Coursera Capstone Project

The Battle of the Neighborhoods

Finding best location to open an ATM in Bangalore, India

IBM DATA SCIENCE PROFESSIONAL SPECIALIZATION BY COURSERA

BY: PADHMA M

TABLE OF CONTENTS

1. INTRODUCTION

Bangalore is one among the fastest growing cities in the world. Bangalore is referred to as the Silicon Valley of India because of its role as the nation's leading Information Technology exporter .It has a population of over ten million, making it a megacity and third most populous city and fifth most populous urban agglomeration in India. Being a demographically diverse city, the needs of the residents are also increasing rapidly. Hence, any new organization or an existing one should keep up with their pace in supplying the needs of the customers.

2. BUSINESS PROBLEM:

Our customer is ABC Bank, which is an International Bank and also a market leader. They have a wide variety of customers all over the country, major of them residing in Metropolitan cities. ABC Bank has received ample amounts of complaints from residents of Bangalore that there aren't sufficient amount of ATM's. Given the extremely large population and the population of the city, our customer wants to identify the best neighborhood area to open more ATM covering the majority of the population. The problem statement will be: **Which neighborhood is most densely populated and has lesser number of ATM's?**

3. DATA:

The data to be used in this project is not readily available. Hence, the data has been obtained from various sources such as

- **Foursquare**, which is a local search-and-discovery mobile app which provides search results for its users. The app provides personalized recommendations of places to go near a user's current location based on users' previous browsing history and check-in history.
- **Wikipedia,** which has the details about the neighborhoods in Bangalore. https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Bangalore
- The geographic coordinates of each location have been obtained through **Geopy,** which makes it easy for Python developers to locate the coordinates of addresses, cities, countries, and landmarks across the globe using third-party geocoders.
- The population data about each neighbourhood has been obtained from: https://www.ichangemycity.com/assembly-constituencies/mahalakshmi-layout https://www.census2011.co.in/census/district/242-bangalore.html

3.1 DATA DESCRIPTION:

As we have to explore and identify the neighborhoods in the city of Bangalore, the Bangalore neighbourhood data is the crucial data for this project. The data about each neighbourhood is not readily available, hence we have to scrape the Wikipedia page and obtain the data. The data also needs information about coordinates which makes it easier for us to cluster the neighborhoods. In order to obtain the coordinates, we make use of geopy library in Python. We also need information about each neighbourhood which will make it simpler to recognise which area has least number of ATM's. Hence, the details about each venue is obtained through FourSquare API. The population about each neighbourhood will let us know which neighbourhood is more preferable. Hence, the demographic data about each neighbourhood is also obtained.