Finding Best Location to start an Asian Restaurant in London

Introduction

- Opening a restaurant is all about location, location, location. However, not every restaurant is suitable for every location, and vice versa.

- It comes down to a combination of restaurant style, target audience, your competitors.

- If you can define your restaurant type and identify your target demographic and its most populated areas, you'll be well on your way to choosing a restaurant location that sets your business up for success.

- There's a lot of work, planning and preparation that goes into opening a restaurant.
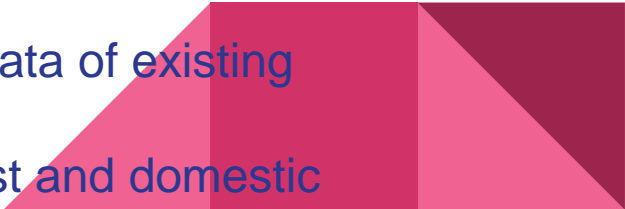
Steps

- Restaurant style

- Your Target Market

- Who are your competitors?

- Create your Menu

- Location and Premises

To solve our problem of finding a best location to start an Asian restaurant in London, we need to datasets based on various parameters such as:

1. Population of target audience in all the boroughs of London based on their :
• Asian ethnicity
• Age
• Gender
• Marital Status
• Employment Status
• Income
2. We also need the data about the required Business floor space and Rateable Value statistics of each borough.
3. Considering the competitors factor, we also need the data of existing Licensed Restaurants in each borough.
4. And lastly we will also consider the borough level tourist and domestic

All the above required information is available at **London Data store, which is a free and** open data-sharing portal where anyone can access data relating to the city.

The data is available in XLS and CSV format, which we can download and can use as-is for solving our problem.

The link for the London Data store - https://data.london.gov.uk/

Along with the above datasets we will also use the **Foursquare location data to solve our** problem.

Methodology

To work on the solution, I have used Pandas library to read the data in XLS format and convert into pandas data frame.

Extensive data exploration analysis is done, where lot of data is cleaned and presented in a suitable format.

Machine Learning Algorithm **Simple Linear Regression** is used to predict the data for Rated Value for the year 2018 for the selected borough.

The dependant variable would be the rated value for year 2018 and the independent variables are the earnings of each borough and the existing
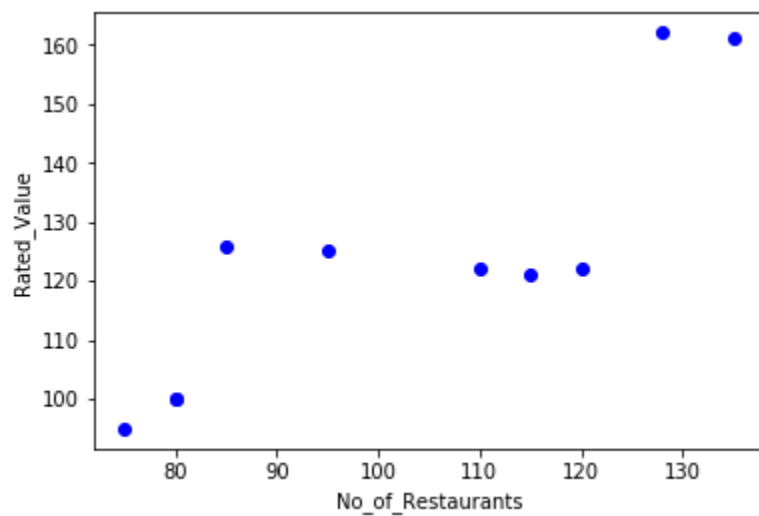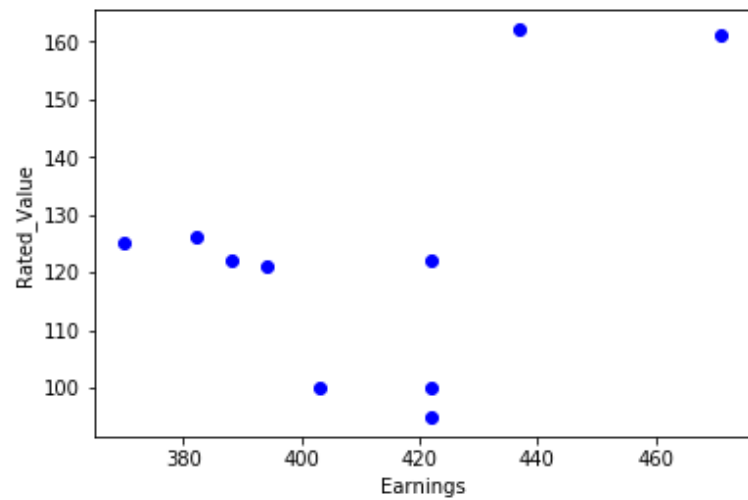
Plots Description:

