Battle of the Neighbourhoods

Heather Tymms

August 22, 2019

Abstract

This report compares two of the financial capitals of their respective countries, New York City, USA and London, UK. The comparison will be around two factors: house prices and crime rates. These are two highly valuable factors to be considered when starting up a business. Crime rate is a useful tool to analyse the security of property and lessen harm to employees. House prices are beneficial to find out as it is a good review of rent and also good for moral of employees.

1 Introduction

The Problem in question is about finding where is better to live in a large City environment. Both New York City, USA, and London, UK, are the financial capitals in their respective countries. This make it desirable to live there, and to start up businesses in and around the cities. This assignment will try to distinguish where would be better to live and to open up a business by considering house prices and crime rates in each location. Analysing the best financially acceptable Borough with consideration for crime rates and house prices. There will be an analysis of coffee shops around each interesting Borough too.

The target audience of this report are prospective business owners and employees moving locations. The result of this report will be useful for businesses because they can see how much it would cost to buy in different boroughs. The crime rate will help figure out if the places would be good for starting up a business. For employees, this report can show which areas would be best to move to while at a job in the big cities, to make sure them and their families would be safe and in budget.

2 Data Collection

The data sets collected were from several sources. For the NYC data, the New York Police Department (NYPD) online database and the Department of Finance (DOF) were used to find each of the data sets needed to solve the problem. For the London data, the "London Borough Profiles" data set was used from

the London DataStore online. These are reputable sources to collect data from because they are official data collected by the authorities of each city. Foursquare was used to find out which has the best reviewed coffee shops around the area.

3 Methodology

3.1 Cleaning

In the initial London Data Frame, there was 84 columns, each with different information regarding attributes of London like: population density and election results from several years. These were not needed for this analysis, two columns were found to be useful, crime rate per thousand and median house sales. The dropna function was used to drop rows that contained any missed data. It was the most reliable way to use the visual packages, because all the missing data was Citywide and not Borough related information.

For New York City's collection, three data sets were used to get the correct information for the analysis, all from either Department of Finance for the USA or from the NYPD itself. Both of these are reliable sources. Using the NYPD database for the arrest rate, identified the correct column, and used the groupby function to group the rows into Boroughs. Each of the values in this column were divided by its population to find the arrests per thousand. The house price went through a similar algorithm: located a specific column (median hose price), then groupby was used.

This equation was how the arrests per thousand column was found:

$$arrestsperthousand = arrests/population * 1000, \tag{1}$$

where we find arrests and population number in each Borough. Although London database had this already implemented.

3.2 Exploratory Data Analysis

After the data was collected and cleaned, the main data was put into a pandas Data Frame to be analysed. There is a difference in the number of Boroughs in each city, NYC has five while London had over 30 Boroughs. Each Data Frame consisting of 3 columns: the index as the Borough name, the Crime rate and the median house price.

In the initial exploratory data analysis, it was found that Manhattan had the highest median house price, and also has the second highest arrest rate. Using the describe attribute, the initial statistical analysis was performed on each of the Data Frames.

Table 1: This is a table to show the NYC.describe() in the Jupyter NoteBooks, showing the descriptive statistics of the NYC data Frame

statistic	Arrest per thousand	Median House Price
count	5.000000	5.000000e+00
mean	7.088502	1.408611e+06
std	2.040000	1.649671e+06
min	5.285755	5.123333e+05
25%	5.344904	5.422230e+05
50%	6.262945	7.546667e+05
75%	9.031504	8.871667e+05
max	9.517404	4.346667e+06

Table 2: This taable shows the descriptive statistics for the London data base for crime rate and house price.

Statistics	Arrest/1000	Median House Price
count	36.000000	3.600000e+01
mean	84.480556	4.448719e+05
std	29.811664	1.922015e+05
min	50.400000	2.099950e+05
Q1	65.300000	3.437500e+05
Q2	77.500000	4.036000e+05
Q3	99.675000	4.775000e+05
max	212.400000	1.200000e+06

4 Results

From the Data Frames, for both cities, many visualisation forms were collected. These include regression between the crime rates and house prices, bar graphs comparing each Borough in a location and line graph comparing the two factors in each city.

The results are in the figures on the next few pages. The regression shows a positive correlation between house prices and crime rates in both New York and London. Figure 1 and Figure 2 are the graphs.

Figure 3 are the bar graphs showing the comparison of each Borough using one factor. Figure 4 and Fig5 are the line graphs showing the comarison of the two factors in each city.

5 Discussion

This report has some interesting results. As you can see, the most crime ridden Boroughs in the USA city is the Bronx and Manhattan. Manhattan is the most expensive place to live in the City as well, this skews the regression to infer that as crime increases, so does the house prices. This may not be accurate to infer as there are many other factors to consider when computing the house prices.

