**Coursera Capstone Project – The Battle of Neighborhoods**

**Week 1**

**Introduction:**

* 1. Background:

Berlin is the capital and the chief urban center of Germany. After the end of World War II, Berlin was able to revive Germany’s economy through the production of products like textiles, clothing, porcelain, bicycles, and machinery. Berlin stands as Germany’s largest industrial town and is known to be the center of technological development and trade, as it attracts companies to set up facilities in the area.

The coordinates of Berlin, Germany are 52.5200 N, 12.4050 E

As of today, Berlin is divided into 12 boroughs as provided below:

* **Charlottenburg-Wilmersdorf** with a population of 319,628
* **Frederikshavn-Kreuzberg** with a population of 268,225
* **Lichtenberg** with a population of 259,881
* **Marzahn-Hellersdorf** with a population of 248,264
* **Mitte** with a population of 332,919
* **Neukolln** with a population of 310,283
* **Pankow** with a population of 366,441
* **Reinickendorf** with a population of 240,454
* **Spandau** with a population of 223,962
* **Steglitz-Zehlendorf** with a population of 293,989
* **Tempelhof-Schoneberg** with a population of 335,060
* **Treptow-Kopenick** with a population of 241,335
  1. Problem:

I would like to open up an Indian restaurant somewhere in Berlin. Which city would I benefit the most from?

* 1. Interest:

Anyone who may have a similar problem can easily modify and change the data that will allow the solution to be more accustomed to their problem

**Data:**

2.1 Data Sources:

In order to find a solution to our problem, we must be able to obtain data from the subjects below:

* Neighborhood data of Berlin from Wiki
* Python geocoder library to get geographical coordinates of neighborhood
* Foursqaure API calls to get venues within a 500 meter radius of a given neighborhood
* Foursquare API calls to get ranks and likes of restaurants by given venue id

2.2 Data Usage:

With the neighborhood location values/data I will be able to analyze Berlin’s geographical structure by using the folium library to place points by the longitude and latitude of the neighborhoods.

I will take the population and population density of each neighborhood into consideration while deciding which neighborhood the owner of the Indian restaurant will best benefit from.

By using the Foursquare API calls I will be able to classify the neighborhood by viewing the venues distribution and counts. This will allow me to find similarities in neighborhoods which will help me to choose a location for opening a new Indian restaurant. This list will also let me see the distribution of possible Indian restaurants in the area and if so, which neighborhoods are lacking any Indian establishments

I will also be viewing the ratings and likes of restaurants around each neighborhood. This data will show me what type of restaurants are preferred by the people in the areas.