## Introduction

My report is for those who are planning to start a new restaurant. I'll try to provide as many suggestions on what all factors need to be considered while starting a new venture.
Here I will be focusing on problems one might face while opening a new restaurant in a city. There are various factors which should be considered while opening a small or big level business. For example say KimJon wants to open a Korean restaurant, first thing she will do is to decide the location for his new restaurant. Now on what basis will KimJon decide his restaurant's location? How much will renting a place cost him? The other key points she will consider is how many Korean restaurants are there in that specific area? If in case there are more than two Korean restaurants which is having good ratings, will it be risky to open new one near these restaurants? What all factors will help KimJon to run his business profitably? .

## Business Problem

In my report I'm going to focus on problems one might face while opening a new restaurant. Well there are many factors to be considered to open a small or big level business. For example say KimJon wants to open an Korean restaurant, first and foremost important point to decide is the location for her new restaurant. On what basis can KimJon decide her restaurant's location? Does KimJon want to rent the place or can she afford to buy it? While selecting the place there are key points to consider like she needs to check out how many Korean restaurants are there in that specific location? If in case there are already two Korean restaurants which have good ratings, will it be risky to open new one near these restaurants? What all factors will help KimJon to run her business above average? (Out of scope for this project: Budget for Kitchen tools, restaurant furniture and decorations, hire new chef's and waiters, budget everybody's salary. Decide on Menu details, restaurant advertisement, publish discount coupons, restaurant website.

## Discussion

Let's discuss the above mentioned problem statements. Firstly KimJon needs to choose a suitable location for her restaurant. Let's say KimJon wants to open a new Korean Restaurant in San Francisco. San Francisco is a famous tourist place, and also its in silicon valley, because of that real estate is very costly. So assuming that KimJon wants to rent a place for her new restaurant. And now she needs to figure out how many restaurants are there in say neighbourhood A, B, C etc. If there are more than two Korean restaurants in a neighbourhood then that would be a great risk to open new restaurant of same cuisine in that neighbourhood. Selecting a place where there is less or no restaurant would be of great choice, considering the rent of neighbourhood too. Why I'm emphasing that there should be less restaurants is, so that KimJon will face less competition with same cuisine restaurants. She needs to look for a place where many people frequently visit so that her business is above average. Places like Downtown, Movie theatre, Malls & Gas stations would help her business running. Restaurants ratings, check in of customers might help in deciding location crowd. I would also suggest that KimJon should check for opening and closing timings of other restaurants. She may try opening her place 30 mins before other restaurants open up and close 30 mins or 1 hour later the other restaurants, this might help to get more customers

## Data Description

As KimJon is planning to open a restaurant in San Francisco, and assuming that she is going to rent a place. So first I took the rent dataset from (https://web.archive.org/web/20180822060556/https://www.rentcafe.com/average-rent-market-trends/us/ca/san-francisco/ , https://www.zillow.com/research/data/) according to neighbourhood wise, so that it's easy for us to check the rent data neighbourhood wise. In this dataset I couldn't get all neighbourhoods

rent information. So I managed to use only those information which I could get from the website. I have cleaned the dataset and I'm going to compare the rent data of year 2018. Because for this project we just need to analyse the current rent range. Since from the webpages I didn't get all neighbourhood's rent data, I planned to test only for the data I have retrieved.

I'm going to use a formula to find which neighbourhood is good to open a new restaurant. Before coming up with a formula, I was wondering what all attributes/factors can we consider because it's really unfair to compare data of 10 years old restaurant with 1 year old restaurant. Like for example, the checkin count of 10 years old restaurant will be more compared to a 1 year old or 6 months old restaurant. And also after analysing data I found that in many restaurants checkin count is zero. I thought check in count would be really be helpful to figure out the number of crowd visiting a particular neighbourhood, but because of data discrepancy I avoided it. Then I concluded that every restaurant would definitely have ratings. Even if 100 customers have visited a 1 year old restaurant, the rating will be out of 5 stars, and same goes for a 10year old restaurant.

## Methodology

To calculate the formula first we have calculated the rent Score. From rent score formula we need maximum and minimum rent of neighbourhood. Maximum and minimum values can be retrieved using max and min built in functions.

*maxrentofN = max(dataframename['rentcolumn'])*

*minrentofN = min(dataframename['rentcolumn'])*

Next in order to calculate ratingScore, we need to retrieve rating of each Korean restaurant neighbourhood-wise. So let's the analyse data. Firstly

we need the co-ordinates of the neighbourhood's, co-ordinates or latitude & longitude can be obtained by passing the Neighbourhood name value through geocoding. The restaurant details can be retrieved using search endpoint of foursquare location. For our project we need only Korean restaurant data, and in search endpoint there is a attribute called category id, i.e for each category(like Korean or Italian or Mexican Restaurant) foursquare has a defined category id which will help us to get the desired data. In this search response, we'll retrieve the venue id of all the Korean venues. And then pass this venue id's through venue_id endpoint to get rating of each Korean restaurant. Let's save the data in a dataframe, for further testing.

Now we need to select a neighbourhood in which we have to reduce the competition for our new restaurant, hence we should test with only good rating restaurants. For this I'm going to consider restaurants which have rating greater than or equal to 7 in foursquare

After getting counts of good rating restaurant in each neighbourhood, we can calculate our rentRating now. Max and min can be calculated using respective functions.

*maxgoodrest = max(dataframename['ratingcountcolumn'])*

*mingoodrest = min(sf_Neighbourhood['ratingcountcolumn'])*

Thus now we can calculate ratingScore using the formula ratingScore = (maxgoodrest-currentrestratingofN)/(maxgoodrest-mingoodrest).

Once we have rentScore and ratingScore we can use our formula to calculate the finalScore. And below is the first 5 rows from the final results dataframe

## Result and conclusion

| | Neighborhood | 2018-07 | Latitude | Longitude | RentScore | GoodRatingRestaurant | RatingScore | FinalScore |
|---|---|---|---|---|---|---|---|---|
| 11 | Stonestown | 2640 | 37.727446 | -122.474895 | 1.000000 | 0 | 1.000 | 1.000000 |
| 12 | Merced Heights | 2640 | 37.717597 | -122.470281 | 1.000000 | 0 | 1.000 | 1.000000 |
| 13 | Lakeside | 2640 | 37.731967 | -122.474257 | 1.000000 | 0 | 1.000 | 1.000000 |
| 10 | Noe Valley | 2618 | 37.751691 | -122.432061 | 0.633800 | 1 | 0.875 | 0.850280 |
| 5 | Marina | 2948 | 37.799793 | -122.435205 | 0.712416 | 0 | 1.000 | 0.827450 |
| 19 | Glen Park | 2967 | 37.733106 | -122.433764 | 0.694676 | 0 | 1.000 | 0.816807 |
| 9 | Pacific Heights | 2973 | 37.792717 | -122.435644 | 0.689076 | 0 | 1.000 | 0.813445 |
| 4 | Western Addition | 2989 | 37.779559 | -122.429810 | 0.674136 | 0 | 1.000 | 0.804482 |

Above is the result of my formula which I used to analyze best neighbourhood's for new Korean restaurant. If you see carefully I have listed Neighbourhood's which has Finalscore greater or equal to 0.8, so that KimJon has more options to choose from. There are 3 Neighbourhood's in above results which are best suited for KimJon as there is no competition, since there is no good rating restaurants and rent is also comparitively low according to our formula