Introduction

When a person is trying to find the best places to live, it's always a good idea to compare cities and if possible, to compare neighbourhoods to see if it suites your taste. After all, when you go to buy a car or a house or any big ticket item, you usually try out a few models or visit a few homes before you decide. The same tactic applies to finding the best places to live. It is always advisable to do it before you start planning your move or to help narrow down your choices.

When thinking about the best place to live or move to, lots of things are considered when trying to make a comparison between cities, towns, or neighbourhood. Some of these includes:

- Neighbourhood Comparison: This looks at neighbourhood comparison and helps one choose the best place to live within any given city. These sites allow you to see some pretty interesting facts about the various communities.
- Crime Rates: Here, the comparison is made to know the crime rates of two cities, then measures them both against the national statistics.
- Overall Comparison: This is a comparison of the same factors for each city, resulting in having a general overview of the two cities. Some of the popular factors include population, cost of living, average rent, crime rate, tax rates, and air quality.
- Cost of Living and Salary Comparison: This takes into account comparing salaries and cost of living within cities for a decision to be made. Some factors for this comparison includes statistics on food, housing, utilities, transportation and more. This is a useful way to find out if your salary will measure up in the new city.
- Compare Schools: This is helpful in finding the best school in a vicinity by doing a comparison between different places. It mostly takes into consideration test scores and teacher and student ratios, including the teacher's experience of the lists schools in the city of your choice.
- Neighbourhood Comparison: This looks at neighbourhood comparison and helps one choose the best place to live within any given city. These sites allow you to see some pretty interesting facts about the various communities .etc.,



Problem and Purpose of this Project

The data set includes the coordinates of the cities/neighbourhood in USA. However, it does not include the venues within these locations. If we had the venue information, we could easily find out more information about the neighbourhoods. For example, how many restaurants are there, are there parks or cinemas? What about banks and glossary stores? If all these information are known, we could better understand or make educated decision about where to move or relocate to.

Hence, the purpose of this project is to, algorithmically, find a way to use the location coordinates and tag each data point into a neighbourhood in two Counties in North Carolina-Wake County and Mecklenburg County. The algorithm used is k-means clustering. The main idea is to determine neighbourhood with venues clustered around each other so that one can make a decision on the right neighbourhood to chose based on the proximity of amenities and venues to each other.

NOTE: While all of these analyses are useful for comparing the neighbourhoods, there is nothing like visiting the actual city, seeing the neighbourhoods, and speaking with residents. If it's possible, an in-person visit is highly recommended before making a big move or relocating decision.

k-means Clustering

The k-means clustering algorithm is an unsupervised clustering technique searches for a pre-determined number of clusters within an unlabelled multidimensional dataset. It accomplishes this using a simple conception of what the optimal clustering looks like:

- The "cluster centre" is the arithmetic mean of all the points belonging to the cluster.
- Each point is closer to its own cluster centre than to other cluster centres in the dataset.

The two assumptions above are presumably the basis of the k-means model.