London has a similar output, Kensington and Chelsea has a high crime rate as well as high house rate. In the regression of the City of London, it also shows a positive regression, as crime increases, so does the

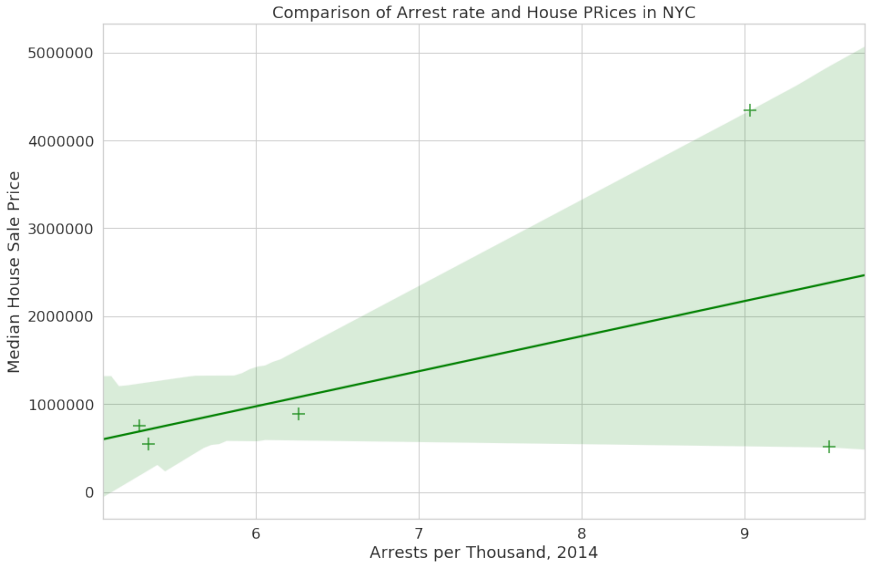
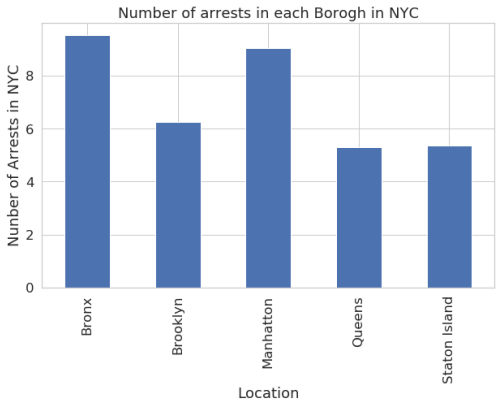
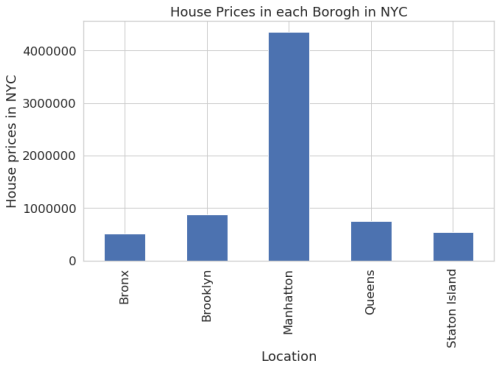


Figure 1: This is the regreesion between house prices(y) and arrest rate (x) in NYC.

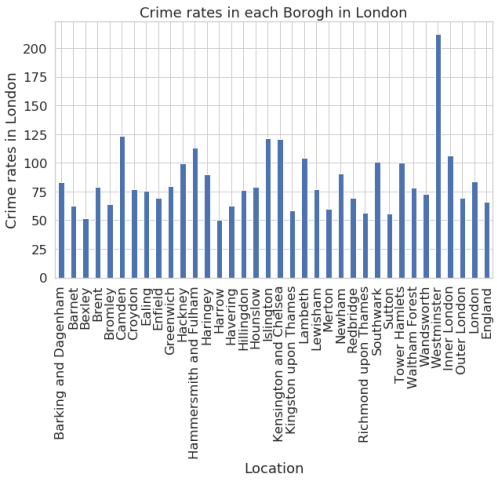
Figure 2: This is the regression bettwen house prices(y) and crime rate(x) in London.



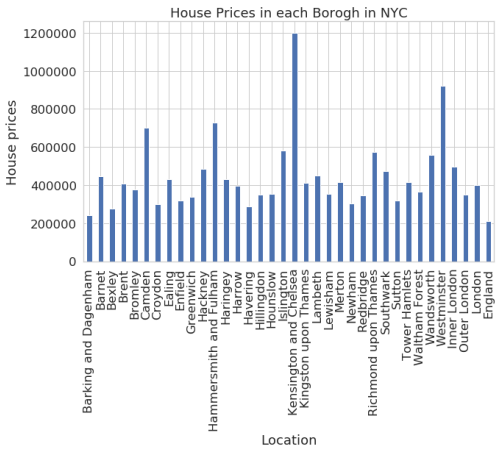
(a) Bar graph of arrest rate in NYC



(b) Bar graph of house prices in NYC



(c) Bar graph of arrest in London



(d) Bat graph of house prices in LDN

Figure 3: Bar graphs showing anount of each attribute in every Borough.

