CAPSTONE PROJECT – THE BATTLE OF NEIGHBORHOODS REPORT

CHAPTER – 1

INTRODUCTION

1.1 BACKGROUND

“Migration is an expression of the human aspiration for dignity, safety and a better future. It is part of the social fabric, part of our very make-up as a human family”

This quote of Ban Ki-Moon states the mere fact of people migrating to places. This carries us to the question: Do individuals move until they discover a spot to settle down where they really feel upbeat, or then again do our needs and needs change after some time, inciting us to in the long run leave a town we when considered home for another zone that will bring us fulfilment? Or on the other hand, do we over and over again move to another territory without knowing precisely what we’re getting into, compelling us to retreat in fear at the principal indication of uneasiness?

To limit the odds of this incident, we ought to consistently do legitimate research when arranging our best course of action throughout everyday life. Consider the accompanying elements when picking another spot to live so you don’t wind up burning through your important time and lucrative a move, you’ll end up lamenting. Wellbeing is a top concern when moving to another territory. On the

off chance that you don’t have a sense of security in your own house, you’re not going to have the option to appreciate living there.

1.2 LOCATION DATA PROVIDERS

Location data is information relating to the position of mobile phones, tablets, and laptops, or structures like historical monuments, buildings, and attractions. The most common attributes of location data are latitude and longitude (lat/long), usually expressed in coordinates which correlate to a geographical position. One of the most notable firms in the area information advertise, Foursquare currently runs a free area information and innovation stage for advertisers and designers. Utilizing Four square’s area information, brands can measure the effect of their media (utilizing approved stops) and find new crowds that they can focus with pertinent promoting, Foursquare additionally offers a Places API and a Pilgrim SDK, which empower area disclosure inside applications and sites, just as continuous area identification.

1.3 NEIGHBORHOOD SEGMENTATION AND CLUSTERING

In order to explore the neighborhoods, segment them, and group them into clusters to find similar neighborhoods clustering is used. Clustering is a form of unsupervised machine learning. K-Means clustering is the most common form of

clustering used. K-means can group data only unsupervised based on the similarity of objects to each other. There are various types of clustering algorithms such as partitioning, hierarchical, or density-based clustering. K-means is a type of partitioning clustering. It divides the data into k non-overlapping subsets or clusters without any cluster internal structure or labels. This means, it's an unsupervised algorithm.

1.4 PROBLEM DEFINITION

The wrongdoing measurements dataset of London found on Kaggle has violations in every ward of London from 2015 to 2016. The year 2016 we will considered and the information of that year which is really old data starting at now. The crime percentages in every precinct may have changed over time. This venture intends to choose the most secure district in London dependent on the all-out wrongdoings, investigate the areas of that ward to locate the 10 most regular settings in every area lastly group the areas utilizing k-mean clustering.

CHAPTER 2

DATA COLLECTION AND SCOURING

2.1 Data COLLECTION

The information obtained for this venture is a blend of information from three sources. The main information wellspring of the task utilizes a London wrongdoing information that shows the wrongdoing per precinct in London. The dataset contains the accompanying segments:

● lsoa_code: code for Lower Super Output Area in Greater London.

● district: Common name for London precinct.

● major_category: High level arrangement of wrongdoing

● minor_category: Low level arrangement of wrongdoing inside significant class.

● esteem: month to month detailed include of all out wrongdoing in given district

● year: Year of detailed checks, 2015-2016

● month: Month of detailed checks, 1-12

The second wellspring of information is scratched from a Wikipedia page that contains the rundown of London wards. This page contains extra data about the wards, coming up next are the sections:

● Borough: The names of the 33 London wards.

● Inner: Categorizing the ward as an Inner London district or an Outer London Ward.

● Status: Categorizing the ward as Royal, City or other district.

● Local power: The neighborhood authority allocated to the precinct.

● Political control: The ideological group that control the precinct.

● Headquarters: Headquarters of the Boroughs.

● Area (sq mi): Area of the ward in square miles.

● Population (2016 est)[1]: The populace in the ward recorded during the year 2016.

● Co-ordinates: The scope and longitude of the wards.

● Nr. in map: The number allotted to every precinct to speak to outwardly on a guide.

The third information source is the rundown of Neighborhoods in the Royal Borough of Kingston upon Thames as found on a Wikipedia page. This dataset is made with the help of Foursquare location data and preparatory utilizing the rundown of neighborhood accessible on the site, coming up next are segments:

● Neighborhood: Name of the area in the Borough.

- Borough: Name of the Borough.
- Latitude: Latitude of the Borough.
- Longitude: Longitude of the Borough.