The Battle of Neighborhoods – week 1

30st April 2020

1.Introduction

1.1 BACKGROUND

According to a survey conducted by Indian government 23% of India's population runs their own business. When a person wants to choose a city of india to start his own food business, he needs to go through various factors like city population, most visited place in the city, etc.,

By analysing various cities with respect to the population, most visited places, transportation facilities, types of restaurants/ food chains, a person can finalise his business location and the cuisine.

1.2 Problem Statement

The dataset consists of the latest population of various cities in India as of 2019. The problem statement is to find out the best cities in India when a person desires to migrate to a city where there is a moderate population and to start his own food business.

This project aims to help the person to choose a city with moderate population and to go through the most common visited venues in that city to choose the best city and choose a best suited cuisine accordingly.

2.Data Acquisition and Pre-Processing

2.1 Data Collection and description

The dataset used in this project is downloaded from the https://simplemaps.com/data/in-cities which consists of city wise India's population data along with the data of location coordinates and migrated population.

The data set consists of 7 columns

1. **City:** Name of the city
2. **Lat:** The city's Latitude
3. **Lng:** The city's Longitude
4. **Country** : India
5. **Iso2** : India's location code
6. **Admin** : state name
7. **Capital**: category of city
8. **Population** : population of city
9. **Migrated population**: the population migrated to that city

2.2 Data Cleaning

The data is downloaded from the url of given website as a dataframe – the raw data before data preprocessing looks like in the image below

	city	lat	lng	country	iso2	admin	capital	population	population_proper
0	Mumbai	18.987807	72.836447	India	IN	Mahārāshtra	admin	18978000.0	12691836.0
1	Delhi	28.651952	77.231495	India	IN	Delhi	admin	15926000.0	7633213.0
2	Kolkata	22.562627	88.363044	India	IN	West Bengal	admin	14787000.0	4631392.0
3	Chennai	13.084622	80.248357	India	IN	Tamil Nādu	admin	7163000.0	4328063.0
4	Bengalūru	12.977063	77.587106	India	IN	Karnātaka	admin	6787000.0	5104047.0

- After the **preprocessing**
    1. Removing NA values
    2. Removing unnecessary columns
    3. Creating new columns for migrated population and location(combination of city, state and country)
    4. Label encoding the categorical value in capital column

