|  |
| --- |
| *Restaurant Recommendation* |
|  |



**Muhammad Faizan Rehman**

**27 April, 2023**

Table of Contents

|  |  |
| --- | --- |
| **Topic** | **Page number** |
| Introduction | 2 |
| Data | 4 |
| Methodology | 8 |
| Result | 12 |
| Discussion | 12 |
| Conclusion | 13 |

1. **Introduction:**

**Problem Description:**

Bangalore, the capital and largest city of Karnataka, India, is renowned for its diverse cuisine that reflects the social and economic diversity of the city. However, for individuals who frequently travel and change locations, finding the best restaurant that offers the most delicious food at a reasonable cost can be challenging. As food is a critical factor in rating a trip and recommending it to others, identifying the right restaurants is essential in attracting people from around the world to try the city's best food.

To address this issue, the manager of Company has allocated a project to me to develop a recommender system that can suggest new restaurants based on rankings compared to previously visited ones. This system will also provide solutions to questions such as the types of food available in the restaurants in a particular area, the nearest restaurant with a good rating, and the number of similar restaurants in the vicinity and their specialties.

The goal of the recommender system is to provide comprehensive answers to these questions and uncover all the perspectives of managing recommendations. Specifically, the system will offer information on the types of restaurants available in a particular area, the location of similar restaurants based on a preference for a specific type of cuisine, and the ranking of different restaurants concerning individual preferences. By addressing these questions and developing a reliable recommender system, individuals who frequently travel to Bangalore will have access to a powerful tool that can help them make informed decisions about where to eat, ultimately enhancing their travel experiences.

In conclusion, the development of a recommender system that can accurately suggest restaurants based on individual preferences is essential in helping individuals make informed decisions about where to eat, particularly when they frequently travel to Bangalore. By leveraging data-driven insights, the system can provide comprehensive and reliable recommendations, ultimately enhancing the culinary experiences of travelers and promoting the city's diverse cuisine to people from around the world.

The target audience for this project is not limited to frequent travelers but includes everyone who is exploring new places or seeking similar ones. The recommender system can cater to different preferences, such as those addicted to specific categories of food or those who prefer the highest-rated restaurants nearby. With the evolution of new food categories and the increasing popularity of hybrid food, access to a vast number of food varieties is essential.

The success rate of this project is expected to increase over time, as the recommender system utilizes machine learning to provide a personalized experience based on individual likes and dislikes. The system can also keep track of evolving food trends and recommend restaurants accordingly. Ultimately, the recommender system will serve as a personal guide, enhancing the user's experience by providing comprehensive answers to questions about the types of restaurants available in a particular area, the location of similar restaurants based on a preference for a particular type of food, and the ranking of different restaurants with respect to individual preferences.

1. **Data :**

In order to build an effective recommender system for restaurants, we need a large and diverse dataset that includes information on the following factors:

1. Geographical Coordinates: The latitude and longitude of each restaurant location will help us create a map displaying all the restaurants with their labels. This will allow users to easily locate restaurants near them and make informed decisions about where to eat.
2. Population of the Neighborhood: The population of the neighborhood where the restaurant is located is a key factor in determining a restaurant's potential for growth and customer base. Restaurants located in more densely populated areas are likely to have more customers and be rated higher as they are accessed by a diverse range of people with different tastes.
3. Average Income of the Neighborhood: Income is another important factor that influences a restaurant's success. People who earn more than the average income in a neighborhood are more likely to spend more on dining out, which could increase the restaurant's revenue. However, this is not always the case, and there are several other factors that contribute to the popularity and success of a restaurant.

In addition to the above factors, we also need data on the types of food available in each restaurant, the price range, ratings and reviews, as well as any other relevant information that could help us build a comprehensive and accurate recommender system.

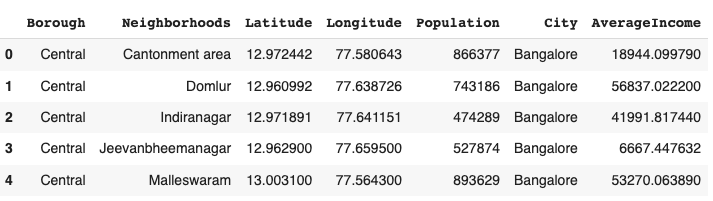
Collecting and analyzing such a large and diverse dataset may require advanced data mining and machine learning techniques. However, with the right tools and expertise, we can build a robust and reliable recommender system that can help people make informed decisions about where to eat and enhance their overall dining experience.

