Choosing the best place to set up a restaurant in London

1.-Introduction

In this report we will analyze a group of datasets about Crime, Restaurants, and Income in London , that aim to solve the question , If i want to put a restaurant in London, Which would be the place to set it up?.

As many entrerpreneurs know, all kind of businesses involve a degree of risk , and without the proper knowledge, one can lose a big quantity of money, effort and time .

But how the risk is measured ...Risk depends of several factors, It's of public knowlegde that people will avoid a venue if is placed in a red zone , based on this business wont flourish due to the threat level. Another way to measure the risk for this report is the number of Restaurants in the area,in a simple approach (Considering low mobility between borough) as more restaurants ,less average people per each one. Also it depends of the category of the restaurant, something different or scarce will drive more people to it

In the past this quantities were difficult to measure , and without analytics and guiding only by local knowledge , the risk of setting up a business was several times greater . However since 2001 and with the Open data initiative of London Government ,has been collecting information on crime, restaurants , income ,

Doing and analysis would take several weeks , however now it could take just a matter of days , and with a probability of giving better results

In this project with this data and an addequate use of statistics techniques ,we will find the best addecuate place to set up a Restaurant, in addition we'll tell the restaurants per category in the borough

2.-Data

London Crime Per Borough Monthly Report of Crimes per LSOA in London from 2008-2016 It could be use to group the number of crimes per borough in a given period of time , find the places with most crime ocurrences , and see how crime rate changes with time

London Licensed Restaurants List of Licensed Restaurants (2001-2017) for number of venues, number of employment. Foursquare

London Unlicensed Restaraunts List of Unlicensed Restaurants (2001-2017) for number of venues, number of employment

London TakeAway food places List of Takeaway Places (2001-2017) , for number of venues, number of employment London Pubs List oF London Pubs London Earnings Residence Borough (2001-2018) for number of venues, number of employment

All of the venues databases will be used to find the number of food places per borough and a more detailed anaylisis will find the borough's were restaurants are more succesful

Population Data To find crime and restaurant density

Foursquare location data To list the food category places in a given borough, this is important in order to give the consumer something new

Libraries

Pandas Numpy Matplotlib Json Folium Beautiful Soup

3.-Methodology

First all the necessary libraries were imported ,as a second step , databases were read by the intepreter using pandas

The first dataframe readed was London Crime ,below is a preview of a few rows

6	E01004563	Wandsworth	Robbery	Personal Property	0	2008	6
7	E01001320	Ealing	Theft and Handling	Other Theft	0	2012	5
8	E01001342	Ealing	Violence Against the Person	Offensive Weapon	0	2010	7
9	E01002633	Hounslow	Robbery	Personal Property	0	2013	4
10	E01003496	Newham	Criminal Damage	Criminal Damage To Other Building	0	2013	9
11	E01004177	Sutton	Theft and Handling	Theft/Taking of Pedal Cycle	1	2016	8
12	E01001985	Haringey	Theft and Handling	Motor Vehicle Interference & Tampering	0	2013	12
13	E01003076	Lambeth	Violence Against the Person	Other violence	0	2015	4
14	E01003852	Richmond upon Thames	Robbery	Personal Property	0	2014	1
15	E01004547	Wandsworth	Violence Against the Person	Offensive Weapon	0	2011	10
16	E01002398	Hillingdon	Theft and Handling	Theft/Taking Of Motor Vehicle	0	2016	2
17	E01002358	Havering	Violence Against the Person	Wounding/GBH	0	2012	2

The first line is the LSOA code , which is group of coordinates or in specific an Area,

The second one is the Borough in wich LSOA is located

The third is the major crime category and the fourth the minor crime category

The fifth is the crime prevalence ocurrence in a month, the number six the year of the report , and number seven the month of the report.

For this specific applications we'll only need the number of crimes per year , so I used pivot table to get the real number of crimes per borough in a range of years

year	2008	2009	2010	2011	2012	2013	2014	2015	2016
borough									
Barking and Dagenham	17656.0	17713.0	17130.0	16686.0	15990.0	15759.0	15426.0	16346.0	16741.0
Barnet	23944.0	23449.0	23416.0	24007.0	23573.0	22510.0	22572.0	24036.0	24684.0
Bexley	15305.0	14445.0	13166.0	11325.0	11669.0	11552.0	11844.0	11990.0	12840.0
Brent	23504.0	24748.0	25512.0	27669.0	26357.0	23759.0	24426.0	24883.0	26693.0
Bromley	24015.0	21858.0	19572.0	19868.0	19622.0	18945.0	19668.0	20637.0	20164.0
Camden	32010.0	31612.0	31804.0	33956.0	31472.0	28435.0	27198.0	29477.0	29183.0
City of London	NaN	NaN	NaN	69.0	157.0	110.0	115.0	151.0	178.0
Croydon	29516.0	29784.0	29373.0	29830.0	30189.0	27444.0	27357.0	28089.0	28712.0
Ealing	28947.0	28617.0	31360.0	29537.0	28832.0	25917.0	24978.0	26247.0	27127.0

I also got the average crimes per Borough,

	borough	count
0	Barking and Dagenham	16605
1	Barnet	23576
2	Bexley	12681
3	Brent	25283
4	Bromley	20483
5	Camden	30571

I also worked with population , number of restaurants and average wage

,I followed the same methodology , and tools merge dataframes, drop columns , drop rows, restart indexes , scraping wikipedia tables , convert string to numbers.

