# The Battle of Neighborhoods

## Problem description and background:

#### "Find out a suitable location in Hong Kong to open a Bollywood themed restaurant"

A very famous restaurant chain in India is interested to expand their global footprint and is interested to open a Bollywood themed restaurant in Hong Kong. They have hired my team to do neighborhood analysis and suggest the best possible location for the restaurant.

This restaurant chain is famous not only for the quality of the cuisine but also for the quirky yet classy Bollywood themed interiors they use. The interior design involves movie cutouts, posters, quirky movie dialogues on table decoration etc. which are very big crowd puller. They also invite movie stars on special occasions. As part of promotional activity, they give out movie tickets or movie hall food coupons etc. to lucky customers during festive season.

After multiple discussions with stakeholders we figured out the below list of requirements for the suitable location:

* Movie going crowd are the primary target customers.
* As the theme of the restaurant is cinema centric so there should be one or more favorite cinemas nearby.
* Having shopping places, amusement parks or other places of interest are a plus as that will ensure more visibility.
* Good public transport is a must. Public transport facilities within 5-minute walking distance is preferable.

The stakeholders emphasized that rental price is not a concern, but the location should be selected based on nearby environment. **Also, they want the restaurant to be setup near one of the favorite cinemas of Hong Kong**. **The list of cinemas were provided by the stakeholders.**

Based on the requirement we did our initial data analysis and came up with **five possible locations** for the restaurant.

Now our challenge is to suggest one among the five possible locations and we will try to find that out the same in the rest of this assignment.

## Data

As per the requirement the following data are needed for analysis:

### 1. Location of Hong Kong Cinemas

As the stakeholders clearly mentioned that the restaurant needs to be setup in the locality of a famous cinema hence, we need the list of Hong Kong cinemas and their geographic coordinates. The data is readily available on the website <https://hkmovie6.com/cinema>. We downloaded the data and converted into datafrmae for analysis. The first five entries are given below:

| **Name** | **ChiName** | **Address** | **Latitude** | **Longitude** |
| --- | --- | --- | --- | --- |
| Emperor Cinemas - Entertainment Building | 英皇戲院 - 娛樂行 | 3/F, Emperor Cinemas Entertainment Building | 22.281453 | 114.154230 |
| The Coronet @ Emperor Cinemas - Entertainment ... | The Coronet @ 英皇戲院 - 娛樂行 | 3/F, Emperor Cinemas Entertainment Building | 22.281453 | 114.154230 |
| Emperor Cinemas - Tuen Mun | 英皇戲院 - 屯門新都商場 | 3/F, New Town Commercial Arcade, 2 Tuen Lee St | 22.390776 | 113.975983 |
| Broadway Circuit - CYBERPORT | 百老匯戲院 - 數碼港 | Shop L1 - 3, Level 1, The Arcade | 22.261067 | 114.129825 |
| Broadway Circuit - PALACE IFC | 百老匯戲院 - PALACE IFC | Podium L1, IFC Mall, 8 Finance Street, Central | 22.285545 | 114.157979 |

### 2. Geographic coordinates of five cinema addresses

Google Map API is used to find geo coordinates of five possible cinemas.

| **Location** | **Address** | **Latitude** | **Longitude** |
| --- | --- | --- | --- |
| L1 | Sau Mau Ping Shopping Centre, Sau Mau Ping | 22.319503 | 114.232187 |
| L2 | Tuen Mun Ferry, Tuen Mun | 22.371780 | 113.966039 |
| L3 | Un Chau Shopping Centre, Cheung Sha Wan | 22.337280 | 114.156457 |
| L4 | Prosperity Millennia Plaza, North Point | 22.291698 | 114.208168 |
| L5 | Tsuen Fung Centre Shopping Arcade, Tsuen Wan | 22.372112 | 114.119317 |

### 3. Favorite cinema list provided by stakeholders

The below is the list of some of the favorite cinemas provided by the stakeholders which will be used to profile the best location for the restaurant.

| **Name** | **Rating** |
| --- | --- |
| Broadway Circuit - MONGKOK | 4.5 |
| Broadway Circuit - The ONE | 4.5 |
| Grand Ocean | 4.3 |
| The Grand Cinema | 3.4 |
| AMC Pacific Place | 2.3 |
| UA IMAX @ Airport | 1.5 |

### 4. Other hangout venues and Public transportation facilities in the locality

The recommended location for the restaurant should have other places of entertainment and hangout venues such as eateries, shopping malls, amusement parks etc. nearby and it should be well connected by public transport.

The following venue categories are used:

* Food, Shop & Service
* Bus Stop
* Metro Station
* Nightlife Spot, Arts & Entertainment

Our team assumed that such venues needs to be within 500 meter of the proposed restaurant which is about 5-minute walking distance.

