

# Capstone project- REPORT

## Contents:

### 1. Introduction Section:

1.1 Discussion of the "background situation" leading to the problem at hand:

1.2 Problem to be resolved

1.3 Audience for this project.

### 2. Data Section:

2.1 Data of Current Situation (current residence place)

2.2 Data required to resolve the problem

2.3 Data sources and data manipulation

# 1. INTRODUCTION

We now discuss the business problem and the audience who will be interested in this project.

## 1.1SCENARIO AND BACKGROUND

I have chosen CHANDIGARH,INDIA (my hometown) as the location to be used as current location for this project.

Being a growing town with many breweries, malls, gyms, clubs, restaurants, shops, cafes opening with a rapid pace, my hometown surely serves a great place to live

## 1.2PROBLEM TO BE RESOLVED

Many youngsters are migrating to foreign countries in search of jobs,better education and thus want to find an apartment atleast similar to one they used to live.

Thus the problem is to find similar apartment in MANHATTAN,NEW YORK,USA with similar amenities which were around in chandigarh

*-as a reference I have included the map of venues in chandigarh.*

## 1.3AUDIENCE FOR THIS PROJECT

I think that this project will be useful to every person considering moving to foreign countries similar to USA ,basically anyone migrating to developed nation.

Also I think this will help travel agencies partner with appropriate entities offering accommodation, homestays, lodging, etc.

Also this will surely develop my data science skills

*The use of foursquare data and mapping techniques along with data analysis and visualization will help solve the core problem*

## 2. DATA SECTION

We now describe the data and its sources used in this project

### 2.1 DATA OF CURRENT PLACE(CHANDIGARH HERE)

I currently reside in neighbourhood of Chandigarh in city Panchkula, Haryana, India.

It is directly connected to Chandigarh and shares its borders with it.

Data from FOURSQUARE will be used to identify the main and appropriate venues near Chandigarh.

This data will be superimposed on the map of Chandigarh, INDIA and thus will serve as reference to find desired location in MANHATTAN, NY.

### 2.2 DATA NEEDED TO SOLVE THE PROBLEM

To make the best prediction of the required apartment in Manhattan, data we need:

- Information of the neighbourhoods in Manhattan with their geodata (coordinates-latitude and longitude)
- Information of the apartments listed for renting out in Manhattan
- List of basic facilities provided at each neighbourhood (top 5)
- List of metro/bus stations near and in each neighbourhood.

### 2.3 DATA MANIPULATION AND DATA SOURCES

Data for Manhattan neighbourhoods will be taken directly from Capstone week-3 lab since it was well cleaned. Also clustering of the data will be shown accordingly

Data of the venues around the current residence i.e. Chandigarh will be gathered from the FourSquare API.

Data of other amenities , places of rent, of metro stations, etc will also be derived from foursquare API.

## 2.4 INFORMATION IN DATA

The data of venues around chandigarh extracted from foursquare consists of information of venue and we thus clean it to get the NAME OF VENUE, CATEGORY OF VENUE, LATITUDE AND LONGITUDE OF VENUE

Coordinates will be used to map the data on the map of Chandigarh.

The data of manhattan neighborhoods consists of Borough name, neighbourhood name , postal code , and coordinates.

The data will be clustered based on the :

- Manhattan subway metro locations
- Manhattan map of places for rent
- Combined maps of Manhattan rent places with subway locations
- Combined maps of Manhattan rent places with subway locations and venues clusters