For getting the most suitable places I considered Crime density and Restaurant Density being

Crime density :Number of crimes /Number of pop in borough

Restaurant density :Number of Restaurants /Numbers of pop in borough

I chose this because is well known that people are less likely to buy / consume in a dangerous place .

And it's very likely to earn more in a place where a good is scarce(As less restaurants more quantity of people for each one .)

But which variable is more important , for the analisis Crime density or Restaurant Density; For this I' ve used variance and mean to find the evolution of each one within the years

4.-Results

## I got a general table where restaurants and crime with populations appears

|	Borough	Population	Crime_Density	Rdensity	ULdensity
0	City of London	7000	0.018429	0.114286	0.069286
1	Harrow	243,372	0.053346	0.001767	0.001089
2	Bexley	236,687	0.053577	0.001753	0.000824
3	Richmond upon Thames	191,365	0.056186	0.002770	0.001672
4	Sutton	195,914	0.057270	0.001914	0.001021
5	Kingston upon Thames	166,793	0.059487	0.002338	0.001349
6	Merton	203,223	0.063231	0.001944	0.001181
7	Havering	242,080	0.063772	0.001921	0.001033
8	Barnet	369,088	0.063876	0.002018	0.001314
9	Bromley	317,899	0.064432	0.001950	0.001101
10	Enfield	320,524	0.067209	0.001654	0.000936
11	Redbridge	288,272	0.070749	0.001682	0.000919
12	Wandsworth	310,516	0.073262	0.002448	0.001498
13	Greenwich	264,008	0.076414	0.001799	0.000985
14	Croydon	372,752	0.077588	0.001918	0.001033
15	Hounslow	262,407	0.079083	0.001905	0.001086
16	Brent	317,264	0.079691	0.001844	0.001135
17	Hillingdon	286,806	0.081229	0.002005	0.001063
18	Ealing	342,494	0.081610	0.001869	0.001124
19	Lewisham	286,180	0.083528	0.001835	0.000978

I got the Crime density
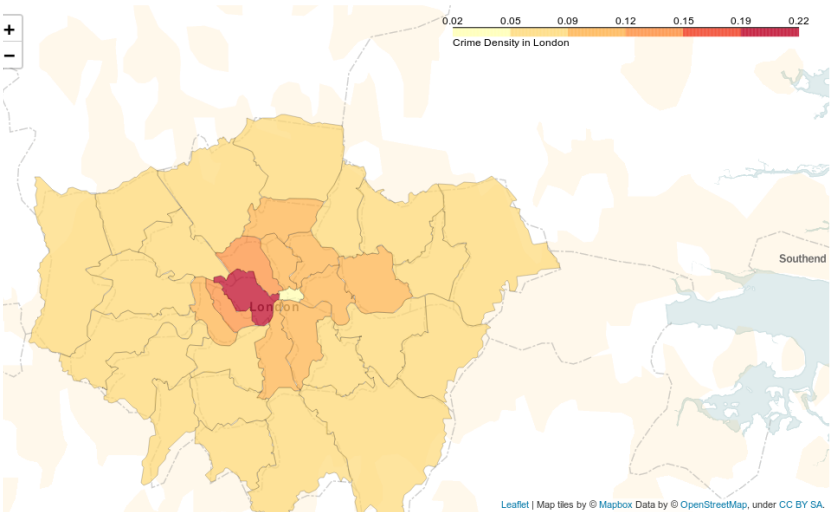
I got the Restaurant Density
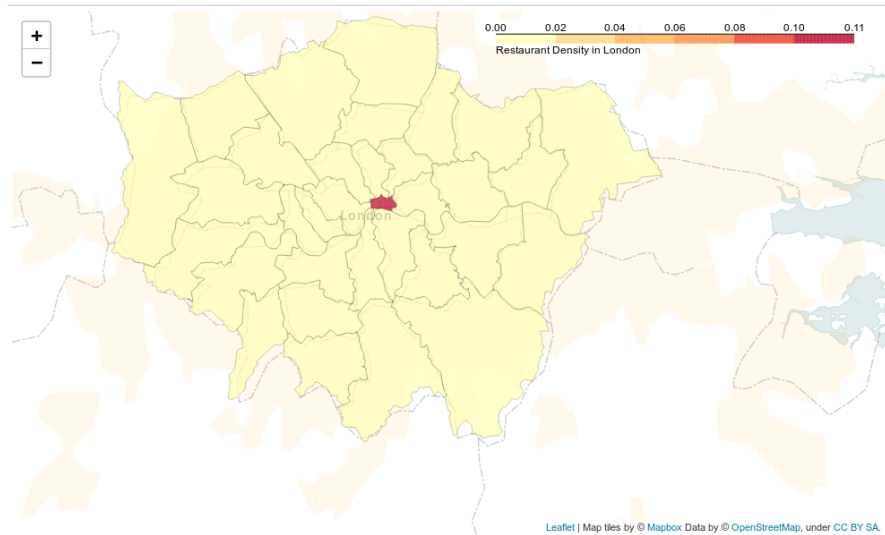
I got the most volatile variable between both (Crime Density and Restaurant density)

