

# Applied Data Science Capstone – Battle of the Neighborhoods

## Introduction/Business Problem

For many aspiring business owners, the decision of where to set down roots is crucial to the business strategy. This location impacts revenue, initial costs, operational costs, customer traffic and advertising. Depending on the city and business, there are also likely to be competitors nearby. Take New York City coffee shops, for example. Throughout NYC there are hundreds of coffee shops, many of them within close proximity to each other. For a potential coffee shop owner, this competition may pose a threat to the business. Why not build a coffee shop in a location where this venue type is scarce? It's not always simple to determine the optimal place to build.

Using Foursquare venue data, locations for a specific venue type can be mapped, and neighborhoods can be ranked by how many venues exist in a specific category, or even what percentage of venues are a specific type. In this project, Foursquare data will be used to map out and guide business-owners towards optimal neighborhoods to open coffee shops in New York City.

## Data

In order to map the coffee shops in NYC, I will combine Wikipedia data of New York City boroughs and neighborhoods with Foursquare location and venue data. I will map all coffee shops in NYC using latitude and longitude coordinates. I will utilize all venues available in the Foursquare dataset in order to calculate a percentage of venues that are coffee shops in each neighborhood, and I will rank neighborhoods based on percentage as well as the count of coffee shops.