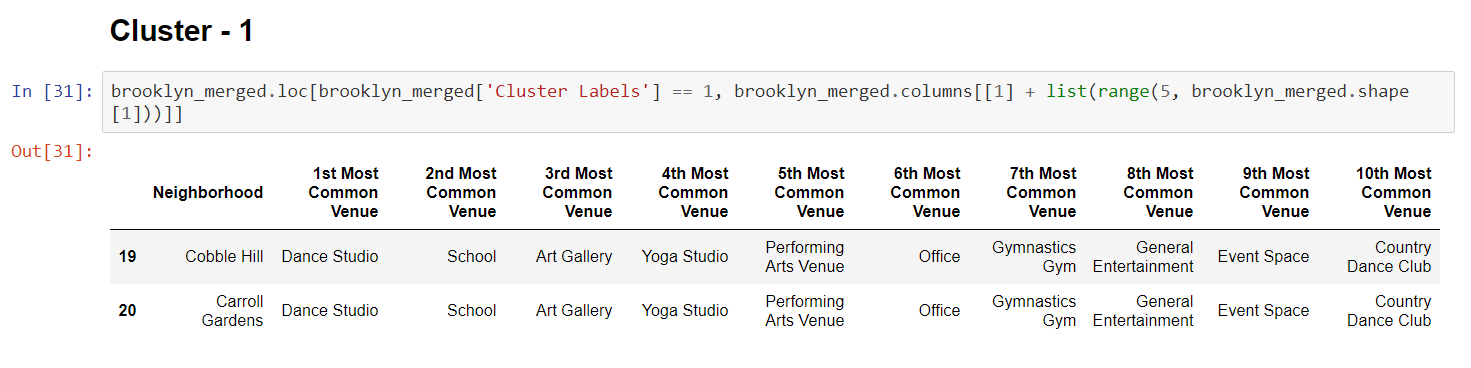
****

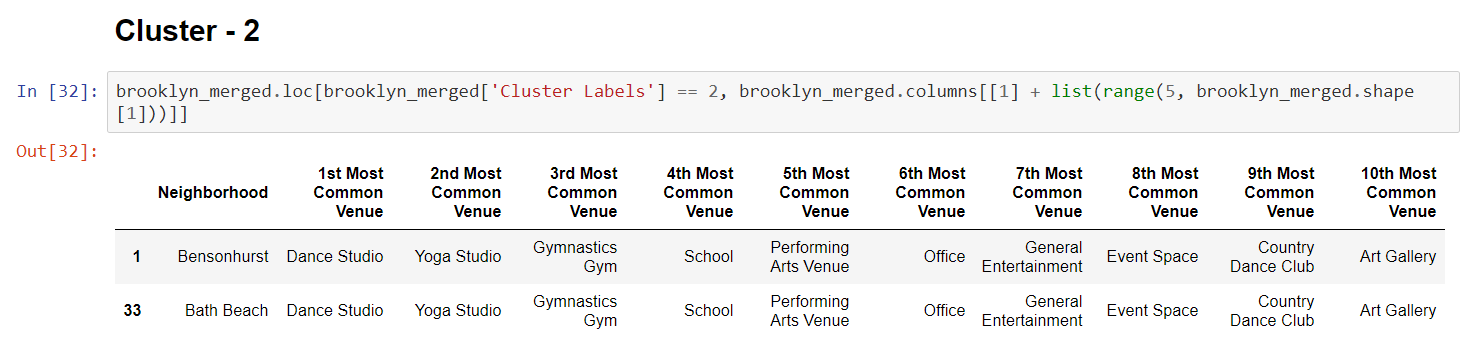
**Top ‘N’ Neighborhood**

****

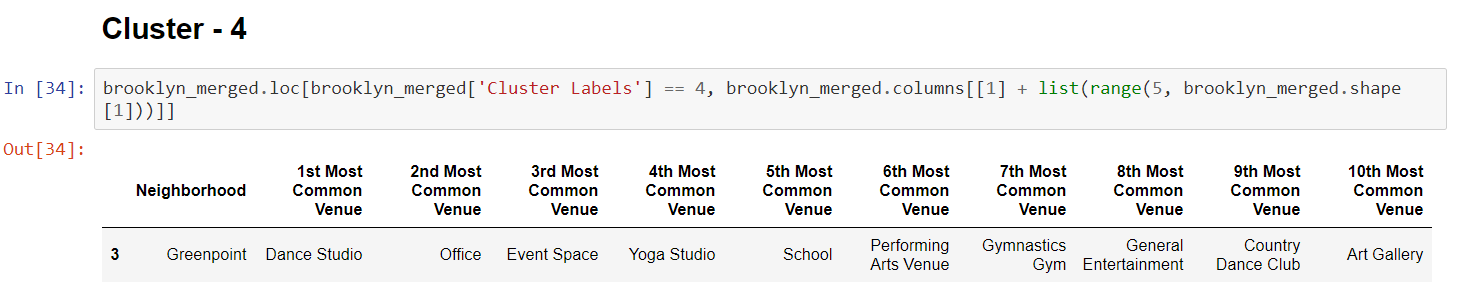
**Dance Studio in Brooklyn**

* **Results:**

****

****

****

****

* **Discussion:**

It is clear now that k-means is a simplistic, yet powerful algorithm and it can be useful for many different types of problems that may arise in analytics. With that said, it may not always be the best choice for our problem and there are some assumptions that the algorithm makes which we need to be aware of if we are going to use it. Probably the biggest assumption and limitation of k-means is that it assumes that the clusters are spherical.

* **Conclusion:**

This presentation is developed for IBM data science capstone project assignment to keep in view the optimal usage of machine learning model for better outcome. Although dataset used is very small and the accuracy of the outcome is compromised but is it a good opportunity to showcase the learning of the course by adhering the use case.