The Battle of the Neighbourhood – Report

1. Introduction and business problem

The city of New York being the most populated city in United States has a lot of business opportunities and also a business friendly environment.

The market here is very competitive, the multicultural environment triggering many various cuisine restaurants being opened here.

Problem description:

Opening a restaurant is challenging because of the competitive environment, the city of NY being famous for its excellent cuisine.

Its food culture includes an array of international cuisines influenced by the city's immigrant history.

The following immigrant cultures are present here:

-Italian immigrants

-Jewish immigrants

-Central and Eastern Europe immigrants

-Chinese ,Asian & Indian restaurants

-Middle Eastern foods

It is obvious that surviving in such a competitive market it is important to have a strategic plan.

In order to decide the opportunity and location of the Italian restaurant many factors need to be taken into account:

-structure of the population

-city demographics

-the competitors

-segmentation of the borough

-untapped/saturated markets

Selecting the best neighbourhood to open a restaurant in New York is involving several steps:

- Collecting data regarding neighbourhoods, farmers market, demographics of the population and cuisine, and venues data.
- Understanding the similarities and differences between the neighbourhoods using Unsupervised K-means clustering algorithm; K-means clustering is a type of unsupervised machine learning, the goal of the algorithm is to find unlabelled groups in the data; data points are clustered based on feature similarity
- Analysing the results of the clustering process and business discussion regarding the opportunities
- Conclusions regarding the decision made

Choice of location is very important and is a prerequisite for the future success of the business.

2. Data

Data regarding New York city was analysed.

The following datasets were further used:

Data 1: Neighborhoods from New York city

- Neighborhood has a total of 5 boroughs and 306 neighbourhoods
- This dataset is already available on the web, the link is:

https://geo.nyu.edu/catalog/nyu_2451_34572

Data 2: Farmers market

- The data is available at the following link: https://data.cityofnewyork.us/dataset/DOHMH-Farmers-Markets-and-Food-Boxes/8vwk-6iz2

Data 3: Demographics and cuisine

Data was taken using web scraping from Wikipedia as from the below links:

https://en.wikipedia.org/wiki/New_York_City

https://en.wikipedia.org/wiki/Economy_of_New_York_City

https://en.wikipedia.org/wiki/Portal:New_York_City

https://en.wikipedia.org/wiki/Cuisine_of_New_York_City

Data 4: Venues

- Data was taken from foursquare.com
- New York city geographical coordinates data will be utilized as input for the Foursquare API, that will be leveraged to provision venues information for each neighborhood. We will use the Foursquare API to explore neighborhoods in New York City