**Data collection:**

* Acquiring geographical coordinates for each neighborhood in Bangalore posed a challenge as open-source data websites, including Wikipedia, India government websites, and census report websites, did not provide such information. Therefore, the decision was made to utilize Google Maps API to retrieve the latitude and longitude data, but due to limitations on the number of calls that could be made with a free account, the process of obtaining the location data for all the neighborhoods would take approximately 15 to 20 days.

The initial step in obtaining the neighborhood data was to scrape the list of neighborhoods using the Beautiful Soup 4 library from the Wikipedia page entitled "List of neighborhoods in Bangalore." The table headings were used as the boroughs, and the data became the neighborhoods. As Bangalore consists of eight boroughs and 64 neighborhoods, each neighborhood's corresponding latitude and longitude data were manually retrieved by conducting a Google search. The resulting data frame containing the geographical coordinates for each neighborhood in Bangalore is presented

below.



* Gathering population data by neighborhood can usually be done with relative ease, given that it is readily available through various government sources. However, in the case of Bangalore, this proved to be a challenge. Despite searching through opensource data websites such as Wikipedia, India gov website, census report websites and others, only a limited amount of population data was found for a few select cities. As a result, for the majority of Bangalore's neighborhoods, population data had to be assumed, and may therefore be inaccurate. It is important to note that while this may affect the accuracy of the model, the primary goal of this project is to demonstrate the functionality of the recommender system. Gathering population data by neighborhood can usually be done with relative ease, given that it is readily available through various government sources. However, in the case of Bangalore, this proved to be a challenge. Despite searching through opensource data websites such as Wikipedia, India gov website, census report websites and others, only a limited amount of population data was found for a few select cities. As a result, for the majority of Bangalore's neighborhoods, population data had to be assumed, and may therefore be inaccurate. It is important to note that while this may affect the accuracy of the model, the primary goal of this project is to demonstrate the functionality of the recommender system. To gather information about nearby venues, the Foursquare API was utilized. This API allows for the identification of the nearest venues within a specific radius (in this case, 500 meters), providing details such as their coordinates, names, and locations. This data was used to form clusters in order to make recommendations for similar venues to users.
* Collecting income data by neighborhood in Bangalore was not possible as it is not readily available on open-source data websites. However, I was able to find income data for the main city on a website that lists Indian cities by GDP per capita. I have assumed that this data is representative of the income levels in the neighborhoods of Bangalore, although it may not be entirely accurate. It is important to note that this project is meant to demonstrate the workings of a recommender model and is not intended to be a comprehensive study of income levels in Bangalore.



