Coursera Capstone Project
IBM Applied Data Science Capstone

M S Ishwarya

29-Nov-2019

Overview

- Introduction
- Business Problem – Background
- Problem Statement
- Target Audience
- Data requirement
- Proposed Methodology
- Results
- Discussion
- Recommendation

Introduction

- This presentation illustrates the capstone project carried out for the course “Applied Data Science Capstone”
- Accordingly, a business problem has to be proposed for which can have a Data Science based solution

Business Problem- Background

- Investors and Property Developers is keen on projects that has high Return of Investments
- Given a city, it is helpful to understand the importance of the location when considered for launching a project
- Further, the importance in the city as well as existing venues are generally non-uniformly distributed.

Problem Statement

- Finding a suitable location in metropolitan city namely Chennai, India for opening a new project namely Shopping Mall

Target Audience

- The stakeholders of this project are of two folds:

a) Property developers or investors on a shopping mall

b) Retailer/ vendors for a stall in shopping mall

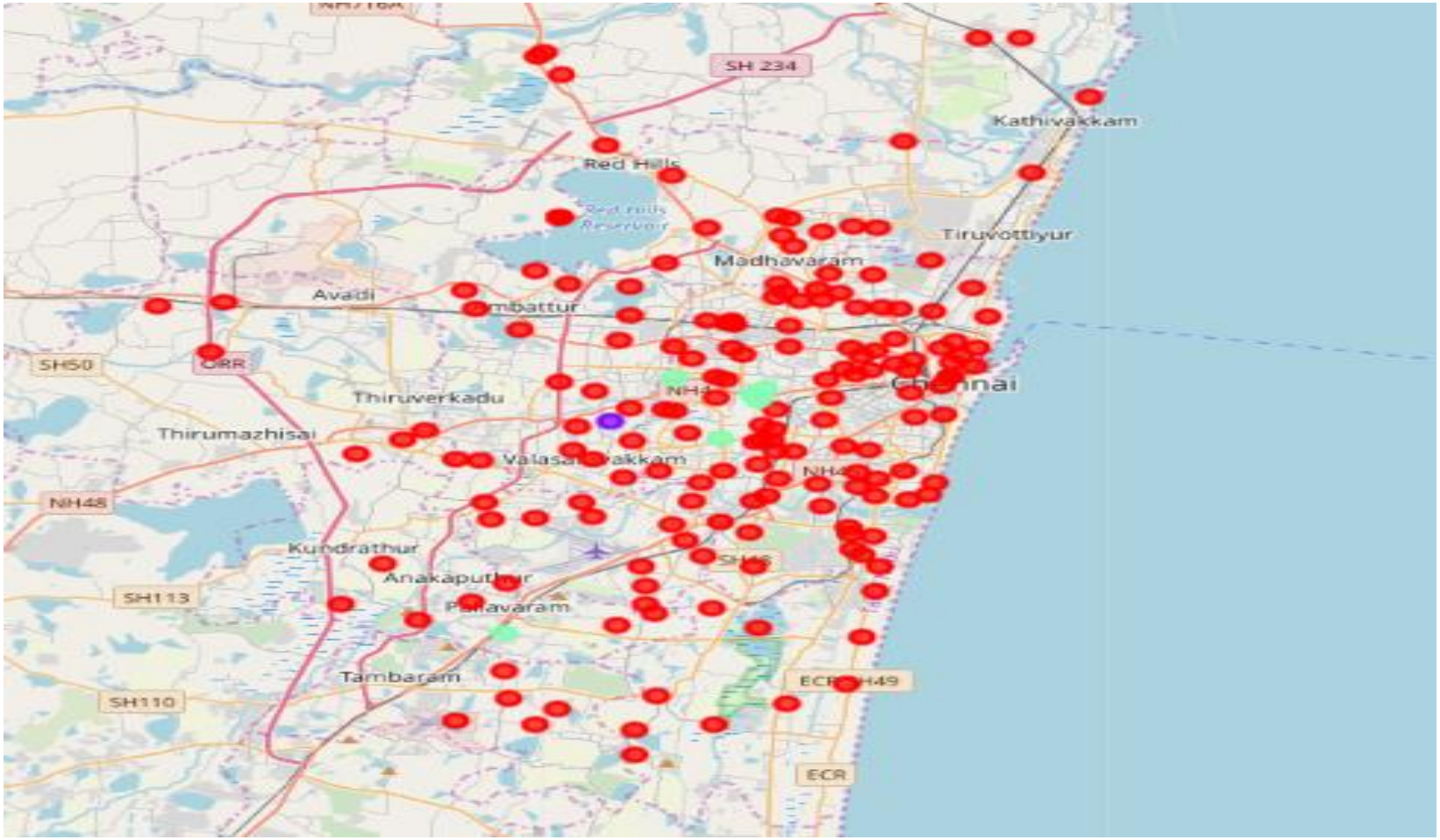
Data Requirement

- **Wikipedia:** List of neighbourhoods in Chennai. Link to source: https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Chennai
- **GeoPy:** Further, these neighbourhoods are analysed based on the geographical coordinates that are inside certain radius. For this, we make use of Geopy package.
- **Four Square API:** These coordinates based neighbourhoods are further analysed from the perspective of venues, in particular, shopping mall. For this, we make use of Four Square API.

Proposed Methodology

- Collecting Data regarding the neighbourhoods in Chennai (Source- Wikipedia)
- Collecting the geographical coordinates based information of the neighbourhoods in Chennai (Source- Geopy)
- Selecting all the neighbourhoods with a specified radius
- Identification of all types of venues with that radius
- Identification of shopping malls with in the radius
- Clustering the neighbourhoods based on the distribution of shopping malls in them
- Find the best cluster that illustrates the potentiality of neighbourhood
- Provide recommendation

Results


SH 234
Kathivakkam
Red Hills
Tiruvottiyur
Madhavaram
Avadi
SH50
ORR
Thiruverkadu
NH4
Thirumazhisai
NH48
Kundrathur
SH113
Tambaram
SH110
ECR
NH49

Results

```
In [61]: df_merged_1.loc[df_merged_1['Cluster Labels'] == 1]
```

Out[61]:

	Neighborhood	Shopping Mall	Cluster Labels	Latitude	Longitude
83	Madhya Kailash Junction	0.166667	1	13.060537	80.176599

```
In [62]: df_merged_1.loc[df_merged_1['Cluster Labels'] == 2]
```

Out[62]:

	Neighborhood	Shopping Mall	Cluster Labels	Latitude	Longitude
22	Chinnakudal	0.043478	2	13.07444	80.22167
170	Thirumangalam, Chennai	0.041667	2	13.08281	80.19703
156	Shenoy Nagar	0.031250	2	13.07732	80.22498
29	Chromepet	0.062500	2	12.95234	80.14411
7	Aminjikarai	0.050000	2	13.07139	80.22256
184	Vadapalani	0.031250	2	13.05228	80.21120

```
In [ ]:
```

Discussions

- The most potential locations for launching a shopping mall project are :

- **First choice:** Madhya Kailash Junction (cluster 1)

- **Second choice:** Chinnakudal/ Thirumangalam/ Shanoy Nagar/ Chromepet/ Aminjikari/ Vadapalani (cluster 2)

Recommendations

Cluster no. 1 (Madhya Kailash Junction) is the most potential candidate for launching a shopping mall considering its neighbourhoods impact and geographical coordinates.

Thank You