# Introduction

The problem I am considering is to compare the neighborhoods of four cities in four different countries. The countries I have selected are Bengaluru in India, Seoul in South Korea, Vancouver in Canada and San Francisco in USA. Attempt will be made to check which neighborhoods of the 4 cities are similar.

The target audience for this project is the owners of a restaurant chain which might already have their franchises set up in Vancouver and San Francisco and who want to enter new markets in Asia. They might consider other prominent tech cities in Asia since their target customer is the tech community. Since Seoul and Bengaluru have many MNCs and have a large tech community, they are the target of this project.

# Data Sources

The neighborhood data of the four cities is taken from Wikipedia pages.

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

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

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

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

Beautiful Soup library will be used to web-scrap from the Wikipedia pages. The geocode library to be used is geopy. Care is taken to limit the calls to be less than 1call/sec to meet the term of use of the library. Folium library will be used to represent the data on maps. And scikit-learn will be used to utilize machine learning. And Foursquare API will be used to gather neighborhood data.