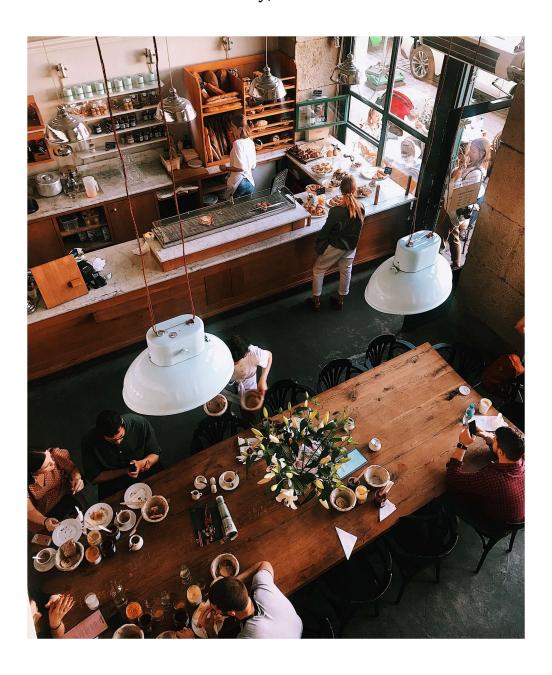# Location of Manhattan's Next Breakfast Spot

IBM Applied Data Science Capstone
Capstone Project

Vineet Raichur
May, 2020

Background

Mr. John Snow is a 35 year old software engineer working in Manhattan, New York. John was born and brought up in the New York city. As a proud New Yorker, John's dream has been to own and operate a breakfast spot in Manhattan. Having saved some money, John has decided to finally retire from the software engineering job and pursue his dream of running a breakfast spot. John is motivated by the concept of a neighborhood breakfast spot which will be a favorite among the locals. John has asked me to help him with finding the location for his breakfast spot.

Business Problem

Manhattan happens to be one of the most expensive real estate markets in the United States [1]. Therefore, choosing the right location will be critical to ensure profitability of the proposed breakfast spot. In choosing the right location, we will need to consider several factors including type of the neighborhood, size of the space, cost of renting/leasing and competition in the neighborhood. To help narrow down the search we decided that we will first identify a few neighborhoods suitable for opening the breakfast spot. Once suitable neighborhoods have been identified, further research will be conducted to identify suitable buildings/spaces.

Audience

John Snow is a persona of an individual whose goal is to own and operate a neighborhood breakfast spot. This analysis can be useful to any individual thinking of owning and operating their small food and beverage business in a city where real estate is expensive. Individuals looking to start and establish a small business, compared to a chain or a franchise, will not have a brand image that they can benefit from. In this situation, small business owners will need to be mindful of the competition from the more established businesses in the vicinity and invest in creating a community around them that will support their business in the long run. Data-based insights can help small businesses in choosing the right location. While this analysis is conducted specifically for a breakfast spot, the overall data science-based approach can be applied in the context of other small food and beverage businesses.

Data

We will use two data elements in this analysis.

1. List of neighborhoods in Manhattan
   NY city has a total of 5 boroughs and approximately 329 neighborhoods [2]. We will use a dataset that contains the 5 boroughs and a list of the neighborhoods along with their latitude and longitude coordinates within each borough. We will use the dataset made available by

NYU Spatial Data Repository [3]. Because we are focusing on Manhattan specifically for this analysis, we will filter out the data for the other four boroughs available in the dataset.

2. Data on the number commercial establishments in each neighborhood of Manhattan
Will use the Foursquare API to gather data on the different types of venues in each neighborhood. We will specifically use information on the category of the venue (e.g., coffee shop, restaurant, park, etc.) and the number of establishments in each category in a given location.

Analysis Plan

We will employ data science skills such as working with APIs, data wrangling, unsupervised machine learning (k-means clustering) and visualizing clusters (map, folium). In identifying suitable neighborhoods for establishing a new breakfast spot will be looking for neighborhoods with low number of existing breakfast spots and high number of certain types of venues that could indicate customer support for food and beverage businesses in the neighborhood. We will use k-means clustering method to group neighborhoods on the basis of the number of venues in certain categories. We will use the silhouette score approach to choose the optimal number of clusters. We will then explore the neighborhoods in each cluster to identify a few options suitable for establishing the new breakfast spot.

References

[1]   CNBC, "Manhattan real estate is the most expensive in the US per square foot with some properties topping $10,000: Study," 2018. [Online]. Available: https://www.cnbc.com/2018/08/11/manhattan-real-estate-is-the-most-expensive-in-the-us-per-square-foo.html. [Accessed: 23-May-2020].

[2]   Baruch College, "New York City (NYC) Neighborhoods - By Borough," 2020. [Online]. Available: https://www.baruch.cuny.edu/nycdata/population-geography/neighborhoods.htm. [Accessed: 24-May-2020].

[3]   NYU Spatial Data Repository, "2014 New York City Neighborhood Names," 2018. [Online]. Available: https://geo.nyu.edu/catalog/nyu_2451_34572. [Accessed: 23-May-2020].