Two scatter plots are plotted between these variables and from these 2 plots, it is observed that the Linear relationship exists between the Restaurants and the Rated Value.

A scatter plot clearly shows the relation between variables where changes in one variable explain or possibly cause changes in the other variable.

Also, it indicates that these variables are linearly related.

Simple Linear Regression fits a linear model with coefficients Ø = (Ø1, Ø2, …. Øn) to minimize the residual sum of squares between the independent

Coefficient and Intercept in the Simple Linear Regression are the parameters of the fit line. Given that it is simple linear regression with only 2 parameters and knowing that the parameters are the intercept and slope of the line, using the python library SciKit Learn, we can estimate them directly from our data. The available data is divided into Train   model and the test da
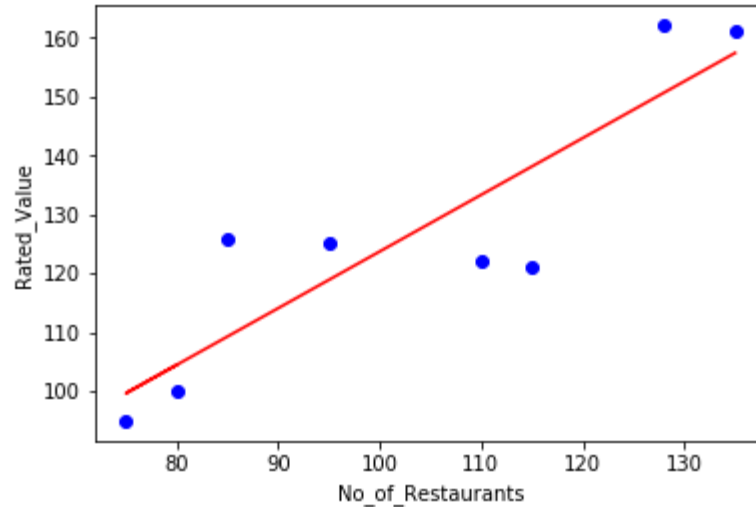
Evaluation of the model is performed using the Evaluation Metrics such as **Mean Absolute Error, Mean Squared Error and R-Squared.**

Due to very less available test data, the R-Squared for our model is not that great, but still we can consider our model for the prediction of the Rated Value for the year 2018.

After the prediction of rated value per sqm of a retail space is completed and when we are convinced that a particular borough will be the preferred location for the restaurant, we have to get the necessary data of that borough.

Discussion:

My observation after doing this analysis is the model we used could have given better results, if we had huge data to train and test our model. In spite of that this model gives us a better insight for our problem and also help us to gain better results. From the clustering results our problem finds a better solution of identifying the best location for the Asian restaurant. We could explore all the neighbourhoods of the borough and could list the most common venues based on their frequency of occurrence. From these results I can strongly recommend the Beckton, Custom house and few other neighbourhoods as a preferred location for our restaurant , as these areas have the restaurant venue as the most common venue

Conclusion:

There is always room for improvement and hence the above solution I have provided can also be improved and the machine learning models can be trained and tested for best results depending upon the data we have.