**Quality of Living assessment for Indian Cities**

**Introduction**

An assessment of quality of life in any place plays an important role in many important decisions people make. One of the key reasons of wanting to know this information is to make right career choice and ask for a compensation that is as per the standards of the city. While assessment of quality of life across various major cities in the worlds is available, courtesy efforts of international agencies like Mercer, such that is not available for smaller cities at country level. In this project our focus will be on assessing the quality of living across major cities in India.

As per Mercer’s quality of living reports, the assessment of quality of living in any place depends on the following availability of factors :

* Consumer goods
* Economic environment
* Housing
* Medical and health considerations
* Natural environment
* Political and social environment
* Public services and transport
* Recreation
* Schools and education
* Socio-cultural environment

In a population heavy country like India, another factor is per person availability of the above mentioned facilities. Thus another factor that may play an important role in defining quality of life would be population of the city.

In this project our objective is to :

“Classify major Indian cities based on different levels of quality of living offered by them to a citizen”

**Data Requirement**

To achieve this objective the data required would be as follows :

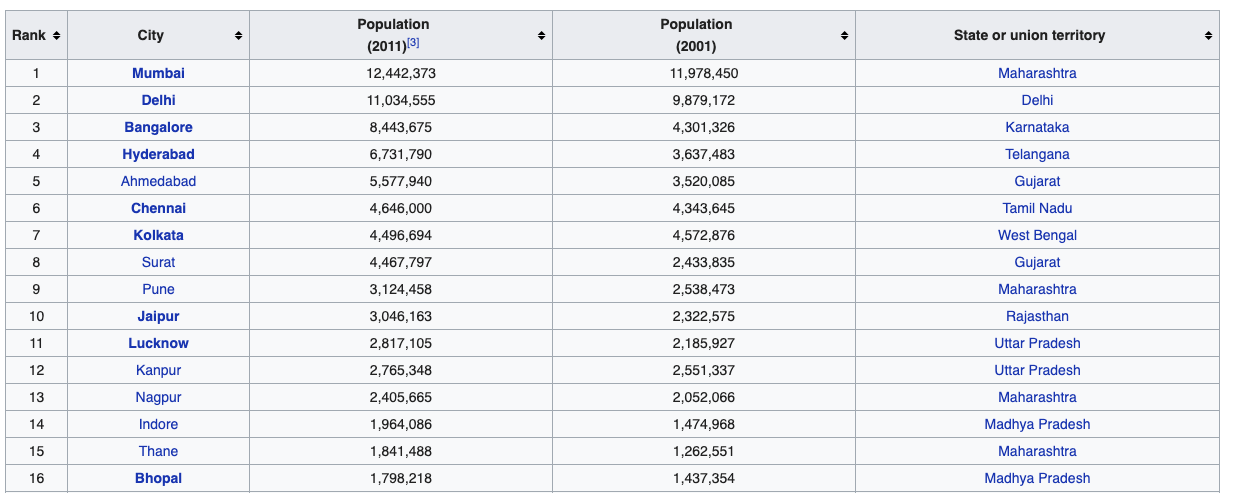
1. List of Indian Cities

To start with we will need a list of various Indian cities that we will be classifying under this project. Such list is available on the following Wikipedia page :

<https://en.wikipedia.org/wiki/List_of_cities_in_India_by_population>

This page gives a list of 300 cities of India along with respective populations as per 2011 census and the states these cities belong to.

A sample screen shot is given below :



We will scrap this data from the Wikipedia page using a python code. For the purpose of this project we will limit our analysis to first 50 cities mentioned in the list.

1. Latitude & Longitude of the cities

We will use the geocoder package in Python to extract latitude and longitude of each city.

1. Foursquare data

Upon entering the details of each city in the Foursquare API we get a list of venues in that city. The venues have different categories such as cafes, diners, schools, parks, metro station, libraries etc. We will use this data and classify the results obtained for each city into following categories. What all venues should be categorised under each of the following categories will be decided after content analysis of all the unique category of venues obtained for all the cities.

* *Recreation*: No of parks, cinema halls, cafes, theatres, restaurants, community centres, zoos etc. for each city.
* *Schools & Education* : No of schools, colleges, universities, learning centres for each city
* *Public Services & Transport* : Number of metro stations, bus stops, taxi stands for each city
* *Medical and health considerations:* No of clinics, hospitals, pharmacies, healing centres, gyms, yoga centres for each city
* *Consumer goods:* Number of shopping centres, supermarkets, shops etc. for each city

1. Political and social environment : An important criteria under political and social environment which can be of great concern to people could be the crime rates in the city. To get this data we will mine the Wikipedia page :

<https://en.wikipedia.org/wiki/List_of_states_and_union_territories_of_India_by_crime_rate>

Since the data given in this page is as state level and not city level, we will map this data to each city depending on the state each city belongs to.

1. Housing : For an assessment of housing prices in the city we will scrape data from the following page :

<https://www.makaan.com/price-trends>

This page provides as the Average price per sqft of an apartment for the major cities in India.

A combination of all the above mentioned data points will then be used in the next stage of analysis to classify cities as per the quality of living offered by them.