

1. Introduction/Business Problem

I am VietNameese , “Phở” is a Vietnamese soup consisting of broth.



In this project we will try to find an optimal location for a **Phở 24h restaurant**. Specifically, this report will be targeted to stakeholders interested in opening an **Phở 24h restaurant** in **New York**.

New York City is one of the most known cities-symbols of the USA. It is known as a global power city and one of the most populous cities of the world, with the population approaching to 8.5 million people. Since there are lots of restaurants in **New York** we will try to detect locations that are not already crowded with restaurants. We are also particularly interested in areas with no **VietNam restaurants** in city. We would also prefer locations as **close to city center** as possible, assuming that first two conditions are met.

We will use our data science powers to generate a few most promissing neighborhoods based on this criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

2. Description of the data

Based on definition of our problem, factors that will influence our decision are:

- * number of existing restaurants in the New York city (any type of restaurant)
- * number of and distance to Vietnam restaurants in the neighborhood, if any
- * distance of neighborhood from city center (Latitude and longitude coordinates are: 40.730610, -73.935242)

We decided to use regularly spaced grid of locations, centered around city center, to define our neighborhoods.

Following data sources will be needed to extract/generate the required information:

- * number of restaurants and their type and location in every neighborhood will be obtained using ****Foursquare API****
- * coordinate of New York center will be obtained using ****Google Maps API geocoding**** of well known New York location (Latitude and longitude coordinates are: 40.730610, -73.935242)

3. Neighborhood Candidates

Let's create latitude & longitude coordinates for centroids of our candidate neighborhoods. We will create a grid of cells covering our area of interest which is approx. 10x10 kilometers centered around New York city center.

Let's first find the latitude & longitude of New York city center, using information at <https://www.latlong.net/place/new-york-city-ny-usa-1848.html> i have latitude =40.730610 and longitude= -73.935242