1. **Methodology:**

**Exploratory Analysis:**

Top of Form

Bottom of Form

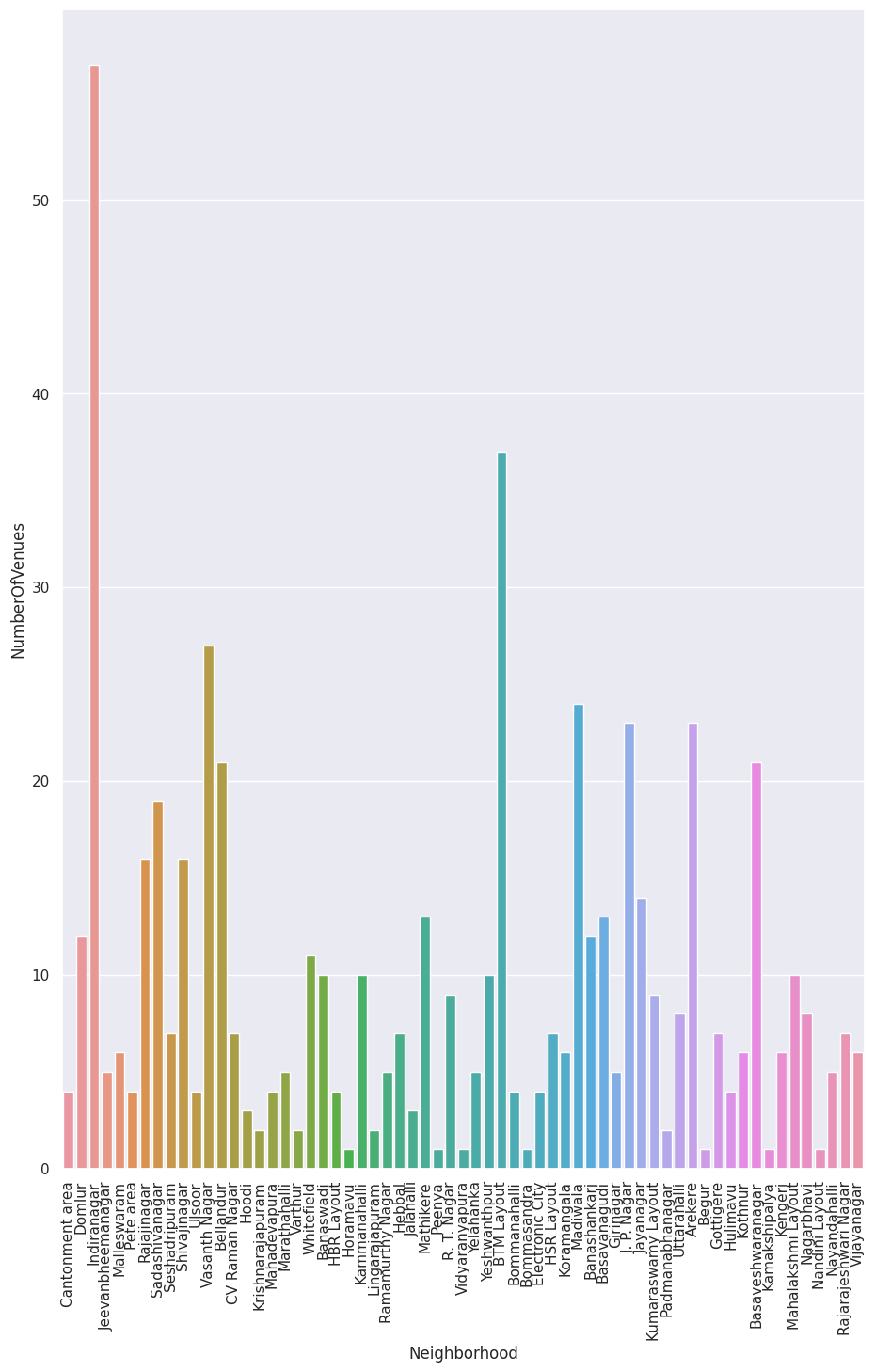
When working on a project that involves collecting data from various sources and combining it into a single dataset, it is essential to first explore the current state of the dataset. This task can be challenging as it requires identifying all the relevant features that need to be collected.

During my exploration of the dataset, I encountered several obstacles, including the absence of readily available data on open-source platforms. I had to resort to using various methods to collect the necessary data, such as web scraping and API calls. However, despite the challenges, exploring the dataset was critical in gaining initial insights and identifying the features required to achieve the project's objectives.

