**COURSERA CAPSTONE**

**OPENING A NEW GYM IN MUMBAI, INDIA**

**INTRODUCTION**

**Mumbai**  is the [capital city](https://en.wikipedia.org/wiki/Capital_city) of the [Indian](https://en.wikipedia.org/wiki/India) [state](https://en.wikipedia.org/wiki/States_and_union_territories_of_India) of [Maharashtra](https://en.wikipedia.org/wiki/Maharashtra). According to [United Nations](https://en.wikipedia.org/wiki/United_Nations), as of 2018, Mumbai was the second most populous city in India after [Delhi](https://en.wikipedia.org/wiki/Delhi) and the seventh most populous city in the world with a population of almost 20 million. As per Indian government population census of 2011, Mumbai was the [most populous city](https://en.wikipedia.org/wiki/List_of_cities_in_India_by_population) in India with an estimated [city proper](https://en.wikipedia.org/wiki/City_proper) population of 12.5 million living under [Municipal Corporation of Greater Mumbai](https://en.wikipedia.org/wiki/Municipal_Corporation_of_Greater_Mumbai). Mumbai is the centre of the [Mumbai Metropolitan Region](https://en.wikipedia.org/wiki/Mumbai_Metropolitan_Region), the sixth most populous metropolitan area in the world with a population of over 23 million. Mumbai lies on the [Konkan](https://en.wikipedia.org/wiki/Konkan) coast on the west coast of India and has a deep [natural harbour](https://en.wikipedia.org/wiki/Natural_harbour). In 2008, Mumbai was named an [alpha world city](https://en.wikipedia.org/wiki/Alpha_world_city).

**BUSINESS PROBLEM**

The objective of this capstone project is to analyse and select the best locations in the city of Mumbai , India to open a new Gym. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business question i.e; In the city of Mumbai, India if a business man is looking to open a new Gym , where would you recommend to open it?

**DATA**

**To solve the problem , we need the following data:**

* List of neighbourhoods in Mumbai.
* Latitude and Longitudes coordnates of those neighbourhoods.
* Venue data, particularly the data related to Gyms.

**Sources of data and methods to extract them**

This Wikipedia page(<https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai>) contains a list of neighbourhoods in Mumbai with a total of 93 neighbourhoods. We will convert this into a panas dataframe.

After this we will use FourSquare API to get the venue data for those neighbourhoods. This project will make use of many data science skills from data prepreocessing to , working with API(Foursquare), data cleaning , data wrangling , to machine learning and map visualization(Folium).