        1. Delhi – Capital of India – 3 (category after encoding)

        2.Ahmedabad – minor – 2

        3. Hyderabad– Capital of state – 1
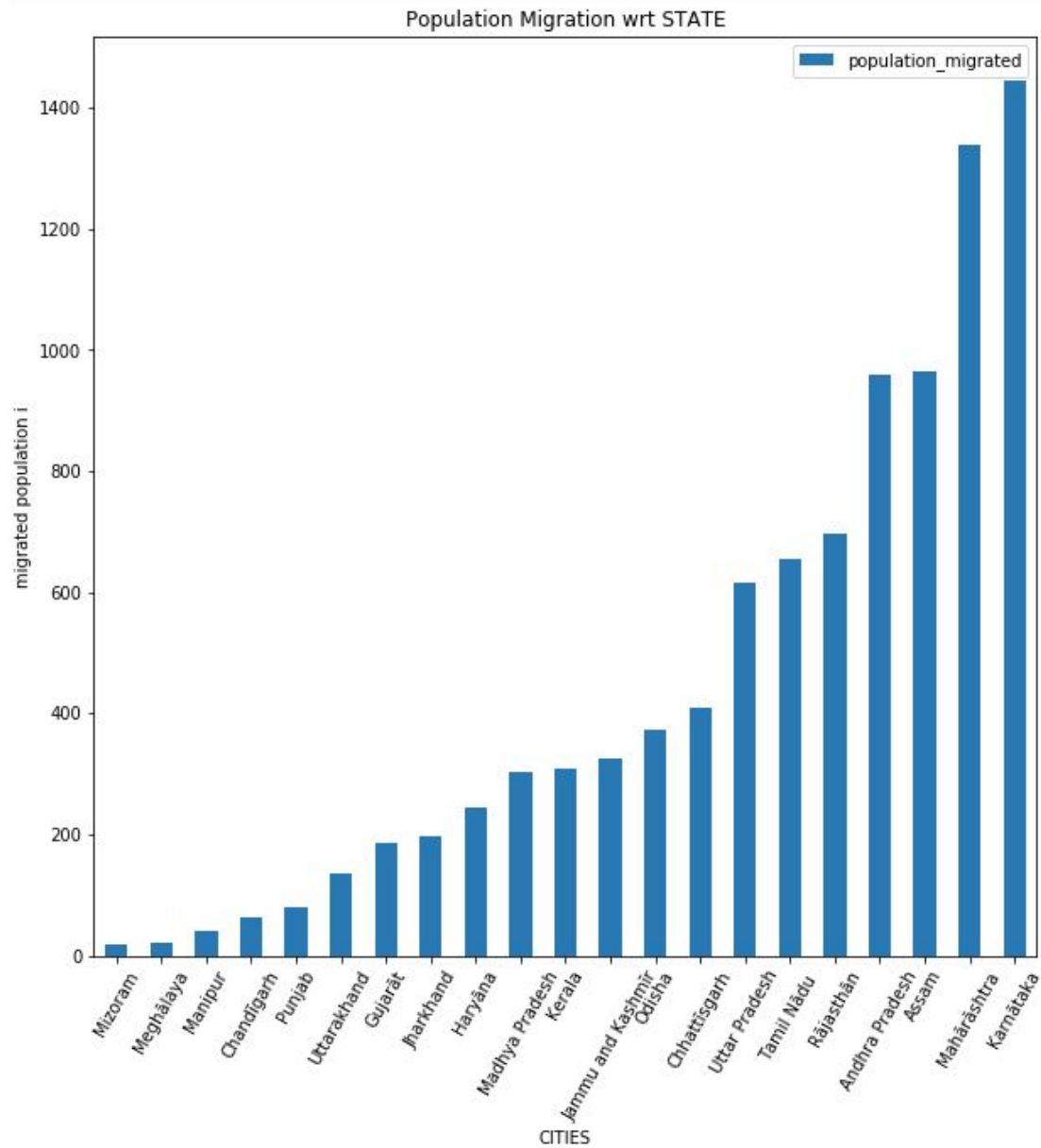
        4. Hauora – non capital city – 0

	city	lattitude	longitude	country	state	capital	population	population_proper	population_migrated	Location
0	Mumbai	18.987807	72.836447	India	Mahārāshtra	1	18978000.0	12691836.0	6286164.0	Mumbai,Mahārāshtra,India
1	Delhi	28.651952	77.231495	India	Delhi	1	15926000.0	7633213.0	8292787.0	Delhi,Delhi,India
2	Kolkata	22.562627	88.363044	India	West Bengal	1	14787000.0	4631392.0	10155608.0	Kolkata,West Bengal,India
3	Chennai	13.084622	80.248357	India	Tamil Nādu	1	7163000.0	4328063.0	2834937.0	Chennai,Tamil Nādu ,India
4	Bengalūru	12.977063	77.587106	India	Karnātaka	1	6787000.0	5104047.0	1682953.0	Bengalūru,Karnātaka,India
5	Hyderabad	17.384052	78.456355	India	Andhra Pradesh	1	6376000.0	3597816.0	2778184.0	Hyderabad,Andhra Pradesh,India

- After filtering data according to the requirements
    1. Filtering Moderate populated cities
    2. Filtering high migrated cities

	city	lattitude	longitude	country	state	capital	population	population_proper	population_migrated	Location
164	Tonk	26.168672	75.786111	India	Rājasthān	0	181734.0	151331.0	30403.0	Tonk,Rājasthān,India
162	Bhusāval	21.043649	75.785058	India	Mahārāshtra	0	183001.0	172366.0	10635.0	Bhusāval,Mahārāshtra,India
161	Nandyāl	15.477994	78.483605	India	Andhra Pradesh	0	188654.0	165337.0	23317.0	Nandyāl,Andhra Pradesh,India
156	Proddatūr	14.750200	78.548129	India	Andhra Pradesh	0	197451.0	177797.0	19654.0	Proddatūr,Andhra Pradesh,India
150	Ongole	15.503565	80.044541	India	Andhra Pradesh	0	202860.0	172872.0	29988.0	Ongole,Andhra Pradesh,India

- Data Visualisation

1. Data of the states with moderate population

Population Migration wrt STATE

2. Data of the cities migrated population greater than average

States with migrated population greater than average