Average variation of deviation for Restaurant : 0.11865445405918966
Average variation of deviation for Crime : 0.07048461173643936

As the average variation of deviation for Crime
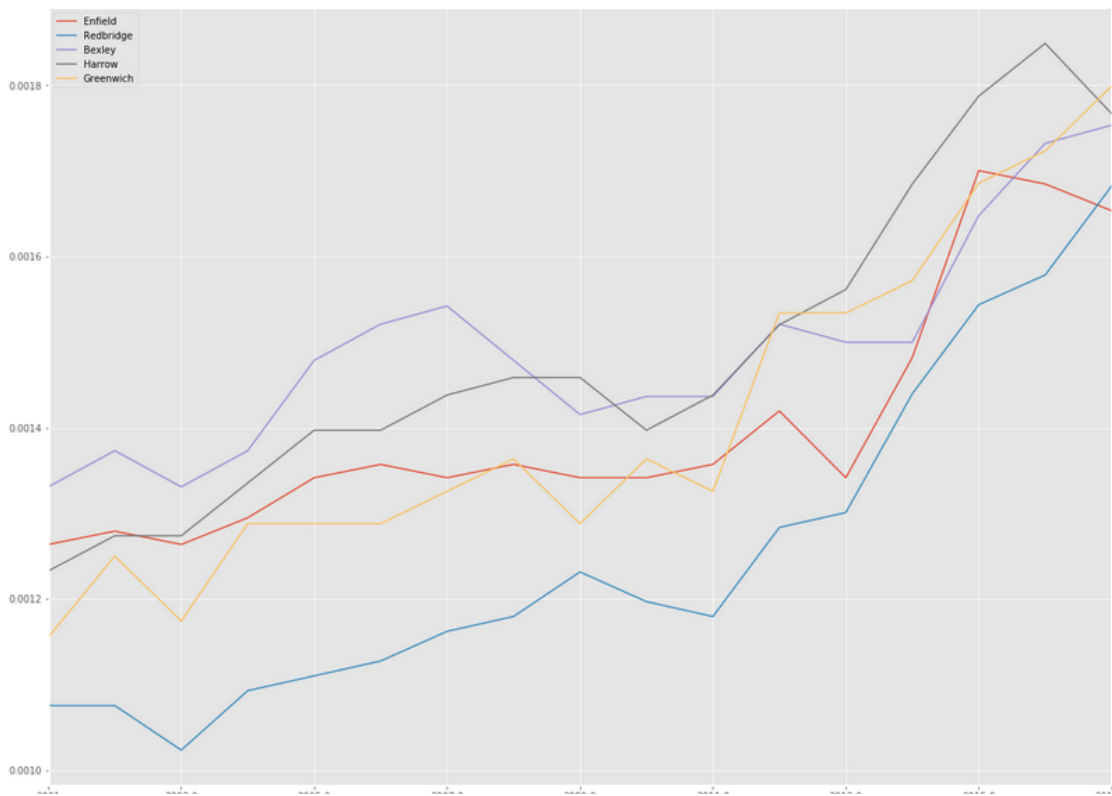
With the 5 less dangeroues places i found the evolution of the number of restaurant within the years

Results from foursquare API for 5 less dangerous places, the number of restaurants didnt match the average restaurant number (~250)

'] :

Venue Category	Borough
African Restaurant	1
American Restaurant	3
Asian Restaurant	1
Chinese Restaurant	4
Coffee Shop	27
English Restaurant	2
Fast Food Restaurant	15
French Restaurant	1
Greek Restaurant	1
Indian Restaurant	17
Italian Restaurant	9
Mediterranean Restaurant	3
Mexican Restaurant	1
Middle Eastern Restaurant	1
Portuguese Restaurant	7
Restaurant	1
Sandwich Place	11
Thai Restaurant	2
Turkish Restaurant	5
Vietnamese Restaurant	1

Average wage of most suitable Boroughs

	Area	2018
0	Greenwich	573.7
1	Redbridge	554.7
2	Harrow	538.3
3	Bexley	513.8
4	Enfield	479.1

Conclussions

I used foursquare but the records of 4Square didnt match wwith the actual numbers so , 4 Square data is not useful for this exercise.

Machine learning algorithms weren't needed

Greenwich was selected as the most suitable place to set up a new restaurant base on business growth and income