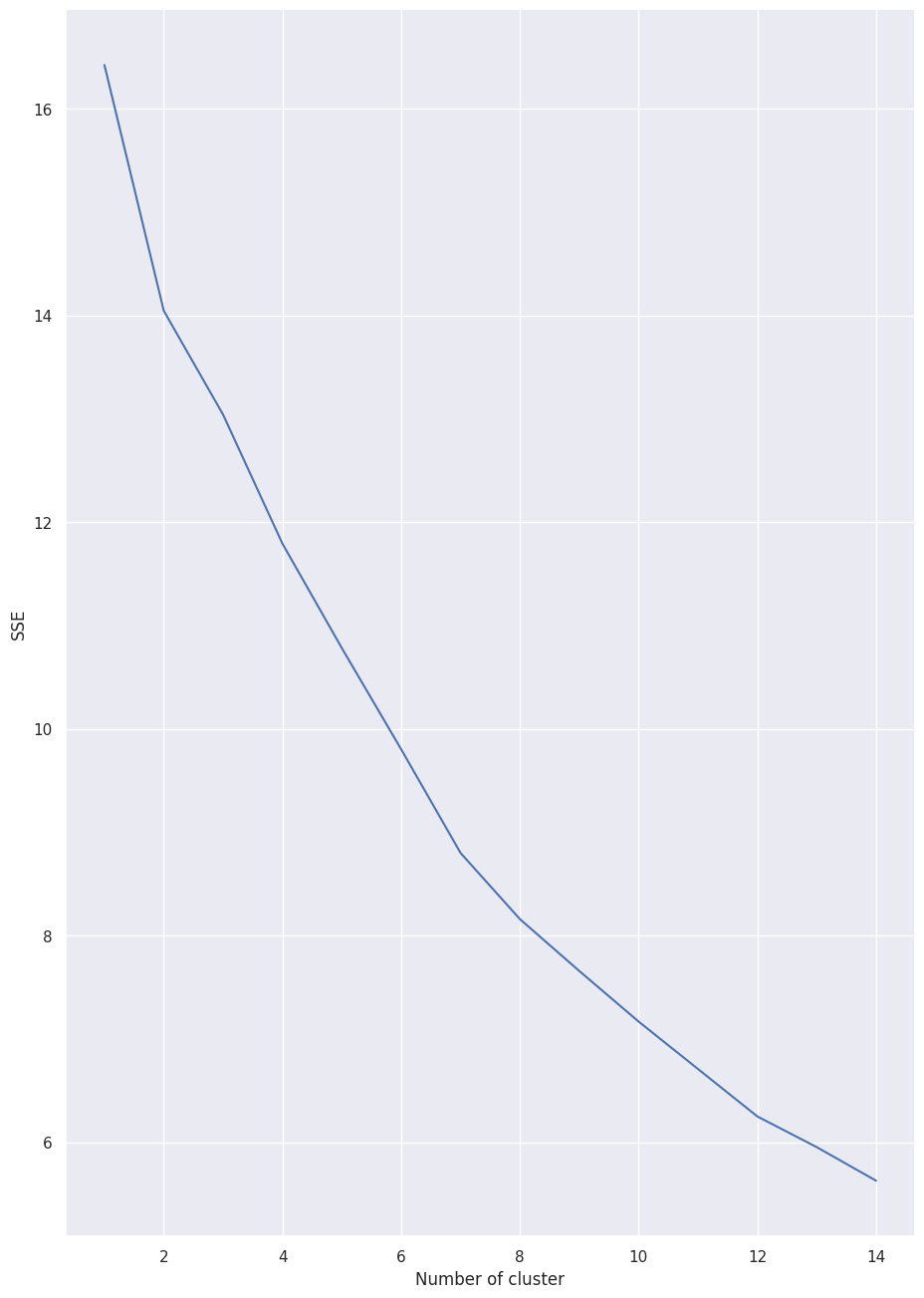
Through the exploration process, I discovered that Indiranagar had the highest number of venues, while Varthur had the least. This finding highlights the importance of data exploration, as it provides valuable information that can help guide subsequent data analysis and modeling.

Top of Form

Bottom of Form



I also plotted an elbow graph to explore the range of values for n clusters and determine the optimal number for clustering.

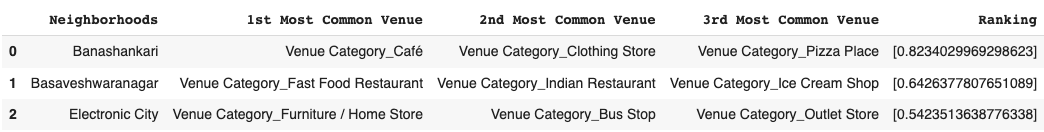


**Inferential Analysis:**

The most important factors for building the recommender system were population and income, as they have a non-linear relationship in the dataset. Inferential analysis was conducted to understand this relationship, which showed that an increase in population does not necessarily result in an increase in average income, and vice versa, as seen in the graph.

1. **Result:**

The recommender system generates a list of top restaurants and popular venue items for users to enjoy based on the input neighborhood. In our simulation, we used 'Whitefield' as the input neighborhood to recommend similar neighborhoods with similar characteristics.



**5. Discussion:**

Based on the analysis conducted, it was observed that there exists a nonlinear relationship between income and population. Therefore, it is essential to perform an inferential approach to identify and understand the relationship between different features. During the clustering process, it is crucial to ensure that similar neighborhoods are correctly assigned to the appropriate cluster, as this will lead to better insights and decision-making.

It is important to note that the choice of the number of clusters can significantly impact the clustering results. Overfitting or underfitting of the data may occur if the number of clusters is not chosen appropriately. To address this issue, the analysis of the number of clusters can be conducted, as demonstrated by the elbow graph in the Methodology section. This will help to identify the optimal number of clusters and ensure that the clustering results are reliable and meaningful.



1. **Conclusion:**

The recommender system is a sophisticated model that takes into account various factors, including population and income, to recommend nearby venues using the Foursquare API. This system leverages the power of data to provide efficient and effective recommendations for users. While the system's efficiency may decrease as the amount of data increases, its accuracy will improve as it continues to learn from user behavior and preferences.

The primary objective of the recommender system is to help users satisfy their hunger by providing them with the best recommendations that meet all their needs. By analyzing various factors, such as the type of cuisine, price range, location, and ratings, the system can make personalized recommendations that match users' preferences and expectations. Ultimately, the recommender system aims to enhance the overall dining experience of users and provide them with the convenience and satisfaction they desire.