For demo purpose we used FourSquare API to query some venues around the first cinema (英皇戲院 - 娛樂行) in the cinema list above. The below is a list of results:

* **Metro Station**

| **Name** | **Latitude** | **Longitude** | **Tips** | **Users** | **Visits** |
| --- | --- | --- | --- | --- | --- |
| MTR Central Station (港鐵中環站) | 22.281911 | 114.158406 | 0 | 0 | 0 |
| MTR Hong Kong Station (港鐵香港站) | 22.284926 | 114.158314 | 0 | 0 | 0 |

* **Bus Stop**

| **Name** | **Latitude** | **Longitude** | **Tips** | **Users** | **Visits** |
| --- | --- | --- | --- | --- | --- |
| Seymour Road / Robinson Road Bus Stop 西摩道／羅便臣道巴士站 | 22.280465 | 114.150347 | 0 | 0 | 0 |
| Douglas Street Bus Stop 德忌利士街巴士站 | 22.283273 | 114.156910 | 0 | 0 | 0 |
| Hang Seng Bank Headquarters / Connaught Road C... | 22.284741 | 114.156404 | 0 | 0 | 0 |
| HSBC Headquarters Bus Stop 匯豐總行巴士站 | 22.280577 | 114.159446 | 0 | 0 | 0 |
| Dr. Sun Yat-Sen Museum Bus Stop 孫中山紀念館巴士站 | 22.279132 | 114.152743 | 0 | 0 | 0 |

* **Food**

| **Name** | **Latitude** | **Longitude** | **Tips** | **Users** | **Visits** |
| --- | --- | --- | --- | --- | --- |
| Mana! Fast Slow Food | 22.282921 | 114.154651 | 0 | 0 | 0 |
| Good Luck Thai Food (鴻運泰國美食) | 22.281165 | 114.155296 | 0 | 0 | 0 |
| Soul Food | 22.281668 | 114.152495 | 0 | 0 | 0 |
| Chiu Lung Fast Food (昭隆美食) | 22.282659 | 114.156753 | 0 | 0 | 0 |
| Sun Hing Fast Food (新興美食) | 22.282521 | 114.156717 | 0 | 0 | 0 |

* **Arts & Entertainment**

| **Name** | **Latitude** | **Longitude** | **Tips** | **Users** | **Visits** |
| --- | --- | --- | --- | --- | --- |
| Tai Kwun Centre for Heritage and Arts (大館古蹟及藝術館) | 22.281668 | 114.154216 | 0 | 0 | 0 |
| Wah Tung China Arts Limited (華通陶瓷藝術有限公司) | 22.283046 | 114.152723 | 0 | 0 | 0 |
| Ravenel Fine Arts Limited 睿芙奧 | 22.281819 | 114.156906 | 0 | 0 | 0 |
| Ben Brown Fine Arts | 22.281853 | 114.157285 | 0 | 0 | 0 |
| KONG Arts Space | 22.281751 | 114.153300 | 0 | 0 | 0 |

## Data Analysis Methodology

Our team decided to use content-based recommendation technique using the above data to find out the most favorable location for the new restaurant.

To do this we will use FourSquare API which will provide how many venues in different category are available near Hong Kong cinemas and a matrix will capture characteristic of venues nearby cinemas. Stakeholder's favorite list will be combined with the matrix to make it a weighted matrix of favorite cinemas.

The weighted matrix can be applied on five target locations with venues information to generate a ranking of each location. Then the location with the top ranking in the list can be recommended to the stakeholder.

To build such a weighted matrix we need to perform the below steps:

### Data Cleansing and Preparation

### This is the pro-processing phase which will take care of following activities:

* We need to prepare the data such a way that duplicate entries be removed. For example, we found some duplicate cinemas which some are basically ‘Special House’ inside an existing cinema. Such records need to be corrected.
* Explore venues around the target locations and show the value count in each category. If insufficient number of entries found for a venue category, then drop that venue.
* Foursquare API returns values such as tips, visit counts etc. which may or may not be enough. We need to drop all unnecessary information and clean the data for processing.

### Data Analysis

Our target is to generates descriptive statistics that summarize the central tendency, dispersion and shape of a dataset's distribution. We will then visualize the data in a map where we will show the list of cinemas, we retrieved along with the list provided by stakeholders and will try to figure out which will be the best location in terms of number and type of venues nearby. We will perform regression analysis and use P-Value to find out the best possible

### Machine Learning

We will use **Content-Based or Item-Item recommendation systems**. The objective is to figure out the new restaurant location by finding the cinema that should be in its locality and counting number of nearby venues and ratings given. We will create a model that will do the analysis for us.