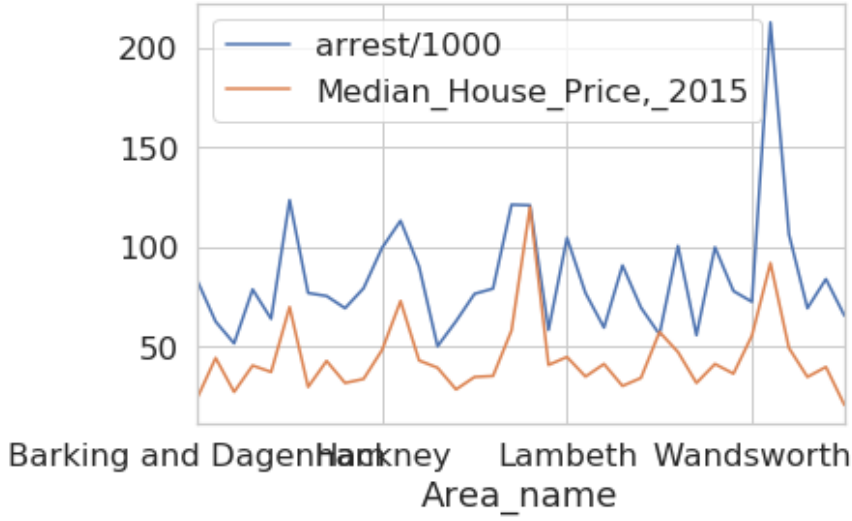


Figure 4: This is a linegraph showing both factors on the same axis in London

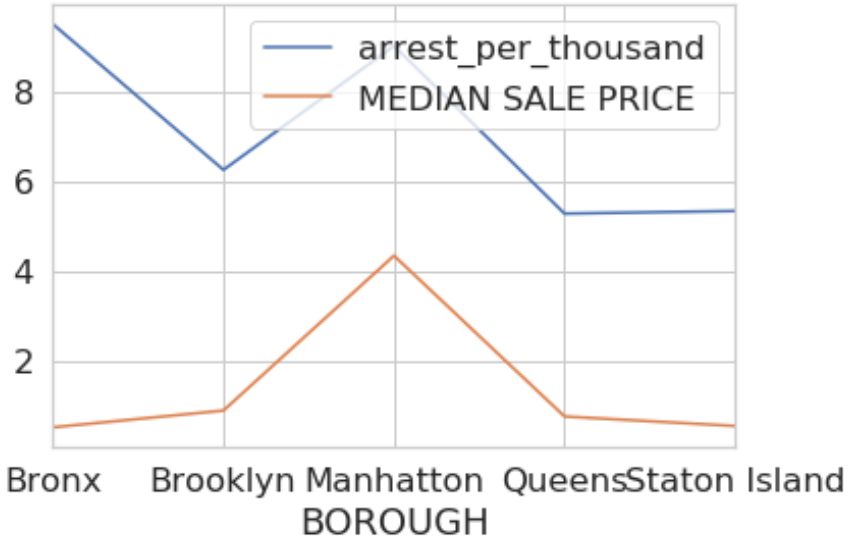


Figure 5: This is a linegraph showing both factors on the same axis in New York City.

house price. This means that this is not the only independent variable that the house price is dependent on.

5.1 Living

Both crime rates and housing prices are factors of moving to a new location. Finding a good place to live, a person would have to balance many factors including the two factors in this report. The less crime rate, the more desirable, and housing within budget. The least arrest rate in New York City is either Queens or Statton Island. There is less likelihood of arrests in both of these areas. Using the Bar Graphs in Figure 3 shows that Queens and Statton Island has a similar arrest rate, but Statton Island is less expensive than Queens. It looks more desirable to live around Statton Island when only considering arrest rate and house price.

In London, there are more Boroughs to choose from. Also using Figure 3, we can see that there are four Boroughs that crime rate is around 50 per thousand: Bexley, Harrow, Richmond-Upon-Thames and Sutton. From the House price Bar graph, Harrow has just below £400,000 price for an average house where Bexley has between £200,000 and £300,000 as an average. Considering only these two factors Bexley seems like the best place to live.

This table shows the best paces to live in both New York City and London with the information given. From this table, using foursquare to find coffee shops in all four locations. It is found that there is more coffee shops in Queens and Becley from the two cites. Comparing these two places, the best place to live/ work is in Queens, New York City.

begintable

Location	City	Arrest per thousand	Median House Price
Statton Island	NYC	5.34	542,000
Queens	NYC	5.28	755,000
Bexley	LDN	51.1	275,000
Harrow	LDN	50.4	396,000

6 Further Reading

There are substantial information on the internet to have a more in depth comparison. Other factors that are useful to compare is the population density and also public transport links. These factors are good to consider. Also next time it is best to compare either arrest rate and crime rate, they are different factors and are not as easy to collate.

7 Summary

In summary, it is best to set up a business in Queens, NYC. This is because there is a balance of lower arrest rate and lower house prices. There are also coffee shops around Queens. Next time, more factors should be used to compare the two locations, including: transport links and population density.