**Capstone:** Finding the best neighborhood in New York City for opening an Indian Restaurant.

**Data Scientist:** Syed Khaleduzzaman

**Data Requirements:**

1. **List of NYC neighborhoods by zip code.**

List of NY Neighborhoods by postal code can be found in the zipatlas. The data can be scrapped from table in the webpage using **Beautiful Soup** package. <http://zipatlas.com/us/ny/new-york/zip-code-comparison/median-household-income.htm>

1. **Geographical location (Longitude, Latitude) data for NYC neighborhoods by zip code**

Longitude, Latitude information can be found in the ‘Location’ column in the data imported in step 1. The information is in [Longitude, Latitude] format. ‘**Location’ column in the data need to be broken down into tow columns using the comma (,) separator.**

1. **Average household income data for NYC neighborhoods by zip code**

Average household income information can be found in the ‘Avg. Income/H/hold’ column in the data imported in step 1.

1. **Population data for NYC neighborhoods by zip code**

Population information can be found in the ‘Population’ column in the data imported in step 1.

1. **We will need a list of all Restaurant Venues from all neighborhoods in NYC by Postal code.** We use 4SQUARE API ([https://api.foursquare.com](https://api.foursquare.com/)) to download all venues for all zip codes in NYC. We will then extract only restaurants include in our data set.
2. We will combine all information into a consolidated working data set to cluster and geo spatial map of the results showing the best neighborhood to open a restaurant in NYC.

Combining number of restaurants, population and household income into one dataset will help us clarify following research questionnaire to determine which zip code will be suitable to open an Indian restaurant.

* Which neighborhoods in NYC have clusters of like restaurants
* We will take top five zip codes for most clustered restaurants and append population to it.
* We will then add household income to the top five zip codes.
* We will then take a decision based on household income and population

**Capstone:** Finding the best neighborhood in New York City for opening an Indian Restaurant.

**Data Scientist:** Syed Khaleduzzaman

**Introduction:**

A friend of mine just migrated to New York City. He used to be the head chef in a famous restaurant in New Delhi, India. We both like to partner to open an Indian restaurant in NYC but not sure which neighborhood presents the best opportunity for us to be successful in the restaurant business.

**Problem Description:**

Which neighborhood should we open our Indian restaurant in NYC? We don’t like to take a chance since almost all of our life savings will be invested in the restaurant business. Before proceeding we like to make sure our restaurant will be profitable from the get go.

**Requirements:**

1. Our restaurant needs to be strategically located with the highest concentration of restaurants in NYC area. Patrons generally goes to areas to dine where multiple choices are available. People goes to shopping mall to buy cloths instead of local boutique store because people have more choices in the mall. Similar to that, people go out for lunch or dinner where clusters of restaurants exist instead of a standalone restaurant so they can choose among various restaurant choices.
2. It is the fact that financially affluent people go out frequently to dine at restaurant compare to people with lower income, therefore, we need to consider additional criteria before we can select a neighborhood. Not only we like to select a rich neighborhood but also a populous one so our restaurant can welcome as many patrons as possible.
   1. The average income for the neighborhood should be greater that the

national average.

* 1. Is the population greater than the comparing zip codes?

1. We need to confirm all the above assumptions before we can finalize a neighborhood to open our Indian restaurant using modeling and testing the data. By visualizing clusters of restaurants in NYC and compare that with income and population information for the neighborhood, we will be able to finalize our neighborhood selection.