

# **Analyzing the Neighborhoods in Chennai for Starting a Restaurant**

## **Applied Data Science Capstone Project**

Date of Submission: June 19, 2021 By: Karthik B

### **Introduction**

Chennai has a wide variety of food to offer, where the Idli Sambhar is a popular dish, which is served as breakfast or dinner. Apart from regular South Indian street food, the city's streets are also filled with several North Indian street food outlets, most of them established by North Indian migrants themselves. It has many restaurants which offer Idli Sambar, Dosa, Uttapam, Paniyaram, Jigarthanda and many more. Using data science concepts I want to find which area is the best to open a new restaurant.

### **Data Collection**

The data required for this project has been collected from multiple sources. A summary of the data required for this project is given below.

## **Neighborhoods Data**

The data of the neighborhoods in Chennai was scraped from [https://en.wikipedia.org/wiki/List\\_of\\_neighbourhoods\\_of\\_Chennai](https://en.wikipedia.org/wiki/List_of_neighbourhoods_of_Chennai). The data is read into a pandas data frame using the `read_html()` method. The main reason for doing so is that the Wikipedia page provides a comprehensive and detailed table of the data which can easily be scraped using the `read_html()` method of pandas.

## **Geographical Coordinates**

The geographical coordinates for Chennai data has been obtained from the GeoPy library in python. This data is relevant for plotting the map of Chennai using the Folium library in python. The geocoder library in python has been used to obtain latitude and longitude data for various neighborhoods in Chennai. The coordinates of all neighborhoods in Chennai are used to check the accuracy of coordinates given on Wikipedia and replace them in our data frame if the absolute difference is more than 0.001. These coordinates are then further used for plotting using the Folium library in python.

## **Venue Data**

The venue data has been extracted using the Foursquare API. This data contains venue recommendations for all neighborhoods in Chennai and is used to study the popular venues of different neighborhoods.