**Finding A Location To Open a Crispy Creme**

Justin Chu

July 12, 2021

**1. Introduction**

**1.1 Background**

A few of my colleagues and myself are looking to start a business venture. We look to open a business close to our homes on Long Island. We’ve decided that opening a Crispy Creme donut shop is a mutual goal.

**1.2 Problem**

We need to look for an opportunistic area that can easily be accessed by other towns of Suffolk County and has few competitors nearby. Placing our business in this area will allow for the local townsfolk to add to our revenue daily and allow customers from other towns availability.

**1.3 Interest**

Obviously this is a personal business problem and is only of interest to my colleagues and I.

**2. Data Acquisition and Cleaning**

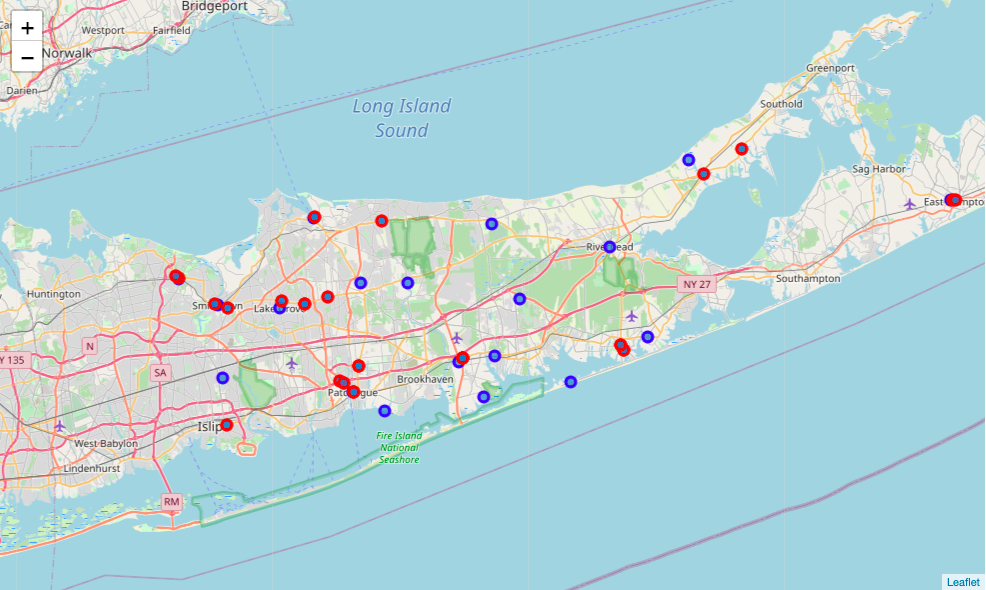
**2.1 Data Sources**

A list of Suffolk County towns can be found at <https://www.longisland.com/towns/>. FourSquare API was used for locating the geological coordinates of the towns and nearby venues.

**2.2 Data Cleaning**

As mentioned prior, the towns and venues of Suffolk County were located using the FourSquare API. Then two data sets were made of both the towns and venues. The venue data frame was filtered based on the type of venue, only keeping venues with the category of donut shop, coffee shop, baker', or dessert shop.

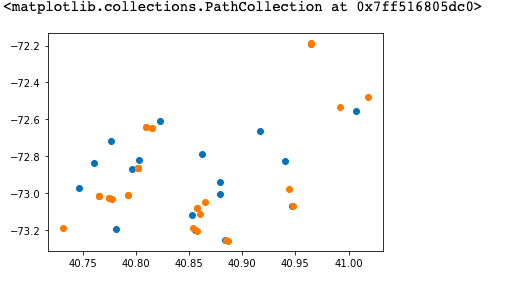
**3. Exploratory Data Analysis**

****

A follium map was created using the coordinates of the towns and venues. Blue points are the towns, red points are the nearby venues that offer donuts. Just this map alone offers great insight into the availability of areas with low competition.

**4. Predictive Modeling**

Although it is unnecessary, I used the KNearestNeighbor algorithm on the data collected. In the case that our business venture takes several more years, this model can be used to predict if new stores opening in any areas will incur more competition to our business.



The coordinates of the towns and venues are illustrated on the plot above. The blue dots are towns, orange dots are nearby venues.

A KNearestNeighbor classifier was then trained on the data above to predict the town it is opening in. After adding the venue to the dataframes we can quickly assess if the business falls under a category that will impede our business.