The Battle of the Neighborhoods

Applied Data Science Capstone by IBM/Coursera

Introduction

In this project, We are trying to find an optimal location for a hospital. This report is targeted to stakeholders who are interested in opening a hospital in Ghaziabad, India.

When the stakeholder is planning to build a hospital in order to help people in emergency situations, it is better to compute distance between existing hospitals and find the areas with no or less hospitals. So we are trying to find location that are not already crowded with hospitals. We are also particularly interested in areas with no hospitals in the vicinity. We would also prefer locations as close to the city center as possible.

We will use our data science powers to generate a few most promising 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.

Data

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

- number of existing hospitals in the neighborhood.
- distance between hospitals in the neighborhood.

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

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

- number of hospitals and their location in every neighborhood will be

obtained using Foursquare API.

- latitude and longitude of an area will be obtained using geocoder python package.